Setting the Stage

Are Humans Rational?

Selmer Bringsjord

September 3 2019

Seats available?;
This is your seat — almost;
reviewing some names.

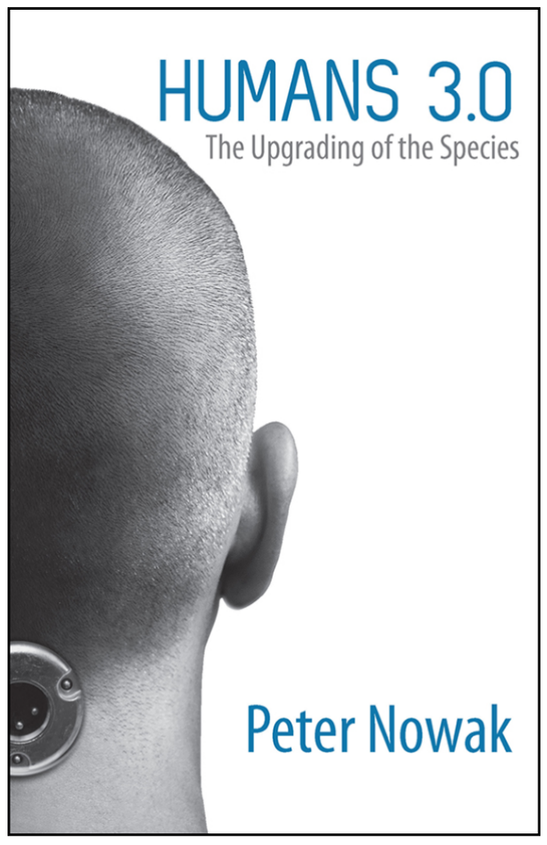
http://kryten.mm.rpi.edu/COURSES/AHR/ahr.html

New version of syllabus now up.

http://kryten.mm.rpi.edu/COURSES/AHR/ahr.html

New version of syllabus now up.

Re. *Humans 3.0: The Upgrading of the Species*
Peter Nowak will speak on September 23, at 7pm, in Sage 3303.



The Balderdash that is *Humans 3.0: The Upgrading of the Species*

Selmer Bringsjord
Sept 25 2019, 12 noon
RPI: Sage 4101
(lunch provided)

Here's an accurate encapsulation, put declaratively, of the book (H3.0) in question:

As a matter of mathematics, religious belief will disappear. Work will be obsolete, but economic well-being will be maximally high across Earth's human population; this will be enabled by AI toiling for us. Science will explain everything, including discovering the "patterns" that are us. With these patterns in our hands, we will be able to repeatedly "upload" to the physical substrate of our choosing, and thereby live forever. Then, by 2045, The Singularity will occur, the moment in time when machine intelligence exceeds human intelligence, and immediately thereafter explodes to higher and higher levels that infinitely exceed our own (relatively speaking) rodent-level one. Conveniently, we will merge with the machines so as to dodge being destroyed by them, and this "hybrid human-machine intelligence" will busy itself with [yada yada yada].

Unfortunately for Nowak (2015), author of H3.0, there is a slight problem: viz., every single claim here is but balderdash, at best. In this talk, I patiently explain this diagnosis, one bound, I know, to be emotionally disturbing to those who take such claptrap seriously.

The Balderdash that is *Humans 3.0: The Upgrading of the Species*

Selmer Bringsjord
Nov 11 2019, 7pm
RPI; Room: Sage 3303
public invited

Here's an accurate encapsulation, put declaratively, of the book (H3.0) in question:

As a matter of mathematics, religious belief will disappear. Work will be obsolete, but economic well-being will be maximally high across Earth's human population; this will be enabled by AI toiling for us. Science will explain everything, including discovering the "patterns" that are us. With these patterns in our hands, we will be able to repeatedly "upload" to the physical substrate of our choosing, and thereby live forever. Then, by 2045, The Singularity will occur, the moment in time when machine intelligence exceeds human intelligence, and immediately thereafter explodes to higher and higher levels that infinitely exceed our own (relatively speaking) rodent-level one. Conveniently, we will merge with the machines so as to dodge being destroyed by them, and this "hybrid human-machine intelligence" will busy itself with [yada yada yada].

Unfortunately for Nowak (2015), author of H3.0, there is a slight problem: viz., every single claim here is but balderdash, at best. In this talk, I patiently explain this diagnosis, one bound, I know, to be emotionally disturbing to those who take such claptrap seriously.

Any questions/comments re syllabus & course mechanics?

Main Claim

Main Claim

$\mathcal{R}$ Humans, at least neurobiologically normal ones, are fundamentally rational, where rationality is constituted by certain logico-mathematically based reasoning and decision-making in response to real-world stimuli, including stimuli given in the form of focused tests; but mere animals are *not* fundamentally rational, since, *contra* Darwin,[1] their minds are provably, fundamentally, qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is also "No." For starters, if x can't read, write, and create, x can't be (presently[2]) rational; neither computing machines/robots nor non-human animals can read nor write nor create; ergo, they aren't fundamentally rational for this reason alone. But the news for non-human animals and computing machines/robots gets much worse, for they have not the slightest chance when they are measured against $\mathcal{H}$.

From Syllabus (Last Time)

From Syllabus (Last Time)

Amtrak-to-Princeton J-L Problem

Suppose that the following two statements are true:

(1) Everyone likes anyone who likes someone.

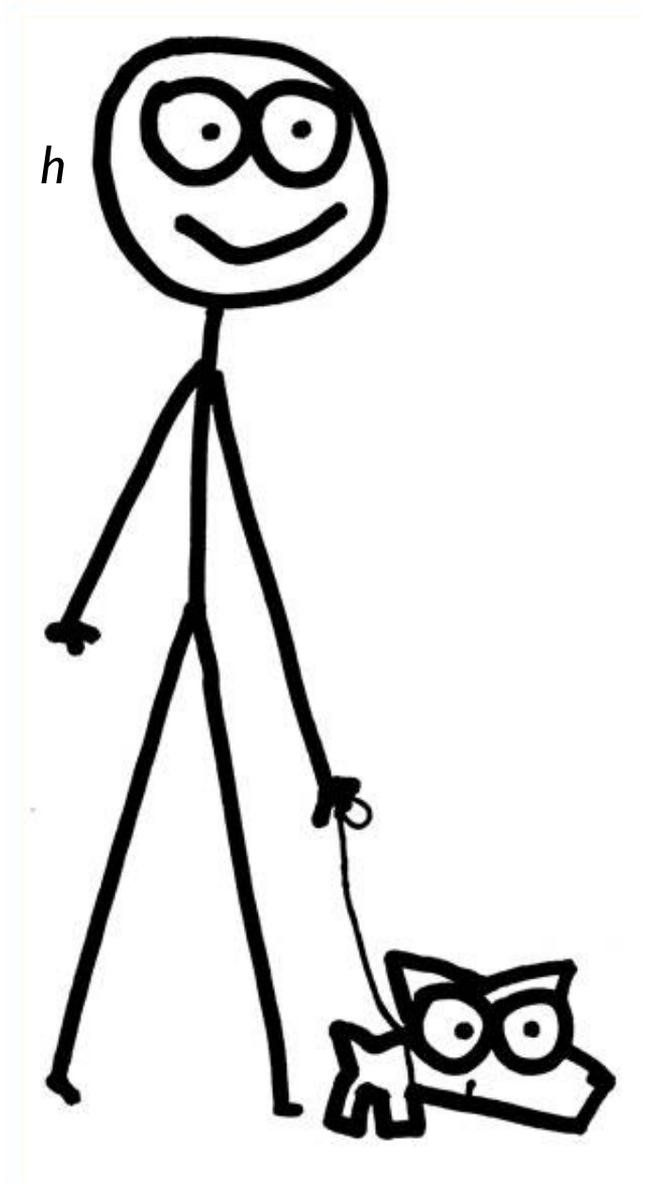
(2) Abigail likes Bruno.

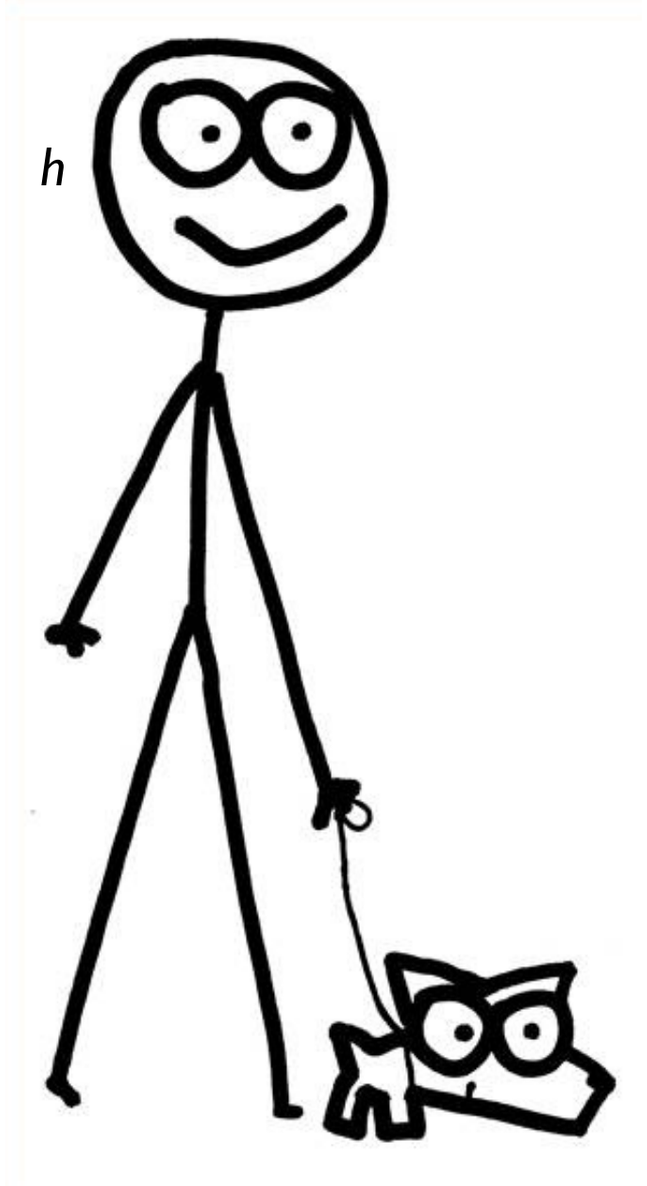
Does it follow deductively that everyone likes Bruno? Prove that your answer is right!

And Supporting Main Claim …

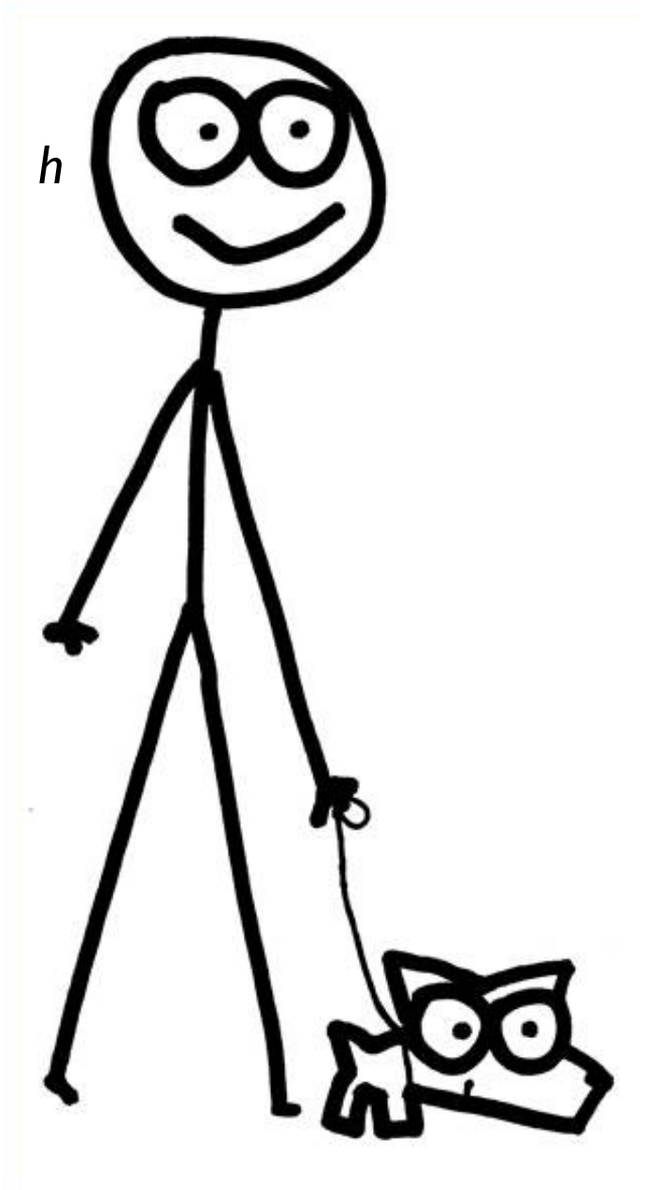
And Supporting Main Claim …

$\mathcal{H}$ Humans have the ability to gain knowledge by reasoning (e.g., deductively) quantificationally and recursively over abstract concepts, including abstract concepts of a highly expressive, including infinitary, nature, expressed in arbitrarily complex natural and formal languages.


h

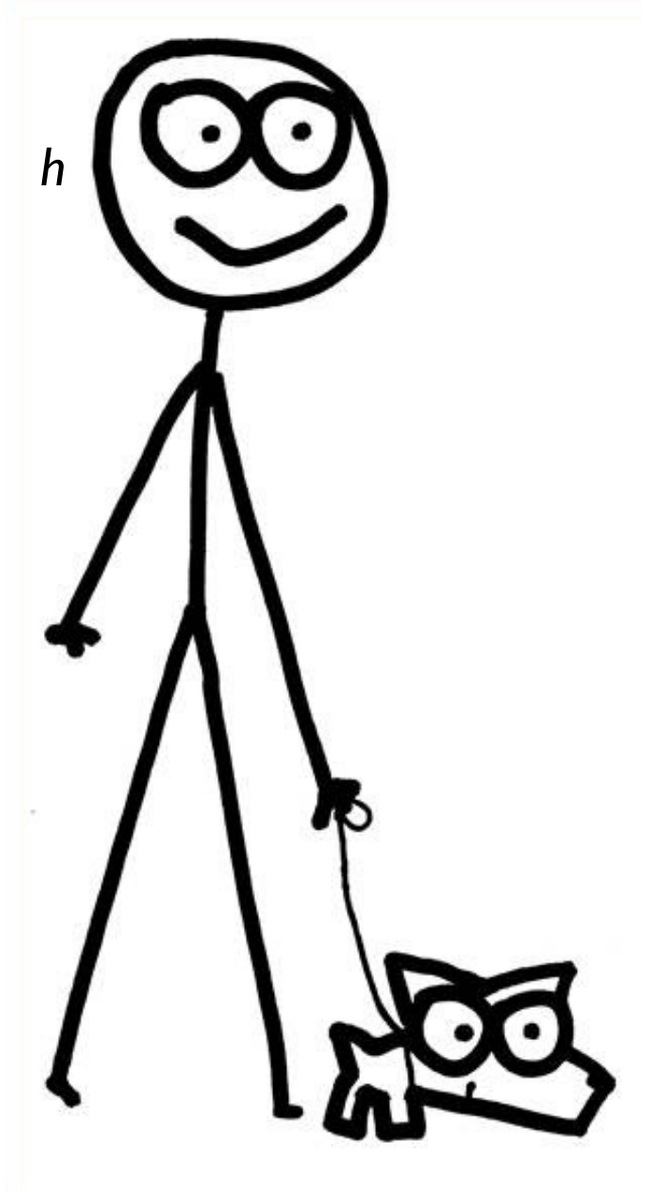


Problem!



Problem!

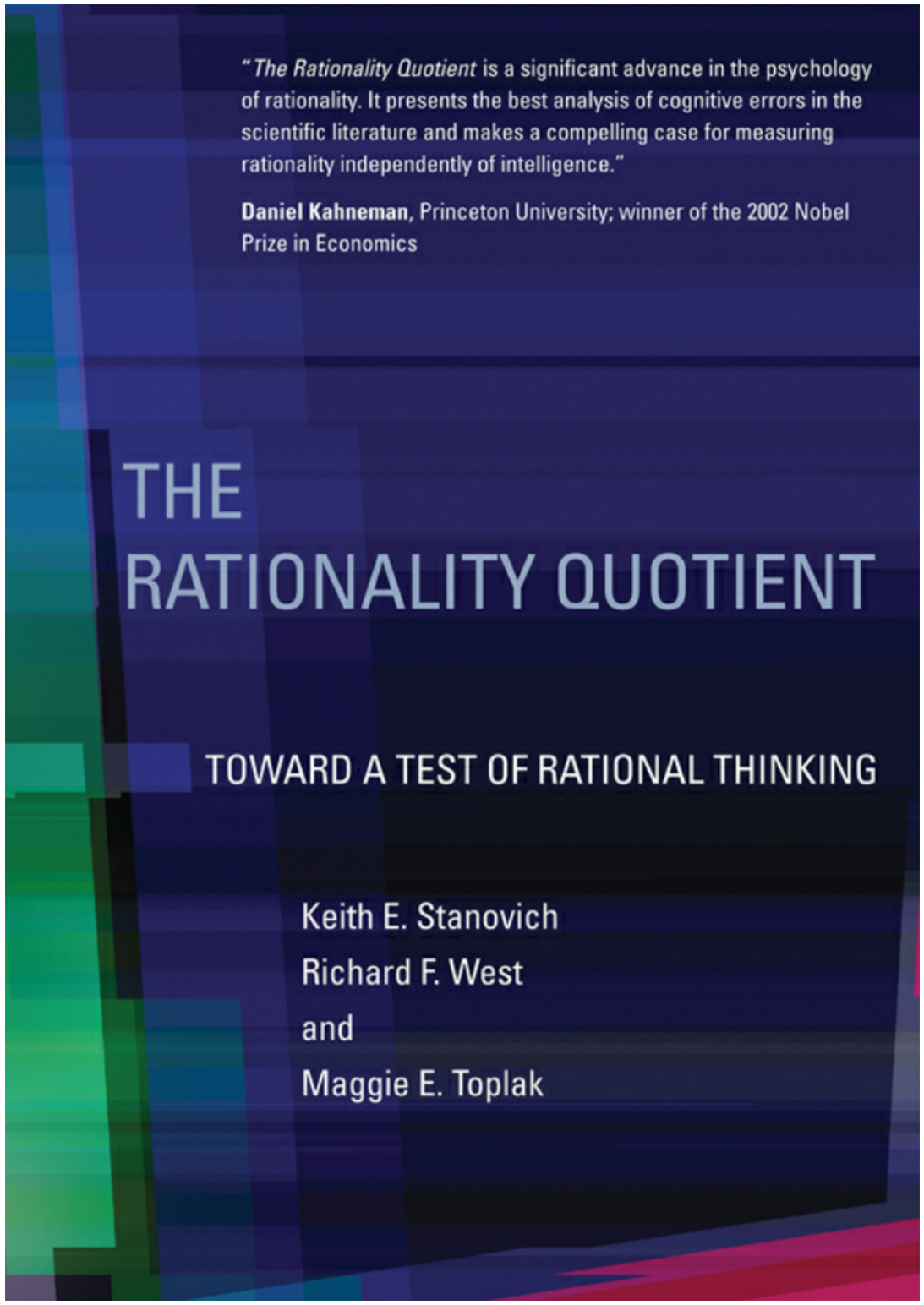
Or test. For an overview of Psychometric AI, see:



Problem!

Or test. For an overview of Psychometric AI, see:

http://www.tandfonline.com/doi/pdf/10.1080/0952813X.2010.502314


"The Rationality Quotient is a significant advance in the psychology of rationality. It presents the best analysis of cognitive errors in the scientific literature and makes a compelling case for measuring rationality independently of intelligence."
Daniel Kahneman, Princeton University; winner of the 2002 Nobel Prize in Economics
THE RATIONALITY QUOTIENT
TOWARD A TEST OF RATIONAL THINKING
Keith E. Stanovich
Richard F. West
and
Maggie E. Toplak

Sample Item from CART

Sample Item from CART

Premises:

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Premise 2: All tulips are carbitops.

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Premise 2: All tulips are carbitops.

Conclusion:

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Premise 2: All tulips are carbitops.

Conclusion:

All tulips are flowers.

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Premise 2: All tulips are carbitops.

Conclusion:

All tulips are flowers.

Assuming that the two premises are true, does the conclusion follow deductively from them?

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Premise 2: All tulips are carbitops.

Conclusion:

All tulips are flowers.

Assuming that the two premises are true, does the conclusion follow deductively from them?

$$\{\forall x(Fx \to Cx), \forall x(Tx \to Cx)\} \nvdash \forall x(Tx \to Fx)$$

Sample Item from CART

Premises:

Premise 1: All flowers are carbitops.

Premise 2: All tulips are carbitops.

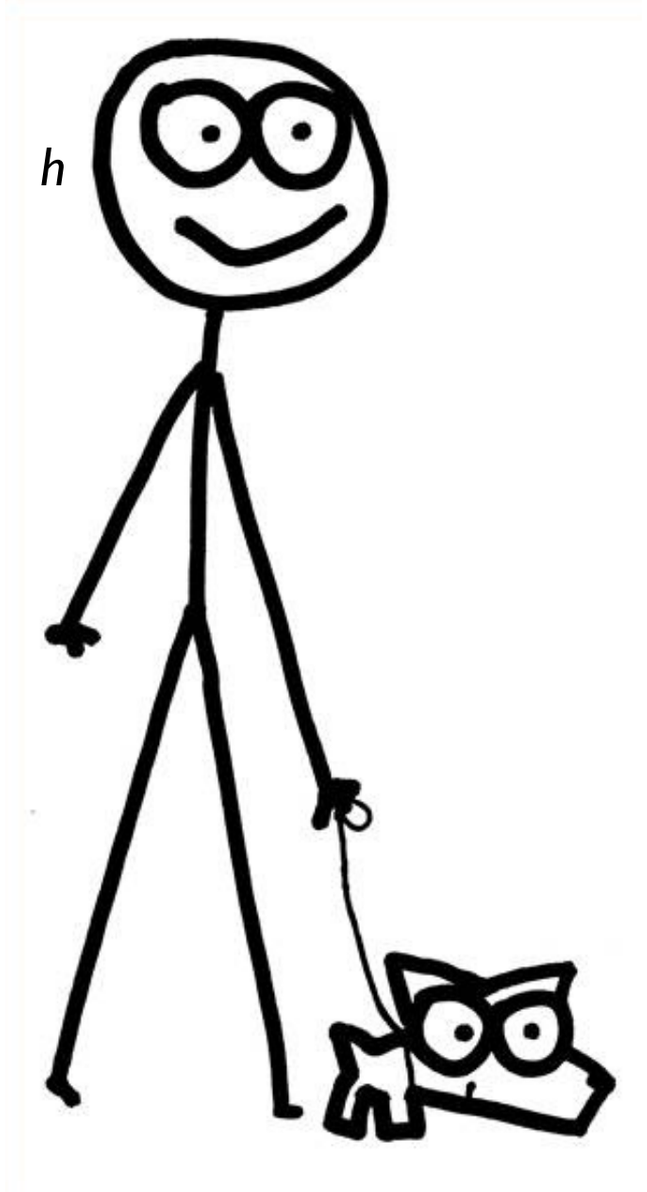
Conclusion:

All tulips are flowers.

Assuming that the two premises are true, does the conclusion follow deductively from them?[13]

$$\{\forall x(Fx \rightarrow Cx), \forall x(Tx \rightarrow Cx)\} \nvdash \forall x(Tx \rightarrow Fx)$$

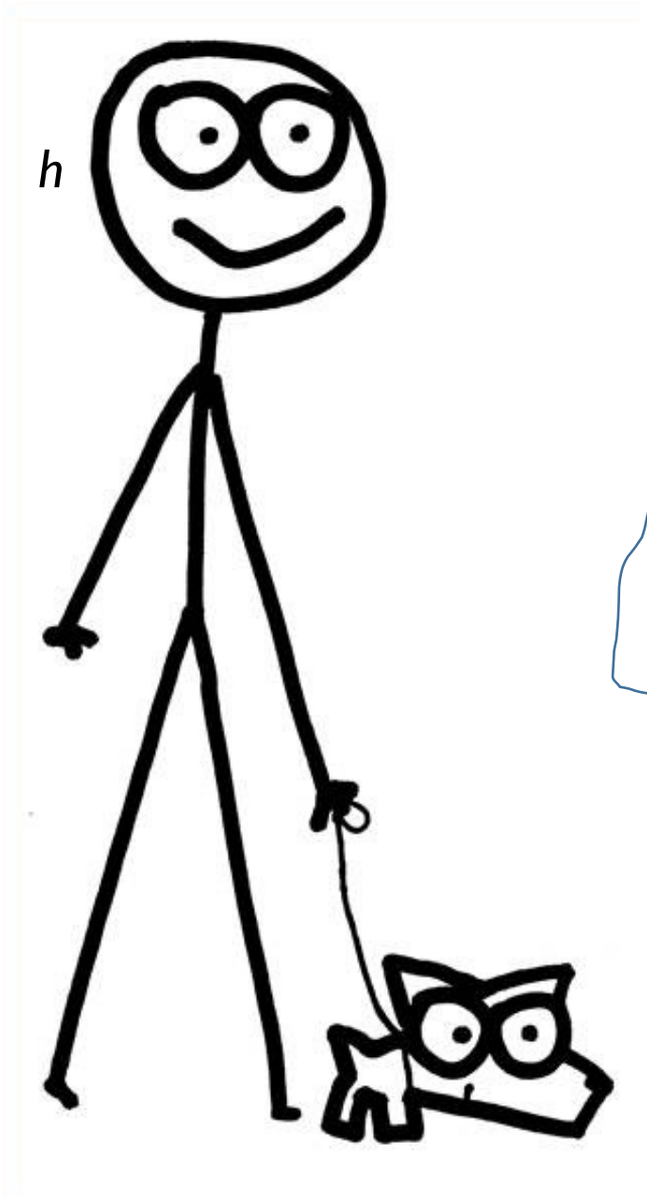
"The Problems that Generate the Rationality Debate are Too Easy, Given What Our Economy Now Demands"

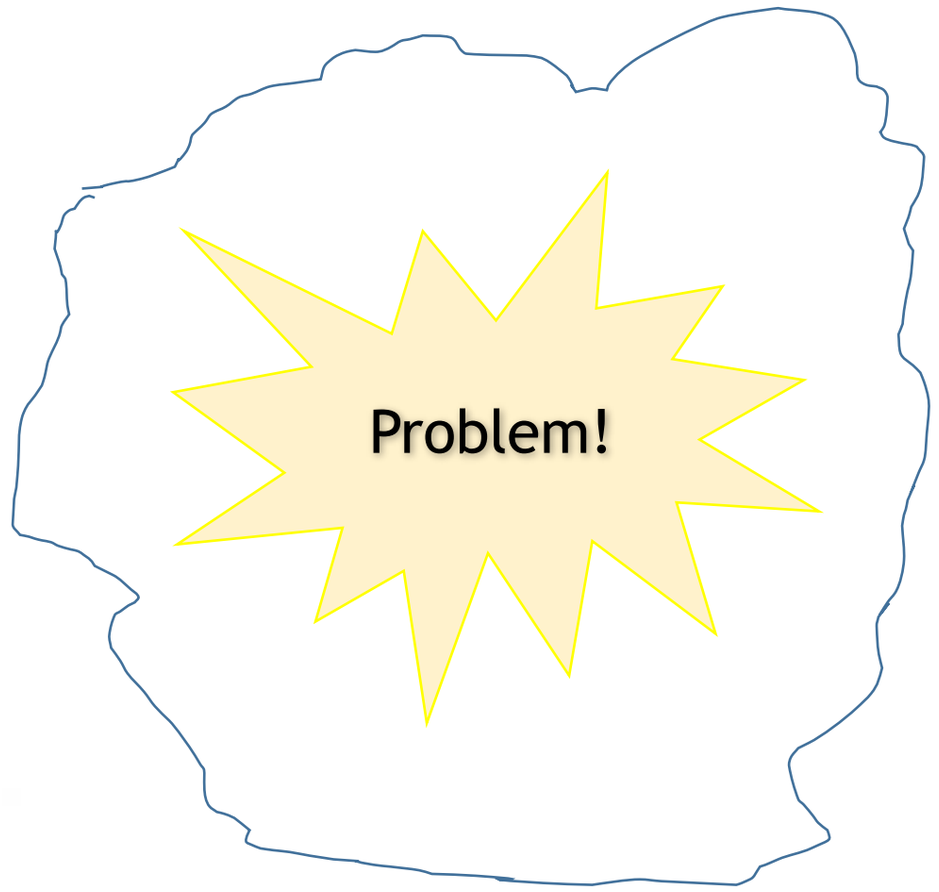


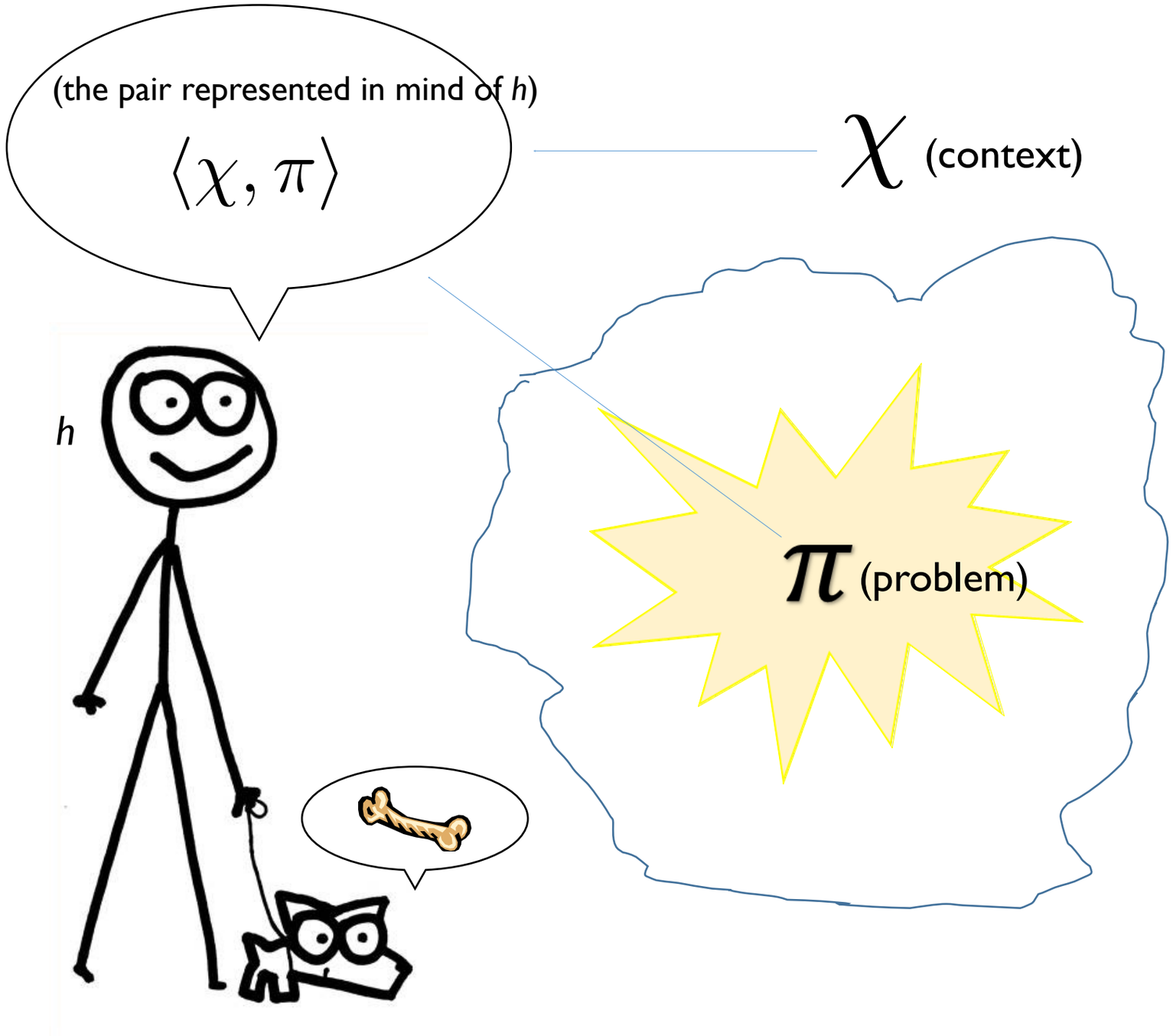
Problem!

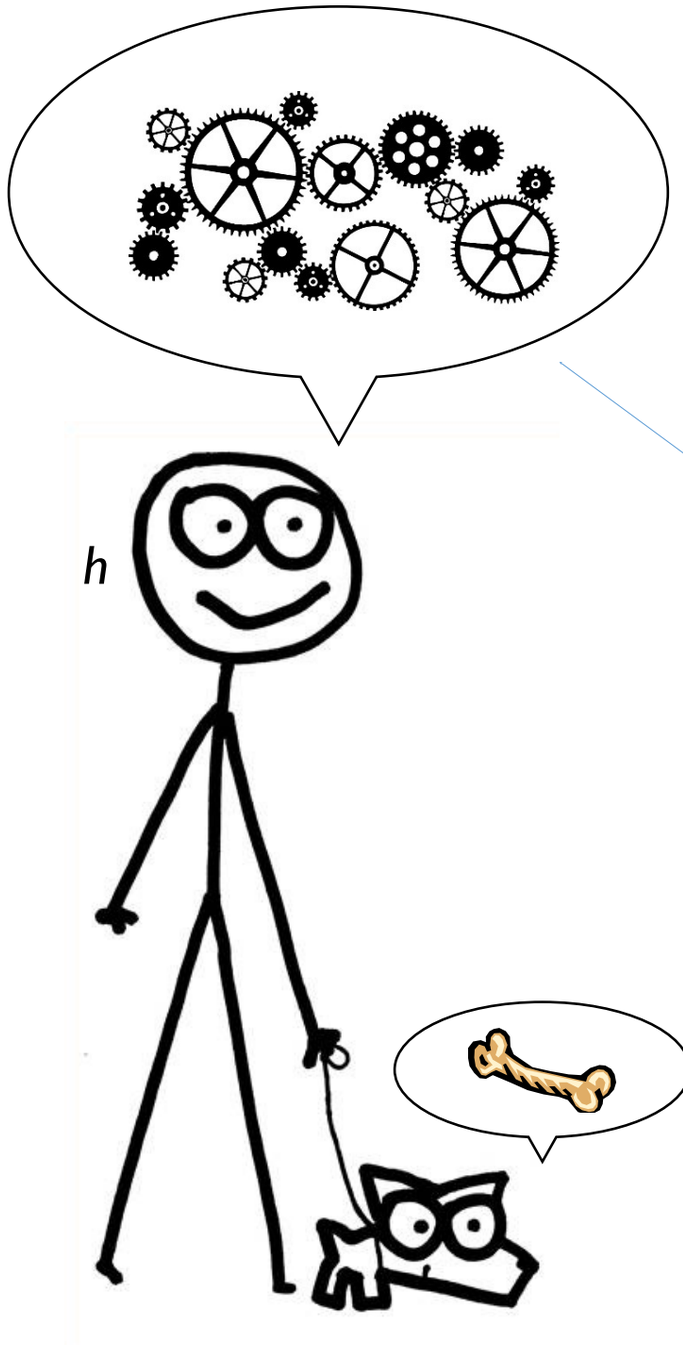
Or test. For an overview of Psychometric AI, see:

http://www.tandfonline.com/doi/pdf/10.1080/0952813X.2010.502314


h

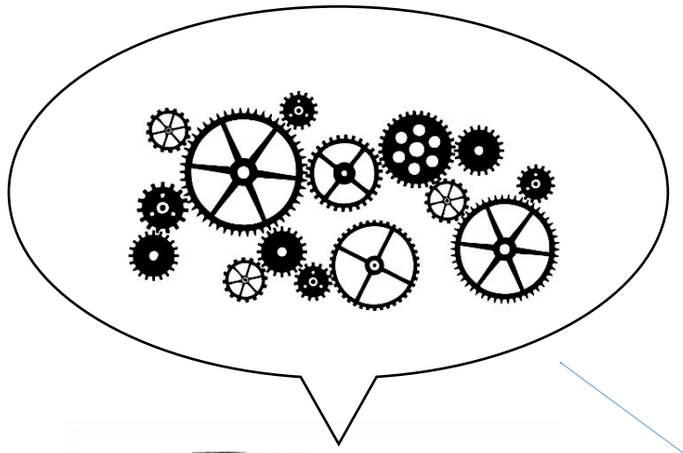

Problem!

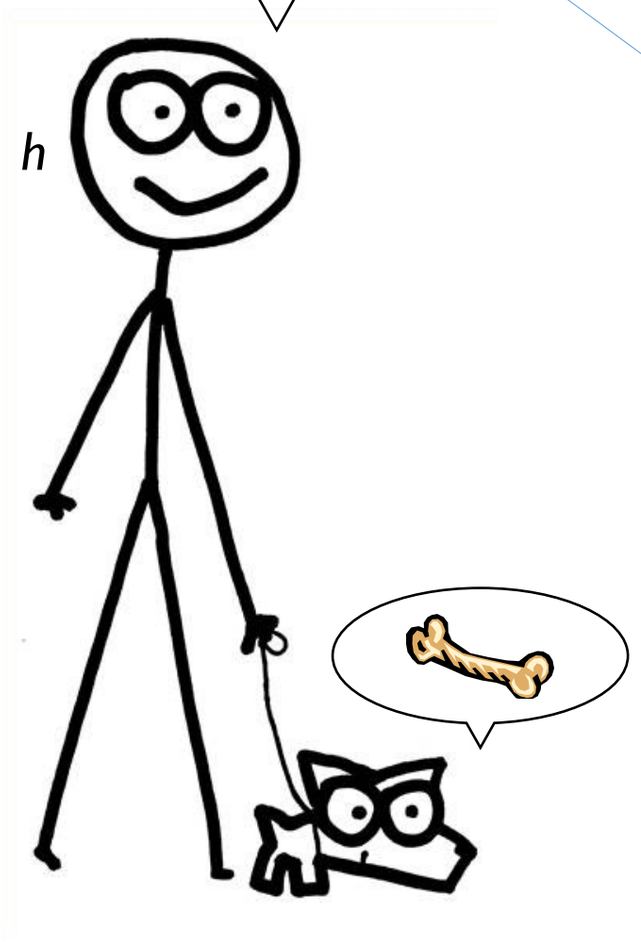

(the pair represented in mind of *h*)
$\langle \chi, \pi \rangle$
χ (context)
h
π (problem)



$\langle \chi, \pi \rangle$

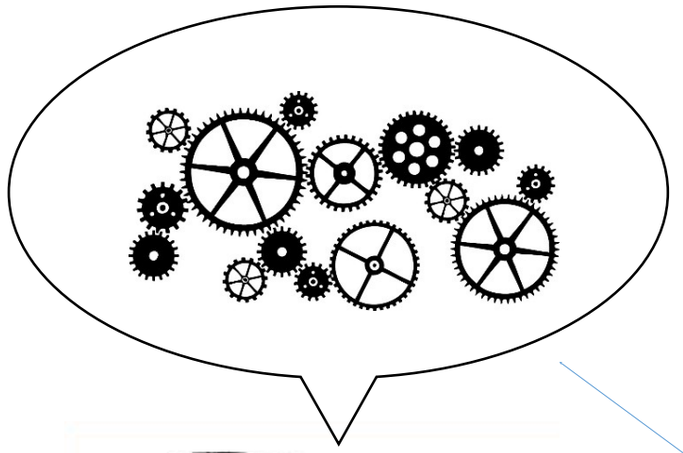
Formalism $_1$
Formalism $_2$
Formalism $_3$
…

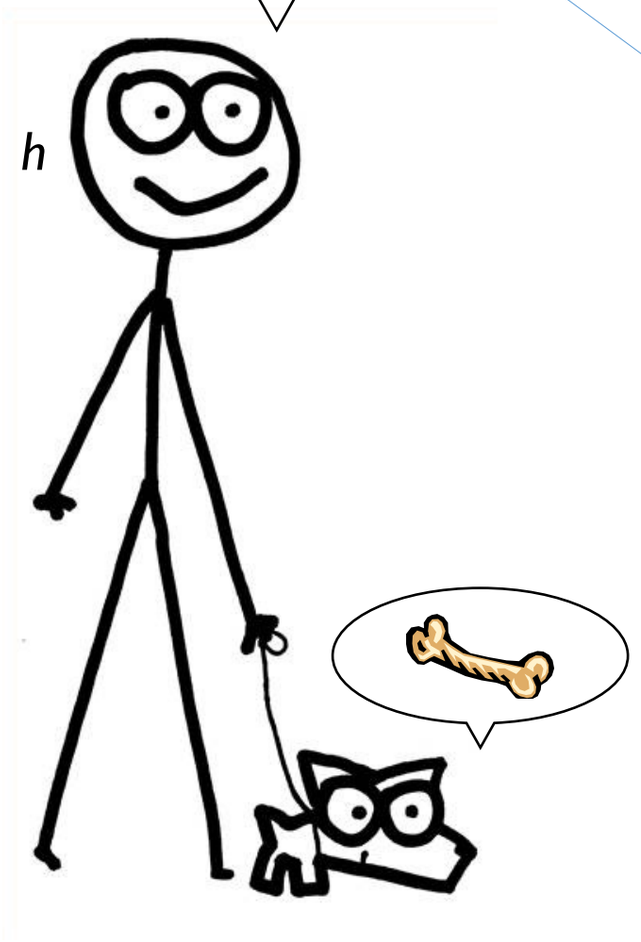




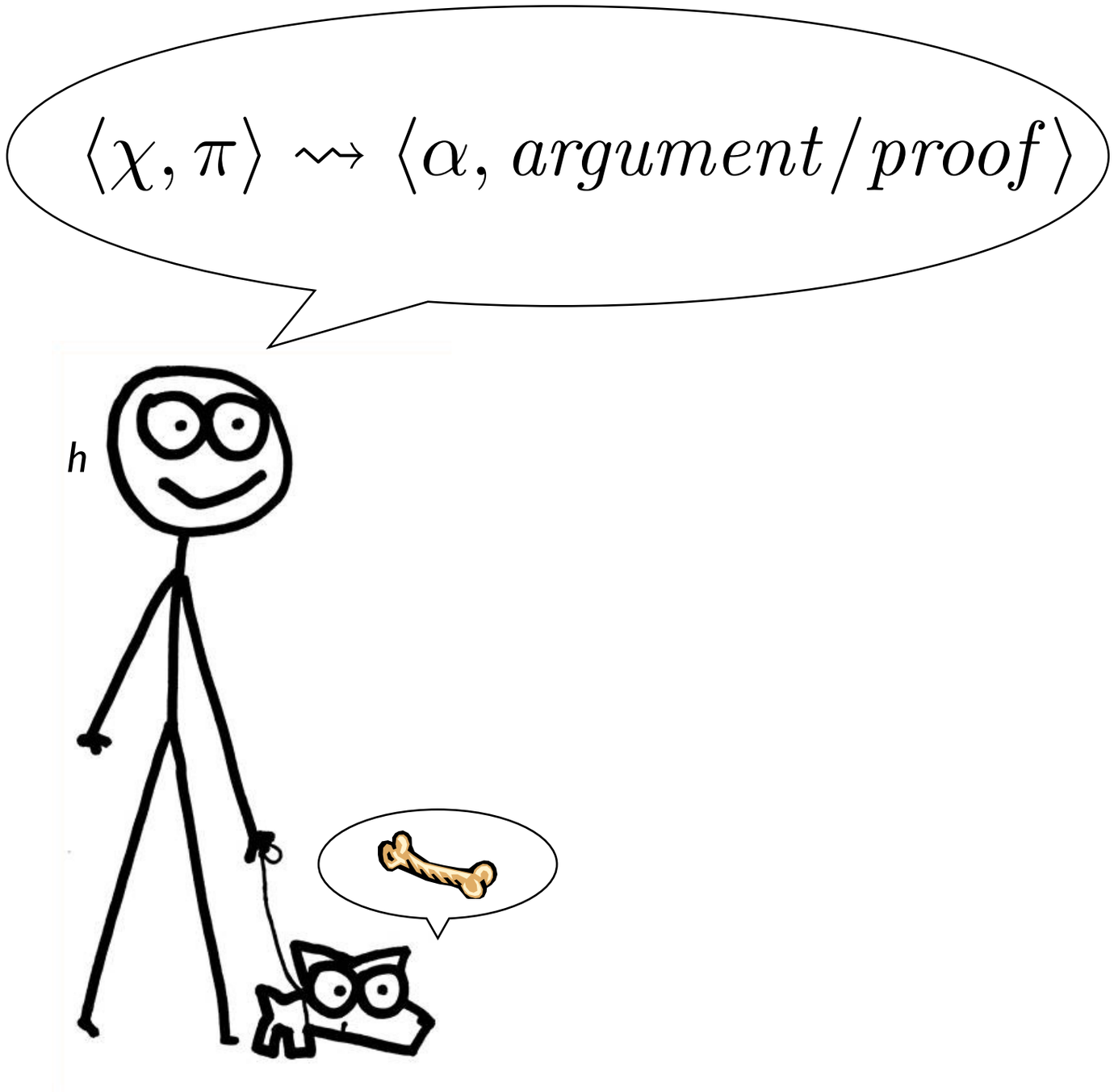
$\langle \chi, \pi \rangle$

Formalism $_1$
Formalism $_2$
Formalism $_3$
…

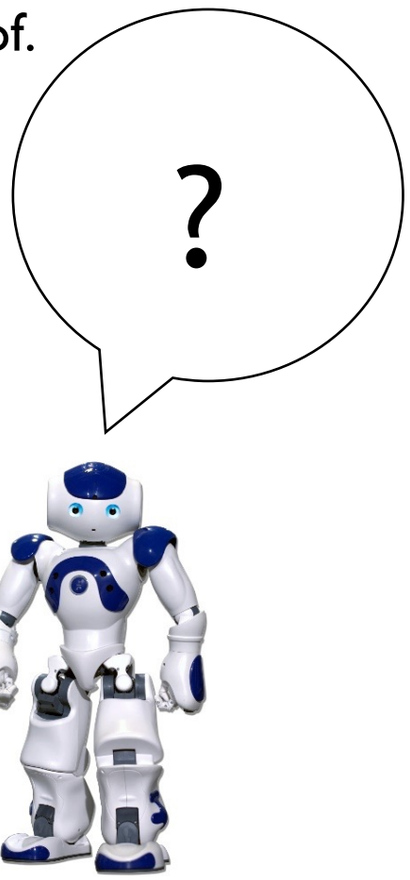


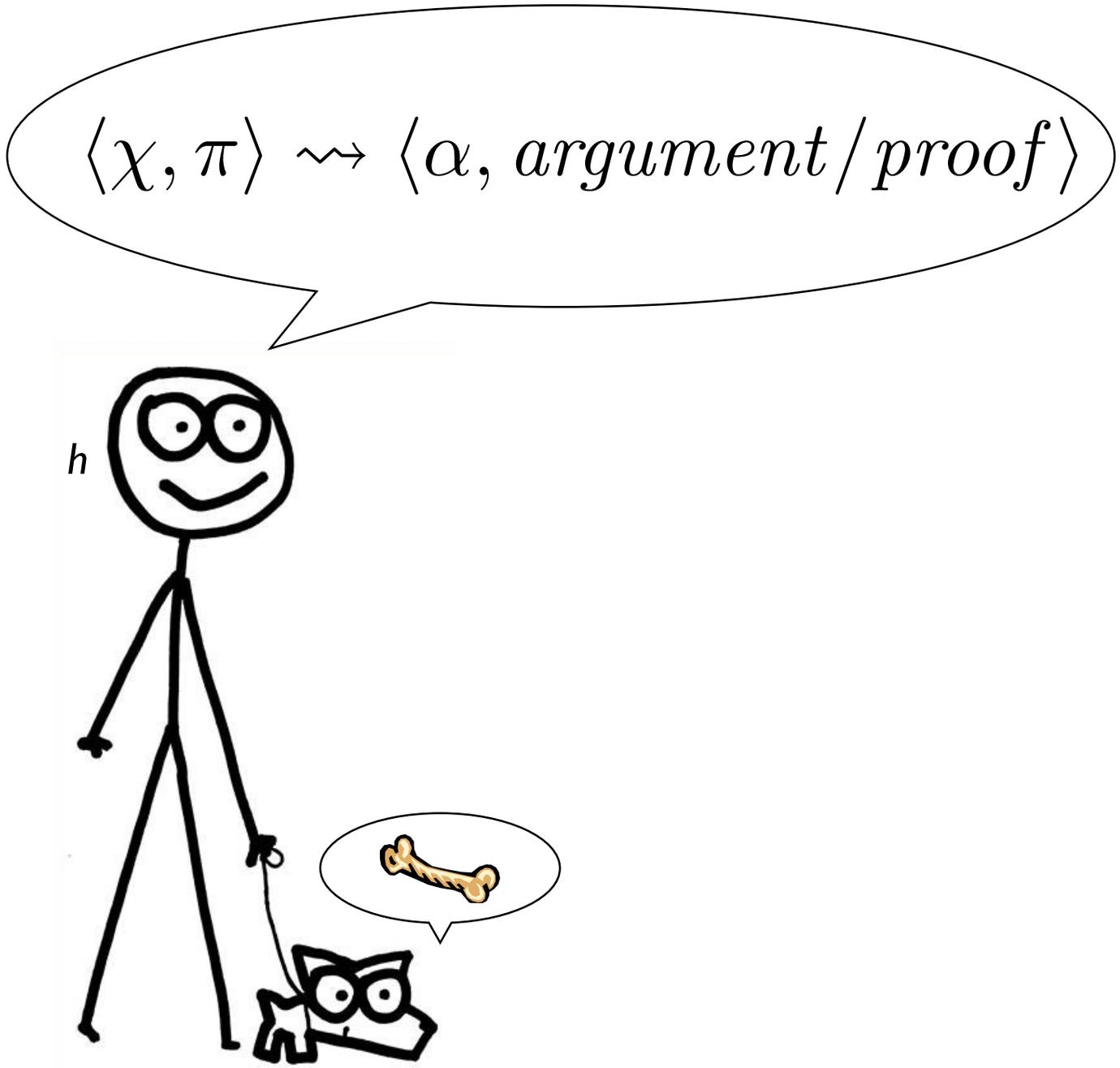


$\langle \chi, \pi \rangle$


$\langle \chi, \pi \rangle \rightsquigarrow \langle \alpha, argument/proof \rangle$
h

Today's (statistical/ connectionist) machine-learning systems are fundamentally incapable of providing the argument/proof.





https://www.darpa.mil/program/explainable-artificial-intelligence

Main Claim

Main Claim

$\mathcal{R}$ Humans, at least neurobiologically normal ones, are fundamentally rational, where rationality is constituted by certain logico-mathematically based reasoning and decision-making in response to real-world stimuli, including stimuli given in the form of focused tests; but mere animals are *not* fundamentally rational, since, *contra* Darwin,[1] their minds are provably, fundamentally, qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is also "No." For starters, if x can't read, write, and create, x can't be (presently[2]) rational; neither computing machines/robots nor non-human animals can read nor write nor create; ergo, they aren't fundamentally rational for this reason alone. But the news for non-human animals and computing machines/robots gets much worse, for they have not the slightest chance when they are measured against $\mathcal{H}$.

Main Claim

$\mathcal{R}$ Humans, at least neurobiologically normal ones, are fundamentally rational, where rationality is constituted by certain logico-mathematically based reasoning and decision-making in response to real-world stimuli, including stimuli given in the form of focused tests; but mere animals are *not* fundamentally rational, since, *contra* Darwin,[1] their minds are provably, fundamentally, qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is also "No." For starters, if x can't read, write, and create, x can't be (presently[2]) rational; neither computing machines/robots nor non-human animals can read nor write nor create; ergo, they aren't fundamentally rational for this reason alone. But the news for non-human animals and computing machines/robots gets much worse, for they have not the slightest chance when they are measured against $\mathcal{H}$.

Definition of Perfect Present Rationality

Let T be the relevant set of tests. Then: An agent a is **perfectly presently rational** if and only if (i) a answers all questions on all $t \in T$, and (ii) provides a sound argument or proof showing that each such answer is correct.

And Supporting Main Claim …

And Supporting Main Claim …

$\mathcal{H}$ Humans have the ability to gain knowledge by reasoning (e.g., deductively) quantificationally and recursively over abstract concepts, including abstract concepts of a highly expressive, including infinitary, nature, expressed in arbitrarily complex natural and formal languages.

Objections

Objections

1. “But aren’t there a lot of people who can’t read and write?”

Objections

1. "But aren't there a lot of people who can't read and write?"
   - ✦ Yes, there are. Unfortunately at a minimum they will not be able to reach the higher levels of so-called 'present' rationality.

Objections

1. "But aren't there a lot of people who can't read and write?"
   - ✦ Yes, there are. Unfortunately at a minimum they will not be able to reach the higher levels of so-called 'present' rationality.
2. "Why are reading and writing so important? Hearing and speaking, why isn't that sufficient?"

Objections

1. "But aren't there a lot of people who can't read and write?"
   - ✦ Yes, there are. Unfortunately at a minimum they will not be able to reach the higher levels of so-called 'present' rationality.
2. "Why are reading and writing so important? Hearing and speaking, why isn't that sufficient?"
   - ✦ Well, the limitations of working memory, for one, means that reading and writing are tremendously powerful, cognitively speaking. Perhaps I once had Hamlet's famous soliloquy praising the human mind memorized, but I don't now — yet I can just look it up!

Objections

1. "But aren't there a lot of people who can't read and write?"
   - ✦ Yes, there are. Unfortunately at a minimum they will not be able to reach the higher levels of so-called 'present' rationality.
2. "Why are reading and writing so important? Hearing and speaking, why isn't that sufficient?"
   - ✦ Well, the limitations of working memory, for one, means that reading and writing are tremendously powerful, cognitively speaking. Perhaps I once had Hamlet's famous soliloquy praising the human mind memorized, but I don't now — yet I can just look it up!
3. "What about the ethics of how to treat beings who aren't on your scheme potentially or presently rational?"

Objections

1. "But aren't there a lot of people who can't read and write?"
   - ✦ Yes, there are. Unfortunately at a minimum they will not be able to reach the higher levels of so-called 'present' rationality.
2. "Why are reading and writing so important? Hearing and speaking, why isn't that sufficient?"
   - ✦ Well, the limitations of working memory, for one, means that reading and writing are tremendously powerful, cognitively speaking. Perhaps I once had Hamlet's famous soliloquy praising the human mind memorized, but I don't now — yet I can just look it up!
3. "What about the ethics of how to treat beings who aren't on your scheme potentially or presently rational?"
   - ✦ Any ethic that includes the notion that humans can be harmed etc because of any cognitive limitations is rejected out of hand (as itself an irrational policy).

Objections

1. "But aren't there a lot of people who can't read and write?"
   - ✦ Yes, there are. Unfortunately at a minimum they will not be able to reach the higher levels of so-called 'present' rationality.
2. "Why are reading and writing so important? Hearing and speaking, why isn't that sufficient?"
   - ✦ Well, the limitations of working memory, for one, means that reading and writing are tremendously powerful, cognitively speaking. Perhaps I once had Hamlet's famous soliloquy praising the human mind memorized, but I don't now — yet I can just look it up!
3. "What about the ethics of how to treat beings who aren't on your scheme potentially or presently rational?"
   - ✦ Any ethic that includes the notion that humans can be harmed etc because of any cognitive limitations is rejected out of hand (as itself an irrational policy).

Objections (con.)

Objections (con.)

1. "You place a heavy emphasis in *R* on reading and writing. But why couldn't we have a society of perfectly rational agents in which all communication and learning happens over purely *auditory* content?"

Objections (con.)

1. "You place a heavy emphasis in *R* on reading and writing. But why couldn't we have a society of perfectly rational agents in which all communication and learning happens over purely *auditory* content?"

   ✦ We should all be able to agree that purely auditory communication, stacked against the written word, would be horribly inefficient. In addition, ambiguity would be widespread and potentially devastating, since sound passing between agents is very hard if not impossible to keep univocal. In the case of symbols, for example then symbol '4,'

notice that '4' is immediately seen as referring to the same

thing. In the case of a sound, even tiny changes lead to ambiguity, lack of understanding, and confusion.

Objections (con.)

1. "You place a heavy emphasis in *R* on reading and writing. But why couldn't we have a society of perfectly rational agents in which all communication and learning happens over purely *auditory* content?"

   ✦ We should all be able to agree that purely auditory communication, stacked against the written word, would be horribly inefficient. In addition, ambiguity would be widespread and potentially devastating, since sound passing between agents is very hard if not impossible to keep univocal. In the case of symbols, for example then symbol '4,'

notice that '4' is immediately seen as referring to the same

thing. In the case of a sound, even tiny changes lead to ambiguity, lack of understanding, and confusion.

Objections (con.)

1. "You place a heavy emphasis in *R* on reading and writing. But why couldn't we have a society of perfectly rational agents in which all communication and learning happens over purely *auditory* content?"

   ✦ We should all be able to agree that purely auditory communication, stacked against the written word, would be horribly inefficient. In addition, ambiguity would be widespread and potentially devastating, since sound passing between agents is very hard if not impossible to keep univocal. In the case of symbols, for example then symbol '4,'

notice that '4' is immediately seen as referring to the same

thing. In the case of a sound, even tiny changes lead to ambiguity, lack of understanding, and confusion.

Further Discussion?