Artificial Intelligence

Deductive Systems

Marco Piastra

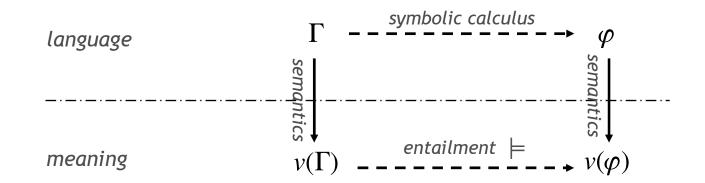
Deductive Systems [1]

Symbolic calculus?

 A wff φ is entailed by a set of wff Γ iff every model of Γ is also model of φ

Formally:

 $\Gamma\models\!\varphi$



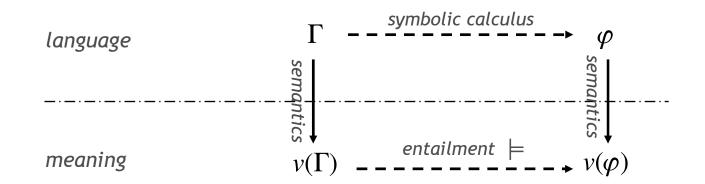
Note that, in the definition above, the set of <u>all</u> possible models is considered

Symbolic calculus?

 A wff φ is entailed by a set of wff Γ iff every model of Γ is also model of φ

Formally:

 $\Gamma\models\!\varphi$



Note that, in the definition above, the set of <u>all</u> possible models is considered

Can we detect entailment by working on wffs only?

For instance, by applying reasoning *schemas*...

Axiomatic method (i.e. Hilbert System, 1899)

Language, axioms and rules of inference

$< L_P, Ax, Inf >$

L_P is a propositional language whose signature is P
Ax a set of wffs, i.e. the axioms
Inf is a set of inference rules

Axioms

The axioms (of a logic) describe the <u>reasoning schemas</u> (of that logic)

Axiom <u>schemas</u> for propositional logic (Lukasiewicz, 1917)

Ax1
$$\vdash \varphi \rightarrow (\psi \rightarrow \varphi)$$
Ax2 $\vdash (\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \chi))$ Ax3 $\vdash (\neg \varphi \rightarrow \neg \psi) \rightarrow (\psi \rightarrow \varphi)$

Each wff obtained by substituting the *meta-variabes* φ , $\psi \in \chi$ with a wff is an axiom The wffs thus obtained are also called *instances* of axiom

Examples:

$$[A \to (\neg A \to A)$$

$$[Ax1: \varphi/A, \psi/\neg A]$$

$$[(\neg (B \lor C) \to \neg D) \to (D \to (B \lor C))$$

$$[Ax3: \varphi/(B \lor C), \psi/D]$$

All axiom instances are *tautologies* (But do not rely on this for the <u>definition</u> of deductive systems)

Inference Rules

Recall that $\{\varphi \rightarrow \psi, \varphi\} \models \psi$ is valid

- Inference rules are fundamental in any symbolic calculus (also known as derivation or deduction rules) They work on the structure of wffs
- For propositional logic, just one inference rule is sufficient
 Modus Ponens (MP):

$$\begin{array}{c}
\varphi \to \psi \\
\varphi \\
\hline
\psi
\end{array}$$

It can be written also in this way:

 $\{\varphi \rightarrow \psi, \varphi\} \models \psi$ (i.e. from $\{\varphi \rightarrow \psi, \varphi\}$, ψ is derivable)

Proofs (also *derivations*)

- A proof (or derivation) of a wff φ from a set of wffs Γ Is a *finite* sequence of steps: $\Gamma \models \phi_1, \Gamma \models \phi_2, ..., \Gamma \models \phi_n$ Admissible steps, at stage *i* :
 - 1) ϕ_i is an instance of an axiom schema Ax_n
 - 2) ϕ_i is in Γ
 - 3) ϕ_i has been obtained from two previous steps, via *Modus Ponens*

In the final step, the wff to be proved is obtained: $\phi_n = \varphi$

The notation is $\Gamma \vdash \varphi$ " φ is derivable Γ "

There must be many different ways to show that $\Gamma \vdash \varphi$, i.e. many different *proofs*

Note that:

$\Gamma \vdash Axn$	(an axiom or axiom instance can be derived from any Γ)
$\vdash Axn$	(an axiom or axiom instance can be derived from an empty Γ)
$\{\varphi,\} \models \varphi$	(any $arphi$ can be derived form a Γ that contains it)

Derivations, an *incomplete* example

■ Same problem ("Harry is happy") $B \lor D \lor \neg (A \land C), B \lor C, A \lor D, \neg B \vdash D$ Rewrite the problem in equivalent terms using only \neg and \rightarrow $C \rightarrow (\neg B \rightarrow (A \rightarrow D)), \neg B \rightarrow C, \neg A \rightarrow D, \neg B \vdash D$

1:
$$\Gamma \vdash \neg B \rightarrow C$$
2: $\Gamma \vdash \neg B$ 3: $\Gamma \vdash C$ 4: $\Gamma \vdash C \rightarrow (\neg B \rightarrow (A \rightarrow D))$ 5: $\Gamma \vdash \neg B \rightarrow (A \rightarrow D)$ 6: $\Gamma \vdash A \rightarrow D$ 7: $\Gamma \vdash \neg A \rightarrow D$ 8*: $\Gamma \vdash (\neg \varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow \psi) \rightarrow \psi)$ 9: $\Gamma \vdash (\neg A \rightarrow D) \rightarrow ((A \rightarrow D) \rightarrow D)$ 10: $\Gamma \vdash (A \rightarrow D) \rightarrow D$ 11: $\Gamma \vdash D$

Any wff implies itself

 $\vdash \varphi \to \varphi$

$$\begin{aligned} 1: & \vdash (\varphi \rightarrow ((\varphi \rightarrow \varphi) \rightarrow \varphi)) \rightarrow ((\varphi \rightarrow (\varphi \rightarrow \varphi)) \rightarrow (\varphi \rightarrow \varphi)) & (Ax2) \\ 2: & \vdash (\varphi \rightarrow ((\varphi \rightarrow \varphi) \rightarrow \varphi))) & (Ax1) \\ 3: & \vdash (\varphi \rightarrow (\varphi \rightarrow \varphi)) \rightarrow (\varphi \rightarrow \varphi) & (MP 1, 2) \\ 4: & \vdash (\varphi \rightarrow (\varphi \rightarrow \varphi)) & (Ax1) \\ 5: & \vdash \varphi \rightarrow \varphi & (MP 3, 4) \end{aligned}$$

Deduction (meta)-theorem

- Deduction theorem $\Gamma \cup \{\varphi\} \models \psi \iff \Gamma \models \varphi \rightarrow \psi$
- (semantic dual of Deduction Theorem)

 $\Gamma \cup \{\varphi\} \models \psi \quad \Leftrightarrow \quad \Gamma \models \varphi \to \psi$

• The order of hypotheses is irrelevant $\vdash (\varphi \to (\psi \to \chi)) \to (\psi \to (\varphi \to \chi))$

1:
$$(\varphi \rightarrow (\psi \rightarrow \chi)), \psi, \varphi \vdash (\varphi \rightarrow (\psi \rightarrow \chi))$$

2: $(\varphi \rightarrow (\psi \rightarrow \chi)), \psi, \varphi \vdash \varphi$
3: $(\varphi \rightarrow (\psi \rightarrow \chi)), \psi, \varphi \vdash \psi \rightarrow \chi$ (*MP* 1,2)
4: $(\varphi \rightarrow (\psi \rightarrow \chi)), \psi, \varphi \vdash \psi$
5: $(\varphi \rightarrow (\psi \rightarrow \chi)), \psi, \varphi \vdash \chi$ (*MP* 3,4)
6: $(\varphi \rightarrow (\psi \rightarrow \chi)), \psi \vdash \varphi \rightarrow \chi$ (*Ded*)
7: $(\varphi \rightarrow (\psi \rightarrow \chi)) \vdash \psi \rightarrow (\varphi \rightarrow \chi)$ (*Ded*)
8: $\vdash (\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow (\psi \rightarrow (\varphi \rightarrow \chi))$ (*Ded*)

■ Double negation implies affirmation $\vdash \neg \neg \varphi \rightarrow \varphi$

1:
$$\models \neg \neg \varphi \rightarrow (\neg \neg \neg \neg \varphi \rightarrow \neg \neg \varphi)$$
(Ax1)2: $\neg \neg \varphi \models \neg \neg \neg \neg \varphi \rightarrow \neg \neg \varphi$ (Ded)3: $\neg \varphi \models (\neg \neg \neg \neg \varphi \rightarrow \neg \neg \varphi) \rightarrow (\neg \varphi \rightarrow \neg \neg \varphi)$ (Ax3)4: $\neg \varphi \models \neg \varphi \rightarrow \neg \neg \varphi$ (MP 3,2)

4:
$$\neg \varphi \vdash \neg \varphi \rightarrow \neg \neg \neg \varphi$$

5: $\neg \neg \varphi \vdash (\neg \varphi \rightarrow \neg \neg \neg \varphi) \rightarrow (\neg \neg \varphi \rightarrow \varphi)$ (Ax3)

$$6: \neg \neg \varphi \vdash \neg \neg \varphi \rightarrow \varphi \qquad (MP 5,4)$$

7:
$$\neg \neg \varphi \vdash \neg \neg \varphi$$
 (MP 6,7)

 8: $\neg \neg \varphi \vdash \varphi$
 (Ded)

A rule is false if the LHS is true and the RHS is false

$$\vdash \varphi \to (\neg \psi \to \neg (\varphi \to \psi))$$

1:
$$\varphi, (\varphi \rightarrow \psi) \vdash \psi$$
(MP)2: $\varphi \vdash (\varphi \rightarrow \psi) \rightarrow \psi$ (Ded)3: $\varphi \vdash ((\varphi \rightarrow \psi) \rightarrow \psi) \rightarrow (\neg \psi \rightarrow \neg (\varphi \rightarrow \psi))$ (Ax3)4: $\varphi \vdash \neg \psi \rightarrow \neg (\varphi \rightarrow \psi)$ (MP 3,2)5: $\vdash \varphi \rightarrow (\neg \psi \rightarrow \neg (\varphi \rightarrow \psi))$ (Ded)

■ From absurd, anything can be derived ("*Ex absurdo sequitur quodlibet*"): $\vdash \varphi \rightarrow (\neg \varphi \rightarrow \psi)$ (vale a dire $\varphi, \neg \varphi \vdash \psi$)

1:	$\varphi, \neg \varphi \vdash \neg \varphi \rightarrow (\neg \psi \rightarrow \neg \varphi)$	(Ax1)
2:	$\varphi, \neg \varphi \vdash \neg \varphi$	
3:	$\varphi, \neg \varphi \vdash \neg \psi \to \neg \varphi$	(<i>MP</i> 1,2)
4:	$\varphi, \neg \varphi \vdash (\neg \psi \rightarrow \neg \varphi) \rightarrow (\varphi \rightarrow \psi)$	(Ax3)
5:	$\varphi, \neg \varphi \vdash \varphi \rightarrow \psi$	(<i>MP</i> 4,3)
6:	$\varphi, \neg \varphi \vdash \varphi$	
7:	$\varphi, \neg \varphi \vdash \psi$	(<i>MP</i> 5,6)
8:	$\varphi \vdash \neg \varphi \to \psi$	(Ded)
9:	$\vdash \varphi \to (\neg \varphi \to \psi)$	(Ded)

A set of wffs that contains a contradiction is called *incoherent* (or *inconsistent*)

From an incoherent set, anything can be derived, including a contradiction like: $\psi \land \neg \psi$

Derivations: example 5 (using theorems)

• When falsity implies contradiction, then it must be true: $\vdash (\neg \varphi \rightarrow \varphi) \rightarrow \varphi$

1:
$$\neg \varphi \rightarrow \varphi, \neg \varphi \vdash \neg \varphi$$

2: $\neg \varphi \rightarrow \varphi, \neg \varphi \vdash \neg \varphi \rightarrow \varphi$
3: $\neg \varphi \rightarrow \varphi, \neg \varphi \vdash \varphi$
4: $\neg \varphi \rightarrow \varphi, \neg \varphi \vdash \varphi \rightarrow (\neg \varphi \rightarrow \neg (\neg \varphi \rightarrow \varphi))$
5: $\neg \varphi \rightarrow \varphi, \neg \varphi \vdash \neg \varphi \rightarrow \neg (\neg \varphi \rightarrow \varphi)$
6: $\neg \varphi \rightarrow \varphi, \neg \varphi \vdash \neg (\neg \varphi \rightarrow \varphi)$
7: $\neg \varphi \rightarrow \varphi \vdash \neg (\neg \varphi \rightarrow \varphi)$
8: $\neg \varphi \rightarrow \varphi \vdash (\neg \varphi \rightarrow \neg (\neg \varphi \rightarrow \varphi)) \rightarrow ((\neg \varphi \rightarrow \varphi) \rightarrow \varphi)$
9: $\neg \varphi \rightarrow \varphi \vdash (\neg \varphi \rightarrow \varphi) \rightarrow \varphi$
10: $\neg \varphi \rightarrow \varphi \vdash \neg \varphi \rightarrow \varphi$
11: $\neg \varphi \rightarrow \varphi \vdash \varphi$
12: $\vdash (\neg \varphi \rightarrow \varphi) \rightarrow \varphi$
(MP 1,5)
(MP 1,5)
(MP 7,8)
(MP 9,10)
(Ded)

Deductive Systems [15]

Derivations: Theorem "X"

• Resolution rule (see the first *incomplete* example): $\vdash (\neg \varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow \psi) \rightarrow \psi)$

$$\begin{array}{ll} 1: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi) \vdash (\neg \varphi \rightarrow \psi) \\ 2: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi) \vdash (\neg \varphi \rightarrow \psi) \rightarrow (\neg \psi \rightarrow \varphi) \\ 3: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi) \vdash (\neg \psi \rightarrow \varphi) \\ 4: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi), \neg \psi \vdash \varphi \\ 5: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi), \neg \psi \vdash \varphi \rightarrow \psi \\ 6: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi), \neg \psi \vdash \psi \\ 7: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi) \vdash (\neg \psi \rightarrow \psi) \\ 8: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi) \vdash ((\neg \psi \rightarrow \psi) \rightarrow \psi) \\ 8: & (\neg \varphi \rightarrow \psi), (\varphi \rightarrow \psi) \vdash \psi \\ 10: & (\neg \varphi \rightarrow \psi) \vdash ((\varphi \rightarrow \psi) \rightarrow \psi) \\ 11: \vdash (\neg \varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow \psi) \rightarrow \psi) \\ \end{array}$$
(Ax3)
(Ax3)
(Ax3)
(MP 1,2)
(Ded)

Rewritten in an equivalent way: $\vdash (\varphi \lor \psi) \rightarrow ((\neg \varphi \lor \psi) \rightarrow \psi)$

Correctness

Correctness

All wffs that are *derivable* from axioms Axn are *tautologies* (valid wffs)

$$\vdash \varphi \Rightarrow \models \varphi$$

It can be verified directly that the axioms schemas Ax1, $Ax2 \in Ax3$ are schemas of tautologies as well

The inference rule of *Modus Ponens* is correct, as it preserves entailment

Any wff entailed by a set of tautologies is a tautology

Completeness

Completeness

All *tautologies* (i.e. *valid* wffs) are *derivable* from axiom schemas Axn

$$\models \varphi \Rightarrow \vdash \varphi$$

Why:

See textbook

Properties of derivations

Coherence (definition)

A set of wffs Γ is *coherent* if it exists at least a wff φ which is not *derivable* from Γ (see theorem 3)

Refutation

```
 \Gamma \cup \{\neg \varphi\} \text{ is incoherent} \Leftrightarrow \Gamma \models \varphi 
 \Gamma \cup \{\neg \varphi\} \text{ is incoherent implies that for any } \psi, \ \Gamma \cup \{\neg \varphi\} \models \psi 
 \text{ In particular} \qquad \Gamma \cup \{\neg \varphi\} \models \varphi 
 \text{ From deduction theorem} \qquad \Gamma \models \neg \varphi \rightarrow \varphi 
 \text{ From Theorem 5} \qquad \Gamma \models (\neg \varphi \rightarrow \varphi) \rightarrow \varphi 
 MP \qquad \Gamma \models \varphi
```

Properties of derivations

• <u>Coherence</u> is equivalent to <u>satisfiability</u>

A set of wffs Γ is *satisfiable* if it is *coherent*

If Γ was incoherent, then it would be possible to derive a contradiction (Theorem 3) But, given that derivability implies entailment, then Γ should be *unsatisfiable*

A set of wffs Γ which is *coherent* is also *satisfiable*

(see textbook)

Syntactic compactness

Consider a set of wffs Γ (not necessarily finite)

 $\Gamma \vdash \varphi \implies$ There exist a *finite* subset $\Sigma \subseteq \Gamma$ such that $\Sigma \vdash \varphi$ (See textbook for a proof)

Properties of derivations

Syntactic monotony

For any Γ and Δ , if $\Gamma \vdash \varphi$ then $\Gamma \cup \Delta \vdash \varphi$

In fact, any derivation of φ from Γ remains valid even if Γ grows larger

Transitivity

If for any $\varphi \in \Sigma$ we have $\Gamma \models \varphi$, then if $\Sigma \models \psi$ then $\Gamma \models \psi$

One can apply the deduction theorem and *MP* repeatedly

Theorems

Given a set of wffs Γ , the **theorems** of Γ is the set of all wff φ that can be derived from $\Gamma \blacktriangleleft (\Gamma \text{ can be empty})$ The set of theorems of Γ is also written as $Th(\Gamma)$ Due to the definition of *derivability*, this means that any such φ descends from $Ax \cup \Gamma$ *This definition has general validity: it applies to any axiomatized logic*

Any *theorem* of Γ is also an *entailment* of Γ $\varphi \in Th(\Gamma) \Rightarrow \Gamma \models \varphi$

Why? (a <u>simple</u> exercise for the reader ...)

Independence of the axiom schemas Axn

Minimality

The proof of completeness requires using them all (see textbook)

Independence

The three schemas are **logically independent**: It is not possible to derive any of them from the other two

 There exist other axiomatizations of propositional logic In one of them, just a single schema is used Using axiom <u>schemas</u>, however, is unavoidable

In other words, using an infinite set of axioms is unavoidable

Theorems, theories, axiomatizations

- Theory = set of wffs (yes, just that) Any set Σ (however defined) is a *theory*
- Theorem = a wff derivable from a set of wffs
 Given a set of wffs Γ, the set of theorems of Γ is the set of all wffs that can be *derived* from Γ

 $Th(\Gamma) = \{\varphi : \Gamma \vdash \varphi\}$

Axiomatizations = a set of wff that describes a *theory*

A set of wffs Γ is an axiomatization of a theory $\Sigma~$ iff

 $\Sigma \equiv Th(\Gamma)$

Axiom schemas Ax*n* describe the *theory* of *valid* wffs in (*classical*) **propositional logic**