Artificial Intelligence

An Advanced Course About Foundations

UNIVERSITÀ DI PAVIA

Symbolic representation and inference: the intuitive idea

Marco Piastra

An Aside: Syntax and Semantics

Symbolic Calculus

- *Solving* quadratic equations

$$x^2 + ax + b = 0$$

$$x^2 + 2(a/2)x + a^2/4 - a^2/4 + b = 0$$

$$(x + a/2)^2 - a^2/4 + b = 0$$

$$(x + a/2)^2 = a^2/4 - b$$

$$x = -a/2 \pm (a^2/4 - b)^{1/2}$$

A sequence of steps: at each step a *transformation rule* is applied

Symbolic Calculus

- *Solving* quadratic equations

$$x^2 + ax + b = 0$$

$$x^2 + 2(a/2)x + a^2/4 - a^2/4 + b = 0$$

$$(x + a/2)^2 - a^2/4 + b = 0$$

$$(x + a/2)^2 = a^2/4 - b$$

$$x = -a/2 \pm (a^2/4 - b)^{1/2}$$

A sequence of steps: at each step a *transformation rule* is applied

From start to end

Start point: a premise (*we assume the truth of something*)

End point: a conclusion (*we state that something else must also be true*)

(*both points are selected by us*)

Symbolic Calculus

- *Solving* quadratic equations

$$x^2 + ax + b = 0$$

$$x^2 + 2(a/2)x + a^2/4 - a^2/4 + b = 0$$

$$(x + a/2)^2 - a^2/4 + b = 0$$

$$(x + a/2)^2 = a^2/4 - b$$

$$x = -a/2 \pm (a^2/4 - b)^{1/2}$$

A sequence of steps: at each step a *transformation rule* is applied

From start to end

Start point: a premise (*we assume the truth of something*)

End point: a conclusion (*we state that something else must also be true*)

(*both points are selected by us*)

Symbolic Calculus

- *Solving* quadratic equations

$$x^2 + ax + b = 0$$

$$x^2 + 2(a/2)x + a^2/4 - a^2/4 + b = 0$$

$$(x + a/2)^2 - a^2/4 + b = 0$$

$$(x + a/2)^2 = a^2/4 - b$$

$$x = -a/2 \pm (a^2/4 - b)^{1/2}$$

A sequence of steps: at each step a *transformation rule* is applied

From start to end

Start point: a premise (*we assume the truth of something*)

End point: a conclusion (*we state that something else must also be true*)

(*both points are selected by us*)

Sequence of steps

Each transformation rule is based on a *semantic equivalence*

Each step must be *correct*, in the above sense

Symbolic Reasoning

Structures in Natural Language

- Natural language is not just a random collection of words:
  it exhibits *systematic patterns* that allow generating and understanding
  an infinite number of sentences from a finite set of elements

Sistematicy

No human being will understand the sentence:

"*Sally* ***likes*** *Harry*"

without being to understand:

"*Harry* ***likes*** *Sally*"

or any other sentences of the kind:

"X ***likes*** Y"

where X and Y can be any nouns or definite descriptions:

"*Ronald's girlfriend* ***likes*** *the cat that Linda brought home yesterday*"

(freely adapted from [Fodor e Phylyshyn, 1988])

Structures in Natural Language

- Natural language is not just a random collection of words:
  it exhibits *systematic patterns* that allow generating and understanding
  an infinite number of sentences from a finite set of elements

Compositionality

The meaning of a sentence is determined by the meanings of its parts
and the rules used to combine them:

"*If* ***it rains***, *(then)* ***the streets will be wet***"

"*If* ***the streets will be wet***, *(then)* ***we better drive carefully***"

where the abstract pattern is:

"***If*** X ***then*** Y"

and X, Y can be any clauses, even composite:

"*If* ***it rains <u>and</u> it's cold***, *(then)* ***people stay indoors***"

(freely adapted from [Fodor e Phylyshyn, 1988])

Reasoning and schemas: *syllogisms*

Many (intuitively) valid arguments obey an abstract schema

All (*humans*) are (*mortals*)
All (*Greeks*) are (*humans*) hence
All (*Greeks*) are (*mortals*)

Abstract schema:

All ***H*** are ***M***
All ***G*** are ***H*** hence
All ***G*** are ***M***

Reasoning and schemas: *syllogisms*

Many (intuitively) valid arguments obey an abstract schema

All (*humans*) are (*mortals*) hence
All (*Greeks*) are (*humans*)
All (*Greeks*) are (*mortals*)

Abstract schema:

All ***H*** are ***M*** hence
All ***G*** are ***H***
All ***G*** are ***M***

CAUTION!
The position of the line is indeed relevant:
in this case the schema does not work

Reasoning and schemas: *syllogisms*

Many (intuitively) valid arguments obey an abstract schema

All (*humans*) are (*mortals*)
All (*Greeks*) are (*mortals*) hence

All (*Greeks*) are (*humans*)

Abstract schema:

All ***H*** are ***M***
All ***G*** are ***M*** hence

All ***G*** are ***H***

CAUTION!
The position with respect to the line is also relevant: in this case the schema does <u>not</u> work

Reasoning and schemas: *syllogisms*

Many (intuitively) valid arguments obey an abstract schema

All (*humans*) are (*mortals*)
All (*Greeks*) are (*humans*) hence
All (*Greeks*) are (*mortals*)

Abstract schema:

All ***H*** are ***M***
All ***G*** are ***H*** hence
All ***G*** are ***M***

The validity of schemas does not depend on *meaning*:

All (*enchanted frogs*) are (*princes*)
All (*princes*) are (*young and beautiful*) hence
All (*enchanted frogs*) are (*young and beautiful*)

Same schema,
different descriptions

Fallacies (*paralogisms*)

Wrong sequence:

All (*humans*) are (*mortals*)
All (*Greeks*) are (*mortals*)

All (*Greeks*) are (*humans*)

The last two sentences have been switched

Referential ambiguities:

(*Nothing*) is better than (*eternal happiness*)
(*One ham sandwich*) is better than (*nothing*)

(*One ham sandwich*) is better than (*eternal happiness*)

'Obscure' subtleties (obscure for now):

All (*enchanted frogs*) are (*princes*)
All (*enchanted frogs*) are (*young and beautiful*)

There is an (*enchanted frog*) which is (*a young and beautiful prince*)

What do we mean by "all"? Do we need the existence of at least one specimen to say "all"?

In such case, the schema would be valid
(*in formal logic it is not, as we will see*)

What is the purpose of *symbolic* logic?

To distinguish correct reasoning
from incorrect reasoning

What is the purpose of *symbolic* logic?

To distinguish correct reasoning
from incorrect reasoning

*by the **formal, symbolic** structure alone*

Truth and Possible Worlds: the quest for *formal semantics*

What is *true*?

- **A world of *cats***

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot	x		x	
Kitty		x	x	
Felix			x	

- **Sentences about this world**

  "Spot likes Tom" and "Tom does not like Spot"
  "Tom likes himself"
  "Kitty likes Spot" and "Spot likes Kitty"
  "Kitty likes herself"
  "Felix likes Kitty"

  *All these sentences are **true**, in the world above*

What is *true*?

- **A world of *cats***

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot	x		x	
Kitty		x	x	
Felix			x	

- **Sentences about this world**

  "Spot likes Felix"
  "Spot likes himself"
  "Kitty does not like herself"
  "Felix likes Spot"

  *All these sentences are **false**, in the world above*

What is *true*?

- **A world of *cats***

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot	x		x	
Kitty		x	x	
Felix			x	

- **Sentences about this world** (with *generalization*)

  "Every cat likes a cat" *is **true***
  "Every cat likes another cat" *is **false***
  "Tom does not like any other cat" *is **true***
  "Kitty is liked by every cat" *is **false***

  "Every cat that likes another cat also likes her/himself" *is **?***

How to make a sentence *true*?

- **Let's start from the other end: one sentence**

  "Kitty likes Spot" and "Spot likes Tom"

  It can be made true in many different ways

  Namely, it may be true in many different *possible worlds*

  Examples (*possible worlds* of cats):

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x			
Kitty		x		
Felix				

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot	x		x	
Kitty		x	x	
Felix			x	

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	x
Spot	x	x	x	x
Kitty	x	x	x	x
Felix	x	x	x	x

How to make a sentence *true*?

- **Consider the sentence**

  "Kitty likes Spot" and "Kitty does not like Spot"

  There is no way to *make it true*

  Namely, there is no *possible world* where this can be true

Sentences may be related

- **Three sentences**

  1) "Every cat that likes Kitty likes Spot as well"
  2) "Tom likes Kitty"

  ---

  3) "Tom likes Spot"

There is no way to make true sentences 1) *and* 2) *without making sentence* 3) *true as well…*

(just give it a try…)

This is just a small subset of ***all*** $2^{16} = 65{,}536$ *possible worlds …*

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x			
Kitty		x		
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x	x	x	
Kitty		x	x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot			x	
Kitty		x		
Felix				x

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x	x	x	x
Kitty			x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom				
Spot				
Kitty			x	
Felix				

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	
Spot	x			
Kitty		x		
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom		x	x	
Spot	x	x	x	
Kitty	x	x		
Felix				

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	
Spot	x		x	
Kitty		x		
Felix		x		x

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	x
Spot		x		
Kitty			x	
Felix				x

likes	Tom	Spot	Kitty	Felix
Tom			x	
Spot	x			
Kitty		x		
Felix		x	x	

Sentences may be related

- **Three sentences**

  1) "Every cat that likes Kitty likes Spot as well"
  2) "Tom likes Kitty"

  ---

  3) "Tom likes Spot"

There is no way to make true sentences 1) *and* 2) *without making sentence* 3) *true as well…*

(just give it a try…)

Sentence 1) is **true** in these worlds

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x			
Kitty		x		
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x	x	x	
Kitty		x	x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot			x	
Kitty		x		
Felix				x

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x	x	x	x
Kitty			x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom				
Spot				
Kitty			x	
Felix				

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	
Spot	x			
Kitty		x		
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom		x	x	
Spot	x	x	x	
Kitty	x	x		
Felix				

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	
Spot	x		x	
Kitty		x		
Felix		x		x

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	x
Spot		x		
Kitty			x	
Felix				x

likes	Tom	Spot	Kitty	Felix
Tom			x	
Spot	x			
Kitty		x		
Felix		x	x	

Sentences may be related

- **Three sentences**

  1) "Every cat that likes Kitty likes Spot as well"
  2) "Tom likes Kitty"

  ---

  3) "Tom likes Spot"

There is no way to make true sentences 1) *and* 2) *without making sentence* 3) *true as well…*

(just give it a try…)

Sentence 1) is **true** in these worlds

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x			
Kitty		x		
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x	x	x	
Kitty		x	x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom	x			
Spot			x	
Kitty		x		
Felix				x

likes	Tom	Spot	Kitty	Felix
Tom				
Spot	x	x	x	x
Kitty			x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom				
Spot				
Kitty			x	
Felix				

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	
Spot	x			
Kitty		x		
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom		x	x	
Spot	x	x	x	
Kitty		x	x	
Felix		x	x	

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	
Spot	x		x	
Kitty		x		
Felix		x		x

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	x
Spot		x		
Kitty			x	
Felix				x

likes	Tom	Spot	Kitty	Felix
Tom			x	
Spot	x			
Kitty		x		
Felix		x	x	

Sentence 2) is **true** in these worlds

Sentences may be related

- **Three sentences**

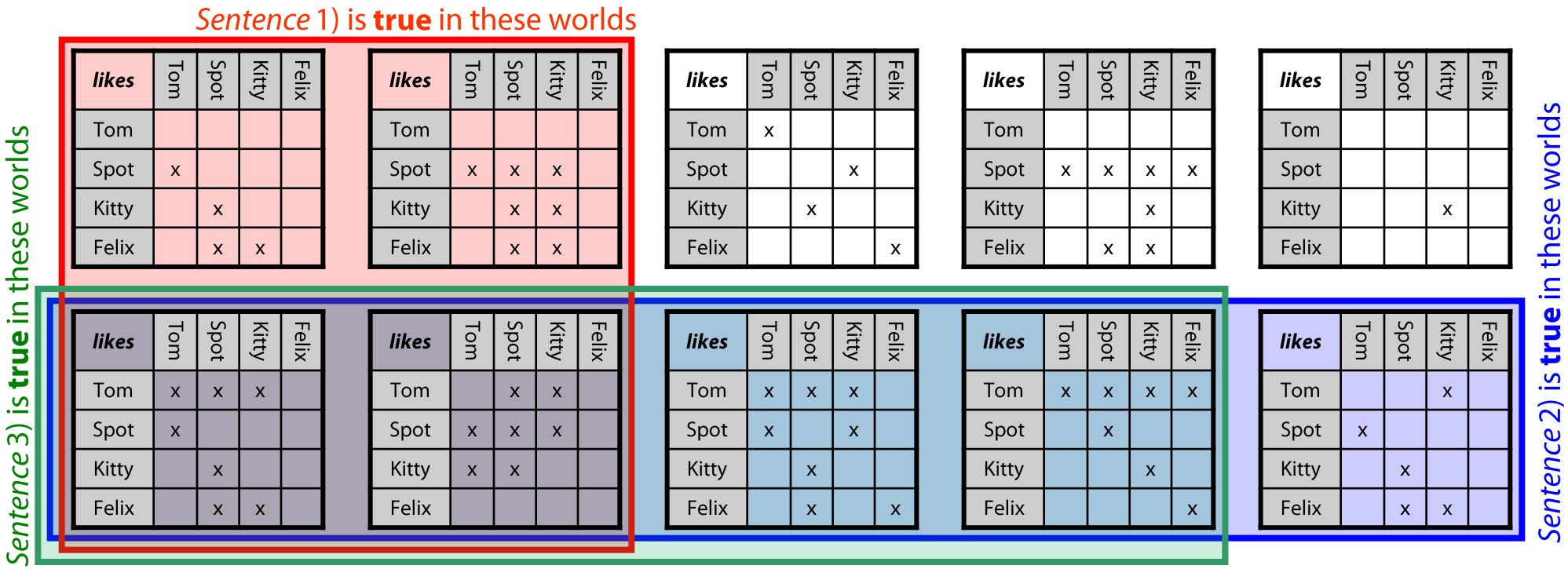
  1) "Every cat that likes Kitty likes Spot as well"
  2) "Tom likes Kitty"

  ---

  3) "Tom likes Spot"

There is no way to make true sentences 1) *and* 2) *without making sentence* 3) *true as well…*

(just give it a try)



Sentences may be related

- **Three sentences**

  1) "Every cat that likes Kitty likes Spot as well"
  2) "Tom likes Kitty"

  ---

  3) "Tom likes Spot"

  There is no way to make true sentences 1) *and* 2) *without making sentence* 3) *true as well…*

  (just give it a try)

Questions:

Was this just coincidence?

How many possible worlds must we consider, to be sure?

Sentences may be related

- **Three sentences**

  1) "Every cat that likes Kitty likes Spot as well"
  2) "Tom likes Kitty"

  ---

  3) "Tom likes Spot"

  There is no way to make true sentences 1) *and* 2) *without making sentence* 3) *true as well…*

  (just give it a try)

Questions:

Was this just coincidence?

How many possible worlds must we consider, be sure?

Hint:

All of them.