

# A Semantics For Inconsistency-Tolerant Agents

Linda Gutsche  
linda.gutsche@inria.fr

**Real  
Bargain**

- I'm paying 10€ per month for the subscription.
- The subscription gives me a 50% reduction on everything.
- I spend approximately 100€ per year on purchases.
- The subscription is saving me money.

**cannot  
be  
true**

Yet someone may still believe it because they did not perform the computations needed to detect the contradiction.

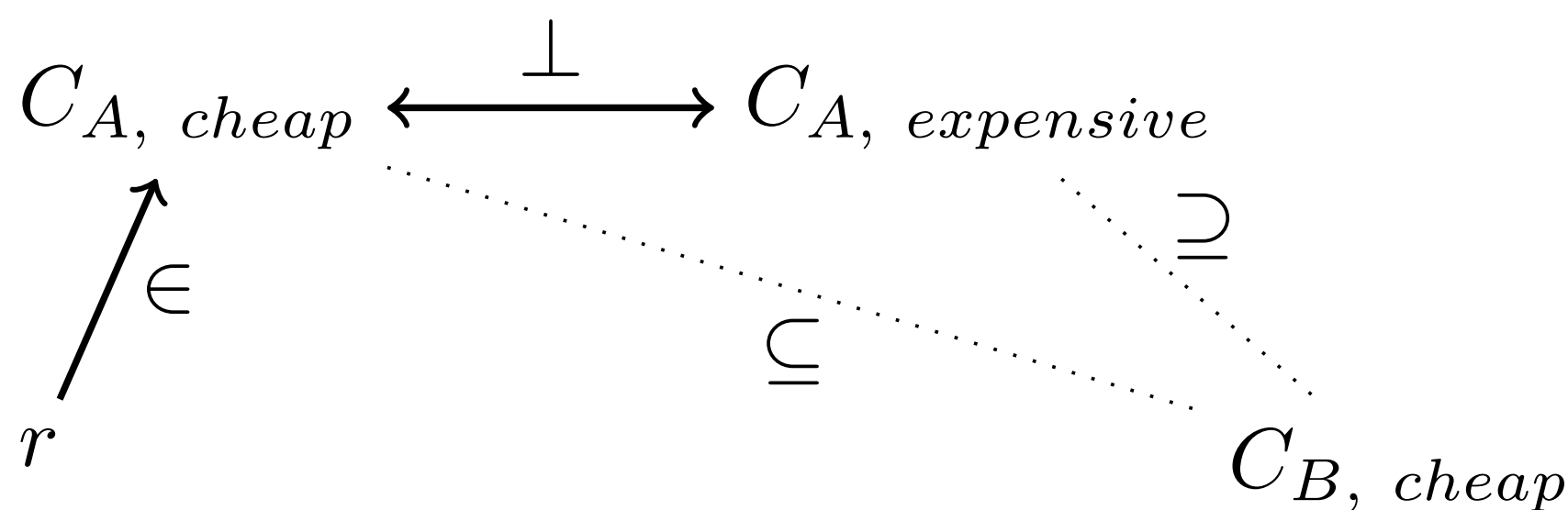
## Classical logic cannot represent such agents

In classical logic, semantics and the deductive systems coincide: what is true is exactly what can be derived. If a statement  $\varphi$  is true in all the models of a set  $\Gamma$ , then  $\varphi$  is derivable from  $\Gamma$  (completeness). Conversely, if  $\varphi$  is derivable from  $\Gamma$ , then it is true in all the models of  $\Gamma$  (soundness).

An inconsistent set of statements has no model. Therefore, a complete reasoner would consider any statement a logical consequence of such a set.

**But believing contradictory things does not make someone believe everything.**

## Inconsistency does not necessarily persist



An example of local inconsistency in ontology alignment. Alice believes that whenever she finds a restaurant to be cheap, Bob agrees. Alice also believes that every time Bob describes a restaurant as cheap, she actually finds it to be expensive.

Experimental results show that, in ontology alignment repair, under the relaxation modality which ensures that every alignment is eventually used, agents always converge towards a fully consistent state.

(Euzenat, 2017)

By revising their beliefs in response to new evidence, agents with incomplete deductive abilities can restore consistency without ever being aware that their beliefs were initially inconsistent.

## Explicit and implicit beliefs

It is useful to distinguish between what agents explicitly hold and what they can derive from it.

$\Delta_a\varphi$ : agent  $a$  *explicitly* believes  $\varphi$

$\Box_a\varphi$ : agent  $a$  *implicitly* believes  $\varphi$

$$\varphi ::= p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Delta_a\varphi \mid \Box_a\varphi$$

$p \in \mathcal{P}, a \in \mathcal{A}$

$\mathcal{P}$ : countable set of atomic propositions

$\mathcal{A}$ : finite set of agents

(Inspired by Lorini, 2020)

## Incomplete Deduction

Agents are not perfect reasoners. Their deductive process does not necessarily coincide with logical consequence. We represent it with a closure function  $Cn_a$ .

$B_a \subseteq \mathcal{L}$ : agent  $a$ 's belief base

$Cn_a$ : agent  $a$ 's deductive closure

$$\Delta_a\varphi \text{ iff } \varphi \in B_a$$

$$\Box_a\varphi \text{ iff } \varphi \in Cn_a(B_a)$$

**Our framework generalises the logic of explicit and implicit beliefs by allowing agents to be incomplete reasoners and to hold inconsistent beliefs.**

When agents are complete reasoners, it coincides with classical semantics.

