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.

Toulouse, 3rd−4th March 2016

Journées en l'honneur de Luis Fariñas del Cerro

Background

1. Arguments are exchanged by human agents in natural language (spoken or written) in discussion, debate, negotiation, persuasion, etc.
2. If we want artificial agents to represent and reason with arguments coming from human agents, then we need formalisms that handle them.
3. If we want to better theories of argumentation, we should compare them against corpora of natural language arguments.
4. The NLP community is interested in identifying arguments and relations between them in natural language.
   - Following from successes in information extraction, sentiment analysis, etc, argument mining is seen is one of the next big challenges for NLP.
   - E.g. 1st ACL Workshop on Argument Mining.
   - E.g. IBM Debating Technologies

Overview

Issues

1. We need an appropriate target language for representing arguments mined from natural language.
2. Given a base of these arguments mined from texts or dialogues (whether obtained by hand or by NLP technology), we want be able combine them, deconstruct them, and to analyse them (for instance to check whether the set is inconsistent).

Proposal

- A formal language for representing some of the structure of arguments.
- A framework for inferencing with the arguments in this formal language.
- This framework is flexible so different sets of inference rules can be used.

Why do we need a new formalism?

As a target language for mined arguments

Abstract argumentation: Each argument is atomic. So insufficient structure for a target language for argument mining.

Logical argumentation: Each argument is a set of formulae for premises, and a formula for a claim. So excessive structure for a target language for argument mining.

As a formalism for reasoning with mined arguments

Neither abstract argumentation nor (in pure form) logical argumentation provides machinery for reasoning with arguments.

Why do we need a new formalism?

Red denotes outer reason-claim coupling, and blue denotes inner reason-claim coupling. Note, outer reason-claim coupling has two reasons for the claim.

⟨claim⟩Heathrow needs more capacity⟨\claim⟩

⟨reason⟩ Heathrow runs at close to 100% capacity. With demand for air travel predicted to double in a generation, Heathrow will not be able to cope without a third runway, say those in favour of the plan. ⟨\reason⟩

⟨reason⟩ ⟨reason⟩ Because the airport is over-stretched, any problems which arise cause knock-on delays. ⟨\reason⟩ ⟨claim⟩ Heathrow, the argument goes, needs extra capacity if it is to reach the levels of service found at competitors elsewhere in Europe, or it will be overtaken by its rivals. ⟨\claim⟩ ⟨\reason⟩

`http://news.bbc.co.uk/1/hi/uk/7828694.stm`

Syntax

Originated with Apothéloz

Formula

A *formula* is an expression of the form

$$(-)\mathcal{R}(y) : (-)\mathcal{C}(x)$$

where each of x and y is $\begin{cases} \textit{either} \text{ an expression of the same form} \\ \textit{or} \text{ a formula of a given logical language } \mathbb{L}. \end{cases}$

The set of formulas is denoted $\mathtt{Arg}(\mathbb{L})$.

Syntax

Argument

An *argument* is a formula of $\mathtt{Arg}(\mathbb{L})$ of the form

$$\mathcal{R}(y) : (-)\mathcal{C}(x)$$

Two types of argument

$$\begin{array}{rll} \mathcal{R}(y) : \mathcal{C}(x) & \text{means that} & \text{“}y \text{ is a reason for concluding } x\text{”} \\ \mathcal{R}(y) : -\mathcal{C}(x) & \text{means that} & \text{“}y \text{ is a reason for not concluding } x\text{”} \end{array}$$

Examples of arguments

1. Paul: Carl will fail his exams (*fe*). He did not work hard ($\neg wh$).
$$\mathcal{R}(\neg wh) : \mathcal{C}(fe)$$
2. Mary: No, he will not fail. The exams will be easy this semester (*ee*).
$$\mathcal{R}(ee) : \mathcal{C}(\neg fe)$$
3. John: Carl is very smart! (*sm*).
$$\mathcal{R}(sm) : -\mathcal{C}(fe)$$

Syntax

Rejection (anti-argument)

A *rejection of an argument* is a formula of $\mathtt{Arg}(\mathbb{L})$ of the form

$$-\mathcal{R}(y) : (-)\mathcal{C}(x)$$

Two types of rejection

$$-\mathcal{R}(y) : \mathcal{C}(x) \text{ means that ``}y\text{ is not a reason for concluding }x\text{''}$$
$$-\mathcal{R}(y) : -\mathcal{C}(x) \text{ means that ``}y\text{ is not a reason for not concluding }x\text{''}$$

Examples with rejections

1. Paul: The fact that Carl is smart is not a reason to stop concluding that he will fail his exams. $-\mathcal{R}(sm) : -\mathcal{C}(fe)$
2. John: Anyway, the fact that Carl did not work hard is not a reason to conclude that he will fail his exams. $-\mathcal{R}(\neg wh) : \mathcal{C}(fe)$
3. Mary: As stress (st) is the reason that Carl will fail his exams, it is not the fact that he did not work hard. $\mathcal{R}(\mathcal{R}(st) : \mathcal{C}(fe)) : \mathcal{C}(-\mathcal{R}(\neg wh) : \mathcal{C}(fe))$
4. Sara: He is not stressed at all. $\mathcal{R}(\neg st) : \mathcal{C}(-\mathcal{R}(st) : \mathcal{C}(fe))$

Syntax

Levels of counterargument

So, for an argument $\mathcal{R}(y) : \mathcal{C}(x)$, there are various levels of counterargument.

1. $\mathcal{R}(z) : \mathcal{C}(\neg x) =$ "z is a reason for concluding $\neg x$"
2. $\mathcal{R}(z) : -\mathcal{C}(x) =$ "z is a reason for not concluding x"
3. $-\mathcal{R}(z) : \mathcal{C}(x) =$ "z is not a reason for concluding x"

Examples

- $\mathcal{R}(\mathit{bird}) : \mathcal{C}(\mathit{fly})$
- $\mathcal{R}(\mathit{dead}) : \mathcal{C}(\neg \mathit{fly})$
- $\mathcal{R}(\mathit{penguin}) : -\mathcal{C}(\mathit{fly})$
- $-\mathcal{R}(\mathit{egglaying}) : \mathcal{C}(\mathit{fly})$

Syntax: Use as a target language

$\langle x_1 \rangle$*Heathrow needs more capacity*$\langle \backslash x_1 \rangle$

$\langle y_1 \rangle$*Heathrow runs at close to 100% capacity. With demand for air travel predicted to double in a generation, Heathrow will not be able to cope without a third runway*$\langle \backslash y_1 \rangle$*, say those in favour of the plan.*

$\langle z_1 \rangle$*Because the airport is over-stretched*$\langle \backslash z_1 \rangle$*,* $\langle z_2 \rangle$*any problems which arise cause knock-on delays*$\langle \backslash z_2 \rangle$*.* $\langle z_3 \rangle$*Heathrow, the argument goes, needs extra capacity if it is to reach the levels of service found at competitors elsewhere in Europe, or it will be overtaken by its rivals*$\langle \backslash z_3 \rangle$*.*

1. $\mathcal{R}(y_1) : \mathcal{C}(x_1)$
2. $\mathcal{R}(\mathcal{R}(\mathcal{R}(z_1) : \mathcal{C}(z_2)) : \mathcal{C}(z_3)) : \mathcal{C}(x_1)$

Advantages (excerpt)

No other logic-based approach to modelling argumentation provides a language for expressing rejection of arguments in the object language.

Example

- We can differentiate between the following where *cr* denotes *"The car is red"* and *bc* denotes *"We should buy the car"*.
  - $-\mathcal{R}(cr) : \mathcal{C}(bc)$ could counter $\mathcal{R}(cr) : \mathcal{C}(bc)$ because we need to consider more than the colour of the car when buying.
  - $\mathcal{R}(cr) : -\mathcal{C}(bc)$ could counter $\mathcal{R}(cr) : \mathcal{C}(bc)$ because we do not like the colour red for a car.
- Even if we identify the rejection $-\mathcal{R}(cr) : \mathcal{C}(bc)$, it is possible that we could identify another argument for buying the car using other criteria such as $\mathcal{R}(ec \wedge sp) : \mathcal{C}(bc)$ where *ec* denotes *"The car is economical"* and *sp* denotes *"The car is spacious"*.

Advantages (excerpt)

Most natural language arguments are enthymemes

- Since most arguments are enthymemes, some premises (and sometimes claim) are implicit.
- Decoding enthymemes from natural language into logic requires
  - extensive background and/or common-sense knowledge.
  - and deep parsing techniques

Our approach handles enthymemes without decoding

For example

- Paul's car is in the park (*pr*) because it is broken (*br*), hence we cannot conclude that Paul is in his office (*of*).

$$\mathcal{R}(\mathcal{R}(br) : \mathcal{C}(pr)) : -\mathcal{C}(of)$$

Reasoning

Consequence relation

The consequence relation $\Vdash$ is the least closure of a set of *inference rules* extended with one *meta-rule*.

Any inference rule can be reversed

$$\frac{\mathcal{R}(y) : \Phi}{-\mathcal{R}(y) : \Psi} \qquad \text{into} \qquad \frac{\mathcal{R}(y) : \Psi}{-\mathcal{R}(y) : \Phi}$$

Meta-rule

Let $i, j \in \{0, 1\}$

$$\frac{-^{(i)}\mathcal{R}(y) : \Phi \qquad \alpha_1 \cdots \alpha_n}{-^{(j)}\mathcal{R}(y) : \Psi} \text{ can be reversed into } \frac{-^{(1-j)}\mathcal{R}(y) : \Psi \qquad \alpha_1 \cdots \alpha_n}{-^{(1-i)}\mathcal{R}(y) : \Phi}$$

Reasoning

Consistency

Let x be a formula in $\mathbb{L}$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x)}{-\mathcal{R}(y) : -\mathcal{C}(x)} \qquad \frac{\mathcal{R}(y) : \mathcal{C}(x)}{\mathcal{R}(y) : -\mathcal{C}(\neg x)}$$

Example

Carl works hard (wh), so he will pass his exams (pe).

$$\frac{\mathcal{R}(wh) : \mathcal{C}(pe)}{-\mathcal{R}(wh) : -\mathcal{C}(pe)} \qquad \frac{\mathcal{R}(wh) : \mathcal{C}(pe)}{\mathcal{R}(wh) : -\mathcal{C}(\neg pe)}$$

Proposition

The inference rules below are derived from (Consistency) and the meta-rule (where x is a formula in $\mathbb{L}$ in the first, third and fourth inference rules).

$$\frac{\mathcal{R}(y) : \mathcal{C}(x)}{-\mathcal{R}(y) : \mathcal{C}(\neg x)} \quad \frac{\mathcal{R}(y) : -\mathcal{C}(x)}{-\mathcal{R}(y) : \mathcal{C}(x)} \quad \frac{\mathcal{R}(y) : \mathcal{C}(\neg x)}{\mathcal{R}(y) : -\mathcal{C}(x)} \quad \frac{\mathcal{R}(y) : \mathcal{C}(\neg x)}{-\mathcal{R}(y) : \mathcal{C}(x)}$$

Example of a reasoning system: Indicative reasoning

Inference rules of indicative reasoning

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \quad \mathcal{R}(x) : \mathcal{C}(y) \quad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(x) : \mathcal{C}(z)} \qquad \textbf{(Mutual Support)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : \mathcal{C}(x)}{\mathcal{R}(y \vee z) : \mathcal{C}(x)} \qquad \textbf{(Or)}$$

$$\frac{\mathcal{R}(y \wedge z) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y) : \mathcal{C}(x)} \qquad \textbf{(Cut)}$$

$$\frac{\mathcal{R}(y \wedge z) : \mathcal{C}(x)}{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))} \qquad \textbf{(Importation)}$$

$$\frac{\mathcal{R}(z) : \mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))}{\mathcal{R}(y \wedge z) : \mathcal{C}(x)} \qquad \textbf{(Exportation)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(\mathcal{R}(z) : \mathcal{C}(x))}{\mathcal{R}(z) : \mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x))} \qquad \textbf{(Permutation)}$$

Example of a reasoning system: Indicative reasoning

The following inferences do not hold for indicative reasoning

$$\frac{\forall x \in \mathbb{L}}{\mathcal{R}(x) : \mathcal{C}(x)} \quad \textbf{(Reflexivity)}$$

$$\frac{y \models x}{\mathcal{R}(y) : \mathcal{C}(x)} \quad \textbf{(Logical Consequence)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad y \models z \qquad z \models y}{\mathcal{R}(z) : \mathcal{C}(x)} \quad \textbf{(Left Logical Equivalence)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad x \models w}{\mathcal{R}(y) : \mathcal{C}(w)} \quad \textbf{(Right Logical Consequence)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad z \models y}{\mathcal{R}(z) : \mathcal{C}(x)} \quad \textbf{(Left Logical Consequence)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y) : \mathcal{C}(x \wedge z)} \quad \textbf{(And)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y \wedge z) : \mathcal{C}(x)} \quad \textbf{(Cautious Monotonicity)}$$

$$\frac{\mathcal{R}(z) : \mathcal{C}(y) \qquad \mathcal{R}(y) : \mathcal{C}(x)}{\mathcal{R}(z) : \mathcal{C}(x)} \quad \textbf{(Transitivity)}$$

Example of a reasoning system: Indicative reasoning

Example to motivate need for failure of (Reflexivity)

Let x stand for "I should have a pay rise".

$$\frac{\forall x \in \mathbb{L}}{\mathcal{R}(x) : \mathcal{C}(x)}$$

Example to motivate need for failure of (Right Logical Consequence)

Let x be "temp in range 39-41C" and let w be "temp in range 36-41C"

$$\frac{\mathcal{R}(flu) : \mathcal{C}(x) \qquad x \models w}{\mathcal{R}(flu) : \mathcal{C}(w)}$$

Example to motivate need for failure of (And)

Let y be "Paul is standing in the middle of the road while a car is approaching", x be "Paul should move forward", and z be "Paul should move backwards".

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(y) : \mathcal{C}(z)}{\mathcal{R}(y) : \mathcal{C}(x \land z)}$$

Example of a reasoning system: Indicative reasoning

Proposition

The following non-trivialization property, $i, j \in \{0, 1\}$, holds for the $\Vdash$ relation:

$$\{-^{(i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x), -^{(1-i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x)\} \nVdash \mathtt{Arg}(\mathbb{L})$$

$$\{-^{(i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x), -^{(i)}\mathcal{R}(y) : -^{(1-j)}\mathcal{C}(x)\} \nVdash \mathtt{Arg}(\mathbb{L})$$

Proposition

Whatever $i, j \in \{0, 1\}$,

$$\nVdash \frac{-^{(i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x)}{-^{(1-i)}\mathcal{R}(y) : -^{(j)}\mathcal{C}(x)}$$

Proposition

The following are properties of the $\Vdash$ relation where Δ is a set of (rejections of) arguments, and α and β are (rejections of) arguments.

$\Delta \Vdash \alpha$ if $\alpha \in \Delta$ **(Reflexivity)**
$\Delta \cup \{\alpha\} \Vdash \beta$ if $\Delta \Vdash \beta$ **(Monotonicity)**
$\Delta \Vdash \beta$ if $\Delta \cup \{\alpha\} \Vdash \beta$ and $\Delta \Vdash \alpha$ **(Cut)**

Representing attacks

Formulas and inferences capturing attacks

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : \mathcal{C}(\neg x)}{\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(\neg x)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \quad \textbf{(Strong Rebuttal)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : -\mathcal{C}(x)}{\mathcal{R}(\mathcal{R}(z) : -\mathcal{C}(x)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \quad \textbf{(Weak Rebuttal)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : \mathcal{C}(\neg y)}{\mathcal{R}(\mathcal{R}(z) : \mathcal{C}(\neg y)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \quad \textbf{(Strong Premise Attack)}$$

$$\frac{\mathcal{R}(y) : \mathcal{C}(x) \qquad \mathcal{R}(z) : -\mathcal{C}(y)}{\mathcal{R}(\mathcal{R}(z) : -\mathcal{C}(y)) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x))} \quad \textbf{(Weak Premise Attack)}$$

$$\mathcal{R}(z) : \mathcal{C}(-\mathcal{R}(y) : \mathcal{C}(x)) \qquad \textbf{(Strong Reason Attack)}$$

$$\mathcal{R}(z) : -\mathcal{C}(\mathcal{R}(y) : \mathcal{C}(x)) \qquad \textbf{(Weak Reason Attack)}$$

$$-\mathcal{R}(y) : \mathcal{C}(x) \qquad \textbf{(Pure Reason Attack)}$$

For comparison with structured argumentation, strong rebuttal captures "rebuttal", strong premise attack captures "assumption attack", and weak reason attack captures Pollock's undercutting.

Representing attacks

Example of strong rebuttal (capturing "rebuttal")

Nixon is quaker (nq) and Nixon is a republican (nr). Is Nixon a pacifist (np)?

$$\frac{\mathcal{R}(nq) : \mathcal{C}(np) \qquad \mathcal{R}(nr) : \mathcal{C}(\neg np)}{\mathcal{R}(\mathcal{R}(nr) : \mathcal{C}(\neg np)) : \mathcal{C}(-\mathcal{R}(nq) : \mathcal{C}(np))}$$

Example of strong premise attack (capturing "assumption attack")

The weather is good (gw) so the bbq will be a success (bs). But, the weather report predicts rain (ra).

$$\frac{\mathcal{R}(gw) : \mathcal{C}(bs) \qquad \mathcal{R}(ra) : \mathcal{C}(\neg gw)}{\mathcal{R}(\mathcal{R}(ra) : \mathcal{C}(\neg gw)) : \mathcal{C}(-\mathcal{R}(gw) : \mathcal{C}(bs))}$$

Example of weak reason attack (capturing Pollock's undercutting)

The object looks red (lr). It is illuminated by red light (il). Thus, we cannot conclude that the object looking red implies it being indeed red (re).

$$\mathcal{R}(\mathcal{R}(il) : \mathcal{C}(lr)) : -\mathcal{C}(\mathcal{R}(lr) : \mathcal{C}(re))$$

Conclusion

Some advantages of our approach

- Target language for mined arguments
- Representation of link between reason and claim
- Explicit representation of support in the object language
- Practical representation of enthymemes
- Representation of rejections (anti-arguments)
- Nesting of arguments and rejections
- Explicit representation of attacks in the object language
- Reasoning systems (inference rules)