Logical Reasoning and Computation: Essays dedicated to Luis Fariñas del Cerro

Pedro Cabalar, Martín Diéguez,
Andreas Herzig and David Pearce

Toulouse, 3rd-4th March 2016

Preface

Near the end of 2015, Luis Fariñas del Cerro officially retired as directeur de recherche in the CNRS and became an Emeritus researcher of the CNRS. The present volume is a Festschrift in his honour to celebrate Luis's achievements in science, both as an outstanding scholar as well as a remarkable and highly successful organiser, administrator and leader in science and technology policy and management.

The volume contains 15 scientific contributions by 21 authors, among them Luis's colleagues, former students and friends. They will be presented at an international workshop, Logical Reasoning and Computation, to be held at IRIT, Université Paul Sabatier, Toulouse, on March 3-4, 2016. The volume includes a short scientific biography, written by Philippe Balbiani and Andreas Herzig, that describes the many different areas of logic and computation where Luis has made significant advances to the field.

Despite setting a tight deadline for contributions, we received a fantastic response from all the scholars we contacted. It became clear that Luis is held in great affection and esteem by his students, co-authors and close collaborators. This is also witnessed by the breadth of Luis's geographical reach: this volume alone includes scholars from 10 different countries and 4 continents. Besides scientific papers, we also received contributions in the form of personal reminiscences, poems and even a song, that will be presented and performed at the celebratory workshop.

Since Luis has been slowly winding down his administrative responsibilities, he has recently been able to dedicate a greater effort to research once again, entering with great enthusiasm new and exciting fields such as computational biology. Luis, we surely speak on behalf of all the contributors here to wish you enormous success and enjoyment in your new role and we look forward to many more years of inspiring cooperation with you in the future.

February 2016

Pedro Cabalar
Martín Dieguez
Andreas Herzig
David Pearce

Table of Contents

Logic, Leadership and Enthusiasm: a Short Biography of Luis Fariñas del Cerro

Philippe Balbiani, Andreas Herzig

Institut de recherche en informatique de Toulouse
Université de Toulouse

Luis studied at the Universidad Complutense de Madrid, where he obtained a *Licenciatura* in mathematics in 1972, and later a PhD in mathematics (1982). In between he had joined in 1977 the *Laboratoire d'Informatique pour les Sciences de l'Homme* of the *Centre National de Recherche Scientifique* (CNRS) as a CNRS researcher. The lab being based in Marseille and headed by Mario Borillo, he prepared his PhD under the supervision of Maurice Nivat at *Université Paris VII* [1]. When he defended his thesis "Déduction automatique et logique modale" in 1981, he and part of the Marseille lab had just moved to Toulouse (1980). The small group soon became part of the *Langages et Systèmes Informatiques* lab (1982) where Luis defended his habilitation in 1985.

The lab was merged in 1989 with three other Toulouse labs into the ***Institut de Recherche en Informatique de Toulouse*** (IRIT). Luis had already served as the head of its scientific board for a couple of years when he became IRIT's director in 1999. During the 12 years of his reign (which ended in 2011) he managed to restructure the lab in depth and to greatly improve its organisation and research output. He managed to bundle the 20+ teams into 7 themes and to enlarge IRIT's patronages to all four public universities of Toulouse. His success was confirmed by the top "A+" mark that was given to IRIT in 2010 by national evaluation agency AERES.

Beyond IRIT, Luis was in charge of the scientific strategy of ***Université Toulouse III Paul Sabatier*** (2008-2012), which hosts the main part of IRIT. He was a driving force in the construction of the *Université de Toulouse* (now *Université Fédérale de Toulouse*) which federates the four already mentioned Toulouse universities and of which he was elected president of the senate in 2012.

Luis made an exceptional career in the **CNRS**: recruited as *chargé de recherche*, he became *directeur de recherche* in 1991 and was subsequently promoted to *première classe* and *classe exceptionnelle*. He served as *directeur adjoint* of the newly created CNRS department *Sciences et Technologies de l'Information et de la Communication* where he was in charge of international relations (2001-2004) and took up service recently as *directeur adjoint scientifique* at the *Institut des Sciences de l'Information et de leurs Interactions* in 2015.

During his career Luis founded the Applied Logic Group, which merged with the "*Langue, Raisonnement, Calcul* group in 2000. The group produced an important number of PhD theses that are detailed below and grew rapidly. Andreas Herzig was recruited as a *chargé de recherche* CNRS in 1990 (*directeur de recherche* since 2004) and Philippe Balbiani in 1991 (*directeur de recherche* since 2007). Olivier Gasquet obtained a *maître de conférences* position at UPS in 1994 (professor since 2005) and Dominique Longin a *chargé de recherche* CNRS position in 2000. When Luis became director of IRIT, the LILaC group was first headed by Andreas Herzig (2000-2004) and then by Philippe Balbiani (2004-2015). It is now lead by Dominique Longin and Emiliano Lorini.

In 1990 Luis founded the Journal of Applied Non-Classical Logics (JANCL) and acted as its Editor-in-Chief until 2014. The JANCL is a major forum for publications covering all aspects of non-classical logic that is well-established in the fields of philosophical logic, mathematical logic, theoretical computer science and artificial intelligence. Since its creation the JANCL was a protagonist in the domain of non-classical logics, promoting the spreading of novel approaches and their application.

Luis was involved in numerous projects on the national and European level, including the ESPRIT Basic Research Actions "Mechanising Deduction in Logics of Practical Reasoning" (MEDLAR) and "Defeasible Reasoning and Uncertainty Management" (DRUMS). He also set up the *Laboratoire Européen Associé* (LEA) "French-Spanish Laboratory for Advanced Studies in Information, Representation and Processing" with *Universidad Politécnica de Madrid*. Luis was elected a member of the *Académie des Sciences Inscriptions et Belles Lettres de Toulouse* in 2014. His international reputation was confirmed by his election as an ECCAI Fellow in 2005.

Luis's work covers many areas of logic, centered around non-classical logics. In the sequel we are going to enumerate the most important topics.

Proof methods and computability of non-classical logics. Everything started with Luis's PhD thesis, where he was the first to extend the resolution method to modal logics. This was prolonged in several publications with Patrice Enjalbert, as well as in the PhD thesis of Marta Cialdea (now professor at *Università degli Studi Roma Tre*) about Herbrand property for modal logics [2], Andreas Herzig's and Olivier Gasquet's PhD theses about the translation into first-order logic [3,4]. By the end of the 90s and together with Olivier Gasquet and Andreas Herzig, he got back to a more traditional proof method for modal logics and started an in-depth investigation of tableaux method. The result was a very general definition of tableaux procedure based on graph rewriting, ultimately leading to the implemented tableaux theorem proving platform LoTREC[1] and to a tableaux-based introductory book to modal logics [5].

Further work of Luis included work on paraconsistency, in the framework of the PhD of Mamede Lima Marques (now professor at *Universidade de Brasilia*) [6] and in collaboration with Walter Carnielli from the *Unversidade de Campinas*, Brazil. Moreover, the proof theory of Epstein's dependence logic was investigated in the PhD thesis of Valérie Lugardon [7].

Logic programming. During his years in Marseille, Luis interacted with Alain Colmerauer and his group who at that time were inventing logic programming and PROLOG. This inspired Luis to investigate extensions of logic programming languages by modal operators. This lead to the metaprogramming framework MOLOG, whose implementation TARSKI was done during the PhD of Jean-Marc Alliot (now professor at IRIT) [8].

Luis had started an in-depth logical investigation of the notion of negation as failure with Philippe Balbiani's PhD thesis [9] whose approach was based on the Gödel-Löb provability logic.

Luis recently got back to the logical foundations of logic programming and more specifically Answer-Set Programming (ASP): during the PhD of Ezgi Iraz Su he investigated the modal logic behind equilibrium logic as well as modal extensions of ASP [10]. Further research was done with his Spanish colleagues Pedro Cabalar, David Pearce and Agustín Valverde [11,12].

In the 90s, Luis obtained several results on modal logic programming together with his Finnish colleague Martti Penttonen [13,14]. A byproduct of this line of work was a general method of producing undecidable modal logics ('grammar logics') [15].

Non-monotonic reasoning, conditional logics, belief revision. The work on negation as failure in logic programming opened a research avenue towards non-monotonic reasoning mechanisms. Luis's approach was based on conditional logics. While standard modal logics have unary modal operators, conditional logics have binary modal operators relating two formulas (an antecedent and a consequent), whence the relation to non-monotonic consequence relations.

The PhD thesis of Philippe Lamarre's (now professor at *Institut National des Sciences Appliquées* in Lyon) provided an embedding of the main existing conditional logics into standard modal logics, which came as a surprise [16].

The PhD thesis of Gabriella Crocco (now professor in the philosophy department of *Aix-Marseille Université* contributed a thorough proof-theoretical analysis of conditional logics and the non-monotonic reasoning principles [17]. This also lead to the publication of an edited volume [18] that became a standard reference.

During these years Luis participated in a group of French researchers who conducted a comparative evaluation of non-monotonic reasoning formalisms under the name Léa Sombé, a rewriting of the French default reasoning statement "*Les A sont B*" ("the *A*s are *B*s") [19,20,21].

Reasoning about time, actions and knowledge. While the modal operators Luis investigated in the 80s were rather abstract, he subsequently started to work on its most important applications, viz. the logical modelling of reasoning about time, actions and knowledge.

Things started with a textbook on temporal logics for program verification Luis wrote together with Eric Audureau and Patrice Enjalbert [22] (*Editions Masson*). and Saïd Soulhi's PhD thesis on reasoning about knowledge and mutual knowledge [23].

The PhD thesis of Pierre Bieber (now researcher at *Office national d'Etudes et de Recherches Aérospatiales* (ONERA)) provided a first integrated account of logics of action and knowledge (precisely, a non-monotonic version: autoepistemic logic) in view of the verification of cryptographic protocols [24].

The PhD thesis of Christel Seguin (now researcher at ONERA) extended the picture towards reasoning about intentions and planning [25]. This line of work was later taken up and applied to speech act theory in the PhD thesis of Dominique Longin, in the framework of a project with *France Télécom* on human-machine dialogue systems [26].

[1] `https://www.irit.fr/Lotrec`

The PhD thesis of Nathalie Chetcuti-Sperandio (now *maître de conférences* at *Université d'Artois*) related the temporal reasoning line of work to proof methods by investigating tableaux for duration calculus [27].

Further work included Data Analysis Logic DAL that Luis defined with Ewa Orlowska and that provided a link with rough set theory [28].

Spatial reasoning. In parallel with his investigations on conditional reasoning and tableaux-based approaches in non-classical logics, in the early 90s Luis became interested by geometrical reasoning. After he presented an unforgettable talk on that subject to his research group, together with Philippe Balbiani he started to investigate the possibility of defining a modal logic of space with points and lines playing the role of possible worlds and with geometrical relationships between them playing the role of accessibility relations. Then, together with Tinko Tinchev and Dimiter Vakarelov, they produced in 1994-1997 the first modal logic of incidence geometries [29]. This modal logic has been, since that time, the starting point of several other modal logics for point-line geometry. At the same time, Luis became interested by qualitative spatial and temporal reasoning (QSTR). At that time, the investigation of QSTR amounted to research on Region Connection Calculus and Allen's calculus. Together with Philippe Balbiani and their PhD student Jean-François Condotta (PhD in 2000, now professor in Artois University), Luis developed many new qualitative frameworks for reasoning about space and time: the rectangle calculus as a two-dimensional variant of Allen's calculus, the block algebra, etc. These qualitative frameworks are, by now, parts of many robot navigation systems based on QSTR [30]. Finally, Luis' investigations in geometrical reasoning has also given rise to the book about the mechanization of geometry *Eléments de géométrie mécanique* [31] written in collaboration with Philippe Balbiani, Vincent Dugat and the PhD student Anne Lopez [32]. Further work included the PhD thesis of Claudio Masolo (now CNR researcher at LOA-ISTC Italy) on the ontology of space and time [33].

Classical logic. Together with Robert Demolombe, Luis worked on abduction problems in classical first-order logic, alias consequence finding, as well as on the notion of topic in first-order logic. They recently applied the resulting techniques to reasoning about metabolic networks, within the framework of the PhD thesis of Naji Obeid [34]. This line of work is currently pursued with Jean-Marc Alliot and Martín Dieguez.

Bibliography

1. Fariñas del Cerro, L.: Déduction automatique et logique modale. PhD thesis, Université Paris VII (1981) (Thèse de doctorat d'état).
2. Cialdea, M.: Une méthode de déduction automatique en logique modale. PhD thesis, Université Paul Sabatier, Toulouse (1986)
3. Herzig, A.: Raisonnement automatique en logique modale et algorithmes d'unification. PhD thesis, Université Paul Sabatier, Toulouse (1989)
4. Gasquet, O.: Déduction automatique en logique multi-modale par traduction. PhD thesis, Université Paul Sabatier, Toulouse (1994)
5. Gasquet, O., Herzig, A., Said, B., Schwarzentruber, F.: Kripke's Worlds - An introduction to modal logics via tableaux. Studies in Universal Logic. Springer-Birkhuser (2014)
6. Lima Marques, M.: De la connaissance à la paraconsistance : un modèle d'application pour la résolution de conflits aériens. PhD thesis, Université Paul Sabatier, Toulouse (1992)
7. Lugardon, V.: Sur les fondements de la notion de dépendance et de son application à la théorie de l'action. PhD thesis, Université Paul Sabatier, Toulouse (1996)
8. Alliot, J.M.: TARSKI, une machine parallèle pour implanter des extensions de PROLOG. PhD thesis, Université Paul Sabatier (1992)
9. Balbiani, P.: Une caractérisation modale de la sémantique des programmes logiques avec négation. PhD thesis, Université Paul Sabatier (1991)
10. Su, E.I.: Extensions of Equilibrium Logic by Modal Concepts. Thèse de doctorat, Université de Toulouse, Toulouse, France (2015)
11. del Cerro, L.F., Pearce, D., Valverde, A.: FQHT: the logic of stable models for logic programs with intensional functions. In Rossi, F., ed.: IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013, IJCAI/AAAI (2013)
12. Cabalar, P., del Cerro, L.F., Pearce, D., Valverde, A.: A free logic for stable models with partial intensional functions. In: Logics in Artificial Intelligence - 14th European Conference, JELIA 2014, Funchal, Madeira, Portugal, September 24-26, 2014. Proceedings. (2014) 340–354
13. Fariñas del Cerro, L., Penttonen, M.: A note on the complexity of the satisfiability of modal horn clauses. J. of Logic Programming **4** (1987) 1–10

14. Fariñas del Cerro, L., Penttonen, M.: Intensional logics for programming (1991)
15. Fariñas del Cerro, L., Penttonen, M.: Grammar logics. Logique et Analyse (1988) 121–122
16. Lamarre, P.: Etude des raisonnements non-monotones : apports des logiques des conditionnels et des logiques modales. PhD thesis, Université Paul Sabatier, Toulouse (1992)
17. Crocco, G.: Fondements logiques du raisonnement contextuel. PhD thesis, Université Paul Sabatier, Toulouse (1993)
18. Crocco, G., Fariñas del Cerro, L., Herzig, A.: Conditionals: From philosophy to computer science (1995) Oxford University Press, Studies in Logic and Computation.
19. Besnard, P., Cordier, M.O., Dubois, D., Fariñas del Cerro, L., Froidevaux, C., Moinard, Y., Prade, H., Schwind, C., Siegel, P.: Raisonnement sur des informations incomplètes en Intelligence Artificielle - Comparaison de formalismes sur un exemple. Teknea, http://www.teknea.com (1989) Groupe Léa Sombé.
20. Besnard, P., Cordier, M.O., Dubois, D., Fariñas del Cerro, L., Froidevaux, C., Moinard, Y., Prade, H., Schwind, C., Siegel, P.: Reasoning under incomplete informatin in artificial intelligence: A comparison of formalisms using a single example. Wiley, http://www.wiley.com/ (1990) Groupe Léa Sombé.
21. Besnard, P., Cholvy, L., Cordier, M.O., Dubois, D., Fariñas del Cerro, L., Froidevaux, C., Lévy, F., Moinard, Y., Prade, H., Schwind, C., Siegel, P.: Revision and Updating in Knowledge Bases. Wiley, http://www.wiley.com/ (1994) Groupe Léa Sombé.
22. Audureau, E., Enjalbert, P., Fariñas del Cerro, L.: Logique temporelle sémantique et validation de programes parallèles. Masson (1990)
23. Soulhi, S.: Un système de déduction automatique pour la connaissance et la croyance. PhD thesis, Université Paul Sabatier, Toulouse (1985)
24. Bieber, P.: Aspects épistémiques des protocoles cryptographiques. PhD thesis, Université Paul Sabatier, Toulouse (1989)
25. Seguin, C.: De l'action à l'intention : vers une caractérisation formelle des agents. PhD thesis, Université Paul Sabatier, Toulouse (1992)
26. Longin, D.: Interaction rationnelle et évolution des croyances dans le dialogue : une logique basée sur la notion de topique. Thèse de doctorat, Université Paul Sabatier, Institut de Recherche en Informatique de Toulouse (IRIT), Toulouse, France (1999)
27. Chetcuti-Sperandio, N.: Dduction automatique en calcul des dures base sur la mthode des tableaux . Thèse de doctorat, Université Paul Sabatier, Toulouse, France (2001)
28. Fariñas del Cerro, L., Orlowska, E.: Automated reasoning in non-classical logic. Logique et Analyse (1985) 110–111 édition d'un numéro special.
29. Balbiani, P., del Cerro, L.F., Tinchev, T., Vakarelov, D.: Modal logics for incidence geometries. J. Log. Comput. **7**(1) (1997) 59–78
30. Condotta, J.F.: Problèmes de satisfaction de contraintes : algorithmes et complexité. Thèse de doctorat, Université Paul Sabatier, Institut de Recherche en Informatique de Toulouse (IRIT), Toulouse, France (2000)
31. Balbiani, P., Dugat, V., Fariñas del Cerro, L., Lopez, A.: Eléments de géométrie mécanique. Hermès, Paris (1994)
32. Lopez, A.: Déduction automatique en Géométrie par réduction de Figures. PhD thesis, Université Paul Sabatier, Toulouse (1995)
33. Masolo, C.: Critères de comparaison et de construction de théories axiomatiques pour la représentation de la connaissance : Ontologies de l'espace et du temps . Thèse de doctorat, Université Paul Sabatier - Universit degli studi di Padova, Toulouse - Padoue (2000)
34. Obeid, N.: MIM-Logic: a Logic for Reasoning About Molecular Interaction Maps. Thèse de doctorat, Université de Toulouse, Toulouse, France (2014)

Mixed Algebras and their Logics

Ivo Düntsch[1⋆], Ewa Orłowska[2⋆⋆], and Tinko Tinchev[3⋆⋆⋆]

[1] Brock University
St. Catharines, Ontario, Canada, L2S 3A1
duentsch@brocku.ca
[2] National Institute of Telecommunications
Szachowa 1, 04–894 Warsaw, Poland
orlowska@itl.waw.pl
[3] Faculty of Mathematics and Computer Science
Sofia University
Sofia, Bulgaria
tinko@fmi.uni-sofia.bg

Dedicated to Luis Fariñas del Cerro

Abstract. We investigate complex algebras of the form $\langle 2^X, \langle R\rangle, [[S]]\rangle$ arising from a frame $\langle X, R, S\rangle$ where $S \subseteq R$, and exhibit their abstract algebraic and logical counterparts.

1 Introduction

Semantics of non-classical logics is provided either in terms of a class of algebras or a class of relational systems (frames. The theme of finding an equivalent (i.e. validating the same formulas) frame (resp. and algebraic) semantics once an algebraic (resp. frame) semantics is given has an extensive literature. One part of the problem – passing from algebraic to frame semantics – is a subject of correspondence theory [1]. Correspondence theory is well developed for logics whose algebraic semantics is based on distributive lattices, possibly with additional operators, and require first order definable relations in the corresponding frames such as standard modal logics, intuitionistic logic, some intermediate and relevant logics. Less is known for logics based on not necessarily distributive lattices. Equivalence of the two semantics can be obtained from a discrete duality between the two underlying classes of structures [2].

In this paper we discuss the problem of finding a frame semantics for logics whose algebraic semantics is based on what we call *PS-algebras*. These are Boolean algebras endowed with a normal and additive operator (a *possibility operator*) and a co-normal and co-additive operator (a *sufficiency operator*).

A special class of PS-algebras are the *mixed algebras* (MIAs). These were introduced in [3] and further investigated in [4]. The possibility part and the sufficiency part are related to each other by a second order property expressed in terms of their respective canonical extensions. We provide an equivalent characterization of mixed algebras in terms of the relations in their canonical frames.

Mixed algebras are not first order definable, and the complex algebras of their corresponding frames are not necessarily MIAs, so, MIAs and their frames treated as semantic structures of a formal language do not provide equivalent semantics for that language. However, for some axiomatic extensions of PS-algebras there are frames such that the equivalence holds. We discuss two of such classes, namely, the class of right ideal MIAs and the class of weak MIAs (wMIAs).We provide several universal-algebraic properties of those classes, in particular, we exhibit the equational class generated by the class of wMIAs.

Furthermore, we present the logic $K^{\#}$ whose algebraic semantics is provided by the class wMIA, and its frame semantics by the class of wMIA-frames. This logic is based on the logic $K^{\sim}$ of [5], presented in a form which makes the connection to the algebras clearer. In turn, $K^{\sim}$ was developed based on the observation that the well known logic K as well as its sufficiency counterpart $K^{\star}$ presented in [6] are lacking in expressive power, and "necessity and sufficiency split the modal theory into two dual branches each of which spreads over less than a half of the Boolean realm" [5]. Finally,

⋆ Ivo Düntsch gratefully acknowledges support by the Natural Sciences and Engineering Research Council of Canada Discovery Grant 250153 and by the Bulgarian National Fund of Science, contract DID02/32/2009.

⋆⋆ Ewa Orłowska gratefully acknowledges partial support from the National Science Centre project DEC-2011/02/A/HS1/00395.

⋆⋆⋆ Tinko Tinchev gratefully acknowledges support by the Bulgarian National Fund of Science, contract DID02/32/2009.

using the copying technique of [7] and the concept of special models of [5], we show that one frame relation suffices for wMIA frames: If $\langle B, f, g\rangle$ is a wMIA, then there is a frame $\langle X, R\rangle$ such that $\langle B, f, g\rangle$ and a subalgebra of $\langle 2^X, \langle R\rangle, [[R]]\rangle$ satisfy the same equations.

2 General definitions and notation

To make the paper more self–contained we recall a few concepts from Universal Algebra. Readers familiar with these concepts may skip straight to Section 3. Let $\mathfrak{F}$ be a signature of algebras, and X be a set of variables. The *set* $T_{\mathfrak{F}}(X)$ *of* $\mathfrak{F}$ *terms over* X is the smallest set such that

1. $X \subseteq T_{\mathfrak{F}}(X)$,
2. Each constant is in $T_{\mathfrak{F}}(X)$,
3. If $t_1, \dots, t_n \in T_{\mathfrak{F}}(X)$ and $f \in \mathfrak{F}$ is $n-$ ary, then $f(t_1, \dots, t_n) \in T_{\mathfrak{F}}(X)$.

In the sequel we assume that $\mathfrak{F}$ is fixed, and we shall just write $T(X)$; we also assume that $T(X) \neq \emptyset$. Furthermore, $T(X)$ will be regarded as the absolutely free algebra over X with signature $\mathfrak{F}$, see [8, p.68].

If t is a term, we write $t(x_1, \dots, x_n)$ if the variables occurring in t are among $x_1, \dots, x_n$. Suppose that $\mathfrak{A}$ is an algebra of type $\mathfrak{F}$. If $t(x_1, \dots, x_n) \in T(X)$, the *term function* $t^{\mathfrak{A}} : \mathfrak{A}^n \to \mathfrak{A}$ is defined as follows:

T_1. If t is the variable x_i, then $t^{\mathfrak{A}}(a_1, \dots, a_n) = a_i$.
T_2. If $f \in \mathfrak{F}$ is $k-$ ary and t has the form $f(t_1(x_1, \dots, x_n), \dots, t_k(x_1, \dots, x_n))$, then

$$t^{\mathfrak{A}}(a_1, \dots, a_n) = f^{\mathfrak{A}}(t_1^{\mathfrak{A}}(a_1, \dots, a_n), \dots, t_k^{\mathfrak{A}}(a_1, \dots, a_n)). \tag{2.1}$$

$t^{\mathfrak{A}}$ is called the *term function* of t (over $\mathfrak{A}$). For later use we mention the following fact:

Lemma 1. *[8, Theorem 10.3] Let $\mathfrak{A}, \mathfrak{B}$ be algebras of the same type and $t(x_1, \dots, x_n)$ be an $n-$ ary term.*

1. *Suppose that $a_i, b_i \in A$ for $1 \le i \le n$ and θ is a congruence on $\mathfrak{A}$. If $a_i \theta b_i$ for all $1 \le i \le n$, then*
$$t^{\mathfrak{A}}(a_1, \dots, a_n) \theta t^{\mathfrak{A}}(b_1, \dots, b_n).$$
2. *If $f : \mathfrak{A} \to \mathfrak{B}$ is a homomorphism, then*
$$f(t^{\mathfrak{A}}(a_1, \dots, a_n)) = t^{\mathfrak{B}}(f(a_1), \dots, f(a_n)).$$

An *equation* (or *identity*, see [8, Definition 11.1]) is an expression of the form $\tau \approx \sigma$, where $\tau, \sigma \in T(X)$. If $\tau, \sigma \in T(X)$ are $n-$ ary and $a_1, \dots, a_n \in A$, then the tuple $\langle a_1, \dots, a_n\rangle$ *satisfies the equation* $\tau \approx \sigma$ if $\tau^{\mathfrak{A}}(a_1, \dots, a_n) = \sigma^{\mathfrak{A}}(a_1, \dots, a_n)$. If $\tau^{\mathfrak{A}}(a_1, \dots, a_n) = \sigma^{\mathfrak{A}}(a_1, \dots, a_n)$ for all tuples $\langle a_1, \dots, a_n\rangle \in A^n$, we say that $\tau \approx \sigma$ *is true in* $\mathfrak{A}$, written as $\mathfrak{A} \models \tau \approx \sigma$.

As no generality is lost, we shall tacitly assume that a class of algebras is closed under isomorphic copies. If $\mathbf{K}$ is a class of algebras of the same type we denote by $\mathbf{H}(\mathbf{K})$ the collection of all homomorphic images of $\mathbf{K}$, by $\mathbf{S}(\mathbf{K})$ the collection of all subalgebras of $\mathbf{K}$, and by $\mathbf{P}(\mathbf{K})$ the collection of all products of elements of $\mathbf{K}$. The equational class $\mathbf{HSP}(\mathbf{K})$ generated by $\mathbf{K}$ is denoted by $\mathbf{Eq}(\mathbf{K})$. $Con(\mathfrak{A})$ is the set of all congruences on the algebra $\mathfrak{A}$.

Suppose that $\mathfrak{B} = \langle B, \wedge, \vee, \neg, 0, 1\rangle$ is a Boolean algebra. With some abuse of language we will usually identify algebras with their base set if no confusion can arise. Note that $a = b$ if and only if $\neg((a \wedge \neg b) \vee (b \wedge \neg a)) = 1$, and thus for each equation $\tau \approx \sigma$ there is an equation $\tau' \approx 1$ such that $B \models \tau \approx \sigma$ if and only if $B \models \tau' \approx 1$.

If $A \subseteq B$ and $f : B \to B$ is a function, then $f[A] = \{f(a) : a \in A\}$ is the *image of* A *under* f. The *dual of* f is the mapping $f^{\partial} : B \to B$ defined by $f^{\partial}(a) = \neg f(\neg a)$.

For the background of universal algebra we refer the reader to [8] and for frame and algebraic semantics of modal logics to [9] or [10].

3 Possibility and sufficiency algebras

In this section we review the concepts of possibility and sufficiency algebras and their canonical extensions.

Traditionally, a *modality* – or an *operator* [11] – on a Boolean algebra is a function $f : B \to B$ which satisfies $f(0) = 0$ (normal), and $f(a \vee b) = f(a) \vee f(b)$ (additive) for all $a, b \in B$. In recent years, however, many more operators with different properties have been considered in the study of modal logics, so that the term may mean almost any intensional operator on B. In this paper we shall be concerned with the two modalities *possibility* and *sufficiency* as well as operators definable from these and the Boolean operators.

A *possibility operator* on B is a normal and additive function $f : B \to B$; its dual f^∂ is called a *necessity operator*. Clearly, a mapping $g : B \to B$ is a necessity operator if and only if $g(1) = 1$ and $g(a \wedge b) = g(a) \wedge g(b)$ for all $a, b, \in B$. If f is a possibility operator on B, the pair $\langle B, f \rangle$ is called a *possibility algebra*. Dually, if u is a necessity operator on B, the pair $\langle B, u \rangle$ is called a *necessity algebra*.

A *sufficiency operator* on B is a function $g : B \to B$ which satisfies $g(0) = 1$ (co–normal), and $g(a \vee b) = g(a) \wedge g(b)$ (co–additive) for all $a, b \in B$. If g is a sufficiency operator on B, the pair $\langle B, g \rangle$ is called a *sufficiency algebra*. To the best of our knowledge, sufficiency operators were first introduced to modal logic by [12]. In some sense, a sufficiency operator is the "complementary counterpart" to a possibility operator. This will be made clearer in the next section.

For a Boolean algebra B, we let $B^c = 2^{\mathrm{Ult}(B)}$ be its canonical extension [11], and $h : B \hookrightarrow B^c$ be the Stone embedding, i.e. $h(a) = \{F \in \mathrm{Ult}(B) : a \in F\}$. If $f, g : B \to B$ are operators on B, then two canonical extensions $f^\sigma, g^\pi : B^c \to B^c$ are defined by

$$f^\sigma(a) = \bigcup\{\bigcap\{h(f(x)) : x \in F\} : F \in a\}, \tag{3.1}$$
$$g^\pi(a) = \bigcap\{\bigcup\{h(g(x)) : x \in F\} : F \in a\}. \tag{3.2}$$

In particular, if $F \in \mathrm{Ult}(B)$, then

$$f^\sigma(\{F\}) = \bigcap\{h(f(x)) : x \in F\}, \tag{3.3}$$
$$g^\pi(\{F\}) = \bigcup\{h(g(x)) : x \in F\}. \tag{3.4}$$

There are representation theorems both for possibility and sufficiency algebras:

Theorem 1. *Suppose that B is a Boolean algebra.*

1. *[11] If f is a possibility operator on B, then, f^σ is a possibility operator on B^c, and the Stone mapping $h : \langle B, f \rangle \hookrightarrow \langle B^c, f^\sigma \rangle$ is an embedding.*
2. *[3] If g is a sufficiency operator on B, then, g^π is a sufficiency operator on B^c, and the Stone mapping $h : \langle B, g \rangle \hookrightarrow \langle B^c, g^\pi \rangle$ is an embedding.*

In particular, $h(f(a)) = f^\sigma(h(a))$ and $h(g(a)) = g^\pi(h(a))$ for all $a \in B$.

If f is a possibility operator on B and g a sufficiency operator, we call the structure $\langle B, f, g \rangle$ a *PS–algebra*, and $\langle B^c, f^\sigma, g^\pi \rangle$ its canonical extension. Theorem 1 tells us that $\langle B^c, f^\sigma, g^\pi \rangle$ is a PS–algebra, and h is an embedding of PS–algebras. For the rest of this section, we suppose that $\langle B, f, g \rangle$ is a PS–algebra.

If g is an operator on B we let $g^*(a) = g(\neg a)$. Note that g^* and g are mutually term definable. Furthermore, g is a sufficiency operator if and only if g^* is a necessity operator.

Theorem 2. *There is a 1 – 1 correspondence between PS – congruences on B and (Boolean) filters which are closed under f^∂ and g^*.*

Proof. It is well known (see e.g. [13]) that each Boolean congruence θ is uniquely determined by a filter F_θ, where

$$F_\theta = \{a \in B : a\theta 1\},$$

and, conversely, each filter F uniquely determines a congruence θ_F on B by

$$a\theta_F b \iff (\exists t)[t \in F \text{ and } a \wedge t = b \wedge t].$$

Furthermore, it was shown in [14] that a Boolean congruence θ preserves a necessity operator m if and only if F_θ is closed under m, i.e. $a \in F_\theta$ implies $m(a) \in F_\theta$. Clearly, θ preserves f if and only if it preserves f^∂, and θ preserves g if and only if it preserves g^*. Since both f^∂ and g^* are necessity operators, the claim follows.

Define a mapping $u : B \to B$ by

$$u(a) = f^{\partial}(a) \wedge g^{*}(a) = f^{\partial}(a) \wedge g(\neg a). \tag{3.5}$$

Since both f^{∂} and g^{*} are necessity operators, so is u.

A filter F of B is called a *u – filter*, if $a \in F$ implies $u(a) \in F$ for all $a \in B$. Theorem 2 now immediately implies

Corollary 1. *There is a 1 – 1 correspondence between congruences on B and u – filters.*

Proof. Let θ be a congruence on $\langle B, f, g\rangle$; then, F_θ is closed under f^{∂} and g^{*} by Theorem 2. Thus, if $a \in F_\theta$, then $f^{\partial}(a), g^{*}(a) \in F_\theta$, and thus, $u(a) = f^{\partial}(a) \wedge g^{*}(a) \in F_\theta$, since F_θ is a filter.

Conversely, let F be a u – filter and $a \in F$. Then, $u(a) = f^{\partial}(a) \wedge g^{*}(a) \in F$ by the hypothesis, and thus, $f^{\partial}(a), g^{*}(a) \in F$ since F is a filter. Hence, θ_F is a PS – congruence, again by Theorem 2.

4 Algebras and frames

The set of all binary relations on a set X is denoted by $\mathrm{Rel}(X)$; if $R_1, \ldots \in \mathrm{Rel}(X)$, the structure $\langle X, R_1, \ldots\rangle$ is called a *frame*. For $x \in X$, we let $R(x) = \{z \in X : xRz\}$. Relational composition and converse are denoted by ; , respectively by $\breve{}$; furthermore, $1'$ is the identity relation.

For $R \in \mathrm{Rel}(X)$, we define two operators on 2^X by

$$\langle R\rangle(S) = \{x : (\exists y)[xRy \text{ and } y \in S]\} = \{x : R(x) \cap S \neq \varnothing\}. \tag{4.1}$$

$$[[R]](S) = \{x : (\forall y)[y \in S \Rightarrow xRy]\} = \{x : S \subseteq R(x)\}. \tag{4.2}$$

We also set

$$[R](S) = \langle R\rangle^{\partial}(S) = \{x : R(x) \subseteq S\}. \tag{4.3}$$

It is well known that $\langle R\rangle$ is a complete possibility operator on the power set algebra of X [11], and that $[[R]]$ is a complete sufficiency operator [3]. Note that

$$[[R]](S) = [-R](X \setminus S), \tag{4.4}$$

so that it may be said that $[R]$ talks about the properties of R, while $[[R]]$ talks about the properties of $-R$ (see also the discussion in [12]).

The structure $\langle 2^X, \langle R\rangle\rangle$ is called the *full possibility (P) complex algebra of* $\langle X, R\rangle$, denoted by $\mathfrak{Cm}_P(X, R)$ or just by $\mathfrak{Cm}_P(X)$ if R is understood. Similarly, $\mathfrak{Cm}_S(X) = \langle 2^X, [[R]]\rangle$ is the *full sufficiency (S) complex algebra of* $\langle X, R\rangle$, and $\mathfrak{Cm}_{PS}(X) = \langle 2^X, \langle R\rangle, [[R]]\rangle$ is the *full PS - complex algebra of* $\langle X, R\rangle$. A *P (S, PS) complex algebra* is an algebra (isomorphic to) a subalgebra of some $\langle 2^X, \langle R\rangle\rangle$ ($\langle 2^X, [[R]]\rangle$, $\langle 2^X, \langle R\rangle, [[R]]\rangle$).

The question arises whether the canonical extension of a possibility or a sufficiency algebra is isomorphic to a structure $\langle 2^U, \langle R\rangle\rangle$ or $\langle 2^U, [[R]]\rangle$ for some frame $\langle U, R\rangle$. In both cases, the answer is positive, and the relation in question is uniquely determined:

Theorem 3. *1. [11, Theorem 3.10] If $\langle B, f\rangle$ is a possibility algebra, then there is, up to isomorphism, a unique relation R_f on $\mathrm{Ult}(B)$ such that $\langle R_f\rangle = f^{\sigma}$. This relation is defined by*

$$FR_fG \Longleftrightarrow F \in f^{\sigma}(\{G\}). \tag{4.5}$$

The structure $\langle \mathrm{Ult}(B), R_f\rangle$ is called the P – canonical frame of $\langle B, f\rangle$.

2. [3, Proposition 7] If $\langle B, g\rangle$ is a sufficiency algebra, then there is, up to isomorphism, a unique relation R_g on $\mathrm{Ult}(B)$ such that $[[R_g]] = g^{\pi}$. This relation is defined by

$$FR_gG \Longleftrightarrow F \in g^{\pi}(\{G\}). \tag{4.6}$$

The structure $\langle \mathrm{Ult}(B), R_g\rangle$ is called the S – canonical frame of $\langle B, g\rangle$.

The *PS – canonical frame* of a PS–algebra $\langle B, f, g\rangle$ is the structure $\langle \mathrm{Ult}(B), R_f, R_g\rangle$. Theorem 1 and Theorem 3 together now give us the following result:

Theorem 4. *Every possibility (sufficiency, PS) algebra is embeddable into the full complex algebra of its canonical frame.*

This shows that the variety of PS–algebras is canonical in the sense of [15].

Finally in this section we mention an alternative description of the relations R_f and R_g of (4.5), respectively, (4.6) which does not explicitly use the canonical extension of B:

Lemma 2. *1.* $\langle F, G\rangle \in R_f \iff f[G] \subseteq F$.
2. $\langle F, G\rangle \in R_g \iff F \cap g[G] \neq \varnothing$.

Proof. 1. "⇒": This has been known for some time, see e.g. [16]. Suppose that $\langle F, G\rangle \in R_f$, i.e. $F \in f^\sigma(\{G\})$. Then, for all $x \in G$, $f(x) \in F$ by (3.3), which implies $f[G] \subseteq F$.

"⇐": Suppose $f[G] \subseteq F$; we need to show that $F \in \bigcap\{h(f(x)) : x \in G\}$. Let $x \in G$; then, $f(x) \in F$ by our hypothesis, and thus, $F \in h(f(x))$.

2. "⇒": Let $\langle F, G\rangle \in R_g$, i.e. $F \in g^\pi(\{G\})$. By (3.4), there is some $x \in G$ such that $g(x) \in F$, in other words, $F \cap g[G] \neq \varnothing$.

"⇐": Let $F \cap g[G] \neq \varnothing$, say, $x \in G$ and $g(x) \in F$. Then, $F \in h(g(x)) \subseteq \bigcup\{h(f(y)) : y \in G\} = g^\pi(\{G\})$.

5 The class MIA

Suppose that $\langle B, f, g\rangle$ is a PS–algebra. In the general definition, there is no relation between f and g, and between their associated canonical frames $\langle \mathrm{Ult}(B), R_f\rangle$ and $\langle \mathrm{Ult}(B), R_g\rangle$. Of course, such connections may exist: Consider, for example, the condition

$$f(a) = \neg g(a). \tag{5.1}$$

It is not hard to see that the corresponding frame $\langle \mathrm{Ult}(B), R_f, R_g\rangle$ satisfies the condition

$$R_f = \mathrm{Ult}(B)^2 \setminus R_g, \tag{5.2}$$

and that the respective representations for algebras satisfying (5.1) and frames satisfying (5.2) hold (see also Proposition 8 of [3]).

While the possibility algebras are the algebraic models of the logic K and the sufficiency algebras are the algebraic models of its sufficiency counterpart $K^\star$ [6], both are limited in their powers of expression if considered separately. For example, $\langle 2^X, \langle R\rangle\rangle$ can express reflexivity by

$$R \text{ is reflexive } \iff Y \subseteq \langle R\rangle(Y),$$

but it cannot express irreflexivity of R. On the other hand, $\langle 2^X, [[R]]\rangle$ can express irreflexivity by

$$R \text{ is irreflexive } \iff [[R]](Y) \subseteq -Y,$$

but not reflexivity. Neither $\langle 2^X, \langle R\rangle\rangle$ nor $\langle 2^X, [[R]]\rangle$ can express antisymmetry on its own, but together they can [4]:

$$R \text{ is antisymmetric } \iff \langle R\rangle([[R]](-Y) \cap Y) \subseteq Y.$$

Thus, it is worthwhile to consider the PS–algebras $\langle 2^X, \langle R\rangle, [[R]]\rangle$ obtained from a frame $\langle X, R\rangle$ with a single distinguished relation. Let us denote the class of complex algebras of this form by **CMIA**.

Next, let us step back and consider a PS–algebra $\mathfrak{B} = \langle B, f, g\rangle$ as a starting point. In [3], $\mathfrak{B}$ was called a *mixed algebra* (MIA), if in its PS – canonical frame $\langle \mathrm{Ult}(B), R_f, R_g\rangle$, the relations R_f and R_g were equal, and therefore, the full complex algebra of its canonical frame was of the form $\langle 2^{\mathrm{Ult}(B)}, \langle R\rangle, [[R]]\rangle$, where $R = R_f = R_g$; in other words, it is in **CMIA**. The following result now follows immediately from Theorem 3:

Theorem 5. *[11,3] Let $\langle B, f, g\rangle$ be a PS–algebra. Then, there is a relation R on $\mathrm{Ult}(B)$ such that $\langle R\rangle = f^\sigma$ and $[[R]] = g^\pi$ if and only if $f^\sigma(\{G\}) = g^\pi(\{G\})$ for all $G \in \mathrm{Ult}(B)$. Furthermore, the relation R is unique with this property.*

The class of mixed algebras is denoted by **MIA**. Note that the MIA condition $R_f = R_g$ is a second order axiom. Indeed, it was shown in [4] that **MIA** is not first order axiomatizable. Observe that $\mathfrak{B}$ is a MIA if and only if for all $F, G \in \mathrm{Ult}(B)$,

$$f[G] \subseteq F \iff F \cap g[G] \neq \varnothing \tag{5.3}$$

by Lemma 2.

Starting with a MIA leads to a canonical frame $\langle \mathrm{Ult}(B), R_f, R_g \rangle$ with $R_f = R_g$. On the other hand, using a frame $\langle X, R \rangle$ as a starting point and considering the complex algebra $\langle 2^X, \langle R \rangle, [[R]] \rangle$ will not necessarily lead to a MIA since not every algebra in **CMIA** is in **MIA**, as the following example shows:

Example 1. This is based on Proposition 14 of [3]: Let X be infinite, and $R = 1'$. If $\mathfrak{Cm}_{PS}(X)$ is a MIA, then, by (4.5) and (4.6), we must have $R_{\langle R \rangle} = R_{[[R]]}$.

Suppose that F, G are ultrafilters of 2^X. Since R is the identity relation on X, $\langle R \rangle(a) = a$ for all $a \subseteq X$, hence, $\langle R \rangle[G] \subseteq F$ if and only if $F = G$. Suppose that $a \in G$, $|a| > 1$. Then, $x \in [[R]](a) \Longleftrightarrow a \subseteq R(x) = \{x\}$, and it follows that $[[R]](a) = \emptyset$. Thus, if G is non principal, then $G \cap [[R]][G] = \emptyset$ and it follows that $\langle 2^X, \langle R \rangle, [[R]] \rangle$ does not satisfy (5.3).

Similarly, if $R = (X \times X) \setminus 1'$, then $\langle R \rangle(a) = X$ for all a with $|a| > 1$, and thus, $\langle R \rangle[G] \subseteq F$ for all non–principal $G \in \mathrm{Ult}(2^X)$ and all $F \in \mathrm{Ult}(2^X)$; in particular, $\langle R \rangle[G] \subseteq G$. On the other hand, $[[R]](a) = X \setminus a$ for all $a \subseteq X$, so that $G \cap [[R]][G] = \emptyset$. □

Thus, not every PS – complex algebra of a structure $\langle X, R \rangle$ is a MIA, and we cannot have a general discrete duality theorem between PS -frames $\langle X, R, R \rangle$ and canonical frames of complex algebras of $\langle 2^X, \langle R \rangle, [[R]] \rangle$.

It is unknown which class of frames $\langle X, R \rangle$ have a full PS–complex algebra in **MIA**. A general characterization needs to be second order, since **MIA** is not first order axiomatizable. The only general property we know which leads to a MIA is that of right ideal frames. Set $\mathbf{1} = X \times X$. A binary relation R on X is called a *right ideal relation*, if $R \,;\, \mathbf{1} \subseteq R$, and the pair $\langle X, R \rangle$ is called a *right ideal frame*. The following observation is already (implicitly) contained in [17], p. 79:

Lemma 3. *R is a right ideal relation if and only if $\langle R \rangle(X) = [[R]](X)$.*

Proof. "⇒": Let $x \in \langle R \rangle(X)$; then, $R(x) \neq \emptyset$. If, say, xRy and $z \in X$, then $xRy\mathbf{1}z$, and $R \,;\, \mathbf{1} \subseteq R$ implies that xRz. Hence, $X \subseteq R(x)$, and thus, $x \in [[R]](X)$. The other direction follows from Lemma 5 below.

"⇐": Suppose that xRy and $z \in X$; we need to show that xRz. Since xRy, we have $R(x) \neq \emptyset$, hence, $x \in \langle R \rangle(X)$. The hypothesis implies that $x \in [[R]](X)$, hence, $X \subseteq R(x)$; in particular, xRz.

A PS–algebra $\langle B, f, g \rangle$ is called a *right ideal algebra* if $f(1) = g(1)$.

Lemma 4. *A right ideal algebra $\langle B, f, g \rangle$ is a MIA.*

Proof. We have to show the "⇒" direction of (5.3): Suppose that F, G are ultrafilters of B, and that $f[G] \subseteq F$. Then, in particular, $f(1) \in F$, and thus, $g(1) \in F$ since B is a right ideal algebra. Now, $1 \in G$ implies that $F \cap g[G] \neq \emptyset$.

Since the complex algebra of a right ideal frame is a right ideal algebra by Lemma 3, we immediately obtain

Theorem 6. *The PS – complex algebra of a right ideal frame $\langle X, R \rangle$ is a right ideal algebra.*

The other part of the duality also holds:

Theorem 7. *The PS – canonical frame of a right ideal algebra $\langle B, f, g \rangle$ is a right ideal frame.*

Proof. Let $X = \mathrm{Ult}(B)$. In view of Lemma 5 below it suffices to show that $\langle R \rangle(X) \subseteq [[R]](X)$. Let $F \in \langle R \rangle(X)$; then, there is some $G \in X$ such that FRG. Since B is a MIA, $R = R_f$, and thus, $f[G] \subseteq F$, in particular, $f(1) \in F$. Since B is a right ideal algebra it follows that $g(1) \in F$ as well. We need to show that $F \in [[R]](X)$, in other words that $X \subseteq R(F)$. Let $H \in X$; then, $1 \in H$ and $g(1) \in F$ shows that $F \cap g[H] \neq \emptyset$, hence, FR_gH.

6 The class wMIA

As the class **MIA** is too narrow to fully describe the properties of the class **CMIA**, let us start with the properties of $\langle 2^X, \langle R \rangle, [[R]] \rangle \in$ **CMIA**. The following observation shows how these algebras differ from MIAs:

Lemma 5. *1. For all $x \in X$,*

$$\langle R \rangle(\{x\}) = [[R]](\{x\}). \tag{6.1}$$

2. Let $A, B \subseteq X$ such that $A \cap B \neq \emptyset$. Then, $[[R]](A) \subseteq \langle R \rangle(B)$.

Proof. 1. "$\subseteq$": let $y \in \langle R \rangle(\{x\})$, i.e. yRx. Then, $\{x\} \subseteq R(y)$, which shows that $y \in [[R]](\{x\})$.

"$\supseteq$": $y \in [[R]](\{x\})$. Then, $\{x\} \subseteq R(y)$, and thus yRx. It follows that $y \in \langle R \rangle(\{x\})$.

2. Let $x \in A \cap B$; then, $\{x\} \subseteq A \cap B$. Since $[[R]]$ is a sufficiency operator, we have $[[R]](A) \subseteq [[R]](\{x\}$, and the fact that $\langle R \rangle$ is a possibility operator implies $\langle R \rangle(\{x\} \subseteq \langle R \rangle(B)$. The conclusion now follows from (6.1).

Note that $\langle R \rangle(\{x\}) = [[R]](\{x\})$ only implies that in the canonical extension of $\mathfrak{Cm}_{PS}(X)$ we obtain that $\langle R \rangle^{\sigma}(F) = [[R]]^{\pi}(F)$ only for principal ultrafilters F of $\mathfrak{Cm}_{PS}(X)$. Example 1 shows that it need not hold for non–principal ultrafilters.

These observations lead to the following definition: A *weak mixed algebra* (wMIA) is a PS–algebra $\langle B, f, g \rangle$ such that

$$(\forall a, b)[a \wedge b \neq 0 \Rightarrow g(a) \leq f(b)]. \tag{6.2}$$

We shall denote the class of weak MIAs by **wMIA**. Note that, unlike **MIA**, the class **wMIA** is first order axiomatizable, indeed, it is a universal class. There are several characterizations of weak MIAs:

Theorem 8. *Let $\langle B, f, g \rangle$ be a PS–algebra. The following are equivalent:*

1. *B is a weak MIA.*
2. *$R_g \subseteq R_f$.*
3. *$g^{\pi}(\{F\}) \subseteq f^{\sigma}(\{F\})$ for all $F \in \mathrm{Ult}(B)$.*
4. *$(\forall a \in B)[a \neq 0 \Rightarrow g(a) \leq f(a)]$.*

Proof. 1. $\Rightarrow$ 2.: Let $F \cap g[G] \neq \emptyset$ and $a \in G$ with $g(a) \in F$. Suppose that $b \in G$; since $a \in G$ as well, we have $a \wedge b \neq 0$. It follows from (6.2) that $g(a) \leq f(b)$, and $g(a) \in F$ now implies that $f(b) \in F$.

2. $\Rightarrow$ 3.: This follows immediately from the definitions of R_f and R_g in (4.5) and (4.6).

3. $\Rightarrow$ 1: Suppose that $a \wedge b \neq 0$, and assume that $g(a) \nleq f(b)$, i.e. $g(a) \wedge \neg f(b) \neq 0$. Then, there are ultrafilters F, G such that $g(a), \neg f(b) \in F$ and $a, b \in G$. Then, $F \cap g[G] \neq \emptyset$, and thus, it follows from Lemma 2(2) and the definition of R_g that $F \in g^{\pi}(\{G\})$. Then, by the hypothesis, $F \in f^{\sigma}(\{G\})$, and it follows from the definition of R_f and Lemma 2(1) that $f[G] \subseteq F$. Since $b \in G$ it follows that $f(b) \in F$, contradicting that $\neg f(b) \in F$.

Finally, we show that 4. $\Rightarrow$ 1., the other direction being trivial: Suppose that 4. holds, and that $a \wedge b \neq 0$. Then, since g is antitone and f is isotone,

$$g(a) \leq g(a \wedge b) \overset{4.}{\leq} f(a \wedge b) \leq f(b).$$

This completes the proof.

Observe that Theorem 8(3) shows that every MIA is a weak MIA. Since $\mathfrak{Cm}_{PS}(\langle X, R \rangle)$ is a weak MIA, Theorem 8(2) shows that for all ultrafilters F, G of 2^X in a weak MIA

$$F \cap [[R]][G] \neq \emptyset \Rightarrow \langle R \rangle[G] \subseteq F. \tag{6.3}$$

Theorem 8(2) suggests that we call a PS–frame $\langle X, R, S \rangle$ a *weak MIA frame*, if $S \subseteq R$. Even though we use two relations, we have a connection between R and S by $S \subseteq R$ which is one direction of the MIA condition. Our next result shows the correspondence between weak MIA frames and weak MIAs:

Lemma 6. *1. The complex algebra of a weak MIA frame is a weak MIA.*
2. The canonical frame of a weak MIA is a weak MIA frame.

Proof. 1. Suppose that $\langle X, R, S \rangle$ is a weak MIA frame. Let $\emptyset \neq Y \subseteq X$ and $x \in [[S]](Y)$. By 8(4) it is sufficient to show $x \in \langle R \rangle(Y)$. Since $x \in [[S]](Y)$, we obtain $Y \subseteq S(x)$, and therefore, $Y \subseteq R(x)$ by the hypothesis. It now follows from $Y \neq \emptyset$ that $R(x) \cap Y \neq \emptyset$, hence, $x \in \langle R \rangle(Y)$.

2. Suppose that $\langle B, f, g \rangle$ is a weak MIA. By Theorem 8(3) $\langle \mathrm{Ult}(B), R_f, R_g \rangle$ is a weak MIA frame.

This gives us the representation theorem:

Theorem 9. *1. Each weak MIA frame is embeddable into the canonical frame of its complex algebra.*
2. Each weak MIA is embeddable into the complex algebra of its canonical frame.

wMIA is closed under subalgebras and homomorphic images, but not under products, as we shall see below.

7 The logic $K^\#$

In this section we shall exhibit a logic whose algebraic models are weak MIAs; it is inspired by the logic $K^{\sim}$ of [5]. The logic $K^\#$ is a Boolean logic with operators $\wedge, \neg, \top$, a countable set Var of propositional variables, and two additional unary modalities $\Box$ and $\boxplus$ with duals $\Diamond$ and $\diamondplus$. Formulas are recursively defined as usual:

1. $\top$ is a formula.
2. Each $p \in \mathsf{Var}$ is a formula.
3. If φ, ψ are formulas, so are $\neg\varphi, \varphi \wedge \psi, \Box\varphi, \boxplus\varphi$.
4. No other string is a formula.

The set of formulas of $K^\#$ is denoted by Fml. If the variables occurring in a formula φ are among $p_1, \ldots, p_n$, we indicate this by writing $\varphi(p_1, \ldots, p_n)$. We use the usual definitions of the Boolean connectives $\bot, \vee, \rightarrow, \leftrightarrow$. The axiom system for the modal part of $K^\#$ is as follows:

$$K - \text{part } \Box\text{:} \begin{cases} \vdash \Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi), \\ \text{If } \vdash \varphi \text{ then } \vdash \Box\varphi. \end{cases} \tag{7.1}$$

$$K^\star - \text{part } \boxplus\text{:} \begin{cases} \vdash \boxplus\neg(\varphi \rightarrow \psi) \rightarrow (\boxplus\neg\varphi \rightarrow \boxplus\neg\psi) \\ \text{If } \vdash \varphi, \text{ then } \vdash \boxplus\neg\varphi. \end{cases} \tag{7.2}$$

$$\text{Connection:} \begin{cases} \text{If } \underbrace{\nvdash (\varphi \wedge \psi) \rightarrow \bot \text{ then } \vdash \boxplus\varphi \rightarrow \Diamond\psi.}_{\textbf{wMIA}} \end{cases} \tag{7.3}$$

If $\Sigma \subseteq \mathsf{Fml}$ and $\varphi \in \mathsf{Fml}$, then a *proof* (or *derivation*) of φ from Σ is a finite sequence $\varphi_1, \ldots, \varphi_n$ of formulas such that $\varphi_n = \varphi$, and each $\varphi_i \in \Sigma$ is the result of an application of a rule to one or more formulas occurring earlier in the sequence. In this case, we write $\Sigma \vdash \varphi$. If $\Sigma = \varnothing$, then we call φ a theorem of $K^\#$, and write $K^\# \vdash \varphi$ or simply $\vdash \varphi$. A set Σ of formulas is called *deductively closed* if and only if $\Sigma \vdash \varphi$ implies $\varphi \in \Sigma$ for all $\varphi \in \mathsf{Fml}$.

7.1 Frame semantics

Frame models have the form $M = \langle X, R, S, v\rangle$ where $S \subseteq R \subseteq X \times X$, and $v : \mathsf{Var} \rightarrow 2^X$ is a valuation (often called *meaning function*) over the propositional variables which is extended over the Boolean operators in the usual way. With respect to the modal operators, v acts as follows:

$$x \in v(\Box\varphi) \Longleftrightarrow R(x) \subseteq v(\varphi), \qquad [R] \tag{7.4}$$

$$x \in v(\boxplus\varphi) \Longleftrightarrow v(\varphi) \subseteq S(x). \qquad [[S]] \tag{7.5}$$

The *base of a model* $M = \langle X, R, S, v\rangle$ is the frame $\langle X, R, S\rangle$; note that the base of a model is a weak MIA frame.

We say that a formula φ *is satisfied in* M *at* $x \in X$ *with respect to* v, written as $x \models_v \varphi$, if $x \in v(\varphi)$. φ is called *valid in* M, written as $M \models_v \varphi$ if $x \models_v \varphi$ for all $x \in X$, i.e. if $v(\varphi) = X$. If $\langle W, R, S\rangle$ is the base of a model of $K^\#$ we say that φ *is true in* $\langle W, R, S\rangle$, written as $\langle W, R, S\rangle \models \varphi$, if $\langle W, R, S, v\rangle \models_v \varphi$ for all valuations based on $\langle W, R, S\rangle$. If φ is true in all models, we write $K^\# \models \varphi$.

If $\langle W, R, S\rangle$ is the base of a model of $K^\#$, we consider its complex algebra $\langle 2^W, \langle R\rangle, [[S]]\rangle$. By Lemma 6, $\langle 2^W, \langle R\rangle, [[S]]\rangle \in \textbf{wMIA}$. For a PS–algebra $\langle B, f, g\rangle$, the structure $\langle \mathrm{Ult}(B), R_f, R_g\rangle$ is a base of a model of $K^\#$ if and only if B is a weak MIA, since $S \subseteq R$ in a model of $K^\#$.

Theorem 10. *$K^\#$ is sound and complete with respect to the class of models based on weak MIA frames.*

Proof. Soundness is straightforward, so we just concentrate on completeness. Let W be the set of all maximal consistent theories of $K^\#$. For each $t \in W$ let

$$\Diamond t = \{\Diamond\varphi : \varphi \in t\}, \tag{7.6}$$

$$\boxplus_\neg(t) = \{\varphi : \boxplus\neg\varphi \in t\}, \tag{7.7}$$

and for $t_1, t_2 \in W$, let

$$t_1 R t_2 \iff \Diamond t_2 \subseteq t_1 \tag{7.8}$$

$$t_1 S t_2 \iff t_1 \cap \{\boxminus\varphi : \varphi \in t_2\} \neq \varnothing. \tag{7.9}$$

Claim 1: $\Diamond t_2 \subseteq t_1$ if and only if $\{\varphi : \Box\varphi \in t_1\} \subseteq t_2$:

"$\Rightarrow$": Let $\Box\varphi \in t_1$, and assume that $\varphi \notin t_2$. Then, $\neg\varphi \in t_2$ by maximality of t_2. The hypothesis implies that $\Diamond\neg\varphi \in t_1$, and therefore, $\neg\Box\varphi \in t_1$ by the definition of $\Diamond$. This contradicts the consistency of t_1.

"$\Leftarrow$": Let $\varphi \in t_2$ and assume that $\Diamond\varphi \notin t_1$. Then, $\neg\Diamond\varphi \in t_1$ by maximality of t_1, and therefore, $\Box\neg\varphi \in t_1$ by definition of $\Diamond$. It follows from the hypothesis that $\neg\varphi \in t_2$, contradicting the consistency of t_2. □

Claim 2: $\boxminus_\neg(t_1) \subseteq t_2$ if and only if $t_1 \cap \{\boxminus\varphi : \varphi \in t_2\} = \varnothing$.

"$\Rightarrow$": Let $\varphi \in t_2$ and $\boxminus\varphi \in t_1$, then $\neg\varphi \in \boxminus_\neg t_1$. Since $\varphi \in t_2$, it follows that $\boxminus_\neg(t_1) \nsubseteq t_2$.

"$\Leftarrow$": Let $\boxminus_\neg(t_1) \nsubseteq t_2$; then, there is some φ such that $\boxminus\neg\varphi \in t_1$ and $\varphi \notin t_2$. By maximality of t_2 we have $\neg\varphi \in t_2$, and therefore, $\boxminus\neg\varphi \in t_1 \cap \{\boxminus\psi : \psi \in t_2\}$. □

Claim 3: $S \subseteq R$:

Let $t_1, t_2 \in W$ and $\langle t_1, t_2\rangle \in S$. By definition of S there is some $\varphi \in t_2$ such that $\boxminus\varphi \in t_1$. We need to show that $t_1 R t_2$; thus, let $\psi \in t_2$. Since $\varphi \in t_2$ we have $\varphi \wedge \psi \in t_2$, and since t_2 is consistent, $\nvdash (\varphi \wedge \psi) \to \bot$. Therefore, $\vdash \boxminus\varphi \to \Diamond\psi$ by (7.3); in particular, $\boxminus\varphi \to \Diamond\psi \in t_1$. Now, $\boxminus\varphi \in t_1$ and $\boxminus\varphi \to \Diamond\psi \in t_1$ imply $\Diamond\psi \in t_1$ by modus ponens. □

For each formula φ, let $v(\varphi) = \{t \in W : \varphi \in t\}$.

Claim 4: v is a valuation:

We only show the claim for the modal operators. For (7.4), let $t_1 \in v(\Box\varphi)$, i.e. $\Box\varphi \in t_1$. If $t_1 R t_2$, then $\varphi \in t_2$ by Claim 1, and thus, $t_2 \in v(\varphi)$.

Conversely, suppose that $t_1 \notin v(\Box\varphi)$; then $\varphi \notin \Box t_1$. Since $\Box t_1$ is deductively closed, $\Box t_1 \cup \{\neg\varphi\}$ is consistent. Hence, there is some maximal consistent theory t_2 such that $\Box t_1 \cup \{\neg\varphi\} \subseteq t_2$. Since $\Box t_1 \subseteq t_2$ we have $\langle t_1, t_2\rangle \in R$, and $\neg\varphi \in t_2$ implies $v(\varphi) \nsubseteq R(t_1)$.

For 7.5 let $t_1 \in v(\boxminus\varphi)$, and suppose that $t_2 \in v(\varphi)$. Then, $\varphi \in t_2$ and $\boxminus\varphi \in t_1$ show that $\langle t_1, t_2\rangle \in S$. Thus, $v(\varphi) \subseteq S(t_1)$. Conversely, let $t_1 \notin v(\boxminus\varphi)$, i.e. $\boxminus\varphi \notin t_1$; then, $\neg\varphi \notin \boxminus_\neg(t_1)$. Since $\boxminus\neg$ is a necessity operator and t_1 is maximal, the set $\boxminus_\neg(t_1)$ is deductively closed, and therefore, $\boxminus_\neg(t_1) \cup \{\varphi\}$ is consistent. Suppose that $t_2 \in W$ contains $\boxminus_\neg(t_1) \cup \{\varphi\}$. Then, $\langle t_1, t_2\rangle \notin S$ and $t_2 \in v(\varphi)$, which implies $v(\varphi) \nsubseteq S(t_1)$. □

It follows from the previous claims that $\langle W, R, S, v\rangle$ is a model of $K^\#$. If $K^\# \nvdash \varphi$, there is a maximal consistent theory t not containing φ, i.e. $t \notin v(\varphi)$. Hence, $\langle W, R, S, v\rangle \nvDash \varphi$, and therefore, $K^\# \nvDash \varphi$.

7.2 Algebraic semantics

Let $T(\mathsf{Var})$ be the term algebra over the language of $K^\#$ with the set Var of variables; in other words, $T(\mathsf{Var})$ is the absolutely free algebra over the type of PS–algebras generated by Var. Thus, each formula $\varphi(p_1, \ldots, p_n)$ of $K^\#$ can be regarded as an element of $T(\mathsf{Var})$.

Lemma 7. *Let $M = \langle X, R, S, v\rangle$ be a model of $K^\#$, and $B_v = \{v(\varphi) : \varphi \in \mathsf{Fml}\}$.*

1. $\mathfrak{B}_v = \langle B_v, \cap, \cup, \varnothing, X, \langle R\rangle, [[S]]\rangle \in \mathbf{wMIA}$.
2. *If $\mathfrak{B}$ is a subalgebra of $\mathfrak{Cm}_{PS}(M)$ and v is a mapping onto a set of generators of $\mathfrak{B}$, then $\mathfrak{B} = \mathfrak{B}_v$.*

Proof. 1. By definition, the extension of v over $T(\mathsf{Var})$ is a homomorphism $T(\mathsf{Var}) \to \mathfrak{Cm}_{PS}(M)$, thus, $\mathfrak{B}_v$ is a subalgebra of $\mathfrak{Cm}_{PS}(M)$. Since $\mathfrak{Cm}_{PS}(M) \in \mathbf{wMIA}$ and $\mathbf{wMIA}$ is a universal class we obtain $\mathfrak{B}_v \in \mathbf{wMIA}$.

2. This follows again from the definition of the extension of v and the fact that v maps Var onto a set of generators.

The system $\langle M, B_v\rangle$ is an instance of a *general frame* of [16], see also Sections 1.4 and 5.5 of [9], in particular, Example 5.61.

If $\varphi(p_1, \ldots, p_n)$ is a formula, its corresponding term function (as defined in T_1 and T_2) is denoted by $\tau_\varphi^{\mathfrak{B}}(x_1, \ldots, x_n)$. We say that $\varphi(p_1, \ldots, p_n)$ *is valid in* $\mathfrak{B}$, written as $\mathfrak{B} \vDash \varphi(p_1, \ldots, p_n)$, if $\tau_\varphi^{\mathfrak{B}}(x_1, \ldots, x_n) \approx 1$. In other words, $\mathfrak{B} \vDash \varphi(p_1, \ldots, p_n)$ if and only if $\tau_\varphi^{\mathfrak{B}}(v(p_1), \ldots, v(p_n)) = 1$ for all mappings $v : \mathsf{Var} \to B$. If $\mathbf{K}$ is a class of algebras, then we define $\mathbf{K} \vDash \varphi(p_1, \ldots, p_n)$ if and only if $\mathfrak{B} \vDash \varphi(p_1, \ldots, p_n)$ for all $\mathfrak{B} \in \mathbf{K}$.

Theorem 11. *For all formulas $\varphi(p_1,\dots,p_n)$,*

$$K^{\#} \vDash \varphi(p_1,\dots,p_n) \text{ if and only if } \mathbf{WMIA} \vDash \varphi(p_1,\dots,p_n).$$

Proof. "⇒": Suppose that $K^{\#} \vDash \varphi(p_1,\dots,p_n)$, and that $\mathfrak{B} = \langle B, f, g\rangle \in \mathbf{wMIA}$. By Theorem 9, we may suppose that $\mathfrak{B}$ is isomorphic to a complex algebra of a weak MIA frame $\langle X, R, S\rangle$. Then, $\langle X, R, S\rangle \vDash \varphi(p_1,\dots,p_n)$ implies $v(\varphi(p_1,\dots,p_n)) = X$ for all valuations $v : \mathsf{Fml} \to 2^X$, and therefore, in particular, $\tau_\varphi^{\mathfrak{B}}(v(p_1),\dots,v(p_n)) = 1$ for all mappings $v : \mathsf{Var} \to \mathfrak{B}$. It follows that $\mathfrak{Cm}(X) \vDash \varphi(p_1,\dots,p_n)$, and therefore, $\mathfrak{B} \vDash \varphi(p_1,\dots,p_n)$.

"⇐": Suppose that $\mathbf{wMIA} \vDash \varphi(p_1,\dots,p_n)$, and that $\langle X, R, S\rangle$ is a weak MIA frame with full complex algebra $\mathfrak{B}$. Since $\mathfrak{B} \in \mathbf{wMIA}$ and $\mathbf{wMIA} \vDash \varphi(p_1,\dots,p_n)$, we have $\tau_\varphi^{\mathfrak{B}}(v(p_1),\dots,v(p_n)) = X$ for all mappings $v : \mathsf{Var} \to \mathfrak{B}$. Since the extension of v over formulas is the term definition of φ this implies $\langle X, R, S, v\rangle \vDash \varphi(p_1,\dots,p_n)$.

Together with Theorem 10 we obtain the following algebraic completeness theorem:

Theorem 12. *If φ is a formula in $K^{\#}$, then $K^{\#} \vdash \varphi$ if and only if* ***Eq***$(\mathbf{wMIA}) \vDash \varphi$.

8 The equational class generated by wMIA

In this section we shall describe the equational class generated by **wMIA** and thus, we can give a full algebraic characterization of the logic $K^{\#}$. First, we show that a weak MIA is a discriminator algebra. Recall the mapping $u : B \to B$ defined in (3.5), namely,

$$u(a) = f^{\partial}(a) \wedge g(\neg a). \tag{8.1}$$

It will turn out that u^{∂} is the unary discriminator. We have chosen to start with u as this mapping will be important later.

Theorem 13. *Let $\langle B, f, g\rangle$ be a PS–algebra. Then, B is a weak MIA if and only if*

$$u(a) = \begin{cases} 1, & \text{if } a = 1,\\ 0, & \text{otherwise.}\end{cases} \tag{8.2}$$

Proof. "⇒": First, consider $a = 1$. Then,

$$u(1) = f^{\partial}(1) \wedge g(0) = \neg f(0) \wedge g(0) = 1,$$

since $f(0) = 0$, and $g(0) = 1$. Next, let $a \neq 1$. Then, $\neg a \neq 0$, and

$$\begin{aligned} g(\neg a) &\leq f(\neg a) && \text{By Theorem 8}\\ \neg f(\neg a) \wedge g(\neg a) &= 0\\ f^{\partial}(a) \wedge g(\neg a) &= 0\\ u(a) &= 0, \end{aligned}$$

"⇐": Suppose that $a \neq 0$. By Theorem 8 it suffices to show that $g(a) \leq f(a)$. From $a \neq 0$ it follows that $\neg a \neq 1$, and thus, $u(\neg a) = 0$ by the hypothesis. Now, by the definition of u,

$$u(\neg a) = 0 \iff f^{\partial}(\neg a) \wedge g(a) = 0 \iff \neg f(a) \wedge g(a) = 0 \iff g(a) \leq f(a).$$

This completes the proof.

Theorem 13 gives us yet another characterization of weak MIAs among PS–algebras.

Corollary 2. *Each weak MIA is a discriminator algebra.*

Proof. Let B be a weak MIA. We show that B has a unary discriminator, i.e. there is a mapping $t : B \to B$ for which

$$t(a) = \begin{cases} 0, & \text{if } a = 0,\\ 1, & \text{otherwise.}\end{cases}$$

Indeed, set $t(a) = u^{\partial}(a) = \neg u(\neg a)$. Then, t fulfills the condition.

Observe that it follows that **wMIA** is not an equational class, since every discriminator algebra is simple. To describe **Eq(wMIA)** we shall relax the condition that u^∂ is the unary discriminator to the fact that u is an S5 necessity operator. Call a PS–algebra a $\langle B, f, g\rangle$ a $K^\#$*–algebra* if u satisfies the following conditions:

$$u(a) \leq a, \tag{8.3}$$
$$u(a) \leq u(u(a)), \tag{8.4}$$
$$a \leq u(u^\partial(a)). \tag{8.5}$$

The class of $K^\#$–algebras is denoted by **KMIA**. The motivation for these algebras come from the axiom system of the logic $K^\sim$ of [5].

It follows immediately from Theorem 13 that a weak MIA satisfies (8.3) – (8.5). Since **KMIA** is an equational class and **wMIA** is not, the inclusion **wMIA** $\subsetneq$ **KMIA** is strict. It may be instructive to present a concrete example:

Example 2. Suppose that $|B| > 2$, and let f be the identity on B and g be the Boolean complement. Then, f is a possibility operator, g is a sufficiency operator, and therefore, $\langle B, f, g\rangle$ is a PS–algebra. Furthermore, $f = f^\partial$, and, for all $a \in B$,

$$u(a) = f^\partial(a) \wedge g(-a) = a \wedge g(-a) = a, \tag{8.6}$$

and thus, $u^\partial = u$. Clearly, u satisfies (8.3) – (8.5), but is not a weak MIA. □

The next result exhibits the precise connection between **wMIA** and **KMIA**:

Theorem 14. ***Eq(*wMIA*)*** $=$ **KMIA***.*

Proof. We shall show that

1. **KMIA** is semisimple, i.e. every subdirectly irreducible $K^\#$ algebra is simple, and
2. The simple elements of **KMIA** are in **wMIA**.

Then, by Birkhoff's Theorem (see e.g. [8, Theorem 11.12]), every $K^\#$ algebra is isomorphic to a subdirect product of weak MIAs, and thus, it is in the equational class generated by **wMIA**. The other direction follows from **wMIA** $\subsetneq$ **KMIA**.

Let $\langle B, f, g\rangle \in$ **KMIA** be subdirectly irreducible. By Corollary 1, the congruences of B are in $1-1$ correspondence with the u $-$ filters of B, and therefore, $\langle B, u\rangle$ is subdirectly irreducible in the class of all Boolean algebras with an additional necessity operator. By (8.3) and (8.4) we have $u(a) = a \wedge u(a) \wedge u(u(a)) \wedge \ldots \wedge u^n(a)$, and therefore,

$$(\exists c \neq 1)(\forall a \neq 1)u(a) \leq c \tag{8.7}$$

by Rautenberg's criterion [18, p. 155]. By (8.4) we may suppose that $u(c) = c$. Assume that $c \neq 0$. Then, $\neg c \neq 1$, and

$$\neg c \underset{(8.5)}{\leq} u(u^\partial(\neg c)) = u(\neg u(c)) \underset{u(c)=c}{=} u(\neg c) \underset{(8.7)}{\leq} c, \tag{8.8}$$

a contradiction. It follows that $u(a) = 0$ for all $a \neq 1$, and, clearly, $u(1) = 1$. Hence, B is in **wMIA** by Theorem 13.

We close this section by showing that **KMIA** is closed under canonical extensions by describing the canonical frames. Call a frame $\langle X, R, S\rangle$ a **KMIA** *frame* if $R \cup -S$ is an equivalence relation.

Theorem 15. *1. Let $\langle B, f, g\rangle$ be in* **KMIA***, and $\langle \mathrm{Ult}(B), R_f, R_g\rangle$ be its canonical frame. Then, $R_f \cup -R_g$ is an equivalence relation.*
2. Let $\langle X, R, S\rangle$ be a **KMIA** *frame. Then, $\langle 2^X, \langle R\rangle, [[S]]\rangle$ is in* **KMIA***.*

Proof. 1. Let w be the dual of u; then, by the properties of u, w is normal additive closure operator in which every open set is closed. It is well known from the properties of S5 that the canonical relation R_w on $\mathrm{Ult}(B)$ is an equivalence relation. Note that $\langle F, G\rangle \in R_w$ if and only if $w[G] \subseteq F$. We are going to show that $R_w = R_f \cup -R_g$:

"$\subseteq$": Assume that this is not true, i.e. that there are $F, G \in \mathrm{Ult}(B)$ such that

1. $\langle F, G\rangle \in R_w$, i.e. $(\forall a)[a \in G$ implies $f(a) \vee \neg g(a) \in F]$.
2. $\langle F, G\rangle \notin R_f$, i.e. $(\exists b)[b \in G$ and $f(b) \notin F]$.
3. $\langle F, G\rangle \in R_g$, i.e. $(\exists c)[c \in G$ and $g(c) \in F]$.

Let $d = b \wedge c$; then, $d \leq b$. Since $d \in G$ we have $f(d) \vee \neg g(d) \in F$. If $f(d) \in F$, then $d \leq b$ implies $f(b) \in F$, contradicting 2. If, on the other hand, $\neg g(d) \in F$, then $d \leq c$ implies $g(c) \leq g(d)$, since g is antitone. It follows from 3. that $g(d) \in F$, also a contradiction.

"$\supseteq$": Let $\langle F, G\rangle \in R_f$; then $f[G] \subseteq F$. Suppose that $a \in G$. Then, $f(a) \in F$ implies $f(a) \vee \neg g(a) \in F$, since F is a filter. It follows that $\langle F, G\rangle \in R_w$, and consequently, $R_f \subseteq R_w$. Next, assume that $-R_g \nsubseteq R_w$. Then, there is some pair $\langle F, G\rangle$ such that $F \cap g[G] = \varnothing$ and $w[G] \nsubseteq F$. The latter implies that there is some $a \in G$ with $w(g) \notin F$. By definition of w this implies in particular that $\neg g(a) \notin F$, thus, $g(a) \in F$, since F is prime. Together with $a \in G$ this contradicts $F \cap g[G] = \varnothing$.

2. Let $\langle X, R, S\rangle$ be a **KMIA** frame, and define the mapping $[U]] : 2^X \to 2^X$ by $[U]](Y) = [R](Y) \cap [[S]](-Y)$. We need to show that $[U]]$ satisfies (8.3) – (8.5):

(8.3): Let $x \in [U]](Y)$. Then, $x \in [R](Y)$ and $x \in [[S]](-Y)$. By definition of $[R]$ and $[[S]]$, this implies $R(x) \subseteq Y$ and $-Y \subseteq S(x)$. Since $R \cup -S$ is reflexive, we obtain xRx or $x(-S)x$. If xRx, then $R(x) \subseteq Y$ implies $x \in Y$. If $x(-S)x$, then $x \notin S(x)$, in particular, $x \notin -Y$.

(8.4): Let $x \in [U]](Y)$. As above, we have $R(x) \cup -S(x) \subseteq Y$. We need to show that $x \in [U]][U]](Y)$, in other words, $x \in [R][U]](Y) \cap [[S]](-[U]](Y))$.

1. $x \in [R][U]](Y)$: Assume not. Then, $R(x) \nsubseteq [U]](Y)$, and so there is some y such that xRy and $y \notin [R](Y) \cap [[S]](-Y)$.
   (a) $y \notin [R](Y)$: Then, $R(y) \nsubseteq Y$, and there is some z such that yRz and $z \notin Y$. This implies that $xRyRz$, and the transitivity of $R \cup -S$ implies that xRz or $x(-S)z$. Since $R(x) \cup -S(x) \subseteq Y$ we obtain $z \in Y$, a contradiction.
   (b) $y \notin [[S]](-Y)$. Then, $-Y \nsubseteq S(y)$, and there is some z such that $z \notin Y$ and $y(-S)z$. Again by transitivity of $R \cup -S$ we obtain $x(R \cup -S)z$ which again by $R(x) \cup -S(x) \subseteq Y$ contradicts $z \notin Y$.

   Thus, $x \in [R][U]](Y)$.
2. $x \in [[S]](-[U]](Y))$: First, note that

$$\begin{aligned} x \in [[S]](-[U]](Y)) &\iff -[U]] \subseteq S(x) \\ &\iff (\forall y)[y \notin [U]](Y) \text{ implies } xSy] \\ &\iff (\forall y)[x(-S)y \text{ implies } y \in [U]](Y)] \\ &\iff (\forall y)[x(-S)y \text{ implies } y \in [R](Y) \cap [[S]](-Y)]. \end{aligned}$$

   Thus, let $x(-S)y$.
   (a) $y \in [R](Y)$: Assume not; then, there exists some z such that yRz and $z \notin Y$. Thus, $x(-S)yRz$, and the transitivity of $R \cup -S$ implies that $\langle x, z\rangle \in R \cup -S$. It follows that $z \in Y$ by $R(x) \cup -S(x) \subseteq Y$, a contradiction.
   (b) $y \in [[S]](-Y)$: Assume not; then, there is some z such that $z \notin Y$ and $y(-S)z$. As in the previous case we obtain $\langle x, z\rangle \in R \cup -S$ which implies $z \in Y$, a contradiction.

   It follows that $[U]]$ satisfies (8.4).

(8.5): Let $x \in Y$; we need to show that $x \in [R][U]]^{\partial}(Y) \cap [[S]][U]](-Y)$.

1. $x \in [R][U]]^{\partial}(Y)$: Assume not; then, there is some y such that xRy and $y \notin [U]]^{\partial}(Y)$. The latter implies that $y \notin \langle R\rangle(Y) \cup -[[S]](Y)$, i.e. $R(y) \cap Y = \varnothing$ and $Y \subseteq S(y)$. Since $R \cup -S$ is symmetric, xRy implies yRx. Now, $R(y) \cap Y = \varnothing$ implies $x \notin Y$, a contradiction.
2. $x \in [[S]][U]](-Y)$: Assume not; then, there is some y such that $y \in [U]](-Y)$ and $x(-S)y$. The first condition implies that $R(y) \cap Y = \varnothing$ and $Y \subseteq S(y)$. Since $x \in Y$, we obtain ySx. On the other hand, $x(-S)y$ implies $y(-S)x$, a contradiction.

Thus, the PS–complex algebra of a **KMIA** frame is in **KMIA**.

This is similar to the situation that S5 is characterized by the class of frames $\langle X, R\rangle$ where R is an equivalence, as well as by the class $\langle X, R\rangle$, where R is the universal relation on X. $K^{\#}$ is characterized by frames $\langle X, R, S\rangle$ where $R \cup -S$ is an equivalence (corresponding to **KMIA**) and also by frames $\langle X, R, S\rangle$ where $R \cup -S$ is the universal relation (corresponding to **wMIA**).

9 Special models

Now we return to the initial theme of the paper, namely, complex algebras $\langle 2^X, \langle R\rangle, [[R]]\rangle$ arising from a PS-frame $\langle X, R, S\rangle$ where $R = S$. We shall call a model $\langle X, R, S, v\rangle$ of $K^{\#}$ with $R = S$ *special.* Two models $M = \langle X, R, S, v\rangle$ and $M' = \langle X', R', S', v'\rangle$ are called *modally equivalent* if for all $\varphi \in \mathsf{Fml}$,

$$M \models \varphi \iff M' \models \varphi. \tag{9.1}$$

The following result was first mentioned for the logic $K^{\sim}$ [5]. Its proof uses the copying method introduced by Vakarelov, see e.g. [7]. We show it for $K^{\#}$ since the construction is slightly simpler.

Theorem 16. *Let $M = \langle X, R, S, v\rangle$ be a model of $K^{\#}$. Then, M is modally equivalent to a special model $\underline{M} = \langle \underline{X}, \underline{R}, \underline{v}\rangle$.*

Proof. Observe that the fact that $M = \langle X, R, S, v\rangle$ is a model of $K^{\#}$ implies that $R \cup -S = X^2$, and we need to find a model $\underline{M} = \langle \underline{X}, \underline{R}, \underline{S}, \underline{v}\rangle$ in which $\underline{R} \cap -\underline{S} = \varnothing$; in this case, $\underline{R} = \underline{S}$ and the model $\underline{M}$ is special.

Let $M' = \langle X', R', S', v'\rangle$ be an isomorphic copy of M with $X \cap X' = \varnothing$, and set $\underline{X} = X \cup X'$. For $x \in X$, denote its corresponding element in X' by x'. The valuation $\underline{v}$ is defined by $\underline{v}(\varphi) = v(\varphi) \cup v'(\varphi)$. The relations $\underline{R}$ and $\underline{-S}$ (and thus, implicitly, the relation $\underline{S}$) on $\underline{X}$ are defined by cases. The idea of the construction of $\underline{R}$ and $\underline{S}$ is to "separate" pairs $\langle x, y\rangle$ which are in the intersection of R and $-S$, i.e. to remove those pairs which prevent $R = S$. Let $x, y \in X$ and consider the following cases:

1. $\langle x, y\rangle \in R \cap -S$: Then, $x\underline{R}y', x'\underline{R}y, x\underline{-S}y, x'\underline{-S}y'$.
2. $\langle x, y\rangle \in S$: Then, $x\underline{R}y, x\underline{R}y', x'\underline{R}y, x'\underline{R}y'$.
3. $\langle x, y\rangle \in -R$: Then, $x\underline{-S}y, x\underline{-S}y', x'\underline{-S}y, x'\underline{-S}y'$.

Since $S \subseteq R$, these are all possibilities for $\langle x, y\rangle \in X^2$. If $s\underline{R}t$ or $s\underline{-S}t$ are not specified above, we suppose the default that $s - \underline{R}t$ and $s\underline{S}t$. Clearly, $\underline{R} \cap -\underline{S} = \varnothing$, i.e. $\underline{R} \subseteq \underline{S}$. Furthermore, it is not hard, if somewhat tedious, to show that $\underline{R} \cup \underline{-S} = \underline{X}^2$, i.e. $\underline{S} \subseteq \underline{R}$, so that altogether $\underline{R} = \underline{S}$. Since M and M' are isomorphic, M and $\underline{M}$ are modally equivalent with respect to the Boolean operators and $\Box$. This was shown, mutatis mutandis, by [7, Lemma 5.1.]. All that is left to show is $x \models_v \boxbar\varphi$ if and only if $x \models_{\underline{v}} \boxbar\varphi$ for all $x \in X$.

"$\Rightarrow$": Let $x \models_v \boxbar\varphi$. Then, by definition, $y \in v(\varphi)$ implies xSy, i.e. $-S(x) \cap v(\varphi) = \varnothing$. Using 1. or 3. above we show that $x \models \underline{v} \boxbar \varphi$. If $\underline{-S}(x) \cap v(\varphi) \neq \varnothing$, then there is some $y \in X$ such that $y \in v(\varphi)$ and $x - Sy$ by 1. or 3. above. This contradicts $x \models_v \boxbar\varphi$. If $\underline{-S}(x) \cap v'(\varphi) \neq \varnothing$, then there is some $y' \in X'$ such that $y' \in v'(\varphi)$ and $x\underline{-S}y'$ by 1. or 3. above. Since $x\underline{-S}y'$ if and only if $x'\underline{-S}y'$ and therefore $x - Sy$ by 1. or 3., we again arrive at a contradiction.

"$\Leftarrow$": Suppose that $x \models_{\underline{v}} \boxbar\varphi$. Then, $\underline{-S}(x) \cap \underline{v}(\varphi) = \underline{-S}(x) \cap (v(\varphi) \cup v'(\varphi)) = \varnothing$, and therefore, since $-S(x) \subseteq \underline{-S}(x)$, we obtain $-S(x) \cap v(\varphi) = \varnothing$.

Since $\underline{S} = \underline{R}$, the model $\underline{M} = \langle \underline{X}, \underline{R}, \underline{S}, \underline{v}\rangle$ is a special model of the form $\langle \underline{X}, \underline{R}, \underline{v}\rangle$.

Based on the previous considerations, we finally show that one canonical relation is enough for the equational class of generated by some wMIA:

Theorem 17. *Let $\mathfrak{B} = \langle B, f, g\rangle \in$ **wMIA**. Then, there is some frame$\langle X, R\rangle$ such that $\langle B, f, g\rangle$ and a subalgebra of $\langle 2^X, \langle R\rangle, [[R]]\rangle$ satisfy the same equations.*

Proof. By the Löwenheim – Skolem Theorem we may suppose that B is at most countable, and by Theorem 4, we may suppose that B is a subalgebra of $\langle 2^X, \langle R\rangle, [[S]]\rangle$ for some weak MIA frame $\langle X, R, S\rangle$.

Let $T = \{a_n : n \in \mathbb{N}\}$ be a set of generators of B, and define $v : \mathsf{Var} \to T$ by $v(p_n) = a_n$. Since T generates B, the extension $\overline{v}$ of v over the Lindenbaum – Tarski algebra $\mathfrak{L}$ is a surjective homomorphism onto $\mathfrak{B}$.

Consider the model $M = \langle X, R, S, v\rangle$, then, $\mathfrak{B} = \mathfrak{B}_v$ in the sense of Lemma 7, and $M \models \varphi$ if and only if $B_v \models \varphi$ for all $\varphi \in \mathsf{Fml}$. Suppose that $M' = \langle X', R', S', v'\rangle$ is modally equivalent to M, where M' is a special frame. Since the theorems of a model correspond to the equational theory of its general frame, it follows that $\mathbf{Eq}(\mathfrak{B}_v) = \mathbf{Eq}(\mathfrak{B}_{v'})$.

Corollary 3. KMIA *is the equational class generated by* **CMIA**.

Bibliography

1. van Benthem, J.: Correspondence theory. In Gabbay, D., Guenthner, F., eds.: Handbook of Philosophical Logic: Volume III. Reidel, Dordrecht (1984) 325 – 408
2. Orłowska, E., Radzikowska, A., Rewitzky, I.: Dualities for Structures of Applied Logics. College Publications (2015)
3. Düntsch, I., Orłowska, E.: Beyond modalities: Sufficiency and mixed algebras. In Orłowska, E., Szałas, A., eds.: Relational Methods in Computer Science Applications, Heidelberg, Physica Verlag (2000) 277–299
4. Düntsch, I., Orłowska, E.: Boolean algebras arising from information systems. Annals of Pure and Applied Logic **127** (2004) 77–98
5. Gargov, G., Passy, S., Tinchev, T.: Modal Environment for Boolean Speculations. In Skordev, D., ed.: Mathematical Logic and Applications, New York, Plenum Press (1987) 253–263
6. Tehlikeli (Passy), S.: An alternative modal logc, internal semantics and external syntax (A philosophical abstract of a mathematical essay). Manuscript (1985)
7. Vakarelov, D.: Modal logics for knowledge representation systems. Theoretical Computer Science **90** (1991) 433–456
8. Burris, S., Sankappanavar, H.P.: A Course in Universal Algebra. Springer-Verlag, New York (1981)
9. Blackburn, P., de Rijke, M., Venema., Y.: Modal Logic. Cambridge University Press (2001)
10. Goldblatt, R.: Mathematics of Modality. Number 43 in CSLI Lecture Notes. CSLI Publications, Stanford (1993)
11. Jónsson, B., Tarski, A.: Boolean algebras with operators I. American Journal of Mathematics **73** (1951) 891–939
12. Humberstone, I.L.: Inaccessible worlds. Notre Dame Journal of Formal Logic **24** (1983) 346–352
13. Koppelberg, S.: General Theory of Boolean Algebras. Volume 1 of Handbook on Boolean Algebras. North–Holland (1989)
14. Sambin, G., Vaccaro, V.: Topology and duality in modal logic. Annals of Pure and Applied Logic **37**(3) (1988) 249–296
15. Goldblatt, R.: Elementary generation and canonicity for varieties of Boolean algebras with operators. Algebra Universalis **34** (1995) 551–607
16. van Benthem, J.F.A.K.: Canonical modal logics and ultrafilter extensions. The Journal of Symbolic Logic **44**(1) (1979) 1–8
17. Tarski, A.: On the calculus of relations. Journal of Symbolic Logic **6** (1941) 73–89
18. Rautenberg, W.: Splitting lattices of logics. Arch. math. Logik **20** (1980) 155–159

Hypothesis Generation in Linear Temporal Logic for Clauses in a Restricted Syntactic Form

Marta Cialdea Mayer

Dipartimento di Ingegneria, Università degli Studi Roma Tre
email: cialdea@ing.uniroma3.it

Abstract. The interactions among the components of a biological system can be given a logical representation, that is useful for reasoning about them. One of the relevant problems that may be raised in this context is finding what would explain a given behaviour of some component; in other terms, generating hypothesis that, when added to the logical theory modeling the system, imply that behaviour. Temporal aspects have to be taken into account, in order to model the causality relationship that may link the behaviour of a given component to that of another one.
This paper presents a hypothesis generation method for linear temporal logic theories whose formulae have a restricted syntactic form, which is however sufficient to model biological systems. The method exploits the duality between hypothesis generation and consequence finding and is based on a resolution system proposed by Cavalli and Fariñas del Cerro in 1984.

1 Introduction

1.1 The true motivation

A few years ago, I had the occasion to meet Luis again after a long time and we said to ourselves that it would have been nice to join again our research work on some common topic. As a matter of fact, Luis tried a couple of times to draw my interest on some of his ongoing works, but without success (my fault!). Recently, he sent me some email messages asking whether I had any idea on how to extend classical abduction algorithms to propositional temporal logic, since it would be useful for his research on the logical modeling of biological systems.

> Comme tu sais nous travaillons sur la modélisation logique des systèmes biologiques, dans ce contexte nous avons besoins de l'abduction. [...]
> J'ai besoins de faire de l'abduction sur une logique propositionnelle avec la possibilité de l'expression du temps. [...]
> Pour la modélisation nous avons besoins du temps, puisque pour chaque réaction nous devons changer d'état.[1]

When I've been invited to submit a paper to LF2015, I have decided to try to do my homework. The paper presents an hypothesis generation method for linear temporal logic theories whose formulae have a restricted syntactic form, slightly extending the very basic form of a simple example given by Luis. To keep even more in theme with the LF2015 symposium, the method is based on a resolution system proposed by Ana R. Cavalli and Luis in 1984.

1.2 A more scientific motivation

In some recent works [1,2], R. Demolombe, N. Obeid and Luis propose a logical model of the series of biochemical reactions that may occur within a cell of a biological system. Cellular and molecular interactions can be graphically represented by means of diagrams called *molecular interaction maps*. Since such networks may involve many proteins and enzymes, and are consequently very complex, it is important to be able to reason about them. A logical model of these networks can be used to perform query answering by use of deduction, but also some form of abductive reasoning. For instance, since the maps used by biologists define causal relationships between different kinds of proteins, one may be interested in finding out which proteins should be activated or inhibited in order to obtain a given effect (e.g., in the case of research about cancer, the effect is to obtain the death of a cancer cell).

The formalism used in [2] to model metabolic networks is based on a fragment of first order logic. However, in the context of hypothesis generation, some form of temporal reasoning is needed. Let us consider a simple example (thanks to R. Demolombe):

[1] "As you know, we are working on the logical modeling of biological systems. In this context we need abduction. [...] I need abduction in a propositional logic with the possibility to express time. [...] We need time because we change state at every reaction."

We may have a protein of type A that can activate a protein of type B, which means that if A is activated, then B is activated. Then, we may have that protein B can inhibit protein A, which means that if B is activated, then A is inhibited.
In this example, if A is activated, then B is activated and A is inhibited. The conclusion is that A is both activated and inhibited, which is physically impossible (and which is logically inconsistent in our model).
The basic reason of this inconsistency is that time is ignored in the model. Indeed, if time is taken into account, the reasoning is: A is activated and next B is activated and next A is inhibited. Then, there is no more inconsistency because A is activated is consistent with next(next(A is inhibited)).

1.3 The temporal language needed to model biological systems

When I asked Luis to better specify the temporal language needed for their aims, his answer, with a photo in attachment (see Figure 1.1), was, in his usual speedy style:

> Voici un exemple. Nous avons le langage avec next et nécessaire, sans axiome d'induction et les règles sont de la forme comme dans la photo.
> L'idée est de trouver quelque chose de type résolution modale ou tableaux pour faire de l'abduction pour ce type des formules. Et, si possible, que ça soit une extension naturelle des méthodes d'abduction pour le calcul propositionnel (après voir si nous pouvons traiter le même problème mais cette fois avec une axiome d'induction entre le next et le nécessaire).[2]
>
> Envoyé de mon iPad

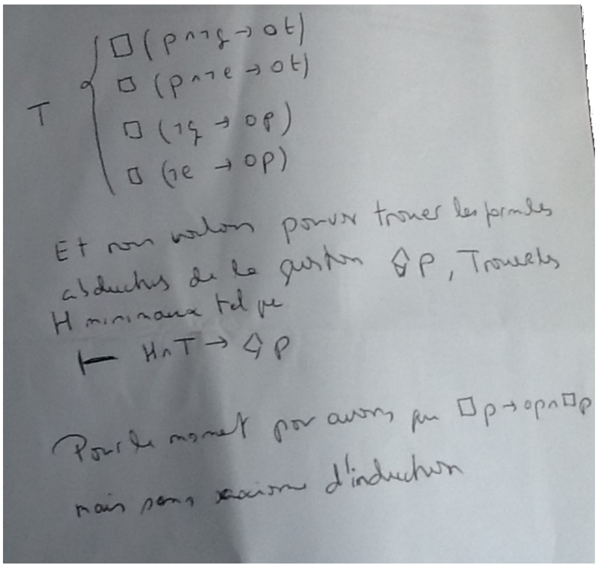


Fig. 1.1. Voici un exemple

This was all I had as the specification of the problem to be addressed: since LF2015 should be a surprise for Luis, I have not dared asking him more information. I asked some clarification to Robert Demolombe, who kindly explained me a little bit why temporal abduction would be of help, but I didn't want to bother him too much.

In the rest of this paper, linear temporal logic is briefly presented (Section 2), followed by a description of hypothesis generation problems (Section 3). Section 4 is the core of this work, presenting the syntactic restrictions on temporal formulae dealt with and the hypothesis generation method for such formulae. Some properties of the underlying resolution system are stated (and proved in the Appendix). Finally, Section 5 concludes this work.

[2] "Here is an example. We have the language with next and always, without induction axiom, and the rules have the form shown in the photo. The idea is to find something like modal resolution or tableaux, in order to perform abduction for this kind of formulae. And, if possible, it should be a natural extension of the abduction methods for propositional calculus (afterwards we shall see if we can face the same problem but with an induction axiom linking next with always)."

2 Linear Temporal Logic

The language of propositional linear temporal logic (LTL) considered in this work contains only unary future time temporal operators: $\Box$ (now and always in the future), $\Diamond$ (either now or sometimes in the future), and $\bigcirc$ (in the next state). Among the propositional connectives, $\neg$ (negation), $\wedge$ (conjunction) and $\vee$ (disjunction) are taken as primitive. The model of time underlying LTL is a countably infinite sequence of states (a *time frame*), that can be identified with $\mathbb{N}$. Its elements are called time points. An interpretation $\mathcal{M}$ is a function mapping each time point i to the set of propositional letters true at i. The satisfiability relation $\mathcal{M}_i \models A$, for $i \in \mathbb{N}$ (A is true at time point i in the interpretation $\mathcal{M}$), is inductively defined as follows:

1. $\mathcal{M}_i \models p$ iff $p \in \mathcal{M}(i)$, for any propositional letter p in the language.
2. $\mathcal{M}_i \models \neg A$ iff $\mathcal{M}_i \not\models A$.
3. $\mathcal{M}_i \models A \wedge B$ iff $\mathcal{M}_i \models A$ and $\mathcal{M}_i \models B$.
4. $\mathcal{M}_i \models A \vee B$ iff either $\mathcal{M}_i \models A$ or $\mathcal{M}_i \models B$.
5. $\mathcal{M}_i \models \Box A$ iff for all $j \geq i$, $\mathcal{M}_j \models A$.
6. $\mathcal{M}_i \models \Diamond A$ iff there exists $j \geq i$ such that $\mathcal{M}_j \models A$.
7. $\mathcal{M}_i \models \bigcirc A$ iff $\mathcal{M}_{i+1} \models A$.

Truth is satisfiability in the initial state: a formula A is true in $\mathcal{M}$ (and $\mathcal{M}$ is a model of A) iff $\mathcal{M}_0 \models A$. Truth of sets of formulae is defined as usual. If S is a set of formulae and A is a formula, A is a logical consequence of S ($S \models A$) iff for every interpretation $\mathcal{M}$ and $k \in \mathbb{N}$: if $\mathcal{M}_k \models S$ then $\mathcal{M}_k \models A$. Two formulae A and B are logically equivalent iff $A \models B$ and $B \models A$.

The standard axiomatic system for LTL is obtained by adding the following axioms A1-A5 and the inference rule R to any axiomatization of classical propositional logic:

$$\begin{array}{ll} A1. & \Box(A \to B) \to (\Box A \to \Box B) \\ A2. & \bigcirc(A \to B) \to (\bigcirc A \to \bigcirc B) \\ A3. & \bigcirc\neg A \equiv \neg\bigcirc A \\ A4. & \Box A \to (A \wedge \bigcirc\Box A) \\ A5. & \Box(A \to \bigcirc A) \to (A \to \Box A) \\ R. & \dfrac{A}{\Box A} \end{array}$$

3 Hypothesis Generation

Hypothesis generation consists in finding what, in the context of a given *background knowledge* T, would explain something that is not a consequence of T. Hypothesis generation is strictly related to abductive reasoning, a natural form of reasoning performed by an agent when new evidence informs her knowledge: an abductive inference is drawn when the truth of the sentence explaining the new evidence is "derived", i.e., it is added to the knowledge base. In general there may be different sentences explaining the same evidence, so some of them are rejected before the inference is performed. In other terms the inference is drawn when the agent is convinced that any other sentence is not plausible enough, compared with the chosen one, to be accepted as an explanation.

In the pure logical account of abduction, an abductive problem (in a given logic $\mathcal{L}$) is specified by a theory T and a sentence F to be explained (the *explanandum*); it is solved by finding a sentence E (among the best ones, according to some given preference criteria), such that $T \cup \{E\} \models_{\mathcal{L}} F$. It is moreover assumed that $T \not\models_{\mathcal{L}} F$ and $T \not\models_{\mathcal{L}} \neg F$, i.e., F is consistent with T. When viewed in these terms, there is an obvious duality between hypothesis generation and consequence finding (and between abduction and deduction): if T is the background theory, in the context of which a given fact F has to be explained, then the problem can be addressed by searching for a formula E such that $T \cup \{\neg F\} \models \neg E$. This amounts to $T \cup \{E\} \models F$, i.e, E explains F in the context of T. In other words, hypothesis generation can be reduced to consequence finding (up to a certain extent).

In both contexts, the formula looked for must satisfy additional conditions. In the logical approach to abduction, such conditions generally include *minimality* with respect to logical consequence: a formula E is a "relevant explanation" of F in the context of a theory T, only if $T \cup \{E\} \models F$ (i.e., E is an explanation of F) and there is no weaker explanation of F: for every formula E', if $T \cup \{E'\} \models F$ and $E \models E'$, then also $E' \models E$. In other terms, weaker explanations are preferred to stronger ones.

On the side of consequence finding, a stronger consequence is preferred to a weaker one: A is a "relevant consequence" of a set S of formulae only if for every consequence B of S, if $B \models A$, then also $A \models B$. Though this may be seen as a "maximality" condition, it amounts to minimality w.r.t. *subsumption* when looking for clauses implied by a given set of clauses S, in classical logic: a "relevant consequence" of S is a clause C that is a consequence of S and is not subsumed by any other clause C' such that $S \models C'$.

Getting back to hypothesis generation, there are additional requirements to be fulfilled by relevant explanations, in order to rule out trivial ones: a relevant explanation E of F in the background theory T is a minimal explanation of F such that

A. $T \not\models \neg E$: if $T \cup \{E\}$ is inconsistent, E is a trivial explanation of any fact.
B. $E \not\models F$: since $T \cup \{F\} \models F$, F is an explanation of F itself in T. If E is a minimal explanation of F in T, then $E \models F$ implies $F \models E$, i.e., E and F are logically equivalent. This would also make E a trivial explanation.

Consequently, when hypothesis generation is reduced to consequence finding, the two conditions A and B must be additionally considered.

Several proof-theoretical methods have been proposed for both consequence and hypothesis generation. Many of them are based on resolution (see, for instance, [3,4,5]), others on semantic tableaux and sequent calculi [6,7].

4 The hypothesis generation method

The problem of computing abductive explanations for LTL theories is very hard to face. In fact, it is not easy to adopt methodologies based on resolution or tableau methods, like can be done for classical logic [6,3,4,5] or even modal logics [7], because both resolution and tableau-based proof systems for temporal logics are among the most complex ones for modal logics. The hypothesis generation method proposed in this work exploits the very simple form of formulae making up theories modeling the behaviour of biological systems, and both *explananda* and explanations. In this restricted context, the problem becomes much simpler.

The method exploits the duality between hypothesis generation and consequence finding, and is based on resolution. In order to properly define the minimality condition, a suitable notion of subsumption for temporal clauses (in the considered restricted syntactical form) is defined. Hypothesis generation is reduced to derive the "relevant" consequences of the set of clauses obtained by adding the negation of the *explanandum* F to the background theory T. Relevant consequences are those which are not subsumed by any other consequence of $T \cup \{\neg F\}$ and are derived by using both $\neg F$ and some clause in T. The latter condition is required in order to satisfy the non triviality conditions A and B given in Section 3.

4.1 The syntactic restrictions

The important restrictions in the LTL language used to state the background theory and the negation of the *explanandum* are:

1. there are no occurrences of the $\Diamond$ operator; and
2. the $\Box$ operator never occurs in the scope of any logical operator.

The next definition introduces the syntactical form allowed for clauses (*flat clauses*) occurring in a resolution proof. The notation $\bigcirc^n$ abbreviates a sequence of n occurrences of the $\bigcirc$ operator.

Definition 1 (Flat clauses). *A* modal literal *is a formula of the form $\bigcirc^n \ell$ where $n \geq 0$ and ℓ is a classical literal (i.e., either an atom or the negation of an atom).*

A flat clause *is either:*

- *an* initial clause*, of the form $L_1 \vee \cdots \vee L_k$ where $L_1, \ldots, L_k$ are modal literals and $k \geq 1$, or*
- *an* always clause*, of the form $\Box(L_1 \vee \cdots \vee L_k)$ where $L_1, \ldots, L_k$ are modal literals and $k \geq 1$.*

Calligraphic lowercase letters (ℓ, p, q, possibly with subscripts) will be used to denote classical literals, while for modal literals the meta-symbols L and M (possibly with subscripts) will be used. As usual, disjunctions of literals are treated as sets, i.e., the order in which literals occur is irrelevant, and they are assumed to contain no repetitions. Flat clauses will sometimes be simply called clauses, when there is no risk of confusion.

Facts to be explained are assumed to have the form $\Diamond(L_1 \wedge \cdots \wedge L_k)$, where $L_1, \ldots, L_k$ are modal literals. Consequently, the negation of an *explanandum* is equivalent to a flat clause.

The subsumption relation for flat clauses is defined next.

Definition 2 (Subsumption). *Let C and C' be flat clauses. The clause C is subsumed by C' if one of the following cases holds (where disjunctions are treated as sets):*

- $C = L_1 \vee \cdots \vee L_k \vee M_1 \vee \cdots \vee M_n$, *for* $n \geq 0$, *and* $C' = L_1 \vee \cdots \vee L_k$;
- $C = \bigcirc^m L_1 \vee \cdots \vee \bigcirc^m L_k \vee M_1 \vee \cdots \vee M_n$, *for some* $m, n \geq 0$, *and* $C' = \Box(L_1 \vee \cdots \vee L_k)$;
- $C = \Box(\bigcirc^m L_1 \vee \cdots \vee \bigcirc^m L_k \vee M_1 \vee \cdots \vee M_n)$, *for some* $m, n \geq 0$, *and* $C' = \Box(L_1 \vee \cdots \vee L_k)$.

For instance, both $\bigcirc p \vee \bigcirc^2\neg p \vee q$ and $\Box(\bigcirc^2 p \vee \bigcirc^3\neg p \vee q)$ are subsumed by $\Box(p \vee \bigcirc\neg p)$. Subsumption between two initial clauses amounts to the subset relationship, like in classical propositional logic. It is worth noting that the subsumption relation is transitive.

4.2 The basic resolution rules for flat clauses

In order to give a compact formulation of the resolution rules, it is useful to define the complement of a modal literal.

Definition 3 (Complement). *Let L be a modal literal. The* complement *of L, $\sim L$, is defined as follows:*

- *if $L = p$, where p is an atom, then $\sim L = \neg p$;*
- *if $L = \neg p$, where p is an atom, then $\sim L = p$;*
- *if $L = \bigcirc M$, where M is a modal literal, then $\sim L = \bigcirc\sim M$. In other terms, $\sim\bigcirc^n \ell = \bigcirc^n \sim\ell$, where ℓ is a classical literal.*

The resolution rules adopted in this work simplify, taking advantage of the flat clause form, the rules introduced by Cavalli and Fariñas del Cerro [8].

Definition 4 (Basic Resolution Rules). *Let $L, L_1, \ldots, L_n, M_1, \ldots, M_m$ be modal literals and $k \geq 0$. The resolution system includes the following rules:*

$$\frac{\Box(L \vee L_1 \vee \cdots \vee L_n) \quad \Box(\bigcirc^k\sim L \vee M_1 \vee \cdots \vee M_m)}{\Box(\bigcirc^k L_1 \vee \cdots \vee \bigcirc^k L_n \vee M_1 \vee \cdots \vee M_m)}\ (R1)$$

$$\frac{\Box(L \vee L_1 \vee \cdots \vee L_n) \quad \bigcirc^k\sim L \vee M_1 \vee \cdots \vee M_m}{\bigcirc^k L_1 \vee \cdots \vee \bigcirc^k L_n \vee M_1 \vee \cdots \vee M_m}\ (R2)$$

$$\frac{L \vee L_1 \vee \cdots \vee L_n \quad \sim L \vee M_1 \vee \cdots \vee M_m}{L_1 \vee \cdots \vee L_n \vee M_1 \vee \cdots \vee M_m}\ (R3)$$

The two modal literals L and $\bigcirc^k\sim L$ in R1 and R2, and L and $\sim L$ in R3 are called the literals resolved upon in the inference.

In addition to the above rules, a simplification rule is added, allowing one to replace empty disjunctions and clauses made of the $\Box$ operator dominating an empty disjunction with $\bot$. Simplification will be applied implicitly. In what follows, R1-R3 will denote the resolution system consisting of the three rules of Definition 4 (and simplification), and the symbol $\vdash_{R1-R3}$ is used to denote derivability in such a system.

Example 1. Consider the simple example of Figure 1.1, where an explanation for $\Diamond p$ is to be looked for in the theory $T = \{\Box(\neg p \vee q \vee \bigcirc t), \Box(\neg p \vee e \vee \bigcirc t), \Box(q \vee \bigcirc p), \Box(e \vee \bigcirc p)\}$. A complete deduction from $T \cup \{\neg\Diamond p\}$ is:

1)	$\Box(\neg p \vee q \vee \bigcirc t)$	(in T)
2)	$\Box(\neg p \vee e \vee \bigcirc t)$	(in T)
3)	$\Box(q \vee \bigcirc p)$	(in T)
4)	$\Box(e \vee \bigcirc p)$	(in T)
5)	$\Box\neg p$	(negation of the *explanandum*)
6)	$\Box q$	(from 3 and 5)
7)	$\Box e$	(from 4 and 5)
8)	$\Box(e \vee \bigcirc q \vee \bigcirc^2 t)$	(from 1 and 4)
9)	$\Box(e \vee \bigcirc e \vee \bigcirc^2 t)$	(from 2 and 4)
10)	$\Box(q \vee \bigcirc q \vee \bigcirc^2 t)$	(from 1 and 3)
11)	$\Box(q \vee \bigcirc e \vee \bigcirc^2 t)$	(from 2 and 3)

The four clauses of lines 8-10, beyond being consequences of T alone, are subsumed by other clauses in the derivation (8 and 11 are subsumed by both 7 and 6, 9 is subsumed by 7 and 10 by 6), so they are ignored for hypothesis generation and should not even be added to the derivation.

The derivation of 6 makes use both of some clause in T and of clause 5, and the same holds for 7, therefore they are used to generate the two explanations $\Diamond\neg q$ and $\Diamond\neg e$, i.e. the negations of clauses 6 and 7, respectively.

Example 2. Consider the theory $T = \{\Box(\neg p \vee q \vee \bigcirc r), \Box(\neg s \vee \bigcirc p)\}$ and the *explanandum* $\Diamond r$. A (complete) deduction from T and the negation of the *explanandum* is the following:

1) $\Box(\neg p \vee q \vee \bigcirc r)$	(in T)
2) $\Box(\neg s \vee \bigcirc p)$	(in T)
3) $\Box\neg r$	(negation of the *explanandum*)
4) $\Box(\neg p \vee q)$	(from 1 and 3)
5) $\Box(\neg s \vee \bigcirc q \vee \bigcirc^2 r)$	(from 1 and 2)
6) $\Box(\neg s \vee \bigcirc q)$	(from either 2 and 4, or 3 and 5)

Clause 5 is a consequence of T, since in its derivation no use is made of clause 3, therefore its negation would be a trivial explanations. On the contrary, 4 and 6 depend both on T and the negation of the *explanandum*, therefore their negations constitute the two explanations found for $\Diamond r$, i.e., $\Diamond(\neg q \wedge p)$ and $\Diamond(s \wedge \bigcirc\neg q)$.

4.3 Refutational completeness

The rules R1-R3 are special cases of the general resolution rules defined in [8]. The latter are defined for formulae in a clause form, that will be here called CF-clauses. CF-clauses include flat clauses, but not vice-versa (CF-clauses are expressively complete for LTL). In order to address the problem of the refutational completeness for flat clauses of the system R1-R3, it is reduced to the resolution calculus defined in [8], that will henceforth be called CF and is given a brief description below, limited to what is relevant for the treatment of flat clauses.

The resolution rule of CF, when restricted to act on to CF-clauses without any occurrence of the $\Diamond$ operator, can be reformulated as follows. If C_1 and C_2 are clauses, $\Sigma(C_1, C_2) \rhd C$ denotes a Σ-*reduction step* of C_1 and C_2 and is recursively defined by the following Σ-reduction rules:[3]

(a) $\Sigma(p, \neg p) \rhd \bot$;
(b) $\Sigma(D_1 \vee D_2, F) \rhd \Sigma(D_1, F) \vee D_2$;
(c) $\Sigma(\bigcirc E, \bigcirc F) \rhd \bigcirc\Sigma(E, F)$;
(d) $\Sigma(\Box E, \nabla F) \rhd \nabla\Sigma(E, F)$ where $\nabla \in \{\Box, \bigcirc\}$;
(e) $\Sigma(\Box E, F) \rhd \Sigma(E, F)$;
(f) $\Sigma(\Box E, F) \rhd \Sigma(\Box\Box E, F)$.

The reflexive and transitive closure of $\rhd$ is denoted by $\rhd^*$: $\Sigma(C_1, C_2) \rhd^* C$ iff C is a clause and there is a sequence of Σ-reduction steps $\Sigma(C_1, C_2) \rhd \cdots \rhd C$. Two clauses C_1 and C_2 are *resolvable* if $\Sigma(C_1, C_2) \rhd^* C$ for some clause C. The *simplification* of a clause C is obtained by recursively replacing F for every subformula of the form $\bot \vee F$, and $\bot$ for every subformula of the form $\Box\bot$ and $\bigcirc\bot$. If C_1 and C_2 are resolvable, then a *CF-resolvent* $R(C_1, C_2)$ of C_1 and C_2 is the simplification of some C such that $\Sigma(C_1, C_2) \rhd^* C$.

When restricted to act on flat clauses, the system CF and R1-R3 are equivalent. Below, the symbol $\vdash_{CF}$ denotes derivability in the calculus CF.

Theorem 1. *If $S \cup \{C\}$ is a set of flat clauses, then:*

1. *if $S \vdash_{CF} C$, then $S \vdash_{R1-R3} C'$ for some flat clause C' that is logically equivalent to C;*
2. *if $S \vdash_{R1-R3} C$, then $S \vdash_{CF} C'$ for some clause C' that is logically equivalent to C.*

As a consequence:

Corollary 1. *The resolution system consisting of the rules R1-R3 is sound and refutationally complete for flat clauses.*

The proof of these results and the following ones can be found in the Appendix.

Example 3. Let us consider, for instance, the unsatisfiable set

$$S = \{p, \Box(\neg p \vee \bigcirc p), \bigcirc^2\neg p\}.$$

The following derivation shows that $S \vdash_{R1-R3} \bot$.

1) $\Box(\neg p \vee \bigcirc p)$	(in S)
2) p	(in S)
3) $\bigcirc^2\neg p$	(in S)
4) $\bigcirc p$	(from 1 and 2)
5) $\bigcirc^2 p$	(from 1 and 4)
6) $\bot$	(from 3 and 5)

[3] In [8], other reduction rules are included, but they all act on CF-clauses containing the $\Diamond$ operator, that is absent in flat clauses.

In general, every clause of the form $\bigcirc^n p$ can be derived from $\{p, \Box(\neg p \vee \bigcirc p)\}$. Therefore any set of the form $\{p, \Box(\neg p \vee \bigcirc p), \bigcirc^n \neg p\}$ can be refuted.

4.4 A weak form of implicational completeness

In order to study the derivational strength of the resolution calculus R1-R3, a translation is defined, mapping each flat clause into a first-order monadic formula. The translation simplifies the standard one (see, for instance, [9]), again taking advantage of the syntactical restrictions. In particular, since the $\Box$ operator only occurs as the outermost logical symbol in a flat clause, it acts as a sort of global modality, and there is no need to use an order relation in the target language.

Definition 5 (Translation). *Let P be a set of propositional atoms, and $\mathcal{L}_P$ the first-order language containing a unary predicate symbol p for each $p \in P$, a constant a and a unary functional symbol f.*

The (auxiliary) mapping τ^* maps a first-order term of $\mathcal{L}_P$ and a disjunction of modal literals into a formula of $\mathcal{L}_P$, and is defined as follows:

- $\tau^*(t, p) = p(t)$, *if* $p \in P$;
- $\tau^*(t, \neg p) = \neg p(t)$, *if* $p \in P$;
- $\tau^*(t, \bigcirc L) = \tau^*(f(t), L)$, *if L is a modal literal;*
- $\tau^*(t, A \vee B) = \tau^*(t, A) \vee \tau^*(t, B)$

Flat clauses are translated into classical first-order clauses of $\mathcal{L}_P$ by means of the mapping τ defined as follows:

- $\tau(L_1 \vee \cdots \vee L_k) = \tau^*(a, L_1 \vee \cdots \vee L_k)$;
- $\tau(\Box(L_1 \vee \cdots \vee L_k)) = \tau^*(x, L_1 \vee \cdots \vee L_k)$.

When translating a set S of flat clauses, a different variable is used for each clause in S.

For instance, if $C = p \vee \bigcirc q \vee \bigcirc^2 \neg r$ (an initial clause), $\tau(C) = p(a) \vee q(f(a)) \vee \neg r(f(f(a)))$, while $\tau(C') = p(x) \vee q(f(x)) \vee \neg r(f(f(x)))$ for the always clause $C' = \Box(p \vee \bigcirc q \vee \bigcirc^2 \neg r)$.

By use of the above defined translation, the relation between classical and temporal subsumption can be established. We recall that, in classical logic, a clause C subsumes a clause C' if there exists a substitution θ such that $C\theta \subseteq C'$.

Theorem 2. *A flat clause C subsumes a flat clause C' if and only if $\tau(C)$ (classically) subsumes $\tau(C')$.*

The correspondence between the temporal and classical settings established by the translation τ applies also to the resolution calculus consisting of the three rules R1-R3. As a matter of fact, such rules are just a rewriting of the classical resolution rule. In what follows, the symbol $\vdash_{FOL}$ denotes derivability by classical resolution in first-order logic. Analogously, while $\models$ denotes logical consequence in LTL, logical consequence in first-order logic is denoted by $\models_{FOL}$. As usual, a classical clause is intended to be universally closed; in particular, $\tau(C_1), \ldots, \tau(C_n) \models_{FOL} \tau(C)$ stands for $\forall\tau(C_1), \ldots, \forall\tau(C_n) \models_{FOL} \forall\tau(C)$, where $\forall A$ is the universal closure of A.

Theorem 3. *If $C_1, \ldots, C_n, C$ are flat clauses, then $C_1, \ldots, C_n \vdash_{R1-R3} C$ if and only if $\tau(C_1), \ldots, \tau(C_n) \vdash_{FOL} \tau(C)$.*

Theorem 3 allows one to exploit results holding for classical logic, such as the *implicational completeness* of resolution:

Theorem 4 (Lee [10]). *Let S be a set of classical clauses and C a non valid clause. If $S \models_{FOL} C$, there is a clause C' subsuming C such that C' is derivable from S by (classical) resolution.*

The strict correspondence between classical resolution and the temporal rules R1-R3 implies a weak form of completeness w.r.t. consequence finding:

Theorem 5. *If $C_1, \ldots, C_n, C$ are flat clauses, C is not valid and $\tau(C_1), \ldots, \tau(C_n) \models_{FOL} \tau(C)$, then there exists a clause C' subsuming C such that $C_1, \ldots, C_n \vdash_{R1-R3} C'$.*

4.5 Coping with the induction axiom

As a consequence of Theorem 5 and Corollary 1, the translation given in Definition 5 enjoys the following general property: if $\tau(C_1), \ldots, \tau(C_n) \vdash_{FOL} \tau(C)$, then $C_1, \ldots, C_n \models C$. The converse, obviously, does not hold: for instance, $p, \Box(\neg p \vee \bigcirc p) \models \Box p$, but $p(a), \forall x(\neg p(x) \vee p(f(x)) \nvdash_{FOL} \forall x p(x)$. As a matter of fact, the rules R1-R3 do not allow one to derive the flat clause $\Box p$ from p and $\Box(\neg p \vee \bigcirc p)$.

So, the system consisting of the three rules R1-R3 only is implicationally incomplete. Such rules in fact do not take into account what makes LTL different from FOL, i.e. the induction axiom (axiom A5 of Section 2): $\Box(A \to \bigcirc A) \to (A \to \Box A)$.

It is worth pointing out that there is no contradiction with Corollary 1: the negation of the induction axiom cannot be refuted in R1-R3 just because it cannot be expressed as a set of flat clauses (i.e. $\{A, \Box(\neg A \vee \bigcirc A), \Diamond\neg A\}$ is not a set of set of flat clauses). However, it could be expressed by use of an infinite disjunction:

$$A \wedge \Box(A \to \bigcirc A) \wedge \bigvee_{n \geq 0} \bigcirc^n \neg A$$

Therefore, when A is a modal literal the negation of the induction axiom could be (infinitely) refuted by refuting all the sets of flat clauses $\{A, \Box(\neg A \vee \bigcirc A), \bigcirc^n \neg A\}$ for every $n \geq 0$. As a matter of fact, each of such sets can be refuted in R1-R3 (see Example 3).

To the aim of gaining implicational completeness, the restricted syntax of flat clauses can be exploited again. Here, the $\Box$ operator only occurs as the outermost logical symbol, and consequence finding is also restricted to flat clauses. In this context, the induction axiom can be taken into account by adding the following induction rule:

$$\frac{L_1 \vee \cdots \vee L_k \quad \Box(\sim L_1 \vee \bigcirc L_1 \vee \cdots \vee \bigcirc L_k) \ \ldots \ \Box(\sim L_k \vee \bigcirc L_1 \vee \cdots \vee \bigcirc L_k)}{\Box(L_1 \vee \cdots \vee L_k)}\,(\text{Ind})$$

It is worth observing that the following inference would also be correct:

$$\frac{L_1 \vee \cdots \vee L_k \vee Q \quad \Box(\sim L_1 \vee \bigcirc L_1 \vee \cdots \vee \bigcirc L_k) \ \ldots \ \Box(\sim L_k \vee \bigcirc L_1 \vee \cdots \vee \bigcirc L_k)}{Q \vee \Box(L_1 \vee \cdots \vee L_k)}$$

but its conclusion is not a flat clause. Analogously, although $\Box(p \vee q), \Box(\neg q \vee \bigcirc q) \models \Box(p \vee \Box q)$, the conclusion could not be derived just because it is not a flat clause.

Let RES be the proof system consisting of the rules R1-R3 and Ind, and let $\vdash_{RES}$ denote derivability in RES. It may be hypothesized that RES is complete for LTL w.r.t. consequence finding restricted to flat clauses.

Conjecture. If $C_1, \ldots, C_n, C$ are flat clauses, C is not valid and $C_1, \ldots, C_n \models C$, then there exists a clause C' subsuming C such that $C_1, \ldots, C_n \vdash_{RES} C'$.

4.6 Non termination

The calculus RES does not enjoy the termination property, even if the generation of clauses subsumed by other clauses in the proof is blocked. A simple example of non-terminating derivation can be extracted from Example 3: if $T = \{\Box(p \to \bigcirc p), p\}$:

$$\cfrac{\Box(\neg p \vee \bigcirc p) \qquad \cfrac{\Box(\neg p \vee \bigcirc p) \qquad \cfrac{\Box(\neg p \vee \bigcirc p) \quad p}{\bigcirc p}\,(R2)}{\bigcirc\bigcirc p}\,(R2)}{\cfrac{\bigcirc\bigcirc\bigcirc p}{\vdots}}\,(R2)$$

Derivations may not terminate even if the application of rule R2 is blocked when the induction rule Ind can be applied. For instance, from $T = \Box(\neg p \vee \bigcirc p)$ every clause of the form $\Box(\neg p \vee \bigcirc^n p)$, for $n \geq 1$, can be generated, and none of them is subsumed by the others:

$$\cfrac{\Box(\neg p \vee \bigcirc p) \qquad \cfrac{\Box(\neg p \vee \bigcirc p) \qquad \cfrac{\Box(\neg p \vee \bigcirc p) \quad \Box(\neg p \vee \bigcirc p)}{\Box(\neg p \vee \bigcirc^2 p)}\,(R1)}{\Box(\neg p \vee \bigcirc^3 p)}\,(R1)}{\cfrac{\Box(\neg p \vee \bigcirc^4 p)}{\vdots}}\,(R1)$$

As a matter of fact, $\Box(\neg p \vee \bigcirc p)$ has an infinite number of logical consequences. They are all implied by $\Box(\neg p \vee \Box p)$, but the latter is not a flat clause.

The hypothesis generation method based on RES has been given a prototype implementation, where termination is forced by setting a bound on the maximal length of sequences of the $\bigcirc$ operator dominating classical literals.

5 Concluding Remarks

This paper presents a hypothesis generation method for linear temporal logic, when formulae are restricted to clauses in a very simple form. The method is based on a resolution system, called RES, whose rules simplify those presented in [8]. The calculus RES is refutationally complete and enjoys a weak form of implicational completeness. Its full implicational completeness is still an open question.

Different resolution methods for LTL have been defined in the literature and some of them could be taken as the basis for a hypothesis generation method. For instance, it would be worth considering the resolution system defined by Abadi and Manna [11], which deals with formulae in non-clausal form. Although in the present setting this might raise unnecessary complications, the approach deserves to be carefully analyzed to see whether the syntactic restrictions on flat clauses could be relaxed.

Other methods are not well suited as a basis for hypothesis generation. For instance, the temporal logic programming system presented in [12] is restricted to act on temporal horn clauses, where only a positive literal may occur, and cannot therefore express temporal clauses as defined in Section 4.

One of the other main resolution systems for LTL has been defined by Fisher et al. (see, for instance, [13]). It deals with clauses in *separated normal form* (SNF), which have one of the following forms:

- initial clauses: $\Box(\mathbf{start} \to \ell_1 \vee \cdots \vee \ell_n)$ where **start** is a distinguished propositional letter whose semantics is given by the condition that $\mathcal{M}_i \models \mathbf{start}$ if and only if $i = 0$;
- step clauses: $\Box(\ell_1 \wedge \cdots \wedge \ell_n \to \bigcirc(\not{p}_1 \vee \cdots \vee \not{p}_k))$
- sometime clauses: $\Box(\ell_1 \wedge \cdots \wedge \ell_n \to \Diamond \ell)$

Here, ℓ_i, $\not{p}_i$ and ℓ are classical literals, which may include $\top$ and $\bot$. Every LTL formula can be rewritten into a conjunction of SNF-clauses. The transformation involves the introduction of new propositional symbols; consequently, in order to use the calculus for hypothesis generation, the explanations extracted from a derivation should be re-converted back to formulae in the original language, and this may not be straightforward.

Assume, however, that the modal literals in the clauses of the background theory T and the negation of the *explanandum* have no nesting of the $\bigcirc$ operator. In this case, the only transformation needed to obtain SNF-clauses would be rewriting formulae of the form $\Box D$, where D is a disjunction of classical literals (without any literal of the form $\bigcirc \ell$) into the conjunction of the two SNF-clauses $\Box(\mathbf{start} \to D)$ and $\Box(\top \to \bigcirc D)$. In particular, if the *explanandum* has the form $\Diamond(\ell_1 \wedge \cdots \wedge \ell_n)$, its negation is rewritten as the conjunction of the two SNF-clauses $\Box(\mathbf{start} \to \overline{\ell_1} \vee \cdots \vee \overline{\ell_n})$ and $\Box(\top \to \bigcirc(\overline{\ell_1} \vee \cdots \vee \overline{\ell_n}))$.

In the hypothesis generation problems considered in this work, the set of SNF-clauses to be considered are only initial and step clauses, since sometime clauses are absent. Therefore, it is sufficient to consider the so-called *step resolution rules* of the resolution calculus given by Fisher et al., i.e.:

$$\frac{\Box(\mathbf{start} \to D_1 \vee p) \quad \Box(\mathbf{start} \to D_2 \vee \neg p)}{\Box(\mathbf{start} \to D_1 \vee D_2)}$$

$$\frac{\Box(C_1 \to \bigcirc(D_1 \vee p)) \quad \Box(C_2 \to \bigcirc(D_2 \vee \neg p))}{\Box(C_1 \wedge C_2 \to \bigcirc(D_1 \vee D_2))}$$

where each C_i is a conjunction of literals and each D_i a disjunction of literals (beyond a merge rule that is not worth stating here).

This calculus is not complete for consequence finding. Consider, for instance the set $S = \{\Box(q \to \bigcirc p), \Box q\}$. The first formula, $C_1 = \Box(q \to \bigcirc p)$, is an SNF-clause and the second one is rewritten into the conjunction of $C_2 = \Box(\mathbf{start} \to q)$ and $C_3 = \Box(\top \to \bigcirc q)$. From these SNF-clauses, no clause set equivalent to $\Box\bigcirc p$ can be derived, although $S \models \Box\bigcirc p$ (and, in fact, $S \vdash_{R1-R4} \Box\bigcirc p$). This is due to the fact that literals resolved upon always occur in the right-hand side of SNF-clauses, so that the only negative occurrence of q (in C_1) cannot be resolved against any of the two positive occurrences of q (in C_2 and C_3).

As far as future work is concerned, the problem of the full implicational completeness of the resolution calculus RES defined in this paper should be addressed. Moreover, possible refinements of RES can be studied, based on implicational complete resolution strategies. Such strategies can be identified by exploiting Theorem 3 and corresponding results in classical logic (for instance, [4,5,14]), as well as the fact that the derivation of any relevant consequence (whose negation is a relevant explanation) must make use of the negation of the *explanandum*. The results that will hopefully be obtained would allow one to enhance the present prototype implementation of the system, so that it can be experimented on complex examples, taken from domains modeling biological systems.

Acknowledgments. I wish to thank first of all Luis, for being always so rich of new and stimulating ideas and for his ability to look ahead and envisage problems and solutions before they show up. He is the exact contrary of typical Spaniards, who, as he once laughing told me, “speak faster than how they think”. And Luis does not speak slowly!

My thanks also go to Robert Demolombe, who was so kind to explain me clearly and concisely the basics of logical modeling of biological systems, and to my friend and colleague Serenella Cerrito, who, notwithstanding her time consuming duties, found the time and the will to give a look at the first draft of this work and pointed out some errors and unclear statements.

Bibliography

1. Demolombe, R., Fariñas del Cerro, L., Obeid, N.: Automated reasoning in metabolic networks with inhibition. In: AI*IA 2013: Advances in Artificial Intelligence. Volume 8249 of Lecture Notes in Computer Science., Springer International Publishing (2013) 37–47
2. Demolombe, R., Fariñas del Cerro, L., Obeid, N.: A logical model for molecular interaction maps. In: Logical Modeling of Biological Systems. John Wiley & Sons, Inc. (2014) 93–123
3. Cox, P.T., Pietrzykowski, T.: Causes for events: their computation and applications. In: Interanational Conference on Automated Deduction (CADE-86). (1986) 608–621
4. Demolombe, R., Fariñas Del Cerro, L.: An inference rule for hypothesis generation. In: Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91), Morgan Kaufmann Publishers Inc. (1991) 152–157
5. Inoue, K.: Linear resolution for consequence finding. Artificial Intelligence **56** (1992) 301–353
6. Cialdea Mayer, M., Pirri, F.: First order abduction via tableau and sequent calculi. Bulletin of the IGPL **1** (1993) 99–117
7. Cialdea Mayer, M., Pirri, F.: Propositional abduction in modal logic. Journal of the IGPL **3**(6) (1995) 907–919
8. Cavalli, A.R., Fariñas del Cerro, L.: A decision method for linear temporal logic. In: 7th International Conference on Automated Deduction. Volume 170 of LNCS. (1984) 113–127
9. Abadi, M.: The power of temporal proofs. Theoretical Computer Science **65**(1) (1989) 35–83
10. Lee, R.C.T.: A completeness theorem and a computer program for finding theorems derivable from given axioms. PhD thesis, University of California, Berkley (1967)
11. Abadi, M., Manna, Z.: Nonclausal temporal deduction. In: Proceedings of the Conference on Logic of Programs, Springer-Verlag (1985) 1–15
12. Abadi, M., Manna, Z.: Temporal logic programming. Journal of Symbolic Computation **8** (1989) 277–295
13. Fisher, M., Dixon, C., Peim, M.: Clausal temporal resolution. ACM Transactions on Computation Logic **2**(1) (2001) 12–56
14. Slagle, J.R., Chang, C.L., Lee, R.C.T.: Completeness theorems for semantic resolution in consequence-finding. In: Proceedings of the 1st International Joint Conference on Artificial Intelligence (IJCAI’69), Morgan Kaufmann Publishers Inc. (1969) 281–285

1 Proofs

Proof of Theorem 1: If $S \cup \{C\}$ is a set of flat clauses, then: (1) if $S \vdash_{CF} C$, then $S \vdash_{R1-R3} C'$ for some flat clause C' that is logically equivalent to C; (2) if $S \vdash_{R1-R3} C$, then $S \vdash_{CF} C'$ for some clause C' that is logically equivalent to C.

Proof. In order to ease readability, the Σ-reduction rules of the calculus CF are reported here:

(a) $\Sigma(p, \neg p) \rhd \bot$;
(b) $\Sigma(D_1 \vee D_2, F) \rhd \Sigma(D_1, F) \vee D_2$;
(c) $\Sigma(\bigcirc E, \bigcirc F) \rhd \bigcirc\Sigma(E, F)$;
(d) $\Sigma(\Box E, \nabla F) \rhd \nabla\Sigma(E, F)$ where $\nabla \in \{\Box, \bigcirc\}$;
(e) $\Sigma(\Box E, F) \rhd \Sigma(E, F)$;
(f) $\Sigma(\Box E, F) \rhd \Sigma(\Box\Box E, F)$.

Let $S \cup \{C\}$ be a set of flat clauses.

1. We first prove that if $S \vdash_{CF} C$, then $S \vdash_{R1-R3} C'$ for some clause C' that is logically equivalent to C. Observe, beforehand, that if $\Sigma(C_1, C_2) \rhd^* C$ for a given clause C, then the sequence of Σ-reduction steps leading to C necessarily ends with an application of the reduction rule (a) to a pair of complementary literals ℓ and $\sim\ell$ occurring in C_1 and C_2, respectively. Such literals ℓ and $\sim\ell$ are called the literals resolved upon in the inference leading from C_1 and C_2 to the simplification of C.
   Furthermore, once the two clauses C_1 and C_2 are given, and the literals ℓ and $\sim\ell$ resolved upon are identified, the clause C such that $\Sigma(C_1, C_2) \rhd^* C$, by menas a sequence of Σ-reduction steps ending with an application of the rule (a) to ℓ and $\sim\ell$, is unique.
   Assume now that C_1 and C_2 are resolvable temporal clauses and that $R(C_1, C_2)$ is a CF-resolvent of C_1 and C_2, obtained by simplifying a clause C such that $\Sigma(C_1, C_2) \rhd^* C$. The following reasoning shows that a clause logically equivalent to C can be derived from C_1 and C_2 by use of the rules R1-R3. Three cases are considered.
   - C_1 and C_2 are both initial clauses. Then C is obtained from C_1 and C_2 by application of the Σ-reduction rules (a)-(c). In particular, rule (b) is applied until a clause of the form $\Sigma(L, M) \vee D$ is obtained, where L and M are modal literals. Then rule (c) is applied until a clause of the form $\bigcirc^n\Sigma(p, \neg p) \vee D$ is obtained, which finally leads to D, by use of rule (a). Therefore, L and M have the forms $\bigcirc^n p$ and $\bigcirc^n \neg p$, respectively. In other terms, they are complementary modal literals, and the resolution rule R3 can be applied to obtain D.
   - C_1 and C_2 are both always clauses. Assume that the literals resolved upon are ℓ and $\sim\ell$. It may be assumed, w.l.o.g., that $C_1 = \Box(\bigcirc^n\ell \vee D_1)$ and $C_2 = \Box(\bigcirc^k\bigcirc^n\sim\ell \vee D_2)$. Then the base step (a) of the Σ-reductions, applied to the given literals, can be reached as follows (where the reduction symbol $\rhd$ is indexed by the applied Σ-reduction rule):

$$\begin{aligned}
\Sigma(C_1, C_2) &\rhd_f^* \Sigma(\Box^k C_1, C_2) = \Sigma(\Box\Box^k(\bigcirc^n\ell \vee D_1), \Box(\bigcirc^k\bigcirc^n\sim\ell \vee D_2)) \\
&\rhd_d \Box\Sigma(\Box^k(\bigcirc^n\ell \vee D_1), \bigcirc^k\bigcirc^n\sim\ell \vee D_2) \\
&\rhd_b \Box(\Sigma(\Box^k(\bigcirc^n\ell \vee D_1), \bigcirc^k\bigcirc^n\sim\ell) \vee D_2) \\
&\rhd_d^* \Box(\bigcirc^k\Sigma(\bigcirc^n\ell \vee D_1, \bigcirc^n\sim\ell) \vee D_2) \\
&\rhd_b \Box(\bigcirc^k(\Sigma(\bigcirc^n\ell, \bigcirc^n\sim\ell) \vee D_1) \vee D_2) \\
&\rhd_c^* \Box(\bigcirc^k(\bigcirc^n\Sigma(\ell, \sim\ell) \vee D_1) \vee D_2) \\
&\rhd_a \Box(\bigcirc^k(\bigcirc^n\bot \vee D_1) \vee D_2)
\end{aligned}$$

     If $D_1 = L_1 \vee \cdots \vee L_m$, by applying the resolution rule R1 to C_1 and C_2 the clause $\Box(\bigcirc^k L_1 \vee \cdots \vee \bigcirc^k L_m \vee D_2)$ is obtained, which is logically equivalent to the last CF-clause of the above Σ-reduction steps.
   - C_1 is an always clause and C_2 an initial one. Assume that the literals resolved upon are ℓ, occurring in C_1, and $\sim\ell$, occurring in C_2. I.e., $C_1 = \Box(\bigcirc^n\ell \vee D_1)$ and $C_2 = \bigcirc^k\bigcirc^n\sim\ell \vee D_2$. Then the the base step (a) of the Σ-reductions, applied to the given literals, can be reached as follows:

$$\begin{aligned}
\Sigma(C_1, C_2) &\rhd_f^* \Sigma(\Box^{k-1}C_1, C_2) = \Sigma(\Box^k(\bigcirc^n\ell \vee D_1), \bigcirc^k\bigcirc^n\sim\ell \vee D_2) \\
&\rhd_b \Sigma(\Box^k(\bigcirc^n\ell \vee D_1), \bigcirc^k\bigcirc^n\sim\ell) \vee D_2 \\
&\rhd_e^* \bigcirc^k\Sigma(\bigcirc^n\ell \vee D_1, \bigcirc^n\sim\ell) \vee D_2 \\
&\rhd_b \bigcirc^k(\Sigma(\bigcirc^n\ell, \bigcirc^n\sim\ell) \vee D_1) \vee D_2 \\
&\rhd_c^* \bigcirc^k(\bigcirc^n\Sigma(\ell, \sim\ell) \vee D_1) \vee D_2 \\
&\rhd_a \bigcirc^k(\bigcirc^n\bot \vee D_1) \vee D_2
\end{aligned}$$

     If $D_1 = L_1 \vee \cdots \vee L_m$, by applying the resolution rule R2 to C_1 and C_2, the clause $\bigcirc^k L_1 \vee \cdots \vee \bigcirc^k L_m \vee D_2$ is obtained, which is logically equivalent to the last clause of the above Σ-reductions.

2. For the other direction, let us assume that $C_1, C_2 \vdash_{R1-R3} C$. The reduction steps shown in the three cases considered above show that $C_1, C_2 \vdash_{CF} C'$ for some clause C' logically equivalent to C. □

Proof of Corollary 1: The resolution system consisting of the rules R1-R3 is sound and refutationally complete for flat clauses.

Proof. If $S \vdash_{R1-R3} C$, then by Theorem 1 $S \vdash_{CF} C$. Since CF is sound, $S \models C$, hence the resolution system consisting of the rules R1-R3 is sound.

Since CF is refutationally complete, if S is unsatisfiable, then $S \vdash_{CF} \bot$. Therefore, if $S \cup \{C\}$ is a set of flat clauses, $S \vdash_{R1-R3} \bot$ follows from Theorem 1 and the fact that $\bot$ can be derived from any contradictory flat clause by use of the simplification rules.

Proof of Theorem 2: A flat clause C subsumes a flat clause C' if and only if $\tau(C)$ (classically) subsumes $\tau(C')$.

Proof. Three cases are considered, according to the forms of C and C'. We consider here only the case when C and C' are both always clauses, the others being similar. In this case, if a clause C is subsumed by C', then they have the forms:

$$\begin{aligned} C &= \Box(\bigcirc^m \bigcirc^{p_1} \ell_1 \vee \cdots \vee \bigcirc^m \bigcirc^{p_k} \ell_k \vee M_1 \vee \cdots \vee M_n) \\ C' &= \Box(\bigcirc^{p_1} \ell_1 \vee \cdots \vee \bigcirc^{p_k} \ell_k) \end{aligned} \tag{1}$$

where $m, p_1, \ldots, p_k \geq 0$ and $\ell_1, \ldots, \ell_k$ are classical.

In this case, $\tau(C)$ and $\tau(C')$ have the forms

$$\begin{aligned} \tau(C) &= \ell_1(f^m(f^{p_1}(x))) \vee \cdots \vee \ell_k(f^m(f^{p_k}(x))) \vee D \\ &= \ell_1(f^{p_1}(f^m(x))) \vee \cdots \vee \ell_k(f^{p_k}(f^m(x))) \vee D \\ \tau(C') &= \ell_1(f^{p_1}(y)) \vee \cdots \vee \ell_k(f^{p_k}(y)) \end{aligned} \tag{2}$$

where $f^n(t)$ stands for the term $f(f(\ldots(f(t))))$ with n applications of the functional symbol f to the term t. If $\theta = \{f^m(x)/y\}$, then $\tau(C')\theta \subseteq \tau(C)$, therefore $\tau(C')$ classically subsumes $\tau(C)$.

For the converse, assume that $\tau(C')$ classically subsumes $\tau(C)$, i.e., $\tau(C')\theta \subseteq \tau(C)$ for some substitution θ. Since C and C' are always clauses, there is a single variable x occurring in C and a single variable y occurring in C'; moreover, every literal in $\tau(C)$ contains x and every literal in $\tau(C)$ contains y. Therefore, $\theta = \{f^m(x)/y\}$ for some m, and $\tau(C)$ and $\tau(C')$ have the forms shown in (2). As a consequence C and C' have the forms given in (1), i.e., C is subsumed by C'. □

Proof of Theorem 3: If $C_1, \ldots, C_n, C$ are flat clauses, then $C_1, \ldots, C_n \vdash_{R1-R3} C$ if and only if $\tau(C_1), \ldots, \tau(C_n) \vdash_{FOL} \tau(C)$.

Proof. The proof is by induction on the lenght of the derivations. The base case is obvious in both directions. For the induction step, it must be proved that C is derivable from C_1 and C_2 by one of the rules R1-R3 if and only if $\tau(C)$ is a classical resolvent of $\tau(C_1)$ and $\tau(C_2)$.

(⇒) Three cases are considered, according to the applied rule. We only show here the treatment of the rule R1, the others being similar. In this case:

$$\begin{aligned} C_1 &= \Box(\bigcirc^p \ell \vee \bigcirc^{p_1} \mathcal{p}_1 \vee \cdots \vee \bigcirc^{p_n} \mathcal{p}_n) \\ C_2 &= \Box(\bigcirc^k \bigcirc^p {\sim}\ell \vee \bigcirc^{q_1} \mathcal{q}_1 \vee \cdots \vee \bigcirc^{q_m} \mathcal{q}_m) \\ C &= \Box(\bigcirc^k \bigcirc^{p_1} \mathcal{p}_1 \vee \cdots \vee \bigcirc^k \bigcirc^{p_n} \mathcal{p}_n \vee \bigcirc^{q_1} \mathcal{q}_1 \vee \cdots \vee \bigcirc^{q_m} \mathcal{q}_m) \end{aligned}$$

(where $\ell, \mathcal{p}_1, \ldots, \mathcal{p}_n, \mathcal{q}_1, \ldots, \mathcal{q}_m$ are classical literals) and

$$\begin{aligned} \tau(C_1) &= \ell(f^p(x)) \vee \mathcal{p}_1(f^{p_1}(x)) \vee \cdots \vee \mathcal{p}_n(f^{p_n}(x))) \\ \tau(C_2) &= {\sim}\ell(f^k(f^p(y))) \vee \mathcal{q}_1(f^{q_1}(y)) \vee \cdots \vee \mathcal{q}_m(f^{q_m}(y)) \\ &= {\sim}\ell(f^p(f^k(y))) \vee \mathcal{q}_1(f^{q_1}(y)) \vee \cdots \vee \mathcal{q}_m(f^{q_m}(y)) \end{aligned}$$

The two classical clauses generate the resolvent

$$\begin{aligned} A &= \mathcal{p}_1(f^{p_1}(f^k(y))) \vee \cdots \vee \mathcal{p}_n(f^{p_n}(f^k(y))) \vee \mathcal{q}_1(f^{q_1}(y)) \vee \cdots \vee \mathcal{q}_m(f^{q_m}(y)) \\ &= \mathcal{p}_1(f^k(f^{p_1}(y))) \vee \cdots \vee \mathcal{p}_n(f^k(f^{p_n}(y))) \vee \mathcal{q}_1(f^{q_1}(y)) \vee \cdots \vee \mathcal{q}_m(f^{q_m}(y)) \end{aligned}$$

by use of the mgu $\theta = \{f^k(y)/x\}$ of $\ell(f^p(x))$ and $\ell(f^p(f^k(y)))$. Since $A = \tau(C)$, we are done.

($\Leftarrow$) First of all, we observe that factorization can never be applied to a clause $\tau(C)$. In fact, if C is an initial clause, $\tau(C)$ contains no variables, and if C is an always clause, all the literals in $\tau(C)$ contain the same variable, so no subset of its literals can be unified (unless it is a singleton).
So, it is sufficient to show that if $\tau(C)$ is a binary resolvent of $\tau(C_1)$ and $\tau(C_2)$, then $C_1, C_2 \vdash_{R1-R3} C$. Assume that

$$\frac{\tau(C_1) \quad \tau(C_2)}{\tau(C)}$$

by use of classical resolution. Different cases are considered, according to whether both C_1 and C_2 are always clauses, or one or both of them are initial clauses. We deal here only with the first case, the others being similar.
If both C_1 and C_2 are always clauses, $\tau(C_1)$ and $\tau(C_2)$ have the forms

$$\begin{aligned}\tau(C_1) &= \not{p}_1(f^{p_1}(x)) \vee \cdots \vee \not{p}_n(f^{p_n}(x))\\ \tau(C_2) &= \not{q}_1(f^{q_1}(y)) \vee \cdots \vee \not{q}_m(f^{q_m}(y))\end{aligned}$$

where $\not{p}_i$ and $\not{q}_j$ are classical literals. We may assume, w.l.o.g., that $\not{p}_1(f^{p_1}(x))$ and $\not{q}_1(f^{q_1}(y))$ are the complementary literals resolved upon, and that $p_1 \leq q_1$. Consequently, $f^{q_1}(y) = f^{p_1}(f^{q_1-p_1}(y))$ and the m.g.u. of $\not{p}_1(f^{p_1}(x))$ and the complement of $\not{q}_1(f^{q_1}(y))$ is $\theta = \{f^{q_1-p_1}(y)/x\}$. Therefore, if $k = q_1 - p_1$, the binary resolvent of $\tau(C_1)$ and $pi(C_2)$ is

$$\begin{aligned}A &= \not{p}_2(f^{p_2}(f^k(y))) \vee \ldots \not{p}_n(f^{p_n}(f^k(y))) \vee \not{q}_2(f^{q_2}(y)) \vee \ldots \not{q}_m(f^{q_m}(y))\\ &= \not{p}_2(f^k(f^{p_2}(y))) \vee \cdots \vee \not{p}_n(f^k(f^{p_n}(y))) \vee \not{q}_2(f^{q_2}(y)) \vee \cdots \vee \not{q}_m(f^{q_m}(y))\end{aligned}$$

Let now $C = \Box(\bigcirc^k \bigcirc^{p_2} \not{p}_2 \vee \cdots \vee \bigcirc^k \bigcirc^{p_n} \not{p}_n \vee \bigcirc^{q_1} \not{q}_1 \vee \cdots \vee \bigcirc^{q_m} \not{q}_m)$. Clearly, $A = \tau(C)$ and, since $q_1 = k + p_1$, C is derivable from C_1 and C_2 by application of the rule R1:

$$\frac{\Box(\bigcirc^{p_1} \not{p}_1 \vee \cdots \vee \bigcirc^{p_n} \not{p}_n) \quad \Box(\bigcirc^k \bigcirc^{p_1} \not{q}_1 \vee \bigcirc^{q_2} \not{q}_2 \vee \cdots \vee \bigcirc^{q_m} \not{q}_m)}{\Box(\bigcirc^k \bigcirc^{p_2} \not{p}_2 \vee \cdots \vee \bigcirc^k \bigcirc^{p_n} \not{p}_n \vee \bigcirc^{q_2} \not{q}_2 \vee \cdots \vee \bigcirc^{q_m} \not{q}_m)}$$

$\square$

The proof of Theorem 5 exploits the following intermediate result.

Lemma 1. *Let $C_1, \ldots, C_n$ be flat clauses. If $\tau(C_1), \ldots, \tau(C_n) \vdash_{FOL} A$, then there exists a flat clause C such that $A = \tau(C)$.*

Proof. The proof is by induction on the length of the derivation of A from $\tau(C_1), \ldots, \tau(C_n)$. If $A = \tau(C_i)$ for some $i = 1, \ldots, n$, then the result trivially holds. For the induction step, it must be shown that for any (classical) resolvent A of two clauses $\tau(C_1)$ and $\tau(C_2)$ there exists a flat clause C such that $A = \tau(C)$.

By definition, any flat clause C has the following property:

(α) either (i) $\tau(C)$ is variable-free, or (ii) $\tau(C)$ contains a single variable x which occurs in every literal of $\tau(C)$.

In fact, when C is an initial clause, $\tau(C)$ is variable-free; if C is an always clause, then (ii) holds.

We first show that property α is preserved:

(1) if A is a classical resolvent of $\tau(C_1)$ and $\tau(C_2)$, then A enjoys α.

If either C_1 or C_2 (or both) are initial clauses, then either $\tau(C_1)$ or $\tau(C_2)$ (or both) are variable-free and so are also their resolvents.

Otherwise, let C_1 and C_2 be always clauses, where the variable x occurs in every literal of $\tau(C_1)$, y occurs in every literal of $\tau(C_2)$, and let $\not{p}(f^n(x)) \in \tau(C_1)$ and $\not{q}(f^k(y)) \in \tau(C_2)$ be the literals resolved upon. We may assume, w.l.g., that $n \leq k$. Then the mgu of the two literals is $\{f^m(y)/x\}$ for $m = k - n$. Consequently, y is the only variable occurring in the resolvent A of $\tau(C_1)$ and $\tau(C_2)$, and it occurs in every literal of A.

Then, we show that

(2) if A is a classical clause satisfying property α, then there exists a flat clause C such that $A = \tau(C)$.

If A is variable free, then it has the form

$$A = \ell_1(f^{p_1}(a)) \vee \cdots \vee \ell_n(f^{p_n}(a))$$

If $C = \bigcirc^{p_1} \ell_1 \vee \cdots \vee \bigcirc^{p_n} \ell_n$, then $A = \tau(C)$.

Otherwise, A satisfies (ii), hence it has the form

$$A = \ell_1(f^{p_1}(x)) \vee \cdots \vee \ell_n(f^{p_n}(x))$$

If $C = \Box(\bigcirc^{p_1}\ell_1 \vee \cdots \vee \bigcirc^{p_n}\ell_n)$, then $A = \tau(C)$.

Finally: if A is a resolvent of $\tau(C_1)$ and $\tau(C_2)$, then (1) implies that it satisfies α. Consequently, by (2), there exists a flat clause C such that $A = \tau(C)$. □

Proof of Theorem 5: If $C_1, \dots, C_n, C$ are flat clauses, C is not valid and $\tau(C_1), \dots, \tau(C_n) \models_{FOL} \tau(C)$, then there exists a clause C' subsuming C such that $C_1, \dots, C_n \vdash_{R1-R3} C'$.

Proof. If C is not valid, then clearly $\tau(C)$ is not valid either. In fact, a pair of complementary classical literals in $\tau(C)$ corresponds to a pair of complementary modal literals in C. Hence, if C is not valid and $\tau(C_1), \dots, \tau(C_n) \models_{FOL} \tau(C)$, then by Theorem 4, there exists a clause A subsuming $\tau(C)$ such that $\tau(C_1), \dots, \tau(C_n) \vdash_{FOL} A$. By Lemma 1, there exists a temporal clause C' such that $A = \tau(C')$, i.e. $\tau(C_1), \dots, \tau(C_n) \vdash_{FOL} \tau(C')$. Since $\tau(C')$ subsumes $\tau(C)$, by Theorem 2, C' subsumes C. Finally, by Theorem 3, $C_1, \dots, C_n \vdash_{R1-R3} C'$. □

Formal Aspects of Architecture of Information

Mamede Lima-Marques[1] and Walter Carnielli[2]

[1] University of Brasília – UnB
Centre for Research on Architecture of Information
mamede@unb.br
[2] State University of Campinas – UNICAMP
Centre for Logic, Epistemology and the History of Science
walter.carnielli@cle.unicamp.br

Abstract. Information is acquiring crucial importance in our time, since the introduction of the so-called *information society* in the late 1970's characterized by the massive presence of information in our everyday life. Science of Information and more specifically the field of Architecture of Information (AI) plays a significant role in the analysis, criticism and in the building of new systems for a suitable use of Information in this new society. We briefly survey the main AI features, and their role in this context, particularly the fundamental aspects of AI proposed by the so-called Brasília Group. We highlight, as an AI-domain solution example, a project dealing with formal ontologies implemented in Classical and Modal Logic Programming. It is a formal language for discourse representation supported by ontological entities, able to handle deductions concerning domain ontologies. From the metamodeling paradigm, the language allows treating heterogeneous ontologies, which can be described as instances of one or more foundational ontologies. The language supports classical and modal features sustained by proof notions based on the Modal Logic Programming paradigm. Finally, it formalizes a systematization of the endurant fragment of the Unified Foundational Ontology (UFO), so as to compose part of the theoretical framework underlying the proposal, and to serve as an example of its instantiation.

Key-words: information architecture, ontologies, modal logic programming, formal language

1 Introduction

Information is acquiring crucial importance in our time, since the introduction of the so-called *information society* in the late 1970's characterized by the massive presence of information in our everyday life. As a result, it is perceived that a scientific understanding of society development in the information age has yet to be unraveled and that, to this moment, a science for, about and by the information society is still nonexistent. In this context, *Science of Information* and more specifically the field of *Architecture of Information* (AI) plays a significant role in the analysis, criticism and the building of new systems for the proper use of information in this new society.

In recent years, the volume of mediated information technology has grown dramatically, and it seems that this growth is far from over, steadily increasing society's technological capacity to store, communicate, and to process information. The way humans interact with the world, what they think, feel and believe, is supported by their sensory perceptions. Much of this background has been and remains built on a rich set of mediated information technology. What we perceive and how we perceive the world has a substantial influence on the construction of what we are, how we behave, and how we relate to other human beings. However, the experience of reality for us is something accessible only through internal mechanisms of perception and thought, which produce a personal and subjective interpretation of objective reality.

The purpose of this paper is to present some conceptual aspects of AI as proposed by Brasília's group in [1] and to report a significant event that has shown unusual applications in AI project of research, development and innovation. This project is an implementation of a formal language that supports discourses based on ontologies. We highlight, as an AI-domain solution example, a project that deals with formal ontologies implemented in classical and modal Logic Programming: a formal language for discourse representation supported by ontological entities, able to handle deductions concerning domain ontologies (cf. [2]). Following the metamodeling paradigm, the language allows the specification of heterogeneous ontologies, which can be expressed as instances of one or more foundational ontologies. The language supports classic and modal constructions sustained by proof notions based on the Modal Logic Programming paradigm. Finally, as an example, it formalizes the endurant fragment of the Unified Foundational Ontology (UFO) (cf. [3]), which can be used to specify domain ontologies based on UFO.

2 Credentials

In the early 1990's, the biologist Peter Marijuán from the University of Zaragoza, Spain, and the biophysicist Michael Earl Conrad (1941 to 2000), from the State University of Wayne, Michigan, started a Community of Information Science. Scientists around the world and from different disciplines have been brought together to discuss the concept of information from a transdisciplinary perspective. Since then this Community has held several international conferences: 1994 in Madrid, 1996 in Vienna, 2005 in Paris, 2010 in Beijing, 2013 in Moscow and the ISIS Summit 2015 in Vienna.

Efforts were made towards creating an organization to focus, develop and promote transdisciplinary approaches to information. In Paris, a group of scientists met in July 2005, at the 3^{rd} International Conference on the Foundations of Information Science (FIS). In the final meeting, participants agreed to extend the work of the FIS-Group, create an institute and further assemble, coordinate and correlate the past and current theoretical work on information. Members decided to call the newly expanded field *Science of Information*, not to be confused with the term "Information Science", which sometimes was known as the advancement "library". The new field takes into account a more recent and larger perspective, covering various academic disciplines and new fields of interest.

In August 2010, during the 4^{th} International Conference on the Foundations of Information Science (FIS), the first scientific conference on the topic *Towards a New Science of Information* was held. A committee was created to prepare the foundation of a Social Information Science Institute (SISI) with the purpose of advancing global and collaborative studies in Science of Information, Information Technology, and the Information Society. Also creating shared conceptual frameworks and implementing them in practice to contribute and to meet the challenges of the information age, and holding conferences in the field (cf. [4]).

The idea of Architecture of Information proposed by M. Lima–Marques follows the same scientific framework. Since its origin, it was designed considering the criticism of the industry's activities. In this general context, the AI was being used in a reductionist way, and applied almost exclusively to website solutions (cf. [5]).

The newly proposed AI is a design methodology applied to any information environment, understood as a space in a particular context, consisting of content flow, and serving a community of users. The purpose of the Arquitetura da Informação is to enable the efficient flow of information through information environments design, (cf. [6]).

AI is *transdisciplinary*[3] and has several professionals involved in its implementation. It applies methods and concepts arising from the new Science of Information and areas such as Philosophy, Mathematics, Logic, Linguistics, Computer Science, Cognitive Science, Business Administration, Economics, Library, Archival, among others. AI is a composition involving process, practice, and domain knowledge. It's a discipline where the practice strengthens and promotes its development (cf. [8]).

These are the objectives of AI: to develop semantically relevant information environments, in particular, contexts to meaning communities; to model information in environments that enable their design, management and sharing by users; and to promote improved communication, collaboration, and exchange of experiences.

AI is based on a humanistic vision where people, the subjects, are central to the creation of solutions, for which technology is, a necessary support. AI should be in agreement with the information requirements of those who use it, and need information at the right time (cf. [9]).

AI is a discipline whose object is information configuration (cf. [2]), i.e., its structure in appropriate phenomena. From a technological perspective, AI can be seen as a set of methods and techniques for designing information environments. The models developed to create AI depart from theoretical concepts and transformed into the information system: a collection of interrelated components – hardware, software, procedures, and databases, among others. These elements work together to act in the information life cycle, characterized by the steps: collection, description, organization, storage, retrieval, access, re-packaging, use, archiving, preservation, prevention, and destruction. Thus, information technology is the infrastructure that materializes each of these perspectives. The construction of such a system requires:

a) the determination of the configuration information in distinct spaces;
b) the implementation of acts that guide the development of necessary transformations; and

[3] *Transdisciplinarity* is a process in which there is a convergence between disciplines, accompanied by mutual integration of individual disciplinary epistemology (cf. [7]).

c) the development of valid solutions within AI models suitable for each particular purpose.

AI, from the perspective of social organizations or of the information society, can be associated with a world view, seen as a set of actions applied to a given information space to turn it into the information system. Currently, the information model domain is vital to an organization's survival. The whole economic structure is mounted on patterns of information. Mastering the information lifecycle from its origin, its organizational patterns, its representations of appropriate models that enable the understanding of phenomena and decision-making, has been the constant endeavor of organizations since the late twentieth century. However, the amount of information has proved too large for the human ability to consume. The profusion of information and its relevance to specific issues stirred up the development of technology and science of information and, more specifically, of AI as a tool to reduce the spread of information and make it more suitable for human understanding.

AI, as a discipline, refers to a systematic effort to identify patterns and to place methods for information space configuration, whose purpose is the representation and manipulation of information; and the establishment of relationships linking its linguistic entities (cf. [8]).

In [10] the authors defend the proposal of AI as a discipline and introduce a methodological approach based on van Gigch and Pipino (cf. [11]). This important step for building a new field announces an international trend in understanding its foundations. AI must assimilate elements of space, time, structure, semantics, and context; deepen the understanding of the nature of information and apply these results to critical problems in society.

While this information society has peculiarities that lead to the creation of a new scientific field, its primary characteristic is still maintained: people, human individuals who share experiences with each other, based on information. Although information is the bias of this new society, its subject matter is the people and the relationship they have with information. One of the ways people express this behavior is through speech, written or spoken, about the world experience.

In this sense, logic studies regarding the reasoning in a manner that take into account particular aspects of individuals and societies, as epistemic and society logics are instruments that promote understanding and the description of information phenomena in the context of human agents. AI combines logics and information to provide systems which enable the representation and manipulation of spaces for designing information environments to support human's common endeavors: speeches about the information phenomena they experience.

Therefore, in the context of AI development models applied to societies, we are building solutions that use modal logics to reason about speeches done by agents in the context of the conceptual modeling of reality. Conceptual Modeling is an area that applies Ontology results to develop descriptions of reality, usually based on some foundational ontology. Mostly, these ontologies operate as a reference and as a base ideology. Discourses concerning reality must be delivered with conceptual support given by these theories about the world. Epistemic logics can be applied to reason with reference to these speeches. Logic programming systems developed from the results of Luiz Fariñas del Cerro with Molog are employed to make an executable version of the results of investigations on ontologies, speeches, and modal logics. Primarily, it uses the works of Linh Anh Nguyen, (cf. [12]), with MProlog to prototype logic systems supporting the description of domain ontologies based on foundational meta-models. MProlog is a complementary approach to Molog that specializes in serial modal and multimodal systems. The characteristic of seriality is especially useful in epistemic reasoning systems suitable for the architectural approach of the information society. The following section describes those results.

3 Ontoprolog

The idea of speech is related to the speech act of the subject (cf. [13]) of a society of agents, about a particular reality. Speech acts are instruments of action on reality. By experiencing the reality of the universe of discourse, the subject is directly influenced by it, and by the speech the subject performs. Whereas the speech itself becomes part of that reality projected by the subject. Aforementioned shows a continuous and corresponding feedback process of objects over reality. So in that context, the speech is not free of ideologies, it can be interpreted as a way to act, to act on the other: it is an instrument of action in the world.

The discourse about reality (speech acts) is a description of objects by the subject. In his speech, information is indistinguishable from the object itself and the content of the communication. Information is the foundational element of speech, whereas information carries the speech relationship with reality. The description of objects, when performed in a systematic manner, is an ontology. Thus, information supports the ontological commitment grounding the definition of an ontology concerning a reality. In this sense, information is also communicated through speech, therefore performing speeches denotes ontology over information.

The systematization of a speech can be induced by a conceptual formalization process, based upon a description of the ontological commitments and standards of meanings shared by the subjects of the speeches. Formal languages are appropriate in this process, as they allow us to highlight concepts and relationships between concepts that are the basis for understanding the meaning of a speech.

The idea behind the Ontologies proposal is that they act as a semantic reference to the discourse. Ontologies should reflect real patterns of meaning, captured from ontological commitments, and sustained by a particular logic. However, to produce a discourse is also necessary to build an ontology (a new one), which should be of a specific domain, and which carries the worldview of the reference ontology. For this reason, it is important that everything on the dependencies among ontologies is semantically clear.

Ontoprolog is a formal language built to describe discourses based on ontologies. Reference ontologies can be constructed from a generic metamodel. Domain specific ontologies can be specified using on general ontologies. The system is anchored on a logical framework of Logic Programming and uses classical and modal reasoning. Ontoprolog features includes:

– a formal textual language, i.e. a basis for well-founded description of speeches covering subject experiences by ontological entities;

– a set of semantic and syntactic rules of the language, based on the proof framework of classical Programming in Logic;

– a set of modal extensions provided by MProlog to the base language for dealing with speeches based on multi-agent semantics as proposed by Nguyen (cf. [14]).

Here we present a solution implemented in Ontoprolog to exemplify its applicability. Let's use an ontology found in the literature: the Unified Foundational Ontology (UFO-A) (cf. [3]). The UFO-A is an ontology used as a reference in the conceptual modeling area for building domain-specific ontologies. In Ontoprolog, it is defined as an instance of a general metamodel. Accordingly, it can be used to describe particular ontology. Therefore, the framework defines a specification of UFO-A, a set of syntactic sugar, built by language, and a set of rules to validate extended universal ontologies and its particular ontology. The results are shown in [2] and are not detailed in this section.

Ontoprolog language is based on Prolog operators. In its grammar, sentences are syntactically designed to be recognized as the standard program. Therefore, it is possible to insert any ontology specifications in any Prolog program, including existing ones. Further, Prolog operators allow language extensions be performed relatively quickly. Therefore, expansions can be created to enrich the expressivity of the language. Regarding pragmatics of language, grammar was designed so that a user can apply the Ontoprolog specification in real conceptual modeling sessions, as intelligible as a *lingua franca* in a community of users.

From a technical point of view, the use of operators in implementing syntax allows representation and direct access to parse trees. Accordingly, lexical analysis of a language internally built on logic programming is automatically created. Thus, it is possible to construct sentences of the language that are equally sentences of a standard Prolog program, including the ability to make unification of sentences based on syntactic tree combined operators. So, initially, a base syntax that is able to express primary language constructs is conceived. This basic syntax is shown in Code 1 through EBNF rules (*Extended Backus-Naur Form*). The rating is based on the standard proposed in [15]:

Code 1. Ontoprolog EBNF syntax definition

```
ONTOLOG_SPECIFICATION ::= SENTENCE+

SENTENCE ::= ( DISJOINT | INSTANCE | PARTIAL_SUBSUMPTION | COMPLETE_SUBSUMPTION |
      PROPERTY_ASSOCIATION | PROPERTY_ASSIGNMENT | SUBSETS | REDEFINES | POWER_TYPE | META ) '.'

DISJOINT ::= 'disjoint' '[' ATOM (',' ATOM)* ']'

INSTANCE ::= ( ATOM | 'disjoint' '[' ATOM (',' ATOM )* ']' ) '::' ( INSTANTIABLE_ENTITY | '['
      ATOM (',' ATOM )* ']' )

PARTIAL_SUBSUMPTION ::= ( ATOM | 'disjoint'? '[' ATOM ( ',' ATOM )* ']' ) ( '::'
      INSTANTIABLE_ENTITY )? ( 'extends' | 'extend' ) ( ATOM | '[' ATOM ( ',' ATOM )* ']' )

COMPLETE_SUBSUMPTION ::= ( ATOM | ('disjoint'? '[' ATOM ( ',' ATOM)* ']' ) ) ( '::'
      INSTANTIABLE_ENTITY )? 'cover' ATOM

PROPERTY_ASSOCIATION ::= 'property' ( ATOM | '[' ATOM (',' ATOM )* ']' ) 'on' (
      ENTITY_WITH_PROPERTY | '[' ENTITY_WITH_PROPERTY (',' ENTITY_WITH_PROPERTY )* ']' )

PROPERTY_ASSIGNMENT ::= PROPERTY_TYPE 'at' ENTITY_IN_RELATION ':=' VALUE
```

```
VALUE ::= ANY_PROLOG_TERM

SUBSETS ::= PROPERTY_TYPE 'at' ENTITY_IN_RELATION 'subsets' PROPERTY_TYPE 'at' ENTITY_IN_RELATION

REDEFINES ::= PROPERTY_TYPE 'at' ENTITY_IN_RELATION 'redefines' PROPERTY_TYPE 'at'
      ENTITY_IN_RELATION

POWER_TYPE ::= 'powertype' ( ENTITY_PT | '[' ENTITY_PT (',' ENTITY_PT )* ']' ) 'classifying' (
      ATOM | '[' ATOM ( ',' ATOM )* ']')

META ::= 'meta' ( META_PROPERTY | '[' META_PROPERTY (',' META_PROPERTY)* ']' ) 'on' ( ( ATOM
      'at' )? ( ( 'extensions' | 'transitive'? 'direct'? 'instances' ) 'of' )? )? (
      ENTITY_WITH_META | '[' ENTITY_WITH_META (',' ENTITY_WITH_META )* ']' )

META_PROPERTY ::= ANY_PROLOG_TERM

INSTANTIABLE_ENTITY ::= ATOM

PROPERTY_TYPE ::= ATOM

ENTITY_IN_RELATION ::= ATOM

ENTITY_WITH_PROPERTY ::= ATOM

ENTITY_WITH_META ::= ENTITY_WITH_PROPERTY

ENTITY_PT ::= ENTITY_WITH_PROPERTY | ATOM 'at' ATOM

ATOM ::= [a-z][a-zA-Z0-9]* | "'" [^']* "'"
```

A valid instance of the grammatical structure presented in Code 1 is called *Ontoprolog Specification*. Therefore, from the presented definition, a specification (`ONTOPROLOG_SPECIFICATION`) is a non empty set of sentences (`SENTENCE`). Sentences can be from different syntactic categories but always end with a full stop. The first, `DISJOINT`, is defined with the operator `disjoint` followed by a list of atoms separated by commas and brackets. The same thought applies to other syntactic categories, which define the EBNF rules. The semantics of an Ontoprolog specification is called *theory* and is given by the structure

$\mathcal{OT} = \langle \mathcal{C}, \mathcal{G}, \mathcal{R}, \mathcal{H} \rangle$, such that:

a) $\mathcal{C}$ is a finite set, and not empty, of entities that hold ontological relations with the world;
b) $\mathcal{G}$ is a finite set, and possibly empty, of logical entities that do not require an ontological commitment to the described universe but are part of the universe of theoretical meta-objects about reality;
c) $\mathcal{R}$ is a finite set, and non-empty, of positive and negative rules defined on relations of $\mathcal{H}$. These rules regulate the notion of logical consequence obtained from the semantic relations;
d) $\mathcal{H}$ is a finite set, and not empty, of relations between instances of objects $\mathcal{C} \cup \mathcal{G}$ represented by a set of primitive predicate symbols. The relations are presented in the form of Horn clauses and are called *semantic relations*.

The assignment of meaning to the syntactic constructs created as instances of the grammatical definitions in Code 1 is performed using translations between such syntax and to a set of semantics relations of $\mathcal{H}$. It is essentially an approach driven by the syntax where semantics is defined by induction on the structure of language syntax. That is, using production rules; semantics is assigned to each possible combination of the syntactic constructs. Translation is a kind of *denotational semantic*, based on [16, p. 91]. Meaning that "there is a semantic clause for each syntactic basic category" and "to each method of constructing composite elements (the syntactic category) there is a semantic clause defined regarding semantic function, this one is applied to the immediate constituents of the composite members."

The definite clauses of $\mathcal{R}$ are divided into two categories of rules: *positive rules* and *semantic validation rules*, these defined as *negative predicates written as positive rules.* The *positive rules* are positive predicates denoting transitive closures and relations derived from the fact base. Derived relations are written in the form $h \leftarrow \varphi$, where h is the consequent of the implication, called the *head*, which consists precisely of an atom, and φ is a schema conjunction as $\psi_1 \wedge \ldots \wedge \psi_n$ for $n > 0$, where ψ_i is literal possibly negative. The set of all clauses of the same predicate symbol p on the head is called the *definition* of p. Therefore, a rule p can be defined as a disjoint set of literals. For the sake of space limitations, all predicative definitions contained in an Ontoprolog theory are not presented in this paper.

Each *semantic validation rule* represents defined queries that attempt to prove that certain unwanted property occurs. Therefore, to check if a particular relation is unreflective, there must be a semantic validity rule that tries to show that there is, at least, one reflective instance of that particular relation. If this rule (or a negative rule written in positive form) cannot be established, then, due to the assumption of the closed world inherent in all program Prolog, it is assumed in a non monotonic manner, the opposite, namely that such a relation is unreflective.

Considering that `a :: b.` is a Ontoprolog sentence, grammar instance of Code 1: $\boxed{\texttt{a :: b} \Longrightarrow dio(a, b)}$

a) `a :: b` is a sentence (syntactic) valid on Prolog where `::` is an infix operator, so that $\boxed{\texttt{a :: b.}}$ is a fact (note the full stop added to the sentence);
b) $dio(a, b)$ is a semantic relation that denotes (partially)[4] the meaning of the sentence `a :: b`, where $dio/2$ is a mnemonic for *direct instance of*; and
c) $\Longrightarrow$ represents a function that translates sentences into semantics relations.

Therefore, there is a clear separation between Ontoprolog syntax and semantics, where syntax is described by EBNF rules, and semantics is provided by combining the following:

a) *semantic translation*: the $\Longrightarrow$ represents a type of denotational semantics given by translation of sentences. They are represented by syntax trees from combining operators. It makes up a group of relations (terms/Horn Clauses), representing the semantics of translated sentences;
b) *semantic validation rules* and *notion of consistency*: the validity of rules only occur on the side of the semantic relations, and govern the validity of the models produced by these relations. Because these rules are written in a positive way, and still denote negative conditions, a theory is consistent when none of the rules succeeds, i.e. when none of them are true, or when neither "is the case".

Under the ontology description language, in fact, "consistency" is not related to a purely classical contradiction in the sense of being present in a statement and its negation (cf. [17]). A discussion on several senses of the notion of consistency is found in [18]. In Ontoprolog, there is not an explicit negation of the claims. Negations take place by the negation as failure of the closed world assumption. However, within the Ontoprolog's meta-Logic, we express the notion of consistency as validity regarding the rules.

In Ontoprolog's meta-logic, the consistency notion is defined as rules created based on definitions that may not be true in a particular model. For example, if we define that NT is a rule that indicates that an appropriate relation can not be transitive, and it is the case that an instance of the Y relation be transitive, therefore that instance of Y relation is "inconsistent", or "incompatible" or "invalid" concerning the NT rule. In other words, if it occurs in a particular model something that makes the NT (or any rule) true, then it is not the case that all rules are false. Therefore, the model is considered "inconsistent" because there is, at least, a "contradiction" with the rules. I see that in this sense, the Ontoprolog's meta-logical system is purely classical, and contradiction implies inconsistency.

However, under the ontology description language, i. e., under the system built on that classic paradigm, we have a non-classical system in which the notion of consistency is different. Although the statements about ontologies do not have negation, they can participate under validity notions that change about ontological models that they are inserted. This situation occurs because they are in a non-monotonic system, i. e., new statements could avoid deductions already made. Moreover, when we employ modalities, a particular ontology relation may be consistent with a set of statements of a particular agent, but may be inconsistent when, from a current world, it has access to statements base of another agent (in other accessible possible world). But it can be valid again when another world becomes accessible, and so on.

Another issue is that it does not say in Ontoprolog that "it is not the case that X and not X". It is expressed, always in positive form, relationships between entities. In turn, these instances of relations, in a certain world, can denote situations that imply a contradiction to some rule. However, for the Ontoprolog system, the rules themselves may be in a non-classical paradigm, as indeed occurs in the modal system, without necessarily changing interpretations of ontologies.

The rewriting of certain cases in the grammar defined in Code 1 is presented to exemplify how the translation semantics is defined. According to syntactic categories, grammar for describing the relation of *disjoint* concepts is represented by the operator `disjoint` (`DISJOINT` in Code 1), and denotes that a list of ontological entities $T_1, \ldots, T_n$ are disjoint. The operator `disjoint` is translated into semantic predicate $dd/1$, as follows:

[4] Add $entity(a)$ and $entity(b)$ in relations denoting the meaning of $\boxed{\texttt{a :: b.}}$

$[\![\texttt{disjoint [T}_1\texttt{, ... , T}_n\texttt{]}]\!]$
$\Longrightarrow$

$dd([\![\texttt{[T}_1\texttt{, ... , T}_n\texttt{]}]\!]_{tl})$
$[\![\texttt{[T}_1\texttt{, ... , T}_n\texttt{]}]\!]_{entity}$

where:

- $[\![\ \texttt{[T}_1\texttt{, ... , T}_n\texttt{]}\ ]\!]_{tl} \Longrightarrow \texttt{[}\ [\![\texttt{T}_1]\!]_t\texttt{, ... , }[\![\texttt{T}_n]\!]_t\ \texttt{]}$
- $[\![\ \texttt{[T}_1\texttt{, ... , T}_n\texttt{]}\ ]\!]_{entity} \Longrightarrow \texttt{entity(}[\![\texttt{T}_1]\!]_t\texttt{).}, \ldots, \texttt{entity(}[\![\texttt{T}_n]\!]_t\texttt{).}$
- $[\![\texttt{T}]\!]_t \Longrightarrow \texttt{T}$, if $atom(T)$ is the case for any T, and $atom/1$ happens, if the only parameter is qualified as atom by the definition of `ATOM` in Code 1, and T does not coincide with a Prolog operator

The Definition 1 exemplifies how a semantic validation rule is defined in the system. In this case, `ngrule/1` is a predicate that defines a *semantic validation rule*. Therefore, if all other semantic validity rules do not succeed, the theory is said to be "consistent with the rules." In [2] one can find all semantic validation rules, and all cases of syntax translation function in the semantic relations of Ontoprolog's basis.

Definition 1 (Rules for $dd/1$).
$dd/1$ imposes constraints on the relations of the $dio/2$ and $deo/2$, so that no entity can instantiate or extend simultaneously, two disjoint types. The following rules define the relations $dd/1$:

- *be I the entities list instantiated by E. No two $\langle i_1, i_2\rangle \in I$ can be disjoint:*

$$\begin{aligned} ngrule(dio\_disjoint\_types) \leftarrow & setof(IT, dio(E, IT), I),\\ & disjoint\_types(I). \end{aligned}$$

- *be I the extended entities lists by E. No pair $\langle i_1, i_2\rangle \in I$ can be disjoint:*

$$\begin{aligned} ngrule(deo\_disjoint\_types) \leftarrow & setof(IT, deo(E, IT), I),\\ & disjoint\_types(I). \end{aligned}$$

In Ontoprolog, the speeches are formalized as theories, instances of a metamodel, which can be an ontology or just a metatheory with basic logic elements. In this case, it is called *hypertypes metatheory* and contains most foundation concepts used to define ontologies. This metatheory is created from the Aristotelian Square and, so it contains, at least, two bases entities: `type`: are elements independent on each other, and `property`: are *moments*, and so, are dependents of types.

As an example, Code 2 is a description of an ontology as expressed in Ontoprolog. It outlines the specification of a single speech, described as an instance of a fundamental metatheory named *hypertypes metatheory*. This example shows `slotProperties`, whose fields are instances of `type`.

Code 2. "Sinatra" specification

```
:- include('../../../ontoprolog/ontoprolog').

% properties
[nome, data_morte] :: slotProperty.

% type theory
criatura :: type.
pessoa :: type extends criatura.

disjoint [vivo, morto] :: type extends criatura.

property nome on pessoa.
property data_morte on morto.

pessoa_viva :: type extends [pessoa, vivo].
pessoa_morta :: type extends [pessoa, morto].

% individual theory
lauro :: pessoa_viva.
nome at lauro := 'lauro cesar'.

sinatra :: [pessoa_morta, vivo].

:- otp_compile,
   check_semantics.
```

Code 3 shows the output of the semantic evaluation of Code 2. The predicate `opt_compile/0` is the implementation of the translation filter, that produces the clauses which indicate the semantic relations. The predicate `check_semantics/0`, in turn, applies the rules of validity produced by the filter and verifies the notion of consistency of the conceptual model. As expected, the evaluation mechanism of the theory denoted by specification finds out that there is an inconsistency about the concept `sinatra`, caught by the rule `dio_disjoint_types` (Definition 1). In this case, `sinatra` can not simultaneously instantiate types `pessoa_morta` and `vivo`, once they are disjoint concepts. The result shows the mechanism that defines the notion of consistency within the theories of Ontoprolog.

Code 3. Output of Code 2

```

Compilation of the ontoprolog specification was successful.

Checking semantics using negative rules of the current Ontoprolog theories...
--------------------------------------------------
Checking ngrule/3...
 sinatra:
   -> "sinatra" instantiates disjoint types "[pessoa_morta,vivo]". (dio_disjoint_types)

--------------------------------------------------
```

Based on the specification and in the theory, it is possible to build programs that use information about ontologies. Applications within systems can use this ontological structure description as input to decisions and evaluations of relevant rules in particular contexts.

The modal extension of Ontoprolog replaces the underlying classical logic evaluation of sentences by one of the serial modal systems from MProlog. In this case, the grammar sentences, defined in Code 1 are added to modal contexts. For reasoning about the discourse of different subjects, it uses the system $KD45_m$ for the doxastic semantics of multiple agents. As a general rule consistency evaluation, it is considered that each speech and its conjunction should be valid on the reference ontology used by subjects. Therefore, the key requirement refers to the fact that discourse should be related to the same universe of discourse, based on the same ontological commitment, to be able to identify consistencies and inconsistencies.

The syntactic form to be used in the modal context specification varies from the deductive system chosen. In this case, the option is to use the serial system $KD45_m$, in which the agents do not have access to the knowledge base of the others. In this system, the modalities have syntactic forms, `bel(i)` γ and `pos(i)` γ. They are interpreted as “the agent i beliefs that γ is true” and “γ is considered possible by the agent i”, respectively[5], where $\diamond$ is a sentence Prolog whatsoever.

Following Code 4, there are two groups of sentences for each speech regarding claims of a subject on the same shared reality. In the code, the first subject, identified as `1`, exposes the concepts of `living_person` and `dead_person`. Moreover, this agent *securely* exposes that `sinatra` is an instance of `dead_person`. This certainty of the first agent is explained by the use of the modal operator `bel`. The second group of statements consists of the agent’s claims idenfied by `2`. The agent `2` adds an important restriction on the concepts `living_person` and `dead_person` by indicating that these are disjoint concepts. Thus, according to the Definition 1, direct or indirect instances of these entities can not instantiate simultaneously `living_person` and `dead_person` entities. Finally, agent `2` says that `sinatra` is an instance of `living_person`. But he does this in a *hesitante* way, so he uses the operator `pos`. Importantly, due to the presence of the (D) axiom, all that is known certainly (`bel`) is also known possibly (`pos`).

Code 5 operates as a modal program controller. Line 2 embeds the framework of Ontoprolog-Modal, consisting of the inclusion of MProlog[6] library and Ontoprolog. Line 5 refers to epistemic modal calculations library included in the `belief.cal`. In this library, the system $KD45_m$ and others are defined. Line 8 effectively queries the modal program, described by Code 4. The lines between 10 and 22 defines the predicate `otp_m_compile/0`, which consists in:

a) producing a list LMA with the sentences Ontoprolog asserted by the agent `1`;

b) producing a list LMB with the sentences Ontoprolog asserted by the agent `2`;

c) producing a set LMB with the union of the lists LMA and LMB. Since it is a set, LMB does not have repeated assertions;

[5] This is the text version for universal ($\square$) and existential ($\diamond$) with doxastic interpretation.

[6] The MProlog is adapted so that it can be used with the framework of Ontoprolog.

d) producing a line LC with all the classical sentences present in the program. This is necessary to retrieve the base of facts, which contains the *Metatheory of hypertypes* underlining all Ontoprolog specifications;

e) producing a set $LCMAB$ with the union of $LMAB$ and LC; and

f) compiling the set of sentences $LCMAB$ as Ontoprolog specification.

Finally, the lines 23:24 effectively run the predicate `otp_m_compile/0` and the predicate `check_semantics/0`, used to verify the consistency of the model produced by the specification. Code 6 contains the output from running Code 5.

Code 4. "Sinatra" Modal specification – MProlog file

```
% Modal Epistemic Logic
:- calculus cKD45m.
:- set_option(current_calculus, cKD45m).

% -----
% 1
% -----
% type theory
[bel(1)] : otp(person :: type).
[bel(1)] : otp([living_person, dead_person] :: type extends person).

% individual theory
[pos(1)] : otp(sinatra :: dead_person).

% -----
% 2
% -----
% type theory
[bel(1)] : otp(person :: type).
[bel(1)] : otp(disjoint [living_person, dead_person] :: type extends person).

% individual theory
[pos(2)] : otp(sinatra :: living_person).
```

Code 5. "Sinatra" Modal specification – Prolog file

```
1  % Include ONTOPROLOG-MODAL libraries
   :- include('../../../ontoprolog/ontoprolog-m').

   % Consult calculi
5  :- consult_calculi('../../../ontoprolog/modal/mprolog2-custom/belief.cal').

   % Consult modal program
   :- mconsult('modal_sentence_cKD45m_sinatra.mpl').
9
   % denec
   otp_m_compile :-
      write('Getting candidate modal sentences...'), nl,
13    findall(A, mcall([pos(1)] : otp(A)), LMA),
      findall(B, mcall([pos(2)] : otp(B)), LMB),
      util_append(LMA, LMB, LMAB),

17    write('Getting candidate classical sentences...'), nl,
      setof(S, otp_classic_sentence(S), LC),
      util_append(LMAB, LC, LCMAB),

21    otp_compile(LCMAB).

   :- otp_m_compile,   % denec
      check_semantics.
```

Code 6. Output of Code 5 semantic

```

Compilation of the Ontoprolog specification was successful.

Checking semantics using negative rules of the current Ontoprolog theories...
--------------------------------------------------
Checking ngrule/3...
 sinatra:
   -> "sinatra" instantiates disjoint types "[dead_person,living_person]". (dio_disjoint_types)
```

The output shows that the consistency check based on the *rules of semantics verification* of the form $ngrule/n$ identifies a consistency problem on the concepts `dead_person` and `living_person`. This kind of result shows the ability to perform *global inconsistency verification* theories denoted by the conjunction of the individual specifications speech of each agent, i.e., agents identified as `1` and `2`. The mechanism assists in the process of reaching agreements about reality, insofar as it exposes the inconsistency in the speeches of the agents.

The sort of modal approach to the treatment of ontologies with Ontoprolog is one of the modal treatment strategies discussed in [2]. In that text, other approaches are discussed, for example relating to the insertion modal semantics, not only the specifications, as in the present case, but also in the context of theories denoted by them.

4 Outcome

Results through the application of modal logic, software engineering, and artificial intelligence, as some presented with Ontoprolog are examples of information models for automated reasoning, collaboration and exchange of experiences amongst people. Thus, this article explored the purpose of developing semantically relevant information environments of Architecture of Information, as a discipline, in which the object is information. In this sense, AI plays a crucial role in such initiatives in order to guide approaches through modeling information.

Acknowledgements

The authors acknowledge Lauro César Araujo, for his invaluable help, without which it would not have been possible to achieve the results presented here. Both authors acknowledge support from the Centre for Research on Architecture of Information of the University of Brasília, a project sponsored by the Ministry of the Environment and the National Space Agency, Brazil. The second author acknowledges support from FAPESP Thematic Project LogCons 2010/51038-0, and the National Council for Scientific and Technological Development (CNPq), Brazil.

Bibliography

1. Lima-Marques, M.: Outline of a theoretical framework of Architecture of Information: a school of Brasilia proposal. In Béziau, J., Coniglio, M.E., eds.: Logic without Frontiers: Festschrift for Walter Alexandre Carnielli on the occasion of his 60th Birthday. Volume 17 of Tribute Series. College Publications, London (2011)
2. Araujo, L.C.: Uma linguagem para formalização de discursos com base em ontologias. PhD thesis, University of Brasilia, Centre for Research on Architecture of Information, Brasília, Brazil (2015)
3. Guizzardi, G.: Ontological Foundations for Structural Conceptual Models. PhD thesis, Centre for Telematics and Information Technology, University of Twente, Enschede, The Netherlands (2005)
4. Doucette, D., Bichler, R., Hofkirchner, W., Raffl, C.: Toward a new science of information. Data Science Journal **6**(0) (2007) 198–205
5. Rosenfeld, L., Morville, P.: Information Architecture for the World Wide Web. O'Reilly, USA (1998)
6. Lima-Marques, M., Oliveira, F.L.: Arquitetura da informação: base para a gestão do conhecimento. In Tarapanoff, K., ed.: Inteligência, informação e conhecimento, Brasília, Ibict – Unesco (2006) 241–256
7. Morillo, F., Bordons, M., Gomez, I.: Interdisciplinarity in science: a tentative typology of disciplines and research areas. Journal of the American Society for Information Science and Technology **54**(13) (2003) 1237–1249
8. de Siqueira, A.H.: A Lógica e a Linguagem como fundamentos da Arquitetura da Informação. Master's thesis, University of Brasilia, Centre for Research on Architecture of Information, Brasília (2008)
9. Ranganathan, S.R.: Documentation and Its Facets: Being a Symposium of 70 Papers by 32 Authors. Asia Publishing House (1963)
10. Oliveira, F.L., Lima-Marques, M.: Information Architecture as a Discipline: a methodological approach. Human-computer Interaction Series. In: Reframing Information Architecture. Springer International Publishing, Switzerland (2014) 1–10
11. van Gigch, J.P., Pipino, L.L.: In search of a paradigm for the discipline of information systems. Future Comput. Syst. **1**(1) (1986) 71–97
12. Nguyen, L.A.: Modal logic programming revisited. Journal of Applied Non-Classical Logics **19**(2) (2009) 167–181
13. Austin, J.L.: How to do things with words. Oxford University Press, Great Britain (1962)
14. Nguyen, L.A.: Multimodal logic programming. Theoretical Computer Science **360**(1–3) (2006) 247 – 288
15. W3C: Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C. (2008)
16. Nielson, H.R., Nielson, F.: Semantics with Applications. Undergraduate topics in Computer Science. Springer-Verlag, London (2007)
17. Carnielli, W.A., Coniglio, M.E., Marcos, J.: Logics of formal inconsistency. In Gabbay, D., Guenthner, F., eds.: Handbook of Philosophical Logic. Volume 14. 2 ed. edn. Springer (2007) 15–107
18. Carnielli, W.: The single-minded pursuit of consistency and its weakness. Studia Logica **97**(1) (2011) 81–100

Temporal Equilibrium Logic with Past Operators

Felicidad Aguado[1], Pedro Cabalar[1], Martín Diéguez[2]
Gilberto Pérez[1] and Concepción Vidal[1]

[1]Department of Computer Science,
University of Corunna (Spain)
{aguado,cabalar,gperez,eicovima}@udc.es

[2]IRIT, University of Toulouse, France
martin.dieguez@irit.fr

Abstract. In this paper we study the introduction of modal past temporal operators in Temporal Equilibrium Logic, an hybrid formalism that mixes linear-time modalities and logic programs interpreted under stable models and their characterisation in terms of Equilibrium Logic.

1 Introduction

Many scenarios of commonsense reasoning require the combination of two central dimensions in Knowledge Representation (KR): Temporal Reasoning and Non-Monotonic Reasoning (NMR). Due to the strong connection between NMR and Logic Programming (LP), one interesting possibility for this aim is to rely on the literature on (temporal) modal extensions of LP. This research area dates back to the eighties with several approaches (see survey [1]), starting with the seminal work by Luis Fariñas del Cerro [2] on the MOLOG system, that introduced different types of modalities into Prolog. Other extensions [3,4,5] were specifically focused on enriching LP with temporal modalities as those handled in *Linear-time Temporal Logic* (LTL) [6,7]: $\Box$ standing for "always," $\Diamond$ standing for "eventually" or $\bigcirc$ standing for "next." However, most of them imposed some syntactic restrictions and disregarded the use of default negation. For instance, the system TEMPLOG [4] introduced a particular syntax where temporal modalities could be used in the rule bodies or as the general scope of the rule conditional in a restricted manner. An example of a TEMPLOG rule is:

$$\Box(p \leftarrow \bigcirc q \wedge \Diamond r) \tag{1.1}$$

As shown in [8], this syntax had the semantic advantage of yielding a unique, least Herbrand model for any TEMPLOG program, as happens with positive logic programs in the non-temporal case. Unfortunately, these syntactic limitations make this type of formalisms not suitable for our original purposes in KR. On the one hand, the absence of default negation is a serious drawback that prevents the representation of defaults and NMR. On the other hand, even if we focus on the temporal perspective, the way in which modal operators are used in TEMPLOG is not natural in terms of a commonsense description of a dynamic domain. Take the rule (1.1) as an example. This expression may make sense under a top-down Prolog reading: at any moment, to fulfil goal p we need to satisfy q at the next state and r at some point in the future. However, if we use a bottom-up reading, more common in causal laws used in action languages, (1.1) would assert that if q holds at the next state and r occurs in a future situation, then p is always caused to be true now. What makes an expression of this kind look unnatural is that, excepting in science fiction scripts[1], commonsense causal laws normally describe the cause-effect relations from past to future, not the other way around. For instance, if we want to express that pushing a button lights a lamp in the next situation, unless we can prove that it is broken, we would require a rule like:

$$\Box(\bigcirc light \leftarrow push \wedge \neg broken) \tag{1.2}$$

which cannot be represented in TEMPLOG, since it does not allow rule heads with $\bigcirc$ or $\Diamond$ operators, and cannot deal with default negation $\neg$.

Part of the syntactic limitations present in the temporal LP approaches from the eighties were mostly due to the fact that a satisfactory semantics for default negation in (non-temporal) LP was not successfully proposed until the last part of the decade. In 1988, Gelfond and Lifschitz [10]

[1] As an interesting formal classification of time-travel narratives, see [9].

defined the *stable model* semantics that eventually gave rise to a new LP paradigm called *Answer Set Programming* (ASP) [11,12], becoming nowadays one of the most successful frameworks for practical problem solving and KR. Moreover, as shown in [13], stable models can be logically characterised in terms of *Equilibrium Logic* [13], a formalism that defines a model selection criterion for the (monotonic) intermediate logic of *Here-and-There* (HT) [14]. The equilibrium logic characterisation has eventually allowed the definition of stable models [15] for arbitrary theories without syntactic limitations.

In principle, the extension of HT to the temporal case could be designed as an example of intuitionistic (or intermediate) modal approach, as in [16] and also studied by Luis Fariñas and Andrés Raggio in [17]. In order to obtain a temporal extension of Equilibrium Logic, one further needs to generalise the model minimisation from the latter to temporal (intuitionistic) interpretations. Such an extension of Equilibrium Logic to incorporate LTL modal operators was, in fact, proposed in a series of papers [18,19,20] under the name of *Temporal Equilibrium Logic* (TEL). TEL defines a temporal stable model semantics for any arbitrary theory, and so, it allows free combinations of the temporal operators and LP constructs. In this way, (1.2) is now representable in TEL and behaves like a standard ASP program containing the rules:

$$light(I+1) \leftarrow push(I), \neg broken(I)$$

for any integer $I \geq 0$. Moreover, we can represent other expressions that are not representable in ASP (unless we add auxiliary atoms) such as:

$$\Box(\bigcirc\Diamond light \leftarrow push \wedge \neg broken)$$

meaning this time that the light will be eventually on, but perhaps with a delay of $n \geq 1$ of situations. Although a prover has been built [21] to compute temporal stable models for arbitrary (propositional) temporal theories, such a syntactic flexibility is not so exploited in practice. If we look at the usual encoding of action scenarios in TEL, rules like (1.1) (directed from future to past) simply do not occur. In fact, this has led to the definition of a particular syntactic subset, called *splittable* [22] temporal logic programs, where formulas are constraints like $\Box(\bot \leftarrow \varphi)$ or have the form:

$$\Box(\bigcirc\alpha \leftarrow \bigcirc\beta \wedge \gamma) \tag{1.3}$$

where α is a disjunction of literals (an atom or its negation), and β and γ are conjunctions of literals. For instance, (1.2) is in splittable form. The temporal stable models for a splittable program can always be represented as LTL models of another temporal theory[2] and so can be computed [23] using an LTL model checker as a backend.

Splittable programs cover most examples of transition-based action domains in the literature and allow an arbitrary use of temporal operators in constraints. However, their expressiveness for describing causal laws is limited to (1.3) where the use of temporal operators is rather restrictive. Moreover, as discussed before, even if they are extended to allow more expressive operators in the rule body, as in (1.1), the expressions we obtain seem awkward because they would describe causation from future to past.

A more natural choice for handling expressive modalities in causal laws would be *using past operators* in the rule bodies (that express the law precondition) and using future operators for the rule heads or for the constraints describing the valid narratives. As an example, suppose that the lamp takes a pair of situations to "warm up" if we pushed the button for the first time:

$$\Box(\bigcirc\bigcirc light \leftarrow push \wedge \boxminus\neg push) \tag{1.4}$$

where $\boxminus$ stands for "it has always been true." Of course, we can represent this example without past operators if we introduce an auxiliary predicate to remember that *push* has been true before:

$$\Box(\bigcirc\bigcirc light \leftarrow push \wedge \neg pushed)$$
$$\Box(\bigcirc pushed \leftarrow push)$$
$$\Box(\bigcirc pushed \leftarrow pushed)$$

However, in the general case, past operators allow much more flexible and compact queries on the past narrative without the need of introducing auxiliary atoms, which may become a potential source of errors in the specification.

[2] It is still unknown whether this property also holds for any TEL theory or not.

Another justification for the introduction of past operators relies on the fact that recent implementations of ASP solvers for incremental [24] and stream reasoning [25] which allow multiple-shot execution of the solver, can exploit the search done in previous shots if the time variable in the rules refers to the current instant in the head and previous instants in the body. For instance, for this purpose, we would rather be interested in representing (1.4) as the equivalent formula:

$$\Box(light \leftarrow \ominus\ominus(push \wedge \boxminus\neg pushed))$$

where $\ominus$ means "in the previous state."

It has been proved [26,27] that LTL with past operators it can be translated into an equivalent pure future formula evaluated at the beginning of the path. Still, as shown in [28], any LTL with past is exponentially more succinct[3] than pure-future LTL.

In this paper, we consider an extension of TEL (and THT) to include past operators and show that this extension can be reduced to pure future TEL by a translation that introduces auxiliary atoms.

2 Temporal Equilibrium Logic with past operators

2.1 Syntax

The logic of *Linear Temporal Here-and-There* (THT) is defined as follows. We start from a finite set of atoms $\mathcal{L}_V$ called the *propositional signature*. The syntax of THT is the one from propositional LTL which we recall below. A temporal formula φ is defined as:

$$\begin{aligned}\varphi ::= {} & \bot \mid p \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \bigcirc\varphi_1 \mid \Box\varphi_1 \mid \Diamond\varphi_1 \mid \varphi_1 \, \mathcal{U} \, \varphi_2 \mid \varphi_1 \, \mathcal{R} \, \varphi_2 \mid \\ & \widehat{\ominus}\varphi_1 \mid \ominus\varphi_1 \mid \boxminus\varphi_1 \mid \Diamond\!\!\!-\varphi_1 \mid \varphi_1 \mathcal{S} \, \varphi_2 \mid \varphi_1 \mathcal{T} \, \varphi_2\end{aligned}$$

where φ_1 and φ_2 are temporal formulas in their turn and p is any atom. Negation is defined as $\neg\varphi \stackrel{\text{def}}{=} \varphi \rightarrow \bot$ whereas $\top \stackrel{\text{def}}{=} \neg\bot$. Note that '$\neg$' will stand for *default negation* in all non-monotonic formalisms described in this paper. Concerning to temporal modalities, the operators can be defined in terms of $\mathcal{U}$, $\mathcal{R}$, $\mathcal{S}$ and $\mathcal{T}$:

$$\begin{array}{ll}\Diamond\varphi \stackrel{\text{def}}{=} \top \, \mathcal{U} \, \varphi & \Box\varphi \stackrel{\text{def}}{=} \bot \, \mathcal{R} \, \varphi \\ \Diamond\!\!\!-\varphi \stackrel{\text{def}}{=} \top \, \mathcal{S} \, \varphi & \boxminus\varphi \stackrel{\text{def}}{=} \bot \, \mathcal{T} \, \varphi\end{array}$$

Operator $\Box$ is read "forever" and $\Diamond$ stands for "eventually" or "at some future point." We define the following notation for a finite concatenation of $\bigcirc$'s and $\ominus$'s operators as follows:

$$\begin{array}{ll}\bigcirc^0\varphi \stackrel{\text{def}}{=} \varphi & \bigcirc^i\varphi \stackrel{\text{def}}{=} \bigcirc(\bigcirc^{i-1}\varphi) \quad (\text{with } i \geq 1) \\ \ominus^0\varphi \stackrel{\text{def}}{=} \varphi & \ominus^i\varphi \stackrel{\text{def}}{=} \ominus(\ominus^{i-1}\varphi) \quad (\text{with } i \geq 1)\end{array}$$

2.2 Semantics

An *LTL-interpretation* is an infinite sequence of sets of atoms $H_0, H_1, \ldots$ with $H_i \subseteq At$, $i \geq 0$. Given two LTL-interpretations $\mathbf{H}$ and $\mathbf{T}$, we write $\mathbf{H} \leq \mathbf{T}$ to stand for $H_i \subseteq T_i$ for all $i \geq 0$. As usual, $\mathbf{H} < \mathbf{T}$ represents $\mathbf{H} \leq \mathbf{T}$ and $\mathbf{H} \neq \mathbf{T}$, that is, the inclusion relation holds in all states but is strict $H_j \subset T_j$ for some $j \geq 0$. A *THT-interpretation* $\mathbf{M}$ is a pair of LTL-interpretations $\mathbf{M} = \langle \mathbf{H}, \mathbf{T} \rangle$, respectively standing for *here* and *there*, such that $\mathbf{H} \leq \mathbf{T}$. An interpretation $\mathbf{M} = \langle \mathbf{H}, \mathbf{T} \rangle$ is said to be *total* when $\mathbf{H} = \mathbf{T}$.

Definition 1 (THT-Satisfaction). *We say that an interpretation $\mathbf{M} = \langle \mathbf{H}, \mathbf{T} \rangle$ satisfies a formula φ at state $k \in \mathbb{N}$, written $\mathbf{M}, k \models \varphi$, when the following recursive conditions hold:*

1. $\mathbf{M}, k \models p$ *iff* $p \in H_k$, *for any* $p \in At$.
2. $\mathbf{M}, k \models \varphi \wedge \psi$ *iff* $\mathbf{M}, k \models \varphi$ *and* $\mathbf{M}, k \models \psi$.

[3] Assuming that no auxiliary atoms are introduced.

3. $\mathbf{M}, k \models \varphi \vee \psi$ *iff* $\mathbf{M}, k \models \varphi$ *or* $\mathbf{M}, k \models \psi$.
4. $\mathbf{M}, k \models \varphi \rightarrow \psi$ *iff for all* $\mathbf{H}' \in \{\mathbf{H}, \mathbf{T}\}$, $\langle \mathbf{H}', \mathbf{T}\rangle, k \not\models \varphi$ *or* $\langle \mathbf{H}', \mathbf{T}\rangle, k \models \psi$.
5. $\mathbf{M}, k \models \bigcirc\varphi$ *iff* $\mathbf{M}, k+1 \models \varphi$.
6. $\mathbf{M}, k \models \widehat{\ominus}\varphi$ *iff* $\begin{cases} \mathbf{M}, k-1 \models \varphi & \textit{if } k > 0 \\ \textit{false} & \textit{if } k = 0 \end{cases}$
7. $\mathbf{M}, k \models \ominus\varphi$ *iff* $\begin{cases} \mathbf{M}, k-1 \models \varphi & \textit{if } k > 0 \\ \textit{true} & \textit{if } k = 0 \end{cases}$
8. $\mathbf{M}, k \models \varphi\ \mathcal{U}\ \psi$ *iff there is* $j \geq k$ *s.t.* $\mathbf{M}, j \models \psi$ *and* $\mathbf{M}, i \models \varphi$ *for all* $i, k \leq i < j$.
9. $\mathbf{M}, k \models \varphi\ \mathcal{R}\ \psi$ *iff for all* $j \geq k$ *s.t.* $\mathbf{M}, j \models \psi$ *or* $\mathbf{M}, i \models \varphi$ *for some* $i, k \leq i < j$.
10. $\mathbf{M}, k \models \varphi\ \mathcal{S}\ \psi$ *iff there is* j, $0 \leq j \leq k$ *s.t.* $\mathbf{M}, j \models \psi$ *and* $\mathbf{M}, i \models \varphi$ *for all* $i, j < i \leq k$.
11. $\mathbf{M}, k \models \varphi\ \mathcal{T}\ \psi$ *iff for all* j, $0 \leq j \leq k$ *s.t.* $\mathbf{M}, j \models \psi$ *or* $\mathbf{M}, i \models \varphi$ *for some* $i, j < i \leq k$.
12. *never* $\mathbf{M}, k \models \perp$. ⊠

In particular, the following LTL valid formulas are also THT valid:

$$\varphi\ \mathcal{U}\ \psi \leftrightarrow \psi \vee (\varphi \wedge \bigcirc(\varphi\ \mathcal{U}\ \psi)) \tag{1.5}$$
$$\varphi\ \mathcal{R}\ \psi \leftrightarrow \psi \wedge (\varphi \vee \bigcirc(\varphi\ \mathcal{R}\ \psi)) \tag{1.6}$$
$$\varphi\ \mathcal{S}\ \psi \leftrightarrow \psi \vee (\varphi \wedge \ominus(\varphi\ \mathcal{S}\ \psi)) \tag{1.7}$$
$$\varphi\ \mathcal{T}\ \psi \leftrightarrow \psi \wedge (\varphi \vee \ominus(\varphi\ \mathcal{T}\ \psi)) \tag{1.8}$$

A formula φ is THT-*valid* if $\mathbf{M}, 0 \models \varphi$ for any $\mathbf{M}$. An interpretation $\mathbf{M}$ is a THT-*model* of a theory Γ, written $\mathbf{M} \models \Gamma$, if $\mathbf{M}, 0 \models \varphi$, for all formula $\varphi \in \Gamma$. It is not difficult to see that THT-satisfaction for a total interpretation $\langle \mathbf{T}, \mathbf{T}\rangle$ collapses to LTL-satisfaction for $\mathbf{T}$. As a result:

Observation 1 $\langle \mathbf{T}, \mathbf{T}\rangle \models \Gamma$ *in THT if and only if* $\mathbf{T} \models \Gamma$ *in LTL.* ⊠

Some total models will be said to be *in equilibrium* if they satisfy the following minimality condition in their "here" component.

Definition 2 (temporal equilibrium model). *A total THT-interpretation* $\langle \mathbf{T}, \mathbf{T}\rangle$ *is a* temporal equilibrium model *of a theory* Γ *if* $\langle \mathbf{T}, \mathbf{T}\rangle \models \Gamma$ *and there is no* $\mathbf{H} < \mathbf{T}$, *such that* $\langle \mathbf{H}, \mathbf{T}\rangle \models \Gamma$. ⊠

Since a temporal equilibrium model is a total model $\langle \mathbf{T}, \mathbf{T}\rangle$, by Observation 1, it corresponds to an LTL model $\mathbf{T}$ we will call *temporal stable model.*

Definition 3 (temporal stable model). *If* $\langle \mathbf{T}, \mathbf{T}\rangle$ *is a temporal equilibrium model of a theory* Γ *then* $\mathbf{T}$ *is called a* temporal stable model *of* Γ *(or TS-model, for short).* ⊠

Observation 2 *Given* $\mathbf{M} = \{\mathbf{H}, \mathbf{T}\}$ *and a pair of formulas* φ, ψ, *if* $\mathbf{M}(\varphi) = \mathbf{M}(\psi)$ *then also* $\mathbf{T}(\varphi) = \mathbf{T}(\psi)$. ⊠

We can alternatively represent any interpretation $\mathbf{M} = \langle \mathbf{H}, \mathbf{T}\rangle$ by seeing each $m_i = \langle H_i, T_i\rangle$ as a three-valued mapping $m_i : V \rightarrow \{0, 1, 2\}$ so that, for any atom p, $m_i(p) = 0$ when $p \notin T_i$ (the atom is false), $m_i(p) = 2$ when $p \in H_i$ (the atom is true), and $m_i(p) = 1$ when $p \in T_i \setminus H_i$ (the atom is undefined). We can then define a valuation for any formula φ, written[4] $\mathbf{M}(\varphi)$, by similarly considering which formulas are satisfied by $\langle \mathbf{H}, \mathbf{T}\rangle$ (which will be assigned 2), not satisfied by $\langle \mathbf{T}, \mathbf{T}\rangle$ (which will be assigned 0) or none of the two (which will take value 1). By $\mathbf{M}_i(\varphi)$ we mean the 3-valuation of φ induced by the temporal interpretation $\mathbf{M}_i$, that is, $\mathbf{M}$ shifted i positions.

Definition 4. *From the definitions in the previous section, we can easily derive the following conditions:*

1. $\mathbf{M}_i(p) \stackrel{\text{def}}{=} m_i(p)$
2. $\mathbf{M}_i(\varphi \wedge \psi) \stackrel{\text{def}}{=} min\{\mathbf{M}_i(\varphi), \mathbf{M}_i(\psi)\}$; $\quad \mathbf{M}_i(\varphi \vee \psi) \stackrel{\text{def}}{=} max\{\mathbf{M}_i(\varphi), \mathbf{M}_i(\psi)\}$
3. $\mathbf{M}_i(\varphi \rightarrow \psi) \stackrel{\text{def}}{=} \begin{cases} 2 & \textit{if } \mathbf{M}_i(\varphi) \leq \mathbf{M}_i(\psi) \\ \mathbf{M}_i(\psi) & \textit{otherwise} \end{cases}$
4. $\mathbf{M}_i(\bigcirc\varphi) \stackrel{\text{def}}{=} \mathbf{M}_{i+1}(\varphi)$

[4] We use the same name $\mathbf{M}$ for a temporal interpretation and for its induced three-valued valuation function – ambiguity is removed by the way in which it is applied (a structure or a function on formulas).

5. $\mathbf{M}_i(\ominus\varphi) \stackrel{\text{def}}{=} \begin{cases} \mathbf{M}_{i-1}(\varphi) & \text{if } i > 0 \\ 2 & \text{if } i = 0 \end{cases}$
6. $\mathbf{M}_i(\widehat{\ominus}\varphi) \stackrel{\text{def}}{=} \begin{cases} \mathbf{M}_{i-1}(\varphi) & \text{if } i > 0 \\ 0 & \text{if } i = 0 \end{cases}$
7. $\mathbf{M}_i(\varphi\ \mathcal{U}\ \psi) \stackrel{\text{def}}{=} max\{\ min\{\mathbf{M}_j(\psi), \mathbf{M}_k(\varphi) \mid i \leq k < j\} \mid j \geq i\}$
8. $\mathbf{M}_i(\varphi\ \mathcal{R}\ \psi) \stackrel{\text{def}}{=} min\{\ max\{\mathbf{M}_j(\psi), \mathbf{M}_k(\varphi) \mid i \leq k < j\} \mid j \geq i\}$
9. $\mathbf{M}_i(\varphi\ \mathcal{S}\ \psi) \stackrel{\text{def}}{=} max\{\ min\{\mathbf{M}_j(\psi), \mathbf{M}_k(\varphi) \mid j < k \leq i\} \mid j \leq i\}$
10. $\mathbf{M}_i(\varphi\ \mathcal{T}\ \psi) \stackrel{\text{def}}{=} min\{\ max\{\mathbf{M}_j(\psi), \mathbf{M}_k(\varphi) \mid j < k \leq i\} \mid j \leq i\}$

Under this alternative three-valued definition, an interpretation $\mathbf{M}$ *satisfies* a formula φ when $\mathbf{M}(\varphi) = 2$. When $\mathbf{M} = \langle \mathbf{T}, \mathbf{T} \rangle$, its induced valuation will be just written as $\mathbf{T}(\varphi)$ and obviously becomes a two-valued function, that is $\mathbf{T}(\varphi) \in \{0, 2\}$. A pair of useful observations:

Observation 3 *For any interpretation* $\mathbf{M}$*,* $\mathbf{M} \models \varphi \leftrightarrow \psi$ *iff* $\mathbf{M}(\varphi) = \mathbf{M}(\psi)$ *whereas,* $\mathbf{M} \models \Box(\varphi \leftrightarrow \psi)$ *iff for all* $i \geq 0$*,* $\mathbf{M}_i(\varphi) = \mathbf{M}_i(\psi)$*.*

Example 1 (from [29]). While in TEL we can express, for instance, that any request is eventually granted:

$$\Box\,(request \rightarrow \Diamond grant)$$

with past-time modalities, we can express that a grant should be preceeded by a request

$$\Box\,(grant \rightarrow \diamondminus request)$$

3 Translating TEL into Quantified Equilibrium Logic

Quantified Equilibrium Logic [30] (QEL) extends Equilibrium Logic to the first-order case. As in the propositional setting, QEL defines a selection of models among those from the monotonic logic of *Quantified Here and There* (QHT).

The definition of QHT is based on a first order language denoted by $\mathcal{L} = \langle C, F, P \rangle$, where C, F and P are three disjoint sets that represent constants, functions and predicates, respectively. Given a domain D we define the sets:

- $At_D(C, P)$ stands for all atomic instances that can be formed from $\langle C \cup D, F, P \rangle$.
- $T_D(C, F)$ all ground terms that can be obtained from $\langle C \cup D, F, P \rangle$.

A *QHT-interpretation*[5] is a tuple $\mathcal{M} = \langle (D, \sigma), I_h, I_t \rangle$ such that

- $\sigma : T_D(C, F) \rightarrow D$ is a mapping from ground terms into elements of the domain satisfying that $\sigma(d) = d$ if $d \in D$
- I_h, I_t are two sets of ground atoms from $At_D(C, P)$ such that $I_h \subseteq I_t$.

Given two QHT interpretations, $\mathcal{M} = \langle (D, \sigma), I_h, I_t \rangle$ and $\mathcal{M}' = \langle (D', \sigma'), I'_h, I'_t \rangle$, we say that $\mathcal{M} \leq \mathcal{M}'$ iff $D = D'$, $\sigma = \sigma'$, $I_t = I'_t$ and $I_h \subseteq I'_h$. If, additionally, $I_h \subset I'_h$ we say that the relation is strict (denoted by $\mathcal{M} < \mathcal{M}'$).

Definition 5 (QHT semantics from [30]). *The satisfaction relation for a QHT interpretation* $\mathcal{M} = \langle (D, \sigma), I_h, I_t \rangle$ *is defined as follows:*

- $\mathcal{M} \models \top$, $\mathcal{M} \not\models \bot$
- $\mathcal{M} \models p(\tau_1, \cdots, \tau_n)$ *iff* $p(\sigma(\tau_1), \cdots, \sigma(\tau_n)) \in I_h$
- $\mathcal{M} \models \tau = \tau'$ *iff* $\sigma(\tau) = \sigma(\tau')$.
- $\mathcal{M} \models \varphi \wedge \psi$ *iff* $\mathcal{M} \models \varphi$ *and* $\mathcal{M} \models \psi$
- $\mathcal{M} \models \varphi \vee \psi$ *iff* $\mathcal{M} \models \varphi$ *or* $\mathcal{M} \models \psi$
- $\mathcal{M} \models \varphi \rightarrow \psi$ *iff* $\mathcal{M} \not\models \varphi$ *or* $\mathcal{M} \models \psi$*, and* $\langle (D, \sigma), I_t, I_t \rangle \models \varphi \rightarrow \psi$
- $\mathcal{M} \models \forall x,\ \varphi(x)$ *iff* $\mathcal{M} \models \varphi(d)$*, for all* $d \in D$
- $\mathcal{M} \models \exists x,\ \varphi(x)$ *iff* $\mathcal{M} \models \varphi(d)$*, for some* $d \in D$ ⊠

[5] We assume here a version of QHT taking *static* domain and *decidable* equality. Briefly, this means that the domain D is common to worlds h and t and that equality is a "decidable" predicate, that is, it satisfies the excluded middle axiom $(x = y) \vee \neg(x = y)$.

As usual, we say that a QHT-interpretation $\mathcal{M}$ is a *model* of a first order theory Γ iff $\mathcal{M} \models \phi$ for all $\phi \in \Gamma$.

Definition 6 (quantified equilibrium model from [30]). *Let φ be a QHT formula. A QHT total interpretation $\mathcal{M}$ is a* first-order equilibrium model *of φ if $\mathcal{M} \models \varphi$ and there is no model $\mathcal{M}' < \mathcal{M}$ of φ.* ⊠

For our purposes, it is convenient to define a particular subclass of QHT theories. We define the fragment of QHT called *monadic here-and-there with inequality*, MHT($\leq$), by syntactically restricting all predicates to monadic, excepting a binary predicate $\leq$. Moreover, we also fix the domain D to be the set of natural numbers $D = \mathbb{N}$ so that $\leq$ captures the standard ordering among them. We consider the time constant 0 to stand for the initial situation. Given that both the domain and the interpretation of $\leq$ are fixed, interpretations will only vary for ground atoms in $At(\mathbb{N}, P)$, that is, those formed with the set of monadic predicates P and elements from $\mathbb{N}$. Then, MHT($\leq$) interpretations can be simply given by pairs $\langle \mathcal{H}, \mathcal{T} \rangle$ with $\mathcal{H} \subseteq \mathcal{T} \subseteq At(\mathbb{N}, P)$.

As usual, we write $x > y$ to stand for $\neg(x \leq y)$. We will also use the following abbreviations:

$$\begin{array}{ll}\forall x \geq t.\, \varphi \stackrel{\text{def}}{=} \forall x(t \leq x \rightarrow \varphi) & \forall x \in [t, z).\, \varphi \stackrel{\text{def}}{=} \forall x(t \leq x \wedge x < z \rightarrow \varphi) \\ \exists x \geq t.\, \varphi \stackrel{\text{def}}{=} \exists x(t \leq x \wedge \varphi) & \exists x \in [t, z).\, \varphi \stackrel{\text{def}}{=} \exists x(t \leq x \wedge x < z \wedge \varphi)\end{array}$$

Fragment MHT($\leq$) imposes exactly the same restrictions on QHT than the so-called *monadic first-order logic with inequality*, FOL($\leq$), does on classical First-Order Logic (FOL). This subclass of FOL was used by Kamp in his famous theorem [6] where he proved that LTL is exactly as expressive as FOL($\leq$), so that we can actually see the former as a fragment of the latter. This result was separated into two directions: proving that LTL can be translated into FOL($\leq$) and vice versa. For the first direction, Kamp defined the following translation from modal formulas into quantified first-order expressions:

Definition 7 (Kamp's translation). *Kamp's translation for a temporal formula φ and a timepoint $t \in \mathbb{N}$, denoted by $[\varphi]_t$, is recursively defined as follows:*

$$\begin{aligned}
[\bot]_t &\stackrel{\text{def}}{=} \bot \\
[p]_t &\stackrel{\text{def}}{=} p(t), \text{ with } p \in At. \\
[\neg\alpha]_t &\stackrel{\text{def}}{=} \neg[\alpha]_t \\
[\alpha \wedge \beta]_t &\stackrel{\text{def}}{=} [\alpha]_t \wedge [\beta]_t \\
[\alpha \vee \beta]_t &\stackrel{\text{def}}{=} [\alpha]_t \vee [\beta]_t \\
[\alpha \rightarrow \beta]_t &\stackrel{\text{def}}{=} [\alpha]_t \rightarrow [\beta]_t \\
[\bigcirc\alpha]_t &\stackrel{\text{def}}{=} [\alpha]_{t+1} \\
[\alpha\ \mathcal{U}\ \beta]_t &\stackrel{\text{def}}{=} \exists x \geq t.\, ([\beta]_x \wedge \forall y \in [t, x).\, [\alpha]_y) \\
[\alpha\ \mathcal{R}\ \beta]_t &\stackrel{\text{def}}{=} \forall x \geq t.\, ([\beta]_x \vee \exists y \in [t, x).\, [\alpha]_y) \\
[\ominus\alpha]_t &\stackrel{\text{def}}{=} [\alpha]_{t-1} \\
[\alpha\ \mathcal{S}\ \beta]_t &\stackrel{\text{def}}{=} \exists\ 0 \leq x \leq t.\, ([\beta]_x \wedge \forall y \in (x, t].\, [\alpha]_y) \\
[\alpha\ \mathcal{T}\ \beta]_t &\stackrel{\text{def}}{=} \forall\ 0 \leq x \leq t.\, ([\beta]_x \vee \exists y \in (x, t].\, [\alpha]_y)
\end{aligned}$$

where $[\alpha]_{t+1}$ and $[\alpha]_{t-1}$ are, respectively, abbreviations of

$$\exists y \geq t.\ \big([\alpha]_y \wedge \neg\exists z(t < z \wedge z < y)\big) \tag{1.9}$$

$$\exists y \leq t.\ \big([\alpha]_y \wedge \neg\exists z(y < z \wedge z < t)\big). \tag{1.10}$$

⊠

Note how, per each atom $p \in At$ in the temporal formula φ, we get a monadic predicate $p(x)$ in the translation.

The effect of this translation on the derived operators $\Diamond$, $\Box$, $\diamondminus$ and $\boxminus$ yields the quite natural expressions:

$$\begin{array}{ll}[\Box\alpha]_t \equiv \forall x \geq t.\, [\alpha]_x & [\Diamond\alpha]_t \equiv \exists x \geq t.\, [\alpha]_x \\ [\boxminus\alpha]_t \equiv \forall x \leq t.\, [\alpha]_x & [\diamondminus\alpha]_t \equiv \exists x \leq t.\, [\alpha]_x\end{array}$$

Definition 8 (THT-MHT($\leq$) interpretation correspondence). *Given a THT interpretation* $\mathbf{M} = \langle \mathbf{H}, \mathbf{T} \rangle$ *on a signature* $\mathcal{L}_V$*, we say that the MHT($\leq$)-interpretation* $\mathcal{M} = \langle \mathcal{H}, \mathcal{T} \rangle$ corresponds *to* $\mathbf{M}$ *iff*

- $p \in H_i$ *iff* $p(i) \in \mathcal{H}$*, for all* $i \in \mathbb{N}$.
- $p \in T_i$ *iff* $p(i) \in \mathcal{T}$*, for all* $i \in \mathbb{N}$. ⊠

We now prove that when considering this model correspondence, Kamp's translation allows us to translate a THT theory into a corresponding QHT one.

Theorem 1. *Let* φ *be a THT formula built on a set of atoms* $\mathcal{L}_V$*,* $\mathbf{M} = \langle \mathbf{H}, \mathbf{T} \rangle$ *a THT-interpretation on* $\mathcal{L}_V$ *and* $\mathcal{M} = \langle \mathcal{H}, \mathcal{T} \rangle$ *its corresponding MHT($\leq$)-interpretation from Definition 8. It holds that:*

$$\forall i \in \mathbb{N},\ \mathbf{M}, i \models \varphi \textit{ iff } \mathcal{M} \models [\varphi]_i.$$

⊠

Proof. We proceed by structural induction.

- If $\varphi = \bot$ then $[\varphi]_i = \bot$ and the result is straightforward.
- If $\varphi = p$ is an atom, then $[p]_i = p(i)$ and we get the chain of equivalent conditions: $\mathbf{M}, i \models p \Leftrightarrow p \in H_i \Leftrightarrow p(i) \in \mathcal{H} \Leftrightarrow \mathcal{M} \models p(i)$.
- If $\varphi = \alpha \wedge \beta$ we get:
$$\begin{aligned} \mathbf{M}, i \models \alpha \wedge \beta &\Leftrightarrow \mathbf{M}, i \models \alpha \text{ and } \mathbf{M}, i \models \alpha \\ &\Leftrightarrow \mathcal{M} \models [\alpha]_i \text{ and } \mathcal{M} \models [\beta]_i \quad \text{(induction on } \alpha, \beta) \\ &\Leftrightarrow \mathcal{M} \models [\alpha]_i \wedge [\beta]_i \\ &\Leftrightarrow \mathcal{M} \models [\alpha \wedge \beta]_i \end{aligned}$$
- The proof for $\varphi = \alpha \vee \beta$ is analogous to the one for $\alpha \wedge \beta$.
- If $\varphi = \alpha \rightarrow \beta$ we get:
$$\mathbf{M}, i \models \alpha \rightarrow \beta \Leftrightarrow \text{ for any } w \in \{\mathbf{H}, \mathbf{T}\},\ \langle w, \mathbf{T} \rangle, i \not\models \alpha \text{ or } \langle w, \mathbf{T} \rangle, i \models \beta$$
Now, since the THT-interpretation $\langle \mathbf{T}, \mathbf{T} \rangle$ also corresponds to the MHT($\leq$) interpretation $\langle \mathcal{T}, \mathcal{T} \rangle$ we can apply induction on subformulas, so that we continue with the equivalent conditions:
$$(\text{ for any } w \in \{\mathcal{H}, \mathcal{T}\},\ \langle w, \mathcal{T} \rangle \not\models [\alpha]_i \text{ or } \langle w, \mathcal{T} \rangle \models [\beta]_i) \Leftrightarrow (\langle \mathcal{H}, \mathcal{T} \rangle \models [\alpha \rightarrow \beta]_i)\,.$$
- If $\varphi = \bigcirc\alpha$ we get the equivalent conditions:
$$\begin{aligned} \mathbf{M}, i \models \bigcirc\alpha &\Leftrightarrow \mathbf{M}, i+1 \models \alpha \\ &\Leftrightarrow \mathcal{M} \models [\alpha]_{i+1} \quad \text{(by induction)} \\ &\Leftrightarrow \mathcal{M} \models [\bigcirc\alpha]_i \end{aligned}$$
- If $\varphi = \ominus\alpha$ we get the equivalent conditions:
$$\begin{aligned} \mathbf{M}, i \models \ominus\alpha &\Leftrightarrow \mathbf{M}, i-1 \models \alpha \\ &\Leftrightarrow \mathcal{M} \models [\alpha]_{i-1} \quad \text{(by induction)} \\ &\Leftrightarrow \mathcal{M} \models [\ominus\alpha]_i \end{aligned}$$
- If $\varphi = \alpha\ \mathcal{U}\ \beta$ we get the equivalent conditions:
$$\begin{aligned} \mathbf{M}, i \models \alpha\ \mathcal{U}\ \beta &\Leftrightarrow \exists k \text{ s.t. } k \geq i \text{ and } \mathbf{M}, k \models \beta \text{ and } \forall j \in \{i, \ldots, k-1\},\ \mathbf{M}, j \models \alpha \\ &\Leftrightarrow \exists k \text{ s.t. } k \geq i \text{ and } \mathcal{M} \models [\beta]_k \text{ and } \forall j \in \{i, \ldots, k-1\},\ \mathcal{M} \models [\alpha]_j{}^{6} \\ &\Leftrightarrow \exists k \text{ s.t. } k \geq i \text{ and } \mathcal{M} \models [\beta]_k \text{ and } \forall j \text{ if } i \leq j < k \text{ then } \mathcal{M} \models [\alpha]_j \\ &\Leftrightarrow \mathcal{M} \models [\alpha\ \mathcal{U}\ \beta]_i. \end{aligned}$$
- The proof for $\varphi = \alpha\ \mathcal{R}\ \beta$ is analogous to the one for $\alpha\ \mathcal{U}\ \beta$.
- If $\varphi = \alpha\ \mathcal{S}\ \beta$ we get the equivalent conditions:
$$\begin{aligned} \mathbf{M}, i \models \alpha\ \mathcal{S}\ \beta &\Leftrightarrow \exists k \text{ s.t. } 0 \leq k \leq i \text{ and } \mathbf{M}, k \models \beta \text{ and } \forall j \in \{k+1, \ldots, i\},\ \mathbf{M}, j \models \alpha \\ &\Leftrightarrow \exists k \text{ s.t. } 0 \leq k \leq i \text{ and } \mathcal{M} \models [\beta]_k \text{ and } \forall j \in \{k+1, \ldots, i\},\ \mathcal{M} \models [\alpha]_j{}^{7} \\ &\Leftrightarrow \exists k \text{ s.t. } 0 \leq k \leq i \text{ and } \mathcal{M} \models [\beta]_k \text{ and } \forall j \text{ if } k < j \leq i \text{ then } \mathcal{M} \models [\alpha]_j \\ &\Leftrightarrow \mathcal{M} \models [\alpha\ \mathcal{S}\ \beta]_i. \end{aligned}$$
- The proof for $\varphi = \alpha\ \mathcal{T}\ \beta$ is analogous to the one for $\alpha\ \mathcal{S}\ \beta$.

Corollary 1. *Let* $\mathbf{T}$ *be a temporal interpretation,* $\mathcal{T}$ *its corresponding first-order interpretation and* φ *some temporal formula. Then,* $\mathbf{T}$ *is a TS-model of* φ *iff* $\mathcal{T}$ *is a stable model of* $[\varphi]_0$. ⊠

[6] Here we apply the induction hypothesis on α and β.

[7] As happens in the proof for the operator $\mathcal{U}$, this step of the proof comes from the application of the induction hypothesis on α and β.

4 Removing past operators

Proposition 1. *For any THT formulas φ and ψ built on the signature $\mathcal{L}_V$, the following formulas are tautologies en THT.*
(T. 1) $(\ominus\varphi \leftrightarrow \top)$
(T. 2) $(\widehat{\ominus}\varphi \leftrightarrow \bot)$
(T. 3) $(\varphi\mathcal{S}\psi \leftrightarrow \psi)$
(T. 4) $(\varphi\mathcal{T}\psi \leftrightarrow \psi)$
(T. 5) $\Box(\bigcirc\ominus\varphi \leftrightarrow \varphi)$
(T. 6) $\Box(\bigcirc(\varphi\mathcal{S}\psi) \leftrightarrow (\bigcirc\psi \vee (\bigcirc\varphi \wedge \varphi\mathcal{S}\psi)))$
(T. 7) $\Box(\bigcirc(\varphi\mathcal{T}\psi) \leftrightarrow (\bigcirc\psi \wedge (\bigcirc\varphi \vee \varphi\mathcal{T}\psi)))$.

Proof. (T. 1)-(T. 4) follow directly from Definition 1, we following prove the cases (T. 5)-(T. 7).
(T. 5)

$$\begin{aligned}\mathbf{M}_0(\Box(\bigcirc\ominus\varphi \leftrightarrow \varphi)) = 2 &\Leftrightarrow \forall i \geq 0, \mathbf{M}_i(\bigcirc\ominus\varphi) = \mathbf{M}_i(\varphi)\\ &\Leftrightarrow \forall i \geq 0, \mathbf{M}_{i+1}(\ominus\varphi) = \mathbf{M}_i(\varphi)\\ &\Leftrightarrow \forall i \geq 0, \mathbf{M}_i(\varphi) = \mathbf{M}_i(\varphi)\\ &\Leftrightarrow \top\end{aligned}$$

(T. 6)

$$\begin{aligned}&\mathbf{M}_0(\Box(\bigcirc(\varphi\mathcal{S}\psi) \leftrightarrow (\bigcirc\psi \vee (\bigcirc\varphi \wedge \varphi\mathcal{S}\psi)))) = 2\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{S}\psi)) = \mathbf{M}_i(\bigcirc\psi \vee (\bigcirc\varphi \wedge \varphi\mathcal{S}\psi))\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{S}\psi)) = max\{\mathbf{M}_{i+1}(\psi), min\{\mathbf{M}_{i+1}(\varphi), \mathbf{M}_i(\varphi\mathcal{S}\psi)\}\}\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{S}\psi)) = max\{\mathbf{M}_{i+1}(\psi), min\{\mathbf{M}_{i+1}(\varphi), \mathbf{M}_{i+1}(\ominus(\varphi\mathcal{S}\psi))\}\}\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{S}\psi)) = max\{\mathbf{M}_{i+1}(\psi), \mathbf{M}_{i+1}(\varphi \wedge \ominus(\varphi\mathcal{S}\psi))\}\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{S}\psi)) = \mathbf{M}_{i+1}(\psi \vee (\varphi \wedge \ominus(\varphi\mathcal{S}\psi))).\end{aligned}$$

Finally, by applying (1.7) we conclude:

$$\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{S}\psi)) = \mathbf{M}_{i+1}(\varphi\mathcal{S}\psi) \Leftrightarrow \top.$$

(T. 7)

$$\begin{aligned}&\mathbf{M}_0(\Box(\bigcirc(\varphi\mathcal{T}\psi) \leftrightarrow (\bigcirc\psi \wedge (\bigcirc\varphi \vee \varphi\mathcal{T}\psi)))) = 2\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{T}\psi)) = \mathbf{M}_i(\bigcirc\psi \wedge (\bigcirc\varphi \vee \varphi\mathcal{T}\psi))\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{T}\psi)) = min\{\mathbf{M}_{i+1}(\psi), max\{\mathbf{M}_{i+1}(\varphi), \mathbf{M}_i(\varphi\mathcal{T}\psi)\}\}\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{T}\psi)) = min\{\mathbf{M}_{i+1}(\psi), max\{\mathbf{M}_{i+1}(\varphi), \mathbf{M}_{i+1}(\ominus(\varphi\mathcal{T}\psi))\}\}\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{T}\psi)) = min\{\mathbf{M}_{i+1}(\psi), \mathbf{M}_{i+1}(\varphi \vee \ominus(\varphi\mathcal{T}\psi))\}\\ \Leftrightarrow\ &\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{T}\psi)) = \mathbf{M}_{i+1}(\psi \wedge (\varphi \vee \ominus(\varphi\mathcal{T}\psi))).\end{aligned}$$

Finally, by applying (1.8) we conclude:

$$\forall i \geq 0\ \mathbf{M}_i(\bigcirc(\varphi\mathcal{T}\psi)) = \mathbf{M}_{i+1}(\varphi\mathcal{T}\psi) \Leftrightarrow \top.$$

Definition 9 (Labelling). *Let γ and χ be two THT formulas in $\mathcal{L}_V$ such that the latter is of the form $\ominus\varphi$, $\widehat{\ominus}\varphi$, $\varphi\mathcal{S}\psi$ or $\varphi\mathcal{T}\psi$. We define $\gamma^{\chi}_{\mathbf{L}_\chi}$ on $V_{\mathbf{L}} = \mathcal{L}_V \cup \{\mathbf{L}_\chi\}$, with $\mathbf{L}_\chi$ being a fresh atom, as follows:*

$$\gamma^{\chi}_{\mathbf{L}_\chi} = \begin{cases} \bot & if\ \gamma = \bot\\ p & if\ \gamma = p \in \mathcal{L}_V\\ \bigcirc\varphi^{\chi}_{\mathbf{L}_\chi} & if\ \gamma = \bigcirc\varphi\\ \varphi^{\chi}_{\mathbf{L}_\chi} \odot \psi^{\chi}_{\mathbf{L}_\chi} & if\ \gamma = \varphi \odot \psi\ and\ \odot \in \{\wedge, \vee, \rightarrow, \mathcal{U}, \mathcal{R}\}\\ \ominus(\varphi^{\chi}_{\mathbf{L}_\chi}) & if\ \gamma = \ominus\varphi\ and\ \gamma \neq \chi\\ \widehat{\ominus}(\varphi^{\chi}_{\mathbf{L}_\chi}) & if\ \gamma = \widehat{\ominus}\varphi\ and\ \gamma \neq \chi\\ (\varphi^{\chi}_{\mathbf{L}_\chi})\ \mathcal{S}\ (\psi^{\chi}_{\mathbf{L}_\chi}) & if\ \gamma = \varphi\mathcal{S}\psi\ and\ \gamma \neq \chi\\ (\varphi^{\chi}_{\mathbf{L}_\chi})\ \mathcal{T}\ (\psi^{\chi}_{\mathbf{L}_\chi}) & if\ \gamma = \varphi\mathcal{T}\psi\ and\ \gamma \neq \chi\\ \mathbf{L}_\chi & if\ \gamma = \chi \end{cases}$$

Broadly speaking, $\gamma^{\chi}_{\mathbf{L}_\chi}$ results from replacing every occurrence of χ by $\mathbf{L}_\chi$ in the subformulas of γ. ⊠

Definition 10. *Given a THT formula χ in $\mathcal{L}_V$, of the form $\ominus\varphi$, $\widehat{\ominus}\varphi$, $\varphi\mathcal{S}\psi$ or $\varphi\mathcal{T}\psi$ and a THT interpretation (in three-valued form) $\mathbf{M}$, we denote by $\mathbf{M}^e$ the following THT interpretation built on $V_{\mathbf{L}} = \mathcal{L}_V \cup \{\mathbf{L}_\chi\}$:*

$$\mathbf{M}_i^e(p) = \begin{cases} \mathbf{M}_i(\chi) \ if \ p = \mathbf{L}_\chi \\ \mathbf{M}_i(p) \ if \ p \in \mathcal{L}_V \end{cases} \tag{1.11}$$

⊠

With χ we define $df(\chi)$ as follows:

$$df(\chi) \stackrel{\text{def}}{=} \begin{cases} \Box(\circ\mathbf{L}_\chi \leftrightarrow \varphi) \wedge (\mathbf{L}_\chi \leftrightarrow \top) & \text{if } \gamma = \ominus\varphi; \\ \Box(\circ\mathbf{L}_\chi \leftrightarrow \varphi) \wedge (\mathbf{L}_\chi \leftrightarrow \bot) & \text{if } \gamma = \widehat{\ominus}\varphi; \\ \Box(\circ\mathbf{L}_\chi \leftrightarrow \circ\psi \vee (\circ\varphi \wedge \mathbf{L}_\chi)\ ) \wedge (\mathbf{L}_\chi \leftrightarrow \psi) & \text{if } \gamma = (\varphi \ \mathcal{S} \ \psi); \\ \Box(\circ\mathbf{L}_\chi \leftrightarrow \circ\psi \wedge (\circ\varphi \vee \mathbf{L}_\chi)\ ) \wedge (\mathbf{L}_\chi \leftrightarrow \psi) & \text{if } \gamma = (\varphi \ \mathcal{T} \ \psi). \end{cases}$$

From Proposition 1 and the definition of $\mathbf{M}^e$, it is easy to determine that $df(\chi)$ is always a tautology in $\mathbf{M}^e$.

Lemma 1. *Let γ and χ be two THT formulas in $\mathcal{L}_V$ such that the latter is of the form $\ominus\varphi$, $\widehat{\ominus}\varphi$, $\varphi\mathcal{S}\psi$ or $\varphi\mathcal{T}\psi$. If $\mathbf{M}$ is a model of γ then the THT interpretation on $V_{\mathbf{L}}$, $\mathbf{M}^e$ defined before satifies that $\mathbf{M} = \mathbf{M}^e \cap V$ and also:*

$$\mathbf{M}^e \models \{\gamma^\chi_{\mathbf{L}_\chi}\} \cup \{df(\chi)\}.$$

Proof. For any atom $p \in \mathcal{L}_V$ and $i \geq 0$, $\mathbf{M}_i^e(p) = \mathbf{M}_i(p)$, thus, the valuations for atoms in $\mathbf{M}$ and $\mathbf{M}^e$ coincide, which means that $\mathbf{M} = \mathbf{M}^e \cap V$. Furthermore, since γ does not have labels and $\mathbf{M} \models \gamma$, this means that

$$\mathbf{M}_0(\gamma) \stackrel{(1.11)}{=} \mathbf{M}_0^e(\gamma) = 2.$$

On the other hand, if $\gamma = \chi$ we get that

$$\mathbf{M}_0^e(\gamma^\chi_{\mathbf{L}_\chi}) = \mathbf{M}_0^e(\mathbf{L}_\chi) = \mathbf{M}_0(\chi) = 2.$$

To prove that $\mathbf{M}^e$ satisfies the translation, it remains to be shown that $\mathbf{M}^e \models df(\gamma)$. This proof comes directly from Proposition 1 and the fact that $\mathbf{M}_i^e(\mathbf{L}_\chi) = \mathbf{M}_i(\chi)$, for all $i \geq 0$.

Lemma 2. *Let γ be a THT formula in $\mathcal{L}_V$ and $\mathbf{M}$ a THT interpretation such that*

$$\mathbf{M}^e \models \{\gamma^\chi_{\mathbf{L}_\chi}\} \cup \{df(\chi)\}.$$

For any THT formula χ of the form $\ominus\varphi$, $\widehat{\ominus}\varphi$, $\varphi\mathcal{S}\psi$ or $\varphi\mathcal{T}\psi$ and any $i \geq 0$, the following property holds:

$$\mathbf{M}_i^e(\gamma^\chi_{\mathbf{L}_\chi}) = \mathbf{M}_i^e(\gamma).$$

Proof. We use structural induction on γ.

1. When the subformula γ has the shape $\top$, $\bot$ or an atom p this is trivial, since $\gamma^\chi_{\mathbf{L}_\chi} = \gamma$ by definition.
2. When $\gamma = \varphi \bullet \psi$ for any connective $\bullet \in \{\wedge, \vee, \rightarrow\}$, then the proof follows from Definition 4 and by applying induction on $\varphi^\chi_{\mathbf{L}_\chi}$ and $\psi^\chi_{\mathbf{L}_\chi}$.

   To finish the proof, notice that $df(\chi)$ is always a tautology in $\mathbf{M}^e$.

3. When $\gamma = \circ\varphi$:
$$\begin{aligned} \mathbf{M}_i^e((\circ\varphi)^\chi_{\mathbf{L}_\chi}) &= \mathbf{M}_i^e(\circ\varphi^\chi_{\mathbf{L}_\chi}) \\ &= \mathbf{M}_{i+1}^e(\varphi^\chi_{\mathbf{L}_\chi}) \\ &= \mathbf{M}_{i+1}^e(\varphi) \qquad \text{(induction)} \\ &= \mathbf{M}_i^e(\circ\varphi) \end{aligned}$$
4. $\gamma = (\varphi \ \mathcal{U} \ \psi)$: we get
$$\begin{aligned} \mathbf{M}_i^e((\varphi \ \mathcal{U} \ \psi)^\chi_{\mathbf{L}_\chi}) &= \mathbf{M}_i^e((\varphi)^\chi_{\mathbf{L}_\chi} \ \mathcal{U} \ (\psi)^\chi_{\mathbf{L}_\chi}) \\ &= max\{min\{\mathbf{M}_j^e(\psi^\chi_{\mathbf{L}_\chi}), \mathbf{M}_k^e(\varphi^\chi_{\mathbf{L}_\chi}) \mid i \leq k < j\} \mid j \geq i\} \\ &= max\{min\{\mathbf{M}_j^e(\psi), \mathbf{M}_k^e(\varphi) \mid i \leq k < j\} \mid j \geq i\} \ \text{(induction)} \\ &= \mathbf{M}_i^e(\varphi\mathcal{U}\psi) \end{aligned}$$

5. The proof for $\gamma = \varphi \mathcal{R} \psi$ is similar to the one presented in the previous case.
6. $\gamma = \ominus\varphi$: here we must consider two cases; when $\gamma \neq \chi$ and $\gamma = \chi$. In the former case proceed as follows:

$$\begin{aligned} \mathbf{M}^e_i((\ominus\varphi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_i(\ominus(\varphi)^{\chi}_{\mathbf{L}_\chi}) && (\gamma \neq \chi)\\ &= \mathbf{M}^e_{i-1}((\varphi)^{\chi}_{\mathbf{L}_\chi}) \\ &= \mathbf{M}^e_{i-1}(\varphi) && \text{(induction)}\\ &= \mathbf{M}^e_i(\ominus\varphi) \end{aligned}$$

while if $\ominus\varphi = \chi$ and $i = 0$, we have that:

$$\begin{aligned} \mathbf{M}^e_0((\ominus\varphi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_0(\mathbf{L}_\chi) && (\chi = \gamma)\\ &= \mathbf{M}^e_0(\top) && df(\chi)\\ &= \mathbf{M}^e_0(\ominus\varphi). && \text{((T. 1) from Prop. 1)} \end{aligned}$$

When $i > 0$:

$$\begin{aligned} \mathbf{M}^e_i((\ominus\varphi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_i(\mathbf{L}_\chi) && (\chi = \gamma)\\ &= \mathbf{M}^e_{i-1}(\bigcirc\mathbf{L}_\chi)\\ &= \mathbf{M}^e_{i-1}(\varphi) && df(\chi)\\ &= \mathbf{M}^e_i(\ominus\varphi) && \text{(Definition 4)} \end{aligned}$$

7. $\gamma = \widehat{\ominus}\varphi$: following the same reasoning as above, for the cases $\gamma \neq \chi$ and $\gamma = \chi$ (with $i > 0$), the argument we used coincides with the case $\gamma = \ominus\varphi$. The proof for the remaining case, $\gamma = \chi$ and $i = 0$ is presented below:

$$\begin{aligned} \mathbf{M}^e_0((\widehat{\ominus}\varphi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_0(\mathbf{L}_\chi) && (\chi = \gamma)\\ &= \mathbf{M}^e_0(\bot) && df(\chi)\\ &= \mathbf{M}^e_0(\widehat{\ominus}\varphi) && \text{((T. 2) from Prop. 1)} \end{aligned}$$

8. $\gamma = \varphi\mathcal{S}\psi$: as in the previous cases, if $\gamma \neq \chi$ we have that:

$$\begin{aligned} \mathbf{M}^e_i((\varphi\ \mathcal{S}\ \psi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_i((\varphi)^{\chi}_{\mathbf{L}_\chi}\ \mathcal{S}\ (\psi)^{\chi}_{\mathbf{L}_\chi}) && (\chi \neq \gamma)\\ &= max\{min\{\mathbf{M}^e_j(\psi^{\chi}_{\mathbf{L}_\chi}), \mathbf{M}^e_k(\varphi^{\chi}_{\mathbf{L}_\chi}) \mid j < k \leq i\} \mid j \leq i\}\\ &= max\{min\{\mathbf{M}^e_j(\psi), \mathbf{M}^e_k(\varphi) \mid j < k \leq i\} \mid j \leq i\} && \text{(induction)}\\ &= \mathbf{M}^e_i(\varphi\mathcal{S}\psi). \end{aligned}$$

On the other hand, if $\gamma = \chi$ and $i = 0$, we have that:

$$\begin{aligned} \mathbf{M}^e_0((\varphi\ \mathcal{S}\ \psi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_0(\mathbf{L}_\chi) && (\chi = \gamma)\\ &= \mathbf{M}^e_0(\psi) && df(\chi)\\ &= \mathbf{M}^e_0(\varphi\mathcal{S}\psi). && \text{((T. 3) from Prop. 1)} \end{aligned}$$

However, in the case $i > 0$ we can only prove that:

$$\begin{aligned} \mathbf{M}^e_i((\varphi\ \mathcal{S}\ \psi)^{\chi}_{\mathbf{L}_\chi}) &= \mathbf{M}^e_i(\mathbf{L}_\chi) && (\chi = \gamma)\\ &= \mathbf{M}^e_{i-1}(\bigcirc\mathbf{L}_\chi)\\ &= \mathbf{M}^e_{i-1}(\bigcirc\psi \vee (\bigcirc\varphi \wedge \mathbf{L}_\chi)) && (df(\chi))\\ &= max\{\mathbf{M}^e_{i-1}(\bigcirc\psi), min\{\mathbf{M}^e_{i-1}(\bigcirc\varphi), \mathbf{M}^e_{i-1}(\mathbf{L}_\chi)\}\}\\ &= max\{\mathbf{M}^e_i(\psi), min\{\mathbf{M}^e_i(\varphi), \mathbf{M}^e_{i-1}(\mathbf{L}_\chi)\}\}. \end{aligned}$$

Unfortunately, we cannot get rid of $\mathbf{L}_\chi$, since χ itself is the formula to be proved in the induction step. To prove that $\mathbf{M}^e_i(\mathbf{L}_\chi) = \mathbf{M}^e_i(\chi)$, we will equivalently show that

$$\forall i \geq 0\ \mathbf{M}^e_i(\mathbf{L}_\chi) = \mathbf{M}^e_i(\varphi\mathcal{S}\psi).$$

For the base case ($i = 0$) we proceed as follows:

$$\begin{aligned} \mathbf{M}^e_0(\mathbf{L}_\chi) &= \mathbf{M}^e_0(\psi) && (df(\chi))\\ &= \mathbf{M}^e_0(\varphi\mathcal{S}\ \psi) && \text{((T. 3) from Prop. 1).} \end{aligned}$$

For the inductive step we have:

$$\begin{aligned} \mathbf{M}^e_i(\mathbf{L}_\chi) &= \mathbf{M}^e_{i-1}(\bigcirc\mathbf{L}_\chi)\\ &= \mathbf{M}^e_{i-1}(\bigcirc\psi \vee (\bigcirc\varphi \wedge \mathbf{L}_\chi)) && (df(\chi))\\ &= max\{\mathbf{M}^e_i(\psi),\ min\{\mathbf{M}^e_i(\varphi), \mathbf{M}^e_{i-1}(\mathbf{L}_\chi)\}\}\\ &= max\{\mathbf{M}^e_i(\psi),\ min\{\mathbf{M}^e_i(\varphi), \mathbf{M}^e_{i-1}(\varphi\mathcal{S}\ \psi)\}\} && (Induction)\\ &= \mathbf{M}^e_i(\psi \vee (\varphi \wedge \ominus(\varphi\mathcal{S}\ \psi))) && \text{(Equivalence (1.7))}\\ &= \mathbf{M}^e_i(\varphi\mathcal{S}\ \psi). \end{aligned}$$

9. The proof for $\gamma = \varphi\mathcal{T}\psi$ is similar to the proof for $\gamma = \varphi\mathcal{S}\psi$.

Theorem 2. *Let γ in $\mathcal{L}_V$ be a formula and χ one of its subformulas whose form is $\ominus\varphi$, $\widehat{\ominus}\varphi$, $\varphi\mathcal{S}\psi$ or $\varphi\mathcal{T}\psi$. It holds that:*

$$\{\mathbf{M} \mid \mathbf{M} \models \gamma\} = \{\mathbf{M}^e \cap V \mid \mathbf{M}^e \models \{\gamma^{\chi}_{\mathbf{L}_\chi}\} \cup \{df(\chi)\}.$$

Proof. The '$\subseteq$' inclusion immediately follows from Lemma 1. For proving the '$\supseteq$' inclusion, suppose that we have $\mathbf{M}^e$ such that $\mathbf{M}^e \models \gamma^{\chi}_{\mathbf{L}_\chi} \wedge df(\chi)$. By Lemma 2 we conclude that $\mathbf{M}^e_0(\gamma) = \mathbf{M}^e_0(\gamma^{\chi}_{\mathbf{L}_\chi}) = 2$, so $\mathbf{M}^e$ is a model of γ. Since γ is a formula in $\mathcal{L}_V$, it follows that $\mathbf{M}^e \cap V \models \gamma$. ⊠

Corollary 2. *Given a past-THT formula γ on $\mathcal{L}_V$, every past operator in γ can be removed by introducing auxiliary atoms.*

5 Conclusions

In this paper we have presented an extension of Temporal Equilibrium Logic (TEL) that introduces the use of past modalities. We have defined the syntax and semantics of this extension and provided a translation that, by introducing auxiliary atoms, allows removing past modalities and using the original version of TEL exclusively dealing with future operators.

The immediate future work will be focused on the implementation of these operators in the tool `STeLP` and the study on complexity of the current translation. We also plan to study the potential application as a high-level language for incremental ASP.

Acknowledgements This work has been partially supported by the Spanish project TIN2013-42149-P. We are especially thankful to Luis Fariñas for his support and guidance not only in our work on temporal logic programming, but also in many other research topics related to computational logics. Pedro Cabalar first met Luis at Santiago de Compostela in 1997 attending to Luis' introductory course on modal logics. That course eventually become the seed of our group's interest in the topic. Luis combines a deep knowledge and large experience on a wide spectrum of topics in non-classical logics (his literature covers specialized results in modal logics, logic programming, epistemic logics, spatial reasoning, temporal logics, intuitionistic and intermediate logics, conditional logics, and the list goes on!) with a surprising humble and friendly manner that makes him a very sociable person. With his permanently youthful nature, Luis has created a wonderful research group and academic environment in the IRIT at Toulouse, where the members of our group have always been received with open arms, not only at the University but also at his home. This is even more evident in the case of Martín Diéguez who wishes to express his gratitude to Luis and all the members of the LiLAC group for their wonderful reception. Luis, we wish you all the best and we hope we will go on enjoying your company and learning from you for many years!

Bibliography

1. Orgun, M.A., MA, W.: An overview of temporal and modal logic programming. In Gabbay, D., Ohlbach, H., eds.: Temporal Logic. Volume 827 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (1994) 445–479
2. Fariñas del Cerro, L.: MOLOG: A system that extends PROLOG with modal logic. New Generation Computing **4**(1) (1986) 35–50
3. Gabbay, D.: Modal and temporal logic programming. In Galton, A., ed.: Temporal Logics and their Applications. Academic Press (1987) 197–237
4. Abadi, M., Manna, Z.: Temporal logic programming. Journal of Symbolic Computation **8** (1989) 277–295
5. Orgun, M.A., Wadge, W.W.: Theory and practice of temporal logic programming. In del Cerro, L.F., Penttonen, M., eds.: Intensional Logics for Programming. Clarendon Press, Oxford (1992) 21–50
6. Kamp, J.A.: Tense Logic and the Theory of Linear Order. PhD thesis, University of California at Los Angeles (1968)
7. Pnueli, A.: The temporal logic of programs. In: 18th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press (1977) 46–57
8. Baudinet, M.: A simple proof of the completeness of temporal logic programming. In del Cerro, L.F., Penttonen, M., eds.: Intensional Logics for Programming. Clarendon Press (1992) 51–83
9. Morgenstern, L.: Representing and reasoning about time travel narratives: Foundational concepts. In: Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014, Vienna, Austria, July 20-24, 2014. (2014)
10. Gelfond, M., Lifschitz, V.: The stable model semantics for logic programming. In Kowalski, R.A., Bowen, K.A., eds.: Logic Programming: Proc. of the Fifth International Conference and Symposium (Volume 2). MIT Press, Cambridge, MA (1988) 1070–1080
11. Niemelä, I.: Logic programs with stable model semantics as a constraint programming paradigm. Annals of Mathematics and Artificial Intelligence **25** (1999) 241–273
12. Marek, V., Truszczyński, M. In: Stable models and an alternative logic programming paradigm. Springer-Verlag (1999) 169–181
13. Pearce, D.: A new logical characterisation of stable models and answer sets. In: Non monotonic extensions of logic programming. Proc. NMELP'96. (LNAI 1216). Springer-Verlag (1996)
14. Heyting, A.: Die formalen Regeln der intuitionistischen Logik. Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse (1930) 42–56
15. Ferraris, P., Lee, J., Lifschitz, V.: A new perspective on stable models. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07). (2007) 372–379
16. Fischer Servi, G.: On modal logics with an intuitionistic base. Studia Logica **36** (1977) 141–149
17. Fariñas del Cerro, L., Raggio, A.R.: Some results in intuitionistic modal logic. Logique et Analyse **26**(102) (1983) 219–224

18. Cabalar, P., Vega, G.P.: Temporal equilibrium logic: a first approach. In: Proc. of the 11th International Conference on Computer Aided Systems Theory, (EUROCAST'07). LNCS (4739). (2007) 241–248
19. Aguado, F., Cabalar, P., Diéguez, M., Pérez, G., Vidal, C.: Temporal equilibrium logic: a survey. Journal of Applied Non-Classical Logics (2013)
20. Cabalar, P.: Stable models for temporal theories - invited talk. In: Proc. of the 13th Intel. Conf. on Logic Programming and Non-Monotonic Reasoning (LPNMR'15). Volume 9345 of Lecture Notes in Artificial Intelligence. Springer (2015) 1–13
21. Cabalar, P., Diéguez, M.: Strong equivalence of non-monotonic temporal theories. In: Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14), Vienna, Austria (2014)
22. Aguado, F., Cabalar, P., Pérez, G., Vidal, C.: Loop formulas for splitable temporal logic programs (2011) (submitted to LPNMR'11)
`http://www.dc.fi.udc.es/~cabalar/lfstlp.pdf`.
23. Cabalar, P., Diéguez, M.: `STeLP` – a tool for temporal answer set programming (2011) (System description, submitted to LPNMR'11)
`http://www.dc.fi.udc.es/~cabalar/stelp.pdf`.
24. Gebser, M., Sabuncu, O., Schaub, T.: An incremental answer set programming based system for finite model computation. AI Commun. **24**(2) (2011) 195–212
25. Gebser, M., Grote, T., Kaminski, R., Obermeier, P., Sabuncu, O., Schaub, T.: Stream reasoning with answer set programming: Preliminary report. In: Principles of Knowledge Representation and Reasoning: Proceedings of the Thirteenth International Conference, KR 2012, Rome, Italy, June 10-14, 2012, AAAI Press (2012)
26. Gabbay, D., Pnueli, A., Shelah, S., Stavi, J.: On the temporal analysis of fairness. In: Proceedings of the 7th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, Las Vegas, Nevada (1980) 163–173
27. Gabbay, D.: The declarative past and imperative future: Executable temporal logic for interactive systems. In: Proc. of the 1st Conf. on Temporal Logic in Specification (TLS'1987), Altrincham, UK. Volume 398 of Lecture Notes in Computer Science. Springer-Verlag (1987) 409–448
28. Laroussinie, F., Markey, N., Schnoebelen, P.: Temporal logic with forgettable past. In: Proc. 17th IEEE Symp. Logic in Computer Science (LICS'2002), Copenhagen, Denmark,, IEEE Computer Society Press (2002) 383–392
29. Markey, N.: Temporal logic with past is exponentially more succinct. EATCS Bulletin **79** (2003) 122–128
30. Pearce, D., Valverde, A.: Quantified equilibrium logic and foundations for answer set programs. In: Proc. of the 24th Intl. Conf. on Logic Programming, ICLP 2008, (Udine, Italy, December 9-13). Volume 5366 of Lecture Notes in Computer Science. Springer (2008) 546–560

On How Kelsenian Jurisprudence and Intuitionistic Logic help to avoid Contrary-to-Duty Paradoxes in Legal Ontologies

Edward Hermann Haeusler[1] and Alexandre Rademaker[2]

[1] Department of Informatics, PUC-Rio
[2] IBM Research, Rio de Janeiro, Brazil

Abstract. In this article we show how Hans Kelsen jurisprudence and Intuitionistic logic are used to avoid the well-known contrary-to-duty (CTD) paradoxes, such as Chisholm paradoxes and its variants. This article uses an intuitionistic version of the ALC description logic, named iALC, to show how an ontology based on individually valid legal statements is able to avoid CTDs by providing models to them.

1 Introduction

Prof. Luis Fariñas del Cerro wrote a bunch of articles reporting the results he obtained by designing logics for very interesting and specific purposes. The elegance of the underlying ideas and the presentation form is out of discussion. Many researchers in logic would like to have his ability to extract logic from facts and their relationships, building new judgments. One of the authors of this article ever tried to be able to have this resulting research. When we were invited to contribute to prof. Luis Fariñas del Cerro Festschrifft this become the opportunity to report to the master maybe the only research that we conduct on the lines of defining a logic for a specific formalization. In this article, we report our results in the last seven years in the designing logics for legal ontologies.

Classical Logic has been widely used as a basis for ontology creation and reasoning in many domains. These domains naturally include Legal Knowledge and Jurisprudence. As we expect, consistency is an important issue for legal ontologies. However, due to their inherently normative feature, coherence (consistency) in legal ontologies is more subtle than in other domains. Consistency, or absence of logical contradictions, seems more difficult to maintain when more than one law system can judge a case, what we call a conflict of laws. There are some legal mechanisms to solve these conflicts such as stating privileged fori or other ruling jurisdiction. In most of the cases, the conflict is solved by admitting a law hierarchy or a law precedence, rather better, ordering on laws. Under these precedence mechanisms, coherence is still a major issue in legal systems. Each layer in this legal hierarchy has to be consistent. Since consistency is a direct consequence of how one deals with the logical negation, negation is also a main concern in legal systems. Deontic Logic, here considered as an extension of Classical Logic, has been widely used to formalize the normative aspects of the legal knowledge. There is some disagreement on using deontic logic, and any of its variants, to this task. Since a seminal paper by Alchourron & Martino [1], the propositional aspect of laws has been under discussion. In [1], in full agreement with Hans Kelsen jurisprudence, they argue that *laws* are not to be considered as propositions. The Kelsenian approach to Legal Ontologies considers the term "ontologies on laws" more appropriate than "law ontology". In previous works, we showed that Classical logic is not adequate to cope with a Kelsenian based Legal Ontology. Because of the ubiquitous use of Description Logic for expressing ontologies nowadays, we developed an Intuitionistic version of Description Logic particularly devised to express Legal Ontologies. This logic is called iALC. In this article, we show how iALC avoids some Contrary-to-duty paradoxes, as *Chisholm* paradox and other paradoxes that appear in deontic logic, such as *the good samaritan* and *the knower*. For these paradoxes, we provide iALC models. Finally, we discuss the main role of the intuitionistic negation in this issue, finding out that its success may be a consequence of its paracomplete logical aspect. This investigation opens the use of other paracomplete logics in accomplishing a logical basis for Kelsenian legal ontologies, as a complementary solution to those based on paraconsistent logics, see [20].

2 A brief discussion on Kelsenian Jurisprudence and its logic

A very important task in jurisprudence (legal theory) is to make precise the use of the term "law", the *individuation problem*, and it is one of the most fundamental open questions in jurisprudence.

It requires firstly answering the question “What is to count as one complete law?” ([26]). There are two main approaches to answer this question. One approach is to consider “the law” as the result of a natural process that yields a set of norms responsible for stating perfect social behavior. Another approach is to consider “the law” as a set of individual legal statements, each of them created to enforce a positively desired behavior in the society. As a consequence, in the first approach, the norms say what are the best morally speaking accepted state of affairs in a particular society, while, in the second approach, each legal statement rules an aspect of the society that the legislature wants to enforce the behavior. The first is more related to what is called *Natural Law* and the last to *Legal positivism*. We can say that the *Legal positivism* is closer to the way modeling is taken in Computer Science. In the *natural* approach to the law, it is even harder to define a system of laws than the *legal positivism*. The natural approach demands stronger knowledge of the interdependency between the underlying legal statements than legal positivism. Because of that, the *natural* approach, in essence, is harder to be shared with practical jurisprudence principles, since they firstly are concerned to justify the law, on an essentially moral basis. This justification is quite hard to maintain from a practical point of view.

The coherence of “the law” in both approaches is essential. A debate on whether coherence is built-in by the restrictions induced by Nature in an evolutionary way, or whether coherence should be an object of knowledge management, seems to be a long debate. Despite that, legal positivism seems to be more suitable to Legal Artificial Intelligence. From the logical point of view, the *natural* approach is harder to deal with than the *positivist* one. When describing a morally desired state-of-affairs, the logical statements take the form of propositions that has as a model best of the moral worlds. Deontic logic is suitable to be used to fulfill this task. However, a legal statement (“a law”) is essentially an individual sentence that can also be seen as an order (mandatory command), and hence, it is not a proposition at all. As a consequence, deontic logic is not appropriate to be used in knowledge bases. Besides that, [30] shows that deontic logic does not properly distinguish between the normative status of a situation from the normative status of a norm (rule). We think that the best jurisprudence basis for Legal ontologies and reasoning is *Legal positivism*. Thus, we will be talking a legal ontology as an ontology about (individual) laws, and not an ontology on “the law”.

Hans Kelsen initialized the *Legal positivism* tradition in 1934 (for a contemporary reference see [18]). He used this positive aspect of the legislature to define a theory of *pure law* and applied it to the problem of transfer citizen’s rights and obligations from one country to other when crossing boarders. He produces a quite good understanding of what nowadays we denominate **Private International Law**. This achievement was so important that in many references on international law, Kelsen jurisprudence is the basis for discussions on conflict-of-laws derived from different statements coming from different *fori*. [3]

In what follows we introduce the main terminology and concepts of Kelsenian jurisprudence that we use in this article. We can summarize Kelsen theory of pure law in three principles:

1. According to what was discussed above, individually valid legal statements are the first-class citizens of our ontology. Thus, only inhabitants of the Legal knowledge base are individual laws, see [17], supra note 5, pp 9-10 [4]. For example, if it is the case that Maria is married with John, and, this was legally celebrated, then “Maria-married-with-John” is an individually valid legal statement, and hence, it is a member of the Legal Ontology;
2. Kelsen also says that that the validity of a legal norm can only be provided concerning the validity of another, and higher, one. So, n_1, a norm, is legally valid if, and only if, it was created or promulgated in agreement with other, and higher, legally valid norm, n_2. This justification induces a precedence relationship between norms that is transitive, that is, if n_1 precedes n_2, and, n_2 precedes n_3, then n_1 recedes n_3; [5]
3. There is a mechanism for relating laws from one Legal system to another, the so-called “choice-of-law rule”. This mechanism is very important to the development of a concept of International Law. Assume that *Mary-is-married-with-John* is an individual legal statement in legal system A. Assume also that Mary is a citizen of a country adopting legal system B. Is there any legal statement in B ensuring that Mary is married in B? Well, this depends on B itself, but there

[3] “It is one of Kelsen’s frequently repeated doctrines that conflict of norms, in the absence of a normative procedure for resolving the conflict, shatters the concept of a unified system”, is highly emphasized in Hughes [14], for example, and it is one of the principles most cited when Kelsen jurisprudence is presented.

[4] Kelsen takes norms and valid norms as synonyms. To say that a legal norm is valid is to say that it exists, is affirmed by Kelsen

[5] See [17], supra note 5, p. 196-7. This can be also found in [16], supra note 5, p. 110-1.

is a way to connected the individual law *Mary-is-married-with-John* in A to *Mary-is-married-with-John* in B. In some legal systems, this is accomplished by what Kelsen denominated “a connection”. As shown in the following quotation from [16], page 247, the connection between the laws of A and B is made by reference, but, in fact, each law belongs to its respective legal system. In this specific case we can consider *Mary-is-married-with-John* in system A is connected to *Mary-is-married-with-John* in legal system B the connection *Lex Loci Celebrationis.*

> ... the law of one State prescribes the application of the law of another State, and the latter does not object or demand it. It has no right to do so since it is not really its own law which is applied by the other State. THe latter applies norms of its own law. The fact that these norms have the same contents as corresponding norms of another State does not concern the latter...Since the specific technique of these norms consists in “referring” to the norms of another system and by so doing incorporating norms of identical contents into their own legal system, it would be more justifiable to call them “reference rules”... The reference rule, that is ... the norm regulating the application of foreign law, may be distinguished from the norm to be applied, that is, the norm referred to. Only the former is a norm of private international law. But from a functional point of view, the one is essentially connected with the other.

Nowadays it is a common terminology in Private International Law the use of the connecting factors or legal connections between individual laws in a different legal system. Only to enumerate some of them: *Lex-Domicilii*, *Lex-Patriae*, *Lex-loci-contratum*, *Lex-loci-solutionis*, and etc.

There is a philosophical problem with the principle 2 above. It demands the existence of basic laws. These basic laws do not have their validity/existence as a consequence of other more basic laws. Kelsen name these basic laws *Grundnorms.* Their validity is based on legislature acts and in a certain sense is derived from the sovereign of the State. It is out of the scope of this article to discuss such problem in Kelsen’s jurisprudence. We take as granted that Kelsen jurisprudence can adequately support most of the existent legal systems, a definitively not an unreal working hypothesis.

From the three principles above, we have some very simple ontological commitments:

commitI Individuals are laws;

commitII There is a transitive and reflexive relationship between individual laws that reflects the natural precedence relationship between laws;

commitIII There are legal connections between individual laws in different legal systems or between different *fori* in the same broader legal system.

From these commitments, we derive the basic constructs of the logic $i\mathcal{ALC}$. In the first place, our legal ontology relates concepts to legal systems. Description logics uses nominals to refer to individuals. So, an expression as $i : A$, stands for i is an individual law, belonging to the legal system A, a concept.

From commitment 2 we consider an expression as $i \preceq j$ standing for the individual law i legally precedes individual law j. The subsumption relationship $A \subseteq B$, from description logic, denotes that A is a legal subsystem of B. One could interpret this relation as the inclusion relationship. [6] We discuss the implications of using negated contents together with Kelsenian jurisprudence in the following. This can be found in [10,11,13,12] too.

Under the classical setting, a negated concept $\neg A$ denotes the “set” of all inhabitants of the domain that do not belong to the interpretation of A. Under ontological commitment 2 there is no individual law that does not exist in, belong to, the domain. Since norms and laws are not propositions, it is a complete nonsense to negate a law. As we already seen, we can negate a concept on laws. Consider the collection of all Brazilian individual laws. Call it BR. In a classical setting $BR \sqcup \neg BR$ is the universe of laws. Thus, any law that it is not in BR has to be a law outside BR, that is, belonging to $\neg BR$. For example, if Peter is 17 years old, it is not liable according to the Brazilian law. Is $Peter-is-liable$ a valid law at all? If so, it has to belong to $\neg BR$. Using Kelsen in a classical setting, individual laws not belonging to a concept automatically belong to its complementary concept. The problem with this is that it is possible to create laws outside a jurisdiction or forum by the very simple act of considering or experimenting a legal situation. Nowadays in Brazil, the parliament is discussing the liability under the 16 years. By the simple fact of discussing the validity of their corresponding individual laws, we are forced to accept they exist outside the Brazilian legal system. We do not consider this feature appropriate to legal ontology definition. Dealing with negations every

[6] In Classical ALC this is just the case, but we shown here that classical reasoning it is not a good choice for dealing with legal ontologies

time we assume the existence of a law may bring unnecessary complexity to legal ontology definition. Because the precedence relationship between laws, cf. ontological commitment 2, there is a natural alternative to classical logic, the intuitionistic logic (IL). According to IL semantics, $i : \neg A$, iff, for each law j, such that $i \preceq j$, it is not the case that $j : A$. This semantics means that i does not provide any legal support for any individual law belong to A, which agrees with Kelsen jurisprudence on the hierarchy of individual laws.

Commitment 2 gives rise to expressions of the form m Lex-Loci-Celebrationis m, where m is *Mary-is-married-with-John* and Lex-Loci-Celebrationis is a legal connection. Thus, if *Abroad* is the concept that represents all laws in Portugal, then the concept $\exists\ Lex - Loci - Celebrationis\ Portugal$ represents the Brazilian individual laws stating that Portuguese marriage is valid in Brazil. The private international law of any country is a collection of laws stated in similar ways for every possible legal connection. In [11] it is shown in detail a judicial case deriving that a renting contract is solving a conflict of laws in space through private international law.

3 Some philosophical discussion on the ontological criteria taken on using Kelsen in legal ontologies

We base our work on two ontological criteria: [7] 1- Ontological Commitment (due to W.Quine), our logical approach is ontologically committed to Valid Legal Statements only. The only nominals occurring in our logic language are valid individual laws, and; 2-Ontological Parsimony, which is strongly related to Quine's ontological commitment too, with a mention of its stronger version also known as Occam's Razor, here denoted as OR. The second criteria is based on: "One 'easy' case where OR can be straightforwardly applied is when a theory T, postulates entities which are explanatorily idle. Excising these entities from T produces a second theory, T*, which has the same theoretical virtues as T but a smaller set of ontological commitments. Hence, according to OR, it is rational to pick T* over T."

We observe that nominals, representing individuals, denote only valid individual laws and nothing in the $i\mathcal{ALC}$ language described in the following section, is committed with non-valid individual laws, according to the second ontological criterion above, we do not have to consider non-valid individual laws. Technically speaking there is no element in the $i\mathcal{ALC}$ language able to denote an invalid individual law in any model of any $i\mathcal{ALC}$ theory. If something is a valid individual law regarded some legal system in some place in the world, then this individual belongs to our semantic universe.

This philosophical basis allows us to have only sets of valid individuals as semantics for $i\mathcal{ALC}$ theories. Thus, as the a reviewer have already observed, this implies that $\neg A$ is the set of individual laws holding outside Brazil, and the classical negation is not adequate to denote this set. If we get $\neg A$ meaning "individual laws that *do not* hold in Brazil", the set of laws being a proper subset of the universe, and A is the conjunctive property "laws + holds in A". Then the complement, $\neg A$ would be all elements of the universe which are either not a valid individual law or do not hold in Brazil. But there is no way to take the semantics in this way, for the semantics we get from our ontological commitment 2 is given by "The individual valid laws holding outside of Brazil".

Finally, concerning contradictory individual laws, they can coexist in the same universe, since they are there because they hold in distinct legal systems. In fact they are apparently contradictory. For example, "There is death penalty" and "Death sentence is not allowed" can coexist, since there are countries where each of these legal statement are valid. Concretely: "There is death penalty":Iran and "Death sentence is not allowed":Brazil.

4 The Logic $i\mathcal{ALC}$

Classical Description Logic has been widely used as a basis for ontology creation and reasoning in many knowledge specific domains, including Legal AI.

An adequate intuitionistic semantics for negation in a legal domain comes to the fore when we take legally valid individual statements as the inhabitants of our legal ontology. This allows us to elegantly deal with particular situations of legal coherence, such as conflict of laws, as those solved by Private International Law analysis. In [12,11,13] we present an Intuitionistic Description Logic, called $i\mathcal{ALC}$ for Intuitionistic $\mathcal{ALC}$ (for Attributive Language with Complements, the canonical classical description logic system). A labeled sequent calculus for $i\mathcal{ALC}$ based on a labeled sequent calculus for

[7] see Quine's "On What there is" article and `http://plato.stanford.edu/entries/simplicity`, for example to a primer ontological criteria

$\mathcal{ALC}$ [24], was also presented. In these previous articles, we discussed the jurisprudence foundation of our system, and show how we can perform a coherence analysis of "Conflict of Laws in Space" by means of $i\mathcal{ALC}$. This conflict happens when several laws can be applied, with different outcomes, to a case depending on the place where the case occurs. Typical examples are those ruling the rights of a citizen abroad.

In [11], the semantics of $i\mathcal{ALC}$ presented followed the framework for constructive modal logics presented by Simpson [28] and adapted to description languages by Paiva [6]. We applied this logic to the problem of formalizing legal knowledge.

Description Logics are an important knowledge representation formalism, unifying and giving a logical basis to the well known AI frame-based systems of the eighties. Description logics are very popular right now. Given the existent and proposed applications of the Semantic Web, there has been a fair amount of work into finding the most well-behaved system of description logic that has the broadest application, for any specific domain. Description logics tend to come in families of logical systems, depending on which concept constructors you allow in the logic. Since description logics came into existence as fragments of first-order logic chosen to find the best trade-off possible between expressiveness and tractability of the fragment, several systems were discussed and in the taxonomy of systems that emerged the $\mathcal{ALC}$ has come to be known as the canonical one. The basic building blocks of description logics are *concepts*, *roles* and *individuals*. Think of concepts as unary predicates in usual first-order logic and of roles as binary predicates, used to modify the concepts.

As discussed in [6], considering versions of *constructive* description logics makes sense, both from a theoretical and from a practical viewpoint. There are several possible and sensible ways of defining *constructive* description logics, whether your motivation is natural language semantics (as in [6]) or Legal AI (as in [12]). As far as *constructive* description logics are concerned, Mendler and Scheele have worked out a very compelling system $c\mathcal{ALC}$ [21], based on the constructive modal logic CK [2]), a favorite[8] system of ours. However in this note we follow a different path and describe a constructive version of $\mathcal{ALC}$, based on the framework for constructive modal logics developed by Simpson (the system IK) in his phd thesis [28] (For a proof-theoretic comparison between the constructive modal logics CK and IK one can see [25]).

Our motivation, besides Simpson's work, is the framework developed by Braüner and de Paiva in [3] for constructive Hybrid Logics. We reason that having already frameworks for constructive modal and constructive hybrid logics in the labelled style of Simpson, we might end up with the best style of constructive description logics, in terms of both solid foundations and ease of implementation. Since submitting this paper we have been told about the master thesis of Clément [5] which follows broadly similar lines. Clément proves soundness and completeness of the system called $i\mathcal{ALC}$ and then provides a focused version of this system, a very interesting development, as focused systems are, apparently, very useful for proof search.

Our Sequent Calculus for $i\mathcal{ALC}$ was first presented in [7] where we briefly described the immediate properties of this system and most importantly we discuss a case study of the use of $i\mathcal{ALC}$ in legal AI.

This article corrects and extends the presentation of $i\mathcal{ALC}$ appearing in all previous articles. It points out the difference between $i\mathcal{ALC}$ and the intuitionistic hybrid logic presented in [6]. Completeness and soundness proofs are revised. A discussion on the computational complexity of $i\mathcal{ALC}$ is also taken.

5 Intuitionistic ALC

The $i\mathcal{ALC}$ logic is based on the framework for intuitionistic modal logic IK proposed in [28,8,23]. These modal logics arise from interpreting the usual possible worlds definitions in an intuitionistic meta-theory. As we will see in the following paragraphs, ideas from [3] were also used, where the framework IHL, for *intuitionistic hybrid logics*, is introduced. $i\mathcal{ALC}$ concepts are described as:

$$C, D ::= A \mid \bot \mid \top \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid C \sqsubseteq D \mid \exists R.C \mid \forall R.C$$

where C, D stands for concepts, A for an atomic concept, R for an atomic role. We could have used distinct symbols for subsumption of concepts and the subsumption concept constructor but this would blow-up the calculus presentation. This syntax is more general than standard $\mathcal{ALC}$ since it includes subsumption $\sqsubseteq$ as a concept-forming operator. We have no use for nested subsumptions, but they do make the system easier to define, so we keep the general rules. Negation could be defined

[8] This system has categorical semantics, which are not very easy to obtain for modal logics.

via subsumption, that is, $\neg C = C \sqsubseteq \bot$, but we find it convenient to keep it in the language. The constant $\top$ could also be omitted since it can be represented as $\neg\bot$.

A constructive interpretation of $i\mathcal{ALC}$ is a structure $\mathcal{I}$ consisting of a non-empty set $\Delta^{\mathcal{I}}$ of entities in which each entity represents a partially defined individual; a refinement pre-ordering $\preceq^{\mathcal{I}}$ on $\Delta^{\mathcal{I}}$, i.e., a reflexive and transitive relation; and an interpretation function $\cdot^{\mathcal{I}}$ mapping each role name R to a binary relation $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ and atomic concept A to a set $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ which is closed under refinement, i.e., $x \in A^{\mathcal{I}}$ and $x \preceq^{\mathcal{I}} y$ implies $y \in A^{\mathcal{I}}$. The interpretation $\mathcal{I}$ is lifted from atomic concepts to arbitrary concepts via:

$$
\begin{aligned}
\top^{\mathcal{I}} &=_{df} \Delta^{\mathcal{I}} \\
\bot^{\mathcal{I}} &=_{df} \emptyset \\
(\neg C)^{\mathcal{I}} &=_{df} \{x \mid \forall y \in \Delta^{\mathcal{I}}.x \preceq y \Rightarrow y \notin C^{\mathcal{I}}\} \\
(C \sqcap D)^{\mathcal{I}} &=_{df} C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
(C \sqcup D)^{\mathcal{I}} &=_{df} C^{\mathcal{I}} \cup D^{\mathcal{I}} \\
(C \sqsubseteq D)^{\mathcal{I}} &=_{df} \{x \mid \forall y \in \Delta^{\mathcal{I}}.(x \preceq y \text{ and } y \in C^{\mathcal{I}}) \Rightarrow y \in D^{\mathcal{I}}\} \\
(\exists R.C)^{\mathcal{I}} &=_{df} \{x \mid \exists y \in \Delta^{\mathcal{I}}.(x, y) \in R^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\} \\
(\forall R.C)^{\mathcal{I}} &=_{df} \{x \mid \forall y \in \Delta^{\mathcal{I}}.x \preceq y \Rightarrow \forall z \in \Delta^{\mathcal{I}}.(y, z) \in R^{\mathcal{I}} \Rightarrow z \in C^{\mathcal{I}}\}
\end{aligned}
$$

Following the semantics of IK, the structures $\mathcal{I}$ are models for $i\mathcal{ALC}$ if they satisfy two frame conditions:

F1 if $w \leq w'$ and wRv then $\exists v'.w'Rv'$ and $v \leq v'$
F2 if $v \leq v'$ and wRv then $\exists w'.w'Rv'$ and $w \leq w'$

The above conditions are diagrammatically expressed as:

$$
\begin{tikzcd}
w' \arrow[r, dotted, "R"] & v' \\
w \arrow[u, "\leq"] \arrow[r, "R"'] \arrow[ur, phantom, "(F1)"] & v \arrow[u, dotted, "\leq"']
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
w' \arrow[r, dotted, "R"] & v' \\
w \arrow[u, dotted, "\leq"] \arrow[r, "R"'] \arrow[ur, phantom, "(F2)"] & v \arrow[u, "\leq"']
\end{tikzcd}
$$

Our setting simplifies [21], since $i\mathcal{ALC}$ satisfies (like classical $\mathcal{ALC}$) $\exists R.\bot = \bot$ and $\exists R.(C \sqcup D) = \exists R.C \sqcup \exists R.D$.

Building up from the Simpson's constructive modal logics (called here IML), in [3], it is introduced *intuitionistic hybrid logics*, denoted by IHL. Hybrid logics add to usual modal logics a new kind of propositional symbols, the *nominals*, and also the so-called *satisfaction operators*. A nominal is assumed to be true at exactly one world, so a nominal can be considered the name of a world. If x is a nominal and X is an arbitrary formula, then a new formula $x : X$ called a satisfaction statement can be formed. The satisfaction statement $x : X$ expresses that the formula X is true at one particular world, namely the world denoted by x. In hindsight one can see that IML shares with hybrid formalisms the idea of making the possible-world semantics part of the deductive system. While IML makes the relationship between worlds (e.g., xRy) part of the deductive system, IHL goes one step further and sees the worlds themselves x, y as part of the deductive system, (as they are now nominals) and the satisfaction relation itself as part of the deductive system, as it is now a syntactic operator, with modality-like properties. In contrast with the above mentioned approaches, ours assign a truth values to some formulas, also called assertions, they are not concepts as in [3], for example. Below we define the syntax of general assertions (A) and nominal assertions (N) for ABOX reasoning in $i\mathcal{ALC}$. Formulas (F) also includes subsumption of concepts interpreted as propositional statements.

$$
N ::= x \colon C \mid x \colon N \qquad\qquad A ::= N \mid xRy \qquad\qquad F ::= A \mid C \sqsubseteq C
$$

where x and y are nominals, R is a role symbol and C is a concept. In particular, this allows $x \colon (y \colon C)$, which is a perfectly valid nominal assertion.

Definition 1 (outer nominal). *In a nominal assertion $x \colon \gamma$, x is said to be the outer nominal of this assertion. That is, in an assertion of the form $x \colon (y \colon \gamma)$, x is the outer nominal.*

We write $\mathcal{I}, w \models C$ to abbreviate $w \in C^{\mathcal{I}}$ which means that entity w satisfies concept C in the interpretation $\mathcal{I}$[9]. Further, $\mathcal{I}$ is a model of C, written $\mathcal{I} \models C$ iff $\forall w \in \mathcal{I}.\mathcal{I}, w \models C$. Finally, $\models C$ means

[9] In IHL, this w is a world and this satisfaction relation is possible world semantics

$\forall \mathcal{I}.\mathcal{I} \models C$. All previous notions are extended to sets Φ of concepts in the usual universal fashion. Given the hybrid satisfaction statements, the interpretation and semantic satisfaction relation are extended in the expected way. The statement $\mathcal{I}, w \models x\colon C$ holds, if and only if, $\forall z_x \succeq^{\mathcal{I}} x \,.\, \mathcal{I}, z_x \models C$. In a similar fashion, $\mathcal{I}, w \models xRy$ holds ,if and only if, $\forall z_x \succeq x.\forall z_y \succeq y.(x_x^{\mathcal{I}}, z_y^{\mathcal{I}}) \in R^{\mathcal{I}}$. That is, the evaluation of the hybrid formulas does not take into account only the world w, but it has to be monotonically preserved. It can be observed that for every w', if $x^{\mathcal{I}} \preceq w'$ and $\mathcal{I}, x' \models \alpha$, then $\mathcal{I}, w' \models \alpha$ is a property holding on this satisfaction relation.

In common reasoning tasks the interpretation $\mathcal{I}$ and the entity w in a verification goal such as $\mathcal{I}, w \models \delta$ are not given directly but are themselves axiomatized by sets of concepts and formulas. Usually we have a set Θ [10] of formulas and the set Γ of concepts. Accordingly:

Definition 2. *We write $\Theta, \Gamma \models \delta$ if it is the case that:*

$$\forall \mathcal{I}.((\forall x \in \Delta^{\mathcal{I}}.(\mathcal{I}, x \models \Theta)) \Rightarrow \forall (Nom(\Gamma, \delta)).\forall \boldsymbol{z} \succeq Nom(\Gamma, \delta).(\mathcal{I}, \boldsymbol{z} \models \Gamma \Rightarrow \mathcal{I}, \boldsymbol{z} \models \delta) \quad (1.1)$$

where $\boldsymbol{z}$ denotes a vector of variables $z_1, \ldots, z_k$ and $Nom(\Gamma, \delta)$ is the vector of all outer nominals occurring in each nominal assertion of $\Gamma \cup \{\delta\}$. x is the only outer nominal of a nominal assertion $\{x\colon \gamma\}$, while a (pure) concept γ has no outer nominal.

A Hilbert calculus for $i\mathcal{ALC}$ is provided following [23,28,8]. It consists of all axioms of intuitionistic propositional logic plus the axioms and rules displayed in Figure 1.1. The Hilbert calculus implements TBox-reasoning. That is, it decides the semantical relationship $\Theta, \emptyset \models C$. Θ has only formulas as members.

0.	all substitution instances of theorems of IPL
1.	$\forall R.(C \sqsubseteq D) \sqsubseteq (\forall R.C \sqsubseteq \forall R.D)$
2.	$\exists R.(C \sqsubseteq D) \sqsubseteq (\exists R.C \sqsubseteq \exists R.D)$
3.	$\exists R.(C \sqcup D) \sqsubseteq (\exists R.C \sqcup \exists R.D)$
4.	$\exists R.\bot \sqsubseteq \bot$
5.	$(\exists R.C \sqsubseteq \forall R.C) \sqsubseteq \forall R.(C \sqsubseteq D)$
MP	If C and $C \sqsubseteq D$ are theorems, D is a theorem too.
Nec	If C is a theorem then $\forall R.C$ is a theorem too.

Fig. 1.1. The $i\mathcal{ALC}$ axiomatization

A Sequent Calculus for $i\mathcal{ALC}$ is also provided. The logical rules of the Sequent Calculus for $i\mathcal{ALC}$ are presented in Figure 1.2. [11] The structural rules and the cut rule are omitted but they are as usual. The δ stands for concepts or assertions ($x\colon C$ or xRy), α and β for concept and R for role. Δ is a set of formulas. In rules p-$\exists$ and p-$\forall$, the syntax $\forall R.\Delta$ means $\{\forall R.\alpha \mid \alpha \in concepts(\Delta)\}$, that is, all concepts in Δ are universal quantified with the same role. The assertions in Δ are kept unmodified. In the same way, in rule p-N the addition of the nominal is made only in the concepts of Δ (and in δ if that is a concept) keeping the assertions unmodified.

The propositional connectives ($\sqcap, \sqcup, \sqsubseteq$) rules are as usual, the rule $\sqcup_2$-r is omitted. The rules are presented without nominals but for each of these rules there is a counterpart with nominals. For example, the rule $\sqsubseteq$-r has one similar:

$$\frac{\Delta, x\colon \alpha \Rightarrow\ x\colon \beta}{\Delta \Rightarrow\ x\colon (\alpha \sqsubseteq \beta)}\ \text{n-}\sqsubseteq\text{-r}$$

The main modification comes for the modal rules, which are now role quantification rules. We must keep the intuitionistic constraints for modal operators. Rule $\exists$-l has the usual condition that y is not in the conclusion. Concerning the usual condition on the $\forall$-r rule, it is not the case in this system, for the interpretation of the a nominal assertion in a sequent is already implicitly universal (Definition 2).

[10] Here we consider only acycled TBox with $\sqsubseteq$ and $\equiv$.

[11] The reader may want to read Proof Theory books, for example, [29,4,22,9].

$$\frac{}{\Delta, \delta \Rightarrow \delta} \qquad \frac{}{\Delta, x\colon \bot \Rightarrow \delta}$$

$$\frac{\Delta, xRy \Rightarrow y\colon \alpha}{\Delta \Rightarrow x\colon \forall R.\alpha}\ \forall\text{-r} \qquad \frac{\Delta, x\colon \forall R.\alpha, y\colon \alpha, xRy \Rightarrow \delta}{\Delta, x\colon \forall R.\alpha, xRy \Rightarrow \delta}\ \forall\text{-l}$$

$$\frac{\Delta \Rightarrow xRy \qquad \Delta \Rightarrow y\colon \alpha}{\Delta \Rightarrow x\colon \exists R.\alpha}\ \exists\text{-r} \qquad \frac{\Delta, xRy, y\colon \alpha \Rightarrow \delta}{\Delta, x\colon \exists R.\alpha \Rightarrow \delta}\ \exists\text{-l}$$

$$\frac{\Delta, \alpha \Rightarrow \beta}{\Delta \Rightarrow \alpha \sqsubseteq \beta}\ \sqsubseteq\text{-r} \qquad \frac{\Delta_1 \Rightarrow \alpha \qquad \Delta_2, \beta \Rightarrow \delta}{\Delta_1, \Delta_2, \alpha \sqsubseteq \beta \Rightarrow \delta}\ \sqsubseteq\text{-l}$$

$$\frac{\Delta \Rightarrow \alpha \qquad \Delta \Rightarrow \beta}{\Delta \Rightarrow \alpha \sqcap \beta}\ \sqcap\text{-r} \qquad \frac{\Delta, \alpha, \beta \Rightarrow \delta}{\Delta, \alpha \sqcap \beta \Rightarrow \delta}\ \sqcap\text{-l}$$

$$\frac{\Delta \Rightarrow \alpha}{\Delta \Rightarrow \alpha \sqcup \beta}\ \sqcup_1\text{-r} \qquad \frac{\Delta, \alpha \Rightarrow \delta \qquad \Delta, \beta \Rightarrow \delta}{\Delta, \alpha \sqcup \beta \Rightarrow \delta}\ \sqcup\text{-l}$$

$$\frac{\Delta, \alpha \Rightarrow \beta}{\forall R.\Delta, \exists R.\alpha \Rightarrow \exists R.\beta}\ \text{p-}\exists \qquad \frac{\Delta \Rightarrow \alpha}{\forall R.\Delta \Rightarrow \forall R.\alpha}\ \text{p-}\forall$$

$$\frac{\Delta \Rightarrow \delta}{x\colon \Delta \Rightarrow x\colon \delta}\ \text{p-N}$$

Fig. 1.2. The System SC_{iALC}: logical rules

Theorem 1. *The sequent calculus described in Fig. 1.2 is sound and complete for TBox reasoning, that is $\Theta, \emptyset \models C$ if and only if $\Theta \Rightarrow C$ is derivable with the rules of Figure 1.2.*

The completeness of our system is proved relative to the axiomatization of $i\mathcal{ALC}$, shown in Figure 1.1. The proof is presented in Section 6.

The soundness of the system is proved directly from the semantics of $i\mathcal{ALC}$ including the ABOX, that is, including nominals. The semantics of a sequent is defined by the satisfaction relation, as shown in Definition 2. The sequent $\Theta, \Gamma \Rightarrow \delta$ is valid if and only if $\Theta, \Gamma \models \gamma$. Soundness is proved by showing that each sequent rule preserves the validity of the sequent and that the initial sequent is valid. This proof is presented in Section 7.

We note that although we have here fixed some inaccuracies in the presentation of the $i\mathcal{ALC}$ semantics in [7], the system presented here is basically the same, excepted that here the propositional rules are presented without nominals. Given that, the soundness of the system proved in [7] can be still considered valid without further problems. Note also that the proof of soundness provides in Section 7 is regarded the full language of $i\mathcal{ALC}$. It considers nominals and assertion on nominals relationship, that is it concerns ABOX and TBOX. The proof of completeness is for the TBOX only. A proof of completeness for ABOX can be done by the method of canonical models. For the purposes of this article, we choose to show the relative completeness proof with the sake of showing a simpler proof concerning TBOX.

6 The completeness of SC_{iALC} system

We show the relative completeness of SC_{iALC} regarding the axiomatic presentation of $i\mathcal{ALC}$ presented in Figure 1.1. To prove the completeness of SC_{iALC} it is sufficient to derive in SC_{iALC} the axioms 1–5 of $i\mathcal{ALC}$. It is clear that all substitution instances of IPL theorems can also be proved in SC_{iALC} using only propositional rules. The MP rule is a derived rule from the SC_{iALC} using the cut rule. The Nec rule is the p-$\forall$ rule in the system with Δ empty. In the first two proofs below do not use nominals for given better intuition of the reader about the use of rules with and without nominals.

Axiom 1:

$$\dfrac{\dfrac{\dfrac{\alpha \Rightarrow \alpha \qquad \beta \Rightarrow \beta}{\alpha \sqsubseteq \beta, \alpha \Rightarrow \beta}\ \sqsubseteq\text{-l}}{\forall R.(\alpha \sqsubseteq \beta), \exists R.\alpha \Rightarrow \exists R.\beta}\ \text{p-}\exists}{\forall R.(\alpha \sqsubseteq \beta) \Rightarrow \exists R.\alpha \sqsubseteq \exists R.\beta}\ \sqsubseteq\text{-r}$$

Axiom 2:

$$\dfrac{\dfrac{\dfrac{\alpha \Rightarrow \alpha \qquad \beta \Rightarrow \beta}{\alpha \sqsubseteq \beta, \alpha \Rightarrow \beta}\ \sqsubseteq\text{-l}}{\forall R.(\alpha \sqsubseteq \beta), \forall R.\alpha \Rightarrow \forall R.\beta}\ \text{p-}\forall}{\forall R.(\alpha \sqsubseteq \beta) \Rightarrow \forall R.\alpha \sqsubseteq \forall R.\beta}\ \sqsubseteq\text{-r}$$

Axiom 3:

$$\dfrac{\dfrac{xRy, y : \bot \Rightarrow x : \bot}{x : \exists R.\bot \Rightarrow x : \bot}\ \exists\text{-l}}{\Rightarrow x : (\exists R.\bot \sqsubseteq \bot)}\ \sqsubseteq\text{-r}$$

Axiom 4:

$$\dfrac{\dfrac{x : \exists R.\alpha \Rightarrow x : \exists R.\alpha}{x : \exists R.\alpha \Rightarrow x : (\exists R.\alpha \sqcup \exists R.\beta)}\ \sqcup_1\text{-r} \qquad \dfrac{x : \exists R.\beta \Rightarrow x : \exists R.\beta}{x : \exists R.\beta \Rightarrow x : (\exists R.\alpha \sqcup \exists R.\beta)}\ \sqcup_2\text{-r}}{x : \exists R.(\alpha \sqcup \beta) \Rightarrow x : (\exists R.\alpha \sqcup \exists R.\beta)}\ \sqcup\text{-l}$$

Axiom 5:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{xRy, y : \alpha \Rightarrow y : \alpha \qquad xRy, y : \alpha \Rightarrow xRy}{xRy, y : \alpha \Rightarrow x : \exists R.\alpha}\ \exists\text{-r} \qquad \dfrac{xRy, y : \alpha, y : \beta, \forall R.\beta \Rightarrow y : \beta}{xRy, y : \alpha, x : \forall R.\beta \Rightarrow y : \beta}\ \forall\text{-l}}{x : (\exists R.\alpha \sqsubseteq \forall R.\beta), xRy, y : \alpha \Rightarrow y : \beta}\ \sqsubseteq\text{-l}}{x : (\exists R.\alpha \sqsubseteq \forall R.\beta), xRy \Rightarrow y : (\alpha \sqsubseteq \beta)}\ \forall\text{-r}}{x : (\exists R.\alpha \sqsubseteq \forall R.\beta) \Rightarrow x : \forall R.(\alpha \sqsubseteq \beta)}\ \forall\text{-r}}{\Rightarrow x : [(\exists R.\alpha \sqsubseteq \forall R.\beta) \sqsubseteq \forall R.(\alpha \sqsubseteq \beta)]}\ \sqsubseteq\text{-r}$$

7 Soundness of SC_{iALC} system

In this section we prove that.

Proposition 1. *If $\Theta, \Gamma \Rightarrow \delta$ is provable in SC_{iALC} then $\Theta, \Gamma \models \gamma$.*

Proof: We prove that each sequent rule preserves the validity of the sequent and that the initial sequents are valid. The definition of a valid sequent ($\Theta, \Gamma \models \gamma$) is presented in Definition 2.

The validity of the axioms is trivial. We first observe that any application of the rules $\sqsubseteq$-r, $\sqsubseteq$-l,$\sqcap$-r,$\sqcap$-l, $\sqcup_1$-r,$\sqcup_2$-r, $\sqcup$-l of SC_{iALC} where the sequents do not have any nominal, neither in Θ nor in Γ, is sound regarded intuitionistic propositional logic kripke semantics, to which the validity definition above collapses whenever there is no nominal in the sequents. Thus, in this proof we concentrate in the case where there are nominals. We first observe that the nominal version of $\sqsubseteq$-r, the validity of the premises includes

$$\forall (Nom(\Gamma, \delta)).\forall \boldsymbol{z} \succeq Nom(\Gamma, \delta).(\mathcal{I}, \boldsymbol{z} \models \Gamma \Rightarrow \mathcal{I}, \boldsymbol{z} \models \delta)$$

This means that Γ holds in any worlds $\boldsymbol{z} \succeq \boldsymbol{x}$ for the vector $\boldsymbol{x}$ of nominals occurring in Γ. This includes the outer nominal x_i in δ (if any). In this case the semantics of $\sqsubseteq$ is preserved, since $\boldsymbol{z}$ includes $z_i \succeq x_i$. With the sake of a more detailed analysis, we consider the following instance:

$$\dfrac{x : \alpha_1, y : \alpha_2 \Rightarrow x : \beta}{\alpha_1 \Rightarrow x : \alpha_2 \sqsubseteq \beta}\ \sqsubseteq\text{-r}$$

Consider an $i\mathcal{ALC}$ structure $\mathcal{I} = \langle \mathcal{U}, \preceq, R^{\mathcal{I}} \ldots, C^{\mathcal{I}} \rangle$ In this case, for any $\mathcal{I}$ and any $z_1, z_2 \in \mathcal{U}^{\mathcal{I}}$ if $z_1 \succeq x^{\mathcal{I}}$, $z_1 \succeq y^{\mathcal{I}}$, such that, $\mathcal{I}, z_i \models \alpha_1$ and $\mathcal{I}, z_i \models \alpha_2$, we have that $\mathcal{I}, z_i \models x : \beta$, since the premise is valid, by hypothesis. In this case, by the semantics of $\sqsubseteq$ we have $\mathcal{I}, z_i \models x : \alpha_1 \sqsubseteq \beta$. The conclusion of the rule is valid too.

The argument shown above for the $\sqsubseteq$-r rule is analogous for the nominal versions of $\sqsubseteq$-r, $\sqsubseteq$-l,$\sqcap$-r,$\sqcap$-l, $\sqcup_1$-r,$\sqcup_2$-r, $\sqcup$-l. Consider the rule $\forall$-r.

$$\frac{\mathit{\Delta}, xRy \Rightarrow\ y\colon \alpha}{\mathit{\Delta} \Rightarrow\ x\colon \forall R.\alpha}\ \forall\text{-r}$$

Since the premise is valid we have that if $\forall z_x \succeq x^{\mathcal{I}}$, $\forall z_y \succeq y^{\mathcal{I}}$, $(z_x, z_y) \in R^{\mathcal{I}}$ then $\forall z_y \succeq y^{\mathcal{I}}.\mathcal{I}, z_y \models \gamma$. This entails that $x^{\mathcal{I}} \in (\forall R.\gamma)^{\mathcal{I}}$, for $x^{\mathcal{I}} \succeq x^{\mathcal{I}}$. We observe that by the restriction on the rule application, y does not occur in $\mathit{\Delta}$, it only occurs in xRy and $y : \alpha$. The truth of these formulas are subsumed by $\forall R.\gamma$. The conclusion does not need to consider them any more. The conclusion is valid too. Another way to see its soundness is to prove that if $xRy \Rightarrow y : \alpha$ is valid, then so is $\Rightarrow x : \forall R.\alpha$. This can be show by the following reasoning:

$$\forall x^{\mathcal{I}} \forall y^{\mathcal{I}} \forall z_x \forall z_y (z_x \succeq x^{\mathcal{I}} \rightarrow (z_y \succeq y^{\mathcal{I}} \rightarrow ((z_x, z_y) \in R^{\mathcal{I}} \rightarrow \mathcal{I}, z_y \models y : \alpha)))$$

that is the same as:

$$\forall x^{\mathcal{I}} \forall y^{\mathcal{I}} \forall z_x \forall z_y (z_x \succeq x^{\mathcal{I}} \rightarrow (z_y \succeq y^{\mathcal{I}} \rightarrow ((z_x, z_y) \in R^{\mathcal{I}} \rightarrow \mathcal{I}, y^{\mathcal{I}} \models \alpha)))$$

Using the fact that $\forall y^{\mathcal{I}}(y^{\mathcal{I}} \succeq y^{\mathcal{I}})$, we obtain:

$$\forall x^{\mathcal{I}} \forall z_x (z_x \succeq x^{\mathcal{I}} \rightarrow \forall y^{\mathcal{I}}((z_x, y^{\mathcal{I}}) \in R^{\mathcal{I}} \rightarrow \mathcal{I}, y^{\mathcal{I}} \models \alpha))$$

The above condition states that $\Rightarrow x : \forall R.\alpha$ is valid.

$$\forall x^{\mathcal{I}} \forall y^{\mathcal{I}} \forall z_x \forall z_y (z_x \succeq x^{\mathcal{I}} \rightarrow (z_y \succeq y^{\mathcal{I}} \rightarrow ((z_x, z_y) \in R^{\mathcal{I}} \rightarrow \mathcal{I}, z_y \models y : \alpha)))$$

Consider the rule $\forall$-l:

$$\frac{\mathit{\Delta}, x\colon \forall R.\alpha, y\colon \alpha, xRy \Rightarrow\ \delta}{\mathit{\Delta}, x\colon \forall R.\alpha, xRy \Rightarrow\ \delta}\ \forall\text{-l}$$

As in the $\forall$-r case, we analyze the simplest validity preservation: if $x : \forall R.\alpha \wedge xRy$ is valid, then so is $x : \forall R.\alpha \wedge y : \alpha \wedge xRy$. The first condition is:

$$\begin{aligned}&\forall x^{\mathcal{I}} \forall y^{\mathcal{I}} \forall z_x (z_x \succeq x^{\mathcal{I}} \rightarrow \forall z_y (z_y \succeq y^{\mathcal{I}} \rightarrow\\ &\qquad ((\mathcal{I}, z_y \models x : \forall R.\alpha) \wedge (\mathcal{I}, z_y \models x : \forall R.\alpha) \wedge ((z_x, z_y) \in R^{\mathcal{I}}) \rightarrow\\ &\qquad\qquad (\mathcal{I}, z_y \models y : \alpha) \wedge (\mathcal{I}, z_x \models y : \alpha))))\end{aligned} \tag{1.2}$$

Using $z_y = y^{\mathcal{I}}$, eliminating z_x from the term, and, using the fact that $\mathcal{I}, z_y \models y : \alpha$ is valid, iff, $\mathcal{I}, y^{\mathcal{I}} \models \alpha$, we obtain

$$\begin{aligned}&\forall x^{\mathcal{I}} \forall y^{\mathcal{I}} \forall z_x (z_x \succeq x^{\mathcal{I}} \rightarrow \forall z_y (z_y \succeq y^{\mathcal{I}} \rightarrow\\ &\qquad ((\mathcal{I}, z_y \models x : \forall R.\alpha) \wedge (\mathcal{I}, z_y \models x : \forall R.\alpha) \wedge ((z_x, z_y) \in R^{\mathcal{I}}) \rightarrow\\ &\qquad\qquad (\mathcal{I}, y \models \alpha))))\end{aligned} \tag{1.3}$$

Consider the semantics of $\exists R.\alpha$:

$$(\exists R.\alpha)^{\mathcal{I}} =_{df} \{x \mid \exists y \in \mathcal{U}^{\mathcal{I}}.(x, y) \in R^{\mathcal{I}} \text{ and } y \in \alpha^{\mathcal{I}}\}$$

and the following rule:

$$\frac{\mathit{\Delta} \Rightarrow\ xRy \qquad \mathit{\Delta} \Rightarrow\ y\colon \alpha}{\mathit{\Delta} \Rightarrow\ x\colon \exists R.\alpha}\ \exists\text{-r}$$

We can see that the premises of the rule entails the conclusion. The premises correspond to the following conditions:

$$\forall x^{\mathcal{I}} \forall y^{\mathcal{I}} \forall z_x (z_x \succeq x^{\mathcal{I}} \rightarrow \forall z_y (z_y \succeq y^{\mathcal{I}} \rightarrow ((z_x, z_y) \in R^{\mathcal{I}})))$$

and

$$\forall y^{\mathcal{I}} \forall z_y (z_y \succeq y^{\mathcal{I}} \rightarrow ((\mathcal{I}, z_y \models y : \alpha)))$$

Instantiating in both conditions $z_y = y^{\mathcal{I}}$ and $z_x = x^{\mathcal{I}}$, this yields $(x^{\mathcal{I}}, y^{\mathcal{I}}) \in R^{\mathcal{I}}$, such that $\mathcal{I}, y^{\mathcal{I}} \models \alpha$, so $\mathcal{I}, z_x \models x^{\mathcal{I}} : \exists R.\alpha$. Thus, $\exists$-r is sound. The soundness of $\exists$-l is analogous to $\forall$-l.

Finally, it is worth noting that, for each rule, we can derive the soundness of its non-nominal version from the proof of soundness of its nominal version. For instance, the soundness of the nominal version of rule $\sqcup$-l depends on the diamond conditions F1 and F2. The soundness of its non-nomimal version, is a consequence of the soundness of the nominal version.

The rules below have their soundness proved as a consequence of the following reasonings in first-order intuitionistic logic that are used for deriving the semantics of the conclusions from the semantics of the premises:

(p-$\exists$) $\forall x(A(x) \wedge B(x) \to C(x)) \models \forall x A(x) \wedge \exists x B(x) \to \exists x C(x)$;
(p-$\forall$) $(A(x) \models B(x))$ implies $\forall y(R(y,x) \to A(x)) \models \forall y(R(y,x) \to B(x))$;
(p-N) if $A \models B$ then for every Kripke model $\mathcal{I}$ and world $x^{\mathcal{I}}$, if $\mathcal{I}, x^{\mathcal{I}} \models A$ then $\mathcal{I}, x^{\mathcal{I}} \models B$.

$$\frac{\Delta, \alpha \Rightarrow \ \beta}{\forall R.\Delta, \exists R.\alpha \Rightarrow \ \exists R.\beta}\text{ p-}\exists \quad \frac{\Delta \Rightarrow \ \alpha}{\forall R.\Delta \Rightarrow \ \forall R.\alpha}\text{ p-}\forall \quad \frac{\Delta \Rightarrow \ \delta}{x\colon \Delta \Rightarrow \ x\colon \delta}\text{ p-N}$$

8 Chisholm paradox in iALC

The paradoxes discussed in this section are known from the literature as contrary-to-duty paradoxes. They are deontic paradoxes under SDL formalization. Usually, there is a primary norm/law/obligation and a secondary norm that comes to effect when the primary obligation is violated. The form of these normative and intuitively coherent situations are in general hard to find a consistent deontic formalization. Because of that they are called paradoxes. A typical example of contrary-to-duty paradox appeared in [19]:

1. It ought to be that Jones goes to the assistance of his neighbors.
2. It ought to be that if Jones does go then he tells them he is coming.
3. If Jones doesn't go, then he ought not tell them he is coming.
4. Jones doesn't go.

– This certainly appears to describe a possible situation. 1-4 constitute a mutually consistent and logically independent set of sentences.
– (1) is a primary obligation, what Jones ought to do unconditionally. (2) is a *compatible-with-duty* obligation, appearing to say (in the context of 1) what else Jones ought to do on the condition that Jones fulfills his primary obligation. (3) is a *contrary-to-duty* obligation (*CTD*) appearing to say (in the context of 1) what Jones ought to do conditional on his violating his primary obligation. (4) is a factual claim, which conjoined with (1), implies that Jones violates his primary obligation.

We firstly remember the deontic approch to law and its logic. Differently of ours, it takes *laws* as propositions. Thus, a norm or *law* is an obligatory proposition, such as "You must pay your debits" or "It is obligatory to pay the debits". As a proposition each norm has a truth value. The underlying logic classical. If ϕ is a proposition then $O\phi$ is a proposition too. $O\phi$ intuitively means ϕ must be the case ,or *It is obligatory that* ϕ. The paradoxes that we discuss in this work appear just when *laws* are taken as propositions. They show them up from the most basic deontic logic Standard Deontic Logic (SDL). SDL is a modal logic defined by von Wright19951 [31] and, according to the modal logic terminology on the names of axioms, it is defined by the following set of axioms. The formulas of SDL include the modality O.

TAUT all tautologies of the language. This means that if ϕ is a propositional tautology then the substitution of p for any SDL formula is an SDL tautology too;
OB-K $O(p \to q) \to (Op \to Oq)$
OB-D $Op \to \neg O \neg p$
MP if $\vdash p$ and $\vdash p \to q$ then $\vdash q$
OB-NEC if $\vdash p$ then $\vdash Op$

SDL is just the normal modal logic D or KD, with a suggestive notation expressing the intended interpretation. From these, we can prove the principle that obligations cannot conflict, *NC of SDL*, $\neg(Op \wedge O\neg p)$ (see [31])

The following set of formulas is a straightforward formalization of Chisholm paradox in SDL.

1. Op

2. $O(p \to q)$
3. $\neg p \to O\neg q$
4. $\neg p$

The intuitive meaning of each formula is according the following table, where p is "Jones go to the assistance of his neighbours" and q is "Jones tells his neighbours he is going".

It ought to be that Jones go to the assistance of his neighbours.	$O(p)$
It ought to be that if Jones does go then he tells them he is going.	$O(p \to q)$
If Jones doesn't go, then he ought not tell them he is going.	$\neg p \to Ob(\neg q)$
Jones doesn't go.	$\neg p$

Using the deductive power of SDL we can perform the following derivation of a SDL contradiction.

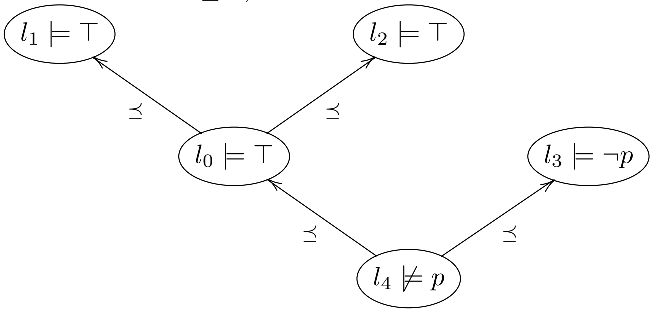
- from (2) by principle *OB-K* we get $Op \to Oq$,
- and then from (1) by *MP*, we get Oq;
- but by *MP* alone we get $O\neg q$ from (3) and (4).
- From these two conclusions, by *PC*, we get $Oq \wedge O\neg q$, contradicting *NC of SDL*.

Assertion 1-4, from Chisholm paradox, leads to inconsistency per SDL. But, 1-4 do not seem inconsistent at all, the representation cannot be a faithful one. We discuss this in the sequel. For reasons that will become clear, we take Chisholm paradox as stated above in natural language, instead of its SDL version. We use the same letters to denote the propositions/laws as used in the deontic representation of the paradox, for a better comparison.

In first law in the paradox, i.e., the law state in item 1 is a nominal in $i\mathcal{ALC}$, and hence it is a Kripke world in our model. The same can be said about item 2. The state-of-affairs, expressed in $i\mathcal{ALC}$, is simply the assertions: $l_1 : \top$ and $l_2 : \top$. Note that this assertions only state that there are two laws l_1 and l_2 in the legal universe. Since a Kripke model for intuitionistic logic is a Heyting algebra, and hence it is a lattice too, there must be the *meet* of these two worlds. This is represented in the model by law l_0, intuitively stating that it is obligatory to do what law l_1 and law l_2 state. Item 3 of the paradox is a conditional that generally states that if some proposition is truth then some law exists. This is a rather hard expression in judicial terms. Laws exist by promulgation only, they do not have their existence conditioned to anything but their own promulgation. This conditional expression can be raised in a legislative discussion only. But even in this exceptional case, the raising of paradoxes, as the one under discussion, advices that such use should be avoided. What item 3 says, instead, is simply that $\neg p$ holds in the world l_3 that is the law cited as the consequent of the conditional. Finally, as the model is a lattice, there must be a world l_4 that represents the law that it is the conjunctive law related to l_0 and l_3, in the same way l_0 is related to l_1 and l_2. Now, in l_4, it is ensured that $\neg p$ holds by the intuitionistic interpretation of the negation. As a result the model depicted in the diagram below is a model for what is known by Chisholm paradox. Thus, it is not a paradox when expressed in $i\mathcal{ALC}$in a kelsenian way.

1. The law l_1, originally Op
2. The law l_2, originally $O(p \to q)$
3. From (3), $\neg p \to O\neg q$, we have $l_3 : \neg p$. If we had $O\neg q \to \neg p$ the translation would be the same. That is, l_3 is $O\neg q$.
4. The law l_0 that represents the infinum of l_1 and l_2.

The following diagram shows the model to Chisholm paradox discussed above. Remember that if $x : A$ then $\forall x' \geq x, x' : A$.



9 Conclusion

In this article, we shown how intuitionistic logic and Kelsen's jurisprudence can be used to express Chisholm paradox faithfully. A key fact in providing a logical model to this paradox is that laws/norms are not taken as propositions. For example, in the explanation above on building the model, if we turn back to deontic expression of laws, we will have that l_1 is Op and l_2 is $O(p \to q)$, but we cannot derive that l_3 is $O(q)$. l_3 is of course the meet ($\sqcap$) between l_1 and l_2, as a meet it is strongly connection to $O(l_1) \wedge O(l_2) \leftrightarrow O(l_1 \wedge l_2)$, which is a SDL valid formula. Thus, l_3 is the norm $O(l_1 \wedge l_2)$, that is an obligation. However, now remembering what norms l_1 and l_2 are in this particular case, l_3 is the meeting $O(p) \wedge O(p \to q)$ that it is $O(p \wedge (p \to q))$. This conclusion, however, does not entail that in l_3 can be identified with $O(q)$, since our implication is the intuitionistic implication. This very last aspect of joining Kelsen jurisprudence and $i\mathcal{ALC}$also helps to avoid other deontic paradoxes.

Jorgensen's Dilemma [15] offers a question, in fact, a dilemma, whether there is, in fact, any deontic logic. The question follows this path: 1) Norms/laws deal with evaluative sentences; 2) Evaluative sentences are not the kind of sentence that can be true or false; 3) Thus, how there is a logic of evaluative sentences? 4) Logic has as goal to define what can be drawn from whatever, and; 5) A sentence follows from a set of sentences on a basis of the relationship between the truth of the sentences in question. Thus, there is no deontic logic. What we have shown in this article, is that deontic logic is possible by considering the logic of norms as a logic on norms, instead. This reading is just what we do in legal ontologies.

Only to estimate the range of our approach for solving semantical (contrary-to-duty) paradoxes. The free choice permission paradox reported on [27]) is derived from Jorgensen dilemma. Consider the tautology $p \to p \vee q$. We have by necessity SDL rule that $O(p \to p \vee q)$ is derivable in SDL by the axiom K we have $O(p) \to O(p \vee q)$. On the other hand, by contra-positive we have $\neg(p \vee q) \to \neg p$, so, and hence $O(\neg(p \vee q)) \to O(\neg q)$, and by contra-positive again, we obtain $\neg O(\neg q) \to \neg O(\neg(p \vee q))$. Taking $\neg O(\neg \phi)$ as permitted ϕ, rather better $P(\phi)$, we draw $P(p) \to P(p \vee q)$. The free choice paradox, based on the fact rule that if $P(p \vee q)$ holds then $P(p) \wedge P(q)$. If we accept the free choice permission, so to say the formula $P(p \vee q) \to then(P(p) \wedge P(q))$. As $p \to (p \vee q)$, then by what was observed above, $P(p) \to P(p \vee q)$, then by the free choice permission, we draw that $P(p) \to P(p \wedge q)$, for any q. In summary, in the presence of the free choice permission axiom, we can derive that $P(p) \to P(q)$ for every q, which should not happen. However, if we use $i\mathcal{ALC}$ together with Kelsen jurisprudence, and hence intuitionistic logic, we cannot derive all the steps above. Anyway, our definition of permission is different from what is discussed in this paragraph, see below. We can see that many paradoxes that are based on the axiom K are not paradoxes anymore.

We have to touch some aspects that are very well-known in the deontic approach. One is the deontic concept of *permission*. This case is modeled by observing that in a society regulated by law, permission is nothing more than an obligation of the State. The State promulgates what is allowed. Concerning prohibitions, the foundation is analogous. However, some subtle and theoretical problems may arise if one wants to recover the definition of forbidden (F) regarding the very well-known duality $F(p) \equiv O(\neg p)$. This discussion will be the subject of another article.

Finally, we would like to comment that professor Farinas del Cerro taught us that the research on logic and AI, mainly the first should be approached by solving part-by-part the problem and elegantly putting everything together. Well, we learned the first part of this technique, by following him by reading his articles. We think that in the first part, so to say finding a good foundation on legal ontologies, one that comes from the domain itself, namely Kelsen jurisprudence. Concerning the second part, that is to put everything together in an elegant way, we known that we are far from it. Another thing that we might have learned from prof. Farinas is that in this case, it is a matter of time to have the work in a more mature stage. We hope we reach this stage.

References

1. C. E. Alchourron and A.A. Martino. Logic without truth. *Ratio Juris*, 3(1):46–67, 1990.
2. G. Bellin, V de Paiva, and E. Ritter. Extended curry-howard correspondence for a basic constructive modal logic. In *Procs of Methods for the Modalities*, 2001.
3. T. Braüner and Valéria de Paiva. Intuitionistic hybrid logic. *JAL*, 4(3):231–255, 2006.
4. Samuel R Buss. *Handbook of proof theory*. Elsevier, 1998.
5. I. Clément. Proof theoretical foundations for constructive description logic. Master's thesis, School of Computer Science. McGill University, Montrel, July 2008.

6. Valeria de Paiva. Constructive description logics: what, why and how. In *Context Representation and Reasoning*, 2006.
7. Valeria de Paiva, Edward Hermann Hausler, and Alexandre Rademaker. Constructive description logic: Hybrid-style. In *Proceedings of Hybrid Logic and Applications (HyLo 2010)*, 2010.
8. G. Fischer-Servi. Axiomatizations for some intuitionistic modal logics. *Rendiconti del Seminario Matematico Università e Politecnico di Torino*, 42, 1984.
9. J.Y. Girard, P. Taylor, and Y. Lafont. *Proofs and Types.* Cambridge Tracts in Theoretical Computer Science 7. Cambridge University Press, 1989.
10. Edward H. Haeusler, Valeria de Paiva, and Alexandre Rademaker. Intuitionistic description logic and legal reasoning. In *Proceedings of International Workshop Data, Logic and Inconsistency with DEXA 2011*, 2011.
11. Edward Hermann Haeusler, Valéria de Paiva, and Alexandre Rademaker. Intuitionistic logic and legal ontologies. In *Proc. JURIX 2010*, pages 155–158. IOS Press, 2010.
12. Edward Hermann Haeusler, Valeria de Paiva, and Alexandre Rademaker. Using intuitionistic logic as a basis for legal ontologies. In *Proceedings of the 4th Workshop on Legal Ontologies and Artificial Intelligence Techniques*, pages 69–76, Fiesole, Florence, Italy, 2010. European University Institute.
13. Edward Hermann Haeusler, Valeria De Paiva, and Alexandre Rademaker. Using intuitionistic logic as a basis for legal ontologies. *Informatica e Diritto*, XIX(1-2):289–298, 2010.
14. Graham Hughes. Validity and the basic norm. *California Law Review*, 59(3):695–715, 1971.
15. Jorgen Jorgensen. Imperatives and logic. *Erkenntnis*, 7:288–296, 1937.
16. H. Kelsen. *General Theory of Law and State.* Harvard University Press, 1946. Translation from german by Wedberg, A. and Kraus, W.H.
17. Hans Kelsen. *Pure theory of law.* Berkeley : University of California Press, 1967. Translated from Reine Rechtslehre, second edition, 1960.
18. Hans Kelsen. *General theory of norms.* Oxford Univ. Press, USA, 1991.
19. Chisholm Roderick M. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24(33–36), 1963.
20. N.M. Peron M.E. Coniglio. A paraconsistent approach to chisholm's paradox. *Principia*, 13(3):299–326, 2009.
21. Michael Mendler and Stephan Scheele. Towards constructive DL for abstraction and refinement. *JAR*, 44(3):207–243, 2010. Proc. 21st International DL Workshop.
22. Sara Negri and Jan Von Plato. *Structural proof theory.* Cambridge University Press, 2008.
23. Gordon Plotkin and Colin Stirling. A framework for intuitionistic modal logics: extended abstract. In *Proceedings of the 1986 conference on Theoretical aspects of reasoning about knowledge*, TARK '86, pages 399–406, San Francisco, CA, USA, 1986. Morgan Kaufmann Publishers Inc.
24. Alexandre Rademaker. *A Proof Theory for Description Logics.* Springer, 2012.
25. K. Ranalter. Embedding constructive k into intuitionistic k. *Electr. Notes Theor. Comput. Sci.*, 262:205–219, 2010.
26. Joseph Raz. Legal principles and the limits of law. *Yale Law Journal*, 81:823–854, 1972.
27. Alf Ross. Imperatives and logic. *Theoria, 7*, 1941.
28. Alex Simpson. *The Proof Theory and Semantics of Intuitionistic Modal Logic.* PhD thesis, University of Edinburgh, December 1993.
29. Gaisi Takeuti. *Proof theory.* Courier Dover Publications, 2013.
30. A. Valente. *Legal knowledge engineering: A modelling approach.* IOS Press, 1995. Amsterdam.
31. G. Von Wright. Deontic logic. *Mind*, 60:1–15, 1951.

A Note on Barcan Formula

Antonio Frías Delgado ⋆

University of Cádiz
antonio.frias@uca.es

Abstract. We present in this note a plea for Barcan Formula. This view connects Barcan Formula with a modal principle that normalizes the ∀-Introduction rule of quantificational logic.

1 Introduction

R. C. Barcan introduced a modal system as "an extension of the Lewis calculus *S2* to include quantification"[1].

The axiom schema 11 of this system is

$\Diamond(\exists\alpha)A \prec (\exists\alpha)\Diamond A$

Hereafter, as it is usual, Barcan Formula (BF) will be the equivalent formula

BF: $\forall x\Box\varphi(x) \rightarrow \Box\forall x\varphi(x)$

In quantified modal logic[1], BF occupies an outstanding and discussed place[3]. One reason lies in the fact that its converse is a theorem of QK[2]. Therefore, in systems in which BF holds, $\forall\Box$ and $\Box\forall$ become equivalent. This equivalence could be regarded as a tool to replace *de re* modalities and protect modal logic from criticism.

The problems BF arises can be separated into three levels:

1. Proof-theory: what modal axioms and/or modal inference rules suffice to prove -or, they are incompatible with- BF.
2. Model-theory: what properties of the models of modal logic yield valid BF.
3. Ontology: what ontological commitments, with regard to individuals and their properties, are assumed in accepting BF.

This last level is what has caused more discussions. Several authors have disputed the philosophical problems -contingency, actualism, possibilism,. . . - directly or indirectly connected with BF[2,4]. It is also the level where the different positions are difficult to reconcile, because the commitments about modalities that are acceptable for some logician or philosopher could be excessive for another one, and the ontological preferences are frequently not founded on logical reasons.

2 A minimum modal principle to prove Barcan Formula

Prior proved that BF is a theorem of QS5[7]. Lemmon -according to Prior- proved that BF is a theorem of a weaker system than QS5: QB, the so called "Brouwersche" modal logic. BF is neither provable in QS4, nor, therefore, in QT. BF cannot be added to QL[6].

However, if Q is a standard first order logic[3], BF is provable in a modal system weaker than QB, because the axiom schema T is not necessary in the deduction of BF. The deduction of BF in [5] uses only:

⋆ Luis Fariñas del Cerro introduced me in modal logics. I thank his kind intellectual support and his friendship.

[1] See [5] for terminology. The modal systems we shall mention here are obtained from the inference rule

N: if $\vdash \alpha$, then $\vdash \Box\alpha$

and the axiom schemata

K: $\Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)$

T: $\Box\alpha \rightarrow \alpha$

B(rouwer): $\alpha \rightarrow \Box\Diamond\alpha$

4: $\Box\alpha \rightarrow \Box\Box\alpha$

5: $\Diamond\alpha \rightarrow \Box\Diamond\alpha$

L(öb): $\Box(\Box\alpha \rightarrow \alpha) \rightarrow \Box\alpha$

PL is a propositional logic. Q is a first order logic.

QK=Q+N+K. QT=QK+T. QB=QT+B. QS4=QT+4. QS5=QT+5. QL=QK+L.

4 is a theorem of QS5 and QL.

[2] $\Box\forall x\varphi(x) \rightarrow \forall x\Box\varphi(x)$ is obtained from the standard rules for quantifiers, N and K.

[3] The condition that Q is a standard first order logic is not superfluous. If Q* is a first order logic without free variables in their axiom schemata, for instance, then BF is unprovable in Q*S5 (Kripke).

(a) An axiom schema and an inference rule for quantifiers (the standard ones),
(b) A modal axiom equivalent to B,
(c) The inference rules:
(c.1) $\vdash \alpha \rightarrow \beta \Rightarrow \vdash \Diamond\alpha \rightarrow \Diamond\beta$
(c.2) $\vdash \Diamond\alpha \rightarrow \Diamond\beta \Rightarrow \vdash \alpha \rightarrow \Box\beta$
It is immediate that (c.1) is proved from N and K, and (c.2) is proved from N and B.
Is B a minimum modal axiom to prove BF? It seems so, if the quantified axioms and rules of Q are:
(UE) $\forall(x)\alpha \rightarrow \alpha$
(UI) $\vdash \alpha \rightarrow \beta \Rightarrow \vdash \alpha \rightarrow \forall x\beta(x)$, x not free in α
We do not present a strict proof, but an argument only. Our reasoning is as follows.
We search for a set $\delta_1, \ldots, \delta_n$ of formulas such that
$\vdash \forall x\Box\varphi(x) \rightarrow \delta_1$
$\vdash \delta_{n-1} \rightarrow \delta_n$
$\vdash \delta_n \rightarrow \Box\forall x\varphi(x)$
At a first step in the deduction of BF, we have
(f) $\vdash \forall x\Box\varphi(x) \rightarrow \Box\varphi(x)$, from (UE).
The $\forall$-introduction in the consequent of (f) firstly requires the $\Box$-elimination using some principle as
(1) $\vdash \mu\Box\alpha \rightarrow \alpha$,
where μ is a sequence of modal operators.
Moreover, if $\mu \neq \emptyset$, the following rule has to be added:
(2) $\vdash \alpha \rightarrow \beta \Rightarrow \vdash \mu\alpha \rightarrow \mu\beta$
From the axiom (1), the rule (2), and (f),
(g) $\vdash \mu\forall x\Box\varphi(x) \rightarrow \varphi(z)$
(h) $\vdash \mu\forall x\Box\varphi(x) \rightarrow \forall x\varphi(x)$, from (g), (UI).
(i)$\vdash \Box\mu\forall x\Box\varphi(x) \rightarrow \Box\forall x\varphi(x)$, from N, K.
Now, for obtaining BF we need a principle equivalent to
(3) $\vdash \alpha \rightarrow \Box\mu\alpha$
What is the minimal modal axiom that satisfies the above principles (1), (2), (3)?

- Case $\mu = \emptyset$
  (1) is the axiom schema T; (2) holds. But (3) is not admissible as a modal principle. T+(3) collapse the modality because $\vdash \alpha \leftrightarrow \Box\alpha$
- Case $\mu = \Box$.
  (1) is obtained from T, and (2) from N. But (3) is not admissible because $\vdash \alpha \leftrightarrow \Box\Box\alpha$
- Case $\mu = \Diamond$
  (2) is obtained from N, and K. (1) and (3) are equivalent. They both express modal principles that may be admitted.

Hence, (1) or (3) seem to be a minimal axiom schema that, added to QK, proves BF.

3 Validity of Barcan Formula: Two Ways

In standard possible world semantics, broadly, if we suppose BF is false we find the following scene:
(i) $\forall x\Box\varphi(x)$ is true at w_0 , and
(ii) $\exists x\neg\varphi(x)$ is true at some w_j, and
(iii) w_0Rw_j .
We have to find ways to generate a contradiction.
From (ii) we have, for a new variable z,
(iv) $\neg\varphi(z)$ is true at w_j .
From (i) we have
(v) $\Box\varphi(z)$ is true at w_0 .
By the necessity rule, we could have
(vi) $\varphi(z)$ is true at w_j,
but nothing guarantees that the entity z denotes in (iv) is the same entity z denotes in (vi) -which is the same entity z denotes in w_0-, unless every individual of the quantificational domain of w_j is included into the quantificational domain of w_0.

The validity of BF requires to have the semantical means to displace a selected individual -or a set of individuals- from a world to another. The scope of a universal quantifier must be moved from the world it appears to every accessible world. This can be achieved in two ways:
(a) By the symmetry of R: $w_jRw_k \Rightarrow w_kRw_j$.

(b) Without restrictive conditions on R, by the inclusion of the domain of each world into the domain of the worlds to which it is accessible: $w_j R w_k \Rightarrow \mathcal{D}(w_k) \subseteq \mathcal{D}(w_j)$.

Condition (a) implies the condition (b).

The validity of BF does not require R is reflexive, according to our proof above, where T was not used. Actually, the validity of BF needs the condition (b) only; symmetry is a way to obtain it.

If we search for a minimum semantic principle that validates BF, the question arises whether (a) or (b) are the most appropriate principle.

The minimum modal requisite for a technical treatment of *de re* modalities forces us only to accept the following principle:

(c) the domain of w_j has to be included into the domain of w_k, if $w_j R w_k$.

The reason is that x has to belong to the domain of every w_k that $w_j R w_k$, for evaluating $\varphi(x)$ as true -false- at w_k, when $\Box\varphi(x)$ is true -false- at w_j.

The (c) principle does not suffice for the validity of BF without either

(1) additional properties of R -e.g. the (a) way above-, or

(2) additional requirements on the domains of accessible worlds -the (b) way.

BF itself allows only an interchange between quantifiers and modal operators, whereas the axiom schema B expresses that all truth is necessarily possible. The *modal* range (i.e. concerning *modal laws*) of BF seems more restricted than the modal range of B. In standard modal logic if we accept B as a means to equal the domains of accessible worlds, we have to accept also $\varphi(t) \rightarrow \Box\Diamond\varphi(t)$ as theorem, for any formula without quantifiers. A fact -or a choice- about quantifiers only (BF) does not have to drag whatever formula. Therefore, it seems rather natural to ask whether BF could be proved in a modal system different from QK+B. The most elemental answer, but lacking in interest, is to add BF to QK. This answer, however, does not clear up what new principles of reasoning, on quantification and modalities, we are actually adding to N and K when BF becomes provable. What is the *logical* principle underlying BF?

4 Barcan Formula and ∀-Introduction Rule

Following ideas introduced by Gödel in his famous paper (1931), we could interpret the N rule as saying that $\Box$ *partially defines* the deducibility[4]. However, we can not say that $\Box$ *strongly defines* some kind of deducibility, because there is no rule that specifies the behavior of $\Box$ with regard to the undeducible formulas. No rule of type $\nvdash \alpha \Rightarrow \vdash \neg\Box\alpha$ exists in standard modal logics. In QT, for example, the rule $\vdash \neg\alpha \Rightarrow \vdash \neg\Box\alpha$ is valid; but we obtain $\neg\Box\alpha$ formulas as theorems, only if we have already $\neg\alpha$ as a theorem, and not, as a general case, if $\nvdash \alpha$ merely. Therefore, $\Box$ only captures partially the undeducibility. This is the case also in QB and QS5.

$\Box$ partially defines the deducibility. This is true in any modal system that has the N rule. We could assume that K formalizes into the modal logic the proof method that corresponds to MP. Similarly, we could think to add a modal axiom that formalizes the ∀-I rule. This modal axiom is the equivalent to K with regard to MP:

G: $\Box(\alpha \rightarrow \beta) \rightarrow \Box(\alpha \rightarrow \forall x\beta(x))$, x not free in α.

If we define Q replacing the ∀-I rule for the axiom schema

$\forall x(\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \forall x\beta(x))$, x not free in α

and the rule

(UI*): $\vdash \varphi(x) \Rightarrow \vdash \forall x\varphi(x)$

then, we add the modal axiom:

(G*): $\Box\varphi(x) \rightarrow \Box\forall x\varphi(x)$

BF is provable from G (G*).

For the case G:

$\Box(\alpha \rightarrow \varphi) \rightarrow \Box(\alpha \rightarrow \forall x\varphi(x)) \vdash \forall x\Box\varphi(x) \rightarrow \Box\forall x\varphi(x)$, x not free in α

A proof follows:

(i) $\varphi(x) \rightarrow (\varphi(z) \rightarrow \varphi(z))$, by PL

(ii) $\Box\varphi(x) \rightarrow \Box(\varphi(z) \rightarrow \varphi(x))$, from N, K

(iii) $\forall x\Box\varphi(x) \rightarrow \Box\varphi(x)$, from ∀-E

(iv) $\forall x\Box\varphi(x) \rightarrow \Box(\varphi(z) \rightarrow \varphi(x))$, from (ii), (iii)

(v) $\forall x\Box\varphi(x) \rightarrow \Box(\varphi(z) \rightarrow \forall x\varphi(x))$, from (iv), QN

(vi) $\forall x\Box\varphi(x) \rightarrow (\Box\varphi(z) \rightarrow \Box\forall x\varphi(x))$, from (v), K

[4] $\mathcal{P}$ partially defines P in M if and only if:
if $P(\alpha)$ holds, then $\vdash_M \mathcal{P}(\alpha)^t$.
$\mathcal{P}$ strongly defines P in M if and only if:
$\mathcal{P}$ partially defines P in M, and if $P(\alpha)$ does not hold, then $\vdash_M \neg\mathcal{P}(\alpha)^t$.
^{t} is some assignment or translation.

(vii) $\forall x \Box \varphi(x) \rightarrow \Box \varphi(z)$, from $\forall$-E
(viii) $\forall x \Box \varphi(x) \rightarrow \Box \forall x \varphi(x)$, from (vi), (vii)
For the case G* the proof is immediate
$\Box \varphi(x) \rightarrow \Box \forall x \varphi(x) \vdash \forall x \Box \varphi(x) \rightarrow \Box \forall x \varphi(x)$
We obtain BF in the modal system QK+G.

What advantage do we get against the QK+B alternative? Since we impose restrictions on the world domains and not on the properties of the accessibility relation, we need fewer commitments on the properties of broad modalities, and quantifier-free modal formulas are not biased when we choose only BF.

Bibliography

1. Barcan (Marcus), R. C. A functional calculus of first order based on strict implication. *Journal of Symbolic Logic* **11** (1946), 1-16.
2. Burgess, J.P. Which Modal Logic Is the Right One? *Notre Dame Journal of Formal Logic* **40** (1) (1999), 81-93.
3. Fitting, M. Barcan Both ways. *Journal of Applied Non-Classical Logics* **9** (2-3), (2012), 329-344.
4. Hayaki, R. Contingent Objects and the Barcan Formula. *Erkenntnis* **64** (1) (2006), 87-95.
5. Hughes, G. E., and Cresswell, M. J. *A New Introduction to Modal Logic* . Routledge, London, 1996
6. Montagna, F. The predicate modal logic of provability. *Notre Dame Journal of Formal Logic* **25** (2) (1984), 179-189.
7. Prior, A. Modality and Quantification in S5. *The Journal of Symbolic Logic* **21** (1956), 60-62.

KD and KT are Nullary

Philippe Balbiani[1] and Çiğdem Gencer[2]

[1]Institut de recherche en informatique de Toulouse
Université de Toulouse
[2]Department of Mathematics and Computer Science
Istanbul Kültür University — Faculty of Science and Letters

Abstract. In the ordinary modal language, KD and KT are the modal logics respectively determined by the class of all serial frames and the class of all reflexive frames. In this paper, we demonstrate that KD and KT are nullary.

1 Introduction

The unification problem in a logical system L can be defined as follows: given a formula $\phi(x_1, \ldots, x_n)$, determine whether there exists formulas $\psi_1, \ldots, \psi_n$ such that $\phi(\psi_1, \ldots, \psi_n)$ is in L. The research on unification was motivated by a closely related and more general decision problem, namely the admissibility problem for rules of inference: given a rule $\frac{\phi_1(x_1,\ldots,x_n),\ldots,\phi_m(x_1,\ldots,x_n)}{\psi(x_1,\ldots,x_n)}$, decide whether for all formulas $\chi_1, \ldots, \chi_n$, if $\phi_1(\chi_1, \ldots, \chi_n)$, ..., $\phi_m(\chi_1, \ldots, \chi_n)$ are in L then $\psi(\chi_1, \ldots, \chi_n)$ is in L. The admissibility problem for rules was put forward by Friedman [15] who asked whether there exists a decision procedure for deciding whether a given rule preserves validity in intuitionistic logic.
Friedman's problem was solved by Rybakov [23,24] who demonstrated that the admissibility problem in intuitionistic logic and modal logic $S4$ is decidable. See also [20,26,30] for a study of unification and inference rules for modal logics. Later on, Ghilardi [17], proving that intuitionistic logic has a finitary unification type, yielded a new solution of Friedman's problem, seeing that deciding whether a given rule preserves validity in intuitionistic logic is equivalent to checking whether the finitely many maximal unifiers of its premises are unifiers of its conclusion. See also [19] for a study of unification and most general unifiers in modal logics. With respect to the complexity issue, Jerábek [21] established the $coNEXPTIME$-completeness of the admissibility problem for several intermediate logics and several $K4$-extensions, in contrast with the admissibility problem for modal logics contained in $K4$ which is undecidable if one considers a language with the universal modality [31]. See also [16] for a study of unifiability in extensions of $K4$.
Is the situation better if the language is restricted? Cintula and Metcalfe [11] considered the negation-implication fragment of intuitionistic logic and proved that the associated admissibility problem was $PSPACE$-complete. Unification of concept terms has been introduced by Baader and Narendran [6] as a tool for detecting redundancies in knowledge bases. In this respect, Baader and Küsters [3] established the $EXPTIME$-completeness of the unification problem in the description logic $\mathcal{FL}_0$ whereas Baader and Morawska [4,5] established the $NPTIME$-completeness of the unification problem in the description logic $\mathcal{EL}$.
Tense logics and epistemic logics provide formalisms for expressing properties about programs, time, knowledge, etc. Within their context, Dzik [13,14] has studied the relationships between the unification type of a fusion of modal logics and the unification types of the modal logics composing this fusion. The unification type of applied non-classical logics such as common knowledge logics and linear temporal logics has also been studied by Babenyshev and Rybakov [7] and Rybakov [27,28,29]. Nevertheless, very little is known about the unification problem in some of the most important description and modal logics considered in Computer Science and Artificial Intelligence. For example, the decidability of the unification problem for the following description and modal logics remains open: description logic $\mathcal{ALC}$, modal logic K, multimodal variants of K, sub-Boolean fragments of modal logics.
In the ordinary modal language, the modal logics KD and KT are the least normal logics respectively containing the formulas $\Box x \rightarrow \Diamond x$ and $\Box x \rightarrow x$. They are also the modal logics determined by the class of all frames (W, R) such that R is serial on W and the class of all frames (W, R) such that R is reflexive on W. Seeing that $\Box\bot$, $\Box\top$, $\Diamond\bot$ and $\Diamond\top$ are, respectively, equivalent in KD and KT to $\bot$, $\top$, $\bot$ and $\top$, it is a well-known fact that KD-unification and KT-unification are in NP. As for the unification type of KD and KT, in this paper, following a line of reasoning suggested by Jerábek [22] within the context of the modal logic K, we demonstrate that KD and KT are nullary.

2 Syntax

Let VAR be an at most countable set of propositional variables (with typical members denoted x, y, etc) and PAR be an at most countable set of propositional parameters (with typical members denoted p, q, etc). In this paper, we will always assume that $VAR \neq \emptyset$. The set $\mathcal{L}$ of all formulas (with typical members denoted ϕ, ψ, etc) is inductively defined as follows:

– $\phi, \psi ::= x \mid p \mid \bot \mid \neg\phi \mid (\phi \vee \psi) \mid \Box\phi$.

We write $\phi(x_1, \ldots, x_n)$ to denote a formula whose variables form a subset of $\{x_1, \ldots, x_n\}$. The Boolean connectives $\top$, $\wedge$, $\rightarrow$ and $\leftrightarrow$ are defined by the usual abbreviations. Let $\Diamond$ be the modal connective defined as follows:

– $\Diamond\phi ::= \neg\Box\neg\phi$.

For all parameters p, the modal connective $[p]$ is defined as follows:

– $[p]\phi ::= \Box(p \rightarrow \phi)$.

For all parameters p, the modal connective $[p]^k$ is inductively defined as follows for each $k \in \mathbb{N}$:

– $[p]^0\phi ::= \phi$,
– $[p]^{k+1}\phi ::= [p][p]^k\phi$.

For all parameters p, the modal connective $[p]^{<k}$ is inductively defined as follows for each $k \in \mathbb{N}$:

– $[p]^{<0}\phi ::= \top$,
– $[p]^{<k+1}\phi ::= [p]^{<k}\phi \wedge [p]^k\phi$.

We adopt the standard rules for omission of the parentheses.

Example 1. $\phi = (x \rightarrow p) \wedge (x \rightarrow [p]x)$ is a readable abbreviation for the less readable formula $\neg(\neg(\neg x \vee p) \vee \neg(\neg x \vee \Box(\neg p \vee x)))$.

The degree of a formula ϕ (in symbols $\deg(\phi)$) is inductively defined as follows:

– $\deg(x) = 0$,
– $\deg(p) = 0$,
– $\deg(\bot) = 0$,
– $\deg(\neg\phi) = \deg(\phi)$,
– $\deg(\phi \vee \psi) = \max\{\deg(\phi), \deg(\psi)\}$,
– $\deg(\Box\phi) = \deg(\phi) + 1$.

A substitution is a function σ associating to each variable x a formula $\sigma(x)$. We shall say that a substitution σ is closed if for all variables x, $\sigma(x)$ is a variable-free formula. For all formulas $\phi(x_1, \ldots, x_m)$, let $\sigma(\phi(x_1, \ldots, x_m))$ be $\phi(\sigma(x_1), \ldots, \sigma(x_n))$. The composition $\sigma \circ \tau$ of the substitutions σ and τ associates to each variable x the formula $\tau(\sigma(x))$.

Example 2. If ϕ is the formula considered in Example 1 and σ_p is the substitution defined by $\sigma_p(x) = p$ then $\sigma_p(\phi) = (p \rightarrow p) \wedge (p \rightarrow [p]p)$.

Example 3. If ϕ is the formula considered in Example 1, $k \in \mathbb{N}$ and σ_k is the substitution defined by $\sigma_k(x) = p \wedge [p]^{<k}x \wedge [p]^k\bot$ then $\sigma_k(\phi) = (p \wedge [p]^{<k}x \wedge [p]^k\bot \rightarrow p) \wedge (p \wedge [p]^{<k}x \wedge [p]^k\bot \rightarrow [p](p \wedge [p]^{<k}x \wedge [p]^k\bot))$.

3 Semantics

A frame is a relational structure of the form $\mathcal{F} = (W, R)$ where W is a nonempty set of states (with typical members denoted s, t, etc) and R is a binary relation on W. A model based on a frame $\mathcal{F} = (W, R)$ is a relational structure of the form $\mathcal{M} = (W, R, V)$ where V is a function associating to each variable x a set $V(x)$ of states and to each parameter p a set $V(p)$ of states. The relation "formula ϕ is true in model $\mathcal{M}$ at state s" (in symbols $\mathcal{M}, s \models \phi$) is inductively defined as follows:

– $\mathcal{M}, s \models x$ iff $s \in V(x)$,
– $\mathcal{M}, s \models p$ iff $s \in V(p)$,
– $\mathcal{M}, s \not\models \bot$,

- $\mathcal{M}s \models \neg\phi$ iff $\mathcal{M}, s \not\models \phi$,
- $\mathcal{M}, s \models \phi \vee \psi$ iff either $\mathcal{M}, s \models \phi$, or $\mathcal{M}, s \models \psi$,
- $\mathcal{M}, s \models \Box\phi$ iff for all states $t \in W$, if sRt then $\mathcal{M}, t \models \phi$.

Let $\mathcal{C}$ be a class of frames. We shall say that a formula ϕ is $\mathcal{C}$-valid (in symbols $\mathcal{C} \models \phi$) if for all frames $\mathcal{F} = (W, R)$ in $\mathcal{C}$, for all models $\mathcal{M} = (W, R, V)$ based on $\mathcal{F}$ and for all states $s \in W$, $\mathcal{M}, s \models \phi$.

Example 4. The following formulas are valid in the class of all frames:

- $[p]p$,
- $[p]^{<k}x \wedge [p]^k\bot \rightarrow [p]([p]^{<k}x \wedge [p]^k\bot)$.

Let $\mathcal{C}$ be a class of frames. We shall say that a substitution σ is $\mathcal{C}$-equivalent to a substitution τ (in symbols $\sigma \simeq_{\mathcal{C}} \tau$) if for all variables x, $\mathcal{C} \models \sigma(x) \leftrightarrow \tau(x)$. We shall say that a substitution σ is more $\mathcal{C}$-general than a substitution τ (in symbols $\sigma \preceq_{\mathcal{C}} \tau$) if there exists a substitution υ such that $\sigma \circ \upsilon \simeq_{\mathcal{C}} \tau$.

4 Unification problem

Let $\mathcal{C}$ be a class of frames. We shall say that a formula ϕ is $\mathcal{C}$-unifiable if there exists a substitution σ such that $\mathcal{C} \models \sigma(\phi)$. In that case, σ is a $\mathcal{C}$-unifier of ϕ.

Example 5. Let $\mathcal{C}$ be a class of frames. If ϕ is the formula considered in Example 1 then the substitution σ_p considered in Example 2 is a $\mathcal{C}$-unifier of ϕ.

Example 6. Let $\mathcal{C}$ be a class of frames. If ϕ is the formula considered in Example 1 and $k \in \mathbb{N}$ then the substitution σ_k considered in Example 3 is a $\mathcal{C}$-unifier of ϕ.

Given a class $\mathcal{C}$ of frames, an important question is the following:

$\mathcal{C}$-unification: given a formula ϕ, decide whether ϕ is $\mathcal{C}$-unifiable.

Let $\mathcal{C}^{KD}$ be the class of all serial frames and $\mathcal{C}^{KT}$ be the class of all reflexive frames.

Proposition 1. *If $PAR = \emptyset$ then $\mathcal{C}^{KD}$-unification and $\mathcal{C}^{KT}$-unification are in NP.*

Proof. Suppose $PAR = \emptyset$.
$\mathcal{C}^{KD}$-unification: Hence, in $\mathcal{C}^{KD}$, every variable-free formula is equivalent to $\bot$ or $\top$. This is a well-known fact. It partly follows from the fact that $\Box\bot$, $\Box\top$, $\Diamond\bot$ and $\Diamond\top$ are, respectively, $\mathcal{C}^{KD}$-equivalent to $\bot$, $\top$, $\bot$ and $\top$. Thus, every closed substitution σ is $\mathcal{C}^{KD}$-equivalent to a substitution τ such that for each variable x, $\tau(x) = \bot$ or $\tau(x) = \top$. Moreover, if a formula ϕ possesses a $\mathcal{C}^{KD}$-unifier then ϕ possesses a closed $\mathcal{C}^{KD}$-unifier. This follows from the fact that for all $\mathcal{C}^{KD}$-unifiers σ of ϕ and for all closed substitutions τ, $\sigma \circ \tau$ is a closed $\mathcal{C}^{KD}$-unifier of ϕ. Consequently, for all formulas ϕ, the following conditions are equivalent: ϕ is $\mathcal{C}^{KD}$-unifiable; there exists a $\mathcal{C}^{KD}$-unifier σ of ϕ such that for all variables x, $\sigma(x) = \bot$ or $\sigma(x) = \top$. Hence, for all formulas $\phi(x_1, \ldots, x_n)$, to decide whether $\phi(x_1, \ldots, x_n)$ is $\mathcal{C}^{KD}$-unifiable, it suffices to nondeterministically guess $\psi_1, \ldots, \psi_n \in \{\bot, \top\}$ and to determine whether $\phi(\psi_1, \ldots, \psi_n)$ is $\mathcal{C}^{KD}$-equivalent to $\bot$ or $\top$. Obviously, this can be done in polynomial time.
$\mathcal{C}^{KT}$-unification: Similar to $\mathcal{C}^{KD}$-unification.

The decidability status of $\mathcal{C}^{KD}$-unification and $\mathcal{C}^{KT}$-unification are unknown when $PAR \neq \emptyset$. Let $\mathcal{C}^{S5}$ be the class of all partitions.

Proposition 2. *If $PAR = \emptyset$ then $\mathcal{C}^{S5}$-unification is in NP.*

Proof. Similar to the proof of Proposition 1.

$\mathcal{C}^{S5}$-unification remains decidable when $PAR \neq \emptyset$. See Balbiani and Gencer [9] for details. Let $\mathcal{C}^{Alt_1}$ be the class of all deterministic frames.

Proposition 3. *$\mathcal{C}^{Alt_1}$-unification is in $PSPACE$.*

Proof. See [10].

Let $\mathcal{C}^{K4}$ be the class of all transitive frames, $\mathcal{C}^{S4}$ be the class of all reflexive transitive frames and $\mathcal{C}^{GL}$ be the class of all transitive well-founded frames.

Proposition 4. *1. $\mathcal{C}^{K4}$-unification is decidable.*
2. $\mathcal{C}^{S4}$-unification is decidable (in NP when $PAR = \emptyset$).
3. $\mathcal{C}^{GL}$-unification is decidable.

Proof. See [18].

As for the decidability status of unification in the class $\mathcal{C}^K$ of all frames, it is unknown.

5 Unification type

Let $\mathcal{C}$ be a class of frames. We shall say that a set Σ of unifiers of a unifiable formula ϕ is complete if for all unifiers σ of ϕ, there exists a unifier τ of ϕ in Σ such that $\tau \preceq_{\mathcal{C}} \sigma$. An important question is the following: when a formula is unifiable, has it a minimal complete set of unifiers? When the answer is "yes", how large is this set? We shall say that a unifiable formula

- ϕ is nullary if there exists no minimal complete set of unifiers of ϕ,
- ϕ is infinitary if there exists an infinite minimal complete set of unifiers of ϕ but there exists no finite one,
- ϕ is finitary if there exists a finite minimal complete set of unifiers of ϕ but there exists no with cardinality 1,
- ϕ is unitary if there exists a minimal complete set of unifiers of ϕ with cardinality 1.

We shall say that

- $\mathcal{C}$ is nullary if there exists a nullary formula,
- $\mathcal{C}$ is infinitary if every unifiable formula is either infinitary, or finitary, or unitary and there exists a infinitary formula,
- $\mathcal{C}$ is finitary if every unifiable formula is either finitary, or unitary and there exists a finitary formula,
- $\mathcal{C}$ is unitary if every unifiable formula is unitary.

Proposition 5. *If $PAR \neq \emptyset$ then $\mathcal{C}^{KD}$ is nullary.*

Proof. See Section 6.

The unification type of $\mathcal{C}^{KD}$ is unknown when $PAR = \emptyset$.

Proposition 6. *$\mathcal{C}^{S5}$ is unitary.*

Proof. See [2].

Proposition 7. *$\mathcal{C}^{Alt_1}$ is nullary.*

Proof. See [10].

Proposition 8. *1. $\mathcal{C}^{K4}$ is finitary.*
2. $\mathcal{C}^{S4}$ is finitary.
3. $\mathcal{C}^{GL}$ is finitary.

Proof. See [18].

As for the unification type of $\mathcal{C}^{K}$, it is nullary [22].

6 Some nullary modal logics

In this section, we will always assume that $PAR \neq \emptyset$. Let $\mathcal{C}$ be a class of frames. Let $\phi = (x \rightarrow p) \wedge (x \rightarrow [p]x)$ be the formula considered in Example 1. Let $\Sigma = \{\sigma_p\} \cup \{\sigma_k : k \in \mathbb{N}\}$ where σ_p is the substitution defined by $\sigma_p(x) = p$ and considered in Example 2 and for all $k \in \mathbb{N}$, σ_k is the substitution defined by $\sigma_k(x) = p \wedge [p]^{<k}x \wedge [p]^k\bot$ and considered in Example 3. By Examples 5 and 6, we know that Σ is a set of unifiers of ϕ.

Lemma 1. *Let $k, l \in \mathbb{N}$. If $k \leq l$ then $\sigma_l \preceq_{\mathcal{C}} \sigma_k$.*

Proof. Suppose $k \leq l$. Let υ be the substitution defined by $\upsilon(x) = x \wedge [p]^k\bot$. The reader may easily verify that $\mathcal{C} \models \upsilon(\sigma_l(x)) \leftrightarrow \sigma_k(x)$. Hence, $\sigma_l \preceq_{\mathcal{C}} \sigma_k$.

Lemma 2. *Let $k, l \in \mathbb{N}$. If $k < l$ then $\sigma_k \not\preceq_{\mathcal{C}} \sigma_l$.*

Proof. Suppose $k < l$ and $\sigma_k \preceq_{\mathcal{C}} \sigma_l$. Let υ be a substitution such that $\sigma_k \circ \upsilon \simeq_{\mathcal{C}} \sigma_l$. Hence, $\mathcal{C} \models \upsilon(\sigma_k(x)) \leftrightarrow \sigma_l(x)$. Thus, $\mathcal{C} \models p \wedge [p]^{<l}x \wedge [p]^l\bot \rightarrow [p]^k\bot$. Consequently, $\mathcal{C} \models p \wedge [p]^l\bot \rightarrow [p]^k\bot$: a contradiction.

Lemma 3. *Let σ be a substitution. The following conditions are equivalent:*

1. $\sigma_p \circ \sigma \simeq_{\mathcal{C}} \sigma$.
2. $\sigma_p \preceq_{\mathcal{C}} \sigma$.
3. $\mathcal{C} \models \sigma(x) \leftrightarrow p$.

Proof. (1. $\Rightarrow$ 2) : By definition of $\preceq_{\mathcal{C}}$.
(2. $\Rightarrow$ 3) : Suppose $\sigma_p \preceq_{\mathcal{C}} \sigma$. Let υ be a substitution such that $\sigma_p \circ \upsilon \simeq_{\mathcal{C}} \sigma$. Hence, $\mathcal{C} \models \upsilon(\sigma_p(x)) \leftrightarrow \sigma(x)$. Thus, $\mathcal{C} \models \sigma(x) \leftrightarrow p$.
(3. $\Rightarrow$ 1) : Suppose $\mathcal{C} \models \sigma(x) \leftrightarrow p$. Hence, $\mathcal{C} \models \sigma(\sigma_p(x)) \leftrightarrow \sigma(x)$. Thus, $\sigma_p \circ \sigma \simeq_{\mathcal{C}} \sigma$.

Lemma 4. *Let $k \in I\!N$. Let σ be a unifier of ϕ. The following conditions are equivalent:*

1. $\sigma_k \circ \sigma \simeq_{\mathcal{C}} \sigma$.
2. $\sigma_k \preceq_{\mathcal{C}} \sigma$.
3. $\mathcal{C} \models \sigma(x) \to [p]^k\bot$.

Proof. (1. $\Rightarrow$ 2) : By definition of $\preceq_{\mathcal{C}}$.
(2. $\Rightarrow$ 3) : Suppose $\sigma_k \preceq_{\mathcal{C}} \sigma$. Let υ be a substitution such that $\sigma_k \circ \upsilon \simeq_{\mathcal{C}} \sigma$. Hence, $\mathcal{C} \models \upsilon(\sigma_k(x)) \leftrightarrow \sigma(x)$. Thus, $\mathcal{C} \models \sigma(x) \to [p]^k\bot$.
(3. $\Rightarrow$ 1) : Suppose $\mathcal{C} \models \sigma(x) \to [p]^k\bot$. Since σ is a unifier of ϕ, therefore $\mathcal{C} \models \sigma(x) \to p$ and $\mathcal{C} \models \sigma(x) \to [p]\sigma(x)$. Hence, $\mathcal{C} \models \sigma(x) \to [p]^{<k}\sigma(x)$. Since $\mathcal{C} \models \sigma(x) \to [p]^k\bot$ and $\mathcal{C} \models \sigma(x) \to p$, therefore $\mathcal{C} \models \sigma(x) \to \sigma(\sigma_p(x))$. Now, we consider the following 2 cases.
Case $k = 0$: Thus, $\mathcal{C} \models [p]^k\bot \to \sigma(x)$.
Case $k \geq 1$: Consequently, $\mathcal{C} \models [p]^{<k}\sigma(x) \to \sigma(x)$.
In both cases, $\mathcal{C} \models \sigma(\sigma_p(x)) \to \sigma(x)$. Since $\mathcal{C} \models \sigma(x) \to \sigma(\sigma_p(x))$, therefore $\mathcal{C} \models \sigma(\sigma_p(x)) \leftrightarrow \sigma(x)$. Hence, $\sigma_p \circ \sigma \simeq_{\mathcal{C}} \sigma$.

Lemma 5. *Let σ be a unifier of ϕ. If $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$ then one of the following conditions holds:*

1. $\sigma_p \preceq_{\mathcal{C}} \sigma$.
2. *There exists $k \in I\!N$ such that $\sigma_k \preceq_{\mathcal{C}} \sigma$.*

Proof. Suppose $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$ and none of the above conditions holds. By Lemmas 3 and 4, $\mathcal{C} \not\models \sigma(x) \leftrightarrow p$ and $\mathcal{C} \not\models \sigma(x) \to [p]^{\deg(\sigma(x))}\bot$. Since σ is a unifier of ϕ, therefore $\mathcal{C} \models \sigma(x) \to p$. Let $\mathcal{F} = (W, R)$ and $\mathcal{F}' = (W', R')$ be frames in $\mathcal{C}$, $\mathcal{M} = (W, R, V)$ and $\mathcal{M}' = (W', R\prime, V')$ be models based respectively on $\mathcal{F}$ and $\mathcal{F}'$ and $s \in W$ and $s' \in W'$ be states such that $\mathcal{M}, s \not\models p \to \sigma(x)$ and $\mathcal{M}', s' \not\models \sigma(x) \to [p]^{\deg(\sigma(x))}\bot$. Hence, $\mathcal{M}, s \models p$, $\mathcal{M}, s \not\models \sigma(x)$, $\mathcal{M}', s' \models \sigma(x)$ and $\mathcal{M}', s' \not\models [p]^{\deg(\sigma(x))}\bot$. Let $t'_0, \ldots, t'_{\deg(\sigma(x))} \in W'$ be states such that $t'_0 = s'$ and for all $i \in I\!N$, if $i < \deg(\sigma(x))$ then $t'_i R' t'_{i+1}$ and $t'_{i+1} \in V'(p)$. Since $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$, therefore without loss of generality, we may assume that $t'_0, \ldots, t'_{\deg(\sigma(x))}$ is the shortest p-path in W' between t'_0 and $t'_{\deg(\sigma(x))}$. Let $\mathcal{M}'' = (W'', R'', V'')$ be the model defined as follows:

– $W'' = W \cup W'$,
– $R'' = R \cup R' \cup \{(t', s)\}$,
– $V'' = V \cup V'$.

Since $\mathcal{M}, s \models p$ and $\mathcal{M}, s \not\models \sigma(x)$, therefore $\mathcal{M}'', s \models p$ and $\mathcal{M}'', s \not\models \sigma(x)$. Since $t'_0, \ldots, t'_{\deg(\sigma(x))}$ is the shortest p-path in W' between t'_0 and $t'_{\deg(\sigma(x))}$, $\mathcal{M}', s' \models \sigma(x)$, $t'_0 = s'$ and for all $i \in I\!N$, if $i < \deg(\sigma(x))$ then $t'_i R' t'_{i+1}$ and $t'_{i+1} \in V'(p)$, therefore $\mathcal{M}'', s' \models \sigma(x)$. Since σ is a unifier of ϕ, therefore $\mathcal{C} \models \sigma(x) \to [p]\sigma(x)$. Since $\mathcal{M}'', s \models p$ and $\mathcal{M}'', s' \models \sigma(x)$, therefore $\mathcal{M}'', s \models \sigma(x)$: a contradiction.

Lemma 6. *If $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$ then Σ is a complete set of unifiers of ϕ.*

Proof. By Lemma 5.

Lemma 7. *If $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$ then there exists no minimal complete set of unifiers of ϕ.*

Proof. Suppose $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$ and there exists a minimal complete set of unifiers of ϕ. Let Γ be a minimal complete set of unifiers of ϕ. Let $\gamma \in \Gamma$ be such that $\gamma \preceq_{\mathcal{C}} \sigma_0$. Since $\mathcal{C} = \mathcal{C}^{KD}$ or $\mathcal{C} = \mathcal{C}^{KT}$, therefore by Lemma 6, let $\sigma \in \Sigma$ be such that $\sigma \preceq_{\mathcal{C}} \gamma$. Now, we consider the following 2 cases.
Case $\sigma = \sigma_p$: Since $\gamma \preceq_{\mathcal{C}} \sigma_0$, therefore $\sigma \preceq_{\mathcal{C}} \sigma_0$. Let υ be a substitution such that $\sigma \circ \upsilon \simeq_{\mathcal{C}} \sigma_0$.

Hence, $\mathcal{C} \models \upsilon(\sigma(x)) \leftrightarrow \sigma_0(x)$. Thus, $\mathcal{C} \models \neg p$: a contradiction.
Case $\sigma = \sigma_k$ for some $k \in \mathbb{N}$: Let $\gamma' \in \Gamma$ be such that $\gamma' \preceq_{\mathcal{C}} \sigma_{k+1}$. Since $\sigma \preceq_{\mathcal{C}} \gamma$, therefore by Lemma 1, $\gamma' \preceq_{\mathcal{C}} \gamma$. Since Γ is a minimal complete set of unifiers of ϕ, therefore $\gamma' = \gamma$. Since $\gamma' \preceq_{\mathcal{C}} \sigma_{k+1}$ and $\sigma \preceq_{\mathcal{C}} \gamma$, therefore $\sigma_k \preceq_{\mathcal{C}} \sigma_{k+1}$. Since $k < k+1$, therefore by Lemma 2, $\sigma_k \not\preceq_{\mathcal{C}} \sigma_{k+1}$: a contradiction.

Finally, we obtain the

Proposition 9. *$\mathcal{C}^{KD}$ and $\mathcal{C}^{KT}$ are nullary.*

Proof. By Lemma 7.

7 Additional comments

In the context of modal logics, classes of frames such as the ones underlying K, KD and KT give rise to quite similar sets of valid formulas for what concerns axiomatization and decidability. Putting known results adapted from [2,12,13,18,22] together with new ones enables us to establish basic facts and outline open problems. See Tab. 1.1. While the study of K, KD and KT has now limited mathematical interest for what concerns axiomatization and decidability, considering unification types in modal logics is justified from applied perspectives: methods for deciding the unifiability of formulas can be used to improve the efficiency of automated theorem provers [8]; deciding the unifiability of formulas like $\phi \leftrightarrow \psi$ helps us to understand what is the overlap between the properties ϕ and ψ correspond to [2]; in description logics, unification algorithms are used to detect redundancies in knowledge-based systems [1]. One readily observes that, while attacking the above-mentioned problems, little, if anything, from the standard tools in modal logics (canonical models, filtrations, etc) is helpful. In order to successfully solve them, new techniques in modal logics must be developed and much remains to be done. The study of unification types in modal logics has still many secrets to reveal.

Class of frame	PAR	Computability	Type
$\mathcal{C}^{KD}$	$= \emptyset$	in NP (Proposition 1)	?
$\mathcal{C}^{KT}$	$= \emptyset$	in NP (Proposition 1)	?
$\mathcal{C}^{S5}$	$= \emptyset$	in NP (Proposition 2)	unitary ([2])
$\mathcal{C}^{Alt_1}$	$= \emptyset$	in $PSPACE$ ([10])	nullary ([10])
$\mathcal{C}^{K4}$	$= \emptyset$	decidable ([25])	finitary ([18])
$\mathcal{C}^{S4}$	$= \emptyset$	in NP (Proposition 2)	finitary ([18])
$\mathcal{C}^{GL}$	$= \emptyset$	decidable ([25])	finitary ([18])
$\mathcal{C}^{K}$	$= \emptyset$	?	nullary ([22])
$\mathcal{C}^{KD}$	$\neq \emptyset$	?	nullary (Proposition 9)
$\mathcal{C}^{KT}$	$\neq \emptyset$	?	nullary (Proposition 9)
$\mathcal{C}^{S5}$	$\neq \emptyset$	decidable ([9])	unitary
$\mathcal{C}^{Alt_1}$	$\neq \emptyset$	in $PSPACE$ ([10])	nullary ([10])
$\mathcal{C}^{K4}$	$\neq \emptyset$	decidable ([25])	finitary ([18])
$\mathcal{C}^{S4}$	$\neq \emptyset$	decidable ([25])	finitary ([18])
$\mathcal{C}^{GL}$	$\neq \emptyset$	decidable ([25])	finitary ([18])
$\mathcal{C}^{K}$	$\neq \emptyset$	?	nullary ([22])

Table 1.1.

Acknowledgement

This paper has been written on the occasion of a 3-months visit of Çiğdem Gencer during the Fall 2015 in Toulouse that was financially supported by the *Paul Sabatier University* ("Professeurs invités 2015") and the *Institut de recherche en informatique de Toulouse* ("Actions spécifiques 2015").

Special acknowledgement

Special acknowledgement is heartly granted to Luis Fariñas del Cerro. His research on modal logic as a general tool for the formalization of human reasoning has exerted a profound influence on our research. To our knowledge, this paper does not relate in any way to Luis' own papers. We hope it can constitute the subject of a joint work in the near future.

Bibliography

1. Baader, F., Borgwardt, S., Morawska, B.: *Extending unification in $\mathcal{EL}$ towards general TBoxes.* In Brewka, G. *et al.* (editors): *Principles of Knowledge Representation and Reasoning.* AAAI Press (2012) 568–572.
2. Baader, F., Ghilardi, S.: *Unification in modal and description logics.* Logic Journal of the IGPL **19** (2011) 705–730.
3. Baader, F., Küsters, R.: *Unification in a description logic with transitive closure of roles.* In Nieuwebhuis, R., Voronkov, A. (editors): *Logic for Programming and Automated Reasoning.* Springer (2001) 217–232.
4. Baader, F., Morawska, B.: *Unification in the description logic $\mathcal{EL}$.* In Treinen, R. (editor): *Rewriting Techniques and Applications.* Springer (2009) 350–364.
5. Baader, F., Morawska, B.: *SAT encoding of unification in $\mathcal{EL}$.* In Fermüller, C., Voronkov, A. (editors): *Logic for Programming, Artificial Intelligence, and Reasoning.* Springer (2010) 97–111.
6. Baader, F., Narendran, P.: *Unification of concept terms in description logics.* Journal of Symbolic Computation **31** (2001) 277–305.
7. Babenyshev, S., Rybakov, V.: *Unification in linear temporal logic LTL.* Annals of Pure and Applied Logic **162** (2011) 991–1000.
8. Babenyshev, S., Rybakov, V., Schmidt, R., Tishkovsky, D.: *A tableau method for checking rule admissibility in $S4$.* Electronic Notes in Theoretical Computer Science **262** (2010) 17–32.
9. Balbiani, P., Gencer, Ç.: *Unifiability and unification types in epistemic logics.* In preparation.
10. Balbiani, P., Tinchev, T.: *Unifiability and unification types in the normal modal logic Alt_1.* In preparation.
11. Cintula, P., Metcalfe, G.: *Admissible rules in the implication-negation fragment of intuitionistic logic.* Annals of Pure and Applied Logic **162** (2010) 162–171.
12. Dzik, W.: *Unitary unification of $S5$ modal logics and its extensions.* Bulletin of the Section of Logic **32** (2003) 19–26.
13. Dzik, W.: *Unification Types in Logic.* Wydawnicto Uniwersytetu Slaskiego (2007).
14. Dzik, W.: *Remarks on projective unifiers.* Bulletin of the Section of Logic **40** (2011) 37–46.
15. Friedman, H.: *One hundred and two problems in mathematical logic.* Journal of Symbolic Logic **40** (1975) 113–129.
16. Gencer, Ç., de Jongh, D.: *Unifiability in extensions of $K4$.* Logic Journal of the IGPL **17** (2009) 159–172.
17. Ghilardi, S.: *Unification in intuitionistic logic.* Journal of Symbolic Logic **64** (1999) 859–880.
18. Ghilardi, S.: *Best solving modal equations.* Annals of Pure and Applied Logic **102** (2000) 183–198.
19. Ghilardi, S., Sacchetti, L.: *Filtering unification and most general unifiers in modal logic.* Journal of Symbolic Logic **69** (2004) 879–906.
20. Iemhoff, R.: *On the admissible rules of intuitionistic propositional logic.* Journal of Symbolic Computation **66** (2001) 281–294.
21. Jerábek, E.: *Complexity of admissible rules.* Archive for Mathematical Logic **46** (2007) 73–92.
22. Jerábek, E.: *Blending margins: the modal logic K has nullary unification type.* Journal of Logic and Computation **25** (2015) 1231–1240.
23. Rybakov, V.: *A criterion for admissibility of rules in the model system $S4$ and the intuitionistic logic.* Algebra and Logic **23** (1984) 369–384.
24. Rybakov, V.: *Bases of admissible rules of the logics $S4$ and Int.* Algebra and Logic **24** (1985) 55–68.
25. Rybakov, V.: *Admissibility of Logical Inference Rules.* Elsevier Science (1997).
26. Rybakov, V.: *Construction of an explicit basis for rules admissible in modal system $S4$.* Mathematical Logic Quarterly **47** (2001) 441–446.
27. Rybakov, V.: *Unification in common knowledge logics.* Bulletin of the Section of Logic **31** (2002) 207–215.
28. Rybakov, V.: *Logical consecutions in discrete linear temporal logic.* Journal of Symbolic Logic **70** (2005) 1137–1149.
29. Rybakov, V.: *Multi-modal and temporal logics with universal formula — reduction of admissibility to validity and unification.* Journal of Logic and Computation **18** (2008) 509–519.
30. Rybakov, V., Terziler, M., Gencer, Ç.: *An essay on unification and inference rules for modal logics.* Bulletin of the Section of Logic **28** (1999) 145–157.
31. Wolter, F., Zakharyaschev, M.: *Undecidability of the unification and admissibility problems for modal and description logics.* ACM Transactions on Computational Logic **9** (2008) 25:1–25:20.

A Multimodal Logic for Closeness

Alfredo Burrieza, Emilio Muñoz-Velasco, and M. Ojeda-Aciego

University of Málaga. Andalucía Tech. Spain.
{burrieza,ejmunoz,aciego}@uma.es

Abstract. We introduce a multimodal logic for order of magnitude reasoning which includes the notions of closeness and negligibility, we provide an axiom system which is sound and complete.

1 Introduction

There are some multimodal logics for order of magnitude reasoning dealing with the relations of negligibility and comparability, see for instance [2,4,8]; however, as far as we know, the only published reference on the notion of closeness in a logic-based context is [6], where the notions of closeness and distance are treated using Propositional Dynamic Logic, and their definitions are based on the concept of qualitative sum; specifically, in [6] two values are assumed to be close if one of them can be obtained from the other by adding a small number, and small numbers are defined as those belonging to a fixed interval.

In this work, we consider a new logic-based alternative to the notion of closeness in the context of multimodal logics. Our notion of closeness stems from the idea that two values are considered to be *close* if they are inside a prescribed area or *proximity interval.* This idea applies to the situations described in the previous paragraph, although it may differ from other intuitions based on distances since it leads to an equivalence relation, particularly, transitivity holds. Neither reflexivity nor symmetry of closeness generate any discussion among the different authors, but transitivity does. The original notion of closeness given by Raiman in [9] allows a certain form of transitivity which he had to tame by using a number of arbitrary limitations to avoid an unrestricted application of chaining. This arbitrariness was criticized in [1], in which a fuzzy set-based approach for handling relative orders of magnitude was introduced. It is remarkable to note that the criticism was made against the arbitrary limitations on chaining the relation, or the impossibility of considering suitable modified versions of transitivity, but not on transitivity per se.

The limitations stated above do not apply to our approach, which can be seen as founded on the notion of granularity as given in [7], which was already suggested in [11]. The main difficulties in accepting closeness as a transitive relation arise in a distance-based interpretation because, then, its unrestricted use would collapse the relation since all the elements would be close. As stated above, our notion will be based not on distance but on membership to a certain element of a given set of proximity intervals, since our driving force is to define an abstract framework for dealing with natural or artificial barriers.

On the other hand, the negligibility notion provided in this paper is a slight generalization of the one given in [5] where, following the line of other classical approaches, for instance [10], the class of 0 is considered to be just a singleton. This choice makes little sense in a qualitative approach, since considering the class of 0 to be just a singleton would require to have measures with infinite precision. Instead, we consider the qualitative class INF of *infinitesimals* which, of course, will be all close to each other. Note that these infinitesimals will be interpreted as numbers indistinguishable from 0 in the sense that their difference cannot be measured, not in the sense of hyperreal numbers.

In this work, we introduce a multimodal logic for order of magnitude reasoning which manages the notions of closeness and negligibility, then an axiom system is introduced which is sound and complete.

2 Preliminary definitions

We will consider a subset of real numbers $(\mathbb{S}, <)$ divided into the following qualitative classes:

$$\begin{array}{lll} \textsc{nl} = (-\infty, -\gamma) & & \textsc{ps} = (+\alpha, +\beta] \\ \textsc{nm} = [-\gamma, -\beta) & \textsc{inf} = [-\alpha, +\alpha] & \textsc{pm} = (+\beta, +\gamma] \\ \textsc{ns} = [-\beta, -\alpha) & & \textsc{pl} = (+\gamma, +\infty) \end{array}$$

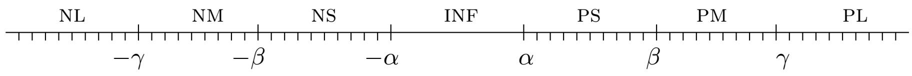


Fig. 1.1. Proximity intervals.

Note that all the intervals are considered relative to $\mathbb{S}$.

The labels correspond to "negative large" (NL), "negative medium"(NM), "negative small"(NS), "infinitesimals"(INF), "positive small" (PS), "positive medium" (PM) and "positive large" (PL). It is worth to note that this classification is slightly more general than the standard one [10], since the qualitative class containing the element 0, i.e. INF, needs not be a singleton; this allows for considering values very close to zero as null values in practice, which is more in line with a qualitative approach where accurate measurements are not always possible.

We will consider each qualitative class to be divided into disjoint intervals called *proximity intervals*, as shown in Figure 1.1. The qualitative class INF is itself one proximity interval.

Definition 1. *Let* $(\mathbb{S}, <)$ *be the set of numbers introduced above.*

- *An* r-proximity structure *is a finite set* $\mathcal{I}(\mathbb{S}) = \{I_1, I_2, \ldots, I_r\}$ *of intervals in* $\mathbb{S}$*, such that:*
  1. *For all* $I_i, I_j \in \mathcal{I}(\mathbb{S})$*, if* $i \neq j$*, then* $I_i \cap I_j = \varnothing$*.*
  2. $I_1 \cup I_2 \cup \cdots \cup I_r = \mathbb{S}$*.*
  3. *For all* $x, y \in \mathbb{S}$ *and* $I_i \in \mathcal{I}(\mathbb{S})$*, if* $x, y \in I_i$*, then* x, y *belong to the same qualitative class.*
  4. INF $\in \mathcal{I}(\mathbb{S})$*.*
- *Given a proximity structure* $\mathcal{I}(\mathbb{S})$*, the binary relation of closeness* $\mathfrak{c}$ *is defined, for all* $x, y \in \mathbb{S}$*, as follows:* $x \,\mathfrak{c}\, y$ *if and only if there exists* $I_i \in \mathcal{I}(\mathbb{S})$ *such that* $x, y \in I_i$*.*

Notice that, by definition, the number of proximity intervals is finite, regardless of the cardinality of the set $\mathbb{S}$. This choice is justified by the applications (the number of values we can consider is always finite) and the nature of the measuring devices that after reaching a certain limit, they do not distinguish among nearly equal amounts; for instance, consider the limits to represent numbers in a pocket calculator, thermometer, speedometer, etc.

The informal notion of negligibility we will use in this paper is the following: x is said to be *negligible* with respect to y if and only if either (i) x is infinitesimal and y is not, or (ii) x is small (but not infinitesimal) and y is *sufficiently large*. Formally:

Definition 2. *The binary relation of negligibility* $\mathfrak{n}$ *is defined on* $(\mathbb{S}, <)$ *as* $x \,\mathfrak{n}\, y$ *if and only if one of the following situations holds:*

(i) $x \in$ INF *and* $y \notin$ INF*,*
(ii) $x \in$ NS $\cup$ PS *and* $y \in$ NL $\cup$ PL*.*

3 A logic for closeness

In this section, we will use as special modal connectives $\overrightarrow{\Box}$ and $\overleftarrow{\Box}$ to deal with the usual ordering $<$, so $\overrightarrow{\Box}A$ and $\overleftarrow{\Box}A$ have the informal readings: *A is true for all numbers greater than the current one* and *A is true for all number less than the current one*, respectively. Two other modal operators will be used, $\boxed{c}$ for closeness, where the informal reading of $\boxed{c}A$ is: *A is true for all number close to the current one*, and $\boxed{n}$ for negligibility, where $\boxed{n}A$ means *A is true for all number with respect to the current one is negligible.*

The alphabet of the language $\mathcal{L}(MQ)^{\mathcal{P}}$ is defined by using a stock of atoms or propositional variables, $\mathcal{V}$, the classical connectives $\neg, \wedge, \vee$ and $\rightarrow$; the constants for milestones $\alpha^-, \alpha^+, \beta^-, \beta^+, \gamma^-, \gamma^+$; a finite set $\mathcal{C}$ of constants for proximity intervals, $\mathcal{C} = \{c_1, \ldots, c_r\}$ [1]; the unary modal connectives $\overrightarrow{\Box}$, $\overleftarrow{\Box}$, $\boxed{n}$, $\boxed{c}$, and the parentheses '(' and ')'. We define the formulas of $\mathcal{L}(MQ)^{\mathcal{P}}$ as follows:

$$A = p \mid \xi \mid c_i \mid \neg A \mid (A \wedge A) \mid (A \vee A) \mid (A \rightarrow A) \mid \overrightarrow{\Box}A \mid \overleftarrow{\Box}A \mid \boxed{n}A \mid \boxed{c}A$$

where $p \in \mathcal{V}$, $\xi \in \{\alpha^+, \alpha^-, \beta^+, \beta^-, \gamma^+, \gamma^-\}$ and $c_i \in \mathcal{C}$. In order to refer to any constant for positive milestones as α^+ we will use ξ^+ and for negative ones as β^- we will use ξ^-.

[1] There are at least as many elements in $\mathcal{C}$ as qualitative classes.

The *mirror image* of a formula A is the result of replacing in A each occurrence of $\overrightarrow{\Box}$, $\overleftarrow{\Box}$, α^+, β^+ and γ^+ respectively by $\overleftarrow{\Box}$, $\overrightarrow{\Box}$, α^-, β^- and γ^- and reciprocally. We will use the symbols $\overrightarrow{\Diamond}$, $\overleftarrow{\Diamond}$, $\langle\mathsf{c}\rangle$, $\langle\mathsf{n}\rangle$ as abbreviations, respectively, of $\neg\overrightarrow{\Box}\neg$, $\neg\overleftarrow{\Box}\neg$, $\neg\boxed{\mathsf{c}}\neg$ and $\neg\boxed{\mathsf{n}}\neg$. Moreover, we will introduce $\mathsf{nl}, \dots \mathsf{pl}$ as abbreviations for qualitative classes, for instance, ps for $(\overleftarrow{\Diamond}\alpha^+ \wedge \overrightarrow{\Diamond}\beta^+) \vee \beta^+$. By means of qc we denote any element of the set $\{\mathsf{nl}, \mathsf{nm}, \mathsf{ns}, \mathsf{inf}, \mathsf{ps}, \mathsf{pm}, \mathsf{pl}\}$.

The cardinality r of the set $\mathcal{C}$ of constants for proximity intervals will play an important role since it, somehow, encodes the granularity of the underlying logic. This implies that, actually, *we are introducing a family of logics which depend parametrically on r.*

Definition 3. *A* multimodal qualitative frame *for $\mathcal{L}(MQ)^{\mathcal{P}}$ (a* frame, *for short) is a tuple $\Sigma = (\mathbb{S}, \mathcal{D}, <, \mathcal{I}(\mathbb{S}), \mathcal{P})$, where:*

1. *$(\mathbb{S}, <)$ is an ordered subset of real numbers.*
2. *$\mathcal{D} = \{+\alpha, -\alpha, +\beta, -\beta, +\gamma, -\gamma\}$ is a set of designated points in $\mathbb{S}$ satisfying $-\gamma < -\beta < -\alpha < +\alpha < +\beta < +\gamma$.*
3. *$\mathcal{I}(\mathbb{S})$ is an r-proximity structure.*
4. *$\mathcal{P}$ is a bijection (called* proximity function*), $\mathcal{P}\colon \mathcal{C} \longrightarrow \mathcal{I}(\mathbb{S})$, that assigns to each proximity constant c a proximity interval.*

Definition 4. *Let Σ be a frame for $\mathcal{L}(MQ)^{\mathcal{P}}$, a* multimodal qualitative model on Σ *(a MQ-*model*, for short) is an ordered pair $\mathcal{M} = (\Sigma, h)$, where h is a* meaning function *(or,* interpretation*) $h\colon \mathcal{V} \longrightarrow 2^{\mathbb{S}}$. Any interpretation can be uniquely extended to the set of all formulas in $\mathcal{L}(MQ)^{\mathcal{P}}$ (also denoted by h) by means of the usual conditions for the classical Boolean connectives and the following conditions:*

$$
\begin{array}{l}
h(\overrightarrow{\Box}A) = \{x \in \mathbb{S} \mid y \in h(A) \text{ for all } y \text{ such that } x < y\} \\
h(\overleftarrow{\Box}A) = \{x \in \mathbb{S} \mid y \in h(A) \text{ for all } y \text{ such that } y < x\} \\
h(\boxed{\mathsf{c}}A) = \{x \in \mathbb{S} \mid y \in h(A) \text{ for all } y \text{ such that } x \,\mathfrak{c}\, y\} \\
h(\boxed{\mathsf{n}}A) = \{x \in \mathbb{S} \mid y \in h(A) \text{ for all } y \text{ such that } x \,\mathfrak{n}\, y\} \\
h(\alpha^+) = \{+\alpha\} \qquad h(\beta^+) = \{+\beta\} \qquad h(\gamma^+) = \{+\gamma\} \\
h(\alpha^-) = \{-\alpha\} \qquad h(\beta^-) = \{-\beta\} \qquad h(\gamma^-) = \{-\gamma\} \\
h(c_i) = \{x \in \mathbb{S} \mid x \in \mathcal{P}(c_i)\}
\end{array}
$$

The definitions of truth, satisfiability *and* validity *are the usual ones.*

Now, we consider the axiom system $MQ^{\mathcal{P}}$ for the language $\mathcal{L}(MQ)^{\mathcal{P}}$, consisting of all the tautologies of classical propositional logic together with the following axiom schemata and rules of inference:

For white connectives

K1 $\overrightarrow{\Box}(A \to B) \to (\overrightarrow{\Box}A \to \overrightarrow{\Box}B)$
K2 $A \to \overrightarrow{\Box}\overleftarrow{\Diamond}A$
K3 $\overrightarrow{\Box}A \to \overrightarrow{\Box}\overrightarrow{\Box}A$
K4 $\big(\overrightarrow{\Box}(A \vee B) \wedge \overrightarrow{\Box}(\overrightarrow{\Box}A \vee B) \wedge \overrightarrow{\Box}(A \vee \overrightarrow{\Box}B)\big) \to \big(\overrightarrow{\Box}A \vee \overrightarrow{\Box}B\big)$

For constants $\xi \in \{\alpha^+, \beta^+, \gamma^+, \alpha^-, \beta^-, \gamma^-\}$

c1 $\overleftarrow{\Diamond}\xi \vee \xi \vee \overrightarrow{\Diamond}\xi$
c2 $\xi \to (\overleftarrow{\Box}\neg\xi \wedge \overrightarrow{\Box}\neg\xi)$
c3 $\gamma^- \to \overrightarrow{\Diamond}\beta^-$
c4 $\beta^- \to \overrightarrow{\Diamond}\alpha^-$
c5 $\alpha^- \to \overrightarrow{\Diamond}\alpha^+$
c6 $\alpha^+ \to \overrightarrow{\Diamond}\beta^+$
c7 $\beta^+ \to \overrightarrow{\Diamond}\gamma^+$

For proximity constants (for all $i, j \in \{1, \dots, n\}$)

p1 $\bigvee_{i=1}^{n} c_i$
p2 $c_i \to \neg c_j$ (for $i \neq j$)
p3 $(\overleftarrow{\Diamond}c_i \wedge \overrightarrow{\Diamond}c_i) \to c_i$
p4 $\overleftarrow{\Diamond}c_i \vee c_i \vee \overrightarrow{\Diamond}c_i$

Mixed axioms (for all $i \in \{1, \dots, n\}$)

m1 $(c_i \wedge \mathsf{qc}) \rightarrow \left(\overleftarrow{\Box}(c_i \rightarrow \mathsf{qc}) \wedge \overrightarrow{\Box}(c_i \rightarrow \mathsf{qc})\right)$
m2 $(c_i \wedge \mathsf{inf}) \rightarrow \left(\overleftarrow{\Box}(\mathsf{inf} \rightarrow c_i) \wedge \overrightarrow{\Box}(\mathsf{inf} \rightarrow c_i)\right)$
m3 $\boxed{c}A \leftrightarrow \left(A \wedge \bigvee_{i=1}^{r}\left(c_i \wedge \overleftarrow{\Box}(c_i \rightarrow A) \wedge \overrightarrow{\Box}(c_i \rightarrow A)\right)\right)$
m4 $\boxed{n}A \leftrightarrow \Bigg(\left(\mathsf{inf} \rightarrow \left(\overleftarrow{\Box}(\neg\mathsf{inf} \rightarrow A) \wedge \overrightarrow{\Box}(\neg\mathsf{inf} \rightarrow A)\right)\right) \wedge$
$$\left((\mathsf{ns} \vee \mathsf{ps}) \rightarrow \left(\overleftarrow{\Box}(\mathsf{nl} \rightarrow A) \wedge \overrightarrow{\Box}(\mathsf{pl} \rightarrow A)\right)\right)\Bigg)$$

The mirror images of **K1, K2** and **K4** are also considered as axioms.

The intuitive meaning of the previous axioms is the following: **K1-K4** (and their mirror images) constitute a fragment of basic linear-time temporal logic; **c1** and **c2** state the existence and the unicity of the milestones in a frame, respectively; **c3-c7** state the ordering of these milestones. Axioms **p1** and **p2** state the existence and unicity, respectively, of proximity intervals; **p3** states that all points denoted by a proximity constant form an interval; **p4** states that every proximity constant denotes some proximity interval. **m1** states that the length of a qualitative class QC fully covers a given proximity interval. **m2** is specific to deal with INF, and states that this class is totally covered by a proximity interval (in combination with **m1**, this axiom implies that INF constitutes itself a proximity interval.) **m3-m4** enable the representation of closeness and negligibility in terms of white connectives and constants; this allows us to use, from now on, only white connectives and constants.

Rules of inference:

(MP) Modus Ponens for $\rightarrow$.
(N$\overrightarrow{\Box}$) If $\vdash A$ then $\vdash \overrightarrow{\Box}A$.
(N$\overleftarrow{\Box}$) If $\vdash A$ then $\vdash \overleftarrow{\Box}A$.

The syntactical notions of *theorem* and *proof* for $MQ^{\mathcal{P}}$ are defined as usual.

Soundness is straightforward, since it is easy to check that all the axioms are valid formulas and the inference results preserve validity.

The completeness follows by the step-by-step method, which is a Henkin-style proof, see [3]. The idea is to show that for any consistent formula A, a model for A can be built, and this is done by successive finite approximations.

Theorem 1 (Completeness). *If A is valid formula of $\mathcal{L}(MQ)^{\mathcal{P}}$, then A is a theorem of $MQ^{\mathcal{P}}$.*

Some words for Luis

- Te conocí hace muchos años y hemos coincidido pocas veces; pero la simpatía que despertaste en mí desde el primer momento no ha hecho sino crecer con los años. Hay algo que me llama la atención especialmente de tu persona, y es cómo has fusionado un sentido del humor jovial con el rigor intelectual. Espero, además, que permanezca eso en ti siempre y que podamos realizar tareas en común en los tiempos venideros, ya que la actividad que realizamos nunca se acaba, sólo se interrumpe.
  Querido Luis, quiero expresarte con estas pocas palabras mi admiración y cariño. Alfredo
- Admiro mucho tu trabajo y, aunque hemos coincidido poco en persona, he leído muchos de tus artículos En particular, tus trabajos relacionados con las cláusulas de Horn para lógicas modales han servido de inspiración para mi investigación actual sobre fragmentos sub-proposicionales para lógicas temporales de intervalos.
  En el Workshop realizado en Málaga el año pasado, pude asistir a tu charla y a tus comentarios en las charlas de los demás, incluida la mía. Me impresionó tu claridad y amplitud de ideas, así como la forma de expresarlos, proponiendo ideas nuevas y cuestiones muy interesantes.
  Espero que sigas vinculado a este mundo de la investigación y nos sigas regalando tu sabiduría. Gracias y enhorabuena. Emilio
- Coincidimos por primera vez hace casi veinticinco años (¡ya ha llovido!, incluso en Málaga). Durante una de nuestras visitas al Imperial College, en la obligada parada en el pub antes de buscar sitio para cenar, se formó un grupo en el que se hablaba, al menos, tres idiomas diferentes. ¿Quién estaba en la intersección? Por supuesto, Luis, al que yo contemplada embelesado, viendo cómo alternaba inglés, francés y español sin mayor problema, en función de la lengua de su

interlocutor. Lo curioso de tal capacidad es que, parece ser, Luis parece disponer de una versión científica, que le posibilita observar un problema desde distintos puntos de vista, de manera aparentemente simultánea, y obteniendo soluciones siempre originales y profundas.
Posteriormente, llegué a conocer algo más su faceta personal y disfrutar con su campechanía (stricto sensu) y con su especialísimo sentido del humor.
¿Qué más puedo decir? Que siempre ha sido un placer coincidir contigo en distintos eventos todos estos años y que, por supuesto, deseo que sigamos coincidiendo, más frecuentemente si cabe , ¿por qué no?, en tu etapa post-Festschrift.
Un fuerte abrazo. Manolo.

Bibliography

1. A. Ali, D. Dubois, and H. Prade. Qualitative reasoning based on fuzzy relative orders of magnitude. *IEEE transactions on Fuzzy Systems*, 11(1):9–23, 2003.
2. P. Balbiani. Reasoning about negligibility and proximity. In *Septièmes Journées d'Intelligence Artificielle Fondamentale*, pages 11–21, 2013.
3. J. P. Burgess. Basic tense logic. In *Handbook of philosophical logic*, pages 89–133. Springer, 1984.
4. A. Burrieza, E. Muñoz-Velasco, and M. Ojeda-Aciego. A multimodal logic approach to order of magnitude qualitative reasoning with comparability and negligibility relations. *Fundamenta Informaticae*, 68(1):21–46, 2005.
5. A. Burrieza, E. Muñoz-Velasco, and M. Ojeda-Aciego. Order of magnitude qualitative reasoning with bidirectional negligibility. In *Current Topics in Artificial Intelligence*, volume 4177 of *Lecture Notes in Artificial Intelligence*, pages 370–378. Springer, 2006.
6. A. Burrieza, E. Muñoz-Velasco, and M. Ojeda-Aciego. Closeness and distance in order of magnitude qualitative reasoning via PDL. In *Current Topics in Artificial Intelligence*, volume 5988 of *Lecture Notes in Artificial Intelligence*, pages 71–80. Springer, 2010.
7. D. Dubois, A. Hadj-Ali, and H. Prade. Granular computing with closeness and negligibility relations. In T. Y. Lin, Y. Y. Yao, and L. A. Zadeh, editors, *Data Mining, Rough Sets and Granular Computing*, pages 290–307. Physica-Verlag GmbH, Heidelberg, Germany, 2002.
8. J. Golińska-Pilarek and E. Muñoz-Velasco. Relational approach for a logic for order of magnitude qualitative reasoning with negligibility, non-closeness and distance. *Logic Journal of IGPL*, 17(4):375–394, 2009.
9. O. Raiman. Order of magnitude reasoning. *Artificial Intelligence*, 51:11–38, 1991.
10. L. Travé-Massuyès, F. Prats, M. Sánchez, and N. Agell. Relative and absolute order-of-magnitude models unified. *Annals of Mathematics and Artificial Intelligence*, 45:323–341, 2005.
11. L. A. Zadeh. Fuzzy sets and information granularity. In G. J. Klir and B. Yuan, editors, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*, pages 433–448. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1996.

Reasoning about Trust and Aboutness in the Context of Communication

Robert Demolombe[1]

Institut de Recherche en Informatique de Toulouse
France
robert.demolombe@orange.fr

Abstract. Trust may have many different informal definitions. In this work formal definitions are proposed in Modal Logic in order to have clear rules for reasoning about trust. We start from trust in some properties of an information source, like sincerity, competence or vigilance, about a given proposition. Then, this definition is extended to trust about all the propositions which are about a given topic (for instance the topic mathematical logic or the topic painting). A further extension is about all the propositions which inform about a given individual (for instance the individual Luis Fariñas or the individual Toledo).
Specific logics are presented for reasoning about the fact that a given proposition is about a given topic and for reasoning about the fact that a proposition informs about a given individual. At the end we give a brief extension to qualitative graded trust.

1 Introduction

Trust plays a significant role in interactions between agents. Here, we have focused on interactions in the context of communication where information sources may not be reliable. For instance, when we access information via the web, when we need to interact with a doctor in order to cure some disease or when we are looking for information about climate change in newspapers.

If we have to reason about trust, for instance when we want to combine pieces of information we have received from several information sources, we have to make clear what kind of logical definition of trust we have adopted. However, many definitions have been proposed in the literature (see [5,7,13,8,4]) and there is no consensus about "the" definition. Nevertheless, most of the definitions are based on the idea that trust is a mental state of an agent which is of the kind of belief.

In [15,14] A.J.I. Jones suggests that trust is a truster's belief about some agent's property such that the truster knows that the agent may fail to fulfil this property, while the truster also believes that in the specific situation where he is the agent will fulfil this property. We agree on this analysis but in order to avoid too complex logical definitions we have accepted that the truster's mental state is just a "standard" belief that complies with the assumption that he is not in an exceptional situation where the agent does not fulfil the property.

Then, we assume that trust is a truster's belief about some trustee's property. Now, the question is: *what kind of property?* In the context of communication these properties involve: *what is the case*, *what agents believe about what is the case* and *what information sources have transmitted.* For instance, such a property may be sincerity which is a relationship between what the trustee believes and what he has transmitted to the truster. Then, a new question is raised: *what is the structure of this relationship?* Some authors [16,4] have proposed that it is a conjunction. For instance, in the case of sincerity it would take the form: the trustee has transmitted some piece of information AND he believes that what he has transmitted is the case. In [7,10] we have claimed that it has the form of a conditional such as : IF the trustee has transmitted some piece of information, THEN he believes that what he has transmitted is the case. Indeed, the negation of sincerity takes the form: the trustee has transmitted some piece of information AND he does NOT believe that what he has transmitted is the case. Since the negation of sincerity is of the form: $\phi \land \neg\psi$, sincerity is of the form: $\phi \to \psi$.

In the following section we give a more detailed definition of the different kinds of properties where we have adopted a very specific notion of trust which is defined in terms of trust only about a given proposition. In the next section this definition is extended to trust about all the propositions that are about a given topic. Then, in the following section, it is extended to all the propositions that inform about a given individual. Finally, we extend the notion of trust to qualitative graded trust.

The general purpose of the paper is to focus on the intuitive ideas that support the definitions of the logics and not to go into technical details which have been presented in referred papers.

2 Trust about one proposition

According to the intuitive definition of trust presented in the introduction, trust has the form:
The truster i believes that (trustee j's property1 entails trustee j' property2)
where *property1* and *property2* may be one of the following properties:[1]

- proposition p is true
- trustee j believes that the proposition p is true
- trustee j has informed truster i about proposition p

We have the notations:

- p: the proposition represented by p is true
- $Bel_i(p)$: agent i believes that proposition p is true
- $Inf_{j,i}(p)$: agent j has informed agent j about proposition p

Then, the general form of trust is: $Bel_i(\phi \to \psi)$ [2].

We have the following properties that may be trusted by the truster.

Trust in Sincerity. $Bel_i(Inf_{j,i}(p) \to Bel_j(p))$
Agent i believes that if j has informed him about p, then j believes p.

Example: Romeo trusts Giullietta in her sincerity about the fact that Giulietta loves Romeo. That is: Romeo believes that if Giuletta has told him that she loves him, then Giuletta believes that she loves him.

Trust in Cooperativity. $Bel_i(Bel_j(p) \to Inf_{j,i}(p))$
Agent i believes that if j believes p, then j has informed him about p.

Example: Romeo trusts Giuletta in her cooperativity about the fact that Giuletta loves Romeo. That is: Romeo believes that if Giuletta loves him, then Giuletta has told him.

Trust in Competence. $Bel_i(Bel_j(p) \to p)$
Agent i believes that if j believes p, then p is true.

Example: Giuletta trusts Romeo about the fact FOL is not decidable. That is: Giuletta believes that if Romeo believes that FOL is not decidable, then FOL is not decidable.

Trust in Vigilance. $Bel_i(p \to Bel_j(p))$
Agent i believes that if p is true, then j believes that p is true.

Example: Giuletta trusts Romeo in his vigilance about the fact that the flight from Sevilla has landed. That is: Giuletta believes that if the flight from Sevilla has landed, then Romeo believes that the flight from Sevilla has landed.

Trust in Validity. $Bel_i(Inf_{j,i}(p) \to p)$
Agent i believes that if j has informed i about p, then p is true.

Example: Romeo trusts Giuletta in her validity about the fact that Giuletta leaves in Cadiz. That is: Romeo believes that if Giuletta has told him that she leaves in Cadiz, then it is true that she leaves in Cadiz.

Trust in Completeness. $Bel_i(p \to Inf_{j,i}(p))$
Agent i believes that if p is true, then agent j has informed him about p.

Example: Giuletta trusts Romeo in his completeness about the fact Romeo is happy. That is: Giuletta believes that if Romeo is happy, then Romeo has told her.

The language for proposition p in these definitions is any nested formula of atomic proposition with conjunction, disjunction, negation and modality Bel or Inf.

For instance, we can have: $Bel_i(Inf_{j,i}(Inf_{k,j}(p)) \to Inf_{k,j}(p))$, which means that i trusts j in his validity when j is reporting that k told him p.

The logic for modality Bel is of the kind KD. It is an open issue to accept positive or negative introspection in this context.

For the modality Inf, in addition to the inference rule of substitutivity of equivalent formulas (EQV), we have accepted the following axiom schema (CONJ) about conjunction:

(EQV) If $\vdash \phi \leftrightarrow \psi$, then $\vdash Inf_{j,i}(\phi) \leftrightarrow Inf_{j,i}(\psi)$

(CONJ) $Inf_{j,i}(\phi) \wedge Inf_{j,i}(\psi) \to Inf_{j,i}(\phi \wedge \psi)$

In this logic we can easily show that trust in validity is a logical consequence of trust in sincerity and trust in competence. However, the converse is not true. Indeed, trust in validity does not entails trust in sincerity since an agent may tell something which is true whereas he believes that it is false.

[1] Of course *property1* and *property2* must be different.

[2] As a matter of simplification conditionals are represented here by material implications.

Also it is worth noting that sincerity and competence are two independent properties. An agent may be sincere and not competent and it may be competent and not sincere.

We have similar relationships between completeness, vigilance and cooperativity.

It is assumed that there is no failure in the communication process between agents. This is formally represented by the following axiom schemas:

(OBS) $Inf_{j,i}\phi \to Bel_i(Inf_{j,i}\phi)$

(OBS') $\neg Inf_{j,i}\phi \to Bel_i(\neg Inf_{j,i}\phi)$

In terms of Speech Act Theory (see [18]) the meaning of these axioms is that locutionary acts always succeed.

Then, reasoning about trust can be illustrated with this example where the assumptions are:

(1) Romeo trusts Giuletta in her validity about the fact she is in the flight arriving from Sevilla.

(2) Romeo trusts the airport announcement in its cooperativity about the fact that the flight from Sevilla is delayed.

(3) Giuletta has inform Romeo about the fact that she is in the flight arriving from Sevilla

(4) The airport announcement has not informed Romeo about the fact that the flight from Sevilla is delayed

We have the notations:

R: Romeo, G: Giuletta, A: airport announcement

p: Giuletta is in the flight arriving from Sevilla

q: the flight from Sevilla is delayed

Then, the formal representation of the assumptions is:

(1) $Bel_R(Inf_{G,R}(p) \to p)$

(2) $Bel_R(q \to Inf_{A,R}(q))$

(3) $Inf_{G,R}(p)$

(4) $\neg Inf_{A,R}(q)$

From (3) and (OBS) we have:

(5) $Bel_R(Inf_{G,R}(p))$

From (5) and (1) we have:

(6) $Bel_R(p)$

From (4) and (OBS') we have:

(7) $Bel_R(\neg Inf_{A,R}(q))$

From (2) we have:

(8) $Bel_R(\neg Inf_{A,R}(q) \to \neg q)$

From (7) and (8) we have:

(9) $Bel_R(\neg q)$

From (6) and (9) we have:

(10) $Bel_R(p \wedge \neg q)$

The intuitive meaning of (10) is that Romeo believes that Giuletta is in the flight arriving from Sevilla and that this flight is not delayed.

3 Trust about propositions about a topic

In most of the situations the field of the propositions a truster trusts about is not restricted to one proposition. For instance, if an agent trusts Romeo for his validity about the fact that FOL is not decidable, he may also trusts him about other sentences which are about the topic of decidability. Then, even if the analysis presented in the previous section is useful to fix clear definitions and reasoning rules it is not enough to cover most of the real situations. Some authors have a stronger position and assume that if an agent trusts another agent, he trusts him for any proposition whatever is its topic. In our view this is an over simplification which is far to real situations.

In this section we have an intermediate position and we formalise the notion of trust about all the propositions which are about a given topic. That requires to answer the question: *what does it mean that a proposition is about a given topic?* A specific logic as been proposed by Demolombe and Jones in [11] to answer this question. A short presentation of this logic is given in the following.

A basic idea of this logic is that the topics a sentence is about are independent of the truth value of this sentence. For instance, it can be assumed that the sentence: *FOL is decidable* and the sentence *FOL is not decidable* are about the same topics, say the topic *decidability* and the topic *logic*.

A consequence of this position is that two logically equivalent sentences of Classical Propositional Calculs (CPC) are not necessarily about the same topics. In particular two tautologies are not necessarily about the same topics. For instance the tautology: *Cadiz is in Spain or Cadiz is not in*

Spain is about the topic geography while the tautology: *Aristoteles is dead or Aristoteles is not dead* is not about geography.

If we also accept that a sentence may be about a topic because a conjunct in this sentence is about this topic (for instance the sentence: *Cadiz is in Spain and Giuletta leaves in Cadiz* is about the topic geography because *Cadiz is in Spain* is about the topic geography), then the topics assigned to a sentence depend on the atomic sentences occurring in this sentence. A consequence is that, for instance, p and $p \wedge (q \vee \neg q)$ are not about the same topics if p and q are not about the same topics.

The formal solution to fulfil these requirements was to define a 3-valued logic with the values: *true*, *false* and *undefined*, such that if a sub-formula is *undefined* the overall formula is undefined. This logic was initially introduced by Bochvar in [2] for a different purpose (see also Buvač, Mason and Mc Carty in [3,17] and Demolombe in [6]).

For the semantics we consider models defined by three sets:

- τ: a set of topics
- S: a set of sentences
- W: a set of worlds

The language is the language of CPC. A topic name is denoted by t and a sentence name is denoted by $'p'$.

In a model M a function I assigns to a topic name a topic in τ, a function J assigns to a sentence name a sentence in S, a function T assigns to a sentence in S a set of worlds in W and a function F assigns to a sentence in S a set of worlds in W. Finally, a function N assigns to a pair of sets of worlds in W a set of topics in τ.

Therefore, $J('p')$ is the sentence the name of which is $'p'$, $T(J('p'))$ is a set of worlds where the sentence $J('p')$ is true and $F(J('p'))$ is a set of worlds where the sentence $J('p')$ is false. Since we have a 3-valued logic the union of the sets of worlds $T(J('p'))$ and $F(J('p'))$ may not be equal to W. The function N assigns to a pair of worlds $< T(J('p')), F(J('p')) >$ a set of topics in τ.

Now, a predicate $A(t,' p')$ is defined in a First Order Language where t is a topic name and $'p'$ is a sentence name, and the intuitive meaning of this predicate is that the sentence named by $'p'$ is about the topic named by t. The satisfiability condition for this predicate in a world w of a model M is:

$M, w \models A(t,' p')$ iff $I(t) \in N(T(J('p')), F(J('p')))$

As usual a valid formula ϕ is denoted by $\models \phi$.

In this semantics we have the rule:

If $\models \phi \leftrightarrow \psi$ and ϕ and ψ **contain the same atoms**, then $\models A(t,' \phi') \leftrightarrow A(t,' \psi')$

The condition "contain the same atoms" guarantees, in particular, that two tautologies which does not contain the same atoms are not necessarily about the same topic.

The following schema can be accepted even if is not valid.

$\models A(t,' \phi') \rightarrow A(t,' \neg\phi')$

It is worth noting that we do **not** have:

$\models A(t,' \phi \wedge \psi') \rightarrow (A(t,' \phi') \vee A(t,' \psi'))$

For instance if the sentence p is *Romeo is married with Giuletta* and s is *Romeo is married with Venus* the sentence $p \wedge q$ is about the topic *bigamy* while neither p nor q are about *bigamy*.

The predicate $A(t,' p')$ can be used to extend the notion of trust to trust about a topic as follows:[3]

$TrustT(i, j, t) \stackrel{\text{def}}{=}$ for every $'\phi'$: if $A(t,' \phi')$, then $Trust(i, j, \phi)$

This is just a notation in the meta language where $Trust(i, j, \phi)$ may be used to denote trust in sincerity, competence or any similar property.

For instance we may have: $TrustT(Giulietta, Romeo, modal\ \ logic)$.

4 Trust about the propositions that inform about an individual

There are many contexts where an agent is interested in getting information about a given individual. This raises the question: *what does it mean that a proposition informs about an individual?*

The answer to this question is not that this proposition is *about* this individual. For instance, we can agree on the fact that the proposition *the painter Sorolla is born in Valencia or the painter Sorolla is not born in Valencia* is about the topic Sorolla, however it does not inform about Sorolla in the sense that even if we know that this proposition is true we ignore wether this painter is born

[3] As a matter of simplification the specific property the trustee j is trusted in is not made explicit here.

in Valencia. At the opposite the proposition *the painter Sorolla is born in Valencia* informs about the city where Sorolla is born.

We think that this example shows that although the word "Sorolla" can be used as well to denote a topic and to denote an individual, the properties: proposition p informs about the individual Sorolla and proposition p is about the topic Sorolla have completely different meanings. Also, it is worth noting that the property of informing about an individual cannot be trivially derived from the fact that the name of this individual occurs in a proposition. It is clear that this is not true in the case of a tautology. In addition, it may be that a proposition informs about an individual although its name does not occur in this proposition. For instance, if it is assumed that Sorolla is a man, the proposition: *every man is mortal* informs about Sorolla.

The purpose of this section is to present a clear answer to the above question on the basis of the work by Demolombe and Fariñas del Cerro in [12].[4]

The first step is to define a property that characterizes the propositions that do **not** inform about an individual (because for technical reason this property is easier to define). This is done in the context of a Classical First Order Logic without function symbols and without equality. The semantics is defined as usual by an interpretation domain D and by an interpretation function which assigns to each constant name and to each variable name an element in D.

Then, we define a variant M' of a given model M with respect to a constant symbol c. The intuitive idea is that propositions represented by formulas which do not inform about c are formulas such that their truth value does change in the models M' where the predicates preserve the same truth value as in M for the tuples that do not contain an interpretation of constant c.

For instance, the formula *Giulietta loves Romeo* does not inform about Venus because the truth value of this formula remains stable if in a model M' we only change the truth value of atomic propositions the tuples of which contain the interpretation of Venus. The only exception is if in M Venus is interpreted by the same individual as Giulietta. In that case we cannot say that in M *Giulietta loves Romeo* does not inform about Venus.

The formal translation of these ideas leads to the rather complex following definition of the variants M' of M with regard to c.

We call variants of M with regard to c the set M^c of interpretations M' defined from M in the following way.

Let $i_{M'}$ and i_M be the interpretation functions in these models. We have

- $D_{M'} = D_M$
- $i_{M'} = i_M$ for every variable symbol and constant symbol,
- $i_{M'}$ is defined from i_M for each predicate symbol as follows: if p is a predicate symbol of arity n
  - if t is a n-tuple of terms of the language that contains no occurrence of the constant symbol c, then $i_{M'}(t) \in i_{M'}(p)$ iff $i_M(t) \in i_M(p)$,
  - if an element $< d_1, \dots, d_n >$ of D^n is such that, for every j in [1,n], $d_j \neq i_M(c)$, then $< d_1, \dots, d_n > \in i_{M'}(p)$ iff $< d_1, \dots, d_n > \in i_M(p)$.

According to the definition of this property, a formula p does not inform about c iff its truth value remains stable in all the variants M' of M in M^c. At the opposite a formula that informs about c is a formula that does not fulfil this property.

If we denote by $AI(F, c)$ the property: formula F is about the individual c we have:

$AI(F, c)\ holds\ \ iff\ \ \exists M(\exists M' \in M^c(M \models F\ and\ M' \not\models F))$

This property can be used to extend the definition of trust to trust about all the propositions which inform about the individual denoted by c in that way:[5]

$TrustI(i, j, c) \stackrel{\text{def}}{=}$ for every closed formula F in a first order language: if $AI(F, c)$, then $Trust(i, j, F)$

For instance we may have: $TrustI(Romeo, Giullietta, Sorolla)$ or $TrustI(Romeo, Giullietta, Cadiz)$.

5 Graded trust

In the previous sections we have assumed that the truster has a strong position with regard to the trustee in the sense that either he trusts him or he does not trust him. However, there are many

[4] In this work the term "individual" has the same meaning as in formal logic, it may refer to any kind of entity, not just to a person.

[5] As a matter of simplification the specific property the trustee j is trusted in is not made explicit here.

situations where the truster has a more refined positions. For instance, he may have several possible levels of trust: it may be strong, medium or low. This raises the new question: *what does it mean that a truster trusts strongly, or weakly, a trustee?*

To answer this question we can take the example of trust in sincerity which is formally represented in general by the formula: $Bel_i(Inf_{j,i}(p) \rightarrow Bel_j(p))$.

A first answer is that the strength of i's belief is high, or is low. If the strength level is a qualitative level g, that can be represented by:

$Bel_i^g(Inf_{j,i}(p) \rightarrow Bel_j(p))$

Another answer is that i strongly trusts j in his sincerity about p means that i believes that, in most of the situations, if j informs i about p, then j believes that p is true. This is a quite different interpretation where the trust level refers to the regularity level of the relationship between $Inf_{j,i}(p)$ and $Bel_j(p)$. That can be represented by: $Bel_i(Inf_{j,i}(p) \Rightarrow^h Bel_j(p))$, where $\Rightarrow^h$ denotes a conditional operator.

Our position is that in general graded trust is a composition of both answers: the strength level of a belief and the regularity level of what is believed. That can be formally represented by:

$$Bel_i^g(Inf_{j,i}(p) \Rightarrow^h Bel_j(p))$$

The two levels are independent since we may have any of the following situations:

- i strongly believes that j is strongly sincere
- i weakly believes that j is strongly sincere
- i strongly believes that j is weakly sincere
- i weakly believes that j is weakly sincere

To make more precise the meaning of the levels g and h we have defined in [8] a logic for reasoning about these qualitative levels. This logic has been slightly improved in [1].

The main idea is that the strength levels of a belief about ϕ and about $\neg\phi$ are not always connected. For instance, if i is almost ignorant about ϕ, his belief levels about ϕ and about $\neg\phi$ are both low although if his belief level about ϕ is high, then his belief level about $\neg\phi$ cannot be high.

At the opposite for the regularity levels there is a relationship between the regularity level of sincerity and unsincerity. That is, if it assumed that in most cases if j informs i about p, then j believes p, it cannot be assumed that in most cases if j informs i about p, then j does not believe p. That is clear when the regularity levels are quantitatively represented by conditional probabilities.

In informal terms the most significant properties of this logic are defined as follows.

Logically equivalent formulas can be substituted in all the rules and axiom schemas.

Levels g and h are unique for a given formula.

(Weak) If $\vdash \phi \rightarrow \psi$, $Bel_i^g(\phi)$ and $Bel_i^{g'}(\psi)$, then we do not have $g' < g$

(ClosDisj) If $Bel_i^{g_1}(\phi_1)$, $Bel_i^{g_2}(\phi_2)$ and $g_3 = Max\{g_1, g_2\}$, then $Bel_i^{g_3}(\phi_1 \vee \phi_2)$

(ClosConj) If $Bel_i^{g_1}(\phi_1)$, $Bel_i^{g_2}(\phi_2)$ and $g_3 = Min\{g_1, g_2\}$, then $Bel_i^{g_3}(\phi_1 \wedge \phi_2)$

(MaxTau) If ϕ is a tautology, then $\vdash Bel_i^{max}\phi$[6]

(Detach) $\vdash (\phi \Rightarrow^h \psi) \rightarrow (\phi \rightarrow \psi^h)$

(Trans) If $n = Max\{Min\{h_1, k_1\}, Min\{h_2, k_2\}\}$, then

$\vdash ((\phi \Rightarrow^{h_1} \psi) \wedge (\phi \wedge \psi \Rightarrow^{k_1} \theta) \wedge (\phi \Rightarrow^{h_2} \neg\psi) \wedge (\phi \wedge \neg\psi \Rightarrow^{k_2} \theta)) \rightarrow (\phi \Rightarrow^n \theta)$

The rule (Trans), even if it is rather complex, is justified by the fact that if h is a numerical level and regularities are conditional probabilities, this rule perfectly fits the detachment rule of conditional probabilities.

6 Conclusions

We have started with a very specific definition of trust which makes explicit the property which is trusted in and the proposition trust is about. Then, we have extended the definition to trust about all the propositions about a given topic and to all the propositions that inform about an individual. The first extension has needed the formalization of a logic for reasoning about aboutness which is a property which is independent of the truth values of the propositions. The second one is based on Classical FOL but requires the definitions of rather complex properties. At the end we have seen how we can formalize a more flexible notion of trust than to trust or not to trust. In our proposal we show that we need different reasoning rules for reasoning about external facts (regularity of a property) and internal facts (strength of a belief).

[6] It is worth noting that $Bel_i^{min}\phi$ does not entail that ϕ is inconsistent.

Trust raises a lot of non trivial questions. For instance, in a first approach it seems clear that trust is transitive in the sense that if the information source s_1 trusts the information source s_2 and s_2 trusts the informations source s_3, then s_1 trusts s_3. However, we have shown in [9] that transitivity holds only if some strong conditions are fulfilled.

In the presented work we have concentrated on the logics for deriving consequences from assumptions about trust in information source. Another kind of problem is to analyze how trust is supported. In [1] it has been shown that the concepts of trust and argumentation that support trust are imbricated. For instance, to trust information source s_1 we may receive arguments from another information source s_2 and these arguments are convincing only if we already trust s_2.

The final conclusion is that there is still a lot of issues to investigate in the field of trust.

Acknowledgements. We would like to thank Luis Fariñas del Cerro for his fruitful cooperation for many years. At the beginning of the eighties he introduces me to Modal Logic which I had the opportunity to apply in many areas like Data Bases or Multi Agent Systems. In our common works I had the opportunity to benefit from his capacity to find very fast what are the basic issues and to suggest original solutions. In addition he was a very good friend.

Bibliography

1. L. Amgoud and R. Demolombe. An argumentation-based approach for reasoning about trust in information sources. *Argument and Computation*, 5(2), 2014.
2. D.A. Bochvar. Two papers on partial predicate calculus. Technical report STAN-CS-280-72, Stanford University, 1972. Translation of two papers from 1938 and 1943.
3. S. Buvač, V. Buvač, and I.A. Mason. Metamathematics of contexts. *Fundamenta Informaticae*, 23(3), 1995.
4. C. Castelfranchi and R. Falcone. *Trust Theory: A Socio-Cognitive and Computational Model.* Wiley, 2010.
5. C. Castelfranchi and Y-H. Tan. *Deception, Fraud and Trust in Agent Societies.* Kluwer, 2001.
6. R. Demolombe. Multivalued logics and topics. In *Proc. of IJCAI workshop on Practical Reasoning and Rationality, Stockolm*, 1999.
7. R. Demolombe. To trust information sources: a proposal for a modal logical framework. In C. Castelfranchi and Y-H. Tan, editor, *Trust and Deception in Virtual Societies.* Kluwer, 1999.
8. R. Demolombe. Graded Trust. In R. Falcone and S. Barber and J. Sabater-Mir and M.Singh, editor, *Proceedings of the Trust in Agent Societies Workshop at AAMAS 2009*, 2009.
9. R. Demolombe. Transitivity and propagation of trust in information sources. An Analysis in Modal Logic. In J. Leite and P. Torroni and T. Agotnes and L. van der Torre, editor, *Computational Logic in Multi-Agent Systems, LNAI 6814.* Springer Verlag, 2011.
10. R. Demolombe. Analytical Decomposition of Trust in Terms of Mental and Social Attitudes. In A. Herzig and E. Lorini, editors, *The Cognitive Foundations of Group Attitudes and Social Interaction.* Springer, 2015.
11. R. Demolombe and A.J.I. Jones. Reasoning about Topics: towards a formal theory. In *American Association for Artificial Intelligence Fall Symposium*, 1995.
12. R. Demolombe and L. Fariñas del Cerro. Towards a logical characterisation of sentences of the kind "sentence p is about object c". In S. Holldobler, editor, *Intellectics and Computational Logic. Papers in Honor of Wolfang Bibel.* Kluwer Academic Press, 2000.
13. R. Falcone and C. Castelfranchi. Trust dynamics: How trust is influenced by direct experiences and by trust itself. In *Proceedings of the 3rd International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-04)*, pages 740–747. New York, ACM, 2004.
14. A.J.I. Jones. On the Attitude of Trust: a Formal Characterization of Trust, Distrust and Associated Notions. In A. Herzig and E. Lorini, editors, *The Cognitive Foundations of Group Attitudes and Social Interaction.* Springer, 2015.
15. A.J.I. Jones, A. Artikis, and J. Pitt. The design of intelligent socio-technical systems. *Artificial Intelligence Review*, 39(1), 2013.
16. E. Lorini and R. Demolombe. From trust in information sources to trust in communication systems: an analysis in modal logic. In J. Broersen and J.-J. Meyer, editors, *International Workshop on Knowledge Representation for Agents and Multi-agent Systems (KRAMAS 2008): Postproceedings*, LNAI. Springer-Verlag, 2009.
17. J. McCarthy and S. Buvač. Formalizing contexts. Technical Report STAN-CS-TN-94-13, Stanford University, 1994.
18. J. R. Searle. *Speech Acts: An essay in the philosophy of language.* Cambridge University Press, New-York, 1969.

Foundations for a Logic of Arguments

Leila Amgoud[1], Philippe Besnard[1], and Anthony Hunter[2]

[1] CNRS, IRIT, Université de Toulouse, Toulouse, France
[2] University College London, London, U.K.
amgoud@irit.fr besnard@irit.fr anthony.hunter@ucl.ac.uk

Abstract. This paper lays the foundations of a *logic of argumentation* in which arguments, as well as attacks and supports among arguments are all defined in a unifying formalism. In the latter, an argument is denoted as a pair displaying a *reason* and a *conclusion* but no condition is required to hold relating the reason to the conclusion. We introduce a series of inference rules relating arguments and show how the resulting logic captures important features of argumentation that hitherto have not been captured by existing formalisms. [3]

1 INTRODUCTION

Argumentation is a common activity in everyday life. Indeed, people frequently justify opinions, decisions or actions by *arguments* in order to increase (or to decrease) their acceptability for an audience. Arguments are therefore of great importance. In order to build argumentation systems that are able to capture natural language arguments, it is fundamental to have a clear understanding and a fair representation of this key notion of argument. For that purpose, the following issues need to be investigated:

- what is an argument?
- what may be the conclusion of an argument?
- what may be the premises of an argument?
- what is the nature of the link between premises and conclusion of an argument?

In the AI literature on argumentation, an argument is viewed as identifying a *reason* for concluding some statement. The reason is a set of premises that somehow lead to the conclusion. Hopefully, linguists and philosophers working on argumentation agree with such an idea. In [3], the linguist Apothéloz argued that the reason is *oriented in favour of* the conclusion to which it *propagates its truth*.

The views of the three communities may differ as to the kind of conclusions that can be justified by arguments and the kind of premises to be used in arguments. In the AI literature, arguments are built in favour of a statement or in favour of its contrary. That is, an agent may argue that a statement holds or that the opposing statement holds. However, it is not possible to build an argument in favour of *not concluding* a statement (this does *not* mean that the statement is false). Yet such arguments are common in natural language argumentation as shown by the next example adapted from [3].

Adam: *Steve is very smart but didn't work hard this term, so it's unclear whether he will pass his exams.*

Through the above argument, Adam *does not commit* to the conclusion "Steve will fail his exams". He simply means that "Steve will pass his exams" cannot be concluded. Notice that Dung's style argumentation [4] cannot capture such a stand-alone argument: In abstract argumentation, failure can only be expressed by means of an attack from an argument over another argument but there is no such attack in the example, as there is a single argument, and it does *not* attack itself.

Adam does not commit either to the opposite conclusion "Steve will pass his exams". His argument expresses that, in the case of a smart person who did not work hard, it is not possible to predict whether he will fail or pass his exams.

More generally, an argument may justify why a statement:

- is true,
- is false,
- is open to doubt.

[3] This paper reviews and extends two previous papers by the authors on this topic [1,2]

The last case actually encompasses two situations: the situation where the opposing statement holds and the situation of *complete indeterminacy* regarding the statement. Importantly, the notion of a statement here is very general as opposed to most proposals for formalizing argumentation, where arguments have a simple format whereby arguments cannot appear in either the premises or conclusions though notable exceptions include [5,6,7]. Apothéloz offers in [3] natural language arguments whose premises (resp., conclusion) may be arguments. Consider the following argument:

> *The fact that Ryan's car is in the car park is not a reason to conclude that Ryan is in his office. Indeed, his car is broken.*

This example displays an argument embedded in another. The first one says that "Ryan is in his office since his car is in the car park". The second argument concludes that the first fails. The premise used for that purpose is: "Ryan's car is broken". Thus, the argument has a simple premise but its conclusion is a *rejection*.

In short, a rejection is a denial of an argument: While the argument being denied offers a reason for a statement, a rejection of the argument expresses that the reason brought forward cannot serve to conclude the statement. Here is an illustration:

Brian: *Steve will fail his exams. It is raining.*
Craig: *Rain is not a reason to infer that Steve will fail his exams.*

Craig's utterance is a rejection of Brian's argument but is not an argument as it merely expresses –without providing any justification– that Brian's argument should be rejected. Though, a rejection need not make any claims whether the conclusion of the denied argument holds. It can actually be the case that rejection acknowledges truth of the conclusion of the denied argument. For example, such is the case should Craig add:

Craig: *Lack of motivation is the reason.*

That is, Steve will indeed fail his exams according to Craig. Thus, Craig agrees with the conclusion of Brian's argument. Still, Craig disagrees with Brian's argument.

Existing models in computational argumentation must resort to an encoding of Brian's argument in order to capture the fact that Craig's utterance is a rejection of Brian's. Moreover, please notice that Brian's argument is certainly *not self-attacking*.

On a general level, in Dung's style argumentation systems [4], arguments may attack each other. Attacks are viewed as reasons for rejecting arguments (in the sense that acceptability of the attacked arguments fails). In argumentation in natural language, such a phenomenon may be expressed by meta-arguments. Indeed, it is possible to offer an argument in favour of rejecting another argument.

As regards premises, a meaningful distinction is as follows:

- *Factual* reasons: The premises are taken as granted. An example is:
  It has been raining all morning, the outdoor tennis tournament this afternoon will be cancelled.
- *Hypothetical* reasons: The premises are not meant to be endorsed. An example is:
  An economic crisis in Germany would be a reason for a declining value of the euro.

Although there seems to be currently no system having special machinery for dealing with hypotheticals, a number of them (e.g., [8,9,10,11]) can deal with hypothetical reasoning by adding hypothetical assumptions to the knowledgebase.

A key feature of an argument is the link between the reason and the conclusion. In existing argumentation systems, the link is *deductive* i.e., the conclusion follows from the reason as an inference in a logic. However, several other kinds of links may be encountered in natural language arguments including *causal*, *analogical*, and others. Here is an illustration:

> *My new phone is the same brand as my former phone. To redial a number, I should probably use the same procedure as with my former phone.*

In the above argument, a reason is given "my new phone is the same brand as my former phone" so that, in order to redial, I should try the same routine as I was used to. In the formalism to be developed in this paper, such an argument is represented with its reason and conclusion, but there is no need to resort to extra premises (presumably rather convoluted) required in a formal derivation of the conclusion.

As to the four questions about the notion of an argument that were listed at the start of the introduction, we can propose the following features to be key to an argument:

- An argument provides a presumptive explanation as to why a statement holds, or why the statement does not hold, or why the statement is doubted.

- A statement is the conclusion of the argument, and can be a simple proposition or a complex one like an argument (or a rejection) or a combination, including nesting, of arguments.
- An argument may resort to factual premises or to hypothetical premises. They play the role of an explanation, and, similarly to the conclusion, range from a simple proposition to a complex one like an argument (or a rejection) or a combination, including nesting, of arguments.
- The nature of the link (relating the premises to the conclusion) can be any among a range of possibilities (e.g. deductive, causal, inductive, analogical, etc).

Taking advantage of the proposal made by Apothéloz in [3], we introduce a logical setting for *representing* and *reasoning* about arguments that enjoy all the features discussed above. We first give a formal definition of argument and rejection of argument. This gives the language $\mathbb{L}_{RC}$ of our logic (in Section 2). We provide a set of *inference rules* that show how arguments (respectively rejections of arguments) are tied together, in the sense that an agent presenting an argument thereby commits himself to other arguments (so an inference rule $\frac{\alpha}{\beta}$ means that holding argument α entails committing to argument β). That is, the inference rules provide us with a notion of equivalence between arguments, they enable us to express that δ is a counter-argument to α, and so on. This gives us the inference system $\Vdash$ of our logic (Section 3). It is essential to notice that $\Vdash$ has nothing to do with evaluation of arguments and does not deal with acceptability of arguments (as is done by acceptability semantics [4]). Instead, $\Vdash$ expresses what one should expect when committing to a given argument. We investigate some properties of the logic and illustrate how *attacks* and *supports* between arguments are expressed as arguments (Section 4). As a result of its high level of expressiveness, the new logic captures important features of argumentation that hitherto have not been captured by existing formalisms. Furthermore, it lays the foundations of a fully fledged *argumentation logic*.

2 FORMAL SYNTAX

We present a formalism to represent arguments, inspired by Apothéloz [3]. It is built upon a classical propositional language $\mathbb{L}$ with the classical connectives $\neg, \vee, \wedge, \rightarrow, \leftrightarrow$. The formalism also uses the symbols $\mathcal{R}$ and $\mathcal{C}$, and additional operators, namely $-, |, \&$ (not, or, and), applying to arguments. Thus, two negation operators are needed: $\neg$ for denying propositional formulas ($\neg x$ denotes that x is false), and $-$ for denying $\mathcal{R}(.)$ and $\mathcal{C}(.)$. Please note that $\neg\neg x$ is identified with x and $- - \mathcal{R}(.)$ is identified with $\mathcal{R}(.)$ (similarly, $- - \mathcal{C}(.)$ is identified with $\mathcal{C}(.)$).

An *argument* gives a reason for concluding a statement. It has two parts: its *premises* (or its reason) and its *conclusion*, following several models (most significantly, [12]) in computational argumentation. An argument is interpreted as follows: its conclusion holds *because* it follows, according to a given notion, from the premises. The notion refers to the nature of the link (for instance, the premises cause the conclusion). Also, a *rejection* is a statement denying an argument. The premises and conclusion occurring in the rejection are those of the denied argument. The difference is that there is "$-$" in front of the (leftmost occurrence of the) $\mathcal{R}$ symbol.

In our formalism, arguments and rejections thereof form the class of *RC-formulas*, denoted $\mathbb{L}_{RC}$.

Definition 1 (RC-formulas). *An* RC-formula *is of the form*

$$(-)\mathcal{R}(y) : (-)\mathcal{C}(x)$$

where x, y are RC-terms, the set of which is defined as the smallest set such that
– a formula of $\mathbb{L}$ is an RC-term,
– an RC-formula is an RC-term,
– if α and β are RC-terms then so are $-\beta$, $\alpha \,|\, \beta$, $\alpha \,\&\, \beta$.

The notation "$(-)$" means that the negation operator "$-$" may, but need not, occur. $\mathcal{R}$ and $\mathcal{C}$ are indicative of the functions of *giving reason* and *concluding*, respectively. Thus, they capture the coupling reason-conclusion. As we will see later, the contents may be true while the functions do not hold and vice versa. Whatever the link between the reason and the conclusion, it is represented by the colon in the definition.

The two symbols $|$ and $\&$ can be used to obtain RC-formulas in a number of ways, examples of RC-formulas include $\mathcal{R}\left(-\mathcal{R}(y) : \mathcal{C}(x)\right) : \mathcal{C}(w)$, $\mathcal{R}(z) : -\mathcal{C}\left(w \,\&\, \left(-\mathcal{R}(y) : \mathcal{C}(x)\right)\right), \ldots$ Contrariwise, examples of expressions that fail to be RC-formulas include $x \,\&\, \mathcal{R}(y) : \mathcal{C}(z)$, $x \mid \mathcal{R}(y) : \mathcal{C}(z) \ldots$

Unlike existing definitions of argument where a conclusion x follows from premises y using a notion of derivation (e.g., [10]), Definition 1 leaves the content of the link unspecified. Accordingly, such a general definition makes it possible to capture links of whatever nature, including non-deductive links, and therefore can offer a way to represent any natural language argument, even somewhat dubious arguments such as:

This paper will be accepted. It's about argumentation.

Taking pa to stand for "this paper will be accepted" and aa to stand for "this paper is about argumentation", $\mathcal{R}(aa) : \mathcal{C}(pa)$ is indeed a representation of the above argument.

Definition 2 (Argument). *An* argument *is an RC-formula of the form*

$$\mathcal{R}(y) : (-)\mathcal{C}(x).$$

The intuitive meaning of the two formal expressions captured by Definition 2 is:

$$\begin{array}{rll} \mathcal{R}(y) : \mathcal{C}(x) & \text{means that} & \text{“}y \text{ is a reason for concluding } x\text{”.} \\ \mathcal{R}(y) : -\mathcal{C}(x) & \text{means that} & \text{“}y \text{ is a reason for not concluding } x\text{”.} \end{array}$$

Example 1. Let sm stand for "Steve is very smart", and wh stand for "Steve worked hard this term", and pe stand for "Steve will pass his exams". Then, Adam's argument "Steve is very smart but didn't work hard this term, so it's unclear whether he will pass his exams" can be captured by the RC-formula

$$\mathcal{R}(sm \wedge \neg wh) : -\mathcal{C}(pe).$$

Let moreover r stand for "it is raining" and lm stand for "Steve is lacking motivation". Reconstructing Craig's "Lack of motivation is *the* reason" to account for concluding that Steve will fail his exams as well as denying rain to account for it, can then be captured by the RC-formula

$$\mathcal{R}(lm) : \mathcal{C}\left(\neg pe \;\&\; -\mathcal{R}(r) : \mathcal{C}(\neg pe)\right).$$

Accordingly, taking x to be a propositional formula, all this faithfully accounts for the distinctions mentioned in the introduction:

- Arguments *in favour* of x, they are of the form $\mathcal{R}(y) : \mathcal{C}(x)$.
- Arguments *against* x, they are of the form $\mathcal{R}(y) : \mathcal{C}(\neg x)$.
- Arguments justifying why x is *doubted*, they are of the form $\mathcal{R}(y) : -\mathcal{C}(x)$.

Please observe that the second item amounts to arguing in favour of $\neg x$ whereas the third item has a sister item, of the form $\mathcal{R}(y) : -\mathcal{C}(\neg x)$, justifying why $\neg x$ is doubted. The case of complete indeterminacy (i.e., when both x and $\neg x$ are doubted) can be identified with both sister items taken together More generally, $\mathcal{R}(y) : -\mathcal{C}(x)$ is right in two kinds of situations: (1) y is a reason for concluding $\neg x$; *e.g.*, being a penguin is not only a reason for not concluding that Tweety can fly, $\mathcal{R}(p) : -\mathcal{C}(f)$, it furthermore is a reason for concluding that Tweety cannot fly, $\mathcal{R}(p) : \mathcal{C}(\neg f)$ and (2) y does not support $\neg x$ but still does not support x either, e.g., being a penguin is neither a reason to conclude that Tweety is a young bird nor a reason to conclude that Tweety is not a young bird.

Definition 3 (Rejection). *A* rejection *is an RC-formula of the form*

$$-\mathcal{R}(y) : (-)\mathcal{C}(x).$$

The intuitive meaning for these formal expressions is as follows:

$$\begin{array}{c} -\mathcal{R}(y) : \mathcal{C}(x) \text{ means that “}y \text{ is not a reason for concluding } x\text{”.} \\ -\mathcal{R}(y) : -\mathcal{C}(x) \text{ means that “}y \text{ is not a reason for not concluding } x\text{”.} \end{array}$$

Example 2. Craig's "Rain is not a reason to infer that Steve will fail his exams" can be captured by the RC-formula

$$-\mathcal{R}(r) : \mathcal{C}(\neg pe).$$

As an argument exhibits a reason, a conclusion and a link over them, an argument can be objected by challenging its reason, or its conclusion, or its link. These three possibilities of objecting to an argument $\mathcal{R}(y) : \mathcal{C}(x)$ are rendered by RC-formulas. Assume, for instance, that $x, y \in \mathbb{L}$:

- The *reason* of the argument *is objected to*, which is achieved by an argument of the form $\mathcal{R}(z) : \mathcal{C}(\neg y)$.
- The *conclusion* of the argument *is objected to*, which is achieved by an argument of the form $\mathcal{R}(z) : \mathcal{C}(\neg x)$.
- The *link* in the argument *is objected to*, which is achieved by a rejection of the form $-\mathcal{R}(y) : \mathcal{C}(x)$.

The link can also be objected to by means of more informed items such as arguments of the form $\mathcal{R}(z) : \mathcal{C}\left(-\mathcal{R}(y) : \mathcal{C}(x)\right)$.

Example 3. Again, on whether Ryan is in his office.

- ro stand for "Ryan is in his office",
- cp stand for "Ryan's car is in the car park",
- bc stand for "Ryan's car is broken".

Then,

Dale: *Ryan is in his office. His car is in the car park.*
Earl: *The car is in the car park because it is broken.*

can be formalized as $\mathcal{R}(cp) : \mathcal{C}(ro)$ for Dale's argument and $\mathcal{R}(bc) : \mathcal{C}(cp)$ for Earl's strict utterance. As a response (or objection to Dale's), Earl's argument can be reconstructed as $\mathcal{R}(\mathcal{R}(bc) : \mathcal{C}(cp)) : -\mathcal{C}(ro)$.

3 INFERENCE

The aim of this section is to introduce the consequence operator $\Vdash$ and some of its properties, where $\Vdash$ is the least closure of a set of *inference rules* extended with one *meta-rule*. We investigate a specific combination of inference rules in this section. We have considered alternative combinations of inference rules previously [1,2].

Of course, w, x, y, z below can be instantiated with RC-terms. These are supposed to obey Boolean identities over $-$ (negation), $|$ (disjunction) and $\&$ (conjunction) such that $- - x = x$, $-(x \,\&\, y) = -x \,|\, -y$, and so on. Also, $-$, $|$ and $\&$ must be understood as $\neg$, $\vee$ and $\wedge$ resp., when applying to RC-terms that are formulas of $\mathbb{L}$.

Importantly, deriving an argument α by means of inference rules does *not* mean that α is accepted. Instead, inferring α means that the argument(s) and/or rejection(s) used as premises for the inference rule(s) applied while deriving α cannot be held without α also being held. Indeed, $\Vdash$ is meant to capture *commitment* between arguments. Hence, if a foolish argument is used as a premise then a foolish α may result: If an agent holds a foolish argument, he henceforth commits to some other foolish arguments.

3.1 Denial of an argument

$\Vdash$ is defined with the requirement that $-(\mathcal{R}(y) : \Phi)$ is identified with $-\mathcal{R}(y) : \Phi$, and similarly for $-(-\mathcal{R}(y) : \Phi)$ with $- - \mathcal{R}(y) : \Phi$. In doing so, we are faithful to Apothéloz who regards them as equivalent [3]. It seems disputable, though. It could be argued that $-\mathcal{R}(y) : \Phi$ disqualifies only *one part* of an argument (i.e., its reason) while $-(\mathcal{R}(y) : \Phi)$ somehow disqualifies the whole argument. Imagine a member of a recruitment committee who presents an argument in favour of a candidate for his own research lab. The argument may be denied by the committee due to conflict of interest. However, such a denial does not (or at least need not) challenge truth of the reason nor its ability to bring about the conclusion of the argument.

3.2 Meta-rule

Rejection $-\mathcal{R}(y) : \mathcal{C}(x)$ means that y is not a reason for x, which is the negation of what $\mathcal{R}(y) : \mathcal{C}(x)$ is supposed to mean, i.e., y is a reason for x. As a consequence, the contrapositive of the fact that $\mathcal{R}(y) : \mathcal{C}(x)$ would entail $-\mathcal{R}(y) : \mathcal{C}(w)$ is that $\mathcal{R}(y) : \mathcal{C}(w)$ would entail $-\mathcal{R}(y) : \mathcal{C}(x)$. Accordingly, the *meta-rule* expresses that we can reverse any inference rule of the form

$$\frac{\mathcal{R}(y) : \Phi}{-\mathcal{R}(y) : \Psi} \qquad \text{into} \qquad \frac{\mathcal{R}(y) : \Psi}{-\mathcal{R}(y) : \Phi}.$$

Of course, the same reversing process takes place whenever "$-$" occurs in front of the leftmost "$\mathcal{R}$" so that, in the general case, an inference rule [4] where $i, j \in \{0, 1\}$

$$\frac{-^{(i)}\mathcal{R}(y) : \Phi \quad \alpha_1 \cdots \alpha_n}{-^{(j)}\mathcal{R}(y) : \Psi} \quad \text{can be reversed into} \quad \frac{-^{(1-j)}\mathcal{R}(y) : \Psi \quad \alpha_1 \cdots \alpha_n}{-^{(1-i)}\mathcal{R}(y) : \Phi}$$

whatever the RC-formulas $\alpha_1, \ldots, \alpha_n$.

[4] $-^{(1)}$ denotes a single occurrence of the hyphen and $-^{(0)}$ the absence of it.

3.3 Inference rules

Certainly, the feature most expected is consistency in terms of arguments:

$$\frac{\mathcal{R}(y) : \mathcal{C}(x)}{-\mathcal{R}(y) : -\mathcal{C}(x)} \qquad \frac{\mathcal{R}(y) : \mathcal{C}(x)}{\mathcal{R}(y) : -\mathcal{C}(-x)} \qquad \textbf{(Consistency)}$$

The leftmost inference rule means that if y is a reason for x then y is not a reason to doubt x. The rightmost inference rule means that if y is a reason for x then it is also a reason to doubt $-x$.

Property 1. The inference rules below derive from (Consistency) and the meta-rule.

$$\frac{\mathcal{R}(y) : \mathcal{C}(x)}{-\mathcal{R}(y) : \mathcal{C}(-x)} \quad \frac{\mathcal{R}(y) : -\mathcal{C}(x)}{-\mathcal{R}(y) : \mathcal{C}(x)} \quad \frac{\mathcal{R}(y) : \mathcal{C}(-x)}{\mathcal{R}(y) : -\mathcal{C}(x)} \quad \frac{\mathcal{R}(y) : \mathcal{C}(-x)}{-\mathcal{R}(y) : \mathcal{C}(x)}$$

Please observe that an instance of the third rule in Property 1 is:

$$\frac{\mathcal{R}(z) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))}{\mathcal{R}(z) : -\mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))}$$

A similar rule, related to $\mathcal{R}$ instead of $\mathcal{C}$, is:

$$\frac{\mathcal{R}(-\mathcal{R}(y) : \mathcal{C}(x)) : \mathcal{C}(w)}{\mathcal{R}(\mathcal{R}(y) : -\mathcal{C}(x)) : \mathcal{C}(w)} \qquad \textbf{(A Fortiori)}$$

The inference rules below are concerned with various principles permitting to infer arguments from other arguments. One such principle is the idea that, if y is a reason for z and vice-versa, then z is a reason for whatever y is a reason for. This motivates the following inference rule.

$$\frac{\mathcal{R}(y) : \mathcal{C}(z) \quad \mathcal{R}(z) : \mathcal{C}(y) \quad \mathcal{R}(y) : \mathcal{C}(x)}{\mathcal{R}(z) : \mathcal{C}(x)} \qquad \textbf{(Mutual Support)}$$

Another principle is that if each of y and z is a reason for x, then the disjunction y or z is a reason for x. Conversely, if y or z is a reason for x then any of y and z must be a reason for x. All this can be expressed by the next rule, as follows.

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : \mathcal{C}(x)}{\mathcal{R}(y\,|\,z) : \mathcal{C}(x)} \qquad \frac{\mathcal{R}(y\,|\,z) : \mathcal{C}(x)}{\mathcal{R}(y) : \mathcal{C}(x)} \qquad \textbf{(Or)}$$

There is also the idea that if a reason can be decomposed into parts, of which one, say y, is a reason for the others (namely, the other parts), then y is enough of a reason. The next inference rule takes care of this.

$$\frac{\mathcal{R}(y \,\&\, z) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y) : \mathcal{C}(x)} \qquad \textbf{(Cut)}$$

The next rule turns an argument whose claim is itself an argument into an argument with decreased depth of nesting in $\mathcal{C}(.)$, as follows.

$$\frac{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))}{\mathcal{R}(y \,\&\, z) : \mathcal{C}(x)} \qquad \textbf{(Exportation)}$$

The last rule expresses how permutation of reasons can take place.

$$\frac{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))}{\mathcal{R}(z) : \mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))} \qquad \textbf{(Permutation)}$$

From now on, $\Vdash$ denotes the system consisting of (Consistency) together with the seven rules above from (A Fortiori) to (Permutation), closed under substitution and the meta-rule. Similarly, "derive" will refer to the usual notion for $\Vdash$ thus defined.

We show that $-\mathcal{R}(y) : \mathcal{C}(x)$ cannot be schematically derived from $\mathcal{R}(y) : \mathcal{C}(x)$ and that $-\mathcal{R}(y) : -\mathcal{C}(x)$ cannot be schematically derived from $\mathcal{R}(y) : -\mathcal{C}(x)$, so that a basic kind of consistency for the consequence operator $\Vdash$ is ensured.

Property 2. There is no $i, j \in \{0, 1\}$ such that

$$\frac{-^{(i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x)}{-^{(1-i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x)}$$

is a derived inference rule.

Property 2 furthermore expresses (using inference rules in Property 1) that neither $\mathcal{R}(y) : \mathcal{C}(-x)$ nor $\mathcal{R}(y) : -\mathcal{C}(x)$ can be schematically derived from $\mathcal{R}(y) : \mathcal{C}(x)$.

The Boolean identity $\alpha \,\&\, \alpha = \alpha$ yields an instance of (Exportation) worth mentioning, that is

$$\frac{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))}{\mathcal{R}(y) : \mathcal{C}(x)}.$$

The converse rule

$$\frac{\mathcal{R}(y) : \mathcal{C}(x)}{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))}$$

seems acceptable as well, in which case $\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))$ could be identified with $\mathcal{R}(y) : \mathcal{C}(x)$.

3.4 Non-Inference

Now we consider some inference rules that do not hold for the consequence relation we are presenting in this paper. First, if there were any axiom, the most likely candidate would be

$$-\mathcal{R}(\top) : \mathcal{C}(\bot).$$

The reader may find it surprising that the list above includes no inference rule induced by logical consequence. Unfortunately, most of the expected rules fail as detailed now.

$$\frac{}{\mathcal{R}(x) : \mathcal{C}(x)}\; x \in \mathbb{L} \qquad \textbf{(Reflexivity)}$$

Key is the difference between being an argument syntactically and being an argument that is held. $\mathcal{R}(x) : \mathcal{C}(x)$ is identified with an argument, by the mere fact that it does conform with Definition 2. Taking $\mathcal{R}(x) : \mathcal{C}(x)$ as an axiom would mean that any agent would be regarded as committed to holding $\mathcal{R}(x) : \mathcal{C}(x)$ for every x. Depending on the nature of the link in the argument, (i.e., the reading of the colon), this might be inappropriate. Think of a recruitment committee member who holds that "*Tracy should be given the position*". Taking x to stand for the statement that Tracy should be given the position, the argument $\mathcal{R}(x) : \mathcal{C}(x)$ is certainly not acceptable. Indeed, what is expected in such committees is to bring independent evidence in favour of candidates.

$$\frac{\models y \to x}{\mathcal{R}(y) : \mathcal{C}(x)}\; x, y \in \mathbb{L} \qquad \textbf{(Logical Consequence)}$$

Inhibiting (Reflexivity) as just argued implies that (Logical Consequence) must also be left out, because (Reflexivity) follows from (Logical Consequence).

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \models y \leftrightarrow z}{\mathcal{R}(z) : \mathcal{C}(x)}\; y, z \in \mathbb{L} \qquad \textbf{(Left Logical Equivalence)}$$

(Left Logical Equivalence) must be left out, again on the grounds that the nature of the link need not conform with logical consequence. Most notably, an effect need not be caused by something logically equivalent to its cause. However, (Mutual Support) can be viewed as a restricted substitute to this purported rule.

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \models z \to y}{\mathcal{R}(z) : \mathcal{C}(x)}\; y, z \in \mathbb{L} \qquad \textbf{(Left Logical Consequence)}$$

This is even more dubious, it actually entails (Left Logical Equivalence) and is then not worth considering any further.

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \models x \to w}{\mathcal{R}(y) : \mathcal{C}(w)}\; w, x \in \mathbb{L} \qquad \textbf{(Right Logical Consequence)}$$

(Right Logical Consequence) cannot be adopted either because being a reason for x is in general more restrictive than having x as a logical consequence. Consider for instance the causal argument in [1]: flu is a reason for your body temperature to be in the range 39° C−41° C. However, the fact that being in the range 36° C−41° C is a logical consequence of being in the range 39° C−41° C does not make flu a reason for your body temperature to be in the range 36° C−41° C (it is the only possible range unless you are dead!).

Interestingly, failure of (Right Logical Consequence) dismisses the seemingly harmless rule below

$$\frac{-\mathcal{R}(y) : \mathcal{C}(w)}{-\mathcal{R}(y) : \mathcal{C}(x\&w)}$$

which is nothing but the contrapositive of an instance of (Right Logical Consequence) —let x be $x \,\&\, w$.

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y) : \mathcal{C}(x \wedge z)} \qquad \textbf{(And)}$$

(And) is inappropriate, too. To start with, (And) opposes most cases dealing with limited resources. Certainly, from the fact that I have one Euro is a reason for me to buy a chocolate bar and is also a reason for me to buy a pastry, it cannot sensibly be held that the fact that I have one Euro is a reason for me to buy *both*. Assume that y stands for "I have one Euro" while x and z stand for "I am to buy a chocolate bar" and "I am to buy a pastry". Definitely, it would be wrong to derive $\mathcal{R}(y) : \mathcal{C}(x \wedge z)$ from $\mathcal{R}(y) : \mathcal{C}(x)$ together with $\mathcal{R}(y) : \mathcal{C}(z)$. Another case against (And), that does not involve limited resources, can be found in [1].

$$\frac{\mathcal{R}(z) : \mathcal{C}(y) \qquad \mathcal{R}(y) : \mathcal{C}(x)}{\mathcal{R}(z) : \mathcal{C}(x)} \qquad \textbf{(Transitivity)}$$

(Transitivity) can be challenged by means of the Ryan example in the introduction. The fact that Ryan's car is broken does not support the conclusion "Ryan is in his office" but precludes it instead. (Transitivity) fails mainly due to $\mathcal{R}$ being non-monotonic in the following sense: It can be the case that y is generally a reason for x although there are some special circumstances where this breaks down.

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y \wedge z) : \mathcal{C}(x)} \qquad \textbf{(Cautious Monotonicity)}$$

(Cautious Monotony), which is adapted from the study of non-monotonic consequence relations [13], is the controversial principle that the reason y in an argument for an x can be expanded with any statement z for which y is a reason. There is an interest in such a principle because it is the converse of (Cut) when both are viewed as principles applying in the context of $\mathcal{R}(y) : \mathcal{C}(z)$. Dismissal of (Cautious Monotony) cannot be escaped, if only from its incompatibility with the converse of (Exportation), namely (Importation).

$$\frac{\mathcal{R}(y \wedge z) : \mathcal{C}(x)}{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))} \qquad \textbf{(Importation)}$$

Actually, in the case that both (Importation) and (Cautious Monotony) were adopted, for every z for which y is a reason, $\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))$ would ensue. In particular, $\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(x) : \mathcal{C}(x))$ would hold for every x for which y is a reason.

Property 3. (Mutual Support) is a restricted version of (Transitivity).

Blocking a reason is different from blocking a conclusion. In symbols:

- $\mathcal{R}(y) : -\mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x)) \quad \not\Vdash \quad \mathcal{R}(y) : -\mathcal{C}(x)$.
- $\mathcal{R}(y) : -\mathcal{C}(x) \quad \not\Vdash \quad \mathcal{R}(y) : -\mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))$.

Consider the following argument.

> *The fact that several European countries have a good economy* (ge) *is a reason for not concluding that an economic crisis (ec) in Spain is a reason for a declining value of the euro* (de).

This has the form $\mathcal{R}(ge) : -\mathcal{C}(\mathcal{R}(ec) : \mathcal{C}(de))$. Please note that $\mathcal{R}(ge) : -\mathcal{C}(de)$ does not necessarily hold since an economic crisis in Germany may lead to a declining value of the euro.

Consider now the informal argument:

> *The fact that Steve did not follow the course (fc) is a reason for his failing his exams.*

It is formally captured as $\mathcal{R}(\neg fc) : \mathcal{C}(fe)$. That this argument is doubted on the grounds that Steve is smart can then be written $\mathcal{R}(sm) : -\mathcal{C}(\mathcal{R}(\neg fc) : \mathcal{C}(fe))$. However, the latter argument, $\mathcal{R}(sm) : -\mathcal{C}(\mathcal{R}(\neg fc) : \mathcal{C}(fe))$, need not hold even in the presence of $\mathcal{R}(sm) : -\mathcal{C}(fe)$ (Steve being smart is a reason not to conclude his failing his exams).

3.5 Properties of the consequence relation

We show that the consequence relation $\Vdash$ meets the minimum requirements as argued by Tarski [14].

Property 4. The following are properties of the $\Vdash$ relation where Δ is a set of RC-formulas, and α and β are RC-formulas.

$$\begin{array}{ll} \Delta \Vdash \alpha \text{ if } \alpha \in \Delta & \textbf{(Reflexivity)} \\ \Delta \cup \{\alpha\} \Vdash \beta \text{ if } \Delta \Vdash \beta & \textbf{(Monotonicity)} \\ \Delta \Vdash \beta \text{ if } \Delta \cup \{\alpha\} \Vdash \beta \text{ and } \Delta \Vdash \alpha & \textbf{(Cut)} \end{array}$$

In addition, the $\Vdash$ consequence relation is *paraconsistent* in the following sense.

Property 5. The following non-trivialization property holds for the $\Vdash$ relation:

$$\{-^{(i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x), -^{(1-i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x)\} \nVdash \mathbb{L}_{RC}.$$

The properties of reflexivity, monotonicity, and cut, mean that with the $\Vdash$ consequence relation, the manipulation of arguments by the inference rules is well-founded. The non-trivialization property means that contradictory arguments can be handled in a straightforward way.

It is worth pointing out that, even though $\Vdash$ is monotonic, it does exhibits non-monotonicity through its object language in the guise of $\mathcal{R}$. Indeed, "being a reason" is a non-monotonic inference relation $\mid\!\sim$ as witnessed by failure of transitivity. However, the fact that $\mathcal{R}$ plays the role of $\mid\!\sim$ in our formalism makes the non-monotonicity confined to failure of inferring $\mathcal{R}(y \wedge z) : \mathcal{C}(x)$ from $\mathcal{R}(y) : \mathcal{C}(x)$. Therefore, this has no effect on the logic. As an aside, the situation is similar to conditional logics because an operator capturing a counterfactual conditional must be non-monotonic (still, conditional logics are monotonic). E.g., "were I to scratch this match, it would ignite" denoted $y \rightarrowtail x$ may hold while "were I to scratch this match, that is wet, it would ignite" denoted $y \wedge z \rightarrowtail x$ fails to hold.

4 EXPRESSIVENESS OF THE LANGUAGE

This section discusses the expressive power of the language, namely the effects of allowing nesting of $\mathcal{R}(.)$ and $\mathcal{C}(.)$ on

- encoding *meta-arguments*,
- expressing various forms of *attacks*, and
- expressing *supports* between arguments.

4.1 Meta-arguments

The next table displays various forms of arguments allowed by Definition 1. Of course, the table is not exhaustive.

Basic arguments	F_1	$\mathcal{R}(y) : \mathcal{C}(x)$
	F_2	$\mathcal{R}(y) : \mathcal{C}(\neg x)$
	F_3	$\mathcal{R}(y) : -\mathcal{C}(x)$
Single-embedding meta-arguments (in reason)	F_4	$\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(y)) : \mathcal{C}(x)$
	F_5	$\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(y)) : \mathcal{C}(\neg x)$
	F_6	$\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(y)) : -\mathcal{C}(x)$
Single-embedding meta-arguments (in conclusion)	F_7	$\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))$
	F_8	$\mathcal{R}(y) : \mathcal{C}(-\mathcal{R}(z) : \mathcal{C}(x))$
	F_9	$\mathcal{R}(y) : -\mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))$
Double-embedding meta-arguments	F_{10}	$\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(y)) : \mathcal{C}(\mathcal{R}(t) : \mathcal{C}(x))$
	F_{11}	$\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(y)) : \mathcal{C}(-\mathcal{R}(t) : \mathcal{C}(x))$
	F_{12}	$\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(y)) : -\mathcal{C}(\mathcal{R}(t) : \mathcal{C}(x))$

Next is a list of arguments showing that each form F_i makes sense.

F_1: Tweety can fly (f). It is a bird (b). $\mathcal{R}(b) : \mathcal{C}(f)$
F_2: Tweety cannot fly. It is a penguin (p). $\mathcal{R}(p) : \mathcal{C}(\neg f)$
F_3: Steve is smart. Thus, it is not possible to conclude that he will fail his exams. $\mathcal{R}(sm) : -\mathcal{C}(fe)$
F_4: That Tweety can fly because it is a bird, is a reason to conclude that Tweety has wings (w). $\mathcal{R}(\mathcal{R}(b) : \mathcal{C}(f)) : \mathcal{C}(w)$

F_5: That Steve will fail his exams because he did not work hard is a reason to conclude that he is not so smart.
$\mathcal{R}(\mathcal{R}(\neg wh) : \mathcal{C}(fe)) : \mathcal{C}(\neg sm)$

F_6: Paul's car is in the park (pr) because it is broken (br), hence we cannot conclude that Paul is in his office (of).
$\mathcal{R}(\mathcal{R}(br) : \mathcal{C}(pr)) : -\mathcal{C}(of)$

F_7: The weather is sunny (su). Thus, rain (ra) will lead to rainbow (rb). $\mathcal{R}(su) : \mathcal{C}(\mathcal{R}(ra) : \mathcal{C}(rb))$

F_8: The fact that Tweety is a penguin is a reason to conclude that being a bird is not a sufficient reason for Tweety being able to fly. $\mathcal{R}(p) : \mathcal{C}(-\mathcal{R}(b) : \mathcal{C}(f))$

F_9: The fact that all European countries have a strong economy (se) is a reason for not concluding that an economic crisis (ec) in Germany is a reason for a declining value of the euro (de). $\mathcal{R}(se) : -\mathcal{C}(\mathcal{R}(ec) : \mathcal{C}(de))$

F_{10}: CFCs (cfc) cause damage to the ozone layer of the atmosphere (do). Man-made pollution (mp) causes global warming (gw). $\mathcal{R}(\mathcal{R}(cfc) : \mathcal{C}(do)) : \mathcal{C}(\mathcal{R}(mp) : \mathcal{C}(gw))$

F_{11}: Stress is the reason that Steve will fail his exams, hence it is not the fact that he did not work hard (st).
$\mathcal{R}(\mathcal{R}(st) : \mathcal{C}(fe)) : \mathcal{C}(-\mathcal{R}(\neg wh) : \mathcal{C}(fe))$

F_{12}: The object looks red (lr). It is illuminated by red light (il). Thus, we cannot conclude that looking red implies the object being indeed red (re). $\mathcal{R}(\mathcal{R}(il) : \mathcal{C}(lr)) : -\mathcal{C}(\mathcal{R}(lr) : \mathcal{C}(re))$

4.2 Expressing attacks

An argument $\mathcal{R}(y) : \mathcal{C}(x)$ may be attacked on any one of its components: conclusion, premises or the function of reason. For instance, the RC-formula below

$$-\mathcal{R}(\mathcal{R}(y) : \mathcal{C}(x)) : \mathcal{C}(x).$$

attacks the premise y in the argument $\mathcal{R}(y) : \mathcal{C}(x)$ because $-\mathcal{R}(\mathcal{R}(y) : \mathcal{C}(x)) : \mathcal{C}(x)$ states that y being a reason for concluding x is not enough to conclude x; therefore y must fail: if y were the case then, that y is a reason for concluding x would lead to conclude x. By contrast, attacking the link in the argument $\mathcal{R}(y) : \mathcal{C}(x)$ is simply the rejection

$$-\mathcal{R}(y) : \mathcal{C}(x).$$

We propose below a set of inference rules which not only show the various forms of attacks that may hold between arguments, but also how to detect attacks (the rules themselves) and how to express attacks as arguments (the part β of a rule α/β).

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : \mathcal{C}(\neg x)}{\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(\neg x)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \qquad \textbf{(Strong Rebuttal)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : -\mathcal{C}(x)}{\mathcal{R}(\mathcal{R}(z) : -\mathcal{C}(x)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \qquad \textbf{(Weak Rebuttal)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : \mathcal{C}(\neg y)}{\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(\neg y)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \qquad \textbf{(Strong Premise Attack)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : -\mathcal{C}(y)}{\mathcal{R}(\mathcal{R}(z) : -\mathcal{C}(y)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \qquad \textbf{(Weak Premise Attack)}$$

$$\mathcal{R}(z) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x)) \qquad \textbf{(Strong Reason Attack)}$$

$$\mathcal{R}(z) : -\mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x)) \qquad \textbf{(Weak Reason Attack)}$$

$$-\mathcal{R}(y) : \mathcal{C}(x) \qquad \textbf{(Pure Reason Attack)}$$

Note that the three attack relations that are distinguished in existing argumentation formalisms are captured in the new setting: *rebuttal* is captured by (Strong Rebuttal), *assumption attack* corresponds to (Strong Premise Attack) and Pollock undercutting is reflected by our notion of (Weak Rebuttal). However, since in those formalisms it is not possible to build arguments for blocking conclusions, the blocking is done in an indirect way as explained in Example 1. Therefore, with our logic of arguments, we can formalize and manipulate attacks explicitly within the logic (which is not possible in other formal systems of argumentation), and we have a wider range of attacks than are considered in other formal proposals for argumentation. For instance, in our formalism the *argumentative orientation* of the reason y towards the conclusion x of an argument $\mathcal{R}(y) : \mathcal{C}(x)$ can be attacked. Consider the following example borrowed from [15].

Floyd: *"A World Apart" is not a good movie. It does not teach us anything new about apartheid.*
Gary: *That's precisely what makes it good.*

Let

- gm stand for "*A World Apart* is a good movie",
- $\neg ta$ stand for "It does not teach us anything new about apartheid",

Then, Floyd's utterance can be captured by $\mathcal{R}(\neg ta) : \mathcal{C}(\neg gm)$. Gary's can be expressed by $\mathcal{R}(\neg ta) : \mathcal{C}(gm)$ and his argument reconstructed as:

$$\mathcal{R}(\mathcal{R}(\neg ta) : \mathcal{C}(gm)) : \mathcal{C}(-\mathcal{R}(\neg ta) : \mathcal{C}(\neg gm)).$$

4.3 Expressing supports

Unlike attacks which express negative links between arguments, supports express positive links. In the existing literature (e.g., [16,17]), such links are captured by a binary relation defined on the set of arguments. In our formalism, such an external relation is not needed since supports can be expressed by arguments of the form

$$\mathcal{R}(\mathcal{R}(y) : \mathcal{C}(x)) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(w))$$

or

$$\mathcal{R}(v) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(w)).$$

Let us return to Steve and his exams:

Hugh: *Steve will pass his exams. He is very smart.*
Ian: *He is well prepared.*

Letting wp stand for "Steve is well prepared", Ian's argument can be formalized as $\mathcal{R}(wp) : \mathcal{C}(\mathcal{R}(sm) : \mathcal{C}(pe))$ (Hugh's is $\mathcal{R}(sm) : \mathcal{C}(pe)$). From $\mathcal{R}(wp) : \mathcal{C}(\mathcal{R}(sm) : \mathcal{C}(pe))$, using the reduction rule, the argument $\mathcal{R}(wp \wedge sm) : \mathcal{C}(pe)$ ensues.

$$(\mathcal{R}(z) : \mathcal{C}(y)) : \mathcal{C}(x)$$

is an even more direct form expressing that $\mathcal{R}(y) : \mathcal{C}(x)$ is supported by $\mathcal{R}(z) : \mathcal{C}(y)$. It is obtained from the more general form above, using reduction. Also, rejection of support has the form

$$\mathcal{R}(z) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x)).$$

5 Cube of opposition

The use of a cube of opposition is an interesting way of organizing complementary notions in the study of logics (e.g. [18]). The idea of opposition plays also an important role in argumentation [19]. Indeed, Apothéloz [3] has pointed out the existence of four basic argumentative forms, where two negations are at work: i) "y is a reason for concluding x", ii) "y is not a reason for concluding x", iii) "y is a reason for not concluding x", and iv) "y is not a reason for not concluding x". These four statements were organized by Salavastru [20] in a square of opposition, which was slightly corrected in [21] (the vertical entailments were put in the wrong way). The following cube of oppositions summarizes the different links between basic forms of arguments and rejections and adapts the previous proposal [21].

A diagonal (whether dashed or not) expresses a contradiction, i.e., its two endpoints cannot be true together and cannot be false together. A thicker edge (whether dashed or not) expresses contrariness, i.e., its two endpoints cannot be true together —but they can be false together. An even thicker edge expresses sub-contrariness, i.e., its two endpoints cannot be false together —but they can be true together. The cube contains some subaltern relations represented by arrows, providing a direction. For instance, the arrow from $\mathcal{R}(y) : \mathcal{C}(x)$ to $-\mathcal{R}(y) : -\mathcal{C}(x)$ means that if the former holds, then so does the latter.

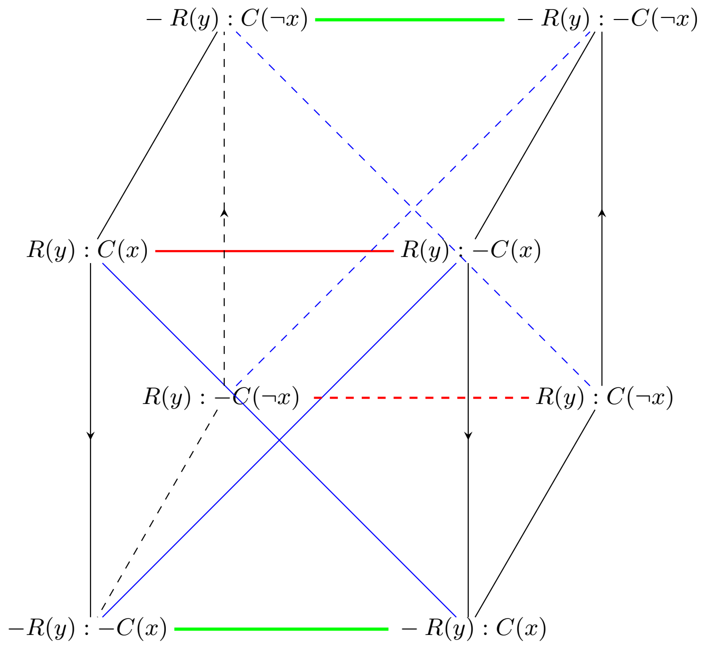


Fig. 1.1. Cube of opposition for RC-formulas.

6 CONCLUSION

This paper proposes a novel logic for representing and reasoning about arguments in a way that is just not possible with the existing formalisms. The logical language is made of arguments and rejections of arguments. The definition of arguments encompasses different roles of reasons (concluding and blocking statements), various forms of reasons (factual and hypothetical) and different kinds of links (deductive, abductive, inductive, ...). Unlike the existing computational models of argumentation where attacks and supports between arguments are expressed by external relations on the set of arguments, in the new logic they are elements of the language. Indeed, every attack (respectively support) is expressed as an argument.

The logic offers key advantages. First, it respects the nature of argument. Indeed, it does not reduce the *meaning* of the statements to a formal derivation between the reason and the conclusion. To say this differently, not any such derivation is a natural argument. Importantly, $\Vdash$ does *not* serve to handle, or cure, inconsistency between arguments, but it provides, in a logical setting, a basis for reasoning between arguments. Second, it can be parameterized for several purposes like reasoning about causal arguments, analogical arguments, decision arguments, etc. Third, it lends itself to encoding *fairly* natural language dialogues. Indeed, one may pass directly from natural language dialogue to the logical setting without intermediate encodings which are often convoluted. Moreover, preferences between arguments could be captured as meta-arguments. Fourth, it provides the basis for a *logic of argumentation*, i.e., a logic in which arguments are represented and *evaluated*. Indeed, in the future, we plan to define on top of $\langle \mathbb{L}_{RC}, \Vdash \rangle$, a logic $\langle \mathbb{L}_{RC}, \Vvdash \rangle$ dedicated to acceptability of arguments (i.e., $\Vvdash$ will return the accepted arguments).

Acknowledgements

The authors are grateful to the reviewers of a former version of this article for their useful remarks.

Bibliography

1. Amgoud, L., Besnard, P., Hunter, A.: Representing and reasoning with arguments mined from texts and dialogues. In Destercke, S., Denoeux, T., eds.: Proc. of the 13th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'2015). Volume 9161 of Lecture Notes in Computer Science., Compiègne, France, Springer (2015) 60–71
2. Amgoud, L., Besnard, P., Hunter, A.: Logical representation and analysis for RC-arguments. In Esposito, A., ed.: Proc. of the 27th IEEE International Conference on Tools with AI (ICTAI'2015), Vietri sul Mare, Italy, IEEE (2015)
3. Apothéloz, D.: The function of negation in argumentation. Journal of Pragmatics (1993) 23–38

4. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2) (1995) 321–357
5. Boella, G., Gabbay, D., van der Torre, L., Villata, S.: Meta-argumentation modelling I: methodology and techniques. Studia Logica **93**(2-3) (2009) 297–355
6. Modgil, S., Bench-Capon, T.: Metalevel argumentation. Journal of Logic and Computation **21**(6) (2011) 959–1003
7. Wooldridge, M., McBurney, P., Parsons, S.: On the meta-logic of arguments. In Parsons, S., Maudet, N., Moraitis, P., Rahwan, I., eds.: Proc. of the 2nd International Workshop on Argumentation in Multi-Agent Systems (ArgMAS'2005). Volume 4049 of Lecture Notes in Computer Science., Utrecht, The Netherlands, Springer (2005) 42–56
8. Bondarenko, A., Dung, P., Kowalski, R., Toni, F.: An abstract, argumentation-theoretic approach to default reasoning. Artificial Intelligence **93** (1997) 63–101
9. Prakken, H.: An abstract framework for argumentation with structured arguments. Argument and Computation **1**(2) (2010) 1–31
10. Besnard, P., Hunter, A.: Elements of Argumentation. MIT Press (2008)
11. García, A., Simari, G.: Defeasible logic programming: an argumentative approach. Theory and Practice of Logic Programming **4**(1-2) (2004) 95–138
12. Verheij, B.: Deflog: on the logical interpretation of prima facie justified assumptions. Journal of Logic and Computation **13**(3) (2003) 319–346
13. Kraus, S., Lehmann, S., Magidor, D.: Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence **44** (1990) 167–207
14. Tarski, A.: On some fundamental concepts of metamathematics. In: Logic, Semantics, Metamathematics. Edited and translated by J. H. Woodger, Oxford U. Press (1956)
15. Quiroz, G., Apothéloz, D., Brandt, P.: How counter-argumentation works. In van Eemeren, F.H., Grootendorst, R., Blair, J.A., Willard, C.A., eds.: Argumentation Illuminated. International Centre for the Study of Argumentation (SICSAT), Amsterdam (1992) 172–177
16. Cayrol, C., Lagasquie, M.: Bipolarity in argumentation graphs: towards a better understanding. International Journal of Approximate Reasoning **54**(7) (2013) 876–899
17. Cohen, A., García, A., Simari, G.: Backing and undercutting in defeasible logic programming. In Liu, W., ed.: Proc. of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'2011). Volume 6717 of Lecture Notes in Computer Science., Belfast, U.K., Springer (2011) 50–61
18. Béziau, J.Y.: New light on the square of oppositions and its nameless corner. Logical Investigations **10** (2003) 218–233
19. Moretti, A.: Argumentation theory and the geometry of opposition (abstract). In van Eemeren, F., Garssen, B., Godden, D., Mitchell, G., eds.: Proc. of the 7th Conference of the International Society for the Study of Argumentation (ISSA'2010), Amsterdam, The Netherlands, International Centre for the Study of Argumentation (SICSAT), Amsterdam (2010)
20. Salavastru, C.: Logique, Argumentation, Interprétation. L'Harmattan, Paris (2007)
21. Amgoud, L., Prade, H.: Towards a logic of argumentation. In Hüllermeier, E., Link, S., Fober, T., Seeger, B., eds.: Proc. of the 6th Conference on Scalable Uncertainty Management (SUM'2012). Volume 7520 of Lecture Notes in Computer Science., Marburg, Germany, Springer (2012) 558–565

Extended Mereotopology Based on Sequent Algebras ⋆

Dimiter Vakarelov

Department of Mathematical Logic
Faculty of Mathematics and Informatics
Sofia University blvd James Bouchier 5
1126 Sofia, Bulgaria
e-mail dvak@fmi.uni-sofia.bg

Essay in honor of Luis Farinas del Cerro for his anniversary

Abstract. Mereotopology is an extension of mereology with some relations of topological nature like contact. An algebraic counterpart of mereotopology is the notion of contact algebra which is a Boolean algebra whose elements are considered to denote spatial regions, extended with a binary relation of **contact** between regions. Although the language of contact algebra is quite expressive to define many usefull mereological relations (**part-of**, **overlap**, **underlap**) and mereotopological relations (**external contact**, **tangential part-of**, **non-tangential part-of**, **self-connectedness**), there are, however, some interesting mereotopological relations which are not definable in it. Such are, for instance, the relation of **n-ary contact**, **internal connectedness** and some others. To overcome this disadvantage we introduce a generalization of contact algebra, replacing the contact with a binary relation $A \vdash b$ between finite sets of regions and a region, satisfying some formal properties of Tarski consequence relation. The obtained system is called **sequent algebra**, considered as an extended mereotopology. We develop the topological representation theory for sequent algebras showing in this way certain correspondence between point-free and point-based models of space. As a bi-product we show how one logical in nature notion - Tarski consequence relation, may have also certain spatial (mereotopological) meaning.

Keywords: mereology, mereotopology, point-free theory of space, sequent algebra, topology, representation theorem.

1 Introduction

Mereotopology is an extension of mereology [13] with some relations of topological nature like contact. Mereotopology is considered also as a kind of point-free theory of space, called also region-based theory of space (RBTS). The roots of RBTS go back to Whitehead [17,18] and the main idea arose from some criticism to the classical point-based approach to the theory of space (see the survey papers [16,2,8] for some historical remarks and motivations on RBTS).

Let us mention that in a sense RBTS had been reinvented in computer science, because of its more simple way of representing qualitative spatial information and in fact it initiated a special field in Knowledge Representation (KR) called *Qualitative Spatial Representation and Reasoning* (QSRR). Survey papers about applications of RBTS and mereotopology in various applied areas are, for instance, [3,8] and the book [9].

An algebraic counterpart of mereotopology is the notion of contact algebra (CA), which is an extension of Boolean algebra with a binary relation C called contact. The elements of the algebra are considered to denote spatial regions, boolean operations are considered as operations for constructions new regions from given ones and also for defining some standard mereological relations between regions as **part-of** relation, **overlap**, **underlap** and others. The intended point models of CA are Boolean algebras of regular-closed subsets of a topological space with aCb iff a and b share a common point (for more details on CA-s and their topological representation see [5,16] and the recent paper [7] describing the topological duality theory of contact algebras). The language of CA is quite expressive. For instance, by means of contact one can define various mereotopological relations between regions: **external contact**, **tangential part-of**, **non-tangential part-of**, **self-connectedness** and others. However, there are some interesting mereotopological relations which are not definable in contact algebras: **n-ary contact**, **internal connectedness**, and others. In order to increase the expressive power of contact algebra, we introduce in this paper a generalization, replacing the contact relation by a new relation $A \vdash b$ between a finite set of regions A and a region b, called sequent relation.

⋆ This paper is sponsored by the Sofia University Science Fund

The obtained system is called **sequent algebra** (S-algebra for short), because the sequent relation satisfies the structural properties of Tarski consequence relation. In topological models with regular closed sets the definition of $\vdash$ is the following: for a given finite set of regions $A = \{a_1, \ldots, a_n\}$ and a region b, $a_1, \ldots, a_n \vdash a$ iff $a_1 \cap \ldots \cap a_n \subseteq b$. By means of the sequent relation one can define ordinary contact by aCb iff $a, b \not\vdash 0$ and n-ary contact by $a_1, \ldots, a_n \not\vdash 0$ where 0 is the zero element of the Boolean algebra (the notion of internal connectedness is also definable - see Section 3.4). We show in this way that Tarski consequence relation, which is a typical logical relation, has also a non-logical, mereotopological meaning.

The structure of the paper is the following. In Section 2 we give abstract definitions of Skott and Tarski sequent systems considered as abstractions from some properties of Gentzen calculi. In Section 3 we introduced the notion of Sequent algebra (S-algebra) as a generalization of contact algebra in which contact is replaced by Tarski consequence relation. Section 4 is devoted to the representation theory of S-algebras. Here we introduce the abstract points of S-algebra, called S-clans, which are generalizations of the abstract points of contact algebras, called clans. The main results of this section, which are also the main results of the paper, are two: Topological representation theorem for S-algebras (Theorem 1), and Discrete representation theorem for S-algebras (Theorem 2).

2 Scott and Tarski sequent systems

In this section we will introduce some abstract versions of Gentzen sequent systems. Let us remaind that a Gentzen sequent is an expression of the form $A \vdash B$ where A and B are finite sets of logical formulas. The intuitive meaning of $A \vdash B$ is "the conjunctions of the formulas of A imply the disjunctions of formulas of B. There are sequent calculi (as for intuitionistic logic) using only sequents in the simplified form $A \vdash b$. The meaning of such sequents is "b follows from the assumptions A", or "the set of formulas A entails the formula b, which is just the meaning of $\vdash$ in the definition of the finitary Tarski consequence relation.

Abstract sequent systems which will be used in this paper are in the form $(W, \vdash)$, where W is an abstract nonempty set (not set of formulas as in traditional Gentzen systems) and $\vdash$ is a binary relation ether in the form $A \vdash B$ or $A \vdash b$, where A and B are finite subsets of W and $b \in W$. The axioms of such systems are certain statements imitating some structural rules of Gentzen calculi. Let us mention that such abstract treating of sequent systems was given for the first time by Dana Scott [12] in an abstract definition of information system with applications to domain theory. Other abstract treating of sequent systems can be found in [19] - again with applications in domain theory, in [15] - with applications to some information modal logics, in [4] - with some topological and categorical characterizations of certain classes of consequence systems. Abstract sequent systems using sequents as a binary relation between finite sets and elements will be called "Tarski sequent systems". The more general abstract sequent systems using sequents between finite sets will be called "Scott sequent systems".

We adopt for later use standard abbreviations for sequents taken from Gentzen sequent calculi. For instance, instead $\{a_1, \ldots, a_n\} \vdash \{b_1, \ldots, b_m\}$ we write $a_1, \ldots, a_n \vdash b_1, \ldots, b_m$, instead $A \cup B \vdash C$ we write $A, B \vdash C$, instead $A \cup \{b_1, \ldots, b_n\} \vdash C$ we write $A, b_1, \ldots, b_n \vdash C$, and some other simplifications of similar kind.

Definition 1. (Scott sequent system) *Let W be a nonempty set and $\vdash$ be a binary relations between finite subsets of W. We say that $\vdash$ is a Scott consequence relation in W provided that the following axioms are satisfied (A, B are finite subsets of W and $x \in W$) :*

*(Ref) $x \vdash x$, **reflexivity***
*(Mono) If $A \vdash B$, then $A, x \vdash B$ and $A \vdash x, B$, **monotonicity**, and*
(Cut) If $A, x \vdash B$ and $A \vdash x, B$ then $A \vdash B$.

*We say that the system $S = (W, \vdash)$, is a **Scott sequent system** (Scott S-system for short), if $\vdash$ is a Scott consequence relation in W.*

Notation: $A \not\vdash B$ for the negation of $A \vdash B$. Expressions in the form $A \vdash B$ are called Scott sequents. The special case $A \vdash b$, where B is equal to the singleton set $\{b\}$ is called Tarski sequent. Expressions $A \vdash B$ and $A \vdash b$ will be used only for finite subsets A, B of W.

The following lemma states some generalizations of the above axioms which will be used later on by their names.

Lemma 1. ***(Properties of Scott consequent relation.)*** *The following conditions hold for arbitrary Scott S-system:*

(i) (REF) If $A \cap B \neq \varnothing$, then $A \vdash B$,
(ii) (MONO) If $A \vdash B$, $A \subseteq A'$ and $B \subseteq B'$, then $A' \vdash B'$,
(iii) (CUT) If $A_1, x \vdash B_1$ and $A_2 \vdash x, B_2$, then $A_1, A_2 \vdash B_1, B_2$.

Note that (REF), (MONO) and (CUT) are considered in the literature as the typical characteristic properties of $\vdash$. It is easy to see that in fact they are equivalent to the more simple (Ref), (Mono) and (Cut), which are more suitable in the verifications of these axioms in some models.

Sequent systems containing only Tarski sequents can be axiomatized separately as follows, taking the obvious restrictions of the axioms (Ref), (Mono) and (Cut) for which we will use the same notations.

Definition 2. *(***Tarski sequent system***) Let W be a nonempty set and $A \vdash b$ be a binary relation with first argument all finite subsets of W and second argument all elements of W. We say that $\vdash$ is a Tarski consequence relation in W provided that the following axioms are true:*

(Ref) $x \vdash x$,
(Mono) If $A \vdash x$, then $A, y \vdash x$,
(Cut) If $A \vdash x$, $A, x \vdash y$, then $A \vdash y$.

We say that the system $(W, \vdash)$ is a Tarski sequent system (Tarski S-system), if $\vdash$ is a Tarski consequence relation in W.

Let us note that for the sake of simplicity we use one and the same notation for Scott and Tarski consequence relations and the names of their axioms.

In the following lemma we list some properties of Tarski sequents which will be used later on by their names. We use again one and the same names for obviously analogical properties for Scott and Tarski sequents.

Lemma 2. ***(Properties of Tarski consequence relation.)*** *The following conditions hold for arbitrary Tarski S-system:*

(REF) For any $b \in W$ and any finite subset $A \subseteq W$: if $b \in A$, then $A \vdash b$,
(MONO) If $A \vdash b$ and $A \subseteq A'$, then $A' \vdash b$,
(CUT) If $A_1 \vdash x$, $A_2, x \vdash b$, then $A_1, A_2 \vdash b$,
(Tran) If $A \vdash b$ and $b \vdash c$, then $A \vdash c$, transitivity,
(TRAN) If $A_1 \vdash x_1$, ..., $A_n \vdash x_n$ and $B, x_1, ..., x_n \vdash b$, then $A_1, ..., A_n, B \vdash b$, extended transitivity.

Example 1. **Example of Scott and Tarski S-system**. Let W be a nonempty set whose elements are subsets of a given set X and let $A = a_1, \ldots, a_n$ and $B = b_1, \ldots, b_m$ be arbitrary finite subsets of W. Define $A \vdash B$ iff $a_1 \cap \ldots \cap a_n \subseteq b_1 \cup \ldots \cup b_m$. Then $(W, \vdash)$ is an S-system. Note that $A = \varnothing$ and $B = \varnothing$ are not excluded and are obtained in the cases $n = 0$ and $m = 0$ respectively. Then the empty intersection is taken to be the set X and the empty union to be the empty set $\varnothing$. Obvious modification of this example gives example of Tarski S-system.

3 Sequent algebras

3.1 Abstract definition

Definition 3. ***(Sequent Boolean algebra.)*** *The system $\underline{B} = (B, 0, 1, \leq, +, ., *, \vdash)$ is called a* **sequent Boolean algebra** *(S-algebra for short) if $(B, 0, 1, \leq, +, ., *)$ is a non-degenerate Boolean algebra, $(B, \vdash)$ is a Tarski S-system and $\vdash$ satisfies the following additional axioms:*

(S1) $x \vdash y$ iff $x \leq y$,
(S2) $\varnothing \vdash y$ iff $y = 1$,
(S3) If $A, x \vdash z$ and $A, y \vdash z$, then $A, x + y \vdash z$,

Lemma 3. ***(Properties of Tarski consequent relation in S-algebras.)*** *Let $\underline{B}$ be a S-algebra. Then the following conditions are true for the relations $\vdash$:*

(i) $A_1, x_1 \vdash y_1$ *and* $A_2, x_2 \vdash y_2$*, then* $A_1, A_2, x_1 + x_2 \vdash y_1 + y_2$*,*
(ii) If $A, x, y \vdash z$*, then* $A, x.y \vdash z$*,*
(iii) $A \vdash 1$*,*
(iv) If $A \vdash x$ *and* $x \leq y$*, then* $A \vdash y$*,*
(v) If $A, x \vdash b$ *and* $A \vdash x + b$*, then* $A \vdash b$*.*

Lemma 4. *Let* $\underline{B}$ *be a S-algebra, A and* $B = \{b_1, \ldots, b_n\}$ *be arbitrary finite subset of* $\underline{B}$*. Define* $A \vdash B$ *iff* $A \vdash b_1 + \ldots + b_n$*. Then this extension of Tarski sequence defines the structure of Scott S-system in* $\underline{B}$*.*

Let us note that in the above lemma the case $B = \varnothing$ is included by the standard assumption that the empty sum is equal to 0, i.e. $A \vdash \varnothing$ iff $A \vdash 0$.

3.2 Comparison with contact algebras

We remaind the abstract definition and some facts about contact algebras from [5,16]

Definition 4. *We say that the algebraic system* $(B, C) = (B, \leq, 0, 1, ., +, *, C)$ *is a contact algebra if* $(B, \leq, 0, 1, ., +, *)$ *is a non-degenerate Boolean algebra and C is a binary relation in B called contact and satisfying the following axioms:*

(C1) If aCb*, then* $a \neq 0$ *and* $b \neq 0$*,*
(C2) If aCb *and* $a \leq a'$ *and* $b \leq b'$*, then* $a'Cb'$*,*
(C3) If $aC(b + c)$*, then* aCb *or* aCc*,*
(C4) If aCb*, then* bCa*,*
(C5) If $a.b \neq 0$*, then* aCb*.*

The elements of B are considered to denote spatial regions and the Boolean operations are considered as operations for constructing new regions from given ones.

The intuitive meaning of the contact relation is that x and y share a common point. Let us note that the Boolean part of the definition of contact algebra can be considered as its mereological part. In this part one can define the standard mereological relations:
• **part-of relation** - $x \leq y$ - this is just the Boolean ordering relation,
• **overlap** - xOy iff $x.y \neq 0$,
• **underlap** (dual overlap) - xUy iff $x + y \neq 1$ iff x^*Oy^*.
By means of the contact relation one can define the following important mereotopological relations:
• **external contact** - $xC^{ext}y$ iff xCy and $x\overline{O}y$,
• **tangential part-of** - $x \ll y$ iff $x\overline{C}y^*$,
• **non-tangential part-of** $x \prec y$ iff $x \leq y$ and $x \not\ll y$
• **self-connectedness** $c(x)$ iff $(\forall y, z)(y \neq 0 \wedge z \neq 0 \wedge x = y + z \rightarrow yCz)$ (see [16,14]).

Examples of contact algebras

Example 2. **Topological example of contact algebra.** Let X be a topological space and for the subset $a \subseteq X$ and let $Int(a)$ and $Cl(a)$ denote correspondingly the topological interior and the closure of the set a. The set a is called a regular closed subset of X if $a = Cl(Int(a))$. It is a well known fact the set $RC(X)$ of regular closed subsets of X is a Boolean algebra under the following definitions of the corresponding Boolean constants and operations: $a \leq b$ iff $a \subseteq b$, 0=$\varnothing$, $1 = X$, $a + b = a \cup b$, $a.b = Cl(Int(a \cap b))$, $a^* = Cl(X \setminus a)$. It is a contact algebra under the definition: aCb iff $a \cap b \neq \varnothing$, i.e. a and b share a common point. This example of contact algebra is typical in the sense that every contact algebra is representable as a subalgebra of the contact algebra $RC(X)$ of certain topological space X. See [5,16] for the representation theory of contact algebras. This theory is based on a special kind of abstract points definable in contact algebras, called clans. In order to compare this notion with the abstract points in S-algebras we present here the definition of clan.

Definition 5. *A subset* $\Gamma \subseteq B$ *is a* ***clan*** *if it satisfies the following conditions:*

(Clan 1) $1 \in \Gamma$, $0 \notin \Gamma$*,*
(Clan 2) If $a \in \Gamma$ *and* $a \leq b$*, then* $b \in \Gamma$*,*
(Clan 3) If $a + b \in \Gamma$*, then* $a \in \Gamma$ *or* $b \in \Gamma$*,*
(Clan 4) If $a, b \in \Gamma$*, then* aCb*.*

Example 3. **Non-topological (discrete) example of contact algebra.** A non-topological example of contact algebra can be obtained as follows. Let (X, R) be a relational system with $X \neq \varnothing$ and R be a reflexive and symmetric relation in X. The Boolean algebra over all subsets of X is a contact algebra, provided the contact is defined as follows: for $a, b \subseteq X$ aCb iff $(\exists x \in a)(\exists y \in b)(xRy)$. This example is also typical because it can be proved that every contact algebra can be isomorphically embedded into the algebra of all subsets of certain relational system (X, R) with reflexive and symmetric relation R. For more details about such a "discrete" representation theory see [6].

3.3 Examples of S-algebras

Example 4. **Topological example of S-algebra.** This example extends the topological example of contact algebras by regular closed subsets of a topological space. For a finite subset A of the set of regular closed sets $RC(X)$ of X and $b \in RC(X)$ define Tarski consequence relation in $RC(X)$ as in the Example 1. Then $(RC(X), \vdash)$ is a Tarski sequent algebra. This kind of S-algebra is a typical one because we will show later on that each S-algebra can be isomorphically embedded into a topological S-algebra over a certain topological space.

Example 5. **Non-topological (discrete) example of S-algebra.** In this example we will extend in a sense the non-topological example of contact algebras as follows. Let (X, Y) be a pair (called a **discrete S-space**) with X a non-empty set and Y a set of non-empty subsets of X containing all singletons of X. Let $B(X, Y)$ be the Boolean algebra of all subsets of X.

We define Tarski consequence relation in $B(X, Y)$ as follows $(a_1, \ldots, a_n, b \in B(X, Y))$:

$a_1, \ldots, a_n \vdash b$ iff $(\forall x_1 \ldots x_n \in X, \forall \Gamma \in Y)(\{x_1 \ldots x_n\} \subseteq \Gamma, x_1 \in a_1, \ldots, x_n \in a_n \rightarrow b \cap \Gamma \neq \varnothing)$.

Then $B(X, Y)$, equipped with the above defined relation, is an S-algebra called the **discrete S-algebra over the discrete S-space** (X, Y)**.** The proof of this fact goes by routine verification of the axioms of S-algebra.

Discrete S-algebras are in a sense characteristic, because every S-algebra is representable as an S-algebra over a discrete S-space, as we will see later on.

3.4 Undefinability of some mereotopological relations in contact algebras and their definability in S-algebras.

Note that in contact algebras one can not express the **n-ary contact** (see [11]). In topological models n-ary contact is definable as follows ($a_1, \ldots, a_n$ are regular-closed subsets of a topological space X):

$C^n(a_1, \ldots, a_n)$ iff $a_1 \cap \ldots \cap a_n \neq \varnothing$, i.e. $a_1, \ldots, a_n$ share a common point.

Having in mind topological models of S-algebras one can see that contact and n-ary contact are definable in S-algebras as follows:

Definition of contact: aCb iff $a, b \nvdash 0$,

Definition of n-ary contact: $C^n(a_1, \ldots, a_n)$ iff $a_1, \ldots, a_n \nvdash 0$.

It can easily be seen that the above defined contact in S-algebra satisfies the axioms of contact, which shows that all S-algebras are contact algebras under the above definition. This implies that definable mereotopological relations in contact algebras are definaqble also in S-algebras.

Another interesting mereotopological relation which is not definable in contact algebras is **internal connectedness**. In topological models this is a property of a region saying that its internal part is topologically connected (in symbols $c^o(a)$).

Internal connectedness is definable in S-algebras as follows:

Definition of internal connectedness: $c^o(a)$ iff $(\forall b, c)(a \leq b+c \wedge a.b \neq 0 \wedge a.c \neq 0 \rightarrow b, c \nvdash a^*)$.

It can be proved that in topological models this formula is true for a regular closed set a iff $Int(a)$ is a connected set in the sense of topology.

Another equivalent definition of $c^o(a)$ in the so called **extended contact algebras** (ECA) is given in [10]. Extended contact algebra is a Boolean algebra with a ternary relation $a, b \vdash c$ which just axiomatizes the restriction of the relation $A \vdash c$ with A being at most two-element non-empty set. The paper [10] contains also a proof that $c^o(a)$ is not definable in contact algebras.

Internal connectedness of a region and n-ary contact were studied with respect to their expressiveness and computational complexity in [11].

4 Representation theory for sequent algebras

4.1 S-filters and S-clans

Definition 6. ***(S-filter.)*** *Let $\underline{B}$ be a S-algebra. A subset $F \subseteq B$ is a* ***S-filter*** *if the following condition is satisfied:*

(S-fil) for any finite subset $A \subseteq B$ and $b \in B$: if $A \subseteq F$ and $A \vdash b$, then $b \in F$.

F is a proper S-filter if $0 \notin F$.

F is a S-clan if F is a proper S-filter and for all $a, b \in B$, if $a + b \in F$, then $a \in F$ or $b \in F$.

F is a maximal S-filter if it is a proper S-filter and for every proper S-filter G: if $F \subseteq G$ then $F = G$.

F is a maximal S-clan if for every S-clan G: if $F \subseteq G$, then $F = G$.

Lemma 5. ***(Properties of S-filters and S-clans.)*** *Let F be an S-filter. Then:*

(i) $1 \in F$, if $a \in F$ and $a \leq b$ then $b \in F$.

(ii) If in addition F is an S-clan, then: $0 \notin F$, if $a + b \in F$ iff $a \in F$ or $b \in F$.

Lemma 6. *Each S-clan is a clan in the sense of contact algebra.*

Lemma 7. *F is a propper S-filter iff $F \neq B$.*

The next lemma shows that the notions of S-filter and S-clan generalize the notions of a (Boolean) filter and (Boolean) ultrafilter.

Lemma 8. *Let $\underline{B}$ be a S-algebra. Then:*

(i) every filter in $\underline{B}$ is a S-filter,

(ii) every ultrafilter in $\underline{B}$ is a S-clan.

The next lemma gives a general construction of S-clans by special sets of ultrafilters.

Lemma 9. ***(Construction of S-clans from ultrafilters.)*** *Let Σ be a nonempty set of ultrafilters satisfying the following condition:*

($\sharp$) $(\forall n)(\forall a_1, \ldots, a_n, b \in B)(\forall F_1, \ldots, F_n \in \Sigma)(a_1 \in F_1 \ldots a_n \in F_n$ and $a_1, \ldots, a_n \vdash b \rightarrow (\exists F \in \Sigma)(b \in F))$.

Let F be the union of the ultrafilters from Σ. Then F is a S-clan.

Note that the construction of S-clans by the above lemma depends on the existence of sets of ultrafilters satisfying the condition ($\sharp$). Later we will show how to construct such sets and that every S-clan can be constructed in this way.

The next lemma presents general constructions of S-filters.

Lemma 10. ***(Constructions of S-filters and S-clans.)*** *Let $\underline{B}$ be a S-algebra.*

(i) Let X be any subset of B and define $\mathbf{F}(X) =_{def} \{b \in B : (\exists finite A \subseteq X)(A \vdash b)\}$. Then $\mathbf{F}(X)$ is the smallest S-filter containing X.

(ii) Let F be an S-filter and $a \in B$. Define $F \oplus a =_{def} \{b \in B : (\exists finite A \subseteq B)(A, a \vdash b)\}$. Then $F \oplus a$ is the smallest S-filter containing F and a.

(iii) Let Γ be a non-empty chain of proper S-filters and let F be the union of the members of Γ. Then F is a proper S-filter.

Lemma 11. *(i) Every proper S-filter is contained in a maximal S-filter.*

(ii) Every maximal S-filter is a S-clan.

(iii) Every proper S-filter is contained in a S-clan.

(iv) Every S-clan is contained in a maximal S-clan.

Proof. (i) By an application of Zorn Lemma.

(ii) Let F be a maximal S-filter and suppose that F is not a S-clan. Then there are a_1 and a_2 such that $a_1 + a_2 \in F$ but $a_1 \notin F$ and $a_2 \notin F$. Let $F_1 = F \oplus a_1$ and $F_2 \oplus a_2$. From here we get $F \subset F_1$ and $F \subset F_2$. It follows by maximality of F that F_1 and F_2 are not proper, hence $0 \in F_1$ and $0 \in F_2$. By the definitions of F_1 and F_2 we obtain: there exist finite subsets A_1, A_2 of F such that $A_1, a_1 \vdash 0$ and $A_2, a_2 \vdash 0$. Then by Lemma 3 (i) we get $A_1, A_2, a_1 + a_2 \vdash 0$. Since $A = A_1 \cup A_2 \cup \{a_1 + a_2\} \subseteq F$, we obtain that $A \vdash 0$, hence $0 \in F$ - a contradiction.

(iii) The statement follows from (i) and (ii).

(iv) The proof follows by an application of Zorn Lemma.

Lemma 12. ***(Separation Lemma for S-filters and ideals)*** *Let $\underline{B}$ be a S-algebra, F_0 be a S-filter, I be an ideal (Boolean) in $\underline{B}$ and $F_0 \cap I = \varnothing$. Then there exists a S-clan F such that $F_0 \subseteq F$ and $F \cap I = \varnothing$.*

Proof. The proof goes by an application of Zorn Lemma. Let $M = \{G : G$ be an S-filter and $G \cap I = \varnothing\}$. Obviously the elements of M are proper S-filters. We have also that $F_0 \in M$, so $M \neq \varnothing$. Let N be a nonempty chain in M and denote by G_N the union of the elements of N. By Lemma 10 G_N is a S-filter which is an upper bound of N which obviously belongs to M. Thus, by the Zorn Lemma M has a maximal element, say F which extends F_0. It remains to show that F is a S-clan. Suppose the contrary, i.e. that there exist a_1 and a_2 in B such that $a_1 + a_2 \in F$ but $a_1 \notin F$ and $a_2 \notin F$. Consider the S-filters $F_1 = F \oplus a_1$ and $F_2 = F \oplus a_2$. Since F is a maximal element in M we get that $F_1 \cap I \neq \varnothing$ and $F_2 \cap I \neq \varnothing$. So, there exist b_1 and b_2 such that $b_1 \in F_1$, $b_1 \in I$ and $b_2 \in F_2$, $b_2 \in I$. These conditions imply that $b_1 + b_2 \in I$, and that there exists $A_1 \subseteq F$ such that $A_1, a_1 \vdash b_1$ and there exists $A_2 \subseteq F$ such that $A_2, a_2 \vdash b_2$. Then by Lemma 3 (i) we obtain $A_1, A_2, a_1 + a_2 \vdash b_1 + b_2$. Obviously the set $A = A_1 \cup A_2 \cup \{a_1 + a_2\}$ is a subset of F, such that $A \vdash b_1 + b_2$ which shows that $b_1 + b_2 \in F$. Consequently $F \cap I \neq \varnothing$ - a contradiction with the fact that $F \in M$.

Lemma 13. ***(Separation Lemma.)*** *Let $\underline{B}$ be a S-algebra, A be a finite subset of B, $b \in B$ and $A \nvdash b$. Then there exists a S-clan F such that $A \subseteq F$ and $b \notin F$.*

Proof. By Lemma 10 (i) $\mathbf{F}(A)$ is an S-filter not containing b. Let $(b] = \{x \in B : x \leq b]$ be the smallest ideal containing b. We have $\mathbf{F}(A) \cap (b] = \varnothing$, otherwise by Lemma 3 (iv) we obtain $A \vdash b$ - a contradiction. Then by Lemma 12 There exists a S-clan F such that $\mathbf{F}(A) \subseteq F$ and $F \cap (b] = \varnothing$. This implies $A \subseteq F$ and $b \notin F$.

Lemma 14. *Let $\underline{B}$ be a S-algebra. Then:*

(i) If F is a S-clan in $\underline{B}$, then its complement $\overline{F}$ is an ideal in $\underline{B}$.

(ii) ***(Interpolation lemma for filters and S-clans.)*** *Let G be a S-clan and F be a filter such that $F \subseteq G$. Then there exists an ultrafilter U such that $F \subseteq U \subseteq F$.*

(iii) Let G be a S-clan and $a inG$. Then there exists an ultrafilter U such that $a \in U \subseteq F$.

(iv) If F is a S-clan in $\underline{B}$. Then F coincides with the union of all ultrafilters contained in F.

(v) Let F be a S-clan in $\underline{B}$ and let $ULT(F)$ be the set of all ultrafilters contained in F. Then $ULT(F)$ satisfies the condition ($\sharp$) from Lemma 9.

(vi) Every S-clan can be obtained by the construction described in Lemma 9.

Proof. (i) follows directly by the definitions of S-clan and an ideal.

(ii) Let G be a S-clan and F be a filter such that $F \subseteq G$. Then by (i) $\overline{F}$ is an ideal and $F \cap \overline{F} = \varnothing$. Then by properties of filters and ideals in Boolean algebra, F can be extended into an ultrafilter U such that $U \cap \overline{F} = \varnothing$, which implies that $F \subseteq U \subseteq F$.

(iii) Let $a \in F$ then the filter $[a) = \{b : a \leq b\}$ is contained in F and by (ii) there exists an ultrafilter U such that $[a) \subseteq U \subseteq F$, which implies the statement.

(iv), (v) and (vi) are direct consequences of (iii) and (iv).

4.2 Topological representation theorem for S-algebras

Some topological notions. In this section we will prove a topological representation theorem for S-algebras. For that purpose we remaind some additional facts from topology. Assume that X is a topological space.

- A set of closed subsets of X is called a base for closed sets if every closed set can be represented as an intersection of sets from the base.
- X is a semiregular space if it has a base for closed sets consisting of regular closed sets.

Separation properties of X:

- T_0-separation axiom: X is a T_0 space if for any two different points $x, y \in X$, there is an open set U such that $x \in U$ and $y \notin U$ or $y \in U$ and $x \notin U$.
- T_1-separation axiom: X is a T_1 space if for any two different points $x, y \in X$, there exists two open sets U and V such that $x \in U$ and $y \notin U$, and $y \in V$ and $x \notin V$. X is a T_1 space if and only if every singleton set is closed.
- T_2-separation axiom: X is a T_2 space (Hausdorff space), if for any two different points $x, yinX$, there exists two open sets U and V such that $x \in U$, $y \in V$, and $U \cap V = \varnothing$.

• Compactness. Let I be a nonempty set of indices and let $\Sigma = \{A_i : i \in I\}$ be a family of closed sets of X. We say that Σ possesses finite intersection property if for every finite subset I_0 of I, the intersection $\bigcap\{A_i : i \in I_0\} \neq \varnothing$. X is said to be a compact space if any collection of closed subsets of X with the finite intersection property has nonempty intersection.

Canonical space of S-algebra Let $\underline{B}$ be a S-algebra. Denote by S-Clans(B) the set of S-clans of $\underline{B}$ and let for $a \in b$, $h(a) = \{F \in S-Cland(B) : a \in F\}$. We define **the canonical topological space** (X^c) of $\underline{B}$ to be the set S-Clans(B) with a topology having the set Cl-base$(X^c) = \{h(a) : a \in B\}$ as a base for closed sets of X^c, and the canonical S-algebra related to $\underline{B}$ to be the S-algebra of regular closed sets $RC(X^c)$ as it is defined in Example 4.

Lemma 15. *(i) $h(0) = 0 = \varnothing$, $h(1) = 1 = X^c$, and $h(a+b) = h(a) \cup h(b)$.*
(ii) $a \leq b \Leftrightarrow h(a) \subseteq h(b)$, $a = b$ iff $h(a) = h(b)$.
(iii) $h(a^*) = Cl - h(a)$,
(iv) $h(a)$ is a regular closed set in X^c,
(v) $a_1, \ldots, a_n \vdash b \Leftrightarrow h(a_1) \cap \ldots \cap h(a_n) \subseteq h(b)$.

Proof. (i) follows by the properties of S-clans.

(ii) $(\Rightarrow)$ is easy. For $(\Leftarrow)$ suppose that $a \not\leq b$. Then by the representation theory of Boolean algebras there exists an ultrafilter U such that $a \in U$ and $b \notin U$. Since ultrafilters are S-clans (8, this implies $h(a) \not\subseteq h(b)$.

In order to proof (iii) we need the following equivalence which is true for all S-clans F: $a^* \in F$ iff $(\forall b)(a + b = 1 \to b \in F)$.

For the implication from left to right: suppose $a^* \in F$ and let $a + b = 1$. Then $a^* \leq b$ and hence $b \in F$. For the converse implication let $a^* \notin F$. We have to show that there exists b such that $a + b = 1$ and $b \notin F$. The element $b = a^*$ will do the job.

Proof of (iii). Let F be an arbitrary S-clan. Then: $F \in Cl - h(a)$ iff (by the definition of Cl) $(\forall h(b) \in Cl - base(X^c))(-h(a) \subseteq h(b) \to F \in h(b))$ iff $(\forall b \in B)(h(a) \cup h(b) = X^c \to b \in F)$ iff $(\forall b \in B)(a + b = 1 \to b \in F)$ iff (by the above equivalence) $a^(*) \in F$ iff $F \in h(a^*)$.

(iv) Applying (iii) two times we get: $ClInth(a) = Cl - Cl - h(a) = Cl - h(a^*) = h(a^{**}) = h(a)$, which shows that $h(a)$ is regular closed set and that h is an embedding in the Boolean algebra $RC(X^c)$.

(v) $(\Rightarrow)$ Let $a_1, \ldots a_n \vdash b$ and suppose that $F \in h(a_1), \ldots F \in h(a_n)$. Then $a_1 \in F$, ..., $a_n \in F$ and since F is a S-clan, then $b \in F$. Hence $h(a_1) \cap \ldots \cap h(a_n) \subseteq h(b)$.

$(\Leftarrow)$) This direction follows directly from Lemma 13.

Lemma 16. *Let $\underline{B}$ be a S-algebra. Then the canonical space X^c of $\underline{B}$ is semiregular, T_0 and compact.*

Proof. To prove that X^c is T_0 let Γ and Δ be two different points. Since Γ and Δ are S-clans, one of them, say Γ, is not included in the other one, Δ. Then there is $a \in \Gamma$ and $a \notin \Delta$. Hence the open set $-h(a)$ contains Γ and not Δ.

Since the set $\{h(a) : a \in B\}$ is a base of the closed sets of X^c, to prove the compactness of X^c it suffices to show the following. Let $I \subseteq B$ be a non-empty set of the algebra and let $A = \bigcap\{h(a) : a \in I\}$. Then, if for every finite set $I_0 \subseteq I$ we have $\bigcap\{h(a) : a \in I_0 \neq \varnothing$ then $A \neq \varnothing$. Indeed, the condition that $\bigcap\{h(a) : a \in I_0\}$ for every finite set I_0 of I guarantees the existence of of an ultrafilter U in the Boolean algebra of subsets of X^c containing the set $\{h(a) : a \in I\}$. We claim that the set $\Gamma = \{a : h(a) \in U\}$ is an S-clan. Suppose $a_1, \ldots a_n \in \Gamma$ and $a_1, \ldots a_n \vdash b$. Then $h(a_1) \cap \ldots \cap h(a_n) \subseteq h(b)$ and $h(a_1) \ldots h(a_n \in U)$. Consequently $h(a_1) \cap \ldots \cap h(a_n) \in U$, and since $h(a_1) \cap \ldots \cap h(a_n) \subseteq h(b)$ we get $h(b) \in U$, so $b \in \Gamma$. Thus Γ is a S-filter. Since $h(0) = \varnothing$ is not in U then $0 \notin \Gamma$. Suppose that $a + b \in \Gamma$ then $h(a + b) = h(a) \cup h(b)$ is in U and since U is ultrafilter, then $h(a) \in U$ or $h(b) \in U$. Consequently $a \in \Gamma$ or $b \in \Gamma$, which finally shows that Γ is a S-clan. hence for every $a \in I$

$a \in I \to h(a) \in U \to a \in \Gamma \to \Gamma \in h(a)$.

This shows that $\Gamma \in \bigcap\{h(a) : a \in I\} = A$ and, consequently, $A \neq \varnothing$.

Theorem 1. ***(Topological representation theorem for S-algebras.)*** *Let $\underline{B}$ be a S-algebra. Then there exists a semiregular and compact T_0 space X and an embedding h of $\underline{B}$ into the S-algebra $RC(X)$ of regular closed subsets of X.*

Proof. The proof follows from Lemma 15 and Lemma 16.

4.3 Discrete representation of S-algebras

The aim of this section is to prove that every S-algebra can be isomorphically embedded into the S-algebra over some S-space (see example 5).

Let $\underline{B}$ be a S-algebra and let $ULT(B)$ be the set of ultrafilters in B, S-Clans(B) be the set of all S-clans of B and for each S-clan F let $ULT(F)$ be the set of ultrafilters contained in F. The canonical S-space $S(B) = (X^c, Y^c)$ of B is defined as follows: $X^c = ULT(B)$, $Y^c = \{ULT(F) : F \in$ S-Clans(B)$\}$. The canonical embedding is defined by $g(a) = \{U \in ULT(B) : a \in U\}$. It is a well known fact by the representation theory for Boolean algebras that g is an embedding of B into the Boolean algebra of all subsets of X^c. It remains to show that g preserves the relation $\vdash$. Following the definition of $\vdash$ between subsets of X^c described in Example 5) and the above notations, we have to prove the following equivalence:

$a_1, \dots a_n \vdash b \Leftrightarrow (\forall F_1, \dots F_n \in X^c)(\forall \Gamma \in Y^c)(F_1 \in \Gamma \& \dots \& F_n \in \Gamma \& F_1 \in h(a_1) \& \dots \& F_n \in h(a_n) \rightarrow h(b) \cap \Gamma \neq \varnothing)$.

Note that the left part of the above equivalence just indicates that $g(a_1), \dots g(a_n) \vdash g(b)$ in the S-algebra over (X^c, Y^c).

($\Rightarrow$) Suppose that $a_1, \dots a_n \vdash b$, and that $F_1, \dots F_n$ be S-clans contained is some set $\Gamma \in Y^c$ and $F_1 \in g(a_1)$ (so $a_1 \in F_1$) and ... and $F_n \in g(a_n)$ (so $a_n \in F_n$). By the definition of Y^c, $\Gamma = ULT(F)$ for some S-clan F. Obviously $a_1, \dots a_n \in F$ and since $a_1, \dots a_n \vdash b$, we get that $b \in F$. Then by Lemma 14 (iii) there exists an ultrafilter U such that $b \in U \subseteq F$. From here we obtain that $U \in h(b)$ and $U \in ULT(F) = \Gamma$, so $h(b) \cap \Gamma \neq \varnothing$.

($\Leftarrow$) We will reason by contraposition. Suppose that $a_1, \dots a_n \not\vdash b$. Then by Lemma 13 there exists a S-clan F, such that $a_1, \dots, a_n$ are in F but b is not in F. Since F is the union of ultrafilters contained in F (Lemma 14), then there are ultrafilters F_1,..., F_n in $ULT(F)$ such that $a_1 \in F_1$ (hence $F_1 \in g(a_1)$,..., $a_n \in F_n$ (hence $F_n \in g(a_n)$). Condition $b \notin F$ implies that for all $U \in ULT(F) b \notin U$ which is the same as $g(b) \cap ULT(F) \neq \varnothing$. Denoting $ULT(F)$ by Γ which is in Y^c we get $g(b) \cap \Gamma = \varnothing$. All this implies that $g(a_1), \dots g(a_n) \not\vdash g(b)$ in the S-algebra over (X^c, Y^c) which finishes the proof.

Theorem 2. ***(Discrete Representation Theorem for S-algebras.)*** *Let $\underline{B}$ be a S-algebra. Then there exists a S-space (X, Y) and an embedding g into the S-algebra of all subsets of (X, Y).*

Proof. Take (X, Y) to be the canonical S-space (X^c, Y^c) over $\underline{B}$ and g to be the canonical embedding into the canonical S-algebra over (X^c, Y^c).

5 Concluding remarks

The results of the present paper were reported for the first time in the Workshop "Spatial and Spatio-temporal Logics" organized by Michael Zakharyaschev in 2007 as one of satellite workshops of the international Conference "Algebraic and Topological Methods in Non-Classical Logics III" (TANCL-07), Oxford, England, 5-9 August, 2007. The present text can be considered as an extended abstract, because proofs of many statements are omitted and will be included in the subsequent full version. We plane to include in the full version topological representation theory of some extensions of S-algebras with new axioms, yielding representations in connected spaces and in compact T_1 and T_2 spaces. Applications to the completeness theorems of some spatial logics based on S-systems also will be given.

Bibliography

1. Ph. Balbiani, T. Tinchev and D. Vakarelov, Modal Logics for the Region-based Theory of Space. *Fundamenta Informaticae*, Special Issue: Topics in Logic, Philosophy and Foundation of Mathematics and Computer Science in Recognition of Professor Andrzej Grzegorczyk, vol. (81), (1-3), (2007), 29-82.
2. B. Bennett and I. Düntsch, Axioms, Algebras and Topology. In: *Handbook of Spatial Logics*, M. Aiello, I. Pratt, and J. van Benthem (Eds.), Springer, 2007, 99-160.
3. A. Cohn and J. Renz. Qualitative spatial representation and reasoning. In: F. van Hermelen, V. Lifschitz and B. Porter (Eds.) *Handbook of Knowledge Representation*, Elsevier, 2008, 551-596.
4. G. Dimov, D. Vakarelov, On Scott Consequence Systems, FI, 33 (1998), 43-70
5. G. Dimov and D. Vakarelov. Contact algebras and region–based theory of space: A proximity approach I. *Fundamenta Informaticae*, vol. 74, No 2-3,(2006) 209-249
6. I. Düntsch and D. Vakarelov, Region-based theory of discrette spaces: A proximity approach. In: Nadif, M., Napoli, A., SanJuan, E., and Sigayret, A. EDS, *Proceedings of Fourth International Conference Journées de l'informatique Messine*, 123-129, Metz, France, 2003. Journal version in: *Annals of Mathematics and Artificial Intelligence*, **49**(1-4):5-14, 2007.

7. R. Goldblatt and M. Grice, Mereocompactness and duality for mereotopological spaces. Preprint, 2015.
8. T. Hahmann and M. Gröuninger. *Region-based Theories of Space: Mereotopology and Beyond*, in Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions, edited by Hazarika, S., IGI Publishing, 2012, 1-62.
9. Shyamanta M. Hazarika (Ed.), *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions*, IGI Global; 1 edition (May 31, 2012)
10. T. Ivanova, Extended contact algebras and internal connectednes, In: Proceedings of the 10th Panhellenic Logic Symposium, June 11-15, 2015, Samos, Greece, pages 87-91.
11. R. Kontchakov, I. Pratt-Hartmann, F. Wolter and M. Zakharyaschev, Topology, connectedness and modal logic. in: C. Areces and R. Goldblatt (EDS) *Advances in Modal Logic*, vol. 7, Colege Publications, 2008, pp 151-176.
12. D. S. Scott, Domains for denotational semantics. LNCS 140 (1982)
13. P. Simons, *PARTS. A Study in Ontology*, Oxford, Clarendon Press, 1987.
14. T. Tinchev and D. Vakarelov, Logics of space with connectedness predicates. Complete axiomatizations. In: L. Beklemishev, V. Goranko, V. Shehtman (EDS) *Advances in Modal Logic*, volume 8, College Publications 2010.
15. D. Vakarelov, Information Systems, Similarity Relations and Modal Logics. In: E. Orlowska (Ed.) *INCOMPLETE INFORMATION: ROUGH SET ANALYSIS*, Phisica-Verlag,1998, 492-550.
16. D. Vakarelov, Region-Based Theory of Space: Algebras of Regions, Representation Theory and Logics. In: Dov Gabbay et al. (Eds.) *Mathematical Problems from Applied Logics. New Logics for the XXIst Century.* II. Springer, 2007, 267-348.
17. A. N. Whitehead, *The Organization of Thought Educational and Scientific.* London: Williams & Norgate, 1917.
18. A. N. Whitehead, *Process and Reality*, New York, MacMillan, 1929.
19. Guo-Qiang ZShang, *Logic of Domains*, Boston-Basel-Berlin, 1991

Dynamic Epistemic Logics: Promises, Problems, Shortcomings, and Perspectives

Andreas Herzig

IRIT,
Université de Toulouse,
Centre National de Recherche Scientifique (CNRS),
Institut National Polytechnique (INPT),
Université Paul Sabatier (UPS),
Université Toulouse 1 Capitole (UT1),
Université Toulouse 2 Jean Jaures (UT2J)
IRIT, 118 Route de Narbonne, 31062 Toulouse Cedex 9, France
`http://www.irit.fr/~Andreas.Herzig`

Abstract. Dynamic epistemic logics provide an account of the dynamics of an agent's belief and knowledge. They became popular about 15 years ago and by now there are numerous publications about it. In this paper I will briefly summarize the existing body of literature and discuss some problems and shortcomings.

1 Introduction

Luis was the supervisor of my PhD thesis that I defended more than 25 years ago. The topic was automated theorem proving methods for modal logics, but I actually spent my first year on a quite different subject: logics for database updates. We had come up with a semantics that was close to Winslett's so-called possible models approach [1]. However, despite months of efforts we were neither able to find an axiomatisation nor a syntactical theorem proving method for that semantics, and Luis wisely decided that it was too much a risk to continue on this. Interestingly, the community did not come up with such characterisations either: people seemed to be happy with the semantics and complexity results [2]. The problem however kept on haunting me and I finally succeeded some years later in solving it [3,4].

The main output of the first year of my thesis was an axiomatisation and a theorem proving method for updates by literals. Such updates can be identified with assignments of truth values to propositional variables. In recent and ongoing work together with several colleagues, we have shown that a dynamic logic whose atomic programs are such propositional assignments has multiple applications in AI [5,6].[1] It in particular captures Winslett's possible models approach [8]. It can be extended by epistemic operators [9,10,11], which is relevant for reasoning with incomplete information in multiagent systems.

The above line of work on assignment programs can be seen as being part of an important recent development in the field of non-classical logics: dynamic epistemic logics (DELs). Research on this family of logics blossomed in the field of non-classical logics during the last 15 years. The success story began with an early paper by Plaza [12] that remained basically unnoticed for 10 years and was only taken up by the end of the 90ies, with foundational papers such as Baltag and Moss's [13] and van Benthem's [14], as well as a rather early textbook [15].

The rise of these logics coincides with Luis's involvement in the administration of science, becoming director of IRIT and serving in several positions and committees on the national level. He consequently lacked time to follow the evolution of the by now huge DEL literature. The present contribution is not only an attempt to briefly sum up the state of the art, but also a critical analysis of the roads the field took in the last years. My main message is that progress was perhaps not as overwhelming as one might think when one sees the hundreds of papers that were published under the DEL label. My conclusion will be that several of the old issues are still open: they were overlooked or just left aside. Luis, several old issues we had worked on together should still be topical!

In the next sections I will start by recalling what DELs are (Section 2). The rest of the paper provides a critical examination of the received views about DELs. It is organised according to the three keywords making up the acronym, in reverse order: I will question the status of DELs as logics (Section 3, point out some weaknesses of the commonly assumed epistemic component (Section 4), and finally discuss some issues with the dynamic component (Section 5). Each time I will point out several mismatches between the discourse about DELs and the state of the art; as I will emphasise, several important issues that are difficult to settle received too little attention up to now.

[1] The paper [6] erroneously asserts ExpTime complexity of the model checking problem. A correction is in [7].

2 DELs in a nutshell

I start by a brief overview of DELs. The presentation is standard: language, semantics and axiomatics.

2.1 Language

The DEL language is an interesting and powerful combinations of two kinds of modal operators: epistemic operators of the form K_i where i is an agent and dynamic operators $[\mathcal{E}]$ where $\mathcal{E}$ is an event. The formula $\mathsf{K}_i\varphi$ reads "i knows that φ" and the formula $\langle\mathcal{E}\rangle\varphi$ reads "$\mathcal{E}$ may occur and φ is true afterwards". Such combinations were first considered in AI more than 30 years ago [16] and are particular cases of multi-dimensional modal logics [17]. In its simplest form, the event $\mathcal{E}$ is the *public announcement* of (the truth of) a formula χ, written $\chi!$. In its most general form, $\mathcal{E}$ is made up of a set of possible events having pre- and postconditions, an accessibility relation on that set, and an actual event. When $\langle e, e'\rangle \in S_i$ then agent i cannot distinguish the occurrence of e from that of e'. The precondition of an event describes the conditions under which the event may occur; the postcondition describes its effects on the world in terms of assignments of formulas to propositional variables: when p is assigned to φ then after the event the truth value of p equals the truth value of φ before the event [9]. For instance, when $\neg p$ is assigned to p then p's truth value gets flipped.

Formally, an event model is a tuple $\langle \mathbb{E}, \{S_i\}_{i\in\mathbb{I}}, \mathsf{pre}, \mathsf{post}, e_0\rangle$ where $\mathbb{E}$ is a non-empty set of possible events, $S_i \subseteq \mathbb{E}\times\mathbb{E}$ for every $i \in \mathbb{I}$, $\mathsf{pre} : \mathbb{E} \longrightarrow \mathcal{L}$ and $\mathsf{post} : \mathbb{E} \longrightarrow (\mathbb{P} \longrightarrow \mathcal{L})$, where $\mathcal{L}$ is either the language of epistemic logic or the language of DEL,[2] and $e_0 \in \mathbb{E}$.

Most of the approaches consider events without postconditions, or rather, with the identity postcondition function post_{id} such that $\mathsf{post}_{id}(e)(p) = p$: all variables keep their truth value. Such events have no effect on the world: they are purely epistemic and only change the agents' epistemic state. Figure 1.1 is a typical example of an event model where p is privately announced to agent 1: there are two possible events $\langle p!, \mathsf{post}_{id}\rangle$ and $\langle\top!, \mathsf{post}_{id}\rangle$; agent 2 believes that $\top!$ happens (i.e., that nothing is learned) and that this is common knowledge; agent 1 believes that $p!$ happens while 2 does not know this. The possible event on the left is the actual event.

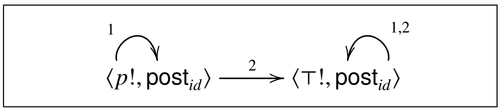


Fig. 1.1. Event model $\mathcal{E}_{1:p!}$ of the private announcement of p to agent 1.

As usual, the formula $[\mathcal{E}!]\varphi$ abbreviates $\neg\langle\mathcal{E}!\rangle\neg\varphi$.

2.2 Semantics

DEL models have accessibility relations for the epistemic operators, one per agent i. Each relation R_i relates worlds that i cannot distinguish based on her knowledge, and $\mathsf{K}_i\varphi$ is true at a possible world w in a model M if φ is true at every world that i cannot distinguish from w in M. While the interpretation of the epistemic operators is thus standard, the interpretation of the dynamic operators deviates from the traditional modal logic setting: there is no accessibility relation for them. So DEL models are nothing but models of the underlying 'static' epistemic logic: tuples of the form $M = \langle W, \{R\}_{i\in\mathbb{I}}, V, w_0\rangle$ where W is a non-empty set of possible worlds, $R_i \subseteq W \times W$ is a binary relation on W for every agent i in the set of agents $\mathbb{I}$, $V : \mathbb{P} \longrightarrow 2^W$ is a valuation function associating to every p in the set of propositional variables $\mathbb{P}$ the set of worlds where p is true, and $w_0 \in$ is the actual world.

Figure 1.2 contains an example of a DEL model where both agent 1 and agent 2 do not know whether p is true or not (and this is common knowledge). The actual world is the lower world $\{p\}$.

The truth condition for the epistemic operators is the standard one: for a model $M = \langle W, \{R\}_{i\in\mathbb{I}}, V, w_0\rangle$,

$$M \Vdash \mathsf{K}_i\varphi \text{ if } \langle W, \{R\}_{i\in\mathbb{I}}, V, w\rangle \Vdash \varphi \text{ for every } w \text{ such that } \langle w_0, w\rangle \in R_i.$$

The dynamic operators are interpreted by means of so-called *model updates*. The simplest case is the model update that is associated to the public announcement $\chi!$. It is interpreted by a partial function on the set

[2] Precisely, the postcondition is either restricted to epistemic formulas or defined by mutual recursion with the truth conditions.

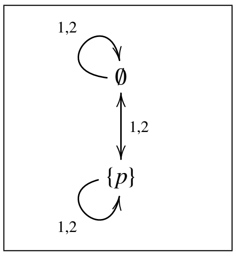


Fig. 1.2. Static epistemic model M_{ign} where both agent 1 and agent 2 do not know whether p (and this is common knowledge).

of pointed Kripke models which restricts the set of possible worlds W of a model M to $\|\chi\|_M$.[3] The accessibility relation and the valuation are restricted in consequence. So when $M \Vdash \mathsf{pre}(e)$ then[4] the update of the epistemic model $M = \langle W, \{R\}_{i \in \mathbb{I}}, V, w_0 \rangle$ by χ is the epistemic model

$$M^{\chi} = \langle \|\chi\|_M, \{R_i^{\chi}\}_{i \in \mathbb{I}}, V^{\chi}, w_0 \rangle$$

where $R_i^{\chi} = R_i \cap (\|\chi\|_M \times \|\chi\|_M)$ and $V^{\chi}(p) = V(p) \cap \|\chi\|_M$. (Precisely, updates are defined by mutual recursion with the truth conditions.) The truth condition for the public announcement operator is:

$$M \Vdash \langle \chi! \rangle \varphi \;\text{ if }\; M \Vdash \chi \text{ and } M^{\chi} \Vdash \varphi.$$

In its most general form, updates are a restricted product between an epistemic model and an event model: the formula $\langle \mathcal{E} \rangle \varphi$ is true at possible world w in model M if the actual event of $\mathcal{E}$ may occur at w in M and φ is true at w in the update $M^{\mathcal{E}}$ of M by $\mathcal{E}$. Formally, for $\mathcal{E} = \langle \mathbb{E}, \{S_i\}_{i \in \mathbb{I}}, \mathsf{pre}, \mathsf{post}, e_0 \rangle$:

$$M \Vdash \langle \mathcal{E} \rangle \varphi \;\text{ if }\; M \Vdash \mathsf{pre}(e_0) \text{ and } M^{\mathcal{E}} \Vdash \varphi$$

where $M^{\mathcal{E}} = \langle W^{\mathcal{E}}, \{R^{\mathcal{E}}\}_{i \in \mathbb{I}}, V^{\mathcal{E}}, (w_0, e_0) \rangle$ with

$$\begin{aligned} W^{\mathcal{E}} &= \{(w, e) \;:\; w \in W,\; e \in \mathbb{E},\; \text{and } M, w \Vdash \mathsf{pre}(e)\} \\ R_i^{\mathcal{E}} &= \{\langle (w, e), (w', e') \rangle \;:\; \langle w, w' \rangle \in R_i \text{ and } \langle e, e' \rangle \in S_i\} \\ V^{\mathcal{E}}(p) &= \{(w, e) \in W^{\mathcal{E}} \;:\; M, w \Vdash \mathsf{post}(e)(p)\} \end{aligned}$$

For example, the update of the static epistemic model M_{ign} of Figure 1.2 by the event model $\mathcal{E}_{1:p!}$ of Figure 1.1 is the static epistemic model $(M_{ign})^{\mathcal{E}_{1:p!}}$ depicted in Figure 1.3. The actual world is on the left. Observe that $(M_{ign})^{\mathcal{E}_{1:p!}}$ does not contain a world $(\emptyset, \langle p!, \mathsf{post}_{id} \rangle)$ because $M_{ign} \nVdash \mathsf{pre}(p!)$. We have $(M_{ign})^{\mathcal{E}_{1:p!}} \Vdash \mathsf{K}_1 p \wedge \neg \mathsf{K}_2 p \wedge \neg \mathsf{K}_2 \mathsf{K}_1 p$. Therefore $M_{ign} \Vdash [\mathcal{E}_{1:p!}](\mathsf{K}_1 p \wedge \neg \mathsf{K}_2 p \wedge \neg \mathsf{K}_2 \mathsf{K}_1 p)$.

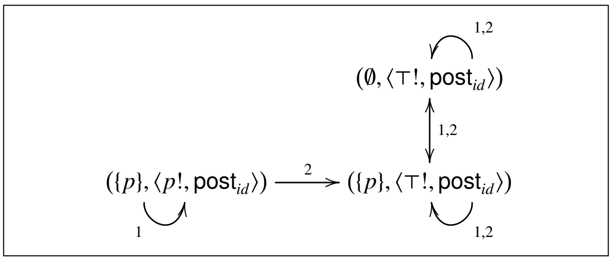


Fig. 1.3. The static epistemic model $(M_{ign})^{\mathcal{E}_{1:p!}}$, after agent 1 has privately learned that p.

[3] As usual, for a given $M = \langle W, \{R\}_{i \in \mathbb{I}}, V, w_0 \rangle$, the notation $\|\chi\|_M$ stands for the set of possible worlds $w \in W$ such that $\langle W, \{R\}_{i \in \mathbb{I}}, V, w \rangle \Vdash \varphi$.

[4] Otherwise the operation $(.)^{\mathcal{E}}$ is undefined.

`EL`	some axiomatics of epistemic logic (e.g. that of `S5`)
`RE`($[\mathcal{E}]$)	$\dfrac{\psi \leftrightarrow \psi'}{[\mathcal{E}]\psi \leftrightarrow [\mathcal{E}]\psi'}$
`Red(p)`	$[\mathcal{E}]p \ \leftrightarrow\ (\mathsf{pre}(e) \to \mathsf{post}(e)(p))$, for p atomic
`Red(¬)`	$[\mathcal{E}]\neg\psi \ \leftrightarrow\ (\mathsf{pre}(e) \to \neg[\mathcal{E}]\psi)$
`Red(∧)`	$[\mathcal{E}](\psi_1 \wedge \psi_2) \ \leftrightarrow\ ([\mathcal{E}]\psi_1 \wedge [\mathcal{E}]\psi_2)$
`Red(K_i)`	$[\mathcal{E}]\mathsf{K}_i\psi \ \leftrightarrow\ (\mathsf{pre}(e) \to \bigwedge_{\langle e,e'\rangle \in S_i} \mathsf{K}_i[\mathcal{E}']\psi)$

Table 1.1. Axiomatisation of DELs, where we suppose that $\mathcal{E} = \langle \mathbb{E}, \{S_i\}_{i\in\mathbb{I}}, \mathsf{pre}, \mathsf{post}, e\rangle$ and $\mathcal{E}' = \langle \mathbb{E}, \{S_i\}_{i\in\mathbb{I}}, \mathsf{pre}, \mathsf{post}, e'\rangle$.

2.3 Axiomatics

An axiomatisation of DELs that is parametrised by the axiomatisation of the underlying epistemic logic is contained in Table 1.1. (It is easy to see that it is equivalent to the axiomatisations one usually finds in papers about DELs, such as in [18,19].) It consists in a complete collection of *reduction axioms*: equivalences whose successive application allows to eliminate all dynamic operators. We therefore end up with an epistemic formula.[5]

3 Logic?

The standard definition of a modal logic is that of a set of modal formulas that contains all classical propositional theorems and that is closed under uniform substitution, modus ponens, and necessitation. This is a rather restrictive definition which excludes DELs. Indeed, they fail to be closed under uniform substitution.[6]

Let us adopt a more liberal position and accept as logics sets of formulas that are not closed under uniform substitution. DELs fail to satisfy a further, fundamental requirement for logics, viz. that language and semantics should be distinct that are only linked via the interpretation function. DELs with event models violate this principle: the event model $\mathcal{E}$ in the formula $[\mathcal{E}]\varphi$ is a semantical object. This was felt to be a problem right from the start, and several proposals for a language allowing to talk about event models were put forward [20,21,22,23,24,25,26]. However, it is only recently that a solution was proposed that I find satisfactory [27].

Finally, there is a further requirement that is natural for logics extending epistemic logic by dynamic operators: the dynamic extension should be a *conservative extension* of the underlying epistemic logic. As will be show in Section 5, this unfortunately fails to hold for the most relevant underlying epistemic logics.

4 Epistemic?

In the literature it is—sometimes explicitly and sometimes tacitly—supposed that DELs provide a satisfactory formalisation of an agent's representation of the world and its evolution. In this section I undertake a critical examination of this claim.

In DELs, the term 'epistemic' refers to an agent's representation of the world. It is understood in a broad sense, not only covering knowledge, but also belief. My main target are the various logics of knowledge and logics of belief underlying the DELs that one can find in the literature. I first discuss logics of knowledge and then logics of belief.

4.1 Knowledge

Following Halpern et al., many authors chose `S5` as their 'official' logic of knowledge. This is however at odds with the philosophical logic literature, where Lenzen and Voorbraak put forward strong arguments against the negative introspection principle of S5 and argued for the weaker modal logic `S4.3`. As I have shown elsewhere together with Philippe Balbiani and Tiago de Lima [28], the replacement of `S5` as the logic of knowledge underlying DEL by `S4.3` leads to serious technical difficulties—failure of preservation of the class of models under relativisation—for which there is currently no good solution.

[5] Observe that there is no reduction axiom for the case of two successive dynamic operators. While it part of the standard axiomatisations of the literature, it is actually not necessary in the presence of the rule of equivalence `RE`($[\mathcal{E}]$).

[6] To see this consider the formula $[p!]p$: it is a theorem of `PAL` because its rewriting with Axiom `Red(p)` results in the classical logic theorem $p \to p$.

4.2 Belief

As to belief, the situation is worse. Some authors choose the fairly uncontroversial logic of belief KD45 as a basis for their DEL. However, dynamic extensions of KD45 face the same technical problems as dynamic extensions of S4.3: the class of models is not preserved under relativisation.

That problem is avoided if one opts for the basic modal logic K as the underlying epistemic logic: K models do not have any constraint to satisfy, every relativisation of a K model is trivially a K model. However, K is a very weak logic of belief allowing an agent to simultaneously hold contradictory beliefs.[7] It was already highlighted by Hintikka [29] that beyond closure under logical truth and modus ponens, consistency is a fundamental property of rational belief. Moreover, epistemic logics allowing for inconsistent beliefs neglect that even if the beliefs of a human or other 'real' agent may sometimes become inconsistent, such agents nevertheless strive to maintain consistency. This leads us to the next problem that DEL extensions of logics of belief face.

I believe that one of the biggest problems for DELs is that currently there are no good accounts of multi-agent belief revision that could be integrated. Indeed, while there is a rich literature on belief revision since the seminal 1985 'AGM' paper by Alchourrón, Gärdenfors and Makinson [30]—see [31] for an overview—, it is fair to say that all of the resulting formalisms are complex and cannot be easily extended beyond classical propositional logic.[8] The situation is worse than in propositional logic because DEL belief revision takes place in the framework of multimodal logics. Indeed, while knowledge is always true (because knowledge implies truth), an agent's beliefs may be false. In PAL, an agent may wrongly believe that some announcement ψ cannot be made because he wrongly believes ψ to be false. In such circumstances, only the left-to-right direction $[\psi!]\mathsf{B}_i\varphi \to (\psi \to \mathsf{B}_i[\psi!]\varphi)$ of the (belief version of the) reduction axiom $\mathtt{Red}(\mathsf{K}_\mathtt{i})$ of Table 1.1 is valid, and trivially so, while the right-to-left direction $(\psi \to \mathsf{B}_i[\psi!]\varphi) \to [\psi!]\mathsf{B}_i\varphi$ is not. While it is clear that a satisfactory doxastic version of PAL clearly should integrate some notion of belief revision, it is basically still an open problem how to do this. Actually most of the DEL extensions of logics of belief fail to account for multiagent belief revision. There exist a few proposals for an integration [32,33,34,35,24,36]. However, it can be argued that all of them are problematic, either as far as foundations or as far as implementability is concerned. One of the problems that has not been addressed up to now in a satisfactory manner is that almost all belief revision theories presuppose some kind of preference information. However, in many cases it is not clear at all where this information comes from. In my opinion, the challenge is to find simpler, probably more modest theories of revision that can be easily and smoothly integrated into logics of belief and action.

People in the field seem to consider that Baltag and Smets's proposal [36] is currently the best solution. However, their account is grounded on the notion of *safe belief*, which is 'almost knowledge': belief that will never be revised. The problem here is that it is not clear how an autonomous agent could ever distinguish safe beliefs from other, non-safe beliefs.

5 Dynamic?

The evolution of an agent's epistemic state has two different causes: first, the agent may get a new piece of information about a static world; second, the world may evolve and our agent observes this event, possibly only in an imperfect way.

5.1 Evolution of the world

The simplest form of the second kind of change, evolution of the world, is when it is the truth value of some propositional variable that changes. This captured in DELs by the postcondition of an event e in an event model $\mathcal{E}$, alias the assignment of a propositional variable at e. Previous accounts of such kinds of change include STRIPS [37] and Reiter's basic action theories [38]. Somewhat surprisingly and as also pointed out by van Benthem [39], the DEL literature largely ignores the existing literature on reasoning about ontic actions. In particular, it is well-known in that field that formalisms are plagued by three tenacious problems: the frame problem, the qualification problem and the ramification problem. The frame problem is how to specify the non-effects of an event. The qualification problem is how to specify the preconditions of an event. The ramification problem is to take into account domain constraints (alias integrity constraints).

[7] Some may still object that this is too strong a logic of belief because in K, an agent's beliefs are closed under logical truth and modus ponens. This is the so-called omniscience problem. We leave it aside here: omniscience can be assumed when we are interested in rational agents or in artificial agents.

[8] Actually the public announcement operators behave just as the expansion operation in AGM theory, which is the simplest belief change operation.

As to the frame problem, I will just say here that DELs' assignments elegantly solve it, as was shown in [10,39].

As to the qualification problem and the ramification problem, it seems that no DEL paper has tried to address them. While some of the approaches could probably be imported (such as [40,41,42,43], one should nevertheless note that—contrarily to the frame problem, where Reiter's solution was largely adopted due to its simplicity—no consensual solution to these two problems exists in the reasoning about actions field.

5.2 Evolution of the agents' epistemic state

The first kind of change is that of a a private announcement to a group of agents J: some formula φ is announced to J in a way that is public for J. In DELs such an announcement is identified with the precondition of an event that takes place; more precisely, for which it is *possible for the agent* that it takes place. This is captured in DEL event models by an event with precondition φ and and an empty postcondition.

This is related to a claim that is often made in presentations (if not in papers), viz. that DELs are logics of communication. However, DELs lack several key ingredients of speech act theory. To start with, it is not easy to come up with a meaningful notion of a speaker. Furthermore, DELs currently do not provide a good account of communicative intention. Unless such concepts can be integrated into DELs, the field will have to wait for a good logic of multiagent systems.

6 Perspectives

Despite the various criticisms that I have put forward in the preceding sections, I believe that DELs are one of the most fruitful recent developments in the domain of nonclassical logics. One of its most striking assets is that their models are very compact. This contrasts with standard temporal and dynamic models whose Kripke models typically contain a huge number of possible worlds even for rather simple applications. This makes that model checking procedures working with Kripke models as they stand are not practically feasible.

It however remains that Kripke models for epistemic logics can become pretty big, too, in particular when there are multiple agents. In the literature on model checking for multiagent systems one can find more compact representations where the epistemic accessibility relation is built from information about observability of propositional variables by agents [44,45]. This perspective was recently imported into the DEL setting [46,47,48,49,50,51,52], and I believe it to be a promising research avenue.

As to the belief revision problem plaguing DELs, it is not clear to me whether and how it will be addressed in the future. A solution that should be achievable without too much effort is to combine the embedding of Dalal's revision operation of [8] with the embedding of the observation-based DEL of [51,52].

References

1. Winslett, M.A.: Reasoning about action using a possible models approach. In: Proc. 7th Conf. on Artificial Intelligence (AAAI'88), St. Paul (1988) 89–93
2. Eiter, T., Gottlob, G.: On the complexity of propositional knowledge base revision, updates, and counterfactuals. Artificial Intelligence J. **57** (1992) 227–270
3. Herzig, A.: The PMA revisited. In Aiello, L.C., Shapiro, S., eds.: Proc. Int. Conf. on Knowledge Representation and Reasoning (KR'96), Morgan Kaufmann (1996) 40–50
4. Herzig, A.: Logics for belief base updating. In Dubois, D., Gabbay, D., Prade, H., Smets, P., eds.: Handbook of defeasible reasoning and uncertainty management. Volume 3 - Belief Change. Kluwer Academic Publishers (1998) 189–231
5. Herzig, A., Lorini, E., Moisan, F., Troquard, N.: A dynamic logic of normative systems. In Walsh, T., ed.: International Joint Conference on Artificial Intelligence (IJCAI), Barcelona, IJCAI/AAAI (2011) 228–233
6. Balbiani, P., Herzig, A., Troquard, N.: Dynamic logic of propositional assignments: a well-behaved variant of PDL. In Kupferman, O., ed.: Logic in Computer Science (LICS), New Orleans, June 25-28, 2013, http://www.ieee.org/, ACM/IEEE (2013) 143–152
7. Balbiani, P., Herzig, A., Schwarzentruber, F., Troquard, N.: DL-PA and DCL-PC: model checking and satisfiability problem are indeed in PSPACE. CoRR **abs/1411.7825** (2014)
8. Herzig, A.: Belief change operations: a short history of nearly everything, told in dynamic logic of propositional assignments. In Baral, C., De Giacomo, G., eds.: Proc. KR 2014, AAAI Press (2014)
9. van Ditmarsch, H.P., van der Hoek, W., Kooi, B.P.: Dynamic epistemic logic with assignment. In Dignum, F., Dignum, V., Koenig, S., Kraus, S., Singh, M.P., Wooldridge, M., eds.: AAMAS, ACM (2005) 141–148
10. van Ditmarsch, H.P., Herzig, A., de Lima, T.: From Situation Calculus to Dynamic Logic. Journal of Logic and Computation **21**(2) (2011) 179–204
11. van Ditmarsch, H.P., Herzig, A., De Lima, T.: Public announcements, public assignments and the complexity of their logic. Journal of Applied Non-Classical Logics **22**(3) (2012) 249–273

12. Plaza, J.A.: Logics of public communications. In Emrich, M., Pfeifer, M., Hadzikadic, M., Ras, Z., eds.: Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems, 201-216 (1989)
13. Baltag, A., Moss, L.S.: Logics for epistemic programs. Synthese **139 (2)** (2004) 165–224
14. van Benthem, J.: One is a lonely number: on the logic of communication. In Chatzidakis, Z., Koepke, P., Pohlers, W., eds.: Logic Colloquium'02. ASL & A.K. Peters, Wellesley MA (2006) 96–129 Tech Report PP-2002-27, ILLC Amsterdam (2002).
15. van Ditmarsch, H., Van Der Hoek, W., Kooi, B.: Dynamic epistemic logic. Volume 337. Springer Verlag (2007)
16. Moore, R.C.: A formal theory of knowledge and action. In Hobbs, J., Moore, R., eds.: Formal Theories of the Commonsense World. Ablex, Norwood, NJ (1985) 319–358
17. Gabbay, D.M., Kurucz, A., Wolter, F., Zakharyaschev, M.: Many-Dimensional Modal Logics: Theory and Applications. Number 148 in Studies in Logic and the Foundations of Mathematics. Elsevier, North-Holland (2003)
18. Wang, Y., Cao, Q.: On axiomatizations of public announcement logic. Synthese **190**(Supplement-1) (2013) 103–134
19. Wang, Y., Aucher, G.: An alternative axiomatization of DEL and its applications. In: IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013. (2013)
20. Gerbrandy, J., Groeneveld, W.: Reasoning about information change. J. of Logic, Language and Information **6**(2) (1997)
21. Gerbrandy, J.: Bisimulations on Planet Kripke. PhD thesis, University of Amsterdam (1999)
22. van Ditmarsch, H.P.: Knowledge games. PhD thesis, Groningen University, ILLC Dissertation Series, Grafimedia (2000)
23. van Ditmarsch, H.P.: Descriptions of game actions. J. of Logic, Language and Information (JoLLI) **11** (2002) 349–365
24. Aucher, G.: Perspectives on Belief and Change. PhD thesis, University of Otago and University Paul Sabatier (2008)
25. Aucher, G.: BMS revisited. In: Proceedings of Theoretical Aspects of Rationality and Knowledge (TARK 2009). (2009) 24–33
26. Bolander, T., van Ditmarsch, H., Herzig, A., Lorini, E., Pardo, P., Schwarzentruber, F.: Announcements to Attentive Agents. Journal of Logic, Language and Information (2016)
27. Hales, J., Tay, E., French, T.: A composable language for action models. In Goré, R., Kurucs, A., eds.: Advances in Modal Logic, College Publications (2014)
28. Balbiani, P., van Ditmarsch, H.P., Herzig, A., de Lima, T.: Some truths are best left unsaid. In Bolander, T., Braüner, T., Ghilardi, S., Moss, L.S., eds.: Advances in Modal Logic, College Publications (2012) 36–54
29. Hintikka, J.: Knowledge and belief. Cornell University Press, Ithaca, N.Y. (1962)
30. Alchourrón, C., Gärdenfors, P., Makinson, D.: On the logic of theory change: Partial meet contraction and revision functions. J. of Symbolic Logic **50** (1985) 510–530
31. Fermé, E.L., Hansson, S.O.: AGM 25 years - twenty-five years of research in belief change. J. Philosophical Logic **40**(2) (2011) 295–331
32. Aucher, G.: A combined system for update logic and belief revision. In Barley, M., Kasabov, N.K., eds.: PRIMA. Volume 3371 of Lecture Notes in Computer Science., Springer (2004) 1–17
33. van Ditmarsch, H.P.: Prolegomena to dynamic logic for belief revision. Synthese **147**(2) (2005) 229–275
34. van Benthem, J.: Dynamic logic for belief revision. Journal of Applied Non-Classical Logics **17**(2) (2007) 129–155
35. Aucher, G.: Internal models and private multi-agent belief revision. In Padgham, L., Parkes, D.C., Müller, J.P., Parsons, S., eds.: AAMAS (2), IFAAMAS (2008) 721–727
36. Baltag, A., Smets, S.: Group belief dynamics under iterated revision: fixed points and cycles of joint upgrades. In Heifetz, A., ed.: TARK. (2009) 41–50
37. Fikes, R., Nilsson, N.J.: STRIPS: A new approach to the application of theorem proving to problem solving. Artif. Intell. **2**(3/4) (1971) 189–208
38. Reiter, R.: Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems. MIT Press (2001)
39. van Benthem, J.: McCarthy variations in a modal key. Artif. Intell. **175**(1) (2011) 428–439
40. Thielscher, M.: Computing ramifications by postprocessing. In: Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95), Montreal, Canada (1995) 1994–2000
41. Thielscher, M.: Causality and the qualification problem. (1996) 51–62
42. Thielscher, M.: Ramification and causality. Artificial Intelligence J. **89** (1997) 317–364
43. Thielscher, M.: Qualified ramifications. (1997) 466–471
44. Lomuscio, A., van der Meyden, R., Ryan, M.: Knowledge in multiagent systems: initial configurations and broadcast. ACM Trans. Comput. Log. **1**(2) (2000) 247–284
45. Su, K., Sattar, A., Luo, X.: Model checking temporal logics of knowledge via OBDDs. Comput. J. **50**(4) (2007) 403–420
46. van der Hoek, W., Iliev, P., Wooldridge, M.: A logic of revelation and concealment. In van der Hoek, W., Padgham, L., Conitzer, V., Winikoff, M., eds.: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS (2012) 1115–1122
47. Balbiani, P., Gasquet, O., Schwarzentruber, F.: Agents that look at one another. Logic Journal of the IGPL **21**(3) (2013) 438–467
48. Herzig, A., Lorini, E.: A modal logic of perceptual belief. In Lihoreau, F., Rebuschi, M., eds.: Epistemology, Context, and Formalism. Springer (2014) 197–211

49. Charrier, T., Schwarzentruber, F.: Arbitrary public announcement logic with mental programs. In: Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), International Foundation for Autonomous Agents and Multiagent Systems (2015) 1471–1479
50. van Benthem, J., van Eijck, J., Gattinger, M., Su, K.: Symbolic model checking for Dynamic Epistemic Logic. In van der Hoek, W., Holliday, W.H., Wang, W.f., eds.: Proceedings of the 5th International Conference on Logic, Rationality and Interaction, Springer-Verlag (2015) 366–378
51. Herzig, A., Lorini, E., Maffre, F.: A poor man's epistemic logic based on propositional assignment and higher-order observation (regular paper). In: International Conference on Logic, Rationality and Interaction (LORI), Taipei, 28/10/2015-31/10/2015, http://www.springer.com, Springer Berlin / Heidelberg (2015)
52. Herzig, A., Maffre, F.: How to share knowledge by gossiping. In Julian Inglada, V.J., Vouros, G., eds.: 3rd International Conference on Agreement Technologies, Athens. LNCS/LNAI, http://www.springerlink.com/, Springer-Verlag (2015)

On the Relation between Possibilistic Logic and Modal Logics of Belief

Mohua Banerjee[1], Didier Dubois[2], Lluis Godo[3], and Henri Prade[2]

[1] Dept. of Mathematics and Statistics, Indian Institute of Technology, Kanpur, India
[2] IRIT-CNRS, Université Paul Sabatier, Toulouse, France
[3] IIIA-CSIC, Bellaterra, Spain

Abstract. Possibility theory and modal logic are two knowledge representation frameworks that share some common features, such as the duality between possibility and necessity, as well as some obvious differences since possibility theory is graded but is not primarily a logical setting. In the last thirty years there have been a series of attempts, reviewed in this paper, for bridging the two frameworks in one way or another. Possibility theory relies on possibility distributions and modal logic on accessibility relations, at the semantic level. Beyond the observation that many properties of possibility theory have qualitative counterparts in terms of axioms of well-known modal logic systems, the first works have looked for (graded) accessibility relations that can account for the behavior of possibility and necessity measures. More recently, another view has emerged from the study of logics of incomplete information, which is no longer based on Kripke-like models. On the one hand, possibilistic logic, closely related to possibility theory, mainly handles beliefs having various strength. On the other hand, in the so-called meta-epistemic logic (MEL) an agent can express both beliefs and explicitly ignored facts (both without strength), by only using modal formulas of depth 1, and no objective ones; its semantics is based on epistemic states. The system MEL$^+$ is an extension of MEL having the syntax of S5. Generalized possibilistic logic (GPL) extends both possibilistic logic and MEL, and has a semantics in terms of sets of possibility distributions. After a survey of these different attempts, the paper presents GPL$^+$, a graded counterpart of MEL$^+$ that extends MEL by allowing objective (sub)formulas. The axioms of GPL$^+$ are graded counterparts of those of S5 modal system, the semantics being based on pairs made of an interpretation (representing the real state of facts) and a possibility distribution (representing an epistemic state). Soundness and completeness are established. The paper also discusses the difference with S5 used as a logic for rough sets that accounts for indiscernibility rather than incomplete information, using also the square of opposition as a common structure underlying modal logic, possibility theory, and rough set theory.

Keywords: Modal logic, possibility theory, epistemic logic, rough sets

1 Introduction

Possibility theory has been introduced by Zadeh [54] as a framework for representing the uncertainty conveyed by linguistic statements. It is based on the notion of possibility distribution π, from which a maxitive possibility measure $\Pi(A)$ is defined as a consistency degree between this distribution representing the available information and the considered event A. This proposal is formally similar to, although fully independent of the one previously developed in economics by Shackle [50] based on the notion of degree of surprise (which corresponds to impossibility).

Although possibility theory has been the basis of an original approximate reasoning theory [56], this setting is not a logical setting strictly speaking. It is only later, in the 1980's, that possibilistic logic, a logic of classical logic formulas associated with certainty levels (thought as lower bounds of a necessity measure) has emerged (see [15,18] for introductions and overviews). Still, in the setting of his representation language PRUF [55] Zadeh discusses the representation of statements of the form "X is A" (meaning that the possible values of the single-valued variable X are fuzzily restricted by fuzzy set A) linguistically qualified in terms of truth, probability, or possibility. Interestingly enough, the representation of possibility-qualified statements led to possibility distributions over possibility distributions, but certainty-qualified statements were not considered at all, just because necessity measures as dual of possibility measures were playing almost no role in Zadeh's view (with the exception of half a page in [57]). Certainty-qualified statements were first considered in [45], and rediscussed in [14] in relation with two resolution principles (respectively involving two certainty-qualified propositions, and one certainty-qualified proposition together with a possibility qualified proposition), whose formal analogy with the inference rules existing in modal logic was stressed.

Such an analogy between possibility theory calculus (including necessity measures) and modal logic was not coming as a surprise since the parallels between $N(A) = 1 - \Pi(A)$ and $\Box p \leftrightarrow \neg\Diamond\neg p$ (duality between necessity and possibility), between $N(A) \leq \Pi(A)$ and $\Diamond p \to \Box p$ (axiom D in modal logic systems), or between the characteristic axiom of necessity measures $N(A \cap B) = \min(N(A), N(B))$ and $(\Box p \wedge \Box q) \leftrightarrow \Box(p \wedge q)$ (a theorem valid in modal system K) had been already noticed. Nevertheless, no formal connection between modal logic and possibility theory existed in those days, even if the idea of graded accessibility relations had been already proposed independently [32] [49] some years before.

The striking parallel between possibility theory and modal logic eventually led to proposals for a modal analysis and encoding of possibility theory, one of which by L. Fariñas and A. Herzig [25], later by Boutilier [5], then extended to multiple-valued propositions [29]. Another more semantically-oriented trend was to build particular accessibility relations [22][31] agreeing with possibility theory. The work in [36,37,38,35] is also worth-mentioning in that respect.

Rather than putting possibility theory under the umbrella of (graded) modal logics, a quite different view has finally emerged by designing a logical system closer to classical logic capable of handling simple certainty- or possibility-qualified statements. This epistemic logic is a two-tiered propositional logic (an idea that first appears in [16]) where propositional combinations of modal formulas of depth 1 can be handled. The resulting logic, called meta-epistemic logic (MEL), when necessity and possibility are binary-valued, proved to be equivalent to a fragment of the normal modal logic system KD [1,3]. MEL can be extended to graded modalities, thus extending possibilistic logic [33,11] (where only conjunctions of certainty- or possibility-qualified statements are allowed) to a generalized possibilistic logic (GPL) [20], where negation and disjunctions of weighted formulas are allowed . The semantics of MEL (resp. GPL) is no longer expressed by means of an accessibility relation, but in terms of a set of sets of models (resp. a set of possibility distributions), which agrees with Zadeh's original semantical view of possibility-qualified statements (applied in his case to linguistic degrees of possibility and thus leading to a fuzzy set of possibility distributions).

MEL has been more recently extended to MEL$^+$ [2] where propositional combinations of objective formulas and modal formulas of depth 1 are allowed. These formulas are semantically evaluated by pairs made of one interpretation (representing the real state of facts) and a non-empty set of interpretations (representing an epistemic state). The axioms of MEL$^+$ are those of propositional logic, modal axioms K (distributivity), and D, plus $\Box p$ if p is a tautology, while MEL^{++} also includes axiom T ($\Box p \to p$). MEL$^+$ and MEL^{++} are respectively equivalent to modal systems KD45 and S5. The purpose of this paper is to extend such a construct to GPL.

The paper is structured as follows. The next two sections provides a detailed background organized in several subsections. Section 2 first covers a square of opposition-based view of modal logic, possibility theory, and rough sets whose logic obey the axioms of modal system S5. Then Section 2 surveys early attempts at bridging possibility theory and modal logics. Section 3 offers overviews of MEL, MEL$^+$ and generalized possibilistic logic. Section 4 is dedicated to the joint extension of MEL$^+$ and GPL in GPL$^+$, and then to the joint extension of MEL^{++} and GPL$^+$ in GPL^{++}; soundness and completeness results are established.

2 Background

This background section is organized into two pieces. First, we indicate how the square of opposition captures and exhibits the roots of the formal similarities underlying modal logic, possibility theory, and rough sets. Then different early attempts at bridging possibility theory and modal logic are reviewed.

2.1 Possibility theory, rough sets and modal logics: a square of opposition viewpoint

Recent studies [19] have pointed out that many artificial intelligence knowledge representation settings are sharing the same structures of opposition that extend or generalize the traditional square of opposition which dates back to Aristotle, and whose logical interest has been rediscovered more than one decade ago [4]. The traditional square involves four logically related statements exhibiting universal or existential quantifications: a statement **A** of the form "every x is p" is negated by the statement **O** "some x is not p", while a statement like **E** "no x is p" is clearly in even stronger opposition to the first statement (**A**). These three statements, together with the negation of the last one, namely **I** "some x is p", give birth to the Aristotelian square of opposition in terms of quantifiers **A** : $\forall x\ p(x)$, **E** : $\forall x\ \neg p(x)$, **I** : $\exists x\ p(x)$, **O** : $\exists x\ \neg p(x)$. This square, pictured in Fig. 1.1, is usually denoted by the letters **A**, **I** (affirmative half) and **E**, **O** (negative half). The names of the vertices come from a traditional Latin reading: **A**ff**I**rmo, n**E**g**O**).

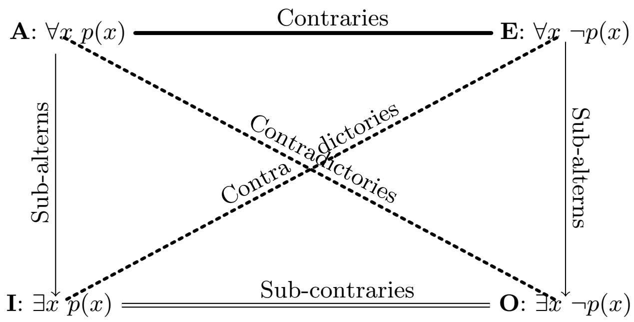


Fig. 1.1. Square of opposition

Note that we assume that some x do exist, thus avoiding existential import problems in Fig. 1.1. The different edges and diagonals of the square exhibits simple logical relations: i) **A** and **O**, as well as **E** and **I** are contraries; ii) **A** entails **I**, and **E** entails **0**; iii) **A** and **E** cannot be true together, while iv) **I** and **O** cannot be false together.

Another well-known instance of this square is in terms of the *necessary* ($\Box$) and *possible* ($\Diamond$) modalities, with the following reading **A** : $\Box p$, **E** : $\Box\neg p$, **I** : $\Diamond p$, **O** : $\Diamond\neg p$, where $\Diamond p =_{def} \neg\Box\neg p$ (with $p \neq \bot, \top$). Then the entailment from **A** to **I** is nothing but the axiom (D) in modal logic, namely $\Box p \to \Diamond p$. This reading has an easy counterpart in terms of binary-valued possibility theory replacing $\Box p$ by $N([p])$ and $\Diamond p$ by $\Pi([p])$ where $[p]$ is the set of models of proposition p [17]. This framework can be extended to graded possibility theory using a graded extension of the square of opposition [8].

A relation-based reading of the square of opposition has been proposed in [7,8]. Let us now consider a binary relation R on a Cartesian product $X \times Y$ (one may have $Y = X$). We assume $R \neq \emptyset$. Let xR denote the set $\{y \in Y \mid (x, y) \in R\}$. We write xRy when $(x, y) \in R$ holds, and $\neg(xRy)$ when $(x, y) \notin R$. Moreover, we assume that $\forall x,\ xR \neq \emptyset$, which means that the relation R is *serial*, namely $\forall x, \exists y$ such that xRy. We further assume that the complementary relation $\overline{R}$ ($x\overline{R}y$ iff $\neg(xRy)$), and its transpose are also serial, i.e. $\forall x,\ xR \neq Y$ and $\forall y,\ Ry \neq X$. These conditions enforce a non trivial relation between X and Y. In the following, set complementations are denoted by means of overbars.

Let S be a subset of Y. We assume $S \neq \emptyset$ and $S \neq Y$. The relation R and the subset S, also considering its complement $\overline{S}$, give birth to the two following subsets of X, namely the (left) images of S and $\overline{S}$ by R

$$R(S) = \{x \in X \mid \exists s \in S, xRs\} = \{x \in X \mid S \cap xR \neq \emptyset\} = \bigcup_{s\in S} Rs \tag{1.1}$$

$$R(\overline{S}) = \{x \in X \mid \exists s \in \overline{S}, xRs\} = \textstyle\bigcup_{s\in\overline{S}} Rs$$

and their complements

$$\overline{R(S)} = \{x \in X \mid \forall s \in S, \neg(xRs)\} = \overline{\textstyle\bigcup_{s\in S} Rs} = \textstyle\bigcap_{s\in S} \overline{Rs} = \textstyle\bigcap_{s\in S} \overline{R}s$$

$$\overline{R(\overline{S})} = \{x \in X \mid \forall s \in \overline{S}, \neg(xRs)\} = \{x \in X \mid xR \subseteq S\} = \overline{\bigcup_{s\in\overline{S}} Rs} = \bigcap_{s\in\overline{S}} \overline{R}s \tag{1.2}$$

The four subsets thus defined can be nicely organized into a square of opposition, see Fig. 1.2. Indeed, it can be checked that the set counterparts of the logical relations existing between the logical statements of the traditional square of opposition still hold here. Namely,

– $\overline{R(\overline{S})}$ and $R(\overline{S})$ are complements of each other, as are $\overline{R(S)}$ and $R(S)$;
  they correspond to the diagonals of the square;
– $\overline{R(\overline{S})} \subseteq R(S)$, and $\overline{R(S)} \subseteq R(\overline{S})$,
  thanks to condition $\forall x,\ xR \neq \emptyset$. These inclusions are represented by vertical arrows in Fig. 1.2;

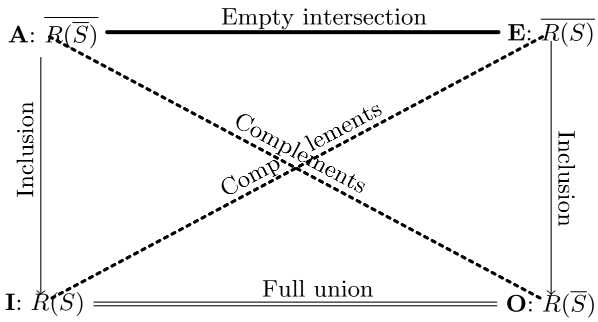


Fig. 1.2. Square of opposition induced by a relation R and a subset S

– $\overline{R(\overline{S})} \cap \overline{R(S)} = \emptyset$;
  this empty intersection corresponds to the thick line in Fig. 1.2,
  and one may have $\overline{R(\overline{S})} \cup \overline{R(S)} \neq Y$;
– $R(S) \cup R(\overline{S}) = X$;
  this full union corresponds to the double thin line in Fig. 1.2,
  and one may have $R(S) \cap R(\overline{S}) \neq \emptyset$.

Conditions (c)-(d) hold also thanks to the X-normalization of R.

Note that this fits with a modal logic reading of this square where R is viewed as an accessibility relation defined on $X \times X$, and S as the set of models of a proposition p. Indeed, $\Box p$ (resp. $\Diamond p$) is true in world x means that p is true at *every* (resp. at *some*) possible world accessible from x; this corresponds to $\overline{R(\overline{S})}$ (resp. $R(S)$) which is the set of worlds where $\Box p$ (resp. $\Diamond p$) is true. Moreover, the entailment from **A** to **I** is the axiom (D) of modal logic which is known to require serial accessibility relations [6].

Note that the relation R is serial if and only if $\overline{R(\overline{S})} \subseteq R(S)$. An interesting instantiation is in terms of rough sets [7], where in the classical case R is an equivalence relation. Then given the above definitions, we recognize that

– $R(S)$ is the upper approximation of S wrt the relation R;
– $\overline{R(\overline{S})}$ is the lower approximation of S wrt the relation R;
– $\overline{R(S)}$ is the exterior region of S;
– $R(\overline{S})$ is the complement of the lower approximation of S.

At this point one may observe that these relationships hold as well for fuzzy rough sets [13], if we replace the approximation operators by fuzzy ones – consider fuzzy box and diamond operators on crisp or fuzzy sets, also studied by Helmut Thiele [51]. A study of fuzzy rough sets in relation to the square of opposition appears in Ciucci et al. [9].

2.2 Early attempts at bridging possibility theory and modal logics

The first attempt at bridging possibility theory with modal logic can be found in a paper co-authored by L. Fariñas [27]. This paper establishes a formal parallel between rough sets and twofold fuzzy sets [12], namely a pair of fuzzy sets of elements that respectively certainly and possibly belong to an ill-known set. Then, taking advantage of the existence of the modal logic DAL for rough sets [26] and of a modal logic view of incomplete information databases [41], the paper discusses some possible options for a modal logic agreeing with possibility theory and with the issue of dealing with incomplete information rather than indiscernibility as in the case of rough sets.

A couple of years later, the idea of building a modal logic from a graded accessibility relation between different incomplete states of knowledge was investigated in detail in the case of binary-valued possibility theory and suggested for the graded case [22]. Then a state of knowledge s_2 is accessible from a state s_1 if and only if the information in state s_1 is consistent with the information in state s_2, but more incomplete (which was formalized as a set inclusion in the binary-valued case). In the general case, the inclusion becomes a matter of degree and the accessibility relation becomes graded. But the underlying axiom system remained an open issue.

Another attempt at the semantical level at bridging uncertainty theories with modal logic can be found in [46,48,47]; it includes the cases of possibility theory [31] and Shafer theory of evidence [30].

In the case of possibility theory, the authors use an accessibility relation assumed to be transitive and complete (connected), which corresponds to modal system S43. Necessity and possibility are built as ratios of the number of worlds in which the corresponding propositions are true.

3 From MEL to GPL

This section completes the background by providing a brief introduction to the meta-epistemic logic MEL, and to MEL$^+$ and then to generalized possibilistic logic GPL.

3.1 MEL and MEL$^+$, two simple epistemic logics

The usual truth values *true* (1) and *false* (0) assigned to propositions are of ontological nature (which means that they are part of the definition of what we call *proposition*), whereas assigning to a proposition a value whose meaning is expressed by the word *unknown* sounds like having an epistemic nature: it reveals a knowledge state according to which the truth value of a proposition (in the usual Boolean sense) in a given situation is out of reach (for instance one cannot compute it, either by lack of computing power, or due to a sheer lack of information). It corresponds to an epistemic state for an agent that can neither assert the truth of a Boolean proposition nor its falsity.

Admitting that the concept of "unknown" refers to a knowledge state rather than to an ontic truth value, we may keep the logic Boolean and add to its syntax the capability of stating that we ignore the truth value (1 or 0) of propositions. The natural framework to syntactically encode statements about knowledge states of classical propositional logic (CPL) statements is modal logic, and in particular, the logic KD. Nevertheless, if one only wants to reason about e.g. the beliefs of another agent, a very limited fragment of this language is needed. The logic MEL [1,3] was defined for that purpose.

Let us consider $\mathcal{L}$ to be a standard propositional language built up from a finite set of propositional variables $\mathcal{V} = \{p_1, \ldots, p_k\}$ along with the Boolean connectives of conjunction and negation $\neg$. As usual, a disjunction $\varphi \vee \psi$ stands for $\neg(\neg\varphi \wedge \neg\psi)$ and an implication $\varphi \to \psi$ stands for $\neg\varphi \vee \psi$. Further we use $\top$ to denote $\varphi \vee \neg\varphi$, and $\bot$ to denote $\neg\top$. Let us consider another propositional language $\mathcal{L}_\Box$ whose set of propositional variables is of the form $\mathcal{V}_\Box = \{\Box\varphi \mid \varphi \in \mathcal{L}\}$ to which the classical connectives can be applied. It is endowed with a modality operator expressing certainty, that encapsulates formulas in $\mathcal{L}$. In other words $\mathcal{L}_\Box = \{\Box\alpha : \alpha \in \mathcal{L}\} \mid \neg\Phi \mid \Phi \wedge \Psi$.

MEL is a propositional logic on the language $\mathcal{L}_\Box$ and with the following semantics. Let Ω be the set of classical interpretations for the propositional language $\mathcal{L}$, i.e. Ω consists of the set of mappings $w : \mathcal{L} \to \{0, 1\}$ conforming to the rules of classical propositional logic. For a propositional formula $\varphi \in \mathcal{L}$ we will denote by $Mod(\varphi)$ the set of $w \in \Omega$ such that $w(\varphi) = 1$. Models (or interpretations) for MEL correspond to epistemic states, which are simply subsets $\emptyset \neq E \subseteq \Omega$. The truth-evaluation rules of formulas of $\mathcal{L}_\Box$ in a given epistemic model E are defined as follows:

- $E \models \Box\varphi$ if $E \subseteq Mod(\varphi)$
- $E \models \neg\Phi$ if $E \not\models \Phi$
- $E \models \Phi \wedge \Psi$ if $E \models \Phi$ and $E \models \Psi$

Note that contrary to what is usual in modal logic, modal formulas are not evaluated on particular interpretations of the langage $\mathcal{L}$ because modal formulas in MEL do not refer to the actual world.

The notion of logical consequence is defined as usual $\Gamma \models \Phi$ if, for every epistemic model E, $E \models \Phi$ whenever $E \models \Psi$ for all $\Psi \in \Gamma$.

MEL can be axiomatized in a rather simple way (see [3]). The following are a possible set of axioms for MEL in the language of $\mathcal{L}_\Box$:

(CPL) Axioms of CPL for $\mathcal{L}_\Box$-formulas
(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
(D) $\Box\varphi \to \Diamond\varphi$
(Nec) $\Box\varphi$, for each $\varphi \in \mathcal{L}$ that is a CPL tautology, i.e. if $Mod(\varphi) = \Omega$.

The only inference rule is modus ponens. The corresponding notion of proof, denoted by $\vdash$, is defined as usual from the above set of axioms and modus ponens.

This set of axioms provides a sound and complete axiomatization of MEL, that is, it holds that, for any set of MEL formulas $\Gamma \cup \{\varphi\}$, $\Gamma \models \varphi$ iff $\Gamma \vdash \varphi$. This is not surprizing: MEL is just a standard propositional logic with additional axioms, whose propositional variables are the formulas of another propositional logic, and whose interpretations are subsets of interpretations of the latter.

MEL has been extended in [2] to allow dealing with not only subjective formulas that express an agent's beliefs, but also objective formulas (i.e. non-modal formulas) that express propositions that hold true in the actual world (whatever it might be). The extended language will be denoted by $\mathcal{L}_\Box^+$, and it thus contains both propositional and modal formulas. It exactly corresponds to the non-nested fragment of the language of usual modal logic.

More precisely, the language $\mathcal{L}_\Box^+$ of MEL$^+$ extends $\mathcal{L}_\Box$ and is defined by the following formation rules:

- If $\varphi \in \mathcal{L}$ then $\varphi, \Box\varphi \in \mathcal{L}_\Box^+$
- If $\Phi, \Psi \in \mathcal{L}_\Box^+$ then $\neg\Phi, \Phi \wedge \Psi \in \mathcal{L}_\Box^+$

$\Diamond\varphi$ is defined as an abbreviation of $\neg\Box\neg\varphi$. Note that $\mathcal{L} \subseteq \mathcal{L}_\Box^+$ and that in $\mathcal{L}_\Box^+$ there are no formulas with nested modalities.

Semantics for MEL$^+$ are given now by "pointed" MEL epistemic models, i.e. by structures (w, E), where $w \in \Omega$ and $\emptyset \neq E \subseteq \Omega$. The truth-evaluation rules of formulas of $\mathcal{L}_\Box^+$ in a given structure (w, E) are defined as follows:

- $(w, E) \models \varphi$ if $w \in Mod(\varphi)$, in case $\varphi \in \mathcal{L}$
- $(w, E) \models \Box\varphi$ if $E \subseteq Mod(\varphi)$
- usual rules for $\neg$ and $\wedge$

Logical consequence, as usual: $\Gamma \models \Phi$ if, for every structure (w, E), $(w, E) \models \Phi$ whenever $(w, E) \models \Psi$ for all $\Psi \in \Gamma$. The following are the axioms for MEL$^+$ in the language of $\mathcal{L}_\Box^+$:
(CPL) Axioms of propositional logic
(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
(D) $\Box\varphi \to \Diamond\varphi$
(Nec) $\Box\varphi$, for each $\varphi \in \mathcal{L}$ that is a CPL tautology, i.e. if $Mod(\varphi) = \Omega$.

The only inference rule is *modus ponens*.[4]

It can be proven that the above axiomatization of MEL$^+$ is sound and complete with respect to the intended semantics, as defined above. Moreover, as it could be expected, if we call MEL^{++} the extension of MEL$^+$ with the axiom:
(T) $\Box\varphi \to \varphi$
then it can be shown that MEL^{++} is complete with respect to the class of *reflexive* pointed epistemic models (w, E), i.e. where $w \in E$.

Actually, MEL, MEL$^+$ and MEL^{++} capture different *non-nested* fragments of the normal modal logics of belief `KD`, `KD4`, `KD45` and `S5` (see e.g. [6] for details). In [2] the following relationships are shown:

- Let φ a formula from $\mathcal{L}_\Box$. Then MEL $\vdash \varphi$ iff $L \vdash \varphi$, for $L \in \{\texttt{KD}, \texttt{KD4}, \texttt{KD45}, \texttt{S5}\}$.
- Let φ a formula from $\mathcal{L}_\Box^+$. Then MEL$^+ \vdash \varphi$ iff $L \vdash \varphi$, for $L \in \{\texttt{KD}, \texttt{KD4}, \texttt{KD45}\}$.
- Let φ a formula from $\mathcal{L}_\Box^+$. Then, MEL$^{++} \vdash \varphi$ iff $\texttt{S5} \vdash \varphi$.

Moreover, by recalling the well-known result that any formula of `KD45` and `S5` is logically equivalent to another formula without nested modalities, the following stronger relationships hold:

- For any arbitrary modal formula φ, there is a formula $\varphi' \in \mathcal{L}_\Box^+$ such that $\texttt{KD45} \vdash \varphi$ iff MEL$^+ \vdash \varphi'$.
- For any arbitrary modal formula φ, there is a formula $\varphi' \in \mathcal{L}_\Box^+$ such that $\texttt{S5} \vdash \varphi$ iff MEL$^{++} \vdash \varphi'$.

3.2 About generalized possibilistic logic

A natural generalization of MEL is to extend epistemic states $E \subseteq \Omega$ to rankings of possible worlds in terms of plausibility. This can be done by means of a mapping $\pi : \Omega \to U$ that assigns to each possible world w a value $\pi(w)$ from a totally ordered uncertainty scale $\langle U, \leq, 0, 1\rangle$ (which we will assume furthermore to be such that $\{0, 1\} \subseteq U \subseteq [0, 1]$ and closed by $n(x) = 1 - x$), with the following conventions:

- $\pi(w) = 1$ if w is fully plausible

[4] An equivalent presentation could be to replace (Nec) by the usual Necessitation rule in modal logics, but restricted to tautologies of propositional logic: if $\varphi \in \mathcal{L}$ is a theorem, derive $\Box\varphi$.

- $\pi(w) = 0$ if w is rejected as a possible world
- $\pi(w) \leq \pi(w')$ if w' is at least as plausible as w.

Such a mapping is called possibility distribution. A possibility distribution $\pi : \Omega \to U$ induces a pair of dual possibility and necessity measures on propositions, defined respectively as:

$$\Pi(\varphi) := \sup\{\pi(w) \mid w \in \Omega, w(\varphi) = 1\}$$
$$N(\varphi) := \inf\{1 - \pi(w) \mid w \in \Omega, w(\varphi) = 0\}\ .$$

They are dual in the sense that $\Pi(\varphi) = 1 - N(\neg\varphi)$ for every proposition φ.

Actually, *possibilistic logic* (see e.g. [11,15,18]), nowadays a well-known uncertainty logic, was initially devised to reason with graded beliefs on classical propositions by means of necessity and possiblity measures. For instance, the necessity fragment of possibilistic logic deals with weighted formulas (φ, r), where φ is a classical proposition and $r \in U$ is a weight, interpreted as a lower bound for the necessity degree of φ. It has a very simple axiomatization:

(CPL) $(\varphi, 1)$, for φ being a tautology of CPL
(GMP) from (φ, r) and $(\varphi \to \psi, s)$ derive $(\psi, \min(r, s))$
(Nes) from (φ, r) derive (ψ, s), if $s \leq r$

A graded extension of MEL capturing possibilistic logic has been proposed under the name *Generalized Possibilistic Logic*, GPL for short, in [20]. To deal with graded possibility and necessity they fix a *finite* scale of uncertainty values $\Lambda = \{0, \frac{1}{k}, \frac{2}{k}, \ldots, 1\}$ and for each value $a \in \Lambda \setminus \{0\}$ introduce a pair of modal operators $\Box_a$ and $\Diamond_a$. In this case models (epistemic states) are possibility distributions $\pi : \Omega \to \Lambda$ on the set Ω of classical interpretations for the language L_1 with values in Λ, and the evaluation of the modal formulas is as follows:

$$\pi \models \Box_a\varphi \qquad \text{if} \qquad N_\pi(\varphi) = \min\{1 - \pi(w) \mid w(\varphi) = 0\} \geq a.$$

The dual possibility operators are defined as $\Diamond_a\varphi := \neg\Box_{s(1-a)}\neg\varphi$, where the superscript $s(a)$ refers to the successor of a in Λ. The semantics of $\Diamond_a\varphi$ is the natural one, i.e. $\pi \models \Diamond_a\varphi$ whenever the possibility degree of φ induced by π, $\Pi(\varphi) = \max\{\pi(w) \mid w(\varphi) = 1\}$, is at least a. A complete axiomatization of GPL is given in [20], an equivalent and shorter axiomatization is given by the following additional set of axioms and rules to those of CPL[21]:

(K) $\Box_a(\varphi \to \psi) \to (\Box_a\varphi \to \Box_a\psi)$
(D) $\Diamond_1\top$
(Nes) $\Box_{a_1}\varphi \to \Box_{a_2}\varphi$, if $a_1 \geq a_2$
(Nec) $\Box_1\varphi$, for each $\varphi \in \mathcal{L}$ that is a CPL tautology.

4 GPL$^+$: extending generalized possibilistic logic with objective formulas

Let again $\Lambda = \{0, \frac{1}{k}, \frac{2}{k}, \ldots 1\}$ where $k \in \mathbb{N} \setminus \{0\}$ be the finite uncertainty scale we will assume. Moreover we let $\Lambda^+ = \Lambda \setminus \{0\}$, and if $a \in \Lambda^+$, we denote by $p(a)$ the value in the scale that preceeds a.

In this section we extend the language of generalized possibilistic logic (GPL) to allow dealing with not only subjective formulas that express an agent's beliefs, but also objective formulas (i.e. non-modal formulas) that express propositions that hold true in the actual world (whatever it might be). The extended language will be denoted by $\mathcal{L}_\Box^{k+}$, and it thus contains both propositional and modal formulas. It exactly corresponds to the non-nested fragment of the language of usual modal logic.

More precisely, the language $\mathcal{L}_\Box^{k+}$ of GPL$^+$ extends the one of GPL, $\mathcal{L}_\Box^+$, and is defined by the following formation rules:

- If $\varphi \in \mathcal{L}$ and $a \in \Lambda^+$ then $\varphi, \Box_a\varphi \in \mathcal{L}_\Box^{k+}$
- If $\Phi, \Psi \in \mathcal{L}_\Box^{k+}$ then $\neg\Phi, \Phi \wedge \Psi \in \mathcal{L}_\Box^{k+}$

$\Diamond_b\varphi$ is defined as an abbreviation of $\neg\Box_a\neg\varphi$, with $b = 1 - p(a)$. Note that $\mathcal{L} \subseteq \mathcal{L}_\Box^+$ and that in $\mathcal{L}_\Box^+$ there are no formulas with nested modalities.

Semantics for GPL$^+$ are given now by "pointed" possibilistic models, i.e. by structures (w, π), where $w \in \Omega$ and $\pi : \Omega \to \Lambda$ such that there is at least one $w \in \Omega$ with $\pi(w) = 1$. For each proposition $\varphi \in \mathcal{L}$, let $N_\pi(\varphi) = \inf_{w \notin Mod(\varphi)} \pi(w)$. The truth-evaluation rules of formulas of $\mathcal{L}_\Box^{k+}$ in a given structure (w, π) is defined as follows:

- $(w, \pi) \models \varphi$ if $w(\varphi) = 1$, in case $\varphi \in \mathcal{L}$

– $(w, \pi) \models \Box_a \varphi$ if $N_\pi(\varphi) \geq a$
– usual rules for $\neg$ and $\wedge$

If we let $\pi_a = \{w \in \Omega \mid \pi(w) \geq a\}$, note that $(w, \pi) \models \Box_a \varphi$ whenever $\pi_{1-p(a)} \subseteq Mod(\varphi)$. Therefore, it becomes clear that each $\Box_a$ operator is a MEL$^+$ modality.

The corresponding logical consequence is defined as usual: $\Gamma \models \Phi$ if, for every structure (w, π), $(w, \pi) \models \Phi$ whenever $(w, \pi) \models \Psi$ for all $\Psi \in \Gamma$.

The following are the axioms for GPL$^+$ in the language of $\mathcal{L}_\Box^{k+}$:
(CPL) Axioms of propositional logic
(K_a) $\Box_a(\varphi \to \psi) \to (\Box_a \varphi \to \Box_a \psi)$, for every $a \in \Lambda^+$
(D_a) $\Box_a \varphi \to \Diamond_1 \varphi$, for every $a \in \Lambda^+$
(Nes) $\Box_a \varphi \to \Box_b \varphi$, where $b \leq a$
(Nec) $\Box_1 \varphi$, for each $\varphi \in \mathcal{L}$ that is a CPL tautology

The only inference rule is *modus ponens*. We will write $\Gamma \vdash \Phi$ to denote that φ can be derived from a set of formulas Γ using the above axioms and modus ponens. Also, in what follows, we will denote by $\vdash_{CPL}$ the notion of proof of classical propositional language on the language $\mathcal{L}_\Box^{k+}$ taking all $\Box$-formulas as new propositional variables.

To prove completeness, we first recall the following useful lemma that allows to express deductions in GPL$^+$ as deductions in CPL.

Lemma 1. *Let $\Gamma \cup \{\Phi\}$ be a set of $\mathcal{L}_\Box^{k+}$-formulas. Then it holds that $\Gamma \vdash \Phi$ iff $\Gamma \cup \{\Box_1 \varphi \mid \vdash_{CPL} \varphi\} \cup \{$instances of axioms ($K_a$), ($D_a$), (Nes) and (Nec)$\} \vdash_{CPL} \Phi$.*

Theorem 1 (Completeness). *For any set of $\mathcal{L}_\Box^{k+}$-formulas $\Gamma \cup \{\Phi\}$, it holds that $\Gamma \vdash \Phi$ iff $\Gamma \models \Phi$.*

Proof. From left to right is easy, as usual. For the converse direction, assume $\Gamma \not\vdash \Phi$. By the preceding lemma and the completeness of PL, there exists a propositional evaluation v on the whole language $\mathcal{L}_\Box^{k+}$ (taking $\Box$-formulas as genuine propositional variables) such that $v(\Psi) = 1$ for all $\Psi \in \Gamma \cup \{\Box_1 \varphi \mid \vdash_{PL} \varphi\} \cup \{$instances of axioms (K), (D) and (Nes)$\}$ but $v(\Phi) = 0$. We have to build a structure (w, π) that it is a model of Γ but not of Φ. So, we take (w, π) as follows:

– w is defined as the restriction of v to $\mathcal{L}$, i.e. $w(\varphi) = v(\varphi)$ for all $\varphi \in \mathcal{L}$.
– For each $a \in \Lambda^+$, let us first define $E_{1-p(a)} = \bigcap \{Mod(\varphi) \mid v(\Box_a \varphi) = 1\}$. Then define $\pi : \Omega \to \Lambda$ as follows: $\pi(w) = \max\{a \in \Lambda^+ \mid w \in E_a\}$, where we adopt the usual convention of taking $\max \emptyset = 0$. In other words, we define π in such a way that each a-cut π_a coincides with E_a.

Note that, since by axioms (D) and (Nec) we have $v(\Diamond_1 \top) = 1$, $E_1 \neq \emptyset$. Then the last step is to show that, for every $\Psi \in \mathcal{L}_\Box^{k+}$, $v(\Psi) = 1$ iff $(w, \pi) \models \Psi$.

We prove this by induction. The case Ψ being a non-modal formula from $\mathcal{L}$ is clear, since in that case $w(\Psi) = v(\Psi)$. The interesting case is when $\Psi = \Box_a \psi$. Then we have:

(i) If $v(\Box_a \psi) = 1$ then, by definition of $E_{1-p(a)}$, $E_{1-p(a)} \subseteq Mod(\psi)$, and hence $(w, \pi) \models \Box_a \psi$.
(ii) Conversely, if $E_{1-p(a)} \subseteq Mod(\psi)$, then there must exist γ such that $v(\Box_a \gamma) = 1$ and $Mod(\gamma) \subseteq Mod(\psi)$. Hence this means that $\gamma \to \psi$ is a PL theorem, and hence we have first, by the necessitation axiom, that $v(\Box_a(\gamma \to \psi)) = 1$, and thus $v(\Box_a \gamma) \leq v(\Box_a \psi)$ holds as well by axiom (K), and therefore $v(\Box_a \psi) = 1$ holds as well.

As a consequence, we have that $(w, \pi) \models \Psi$ for all $\Psi \in \Gamma$ but $(w, \pi) \not\models \Phi$.

Similar to the non graded case of MEL$^+$, we may consider an S5-like extension of GPL$^+$, capturing the pointed possibilistic epistemic models (w, π), where the ‘actual world’ w is one of the non-discarded possible worlds by π. In this case, the higher $\pi(w)$ is, the more the actual world w belongs to the set of plausible worlds, and hence we can speak of a notion of graded reflexive pointed possibilistic epistemic models (w, π).

Definition 1. *Let (w, π) be a pointed possibilistic structure and let $a \in \Lambda^+$. We call (w, π) to be a-reflexive when $\pi(w) \geq a$.*

Let us define GPL$_a^{++}$ to be the axiomatic extension of GPL$^+$ with the following generalized (T) axiom:
(T_a) $\Box_a \varphi \to \varphi$

One can check that (T_a) is valid in all b-reflexive pointed possibilistic structures, with $b = 1-p(a)$. Indeed, if $(w, \pi) \models \Box_a \varphi$ then $N_\pi(\varphi) \geq a$, and thus $\pi_{1-p(a)} \subseteq Mod(\varphi)$. But if (w, π) is b-reflexive, we have $\pi(w) \geq 1 - p(a)$, and hence $w \in \pi_{1-p(a)} \subseteq Mod(\varphi)$. Therefore $(w, \pi) \models \varphi$ as well.

Theorem 2. *GPL_a^{++} is complete with respect to the class of $(1-p(a))$-reflexive pointed possibilistic structures.*

Proof. The proof is analogous to that of Theorem 1.

It is interesting to point out that Liau and Lin [36,37] propose a language similar to GPL$^+$, albeit using $[0,1]$ as a possibility scale (which forces them to introduce additional multimodal formulas to deal with strict inequalities) and graded accessibility relations. Their tableau-based proof methods could be of interest to develop inference techiques for GPL.

5 Concluding remarks

In this paper, following the fact that the fragment MEL$^+$ (resp. MEL^{++}) of the KD45 (resp. S5) logic, the richest of doxastic (resp. epistemic) logics, involving modal formulas of depth 0 or 1 can have simplified semantics, we show that this state of facts extends to graded modalities with the extensions GPL$^+$ and GPL^{++} of the generalized possibilistic logic GPL.

Besides, it has been recently shown that the graded notion of guaranteed possibility can be expressed in GPL enabling us to express "all I know" statements [21] (see also [3] for the crisp case). This result calls for for a deeper comparison with the modal logic presented in [10] that involves the classical modalities of the possible and the necessary together with the nonstandard modalities that are the guaranteed possibility and its dual, having also in mind that these four modalities and their negations makes a cube of opposition [8] that generalizes the square of opposition.

Dedication

This article is particularly dedicated to Luis Fariñas del Cerro. It perfectly illustrates one of the topics at the junction of our respective subjects of interest, namely modal logic and possibility theory. Discussions along 35 years of friendship have repeatedly triggered two of the authors to dig more and more about the relations between these two knowledge representation frameworks, thanks also to the help of the two other authors of this note. Interestingly enough, while gaining mutual understanding of our respective reference theories, each of us has remained a supporter of one's own theory. Let us hope that in the long range, the now obvious bridge between the two formalisms will become routine knowledge so that both can be used appropriately by the same people according to the particulars of the applications at hand.

Bibliography

1. M. Banerjee, D. Dubois. A simple modal logic for reasoning about revealed beliefs. *Proc. 10th Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, (C. Sossai and G. Chemelo, eds.), Verona, Springer, LNAI 5590, Springer, 805-816, 2009.
2. M. Banerjee, D. Dubois and L. Godo. Possibilistic vs. relational semantics for logics of incomplete information. *Proc. 15th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'14)*, (A. Laurent, O. Strauss, B. Bouchon-Meunier, R. R. Yager, eds.), Montpellier, July 15-19, Part I, Springer, CCIS 442, 335-344, 2014.
3. M. Banerjee, D. Dubois, A simple logic for reasoning about incomplete knowledge *Int. J. Approximate Reasoning*, 55 (2), 639-653, 2014.
4. J. Y. Béziau. New light on the square of oppositions and its nameless corner. *Logical Investigations*, 10, 218-233, 2003.
5. C. Boutilier. Modal logics for qualitative possibility theory, *Int. J. Approximate Reasoning*, 10, 173-201, 1994.
6. B. F. Chellas. *Modal Logic: an Introduction.* Cambridge University Press,1980.
7. D. Ciucci, D. Dubois and H. Prade. The structure of oppositions in rough set theory and formal concept analysis - Toward a new bridge between the two settings. *Proc. 8th Int. Symp. on Foundations of Information and Knowledge Systems (FoIKS'14)*, (C. Beierle, C. Meghini, eds.), Bordeaux, Mar. 3-7, Springer, LNCS 8367, 154-173, 2014.
8. D. Ciucci, D. Dubois and H. Prade. Structure of opposition induced by relations. The Boolean and the gradual cases. *Annals of Mathematics and Artificial Intelligence*, DOI:10.1007/s10472-015-9480-8, 2015.
9. D. Ciucci, D. Dubois and H. Prade. Structures of Opposition in Fuzzy Rough Sets. Fundamenta Informaticae, vol. 142, no. 1-4, pp. 1-19, 2015
10. D. Dubois, P. Hajek and H. Prade. Knowledge-driven versus data-driven logics. *J. of Logic, Language, and Information*, 9, 65-89, 2000.

11. D. Dubois, J. Lang and H. Prade. Possibilistic logic. In: *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3*, (D. M. Gabbay, C. J. Hogger, J. A. Robinson, D. Nute, eds.), Oxford University Press, 439-513, 1994.
12. D. Dubois and H. Prade. Twofold fuzzy sets and rough sets - Some issues in knowledge representation. *Fuzzy Sets and Systems*, 23 (1), 3-18, 1987.
13. D. Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets, Int. J. of General Systems, 17(2-3), 191-209, 1990
14. D. Dubois and H. Prade. Resolution principles in possibilistic logic. *Int. J. of Approximate Reasoning*, 4 (1), 1-21, 1990.
15. D. Dubois, H. Prade. Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems*, 144, 3-23, 2004.
16. D. Dubois, H. Prade. Toward multiple-agent extensions of possibilistic logic. *Proc. IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'07)*, London, July 23-26, 187-192, 2007.
17. D. Dubois and H. Prade. From Blanché's hexagonal organization of concepts to formal concept analysis and possibility theory. *Logica Univers.*, 6, 149-169, 2012.
18. D. Dubois and H. Prade. Possibilistic logic - An overview. In: *Handbook of the History of Logic, vol. 9 Comput. Logic*, (D. Gabbay, J. Siekmann, J. Woods, eds.), 283-342, Elsevier, 2014.
19. D. Dubois, H. Prade and A. Rico. The cube of opposition: A structure underlying many knowledge representation formalisms. *Proc. 24th Int. Joint Conf. on Artificial Intelligence (IJCAI'15)*, Buenos Aires, July 25-31 (Q. Yang, M. Wooldridge, eds.), AAAI Press, 2015, 2933-2939.
20. D. Dubois, H. Prade and S. Schockaert. Stable models in generalized possibilistic logic. *Proc.13th Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'12)*, (G. Brewka, T. Eiter, S. A. McIlraith, eds.), Rome, June 10-14, AAAI Press, 519-529, 2012.
21. D. Dubois, H. Prade and S. Schockaert. Reasoning about uncertainty and explicit ignorance in generalized possibilistic logic. *Proc. 21st Europ. Conf. on Artificial Intelligence (ECAI'14)*, (T. Schaub, G. Friedrich, B. O'Sullivan, eds.), Aug. 18-22, Prague, IOS Press, 261-266, 2014.
22. D. Dubois, H. Prade, and C. Testemale. In search of a modal system for possibility theory. *Proc. 8th Europ. Conf. on Artificial Intelligence (ECAI'88)*, Munich, Aug. 1-5, 501-506, 1988.
23. R. Fagin and M. Y. Vardi. An internal semantics for modal logic. *Proc. 17th ACM Symposium on Theory of Computing (STOC'85)*, Providence, RI, May 6-8, 305-315, 1985.
24. J. Y. Halpern, R. Fagin, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge.* MIT Press (Revised paperback edition), 2003.
25. L. Fariñas del Cerro and A. Herzig. A modal analysis of possibility theory. In: *Fundamentals of Artificial Intelligence Research (FAIR'91)*, (Jorrand Ph. and Kelemen J., Eds.), LNCS 535, Springer Verlag, Berlin, 11-18, 1991.
26. L. Fariñas del Cerro and E. Orlowska. DAL- A logic for data analysis. *Theoretical Computer Science*, 36, 251-264, 1985.
27. L. Fariñas del Cerro and H. Prade. Rough sets, twofold fuzzy sets and modal logic. Fuzziness in indiscernibility and partial information. In: *The Mathematics of Fuzzy Systems*, (A. Di Nola, A. G. S. Ventre, eds.), Vol. 88 in Interdisciplinary Systems Research series, Verlag TÜV Rheinland, 103-120, 1986.
28. P. Hájek. A qualitative fuzzy possibilistic logic, *Int. J. of Approximate Reasoning*, 12, 1-19, 1994.
29. P. Hájek P., D. Harmancová, F. Esteva, P. Garcia and L. Godo. On modal logics for qualitative possibility in a fuzzy setting. *Proc. 10th Conf. on Uncertainty in Artificial Intelligence (UAI'94)*, (R. Lopez de Mántaras and D. Poole, eds.), Seattle, July 29-31, Morgan Kaufmann, 278-285, 1994.
30. D. Harmanec and G. J. Klir. On modal logic interpretation of Dempster-Shafer theory of evidence. *Int. J. of Intelligent Systems*, 9, 941-951, 1994.
31. G. J. Klir and D. Harmanec. On modal logic interpretation of possibility theory. *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems*, 2 (2), 237-245, 1994.
32. G. Lakoff. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *J. of Philosophical Logic*, 2, 458-508, 1973.
33. J. Lang, D. Dubois and H. Prade. A logic of graded possibility and certainty coping with partial inconsistency. *Proc. 7th Conf. on Uncertainty in Artificial Intelligence (UAI'91)*, UCLA, Los Angeles, July 13-15, Morgan-Kaufmann, Palo Alto, Ca., 188-196, 1991.
34. J. Lang and P. Marquis. Reasoning under inconsistency: A forgetting-based approach *Artificial Intelligence* 174, pp. 799-823, 2010.
35. C.-J. Liau, T.-F. Fan. A modal logic for reasoning about possibilistic belief fusion. *Proc. 19th Int. Joint Conf. on Artificial Intelligence (IJCAI'05)*, (L. Pack Kaelbling, A. Saffiotti, eds.), Edinburgh, July 30-Aug. 5, 1756-1757, 2005.
36. C.-J. Liau and B. I-Peng Lin. Quantitative modal logic and possibilistic reasoning. *Proc. 10th Europ. Conf. on Artificial Intelligence (ECAI'92)*, (B. Neumann, ed.), Vienna, August 3-7, John Wiley and Sons, Chichester, 43-47, 1992.
37. C.-J. Liau and B. I-Peng Lin. Proof methods for reasoning about possibility and necessity. *Int. J. Approx. Reasoning*, 9 (4), 327-364, 1993.
38. C.-J. Liau and B. I-Peng Lin. Reasoning about higher order uncertainty in possiblistic logic. Proc. 7th Int. Symp. on Methodologies for Intelligent Systems (ISMIS'93), (H. J. Komorowski, Z. W. Ras, eds.), Trondheim, June 15-18, LNCS 689, Springer, 316-325, 1993.

39. V. Lifschitz. Minimal belief and negation as failure. *Artificial Intelligence* 70, 53-72, 1994.
40. F. Lin and Y. Shoham. A logic of knowledge and justified assumptions. *Artificial Intelligence* 57, 271-289, 1992.
41. W. Lipski. On databases with incomplete information. *J. of ACM*, 28 (1), 41-70, 1981.
42. E. Orlowska. A logic of indiscernibility relations. In: *Computation Theory*, LNCS 208, Springer-Verlag, Berlin, 177-186, 1985.
43. Z. Pawlak and A. Skowron. Rudiments of rough sets. *Inf. Sciences* 177, 3-27, 2007.
44. A. Petruszczak. Simplified Kripke style semantics for modal logics K5, KB5 and KD45. *Bulletin of the Section of Logic* 38:163-171, 2009.
45. H. Prade. Reasoning with fuzzy default values. *Proc. 15th IEEE Inter. Symp. Multiple-Valued Logic*, Kingston, On., Canada, May 28-30, 191-197, 1985.
46. G. Resconi, G. J. Klir and U. St. Clair. Hierarchical uncertainty metatheory based upon modal logic. *Int. J. General Systems*, 21, 23-50, 1992.
47. G. Resconi, G. J. Klir, D. Harmanec and U. St. Clair. Interpretations of various uncertainty theories using models of modal logic: a summary. *Fuzzy sets and Systems*, 80, 7-14, 1996.
48. G. Resconi, G. J. Klir, U. St. Clair and D. Harmanec. On the integration of uncertainty theories. *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems*, 1, 1-18, 1993.
49. P. K. Schotch. Fuzzy modal logic. *Proc. 5th IEEE Int. Symp. on Multiple-Valued Logic*, Bloomington, IN, May 13-16, 176-182, 1975.
50. G. L. S. Shackle. *Decision, Order, and Time in Human Affairs.* Cambridge Univ. Press, 1961.
51. H. Thiele, Fuzzy Rough Sets versus Rough Fuzzy Sets Ñ. An Interpretation and a Comparative Study using. Concepts of Modal Logics. Research report No. CI-30/ 98, University of Dortmund, Germany, 1998.
52. Y. Yao and T. Lin. Generalization of rough sets using modal logic. *Automation and Soft Computing* 2(2), 103-120, 1996.
53. Y. Yao, S. Wang and T. Lin. A review of rough set models. In: *Rough Sets and Data Mining: Analysis for Imprecise Data*, (L. Polkowski, A. Skowron, eds.), Kluwer Acad. Publ., 47-75, 1997.
54. L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* **1** (1978) 3–28.
55. L. A. Zadeh. PRUF - A meaning representation language for natural languages. *Int. J. Man-Machine Studies* **10** (1978) 395–460.
56. L. A. Zadeh. A theory of approximate reasoning. In: *Machine Intelligence, vol. 9*, (J. E. Hayes, D. Mitchie, L. I. Mikulich, eds.), Ellis Horwood, 149–194, 1979.
57. L. A. Zadeh. Fuzzy sets and information granularity. In: *Advances in Fuzzy set Theory and Applications*, (M. M. Gupta, R. K. Ragade, R. R. Yager, eds.), North-Holland, 3–18, 1979.

Society Semantics and the Logic way to Collective Intelligence

Walter Carnielli[1] and Mamede Lima-Marques[2]

[1] State University of Campinas – UNICAMP
Centre for Logic, Epistemology and the History of Science
and Department of Philosophy
walter.carnielli@cle.unicamp.br
[2] University of Brasília – UnB
Centre for Research on Architecture of Information
mamede@unb.br

Abstract. The so-called phenomenon of collective intelligence is now a burgeoning movement ([1]), with several guises and examples in many areas. We briefly survey some relevant aspects of collective intelligence in several formats, such as social software, crowdfunding and convergence, and show that a formal version of this paradigm can also be posed to logic systems, by means of the idea of logic societies (cf. [2]). The paradigm of logical societies has lead to a new notion of distributed semantics, the *society semantics*, with theoretical advances in defining new forms of n-valued semantics in terms of k -valued semantics, for $k < n$, and applications as in [3] to flying security protocols. We summarize the main advances of society semantics, commenting on their general case, the *possible-translations semantics* ([4], [5]) and pointing to some conceptual points and to some problems and directions still to be explored.

1 Logic societies and collective intelligence

In a reference paper that tries to characterize the movement of collective intelligence, broadly defined as 'groups of individuals doing things collectively that seem intelligent' ([1]), a genomic approach to crowd intelligence is attempted, in an effort to explain what makes groups smart and how teams can be made more effective. Although examples like Google search and Wikipedia are easily seen to be the result of efforts of distributed contributors, voting, stock market predictions, reputation systems, question-and-answer sites, predicting the likelihood of low-probability events and several other initiatives take profit of the common sense of groups, rather than individuals. Another case of collective intelligence is crowdfunding (in the sense of crowd financing or collective funding), the effort of a coalition of individuals involved in financing a third-party project. Crowdfunding is seen as an entrepreneurial phenomenon that is revolutionizing traditional sources of finance, abandoning fund raising usually done through banks, government agencies or investors in favor of collective funding. What is interesting is that in the crowdfunding process people is not necessarily looking for direct money profit, but in many cases they look for the product they would like to buy, or for the movie they would like to see. In this way, crowdfunding can be regarded as a case of collective monetary intelligence that subverts the traditional economic model, since people may be ultimately paying for someone to make profit on them.

Collective intelligence is not any panacea, though. There is always the danger of transforming collective intelligence into "collective stupidity", generating cognitive biases that may distort the judgment of individuals in a systematic way. A good discussion about cases of biases and distorted heuristics that may influence judgments under uncertainty is found in [6].

Criteria for positive and useful collective intelligence are proposed in [7], condensed into four categories: diversity of opinions, independence, decentralization, and aggregation. Thinking from the point of view of a group of agents (human or machines), diversity of opinions means that each agent in the group should draw conclusions from her/its private information; independence means that each agent's opinion should not be influenced by other group members; decentralization refers more concretely to the ability of each agent to draw from own sources, and aggregation refers to the availability of mechanisms to lead to a global group conclusion. What is interesting, from our perspective, is that these criteria referring to diversity of opinions, independence, decentralization, and aggregation are present, in a formalized way, in the central idea of society semantics (first connected to computer science in [3] [3]). Such semantics models the highly complex process of combining

[3] A Ph.D. thesis jointly supervised by L. Fariñas del Cerro and W. A. Carnielli.

(sometimes contradictory) information or behavior originating from different sources and drawing sensible conclusions from them.

Another interesting case of collective intelligence is the merging of distinct technologies coming from engineering, physical sciences, human sciences and life sciences into a common, more complex model. The so-called NBIC (nano-bio-info-cogno) model, for instance, works by transforming engineering and physical sciences taking advantage from biological models. Information technology, nanotechnology, quantum mechanics and statistics, combined with modeling and simulation are also transforming life sciences and physical sciences.

This mode of integration of disciplinary approaches that were previously viewed as watertight compartmentalized now leads to what is considered to be a "Third Revolution" in science (cf. [8], a white paper signed by twelve scientists), a new society of sciences that does not rest on a particular scientific advance, but collectively takes profit on the combination of collective strategies, methods and views.

Some intrinsic connections between crowd intelligence and the paraconsistent reasoning paradigm are found for instance in [9], where some techniques known as "consider the opposite" or "dialectical bootstrapping" are investigated, by which the wisdom of crowds can be simulated by a single mind averaging its own conflicting opinions (p. 233):

> " which prompt people to consider knowledge that was previously overlooked, ignored, or deemed inconsistent with current beliefs by, for example, asking them to think of reasons why their first judgment might be wrong".

An interesting consequence is that dialectical bootstrapping, taking into account contradictory (or inconsistent, as some authors prefer) information, can lead to higher accuracy than the standard case. This works as a formidable evidence in favor of paraconsistent reasoning, where the presence of contradictions, instead of destroying reasoning, may enhance the rational capacity. More on this topic in Sections 2 and 3

Society semantics, the logic fashion of this paradigm, are a direct heir of the possible-translations semantics ([4], [5], better reworked in [10]), designed as a tool to devise acceptable (in the sense of being at the same time strongly adequate and intuitively palatable) semantics for non-classical logics. The main components of the possible-translations semantics are logic systems, taken to be sets (of sentences) endowed with a consequence relation, and translations understood as morphisms between logics, i.e. maps preserving their consequence relations. The idea behind possible-translations semantics is to base an interpretation to a given sentence α of a logic L on the combination of an appropriate set of translations $t(\alpha)$ into a class of logics with known semantics.

Several paraconsistent logics, despite not being characterizable by finite matrices, can be characterized by a suitable combination of three-valued logics. For those logics, a decidable procedure is immediate, since the evaluation of a given sentence amounts to evaluating a finite collection of its possible translations using the three-valued matrices. In certain particular cases, many-valued logics themselves can be seen as suitable combinations of copies of classical logic, thus giving rise to the concept of society semantics. Possible-translations semantics can be naturally seen as a way to decompose a complex logic into its ingredients. This way of analyzing a complex logic into less complex components is called "splitting logics", and provides a powerful tool to combine logics (a full theoretical account of combination of logics with lots of examples is given in the book [11]).

Society semantics and possible-translations semantics are based on the intuition that the reasoning of a society can be substantially distinct from the reasoning of its members, and that in both levels a formal calculus can be used to express such reasoning.

Society semantics has been successfully used as a formal account of distributed control in a multi-agent scenario in the case of aircraft collision control: since pilots usually do not possess complete knowledge about each other, their reasoning involve concepts of mutual belief, knowledge about knowledge, time planning, and so on, but they have to be prepared to reason in the presence of contradictions. A failure of being unable to reason taking contradictions into account may be disastrous. Flight instruments are not always reliable, and contradictory warnings from the aircraft's flight computer may cause pilots to believe, for instance, that they are at safe altitude, when this is not the case, due to lack of protocols able to analyze contradictory situations. Many flying accidents could be explained by this kind of attitude. In the famous case of Air France Flight 447 from Rio de Janeiro to Paris which crashed on 1 June 2009 killing aircrew, cabin crew and all 228 passengers, crystal ice was blocking the *pitot* tube intake, causing a contradiction in instrument indications. The crew supposedly reacted incorrectly in the face of such contradictions and ultimately led the aircraft to an aerodynamic stall which caused it to impact the ocean.

In order to use ideas behind society semantics to improve flight security under conflict situation, a natural proposal was to integrate the society of agents in a multi-modal system, resulting into the systems $\mathsf{KB}\bigcirc$ and $\mathsf{SKB}\bigcirc$ (see [3]). The implemented logic program was based on a modal resolution paradigm using TIM (Toulouse Inference Machine) (cf. [12] and [13]).

In the society formed by perfectly rational agents, contradictions may arise because the (local) information produced by the agents must be processed in a higher (global) level. In this way, the society of agents must be able to identify and support contradictions and to draw sensible inferences taking such contradictions into account. In order to accomplish this task, the society should be endowed with a formal system which incorporates the logics of the agents, expressing their mutual contradictions and making possible to deduce sound information in such an environment.

In intuitive terms, society semantic offers a clear and natural interpretation for the logical behavior of groups of agents where any member holds or deducts a particular assertion, and another may deduct or hold the negation of the same assertion. In a sense, something of a similar nature was also proposed by S. Jaśkowski in his suggestions for a discussive logic (cf. [14]), where he defended the interest for contradictory deductive systems in terms of combinations of different opinions into a single system. Not surprisingly, Jaśkowski's better known logic D_2 is an LFI (that is, a paraconsistent member of the family of logics of formal inconsistency, see [10] for a detailed explanation).

Society semantics are able to express the fact that the reasoning of an ensemble of classical agents is not necessarily classical. Although an agent could reason with classical logic as an individual, her/its social behavior in a society of agents could be expressed by a different logic. For example, if a classical agent holds an assertion A and another holds the assertion $\neg A$, the society where they are inserted must be prepared to cope with this situation. A collection of societies is, by its turn, another society, and hence a hierarchy can be formed, each level equipped with a distinct logic. The most interesting cases occur when such logics are restricted to finite-valued logics.

2 Formalizing society semantics: from credulity to scepticism

In formal terms, a *society* S is composed by a denumerable (not necessarily finite) set of *agents* $S = \{Ag_1, Ag_2, \cdots, Ag_n, \cdots\}$ where each agent Ag_i is a pair $Ag_i = (C_i, L_i)$ formed by a collection C_i of sentences (in particular, propositional variables) in a formal language (intuitively interpreted as the set of propositions accepted by the agent) and by an underlying logic L_i.

In the case where all agents are subjected to the laws of classical propositional logic (we call them *classical agents*) the logic Lg of the society coincides with the deductive machinery of Classical Propositional Calculus (CPC) and C_i are sentences (or propositional variables) in the language generated by propositional variables and connectives $\to$, $\wedge$, $\vee$, $\neg$. Each agent can be naturally identified with the set of propositions which it accepts.

In such a particular case where all agents are subjected to CPC, each agent can be regarded as being completely "rational" (that is, classical). It is interesting to note, however, that even a group of classical reasoners may present a non-standard reasoning capacity, depending on the rules governing their mutual behavior.

A society is said to be *biassertive* if negation is not truth-functional (that is, if the truth value of $\neg A$ does not depend functionally on the truth-value of A). In this case, A and $\neg A$ can occur as primitive, for certain types of formulas A.

An agent Ag *accepts* a formula A, denoted by $Ag \models A$, if all classical valuations which satisfy the sentences of Ag also satisfy A. Interesting examples occur in voting schemas, as in the process of refereeing papers for a scientific congress. Suppose that a committee (that is, a society of referees) has adopted an *open policy*, in the sense of supporting the opinion of accepting a paper if at least one of the referees votes positively, and supporting the opinion of rejecting it if at least one of them gives a negative vote. We can express a referee's opinions by atomic formulas, and the committee's final opinions by non-atomic formulas. Under such rules this society may appear to be itself contradictory, although each agent is perfectly non-contradictory. This incoherence, however, expresses exactly the information that at least two agents disagree about one given task (in the case, in its opinion about a given paper).

The *satisfiability relation* between a society S and an arbitrary sentence A is inductively defined in the usual way for non-atomic formulas, and for atomic p is defined as given below.

DIfferent kinds of society can be defined. The behavior of a society can range from credulity to scepticism, depending on the totality of agents to be taken into account for the society's final decision: the more agents are consulted, the more sceptical a society will be said to be. The fewer agents are consulted, the more credulous a society will be. The extreme case where unanimity is required gives rise to "closed" societies; the antipode case where any singular opinion deserves to be represented gives rise to "open" societies [4].

A society is *open* if it accepts a sentence in case any of its agents does. Open societies are denoted by S_+. Formally:

(OBS-1) $S_+ \models p$ iff there exists an agent Ag in S such that $p \in Ag$

(OBS-2) $S_+ \models \neg p$ iff there exists an agent Ag in S such that $p \notin Ag$

(OBS-3) $S_+ \models A \to B$ iff $S_+ \not\models A$ or $S_+ \models B$

(OBS-4) $S_+ \models A \wedge B$ iff $S_+ \models A$ and $S_+ \models B$

(OBS-5) $S_+ \models A \vee B$ iff $S_+ \models A$ or $S_+ \models B$

(OBS-6) $S_+ \models \neg A$ iff $S_+ \not\models A$ for A not the form p.

A sentence A is *satisfiable in an open society* if there exists S_+ such that $S_+ \models A$, and is said to be an *open-tautology* if it is satisfiable in every open society.

It is clear that $S_+ \not\models (p \wedge \neg p) \to q$ in the case where $S_+ \models p$, $S_+ \models \neg p$ and $S_+ \not\models q$, so $(p \wedge \neg p) \to q$ is not an open-tautology. Similarly, $p \to (\neg p \to q)$ is not an open-tautology.

Thus, although the internal logic of the agents is classical, the external logic of open societies supports contradictions without crashing into trivialization (in particular, the implication in an open society has a paraconsistent and relevant character).

Now, if in the example above, the committee adopts a *closed policy*, in the sense of supporting the opinion of accepting a paper only if all the referees vote positively, and supporting the opinion of not accepting it only if all of them give negative votes, a similar explanation justifies the definition to follow.

A society is *closed* if it accepts a formula only in case all of its agents does. Closed societies are denoted by S_-. We define only the atomic cases (other clauses are defined similarly as above):

(CBS-1) $S_- \models p$ iff for every agent Ag in S, $p \in Ag$

(CBS-2) $S_- \models \neg p$ iff for every agent Ag in S, $p \notin Ag$

A formula A is *satisfiable in a closed society* if there exists S_- such that $S_- \models A$, and is said to be a *closed-tautology* if it is satisfiable in every closed society.

An interesting question concerns the notion of *representativeness*: for a finite number of agents, it is relevant to determine, for closed and open logic societies, how many agents can represent the whole society, or in other words, what is the effect of the cardinality of agents in the society. It can be proved that, for both open and closed societies, the rules adopted always make it possible to separate the agents into two blocks, which means that societies of these kinds can be replaced by societies having only two agents:

Theorem 1. *Let S_+ (respectively, S_-) be an open (respectively, closed) biassertive society. Then there exists an open (respectively, closed) biassertive society S_{2+} (respectively, S_{2-}) containing at most two agents such that $S_{2+} \models A$ iff $S_+ \models A$ (respectively, $S_{2-} \models A$ iff $S_- \models A$) for every formula A.*

Proof. A detailed argument is found in [2]. □

The above theorem permits to establish close connections between logic societies and finite-valued logics, by means of introducing hierarchies of societies.

3 Many-valued logics and logic societies

The three-valued system P^1 was introduced in [15] with the intention to define the simplest possible paraconsistent calculus. P^1 is a subsystem of CPC, and is maximal in the sense that, by adding to its axioms any classical tautology which is not a P^1-tautology, the resulting system collapses into CPC.

Axiomatically P^1 is characterized in the following way, in the language of CPC:[5]

[4] No connections to K. Popper's critique of historicism and his defense of liberal democracy, as in his *The Open Society and Its Enemies*.

[5] The original formulation of [15] includes an extra axiom which can be deduced from the ones given here.

(P^1-1) $A \to (B \to A)$

(P^1-2) $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

(P^1-3) $(\neg A \to \neg B) \to ((\neg A \to \neg\neg B) \to A)$

(P^1-4) $(A \to B) \to \neg\neg(A \to B)$

and *modus ponens* is the only inference rule. P^1 can be proved (cf. [15]) to be semantically characterized with respect to the following matrices, where $\to$ and $\neg$ are primitive, and $\wedge$ and $\vee$ are defined. The truth-values are T, T^*, F, of which T, T^* are distinguished. Intuitively, T and F mean plain truth and falsity, whereas T^* can be understood as "truth by default", or "by the lack of evidence to the contrary".

	T	T^*	F
$\overset{P^1}{\neg}$	F	T	T

$\overset{P^1}{\to}$	T	T^*	F
T	T	T	F
T^*	T	T	F
F	T	T	T

The primitive negation of P^1 is paraconsistent in the sense that, for example, $A \to (\neg A \to B)$ is not a P^1tautology, as it can easily be checked from the given matrices assigning the truth-value T^* to A and F to B. It is possible, however, to define in P^1 a *strong negation* $\bar{\neg} A$ which recovers the full power of classical negation: $\bar{\neg} A =_{def} \overset{P^1}{\neg}(\overset{P^1}{\neg} A \overset{P^1}{\to} A)$, giving the following table:

	T	T^*	F
$\bar{\neg}$	F	F	T

Using the strong negation, we can also define conjunction $A \overset{P^1}{\wedge} B$ and disjunction $A \overset{P^1}{\vee} B$ in P^1 as follows:

$$A \overset{P^1}{\wedge} B =_{def} \neg(A \to \bar{\neg} B)$$

$$A \overset{P^1}{\vee} B =_{def} (\bar{\neg} A \to B)$$

$\overset{P^1}{\wedge}$	T	T^*	F
T	T	T	F
T^*	T	T	F
F	F	F	F

$\overset{P^1}{\vee}$	T	T^*	F
T	T	T	T
T^*	T	T	T
F	T	T	F

P^1 has a dual, the system I^1 introduced in [16] as a three-valued counterpart of the system P^1. The truth values of I^1 are T, F^*, F, of which only T is distinguished. Intuitively, again T and F mean plain truth and falsity, whereas F^* can be understood as "false by default", or "by the lack of positive evidence".

The system I^1 instead of paraconsistent, is paracomplete in the sense that, for example, $\neg\neg A \to A$ is not an I^1 tautology, as can be checked from the matrices below, assigning the truth-value F^* to A. In I^1 all the axioms of the well-known Heyting system for intuitionistic logic are valid, and the law of excluded middle is not valid (for the disjunction defined below).

The axioms of I^1 (in the same language of CPC, having *modus ponens* as the only rule) are:

(I^1-1) $A \to (B \to A)$

(I^1-2) $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

(I^1-3) $(\neg\neg A \to \neg B) \to ((\neg\neg A \to B) \to \neg A)$

(I^1-4) $\neg\neg(A \to B) \to (A \to B)$

I^1 can be shown (cf. [16]) to be semantically characterized with respect to the matrices below, where $\to$ and $\neg$ are primitive connectives. As mentioned before, the truth-values are T, F^*, F, and T is the only distinguished value:

	T	F^*	F
$\overset{I^1}{\neg}$	F	F	T

$\overset{I^1}{\to}$	T	F^*	F
T	T	F	F
F^*	T	T	T
F	T	T	T

It is possible to define in I^1 a *dual strong negation* $\tilde{\neg}A$ which has all the properties of classical negation: $\tilde{\neg}A =_{def} A \overset{I^1}{\to} \overset{I^1}{\neg} A$ giving the following table:

	T	F^*	F
$\tilde{\neg}$	F	T	T

We can also define conjunction $A \overset{I^1}{\wedge} B$ and disjunction $A \overset{I^1}{\vee} B$ for this system in the following way:

$$A \overset{P^1}{\wedge} B =_{def} \neg(A \to \tilde{\neg}B)$$

$$A \overset{P^1}{\vee} B =_{def} (\tilde{\neg}A \to B)$$

$\overset{I^1}{\wedge}$	T	F^*	F
T	T	F	F
F^*	F	F	F
F	F	F	F

$\overset{I^1}{\vee}$	T	F^*	F
T	T	T	T
F^*	T	F	F
F	T	F	F

It can be shown that biassertive societies are essentially equivalent to three-valued logics. A detailed proof of the next theorem can be found in [2].

Theorem 2. *The logic of biassertive open (respectively, closed) societies is P^1. (respectively, I^1).*

Society semantics can be seen as a way to restore bivalence to certain logics, at the cost, however, of losing truth-functionality. The Polish logician Roman Suszko had already suggested that Lukasiewiczs logic L_3 could have a bivalent semantics, and in a famous paper in the 1970s even accuses Lukasiewicz of having perpetrated a fraud (or a humbug, in his words), but had not proposed a method for finding it. Several other results by R. Wójcicki, A. Lindenbaum N. da Costa and D. Scott show that any many-valued semantics can be reduced to a two-valued one (see [17] for an account of this chapter in the history of logic, and for a method to reduce many-valued semantics to two-valued equivalents).

As a semantic tool, society semantics are by no means devoted to paraconsistency or paracompleteness. Other logics can be seen to be characterized by society semantics: a nice example is Lukasiewiczs logic L_3. Based on [18] we define a *lukasiewiczian society* S_L as a closed society for atomic sentences, with the following clauses for non-atomic sentences:

(OBS-3) $S_L \models A \to B$ iff $S_L \models \neg A$ or $S_L \models B$, or

$S_L \not\models A, S_L \not\models \neg A, (S_L \not\models B, S_L \not\models \neg B$

(OBS-4) $S_L \models \neg(A \to B)$ iff $S_L \models A$ and $S_L \models \neg B)$

(OBS-5) $S_L \models \neg\neg A$ iff $S_L \models A$.

The proof in [18], although not complex, is worth commenting. By adapting Theorem 1 above for the case of closed societies, it is enough to consider at most two agents. Indeed, by defining agents $Ag1$ and $Ag2$ such that the variables accepted by the agents are: $Var1 = \{p : S_L \models p\}$ and $Var2 = \{p : S_L \models \neg p\}$, it is easy to check that $S_{2L} = \{Ag1, Ag2\}$ has the desired property. So no lukasiewiczian society needs to be represented by more than two agents.

Following [18], we call an L_3-*sociotautology* any sentence A such that for any lukasiewiczian society S, $S \models A$ holds. The next theorem establishes that L_3-sociotautologies coincide with the tautologies of the Lukasiewiczs logic L_3, considering the usual set of values $\{0, \frac{1}{2}, 1, \}$, the unary set of designated truth-values as $D = \{1\}$ and the usual definitions of implication and negation in L_3.

Theorem 3. *Given a lukasiewiczian society S, one may define a three-valued valuation v such that $v(A) \in D$ iff $S \models A$ for every sentence A in the language of L_3. Conversely, given a three-valued valuation v, one may define a lukasiewiczian society S such that $S \models A$ iff $v(A) \in D$ iff $S \models A$ for every sentence A in the language of L_3.*

Proof. A detailed argument is carried out in [18] divided into two propositions called, respectively, S-Convenience and S-Representability. □

It does not seem to be difficult to extend this to other many-valued logics in the hierarchy of Lukasiewicz, but a systematic study in this direction is still missing.

Several generalizations of the idea of society semantics, reformulating intuitions of society semantics in a wider setting, are done in [19], by characterizing hierarchies of paraconsistent logics called ***Pn*** and paracomplete logics ***In*** (for ***n*** any natural number). The paper also shows that a new hierarchy of logics called ***InPk***, which are simultaneously paraconsistent and paracomplete, can be characterized by society semantics as well.

4 Social choice theory and social software

Many-valued logics are not regarded nowadays as just a futile mathematical exercise or any abstruse philosophical delirium, but find practical use, for instance, in information processing, The idea of using logics with finite-valued semantics (especially three-valued and four-valued ones) in information processing is not new, and is justified by the practical necessity to deal with vast amounts of information, often incomplete or contradictory, provided by a multitude of various sources. The analysis in [20] shows that standard existential information processing (EIP) structures are not enough since there are EIP processor valuations which cannot be obtained from any finite number of source valuations.

The authors consider their work an extension of society semantics, and call attention to a point of further interest for the topic of the present paper, namely, the questions of collective judgment as connected to social choice theory, as they put it in [20], p. 1027:

> Further, the problems considered in our work bear an obvious relationship to the works on social choice ... where a group of individuals aggregates their individual judgments on some interconnected propositions into the corresponding collective judgment. Hence, another direction of future work would be to try to apply our approach to the problems of social choice

There is indeed much to be investigated concerning logic and judgment aggregation, as in [21] where the problem of manipulating rules for aggregating judgments on logically connected propositions is analyzed. Regardless of specific details, it seems that a logical approach to social situations presents itself as an important perspective, paraconsistent logic in particular for the cases where contradictory judgments may arise – and one should not forget, of course, that the contrast of judgments and opinions lies in the kernel of true democracy.

The term social software refers to the interest of analyzing social procedures by means of the formal methods of logic and computer science, as coined by Rohit Parikh in [22]. Examples of such formal methods applied to social questions are cake-cutting algorithms: for two people, the well-known algorithm "'I cut, you choose" is a simple method to ensure fairness, and similar strategies can be extended to more than two people. Voting procedures are another example, and connect the area with game theory, social choice theory and behavioral economics.

SOme modal logics (so-called coalition logics) intend to formalize reasoning about effectivity in game contexts, where the modality $[C]\varphi$ expresses that coalition C is effective for φ is proposed in [23]. Paraconsistency and its connection to social software is investigated in [24], where a pluralistic view on the topic is motivated by examples like Parrondo's paradox in game theory, where in some cases a combination of losing strategies can become a winning strategy. Suppose that there are two games to be played: Game 1, where you inescapably lose \$1 every time you play, and Game 2, which is equally perverse: if, after a move, you are left with an even number of dollars, you win \$3, but if you are left with an odd number of dollars, you lose \$5. If you start playing this game with \$10, for example, by playing Game 1 you will lose all your money in 10 rounds. On the other hand, by playing Game 2, you will also lose all your money in 10 rounds according to the sequence:

10 13 8 11 6 9 4 7 2 5 0.

However, if you play both games in the order of "Game 2 - Game 1 - Game 2 - Game 1 - ...., then you will always win according to the sequence:

10 13 12 15 14 17 . . .

The paradoxical result is the fact that by combining two losing strategies, an unexpected winning strategy emerges. By the point of view of society semantics, if yourself, acting as Agent 1 play Games 1 and acting as Agent 2 play Games 2, the society formed by the two agents (your strategic moves) completely change the situation. The collective intelligence formed by the two agents, in this case, cooperate in a formidable way, taking profit from an ostensively contradictory configuration.

5 ND-semantics and society semantics as specialized possible-translations semantics

The idea of non-deterministic (ND) semantics, introduced and developed by A. Avron and collaborators ([25], [26]), is a natural generalization of the notion of truth-functionality by means of more generalized matrices called Nmatrices. In a Nmatrix, the truth-value of a complex formula is chosen non-deterministically from some non-empty (usually finite) set of options. A legal valuation in a Nmatrix is a mapping from formulas to logical values satisfying some constraints. Nmatrices are a powerful tool, as they somehow conserve the advantages of ordinary many-valued matrices, while being applicable to a much wider range of logics; see, for instance, [27] where ND-semantics are, quite surprisingly, applied to characterize certain modal logics, known to be uncharacterizable by means of finite collections of finite matrices. Nmatrices are a descendant of the possible-translations semantics, introduced in [4], and can also be applied to LFIs and paraconsistent logics in general. Indeed, a proof showing that ND-semantics are a particular case of possible-translations semantics can be found in [28]. ND-semantics are thus sisters of society semantics ([2]) and are both special cases of possible-translations semantics, whose strong aspect lies in its generality. If not carefully treated, however, possible-translations semantics may be too general, and an ongoing project is to characterize relevant subclasses of such semantics. From this perspective, society semantics, ND-semantics and even dyadic semantics ([17]) can be characterized as particular forms of possible-translations semantics with specific restrictions on the translations.

6 Closing

We defended the view that the idea of society of things permeates several phenomena such as collective intelligence in several guises, social software, crowdfunding and convergence, the merging of approaches coming from engineering, physical sciences, human sciences and life sciences into a common model. Several examples helped to illustrate this tendency, which as we argue, can be in some cases formalized by means of the idea of society semantics, regarded as an ensemble of independent agents whose behavior defines the context of their collective aggregate. A survey of the main features of society semantics, their connections to non-deterministic semantics, to dyadic semantics and their common ancestrality from the broad notion of possible-translations semantics was discussed, as well as some connections between crowd intelligence and the paraconsistent reasoning paradigm, in view of the inherent risk of contradictions present in such aggregates.

The potentialities of such semantic aggregates (namely, society semantics, non-deterministic semantics, dyadic semantics and possible-translations semantics) for logic itself have been emphasized, and many problems connected to this area await to be tackled – perhaps one of most relevant being the relationship between possible-translations semantics and possible-worlds semantics: are those general semantic paradigms reduced one another? Are society semantics reducible to cases of possible-worlds semantics with a finite number of finite worlds? Those are questions for logic, but also for the foundations of collective intelligence, that we hope to have contributed to make more perspicuous.

Acknowledgments: Both authors acknowledge support from the Centre for Research on Architecture of Information, University of Brasilia, project sponsored by the Ministry of the Environment. The first author acknowledges support from FAPESP Thematic Project LogCons 2010/51038-0, Brazil and from the National Council for Scientific and Technological Development (CNPq), Brazil. We dedicate this paper to Luis Fariñas del Cerro for his unfaltering inquiry on the role of contradictions in computer science.

Bibliography

1. Malone, T.W., Laubacher, R., Dellarocas, C.: Harnessing crowds: Mapping the genome of collective intelligence. http://ssrn.com/abstract=1381502 (2010)
2. Carnielli, W., Lima-Marques, M.: Society semantics for multiple-valued logics. In Carnielli, W.A., D'Ottaviano, I.M.L., eds.: Advances in Contemporary Logic and Computer Science. Volume 235 of Contemporary Mathematics Series., American Mathematical Society (1999) 33–52
3. Lima–Marques, M.: De la connaissance à la paraconsistance : un modèle d'application pour la résolution de conflits aériens. PhD thesis, Université Paul Sabatier, Toulouse, France (1992)
4. Carnielli, W.: Manyvalued logics and plausible reasoning. In: Proceedings of the XX International Congress on ManyValued Logics, Charlotte, NC, USA, IEEE Computer Society (1990) 328–335

5. Carnielli, W.: Possible-translations semantics for paraconsistent logics. In Batens, D., Mortensen, C., Priest, G., Van Bendegem, J.P., eds.: Frontiers of Paraconsistent Logic: Proceedings of the I World Congress on Paraconsistency. Logic and Computation Series, Baldock: Research Studies Press, King's College Publications (2000) 149–163
6. Kahneman, D., Slovic, P., Tversky, A.: Judgment under Uncertainty: Heuristics and Biases. Cambridge University Press (1982)
7. Surowiecki, J.: The Wisdom of Crowds. Anchor (2005) Reprint edition.
8. Sharp, P.A., Conney, C.L., Kastner, M.A., Lees, J., Sasisekharan, R., Yaffe, M.B., Bhatia, S.N., Jacks, T.E., Lauffenburger, D.A., Langer, R., Hammond, P.Y., Sur, M.: The third revolution: The convergence of the life sciences, physical sciences and engineering. MIT Report, `http://dc.mit.edu/sites/dc.mit.edu/files/MIT%20White%20Paper%20on%20Convergence.pdf` (2011)
9. Herzog, S.M., Hertwig, R.: The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. Psychological Science **20** (2009) 231–237 doi:10.1111/j.1467-9280.2009.02271.x.
10. Carnielli, W.A., Coniglio, M.E., Marcos, J.: Logics of formal inconsistency. In Gabbay, D., Guenthner, F., eds.: Handbook of Philosophical Logic. Volume 14., Amsterdam, Springer-Verlag (2007) 1–93
11. Carnielli, W.A., Coniglio, M.E., Gabbay, D., Gouveia, P., Sernadas, C.: Analysis and Synthesis of Logics. Springer (2007)
12. Balbiani, P., Herzig, A., Lima–Marques, M.: TIM: The Toulouse Inference Machine for non-classical logic programming. In Richter, M., Boley, H., eds.: Processing Declarative Knowledge. Number 567 in Lecture Notes in Artificial Intelligence. Springer-Verlag (1991) 365–382
13. Alliot, J.M., Herzig, A., Lima–Marques, M.: Implementing Prolog extensions: a parallel inference machine. In: International Conference on Fith Generation Computer Systems, Tokyo, Institut for New Generation Computer Technology - ICOT (1992)
14. Jaśkowski, S.: Un calcul de propositions pour les systèmes deductifs contraditoires. Studia Soc. Scien. Torunensis (1948) 55–77 English translation in Studia Logica v. 24, 1968, 143-15.
15. Sette, A.M.: On the propositional calculus P^1. Mathematic Japonicæ **18**(13) (1973) 173–180
16. Sette, A.M.A., Carnielli, W.A.: Maximal weakly-intuitionistic logics. Studia Logica **55** (1995) 181–203
17. Caleiro, C., Carnielli, W.A., Coniglio, M.E., Marcos, J.: Two's company: "the humbug of many logical values". Logica Universalis (2005) 169–189
18. Marcos, J.: Many values, many semantics. Draft (2000)
19. Fernández, V., Coniglio, M.E.: Combining valuations with society semantics. Journal of Applied Non-Classical Logics **13**(1) (2003) 21–46
20. Avron, A., Konikowska, B.: Finite-valued logics for information processing. Fundamenta Informaticae **21** (2001) 1001–1028
21. Dietrich, F., List, C.: Strategy-proof judgment aggregation. Economics and Philosophy **23** (2007) 269–300
22. Parikh, R.: Social software. Synthese **132** (2002) 187–211
23. Pauly, M.: A modal logic for coalitional power in games. Journal of Logic and Computation **12** (2002) 149–166
24. Baskent, C.: A non-classical logical approach to social software. In Rohit Parikh on Logic, Language and Society, Edited by C. Baskent, L. Moss and R. Ramanujam, Springer, forthcoming, available at `http://www.canbaskent.net/logic/socialsoftware.pdf` (2016)
25. Avron, A.: Non-deterministic matrices and modular semantics of rules. Logica Universalis (2007) 155–173
26. Avron, A., Zamansky, A.: Non-deterministic semantics for logical systems. In Gabbay, D., Guenthner, F., eds.: Handbook of Philosophical Logic. Volume 16., Amsterdam, Springer-Verlag (2011) 227–304
27. Coniglio, M.E., del Cerro, L.F., Peron, N.M.: Finite non-deterministic semantics for some modal systems. Journal of Applied Non-Classic **25** (2015) 149–166
28. Carnielli, W.A., Coniglio, M.E.: Splitting logics. In Artemov, S., Barringer, H., Garcez, A., Lamb, L., Woods, J., eds.: We Will Show Them! Essays in Honour of Dov Gabbay. Volume 1., London, College Publications (2005) 389–414

Stable Reasoning

Pedro Cabalar[1], David Pearce[2], and Agustín Valverde[3]

[1] University of Corunna, Spain, cabalar@udc.es
[2] Universidad Politécnica de Madrid, Spain, david.pearce@upm.es
[3] Universidad de Málaga, Spain, a_valverde@ctima.uma.es

Abstract. We give an account of stable reasoning, a recent and novel approach to problem solving from a formal, logical point of view. We describe the underlying logic of stable reasoning and illustrate how it is used to model different domains and solve practical reasoning problems. We discuss some of the main differences with respect to reasoning in classical logic and we examine an ongoing research programme for the rational reconstruction of human knowledge that may be considered a successor to the logical empiricists' programme of the mid-20th Century.

1 INTRODUCTION

For most of the 20th Century logic as a scientific discipline was dominated by the paradigm of Frege and Hilbert and the aim of providing mathematics with a secure logical foundation. This challenging goal continued to dominate research programmes in logic long after Gödel's incompleteness theorems showed that the initial expectations were unreachable. Above all the Frege-Hilbert paradigm was and still is based on a standardised, classical conception of logic, on the axiomatic method as the basis for reconstructing mathematical knowledge and on the method of deduction as the central element in logical theorem proving.

The paradigm of Frege and Hilbert was challenged by various critics throughout its lifetime but only in recent years has it been eclipsed by more progressive research programmes within computational logic. A lively and searing critique of the Frege-Hilbert paradigm was published by the logician Carlo Cellucci in 1998 [6]. Cellucci calls the standard paradigm *mathematical logic* and discusses at length many of its significant features. Besides its emphasis on providing a secure foundation for mathematics, key features include the prominence of the axiomatic method, the idea of theories as embodying certain truths and theorems as logical deductions from these. Cellucci questions many of the assumptions of the paradigm, especially the manner in which it focuses on closed conceptual systems and problems of justification, while paying almost no attention to problems of (mathematical) discovery that may involve hypothetical reasoning, induction, abduction, analogy, heuristics and other methods. As a contrast to the axiomatic method Cellucci devotes much attention to describing and motivating the analytic method in mathematical discovery and problem solving[4] and justifying the importance of treating open systems and fallible reasoning.

While Cellucci doesn't offer a detailed description of modern *computational logic*, it is evident that some of the positive features of the analytic method can be found in computational approaches to logic and logic-based programming languages. He himself cites with approval six different features of Prolog that mirror aspects of the analytic method and reasoning with open systems (while at the same time noting that there are other important features of the analytic method that are not captured in Prolog). Since the publication of [6] computational logic has made many advances. One of them has been the elaboration of a new approach to logic-based programming known as *answer set programming* or ASP and the development of its underlying logical paradigm which we will call *stable reasoning*. Unlike Prolog, this is not a fully-fledged programming language, but rather a general approach to logic-based problem solving that can also be efficiently implemented in (answer set) solvers.

We cannot claim that stable reasoning currently satisfies all the requirements proposed by Cellucci for constituting a new and wholly adequate paradigm for logic. For one thing, the focus in [6] is on mathematics and mathematical problem-solving and discovery, while stable reasoning has a broad range of applications to many areas of inquiry. However, we do suggest that

1. it is distinctly different from the Frege-Hilbert paradigm of mathematic logic;

[4] Closely associated with Plato and other classical scholars.

2. its approach to problem-solving has significant points in common with the analytic method;
3. it is able to embrace various aspects and methods of discovery;
4. it is able to deal with dynamical and open systems.

In this paper we will discuss stable reasoning, focussing mainly on items 1 and 2 above, while mentioning 3 and 4 briefly towards the end.

Stable reasoning is a recent and novel approach to problem solving from a formal, logical point of view. An important difference compared to reasoning based on classical logic is that stable reasoning can take account of default assumptions and conclusions that follow from them. It is not based on two-valued classical logic, since this does not allow for the distinction between certain truth and truth-by-default or the kind of truth that can be assumed when there is no evidence to the contrary. To account for defaults in the setting of classical logic usually special syntactic devices are employed or one has to distinguish between different kinds of inference, some defeasible others not. Stable reasoning does not require any special inference rules or syntactic devices because it is based on a many-valued logic where precisely the distinction between certain truth and truth-by-default can be made. There is just one basic kind of negation and one kind of inference.

There is another important difference compared to the axiomatic tradition of Frege and Hilbert that dominated the methodology of formal logic for much of the 20th Century. Stable reasoning is closer to what might be termed a problem-solving approach to formal reasoning. In the axiomatic tradition a mathematical or empirical domain is formalised by introducing a language or vocabulary and a set of sentences (axioms) of the language intending to capture once and for all the entirely of knowledge governing that domain. Mathematical theorems are inferred through logical deduction from the axioms. Predictions or explanations from empirical theories are also deduced from their axioms once the initial conditions of a system are specified.

By contrast, stable reasoning is problem-driven. One wishes to find a solution or several possible solutions to a certain problem. One describes the problem domain by specifying the entities and relations that govern the domain. This description need not be complete nor need it capture the entirety of knowledge once and for all, but merely offer hypotheses sufficient to produce adequate answers or solutions. Some descriptions may indeed be robust and reusable, others may be fragile and of temporary validity/use. As in the classical, axiomatic case, answers will be produced once certain facts or initial conditions are specified. These answers or solutions are in general model-based. They are produced not by means of logical deduction or inference from axioms plus initial conditions, but rather by computing a model or state of affairs that embodies the solution in some obvious way.

To illustrate the difference between the two orientations, consider the following well-known graph-theoretical problem.

Example 1 (Hamiltonian cycles). A *Hamiltonian cycle* of a graph G is a cyclic path that traverses each node in G exactly once. The same graph may have different Hamiltonian cycles or none at all. □

Suppose we want to obtain the Hamiltonian cycles of graph G in Figure 1.1. To this aim we must decide:

(a) how to represent the graph;
(b) how to represent the cycles;
(c) the type of logical reasoning task to obtain the desired solution.

Regarding (a), the obvious solution in Predicate Calculus is just using a pair of predicates, say $Node(x)$ and $Edge(x, y)$, to respectively describe the nodes and edges of G. Still, even at this elementary level, a first difference between the classical approach and stable reasoning already arises. In classical logic, if we just list the set of ground atoms:

$$\begin{aligned}&Node(0), \quad Node(1), \quad Node(2), \quad Node(3),\\ &Edge(0,1), \quad Edge(1,2), \quad Edge(1,3), \quad Edge(2,0),\\ &Edge(2,3), \quad Edge(3,2), \quad Edge(3,0).\end{aligned} \tag{1.1}$$

this will leave free the possibility for other objects of becoming nodes or edges – we would have models where, for instance, $Edge(0, 2)$ could become true. In order to precisely capture the nodes and edges we

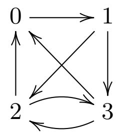


Fig. 1.1. A simple graph with a pair of Hamiltonian cycles.

should actually use instead the pair of formulas[5]:

$$\begin{aligned} Node(x) &\leftrightarrow x = 1 \vee x = 2 \vee x = 3 \vee x = 4 && (1.2) \\ Edge(x, y) &\leftrightarrow (x = 0 \wedge y = 1) \vee (x = 1 \wedge y = 2) \\ &\quad \vee (x = 1 \wedge y = 3) \vee (x = 2 \wedge y = 0) \\ &\quad \vee (x = 2 \wedge y = 3) \vee (x = 3 \wedge y = 2) \vee (x = 3 \wedge y = 0) && (1.3) \end{aligned}$$

In the stable reasoning approach, however, the set of facts (1.1) would suffice because predicates are subject to the so-called *Close World Assumption* (CWA). Informally speaking, this means that anything not explicitly stated will be false by default. We could, of course, still use the stronger version (1.2) $\wedge$ (1.3) for stable reasoning, but the set of facts (1.1) is more flexible in the sense that new nodes or edges can be included by the simple *addition* of new formulas, rather than manipulating the existing ones (this feature is often referred as *elaboration tolerance* [20]). Predicates like $Node(x)$ or $Edge(x, y)$ whose extensions are defined by enumerating their lists of (true) atoms conform what is usually called the *extensional database* (a term inherited from Database theory) as opposed to those described by additional conditional formulas (or *rules* in Logic Programming) which receive the name of *intensional predicates*.

Another important feature about CWA is that, if necessary, we can remove this assumption in a selective way for any formula ϕ or, in particular, any atom $P(\overline{x})$. To this aim, it just suffices with adding an axiom like:

$$P(\overline{x}) \vee \neg P(\overline{x}) \qquad (1.4)$$

Note that although (1.4) has the form of a classical tautology (the so-called *excluded middle axiom*), it is not a tautology for stable reasoning, since negation here has a different meaning. As said before, the effect of (1.4) is that CWA for predicate P is removed, so it behaves in a "classical" way. For instance, adding the axiom $Node(x) \vee \neg Node(x)$ would remove $Node$ to be false by default.

Let us move now to consider problem (b), that is, how to represent Hamiltonian cycles. As any of them visits all nodes in the graph, we obviously must refer to the set of edges involved in the cycle to differentiate one from each other. In classical logic, we would typically represent sets of edges using some additional notation: assume, for instance, that standard set terms are allowed. Then, we would use some predicate, say $HamCycle(s)$, to represent that the set s of edges constitutes a Hamiltonian cycle for the graph, including a hypothetical set of axioms $\Gamma(HAM_G)$ that includes the graph description (1.2) $\wedge$ (1.3) and the meaning of predicate $HamCycle$. Under the axiomatic method, our problem (c) would then reduce to decide for which sets s of edges we can derive $\Gamma(HAM_G) \vdash HamCycle(s)$. For instance, in our example, we should conclude that this holds for $s = \{(0, 1), (1, 3), (3, 2), (2, 0)\}$ and for $s = \{(0, 1), (1, 2), (2, 3), (3, 0)\}$. It is worth to notice that as these two facts for predicate $HamCycle(s)$ are derived as theorems, they will be simultaneously true in *all models* of $\Gamma(HAM_G)$. In the general case, the extension of $HamCycle$ would contain the whole set of Hamiltonian cycles in *any* model of $\Gamma(HAM_G)$.

Under the stable reasoning approach, tasks (b) and (c) become model-oriented, rather than theorem-based. One of the main features of this methodology is that we identify each solution of the original problem with (a distinguished part of) each model of our logical representation. For instance, in our example, rather than collecting all cycles in a single predicate $HamCycle(s)$, our purpose is obtaining a *different model* per each possible Hamiltonian cycle of the original graph. Using this approach, we need not dealing with sets any more: we can just collect the edges *in* the current cycle (that is, the cycle represented inside the current model) as the extent of some predicate, call it $In(x, y)$. Thus, if we call again $\Gamma(HAM_G)$ to our logical representation of the problem, we are interested in obtaining the different models M such that $M \models \Gamma(HAM_G)$ so that each Hamiltonian cycle can be directly obtained by selecting some relevant information in each obtained model

[5] We assume that free variables, like x, y here, are implicitly universally quantified.

M. In our example, we should obtain a pair of models M_1 and M_2 such that:

$$M_1 \models In(0,1) \wedge In(1,3) \wedge In(3,2) \wedge In(2,0)$$
$$M_2 \models In(0,1) \wedge In(1,2) \wedge In(2,3) \wedge In(3,0)$$

and no other atoms for $In(x, y)$ hold in each model.

Let us consider now how to specify $\Gamma(HAM_G)$ under the stable reasoning approach. Typically, we would first consider models for all possible subsets of edges, and then include additional formulas that rule out the ones that do not correspond to Hamiltonian cycles. In a first attempt, we could include:

$$Edge(x, y) \rightarrow \big(In(x, y) \vee \neg In(x, y)\big) \tag{1.5}$$
$$In(x, y) \wedge In(x, z) \rightarrow y = z \tag{1.6}$$
$$In(y, x) \wedge In(z, x) \rightarrow y = z \tag{1.7}$$

together with a graph description, that is, an extensional database like (1.1).

As a first important remark, note that the consequent $In(x, y) \vee \neg In(x, y)$ in (1.5) has the form of an excluded middle formula like (1.4). As we explained before, this means that predicate $In(x, y)$ will not be subject to CWA, provided that x, y is a pair of nodes forming an edge. As a result, the effect of (1.5) is that we would have a model per each possible subset of edges in the graph. Formula (1.6) (resp. (1.7)) specifies that we never pick two different outgoing (resp. incoming) edges for a given node x. These two formulas can also be represented by using $\bot$ in the consequent (formulas like these receive the name of *constraints*):

$$In(x, y) \wedge In(x, z) \wedge y \neq z \rightarrow \bot$$
$$In(y, x) \wedge In(z, x) \wedge y \neq z \rightarrow \bot$$

This first attempt (1.5)-(1.7), however, is not enough to capture Hamiltonian cycles yet. We can still get disconnected groups of edges like, for instance, $\{(0, 1), (2, 3), (3, 2)\}$ or even nodes that are not connected at all. To rule out these cases, we should further specify that any node can be *reached* from another one. Curiously, this property (graph reachability) is a well-known example of a problem that *cannot be represented* in classical First Order Logic. So, in fact, if we tried to represent Hamiltonian cycles using the classical axiomatic method, $\Gamma(HAM_G)$ should be represented in Second Order Logic. Under stable reasoning, however, we can easily capture reachability with an auxiliary predicate $Reach(x, y)$ defined with the pair of formulas:

$$In(x, y) \rightarrow Reach(x, y) \tag{1.8}$$
$$In(x, z) \wedge Reach(z, y) \rightarrow Reach(x, y) \tag{1.9}$$

Since $Reach(x, y)$ is subject to CWA, these two formulas behave as an inductive definition, as those frequently used in Mathematics. In other words, $Reach(x, y)$ is true if: (1) we have taken the edge $In(x, y)$; or (2) we took an edge to some z and we can reach y from that z, $Reach(z, y)$. Otherwise, $Reach(x, y)$ will be false, due to CWA. A predicate like this, whose extension is determined by some rules like those above in combination with CWA, receives the name of *intensional* predicate.

To complete our example, we would just further need to rule out models where we can find a pair of nodes that are not reachable one from each other. This can be done using the formula:

$$\neg\exists x, y(\ Node(x) \wedge Node(y) \wedge \neg Reach(x, y)\ )$$

or its equivalent representation as a constraint:

$$Node(x) \wedge Node(y) \wedge \neg Reach(x, y) \rightarrow \bot \tag{1.10}$$

Again, it is worth to note that, while (1.10) is classically equivalent to:

$$Node(x) \wedge Node(y) \rightarrow Reach(x, y)$$

the latter has a quite different meaning under the stable reasoning approach – we would read it as a third option for defining $Reach(x, y)$ facts, and in this case, we would get that any pair of nodes would always be reachable, regardless the edges we had in the graph.

Proposition 1. *Let $\Gamma(G)$ be the set of formulas* (1.5)-(1.10) *plus an extensional database for graph G like* (1.1)*. Then, M is a stable model of theory $\Gamma(G)$ iff $\{(x, y) \mid M \models In(x, y)\}$ is a Hamiltonian cycle for G.* □

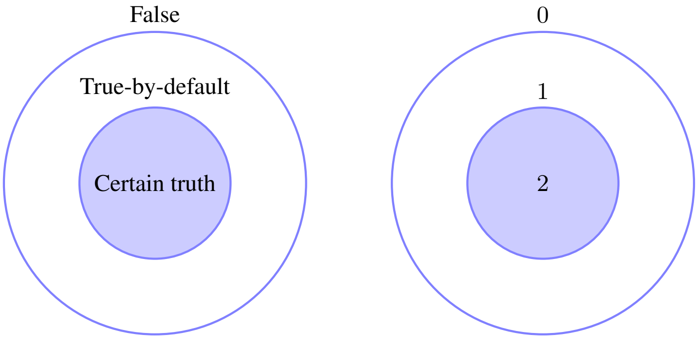


Fig. 1.2. Kinds of truth in stable reasoning.

$\wedge$	0	1	2
0	0	0	0
1	0	1	1
2	0	1	2

$\vee$	0	1	2
0	0	1	2
1	1	1	2
2	2	2	2

$\to$	0	1	2
0	2	2	2
1	0	2	2
2	0	1	2

	$\neg$
0	2
1	0
2	0

Fig. 1.3. Truth tables for the conectives.

2 A logic from first principles

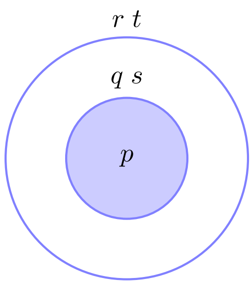
Stable reasoning has very simple logical underpinnings. Its base logic can be constructed from first principles in a few easy steps and with only a few fundamental assumptions. The first idea is that formulas may be true or false, but that there are two kinds of truth: certain truth and truth-by-default. We can picture this as shown in the left hand side of the Figure 1.2.

Suppose that the disk represents the collection of true formulas; while outside the outer circle formulas are false. True formulas come in two kinds: those that are only weakly true or more particularly *true by default* lie within the unshaded, outer part of the disk, those lying within the shaded, inner circle are *certain*, or *true in the strong sense*. Truth in the weak or general sense covers both certain truth and truth by default, so the outer circle contains all the formulas of the disk, both inner and outer parts.

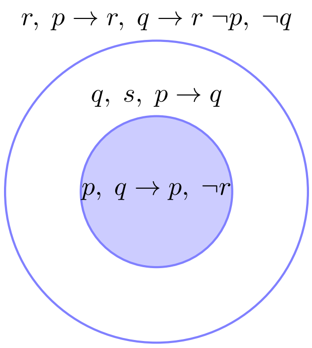
We can assign numbers to these different semantic properties, say 0 for falsity and 2 for certain truth. We assign the value 1 to formulas that lie within the outer but not the inner circle: they are only true in the weaker but not the stronger sense. The pictures looks like as shown in the right hand side of the figure 1.2.

Let us say that our base logic comprises, in the propositional case, the usual set of logical connectives $\{\wedge, \vee, \to, \neg\}$, standing for conjunction, disjunction, implication and negation. We may also make use of the falsum constant, $\perp$. There is an infinite set $Prop$ of propositional atoms from which formulas are constructed in the usual way, as in classical or intuitionistic logic. Let $p, q, r, \ldots$ stand for atomic propositions. As prescribed, the semantics is given by the three truth-values, 2, 1 and 0.

In the picture on the right p is certain, q and s are true by default and r and t are false. The values for complex formulas are assigned according to the usual meaning of connectives. In the picture $p \wedge q$ takes the value 1, since only p is certain, while q is not. On the other hand the fact that p is certain is sufficient for assigning $p \vee q$ and even $p \vee r$ the value 2. Since r is false, also $p \wedge r$ and $q \wedge r$ must be false, while $q \vee r$ takes the value 1.



Following this reasoning we can build the complete truth tables for conjunction and disjunction (see Figure 1.3); they are similar to those of other well-known logics.



Some care has to be taken with the values of formulas of form $A \to B$. In the picture on the left we would expect both $p \to r$ and $q \to r$ to be false since in both cases we have a true antecedent and false consequent. Notice that although q is not certain, we cannot put $q \to r$ in the inner circle since that would make it strongly true and yet false at the same time. Since r is false, evidently both $r \to q$, $r \to p$ and $r \to t$ are true, and even in the strong sense. Since p is certain, it is also true in the general sense and therefore $p \to q$ must be at least weakly true. But it is not certain, since q is not. By contrast, since p is certain, $q \to p$ is also. Lastly, consider the case of two formulas q, s that lie only in the outer circle. Since neither atom is certain, but neither is false, the implications $q \to s$ and $s \to q$ can safely be placed in the inner circle of certainty.

The complete truth matrix for implication is given in the Figure 1.3. The case of negation is straightforward. Negating a false proposition such as r produces a certain proposition $\neg r$; while negating either kind of true formula produces a false one: in the picture both $\neg p$ and $\neg q$ will be false.

Equivalently we can regard negation to be definable, by $\neg p := p \to \bot$, where $\bot$ takes the constant value 0. Then the table for negation is derivable from that for implication.

Our base logic is by no means new. It was first introduced by Heyting in his study of intuitionistic logic [11]. Shortly after it reappeared in Gödel's paper [10] showing that intuitionistic logic is not tabular (finite-valued). Heyting provided the truth matrices while subsequently Łukasiewicz [13] studied the logic in greater depth and gave the first axiomatisation based on the axioms and rules of the intuitionistic calculus extended by the addition of a weaker form of Peirce's Law, viz:

$$(\neg\alpha \to \beta) \to (((\beta \to \alpha) \to \beta) \to \beta). \tag{1.11}$$

Subsequently, the logic was studied by Smetanich [18] and Umezawa [19] who gave an alternative axiom to that of Łukasiewicz:

$$\alpha \vee \neg\beta \vee (\alpha \to \beta)$$

The completeness of this system was then proved by Hosoi [12].

Although Heyting made use of the logic as a technical device, it was clear from the start that it was of interest not only from a purely formal point of view. Łukasiewicz made a detailed comparison with his own 3-valued logic and Heyting already provided a natural interpretation of the third truth-value, claiming that it applies to a correct proposition that cannot be false but whose correctness cannot be proved. If we re-phrase this in terms of truth, then the interpretation 'true-by-default' is a natural one since it conveys the idea of a judgement that is accepted although not formally derivable in the system at hand.[6]

Our logic is known under a variety of names, we prefer to call it the logic of *here-and-there*, in symbols **HT**, whose etymology is easily explained by looking at an alternative way to describe the semantics. It is well-known that intuitionistic logic is complete for possible world models, that is triples of the form $\langle W, \leq, I\rangle$, where W is a non-empty set (of points, states or worlds), $\leq$ is a partial ordering on W, and I is an interpretation function assigning a set of (verified) atoms to each $w \in W$, such that $I(x) \subseteq I(y)$ whenever $x \leq y$. The logic **HT** is also complete for a class of such models, but of an especially simple form: W comprises just two points, say h ('here') and t ('there'), with $h \leq t$. It follows that we can represent a model simply as an ordered pair $\langle H, T\rangle$ of sets of atoms, where $I(h) = H$ and $I(t) = T$. Evidently we always have $H \subseteq T$. These are then *here-and-there* models. Applying the usual semantics to evaluate formulas at worlds one can easily verify the correspondence to our earlier truth-tables. Note that in a here-and-there model $\langle H, T\rangle$, H represents the certain atoms having value 2, while T represents the non-false atoms (of value 1 or 2). Those in $T \setminus H$ are true-by-default, ie take the value 1. The complement of T in $Prop$ is the set of atoms that are false in the model, ie those corresponding to the value 0 in our truth tables. We adopt the usual convention and say that a formula φ is satisfied or holds in a model $\langle H, T\rangle$, in symbols $\langle H, T\rangle \models \varphi$, if it is strongly true or certain in the model, i.e. is satisfied at the world h.

As a simple example, consider how to evaluate an implication $p \to q$ in a model $\langle H, T\rangle$, for atomic p, q. By the possible world semantics, this formula is true at t if $q \in T$ or $p \notin T$. In terms of truth-values it means

[6] Strictly speaking we should say that the base logic gives a monotonic approximation to reasoning by default. There is an additional aspect to defaults that emerges in the nonmonotonic extension of the logic.

that $p \to q$ takes a value of at least 1 if $q \in T$ no matter what value is given to p. However in the case that $p \notin T$, then also $p \notin H$. This means that $p \to q$ is true also at h and so the value of $p \to q$ is then 2.

Much is known about $\mathbf{HT}$ and its properties. It is one of 7 superintuitionistic (SI) logics having the interpolation property, [14], and it is also the strongest SI-logic properly contained in classical logic, all other SI-logics being properly contained in it.

3 Basic stable reasoning

The core logic of stable reasoning is based on $\mathbf{HT}$ but is actually a non-monotonic extension of it. It can be characterised in terms of a preferential entailment relation in the sense of Shoham (1988) defined on here-and-there models. The preference relation can be explained by considering some features of defaults.

When we build up a partial description of the world or of our problem domain we specify a set of formulas that we suppose, at least hypothetically, to be true in the certain sense. Call this our 'theory'. When we then consider the three-valued $\mathbf{HT}$-models of this theory (assuming it is consistent), they in turn generally admit formulas that are true only in the weak or default sense. We typically have models $\langle H, T\rangle$ where $H \subset T$ and then some formulas will be true only in the default sense, for instance all the propositions in $T \setminus H$. Equally, by consistency, we will also have models of the form $\langle T, T\rangle$ where all truths are certain. Our preference condition is to select just those models $\langle T, T\rangle$ where all truths are certain *and our theory does not admit any model with uncertainty whose true atoms are exactly* T; in other words there is no model of our theory of the form $\langle H, T\rangle$, where $H \subset T$. In this sense our theory justifies the choice of T by not accepting any model where T forms the set of truths in the general sense and some of them are only true in the weak sense.

So here is our definition in full. A model $\langle H, T\rangle$ of a theory $\mathcal{T}$ is said to be an *equilibrium* model of $\mathcal{T}$, if (i) $H = T$ and (ii) for any H such that $H \subset T$, $\langle H, T\rangle \not\models \mathcal{T}$. The term equilibrium model derives from [15], and the ensuing logic associated with this is equilibrium logic. However there is complete agreement between equilibrium models and the stable models of logic programs as defined by Gelfond and Lifschitz (1988). That is to say, if the formulas of a theory have precisely the shape of rules of logic programs, then a set of atoms T is a stable model of the theory if and only if $\langle T, T\rangle$ is an equilibrium model of it.[7] For this reason equilibrium logic can serve as a foundation for stable reasoning.[8] Though they are defined differently, we often use the terms stable model, answer set and equilibrium model inter-changeably.

4 Implementation of stable reasoning

Just as classical deduction is implemented in automated provers such as Prover9 or SAT-solvers, so stable reasoning has been implemented in various systems, collectively known as *answer set solvers*. These have obtained a fairly high degree of efficiency and sophistication and can be used to model and solve real-world problems in domains such as software verification, security and configuration management, model checking, agent technologies, constraint satisfaction, reasoning for the semantic web, software synthesis from specifications, knowledge representation, data and information integration, planning and diagnosis. The inputs to answer set solvers are called *answer set programs*, the branch of computer science and programming dealing with these is called *answer set programming*, or ASP for short. A closely related domain is that of DATALOG and deductive database systems. Indeed data and information management is one of the principal application areas of ASP, and one that currently enjoys some commercial success.

While they are successful in supporting real-world problem solving, ASP systems do not implement deduction in a formal logic. They are not designed to deduce theorems or prove the correctness of logical inferences, even though the solutions they compute can be precisely understood in terms of formal models of a logical system of deduction. This correspondence to logic however was a more recent discovery and did

[7] In logical notation the rules of a (disjunctive) logic program have the form

$$b_1 \wedge \ldots \wedge b_m \wedge \neg b_{m+1} \wedge \ldots \wedge \neg b_n \to a_1 \vee a_2 \vee \ldots \vee a_k$$

where the a_i, b_j are atoms. If $k = 1$ everywhere the program is said to be *normal*.

[8] Since the late 1990s the stable model semantics has been systematically extended to embrace wider classes of formulas, more recently including arbitrary propositional and even first-order theories. All the widely accepted extensions have coincided with equilibrium logic.

not directly guide the initial ASP implementations. While in traditional logic programming, systems of logical inference came first and computer programming applications came after, in ASP this order was reversed. The theory preceded the computer implementations that in turn preceded the logic. However, nowadays the logic exerts a growing influence on the development and comprehension of new systems, especially extensions of the initial ASP family of languages. It is also fundamental to understanding stable reasoning from a foundational point of view.

5 Practical examples

Let us now consider some additional examples of typical commonsense reasoning tasks that can be easily represented under the stable reasoning approach. We also show in some cases how these examples can be implemented using existing solvers. For instance, Figure 1.4 shows a possible implementation of the Hamiltonian cycles problem for graph in Figure 1.1, using the input language of the `DLV` solver[9]. The column in the right shows the correspondence to each formula in $\Gamma(HAM_G)$. The translation of our example theory into DLV language follows some standard syntactic conventions from Logic Programming and, in this case, is quite straightforward. Variables begin with uppercase letters and predicates with lowercase letters. Implications are reversed, so that α :- β stands for $\beta \to \alpha$. When $\alpha = \bot$ it is just omitted. Finally, a comma represents a conjunction, the symbol '`v`' represents a disjunction and all formulas are ended by a full stop.

```
node(0). node(1). node(2). node(3).
edge(0,1). edge(1,2). edge(1,3).                (1.1)
edge(2,0). edge(2,3). edge(3,2). edge(3,0).

in(X,Y) v out(X,Y) :- edge(X,Y).                (1.5)

:- in(X,Y), in(X,Z), Y!=Z.                      (1.6)
:- in(Y,X), in(Z,X), Y!=Z.                      (1.7)

reach(X,Y) :- in(X,Y).                          (1.8)
reach(X,Y) :- in(X,Z), reach(Z,Y).              (1.9)

:- not reached(X,Y), node(X), node(Y).          (1.10)
```

Fig. 1.4. Hamiltonian cycle representation $\Gamma(HAM_G)$ for the graph in Figure 1.1 written for `DLV`.

The only newly introduced feature with respect to the first order theory $\Gamma(HAM_G)$ is the use of an auxiliary predicate $Out(x, y)$ to replace the formula $\neg In(x, y)$ since `DLV` does not accept negation to the left of :- operator. This replacement is correct provided that $Out(x, y)$ is not used elsewhere. Otherwise, to fully replace the original formula we should include the pair of rules:

```
out(X,Y) :- not in(X,Y), edge(X,Y).
:- in(X,Y), out(X,Y), edge(X,Y).
```

which are logically equivalent to:

$$Edge(x, y) \to \big(\, Out(x, y) \leftrightarrow \neg In(x, y)\,\big)$$

A fundamental property of stable reasoning we have not exploited yet in the previous example is the use of default negation for non-monotonic reasoning (NMR). To illustrate this concept, consider the following classical example in the NMR literature. We want to capture the default "birds typically fly" and the exception "penguins are birds but do not fly." These two assertions can be respectively encoded as:

$$Bird(x) \wedge \neg CannotFly(x) \to Flies(x) \quad (1.12)$$

$$Penguin(x) \to Bird(x) \wedge CannotFly(x) \quad (1.13)$$

[9] `http://www.dlvsystem.com`

Suppose we have theory Γ' containing the two formulas above plus the ground atom $Bird(Tweety)$. As there is no evidence that $Tweety$ is a penguin, we derive the literal $\neg Penguin(Tweety)$ by CWA. This falsifies the antecedent of (1.13), and so no evidence on $CannotFly(Tweety)$ can be obtained from (1.13). In fact, no evidence on $CannotFly(x)$ can be obtained from (1.12) either. This is because implication in stable reasoning acquires a kind of directionality. In particular, (1.12) should be read as a ***definition rule*** for $Flies(x)$, saying that it will hold when x is a bird for which we have no evidence on $CannotFly(x)$. As a result, we conclude $\neg CannotFly(Tweety)$ by CWA and then, $Flies(Tweety)$ from (1.12).

This simple example became a challenge for NMR approaches because a wrong predicate minimisation policy could easily lead to models where $\neg Flies(Tweety)$ was decided first and then $CannotFly(Tweety)$ was derived by applying (1.12) as the classically equivalent formula:

$$Bird(x) \wedge \neg Flies(x) \rightarrow CannotFly(x)$$

In stable reasoning, such an equivalence does not hold. In fact, the formula above has a quite different reading from (1.12): it states that a bird x cannot fly if we cannot find any evidence for $Flies(x)$.

To complete the example, consider now the extended theory

$$\Gamma'' = \Gamma' \cup \{Penguin(Tweety)\}.$$

Since we do obtain now evidence on $CannotFly(Tweety)$ from (1.13), we lose the justification for $Flies(Tweety)$ we obtained before from (1.12). As a result, $Flies(Tweety)$ is not derived any more. This shows the non-monotonic nature of stable inference, since the addition of a new formula, $Penguin(Tweety)$, has made a previous conclusion $Flies(Tweety)$ to be retracted.

The representation of this example in `DLV` is shown below:

```
flies(X) :- bird(X), not cannotfly(X).
bird(X) :- penguin(X).
cannotfly(X) :- penguin(X).
bird(tweety).
penguin(tweety).
```

6 Stable reasoning, open systems and the analytic method

In this final section we look briefly at one of the directions in which stable reasoning has developed into a programme for the logical reconstruction of human knowledge in a practical setting. Then we return to some of the conditions suggested by Cellucci [6] that should be fulfilled by a logical paradigm to replace the axiomatic tradition of mathematical logic. We consider briefly some of the ways in with stable reasoning conforms to these requirements.

6.1 Gelfond's programme

The rise of the axiomatic method and formal reasoning based on classical logic in the 20th Century was closely linked to two philosophical schools based in central Europe: the Lwow-Warsaw School of Logic and Philosophy and the logical empiricist movement of the Vienna Circle. Particularly the latter school endorsed a programme of rational reconstruction of scientific and other forms of knowledge as well as adopting the idea of explicating philosophically important concepts typically by formalising them within a logico-mathematical system. The limits of the logical empiricist programme and the growing criticism it faced from the 1960s onwards are well-known and well-documented. Especially vulnerable was the ideal to reconstruct (ultimately all of) scientific knowledge in formal languages governed by the classical laws of deduction. A pivotal point of this criticism was the claim that logic and experience are not sufficient to explain the rationality of science, catalogue its methods and reconstruct the knowledge it generates.

With hindsight there is an irony to this criticism and to the ultimate downfall of the empiricist programme. It came about just at a moment in time when logic was about to undergo a radical change and a major shift in its boundaries. This change was led by Montague in natural language processing, by Hintikka in the study of propositional attitudes of knowledge and belief, by McCarthy in artificial intelligence and commonsense reasoning, by Simon in learning and scientific discovery, and by a range of applications in computer science and programming.

Stable reasoning has spawned a successor to the logical empiricist programme of rational reconstruction. It is a research programme that aims to reconstruct some of the most basic forms of human knowledge and to exploit this knowledge for practical problem solving. While logical empiricist efforts were largely theoretically oriented, this programme deals with a mix of theory and practice. It combines scientific and engineering knowledge of real systems (....) with practical human skills and abilities and commonsense reasoning. It deals with both static and dynamic domains. The scientist most closely identified with this programme is Michael Gelfond, also a co-founder of stable reasoning itself. His programme combines the physicalist language of engineering and physical systems with epistemic notions such as belief, agency and action. The new programme of rational reconstruction is much less self-conscious than its predecessor. The latter formed part of a manifesto with a clear philosophical, and sometimes political, message. It was stated and re-stated many times. The new programme is scarcely articulated and hardly known. Nevertheless it's goals and methodology are largely clear, if sometimes buried in technical articles and lectures.[10]

Gelfond's programme for representing and reasoning about knowledge has two main objectives. First it aims to achieve an understanding of "basic commonsense notions we use to think about the world: beliefs, knowledge, defaults, causality, intentions, probability, etc., and to learn how one ought to reason about them." Secondly it aims "to understand how to build software components of agents – entities which observe and act upon an environment and direct its activity towards achieving goals." These goals shape the criteria used to evaluate and select languages for Knowledge Representation (KR). In particular he endorses four main adequacy criteria:

1. Clarity: the logical vocabulary should have a clear and intuitive meaning.
2. Elegance: the corresponding mathematics should be simple and elegant.
3. Expressiveness: the KR language should suggest systematic and elaboration tolerant representations of a broad class of phenomena of natural language, including belief, knowledge, defaults, causality and others.
4. Relevance: a large number of interesting computational problems should be reducible to reasoning about theories formulated in this language.

It is interesting to compare these criteria with the requirements that Carnap proposes for the adequate explication of concepts [5]. Clearly criterion 2 is close to Carnap's requirement for the task of concept explication that "the explicatum should be as *simple* as possible." Criterion 4, on the other hand, and to a somewhat lesser extent 3, are close to Carnap's idea that an explicatum is to be a *fruitful* concept, "that is, useful for the formulation of many universal statements." Although Elaboration Tolerance is a more modern idea, one may suppose that it would also feature among the properties of fruitfulness. Carnap's suggestion that the "explicatum is to be *similar to the explicandum*" in many cases of usage does not appear explicitly in Gelfond's list. However since Gelfond's aim is to reconstruct commonsense knowledge and practical reasoning one can assume that Carnap's requirement is one he would also endorse and is somehow implicit in his programme.[11]

It is also revealing to consider some adequacy criteria for KR languages that Gelfond *rejects*. Among these are two that in the past were often considered sacrosanct in the AI community of Knowledge Representation. One is the idea that any KR language should be *supra-classical* ie. extend first-order classical logic. The other is the requirement of *efficiency* understood in a computational sense. We have seen already that the underlying logic of stable reasoning is not classical, and nor does it become so when additional features and functionalities are provided. On the other hand, the expressiveness of the basic language – a positive feature – also results in it being less efficient computationally than some other languages.[12]

[10] But see specially [9] for an overview.
[11] All quotations above are from [5], Introduction.
[12] Generally speaking the complexity of stable reasoning lies at the second level of the polynomial hierarchy. Neverthless answer set solvers are relatively efficient in being able to deal with quite large amounts of data. Notice that from a representational point of view answer set programs are very efficient in being able to encode complex problems in a concise manner.

6.2 Open conceptual systems

Let us return to the idea of open conceptual systems that we mentioned in the introduction. According to Cellucci there are many similarities between open conceptual systems and the notion of open physical systems. Here is a summary of some of the basic features of the former taken from [6] (pp. 313–315).

1. Open conceptual systems take account of the manner in which the solution of a problem is to be arrived at.
2. Unlike in the axiomatic approach there is no unifying idea that serves as a foundation once and for all.
3. The unifying impulse for problem solving is data driven rather than reductionist.
4. Open conceptual systems tackle the problem to be solved directly, from first principles. There is no *a priori* set of concepts and principles that precede the problem formulation.
5. The rules of the game are not given at the beginning and fixed once and for all, but may be introduced and changed during the course of the game.
6. Open systems are dialogical, since partially given information may be extended through interaction with other systems.
7. The rules of the system give only a partial and dynamically changing representation of knowledge.

Besides these characteristics, Cellucci emphasises the ampliative nature of logical inference, as well as the need to deal with with global inconsistencies and incoherences. There are many other features of open systems and the analytic method discussed in [6] and a detailed examination would take us beyond the scope of this paper. Likewise, here we have described only some core features of stable reasoning and ASP. Many other features emerge in the practical development of systems and their application to problem solving. We conclude with a shortlist of some of the characteristics that may bear on Cellucci's challenge to develop an alternative logical paradigm.

– Stable reasoning in its basic form already deals with weak and strong exceptions to defaults. However to deal with contradictions that arise indirectly as consequences of default conclusions, an extension of ASP with consistency-restoring rules (CR-Prolog) was developed and applied by Balduccini and Gelfond [1]. This is essentially an abductive mechanism.
– Another feature of Gelfond's programme has been the aim to reason about the degrees of belief of a rational agent. This has led to a system (P–log) that combines logical and probabilistic reasoning based on ASP [2].
– Although in its basic form stable reasoning is highly declarative, when ASP is used in practice the problem representation takes account of the way in which the solver will successfully and economically reach a solution. Features such as cardinality and integrity constraints and more generally *aggregates* are employed to direct the computational mechanism and possibly enhance efficiency [8].
– Basic ASP already deals with some problems of temporal projection. However, to deal with a wider range of problems for dynamically changing domains, a temporal version of equilibrium logic and ASP has been developed and studied, following [4].
– A central property of open systems is the necessity to interact with other systems. Scholars have developed different logical semantics based on ASP that facilitate this interaction. In particular, logic program rules may contain concepts whose meanings are partially determined by external data sources such as knowledge bases or ontologies [7,16]. This gives rise to hybrid theories that mix different reasoning systems (e.g. monotonic and non-monotonic). Equilibrium logic can be applied to give a simple and uniform treatment of such theories [3].
– To deal with problem solving in a dynamically evolving setting, it is important to consider the problem of updating knowledge in light of new data and knowledge discovery. This has led to the study of theory and program updates in the framework of ASP [17].

7 Acknowledgements

Work on this paper has been partially supported by the projects MERLOT (TIN 2013-42149-P) and FOREST. It is a great pleasure to dedicate this paper to our friend Luis Fariñas del Cerro who has given us generous encouragement and support over a long period. After devoting many years to the advancement of logic and its applications, Luis has recently become also a major contributor to the programme of stable reasoning. We look forward to many more years of intense and fruitful cooperation.

Bibliography

1. M. Balduccini and M. Gelfond. Logic Programs with Consistency-Restoring Rules. In P.Doherty, J. McCarthy, and M-A. Williams, eds, *International Symposium on Logical Formal- ization of Commonsense Reasoning*, AAAI 2003 Spring Symposium Series, pp. 9-18, 2003.
2. C. Baral, M. Gelfond, and N. Rushton. Probabilistic Reason- ing with Answer Sets. In *Proceedings of LPNMR-7*, 2004.
3. J. de Bruijn, D. Pearce. A. Polleres and A. Valverde. A semantical framework for hybrid knowledge bases, *Knowledge and Information Systems* (2010) 25: 81–104
4. P. Cabalar, G. Vega, Temporal equilibrium logic: A first approach, in: *Proc. 11th EUROCAST*, LNCS 4739, Springer, 2007, pp. 241–248.
5. R. Carnap. *The Logical Foundations of Probability*, University of Chicago Press, 1950.
6. C. Cellucci, *Le Ragioni della Logica*, Laterza, Roma-Bari, 1998.
7. T. Eiter, G. Ianni, R. Schindlauer, and H. Tompits. A uniform integration of higher-order reasoning and external evaluations in answer-set programming. In: *IJCAI 2005*, pp 90-96.
8. M Gebser, R Kaminski, B Kaufmann, and T Schaub. *Answer set solving in practice*. Synthesis Lectures on Artificial Intelligence and Machine Learning 6 (3), pp. 1-238, 2012.
9. M. Gelfond, Personal Perspective on the Development of Logic Programming Based KR Languages (2011). Available online at http://www.depts.ttu.edu/cs/research/krlab/papers.php
10. K. Gödel. Zum intuitionistischen aussagenkalkül. *Anzeiger der Akademie der Wissenschaften Wien, mathematisch, naturwissenschaftliche Klasse*, 69:65–66, 1932.
11. A. Heyting. Die formalen regeln der intuitionistischen logik. *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse*, pages 42–56, 1930.
12. T. Hosoi. The axiomatization of the intermediate propositional systems S_n of Gödel. *Journal of the Faculty of Science of the University of Tokio*, 13:183–187, 1966.
13. J. Łukasiewicz. Die Logik und das Grundlagenproblem, in *Les Entreties de Zürich sur les Fondaments et la Méthode des Sciences Mathématiques* 6-9, 12 (1938), Zürich, 1941, 82-100.
14. L. Maksimova. Craig's interpolation theorem and amalgamable varieties, *Doklady Akademii Nauk SSSR*, 237, no. 6, (1977), 1281-1284 (translated as *Soviet Math. Doklady*).
15. D. Pearce. A new logical characterisation of stable models and answer sets. In *Non-Monotonic Extensions of Logic Programming, NMELP 96*, LNCS 1216, pages 57–70. Springer, 1997.
16. R. Rosati. DL+log: tight integration of description logics and disjunctive datalog. In: *Proceedings KR 2006*, pp. 68-78.
17. M. Slota and J. Leite. On semantic update operators for answer-set programs. In Coelho, H.; Studer, R.; and Wooldridge, M., eds., *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI 2010)*, volume 215 of Frontiers in Artificial Intelligence and Applications, pp. 957-962. IOS Press, 2010.
18. Ya. S. Smetanich. On Completeness of a Propositional Calculus with an additional Operation of One Variable (in Russian). *Trudy Moscovskogo matrmaticheskogo obshchestva*, 9 (1960), 357-372.
19. T. Umezawa. On intermediate many-valued logics. *Journal of the Mathematical Society of Japan* 11(2) (1959).
20. John McCarthy, 'Elaboration Tolerance', in *Proc. of the 4th Symposium on Logical Formalizations of Commonsense Reasoning (Common Sense 98)*, pp. 198–217. London, UK (1998).

Index