

Artificial Intelligence

A Course About Foundations



Plausible Reasoning

Marco Piastra

Plausible (defeasible) reasoning

Why plausible reasoning?

Consider a generic entailment problem $\Gamma \models \varphi$?

Four possible answers:

1. $\Gamma \models \varphi$
 $\Gamma \not\models \neg\varphi$
2. $\Gamma \not\models \varphi$
 $\Gamma \models \neg\varphi$
3. $\Gamma \models \varphi$ ————— * This case occurs only when Γ is contradictory, i.e. unsatisfiable
 $\Gamma \models \neg\varphi$
4. $\Gamma \not\models \varphi$
 $\Gamma \not\models \neg\varphi$

Case 4. is quite frequent: "our knowledge Γ does not allow deciding about φ "

Plausible (defeasible) reasoning

A reasoning process where the **relation** between formulae is rationally plausible yet not necessarily correct (in the classical logical sense)

Notation:

$\Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi$ says that φ is a **plausible** derivation from Γ in $\langle \text{SysLog} \rangle$

i.e. a specific reasoning method

Properties of $\vdash_{\langle \text{SysLog} \rangle}$

$$\Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi \Rightarrow \Gamma \not\vdash_{\langle \text{SysLog} \rangle} \neg \varphi$$

(coherence)

$$\Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi \Rightarrow \Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi$$

(compatibility with derivation)

$$\Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi \not\Rightarrow \Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi (\Rightarrow \Gamma \models \varphi)$$

(not necessarily correct)

It occurs very often in practice:

“The train schedule does not report a train to Milano at 06:55,
therefore we assume that such a train does not exist”

Most databases contain positive information only
Negative facts are typically derived ‘by default’

Closed-World Assumption (CWA)

$$\{\Gamma \not\models \alpha\} \vdash_{CWA} \neg\alpha \quad (\alpha \text{ is an atom})$$

Example (a program):

$$\Pi \equiv \{\{Philosopher(socrates)\}, \{Philosopher(plato)\}, \{Cat(felix)\}\}$$

The program Π can be rewritten in L_{FO} as:

$$\forall x ((x = socrates) \rightarrow Philosopher(x))$$

$$\forall x ((x = plato) \rightarrow Philosopher(x))$$

$$\forall x ((x = felix) \rightarrow Cat(x))$$

The *Closed-World Assumption* (CWA) means completing (i.e. extending) the program Π :

$$\forall x ((x = felix) \leftrightarrow Cat(x))$$

$$\forall x ((x = socrates \vee x = plato) \leftrightarrow Philosopher(x))$$

Notice the double implication

Then these plausible inferences become sound:

$$\Pi \vdash_{CWA} \neg Cat(socrates)$$

$$\Pi \vdash_{CWA} \neg Cat(plato)$$

$$\Pi \vdash_{CWA} \neg Philosopher(felix)$$

Plausible (defeasible) reasoning

- Inference in *defeasible reasoning* is

Non-monotonic

$$\Gamma \vdash_{\langle \text{SysLog} \rangle} \varphi \not\Rightarrow \Gamma \cup \Delta \vdash_{\langle \text{SysLog} \rangle} \varphi$$

The arrival of new pieces of information may falsify inferences that used to be justified

e.g. an extra train to Milano at 06:55 is announced ...

Systemic

In classical logic, the soundness of all inferences schema depend only on the few formulae involved.

e.g. $\varphi \rightarrow \psi, \varphi \vdash \psi$

In defeasible reasoning, inferences are justified by an entire theory Γ

One must check the entire database (see CWA): $\Gamma \not\vdash \varphi \vdash_{\langle \text{SysLog} \rangle} \neg \varphi$

Inference and reasoning (according to C. S. Peirce, 1870 c.a.)

■ Different types of reasoning

Deductive inference (sound)

*Derive only what is justified in terms of **entailment***

"All beans in this bag are white"

"This handful of beans comes from this bag"

"This is a handful of white beans"

$$\frac{\begin{array}{l} \forall x \varphi(x) \rightarrow \psi(x) \\ \varphi(a) \end{array}}{\psi(a)}$$

Inductive inference (plausible)

From repeated occurrences, derive rules

"This handful of beans comes from this bag"

"This is a handful of white beans"

"All beans in this bag are white"

$$\frac{\begin{array}{l} \psi(a) \\ \varphi(a) \end{array}}{\forall x \varphi(x) \rightarrow \psi(x)}$$

Abductive inference (plausible)

From rules and outcomes, derive premises

"All beans in this bag are white"

"This is a handful of white beans"

"This handful of beans comes from this bag"

$$\frac{\begin{array}{l} \forall x \varphi(x) \rightarrow \psi(x) \\ \psi(a) \end{array}}{\varphi(a)}$$