Social logics from the point of view of knowledge representation and reasoning

Andreas Herzig
CNRS-IRIT
Toulouse

Sada, Sep. 1, 2022

Knowledge representation and reasoning in AI

- the KR hypothesis [Brian C. Smith, 1985]:
  Any mechanically embodied intelligent process will be comprised of structural ingredients that a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits, and b) independent of such external semantic attribution, play a formal but causal and essential role in engendering the behavior that manifests that knowledge.
- since ∼1985: dynamic turn
- since ∼2000: social turn

How to use logic in AI

- ‘logic engineering’ [Gabbay, ∼1990]:
    1. identify core concepts
    2. logical language
    3. semantics
    4. reasoning methods
    5. reasoners
    6. benchmarks, competitions
- . . . in that order
    - computation doesn’t matter in stages 1-3
    - but does so in stages 4-6
- application to social phenomena?
    - other successful applications of logic in CS and AI?

Logic in CS and AI: success stories

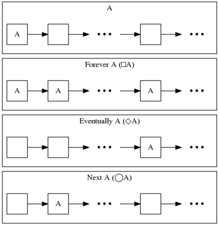
DELs: a good basis for social logics?

Which core concepts?

Lightweight logics of belief and action

Success story: temporal logic

- goal: reason about program behaviour
  $\Longrightarrow$ prove properties (liveness, safety,...)
  $\Longrightarrow$ motivated by applications: verification
- core concepts: 'next', 'eventually', 'always', 'until'
- semantics: Linear-time Temporal Logic LTL
  - consensical [Vardi "Branching vs. linear time: Final showdown", TACAS 2001]
  - rest is exotic (CTL, ATL, strategy logics,...)
- reasoning: model checking
  - consensical (though ExpTime)
  - rest is exotic (i.e., validity checking)



Success story: description logics DLs

- goal: go beyond relational databases (CWA, UNA)
  $\Longrightarrow$ find useful decidable fragments of FOL
  - expressiveness better than propositional logic
  - computation better than FOL
  - motivated by applications: databases, ontologies for the semantic web
- core concepts: unary and binary predicates
- restrict quantifiers to stay decidable

$$\forall R.C = \forall y(R(x,y) \rightarrow C(y))$$
$$\exists R.C = \exists y(R(x,y) \wedge C(y))$$

  - ALC = "Attributive Concept Language with Complements" [Schmidt-Schauß&Smolka, 1991]
  - is nothing but multimodal K [Schild, IJCAI 1995]

Success story: description logic engineering

1. map out logics of the core concepts
   - counting quantifiers, transitive closure of a relation,. . .
   - categorisation, (partial) lattice (cf. modal cube)
   - reasoning problems typically between PSpace and ExpTime
2. beyond core theory: time, real numbers, default reasoning, typicality,. . .
3. implemented systems, benchmarks, competitions
4. for some applications reasoning tasks such as query answering are too complex
   - 'too complex' $\approx$ beyond PSPACE
   - impose more restrictions on the syntax
     - no disjunctions, no negations
     - still sufficient for applications (SNOMED CT)
   - lightweight fragments of DLs

Success story: description logic engineering

Description Logic Complexity Navigator (by Evgeny Zolin) http://www.cs.man.ac.uk/~ezolin/dl/

Complexity of reasoning in Description Logics

Note: the information here is (always) incomplete and updated often

Base description logic: $\mathcal{A}$ttributive $\mathcal{L}$anguage with $\mathcal{C}$omplements

$\mathcal{ALC} ::= \bot \mid \top \mid A \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists R.C \mid \forall R.C$

Concept constructors:

- ☐ $\mathcal{F}$ – functionality[2]: $(\leq 1\ R)$
- ☐ $\mathcal{N}$ – (unqualified) number restrictions: $(\geq n\ R)$, $(\leq n\ R)$
- ☐ $\mathcal{Q}$ – qualified number restrictions: $(\geq n\ R.C)$, $(\leq n\ R.C)$
- ☐ $\mathcal{O}$ – nominals: $\{a\}$ or $\{a_1, \ldots, a_n\}$ ("one-of")
- ☐ μ – least fixpoint operator: $\mu X.C$

Forbid complex roles[5] in number restrictions[6]

Role constructors: trans reg

- ☐ $\mathcal{I}$ – role inverse: R^-
- ☐ $\cap$ – role intersection[3]: $R \cap S$
- ☐ $\cup$ – role union: $R \cup S$
- ☐ $\neg$ – role complement: $\neg R$ full
- ☐ $\circ$ – role chain (composition): $R \circ S$
- ☐ $*$ – reflexive-transitive closure[4]: R^*
- ☐ id – concept identity: $id(C)$

TBox (concept axioms):

- ◉ empty TBox
- ○ acyclic TBox ($A \equiv C$, A is a concept name; no cycles)
- ○ general TBox ($C \sqsubseteq D$, for arbitrary concepts C and D)

RBox (role axioms): OWL-Lite OWL-DL OWL 1.1

- ☐ $\mathcal{S}$ – role transitivity: $Tr(R)$
- ☐ $\mathcal{H}$ – role hierarchy: $R \sqsubseteq S$
- ☐ $\mathcal{R}$ – complex role inclusions: $R \circ S \sqsubseteq R$, $R \circ S \sqsubseteq S$
- ☐ s – some additional features (click to see them)

Reset You have selected a Description Logic: $\mathcal{ALC}$

Complexity[7] of reasoning problems[8]		
Concept satisfiability	**PSpace-complete**	• Hardness for $\mathcal{ALC}$: see [80]. • Upper bound for $\mathcal{ALCQ}$: see [12, Theorem 4.6].
ABox consistency	**PSpace-complete**	• Hardness follows from that for concept satisfiability. • Upper bound for $\mathcal{ALCQO}$: see [17, Appendix A].
Important properties of the Description Logic		
Finite model property	**Yes**	$\mathcal{ALC}$ is a notational variant of the multi-modal logic $\mathbf{K}_m$ (cf. [77]), for which the finite model property can be found in [4, Sect. 2.3].
Tree model property	**Yes**	$\mathcal{ALC}$ is a notational variant of the multi-modal logic $\mathbf{K}_m$ (cf. [77]), for which the tree model property can be found in [4, Proposition 2.15].

Maintained by: Evgeny Zolin
Please see the list of updates

Any comments are welcome:
EZolin@cs.man.ac.uk

Success stories

- lessons to learn:
  1. consensus on a small number of core concepts
  2. consensus on core semantics
  3. complexity too high? $\Longrightarrow$ find ‘good’ fragments
- “logic as a Swiss knife”
  - vs. “logic as a toothbrush” view in AI and philosophical logic
- could DELs serve as a Swiss knife?

Logic in CS and AI: success stories

DELs: a good basis for social logics?

Which core concepts?

Lightweight logics of belief and action

DELs: a good basis for a logic for social phenomena?

- overcome a restrictive hypothesis in previous logical approaches to action and knowledge:

  > "we suppose that all agents observe all action occurrences"
  > (SitCalc@Toronto, action languages@Texas)

- basic idea: event model = Kripke model
  - possible world = possible event
    - announcement of a formula
    - assignments of several prop.var.s
  - indistinguishability relations model agents' observation of events
  - product update: static model $\otimes$ event model

Dynamic Epistemic Logics: language

1. epistemic operators: “agent believes/knows proposition”

$$\mathbf{K}_{\mathit{Pedro}}\, \mathit{Sunny}$$
$$\mathbf{B}_{\mathit{Pedro}}\, \neg \mathbf{K}_{\mathit{Andreas}}\, \mathit{Sunny}$$

2. dynamic operators: “proposition is true after event”

$$\langle \mathit{Event} \rangle\, \mathit{Sunny}$$

where *Event* is a *Kripke model*

- world $\hat{=}$ announcement and assignments
- accessibility relations: model agents' observation of the event

Dynamic Epistemic Logics: semantics

- relation of indistinguishability between possible worlds
  - equivalence relation (but criticisable, cf. [Lenzen; Voorbraak])
- truth conditions:

$M, w \models \mathbf{K}_{Pedro}\, Sunny$ iff for every w' Pedro cannot distinguish from w in E_{Sunny} : $M, w' \models Sunny$

$M, w \models \langle E_{Sunny} \rangle \varphi$ iff $M, w \models \varphi$ and $M \otimes E_{Sunny}, w \models \varphi$

where $M \otimes E_{Sunny}$ is the update of M by E_{Sunny}:
eliminate from M all worlds where *Sunny* is false

Dynamic Epistemic Logics: computation

- difficult
  - SAT for individual knowledge: PSpace
  - SAT for group knowledge: ExpTime
  - planning: undecidable [Bolander; Aucher; Schwarzentruber;. . .]

Dynamic Epistemic Logics?

1. event models amalgamate syntax and semantics
   - Theorem [French, Hales&Tay, AiML 2014]: all event models can be constructed from
     1.1 private announcements to groups

$$\mathit{Sunny}!_{\{\text{Andreas},\text{Pedro},\ldots\}}$$

     1.2 the PDL program operators
2. DELs almost always fail to be a conservative extension of the underlying epistemic logic [Balbiani et al., AiML 2012]
   - existential properties not preserved under world elimination
   - "we choose modal logic K for the sake of generality"
   - "we choose the standard logic of knowledge S5"

Dynamic Epistemic Logics?

1. difficult to replace S5 by more realistic logics of knowledge such as S4.2
   - conservativity fails (v.s.)
2. even more difficulties with belief
   - conservativity fails (v.s.)
   - requires extension by (multiagent) belief revision:

     $$\models \mathbf{B}_{Andreas}\,\neg Sunny \rightarrow \langle Sunny!_{\{\text{Andreas},\text{Pedro},\ldots\}}\rangle \mathbf{B}_{Andreas}\,\bot$$

   - only few approaches: [van Ditmarsch 2006, Aucher 2007, Baltag&Smets 2012]

Dynamic Epistemic Logics?

1. DEL and the classical problems in reasoning about actions [McCarthy; Reiter;. . .]
   - frame problem solved by assignments
     - corresponds to Reiter's solution to the frame problem [van Ditmarsch et al., JLC 2011]

$$\{p_1 \leftarrow \varphi_{p_1}, \cdots, p_n \leftarrow \varphi_{p_n}\}$$

   - no account of qualification problem
   - no account of ramification problem
2. not obvious how to represent knowledge about action *types*
   - models typically have to be infinite (in order to account for all possible observational situations)
   - DELs rather suited for action tokens
3. does not provide an account of agency
   - action has author
   - action theory requires intentions

Conclusion on DELs: not a good basis for a logic of social concepts

- beautiful theory
- important conceptual problems:
  - action types??
  - belief revision??
  - goals and intentions??
- computation difficult

Logic in CS and AI: success stories

DELs: a good basis for social logics?

Which core concepts?

Lightweight logics of belief and action

Logical modelling of social phenomena: which concepts?

- quite some:
  - knowledge and belief
  - preferences, utilities, goals, intentions; norms, obligations
  - action, ability, agency, strategy, speech acts, arguments
- state of domain:
  - too many concepts
  - no consensus which should be core

Logical modelling of social phenomena: which semantics?

- manyfold modelling choices
  - knowledge or belief? degrees of belief? omniscience? groups?
  - actions: joint? profiles? temporally extended?
  - theory of intention?
    - complicated ([Shoham&Leyton Brown, 2008], Section 14.4.2: "The road to hell: elements of a formal theory of intention")
    - gap between theory and implemented BDI agents
  - deontic logic?
    - still struggles with old problems: paradoxes, contrary-to-duty obligations
  - nonmonotonic extensions
    - default reasoning, typicality, arguments,...
    - no good account (exception ASP)
- summary:
  - no consensus
  - too much focus on knowledge (concept needed for intelligent machines: beliefs)

Logical modelling of social phenomena: which reasoning?

- model checking/theorem proving/deduction
- planning & strategic reasoning
  - computation of game-theoretic equililbria
- nonmonotonic deduction (computation of argument framework extensions,. . .)
- summary:
  - many reasoning modes, many prover options
  - no complexity navigator for epistemic/action/deontic/. . . /logic

Logical modelling of social phenomena: a shortlist of core concepts

hypotheses:

H1: socially intelligent machines need theory of mind (ToM)

- ToM = representation of others' beliefs and goals
- Sally-Ann Test [Baron Cohen] $\Longrightarrow$ in DEL [Bolander]

H2: intelligent interaction requires group belief

- logics of common belief; logics of common ground

H3: account of actions needed

- necessary ingredients: belief logic + action logic + revision
- nice to have: group belief, joint action, norms, strategies,...

Logical modelling of social phenomena: a plea for lightweight logics

- logic of belief&action&revision: already a complicated beast
  1. conceptually
  2. computationally
- get better computational properties: restrict static doxastic language further
  $\Longrightarrow$ *lightweight* doxastic logic

Logic in CS and AI: success stories

DELs: a good basis for social logics?

Which core concepts?

Lightweight logics of belief and action

Lightweight logics of knowledge: ‘knowing-that’ literals

[Demolombe&Pozos Parra; Lakemeyer&Lespérance 2012; Muise et al. 2015; 2021]

$$\lambda ::= p \mid \neg\lambda \mid \mathbf{K}_i\lambda$$

- formula = boolan combination of epistemic literals
  - no conjunction or disjunction in scope of epistemic operators
- complexity: same as propositional logic
  - view epistemic atoms as propositional variables
  - plus theory: $\neg(\mathbf{K}_i\lambda \land \mathbf{K}_i\neg\lambda)$, $\mathbf{K}_i\mathbf{K}_i\lambda \leftrightarrow \mathbf{K}_i\lambda$, etc.
- cannot express “I know you know more than me”

$$\neg\mathbf{K}_i p \land \neg\mathbf{K}_i\neg p \land \mathbf{K}_i(\mathbf{K}_j p \lor \mathbf{K}_j\neg p)$$

  but is fundamental in interaction (precondition of questions)
- sequel: ‘knowing-whether’ primitive instead [Lomuscio; van der Hoek et al.; Gattinger et al.]

Lightweight logics of knowledge: background on ‘knowing-whether’

- standard modalities of epistemic logic since [Hintikka, 1962]:

$$\mathbf{K}_i\varphi \quad = \quad \text{“agent } i \text{ knows that } \varphi\text{”}$$
$$\mathbf{B}_i\,\varphi \quad = \quad \text{“agent } i \text{ believes that } \varphi\text{”}$$

- motivation: ‘know whether’ more primitive than ‘know that’
  - knowing the truth value of a proposition more basic than knowing that the truth value equals 1

    “To know is to know the value of a variable” [Baltag, 2016]

- related to:
  - non-contingency logics [Montgomery and Routley, 1966, Humberstone et al., 1995]
  - logic of ignorance [Kubyshkina and Petrolo, 2019]
  - Yanjing Wang’s “beyond knowing-that” research program
- benefit: new lightweight fragments

Knowledge/belief about a proposition

- ‘know whether’ has no belief-counterpart in natural language (just as the other ‘know wh’ modalities) [Egré, 2008]
- therefore:

$$\mathbf{KA}_i\varphi \quad = \quad \text{“agent } i \text{ has knowledge about } \varphi\text{”}$$
$$\mathbf{BA}_i\varphi \quad = \quad \text{“agent } i \text{ has belief about } \varphi\text{”}$$

alternatively: “i is opinionated about φ”

‘About’ modalities: expressivity

1. ‘belief about’: weaker [Fan et al., 2015]

$$\mathbf{BA}_i\varphi \leftrightarrow \mathbf{B}_i\,\varphi \vee \mathbf{B}_i\,\neg\varphi$$
$$\mathbf{B}_i\,\varphi \leftrightarrow ?$$

2. ‘knowledge about’: equi-expressive

$$\mathbf{KA}_i\varphi \leftrightarrow \mathbf{K}_i\varphi \vee \mathbf{K}_i\neg\varphi$$
$$\mathbf{K}_i\varphi \leftrightarrow \varphi \wedge \mathbf{KA}_i\varphi$$

but:

- ‘knowledge about’ can express things more succinctly [van Ditmarsch et al., 2014]
- equivalent presentations may lead to new insights
  - cf. Kosta Došen: “Had Gentzen used Tarski’s consequence operator $Cn(\Gamma)$, he wouldn’t have found the cut rule”

‘Knowledge about’ atoms [Herzig et al., 2015, Cooper et al., 2021]

- grammar:

$$\alpha ::= p \mid \mathbf{KA}_i \alpha$$

  where $p \in Prop$
- formula = boolan combination of epistemic atoms
- can express some disjunctions in scope of epistemic operator:

$$\mathbf{K}_i(\mathbf{K}_j p \vee \mathbf{K}_j \neg p)$$

  expressed as

$$\begin{aligned} & \mathbf{K}_i \, \mathbf{KA}_j p \\ = \; & \mathbf{KA}_j p \wedge \mathbf{KA}_i \mathbf{KA}_j p \end{aligned}$$

'Knowledge about' atoms: computation

- basically: epistemic atoms can be viewed as propositional logic variables
  - take care of introspection: $\mathbf{KA}_i\mathbf{KA}_i\alpha$ valid
  - simple solution: forbid repetitions
- complexity of reasoning: same as propositional logic
  - satisfiability NP-complete
- can be extended by an operator of common knowledge [Herzig&Perrotin, AiML 2020; forthcoming]
  - replace greatest fixed-point axiom for common knowledge

$$(p \wedge \mathbf{CK}\,(p \to \bigwedge_i \mathbf{K}_i p)) \to \mathbf{CK}\,p \quad (1)$$

$$\bigwedge_i \mathbf{CK}\,\mathbf{KA}_i p \to \mathbf{CKA}\,p \quad (2)$$

  - not valid for belief

Lightweight logics of knowledge: dynamics

- 'dual use' of knowledge about atoms [Cooper et al., AIJ 2020]:
  - $\mathbf{KA}_i\alpha$ = agent i sees truth value of α
  - $\mathbf{KA}_i\alpha$ = agent i sees truth value changes of α (except if action makes $\mathbf{KA}_i\alpha$ false)
- STRIPS-like actions: preconditions + pos./neg. effects
- complexity of planning: same as propositional logic
  - plan existence PSPACE-complete

Lightweight logics of belief?

- knowledge-about atoms 'work' because there are 4 independent combinations of p and $\mathbf{KA}_i p$:

$p \wedge \mathbf{KA}_i p$	$\neg p \wedge \mathbf{KA}_i p$
$p \wedge \neg\mathbf{KA}_i p$	$\neg p \wedge \neg\mathbf{KA}_i p$

  - in terms of knowledge-that:

$p \wedge \mathbf{K}_i p$	$\neg p \wedge \mathbf{K}_i \neg p$
$p \wedge \neg\mathbf{K}_i p \wedge \neg\mathbf{K}_i \neg p$	$\neg p \wedge \neg\mathbf{K}_i p \wedge \neg\mathbf{K}_i \neg p$

- for belief: 6 possible doxastic situations

$p \wedge \mathbf{B}_i\, p$	$\neg p \wedge \mathbf{B}_i\, \neg p$
$p \wedge \neg\mathbf{B}_i\, p \wedge \neg\mathbf{B}_i\, \neg p$	$\neg p \wedge \neg\mathbf{B}_i\, p \wedge \neg\mathbf{B}_i\, \neg p$
$p \wedge \mathbf{B}_i\, \neg p$	$\neg p \wedge \mathbf{B}_i\, p$

  - requires 3 dimensions $\Longrightarrow$ cannot be independent

Which epistemic-doxastic situations?

- 8 possible situations:

$p \wedge \mathbf{K}_i p$	$\neg p \wedge \mathbf{K}_i \neg p$
$p \wedge \mathbf{B}_i\, p \wedge \neg \mathbf{K}_i p$	$\neg p \wedge \mathbf{B}_i\, \neg p \wedge \neg \mathbf{K}_i \neg p$
$p \wedge \neg \mathbf{B}_i\, p \wedge \neg \mathbf{B}_i\, \neg p$	$\neg p \wedge \neg \mathbf{B}_i\, p \wedge \neg \mathbf{B}_i\, \neg p$
$p \wedge \mathbf{B}_i\, \neg p$	$\neg p \wedge \mathbf{B}_i\, p$

- cf. Kanger-Lindahl theory of normative positions [Sergot and Richards, 2001, Sergot, 2001]; act positions [Demolombe and Jones, 2002]
- $8 = 2^3 \implies$ which are the 3 dimensions?

Which epistemic-doxastic situations?

- two new modalities:

$$\begin{aligned}
\mathbf{TBA}_i\, p &= (p \wedge \mathbf{B}_i\, p) \vee (\neg p \wedge \mathbf{B}_i \neg p) \\
&= \text{“}i \text{ has a } \textbf{true} \text{ belief about } p\text{”} \\
\mathbf{MBA}_i\, p &= (\mathbf{B}_i\, p \wedge \neg \mathbf{K}_i p) \vee (\mathbf{B}_i \neg p \wedge \neg \mathbf{K}_i \neg p) \\
&= \text{“}i \text{ has a } \textbf{mere} \text{ belief about } p\text{”} \\
&= \text{“}i \text{ has a falsifiable belief about } p\text{”} \\
&= \text{“}i \text{ has a belief about } p \text{ but does not know whether } p\text{”}
\end{aligned}$$

- insensitive to negation:

$$\mathbf{TBA}_i \neg p \leftrightarrow \mathbf{TBA}_i\, p$$

$$\mathbf{MBA}_i \neg p \leftrightarrow \mathbf{MBA}_i\, p$$

Epistemic-doxastic situations: 3 dimensions

- 2^3 epistemic-doxastic situations:

$p \wedge \mathbf{TBA}_i\, p \wedge \neg\mathbf{MBA}_i\, p$	$\neg p \wedge \mathbf{TBA}_i\, p \wedge \neg\mathbf{MBA}_i\, p$
$p \wedge \mathbf{TBA}_i\, p \wedge \mathbf{MBA}_i\, p$	$\neg p \wedge \mathbf{TBA}_i\, p \wedge \mathbf{MBA}_i\, p$
$p \wedge \neg\mathbf{TBA}_i\, p \wedge \neg\mathbf{MBA}_i\, p$	$\neg p \wedge \neg\mathbf{TBA}_i\, p \wedge \neg\mathbf{MBA}_i\, p$
$p \wedge \neg\mathbf{TBA}_i\, p \wedge \mathbf{MBA}_i\, p$	$\neg p \wedge \neg\mathbf{TBA}_i\, p \wedge \mathbf{MBA}_i\, p$

- needs getting used to, but is natural. . .

Example: the Sally-Ann Test

false belief task
[Wimmer and Perner, 1983, Baron-Cohen et al., 1985]

1. Sally puts the marble in the basket

$$\mathbf{TBA}_S\, \mathrm{b} \land \neg \mathbf{MBA}_S\, \mathrm{b}$$

2. Sally goes out for a walk

$$\mathbf{TBA}_S\, \mathrm{b} \land \mathbf{MBA}_S\, \mathrm{b}$$

3. Ann takes the marble out of the basket and puts it into the box

$$\neg \mathbf{TBA}_S\, \mathrm{b} \land \mathbf{MBA}_S\, \mathrm{b}$$

Full expressivity

- knowledge:

$$\mathbf{KA}_i\varphi \leftrightarrow \mathbf{TBA}_i\,\varphi \wedge \neg\mathbf{MBA}_i\,\varphi$$

$$\mathbf{K}_i\varphi \leftrightarrow \mathbf{TBA}_i\,\varphi \wedge \neg\mathbf{MBA}_i\,\varphi \wedge \varphi$$

- belief:

$$\mathbf{BA}_i\varphi \leftrightarrow \mathbf{TBA}_i\,\varphi \vee \mathbf{MBA}_i\,\varphi$$

$$\mathbf{B}_i\,\varphi \leftrightarrow (\varphi \wedge \mathbf{TBA}_i\,\varphi) \vee (\neg\varphi \wedge \neg\mathbf{TBA}_i\,\varphi \wedge \mathbf{MBA}_i\,\varphi)$$

...remember: $\mathbf{B}_i\,\varphi$ cannot be expressed with $\mathbf{BA}_i$ alone

An epistemic-doxastic logic

- logic:

$\mathtt{KD5}(\mathbf{B})$	the principles of modal logic KD5 for $\mathbf{B}_i$
$\mathtt{S4}(\mathbf{K})$	the principles of modal logic S4 for $\mathbf{K}_i$
KiB	$\mathbf{K}_i\varphi \to \mathbf{B}_i\,\varphi$
BiKB	$\mathbf{B}_i\,\varphi \to \mathbf{K}_i\mathbf{B}_i\,\varphi$
BiBK	$\mathbf{B}_i\,\varphi \to \mathbf{B}_i\,\mathbf{K}_i\varphi$

- belief definable from knowledge [Lenzen, 1978, Lenzen, 1995]:

$$\mathbf{B}_i\,\varphi \leftrightarrow \neg\mathbf{K}_i\neg\mathbf{K}_i\varphi$$

  - alternative axiomatisation: $\mathtt{S4.2}(\mathbf{K})$ plus $\mathbf{B}_i\,\varphi \leftrightarrow \neg\mathbf{K}_i\neg\mathbf{K}_i\varphi$
- complexity of satisfiability: PSPACE-complete [Shapirovsky, 2004, Chalki et al., 2021]

Reduction of 'about' modalities

- reduction of consecutive modal operators to length 1:

$$\mathbf{TBA}_i\,\mathbf{TBA}_i\,\varphi \leftrightarrow \mathbf{TBA}_i\,\varphi \vee \neg\mathbf{MBA}_i\,\varphi$$
$$\mathbf{MBA}_i\,\mathbf{TBA}_i\,\varphi \leftrightarrow \mathbf{MBA}_i\,\varphi$$
$$\mathbf{TBA}_i\,\mathbf{MBA}_i\,\varphi \leftrightarrow \neg\mathbf{MBA}_i\,\varphi$$
$$\mathbf{MBA}_i\,\mathbf{MBA}_i\,\varphi \leftrightarrow \mathbf{MBA}_i\,\varphi$$

$\Longrightarrow$ suppose formulas are 'repetition-free'

- no $\cdots\mathbf{TBA}_i\,\mathbf{TBA}_i\,\cdots p$
- no $\cdots\mathbf{TBA}_i\,\mathbf{MBA}_i\,\cdots p$
- no $\cdots\mathbf{MBA}_i\,\mathbf{TBA}_i\,\cdots p$
- no $\cdots\mathbf{MBA}_i\,\mathbf{MBA}_i\,\cdots p$

Lightweight epistemic-doxastic fragments: the idea

- epistemic-doxastic (‘epidox’) atoms:

$$\alpha ::= p \mid \mathbf{TBA}_i\, \alpha \mid \mathbf{MBA}_i\, \alpha$$

- repetition-free

Theorem

If φ is a boolean combination of (repetition-free) epidox atoms then the following are equivalent:

- *φ is valid in epistemic-doxastic logic;*
- *φ is propositionally valid.*

Corollary

Satisfiability of boolean combinations of epidox atom is in NP.
Plan existence is in PSpace.

A lightweight logic of action

- action = precondition + (conditional) effects
  - precondition = boolean combination of epidox atoms
  - effects = add/delete epidox atoms
  - cf. STRIPS
- simple epistemic-doxastic planning problems
  - initial state = boolean combination of epidox atoms
  - goal = boolean combination of epidox atoms
  -
  - solvability of a planning task in epistemic-doxastic logic
    - reduces to solvability in propositional logic
- examples:
  - Sally-Ann test as a planning task (goal = induce Sally's false belief)
  - variants of the grapevine domain
  - ...

Conclusion: a lightweight logic of belief and action

- lightweight fragment of epistemic-doxastic logic
  - ‘true belief about’ and ‘mere belief about’ modalities
  - repetition-free epistemic-doxastic atoms
  - same complexity as propositional logic
- rest of talk (if there is time): lightweight logic of agency

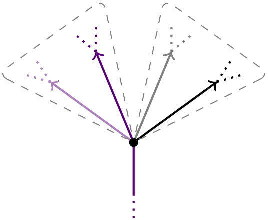
Logics of seeing-to-it-that: concept

- important:
  - agency = relation between individual and proposition
    - $[i\,\mathrm{stit}]\varphi$ = "i sees to it that φ" [Horty&Belnap, 1995]
  - fundamental for the analysis of causality, responsibility, influence, social emotions,...
- can be combined with:
  - groups ('coalitions')
  - time
  - obligation, knowledge,...

Logics of seeing-to-it-that: semantics and reasoning

BT+AC models: Branching Time + Agent Choice

- branching timelines
- agent choices = partition of possible timelines



complex:

- theorem proving is difficult
  - logic of just $[i\,\mathrm{stit}]\varphi$ already NExpTime-complete [Balbiani, Herzig&Troquard, JPL 2008]
  - 2ExpTime-complete with the temporal 'next' [Boudou&Lorini, AAMAS 2018]
  - undecidable and non-axiomatisatible if there are coalitions [Herzig&Schwarzentruber, AiML 2008]
- some fragments do better, but are not very interesting [Schwarzentruber, Studia Logica 2012]
- model checking is unfeasible
  - typically infinite models

Agency based on control and attempt [Herzig et al., IJCAI 2022]

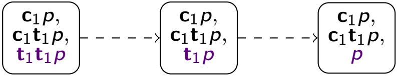
- concepts:
  - $\mathbf{c}_i p$ = control of propositional variable p by agent i
  - $\mathbf{t}_i p$ = attempt to change p by agent i
    - successful if the agent controls p
- features:
  - higher-order (control of attempts, attempt to control,...)
  - temporal operators of LTL

$$\mathbf{X}p \stackrel{\text{def}}{=} \left(\bigvee_i (\mathbf{c}_i p \wedge \mathbf{t}_i p)\right) \vee \left(p \wedge \neg \bigvee_i (\mathbf{c}_i p \wedge \mathbf{t}_i p)\right)$$

  - coalitions of agents

Control and attempt: semantics

- model = truth values of prop.var.s, control and attempt atoms
- future states determined by control & attempts



- agent choices = variations on attempts
- results:
  - LACA is a fragment of standard stit logic
  - model checking is PSpace-complete

Conclusion: a lightweight stit logic

- simple stit logic: LACA
  - agency restricted to propositional variables
  - better complexity results (PSpace-complete model checking)
- application to stit-based representation of influence [Herzig et al., IJCAI 2022]

Logic in CS and AI: success stories

DELs: a good basis for social logics?

Which core concepts?

Lightweight logics of belief and action

Conclusion of talk

- two lightweight logics:
  - belief based on true and mere belief about a proposition
  - action based on control and attempt
- combine independently:
  - $p \wedge \mathbf{TBA}_i\, p \wedge \mathbf{MBA}_i\, p \wedge \mathbf{c}_i p \wedge \mathbf{t}_i p$
  - $p \wedge \mathbf{TBA}_i\, p \wedge \mathbf{MBA}_i\, p \wedge \mathbf{c}_i p \wedge \neg \mathbf{t}_i p$
  - ...

  $\Longrightarrow$ SAT still in NP

  $\Longrightarrow$ plan existence still in PSpace
- missing: group belief
  - formalisation of common knowledge does not transfer
- missing: goals, intentions

Baltag, A. (2016).
To know is to know the value of a variable.
In Beklemishev, L. D., Demri, S., and Maté, A., editors, *Advances in Modal Logic 11, proceedings of the 11th conference on "Advances in Modal Logic," held in Budapest, Hungary, August 30 - September 2, 2016*, pages 135–155. College Publications.

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985).
Does the autistic child have a theory of mind?
Cognition, 21(1):37–46.

Chalki, A., Koutras, C. D., and Zikos, Y. (2021).
A note on the complexity of S4.2.
J. Appl. Non Class. Logics, 31(2):108–129.

Cooper, M., Herzig, A., Maffre, F., Maris, F., Perrotin, E., and Régnier, P. (2021).
A lightweight epistemic logic and its application to planning.
Artificial Intelligence, 298:103437.

Demolombe, R. and Jones, A. J. (2002).
Actions and normative positions.
In Jacquette, D., editor, *A companion to philosophical logic*, pages 355–372. Blackwell Publishing.

Egré, P. (2008).
Question-embedding and factivity.
Grazer Philosophische Studien, 77(1):85–125.

Fan, J., Wang, Y., and van Ditmarsch, H. (2015).
Contingency and knowing whether.
Rew. Symb. Logic, 8(1):75–107.

Herzig, A., Lorini, E., and Maffre, F. (2015).
A poor man's epistemic logic based on propositional assignment and higher-order observation.
In van der Hoek, W., Holliday, W. H., and Wang, W.-f., editors, *Proceedings of the 5th International Conference on Logic, Rationality and Interaction (LORI 2015)*, pages 156–168. Springer Verlag.

Hintikka, J. (1962).
Knowledge and Belief: An Introduction to the Logic of the Two Notions.
Cornell University Press.

Humberstone, I. et al. (1995).
The logic of non-contingency.
Notre Dame Journal of Formal Logic, 36(2):214–229.

Kubyshkina, E. and Petrolo, M. (2019).
A logic for factive ignorance.
Synthese, pages 1–12.

Lenzen, W. (1978).
Recent work in epistemic logic.
North Holland Publishing Company, Amsterdam.

Lenzen, W. (1995).
On the semantics and pragmatics of epistemic attitudes.
In Laux, A. and Wansing, H., editors, *Knowledge and belief in philosophy and AI*, pages 181–197. Akademie Verlag, Berlin.

Montgomery, H. and Routley, R. (1966).
Contingency and non-contingency bases for normal modal logics.
Logique et analyse, 9(35/36):318–328.

Sergot, M. J. (2001).
A computational theory of normative positions.
ACM Trans. Comput. Log., 2(4):581–622.

Sergot, M. J. and Richards, F. (2001).
On the representation of action and agency in the theory of normative positions.

Fundam. Informaticae, 48(2-3):273–293.

Shapirovsky, I. (2004).
On PSPACE-decidability in transitive modal logic.
In Schmidt, R. A., Pratt-Hartmann, I., Reynolds, M., and Wansing, H., editors, *Advances in Modal Logic 5, papers from the fifth conference on "Advances in Modal logic", held in Manchester, UK, 9-11 September 2004*, pages 269–287. King's College Publications.

van Ditmarsch, H., Fan, J., van der Hoek, W., and Iliev, P. (2014).
Some exponential lower bounds on formula-size in modal logic.
In Goré, R., Kooi, B. P., and Kurucz, A., editors, *Advances in Modal Logic 10, invited and contributed papers from the tenth conference on "Advances in Modal Logic," held in Groningen, The Netherlands, August 5-8, 2014*, pages 139–157. College Publications.

Wimmer, H. and Perner, J. (1983).
Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception.
Cognition, 13(1):103–128.