Artificial Intelligence

Symbolic representation and inference: the intuitive idea

Marco Piastra

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

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)

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)

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

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 whole sequence 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**

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 not work

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

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



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

	T	مکم	o t Kit	tyFel	i)
likes	101			دي. د.	
Tom	Х				
Spot	X		Х		
Kitty		х	х		
Felix			Х		

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

	-To ı	mSp	o K it	tvEol	iχ
likes	101			٠,٠ ٠.	
Tom	X				
Spot	X		х		
Kitty		Х	х		
Felix			Х		

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):

	Ta	mSp	<u> ⊲t⁄it</u>	tvEal	iv
likes	101		Jene	cyi Ci	
Tom					
Spot	X				
Kitty		X			
Felix					

	To	mSp	o t Kit	tvEol	ix
likes	101			٠,٠٠٠	
Tom	Х				
Spot	Х		х		
Kitty		х	х		
Felix			Х		

	<u>_</u> T_	mSp	∆t Kit	tyEo	iy
likes	101			cyi Ci	
Tom	Х	Х	Х	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*"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	TOI	n Spo	rtKiti	yFei
Tom				
Spot	х			
Kitty		х		
Felix		х	х	

	Tor	п Зро	otKit	v Fel	İx
likes				,	
Tom					
Spot	х	х	х		
Kitty		х	х		
Felix		х	х		

likes	Toi	пЅро	otKiti	y Fel
Tom	х			
Spot			х	
Kitty		х		
Felix				х

likes	Toi	пЗро	ikiti	yFei
Tom				
Spot	х	х	х	х
Kitty			х	
Felix		х	х	

	likes	Toi	n Sp o	rtKiti	yFel
Ì	Tom				
Ī	Spot				
I	Kitty			х	
	Felix				

likes	Toi	n S po	tΚiτι	yFeli
Tom	х	х	х	
Spot	х			
Kitty		х		
Felix		х	х	

	Tou	n Spo	ni Kir	WEE	j)
likes	101	ПЭР	PCIXICI	y r Ci	ľ
Tom		х	х		
Spot	х	х	х		
Kitty	х	х			
Felix					

	Tor	пЗро	ni Kir	vEel	İ×
likes	101	ПЭР	Perkit	,y	ĺ
Tom	х		х		
Spot	х				
Kitty		х			
Felix		х		х	

likes	Toi	пЗро	νtΚiτι	yFel	į>
Tom			х		l
Spot	х				
Kitty		х			
Felix		х	х		

likes	Tol	пЅро	ρίΚΙτι	yFel
Tom	х	х	х	х
Spot		х		
Kitty			х	
Felix				х

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

1	Tom Speckitty Feli						
	likes						
	Tom						
	Spot	х					
	Kitty		х				
	Felix		х	х		l	

	Tor	n Spo	otKitt	yFel	į
likes		,		,	
Tom					
Spot	х	х	х		
Kitty		х	х		
Felix		х	х		l

likes	101	пэрс	JUNIU	y Fel	X
Tom	х				
Spot			х		
Kitty		х			
Felix				х	

likes	Toi	пSр	otKiti	yFel
Tom				
Spot	х	х	х	х
Kitty			х	
Felix		х	х	

likes	Toi	пSр	otKiti	yFel
Tom				
Spot				
Kitty			х	
Felix				

likes	Tol	пЅро	ιKiti	yFel	×
Tom	х	х	х		
Spot	х				
Kitty		х			
Felix		х	х		

	Tor	п Ѕрс	νίΚίτι	yFeli
likes				
Tom		х	х	
Spot	х	х	х	
Kitty	х	х		
Felix				

likes	Toi	n Sp o	iKlu	yFel
Tom	х		х	
Spot	х			
Kitty		х		
Felix		х		х

likes	Toi	пЗро	νtΚiτι	yFel	į>
Tom			х		l
Spot	х				
Kitty		х			
Felix		х	х		

likes	Toi	n Spo	λίΚΙτι	yFel
Tom	х	х	х	х
Spot		х		
Kitty			х	
Felix				х

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

	Tor	п Ѕрс	ριΚίτι	yFel	İX
likes					
Tom					
Spot	х				
Kitty		х			
Felix		х	х		

likes	Toi	пЅрс	otKiti	y Feli x
Tom				
Spot	х	х	х	
Kitty		х	х	
Felix		х	х	

likes	Ťoi	пЅро	otKiti	y Fel
Tom	х			
Spot			х	
Kitty		х		
Felix				х

1		Tor	пSр	πKit	vFeli
	likes				,
	Tom				
	Spot	х	х	х	х
	Kitty			х	
	Felix		х	х	

	_			
likes	101	пэрс	JUNIU	уге
Tom				
Spot				
Kitty			х	
Felix				

likes	Toi	n Sp o	iKiti	yFel
Tom	х	х	х	
Spot	х			
Kitty		х		
Felix		Х	Х	

likes	Toi	n Spo	ιΚία	yFeli
Tom		х	х	
Spot	х	х	х	
Kitty	х	х		
Felix				

1		Total	nSpo	oi Kii	v Fel	İ١
	likes	101	ПЭР	Perkit	,y	ľ
	Tom	х		х		
	Spot	х				
	Kitty		х			
	Felix		х		х	

1			- S-1	at Kita	w Eur	İX
	likes	101	пэрс	CINIC	yr Ci	Î
	Tom			х		
	Spot	х				
	Kitty		х			
	Felix		х	х		

		Sin	HK1	vFel
likes	5	napo	CIXICI	yr Ci
Tom	х	х	х	х
Spot		х		
Kitty			х	
Felix				х

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

		Toi	пЗро	ptKit	yFe	X		Toi	пSр	ριΚίτι	y Feli	×		Toi	пЗро	τKit	yFel	x		Tor	пбр	φιΚίτ	yFel	X		Toi	пбро	ριΚίτ	yFeli
	likes					l	likes						likes						likes						likes				
	Tom					1	Tom						Tom	х					Tom					Ì	Tom				
	Spot	х				1	Spot	х	х	х			Spot			Х			Spot	х	х	х	х		Spot				
	Kitty		х				Kitty		х	х			Kitty		х				Kitty			х			Kitty			х	
	Felix		х	х			Felix		х	х			Felix				х		Felix		х	х			Felix				
						_					_																		
	likes	Tor	n Spo	νίΚΙι	iy Fel	X	likes	Tor	n Spo	τΚίτι	y Feli x	×	likes	Tor	n Spo	ιKlu	y Fel	İX	likes	Tor	n Spi	νίΚΙι	yFel	х	likes	Tor	n Spo	ρtΚiτ	yFeli:
	Tom	х	х	х		1	Tom		х	х			Tom	х	х	х			Tom			х		l	Tom	х	х	х	х
3) is	three i	n×tŀ	ies	e w	orl	ds	Spot	х	х	х			Spot	х					Spot	х				ĺ	Spot		х		
	Kitty		х				Kitty	х	х				Kitty		х				Kitty		х				Kitty			х	
	Felix		х	х			Felix						Felix		х		х		Felix		х	х			Felix				х
ser	tence											_									Se	ntei	าce	2) i	s true i	n th	nes	e w	orld

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)

Ouestions:

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