Artificial Intelligence

Symbolic representation and inference: the intuitive idea

Marco Piastra

An aside: equations

- 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

An aside: equations

- 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 (*i.e. we assume the truth of something*)

End point: a conclusion (*i.e. we state that something else must <u>also</u> be true*)

(*both points are decided by us*)

An aside: equations

- 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 (*i.e. we assume the truth of something*)

End point: a conclusion (*i.e. we state that something else must <u>also</u> be true*)

(*both points are decided by us*)

An aside: equations

- 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 (*i.e. we assume the truth of something*)

End point: a conclusion (*i.e. we state that something else must <u>also</u> be true*)

(*both points are decided by us*)

Sequence of steps

Each transformation rule is based on *equivalence*

Each step must be *correct*, in this particular sense

An aside: equations

- 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 (*i.e. we assume the truth of something*)

End point: a conclusion (*i.e. we state that something else must <u>also</u> be true*)

(*both points are decided by us*)

Sequence of steps

Each transformation rule is based on *equivalence*

Each step must be *correct*, in this particular sense

Abstraction and *Correctness*

Symbols like x, a and b represent any number, *equivalences* are valid for *all* of them

What does it mean for the <u>whole sequence</u> to be *correct*?

Symbolic Reasoning

Symbolic descriptions can *abstract*

- Symbolic descriptions have *abstraction* capabilities:
  for instance, many linguistic phenomena are *systemic*
  (i.e. their complexity goes beyond a *pattern-matching*)

  Someone may understand English or not.

  But no human being can understand the sentence:
    *"Sally **likes** Harry"*

  without being able 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])

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: *syllogism*

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 relevant:
in this case the schema does <u>not</u> work

Reasoning and schemas: *syllogism*

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 ordering of sentences IS relevant:
in this case the schema does <u>not</u> work

Reasoning and schemas: *syllogism*

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 (*princesses*)
All (*princesses*) are (*young and beautiful*)   hence
_____
All (*enchanted frogs*) are (*young and beautiful*)

*Same schema,
different impression*

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 (*princesses*)

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

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

What do we mean by "all"? Do we need 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 structure

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** (i.e. 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?*

- **Consider the sentence**

  "Kitty likes Spot" and "Spot likes Tom"

  It can be made true in many different ways

  "It may be true in many different *possible worlds*"

  Examples (i.e. of other *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 <u>not</u> like Spot"

    There is no way to *make it true*
    "There is no *possible world* where this can be true"

- **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 true sentence* 3) *as well…*
  (just give it a try…)

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	
Spot	x			
Kitty		x		
Felix		x		x

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

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	x
Spot		x		
Kitty			x	
Felix				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 true sentence 3) as well…

    (just give it a try…)

Sentence 1) is true in these worlds

Row 1

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				

Row 2

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	
Spot	x			
Kitty		x		
Felix		x		x

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

likes	Tom	Spot	Kitty	Felix
Tom	x	x	x	x
Spot		x		
Kitty			x	
Felix				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 true sentence* 3) *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	
Spot	x			
Kitty		x		
Felix		x		x

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

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

Sentence 2) is true in these worlds

- **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 true sentence* 3) *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				

3) is true in these worlds

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			
Kitty		x		
Felix		x		x

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

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

Sentence

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 true sentence* 3) *as well…*
      (just give it a try)

Questions:
Was this just coincidence?
How many possible worlds must we consider in order to be sure?