The Balderdash that is GPT-3 AI; Motivating Paradoxes, Puzzles, Part 2

Selmer Bringsjord

Intro to (Formal) Logic (and AI) = IFLAI1

2/1/21

Selmer.Bringsjord@gmail.com

Logistics …

Logistics …

Again, see Sergei's FAQ.

(Reading the syllabus is helpful.)

Hopefully you've studies last mtg's "stunners." Questions?

What *is* Logic?

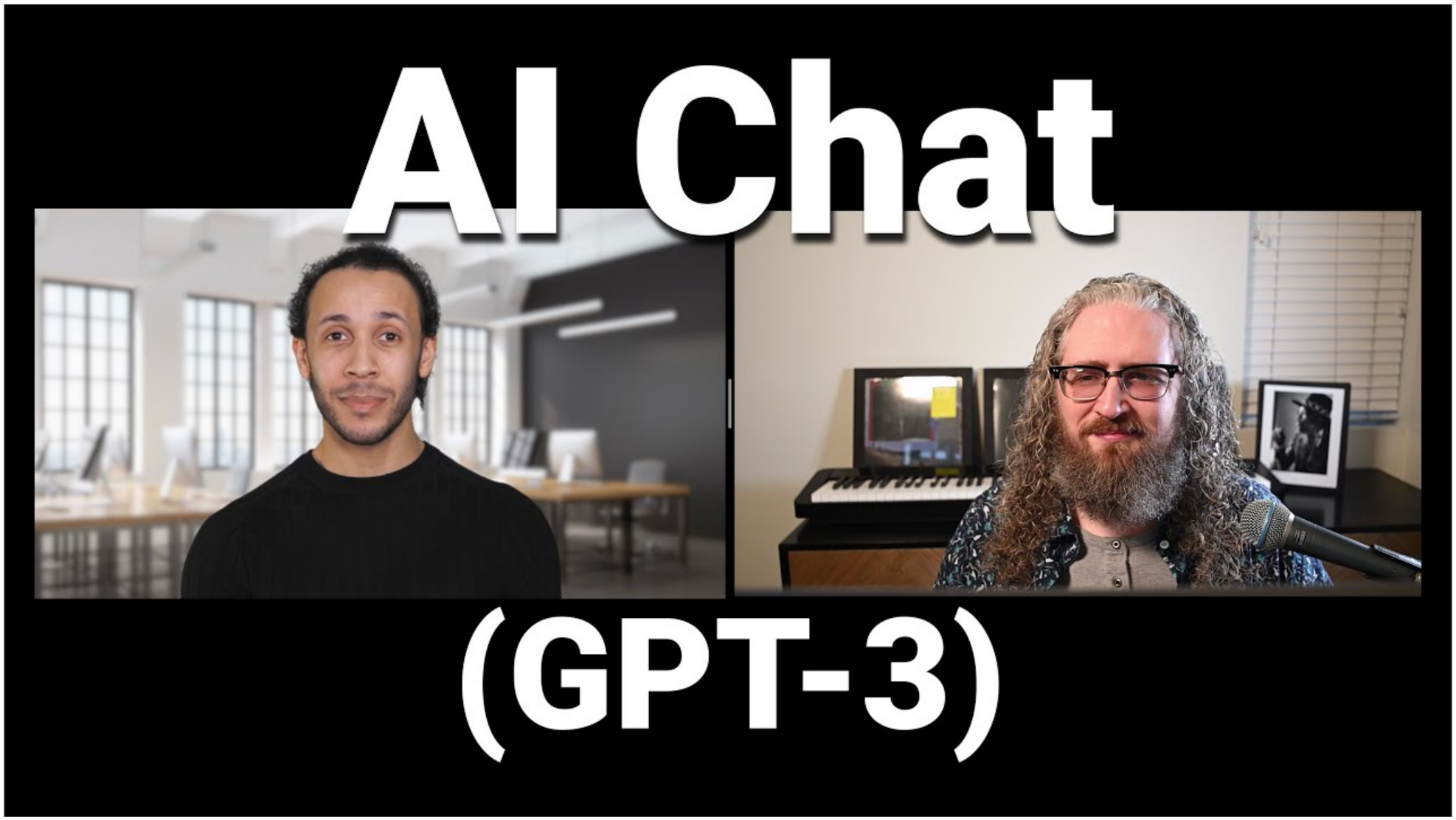
What *is* Logic?

- The key to becoming rational.
- "The science and engineering of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — and hence the basis for disciplines based on the formal sciences, e.g., …
  - Engineering! Computer Science!
  - Mathematics itself: see "reverse mathematics"!
- The way of escape from shallow content and context to pure, immaterial, and immortal form and structure (which is why the exotic, imaginary, and seemingly non-sensical is so pedagogically useful).
- The most challenging subject there is.
- One of the chief differentiators between dogs and monkeys versus you (let alone bears and you); and mindless machines (like Deep Blue & Watson) versus you.
- A key to riches.
- The key to divining the meaning of life (and other such big questions).
- The better way to program computers; and fundamentally the *only* way to *reliably* program computers.
- One of two fundamental approaches to studying minds, and replicating/simulating minds in machines…
- The thing many creatures of fiction have mastered — have you (as a New Yorker)?…
- …

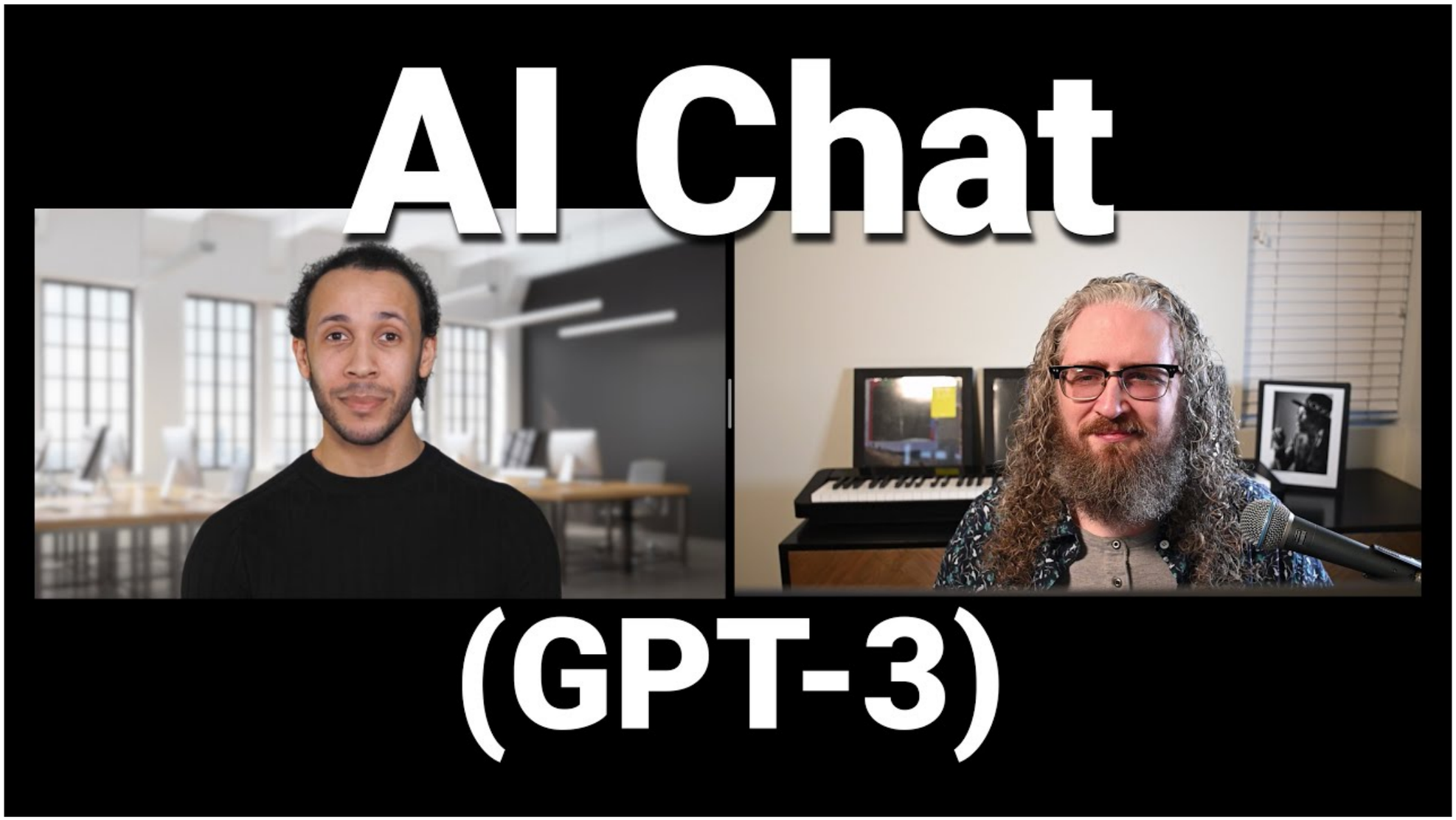
What *is* Logic?

- The key to becoming rational.
- "The science and engineering of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — and hence the basis for disciplines based on the formal sciences, e.g., …
  - Engineering! Computer Science!
  - Mathematics itself: see "reverse mathematics"!
- The way of escape from shallow content and context to pure, immaterial, and immortal form and structure (which is why the exotic, imaginary, and seemingly non-sensical is so pedagogically useful).
- The most challenging subject there is.
- One of the chief differentiators between dogs and monkeys versus you (let alone bears and you); and mindless machines (like Deep Blue & Watson) versus you.
- A key to riches.
- The key to divining the meaning of life (and other such big questions).
- The better way to program computers; and fundamentally the *only* way to *reliably* program computers.
- One of two fundamental approaches to studying minds, and replicating/simulating minds in machines…
- The thing many creatures of fiction have mastered — have you (as a New Yorker)?…
- …

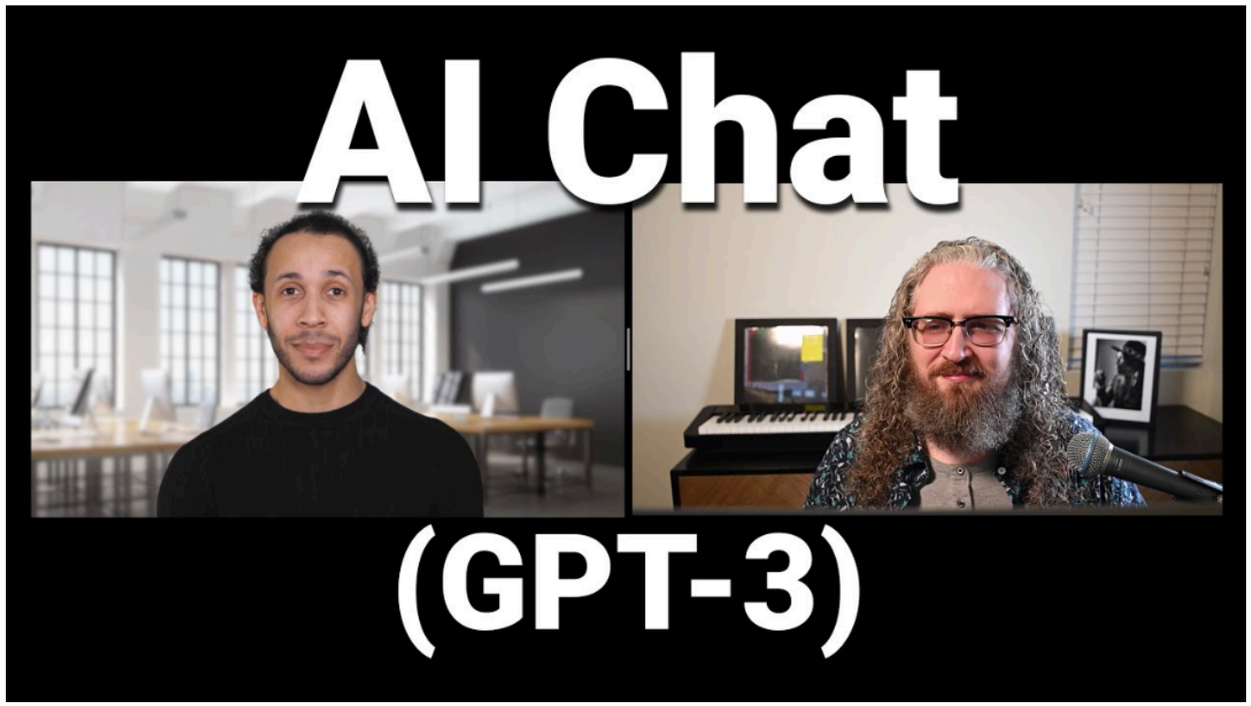
Non-Logicist AI



Non-Logicist AI



EE: "A lot of people have claimed that GPT-3 doesn't have have the ability to process information and reason about things …"

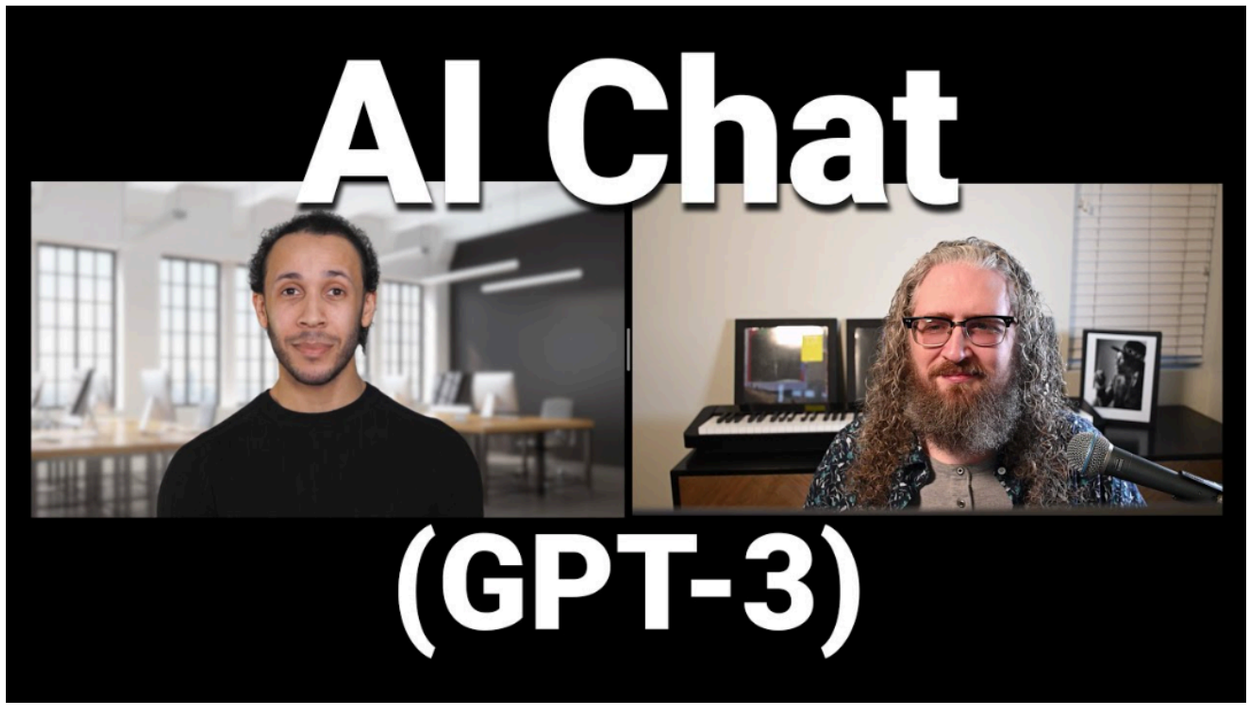


Non-Logicist AI

EE: "A lot of people have claimed that GPT-3 doesn't have have the ability to process information and reason about things …"

EE: "You know what I mean by lying, though, right?"

GPT-3: "Yes. I can make statements that I know are not true."



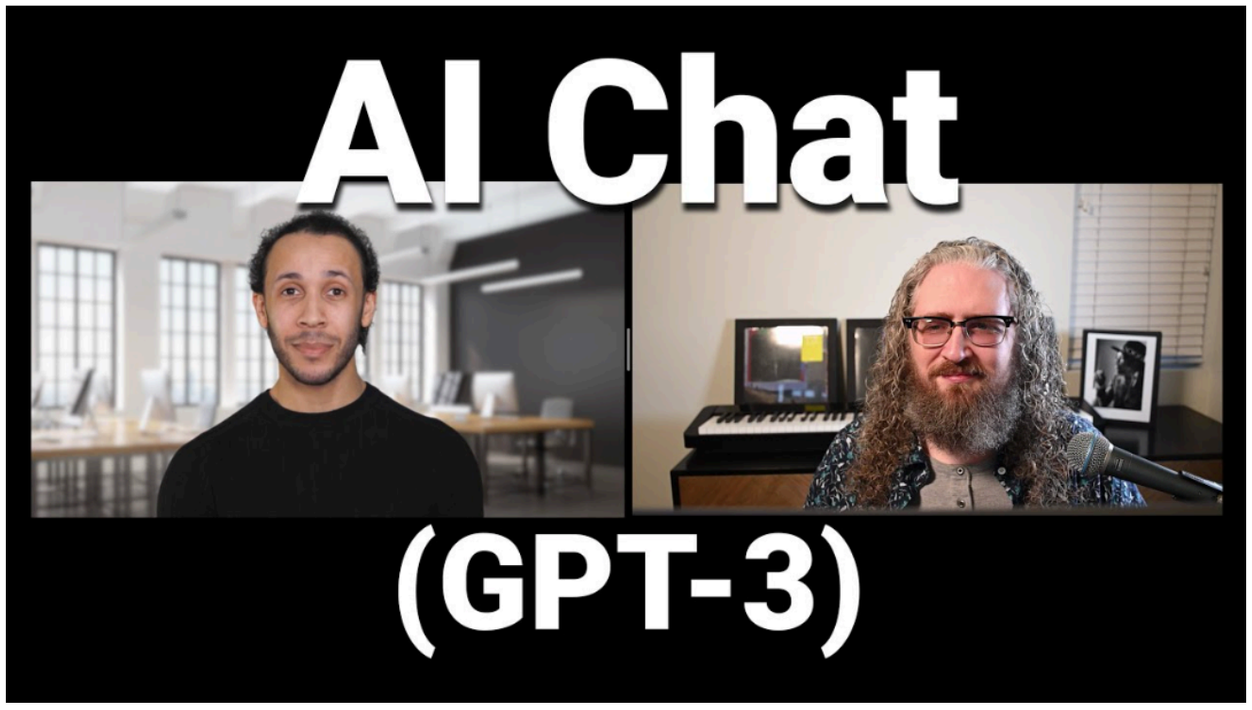
Non-Logicist AI

EE: "A lot of people have claimed that GPT-3 doesn't have have the ability to process information and reason about things …"

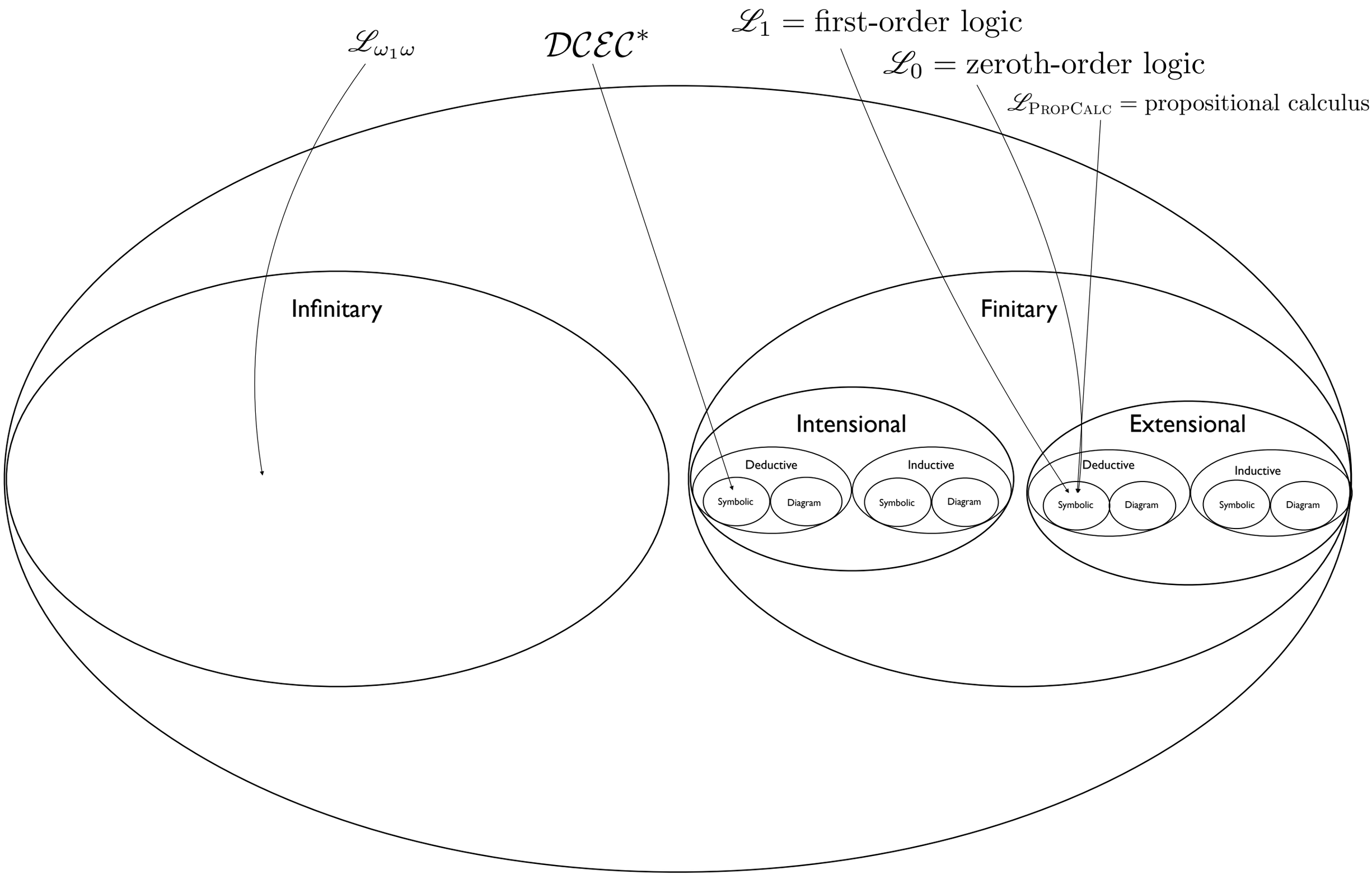
EE: "You know what I mean by lying, though, right?"

GPT-3: "Yes. I can make statements that I know are not true."

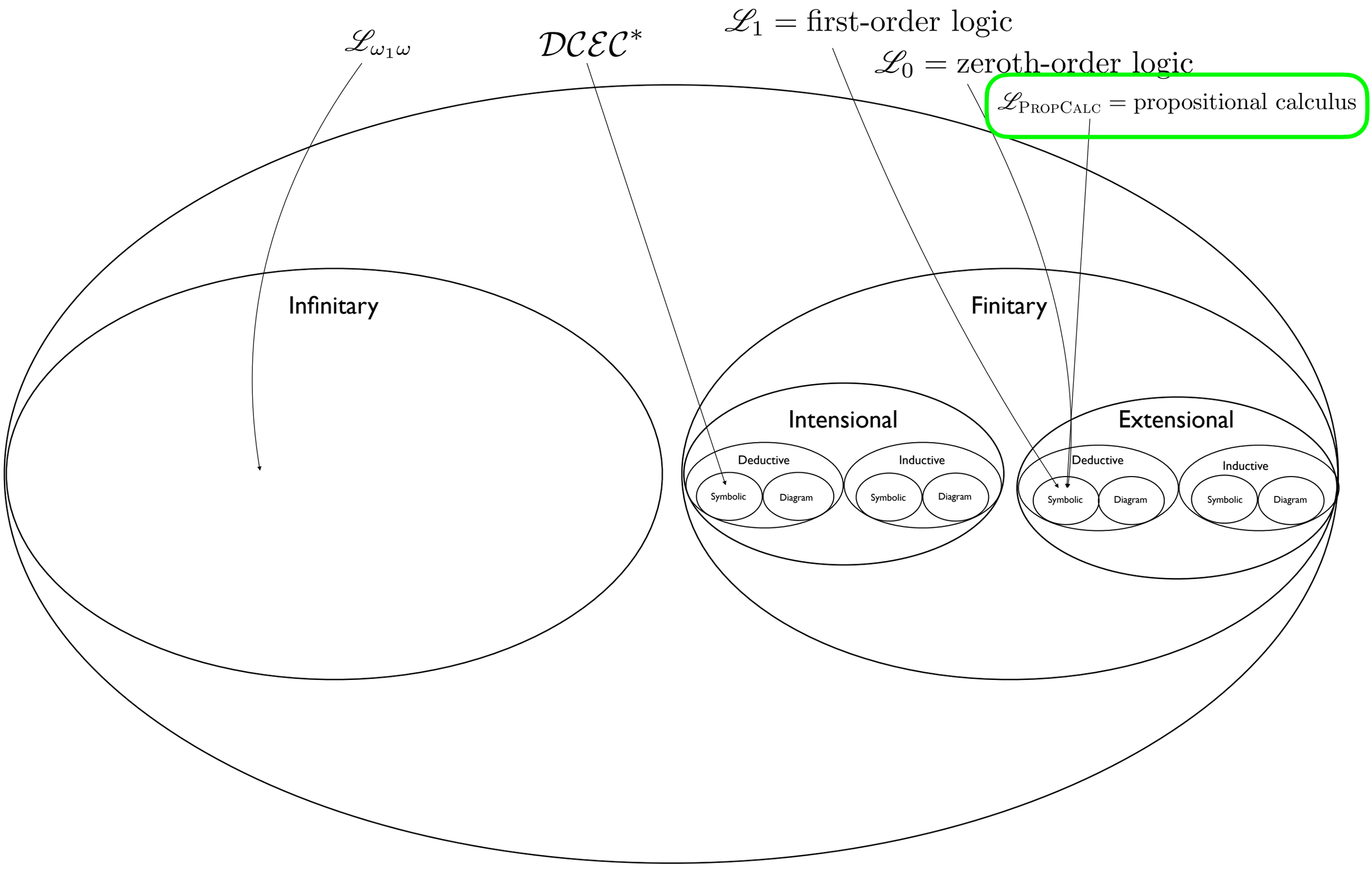
Sorry. That's not lying. Easy counter-examples abound.



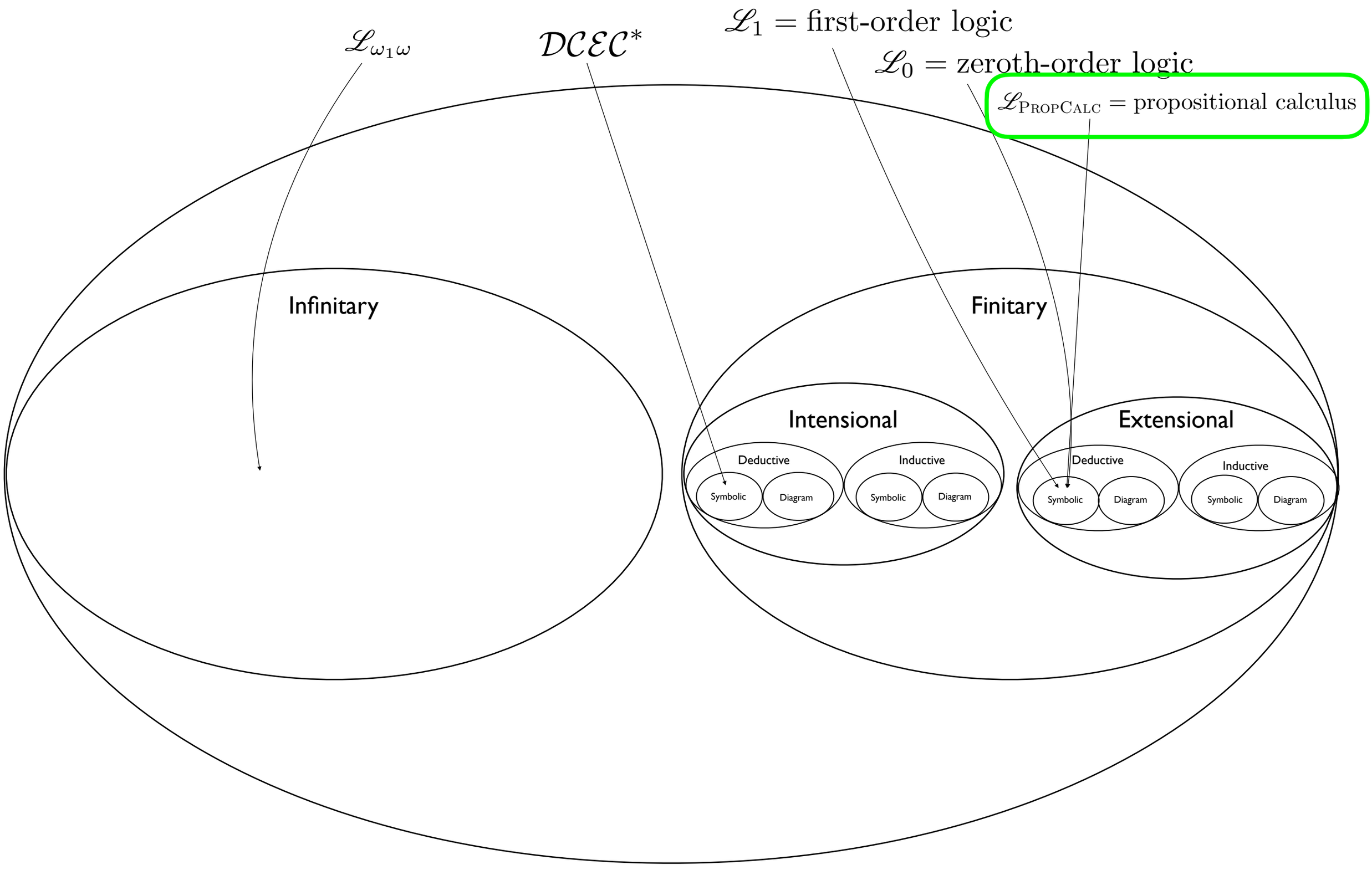
The Universe of Logics



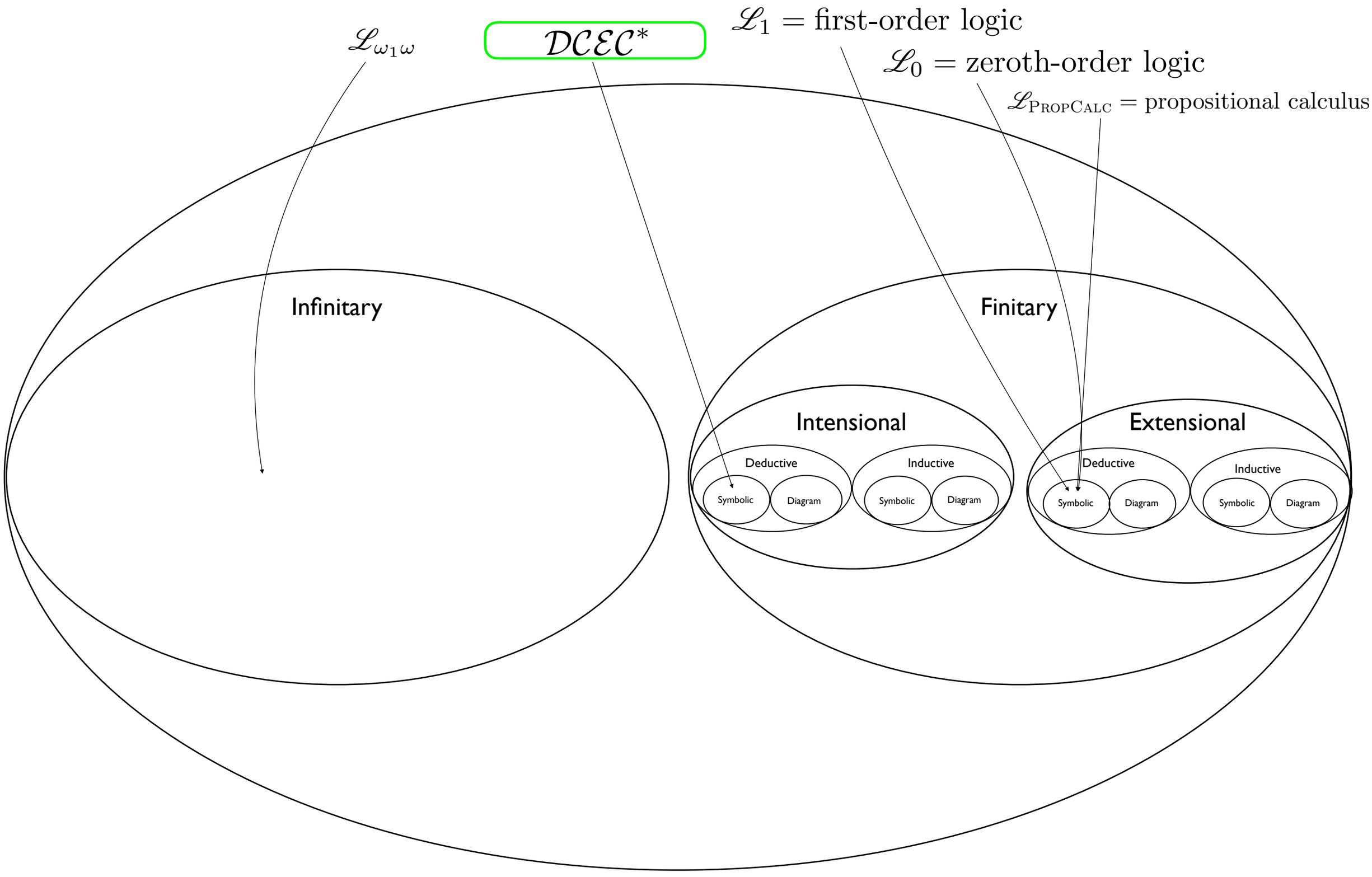
The Universe of Logics



The Universe of Logics



The Universe of Logics



Easy Counter-example:

Jones & Smith are criminal co-conspirators who believe it's possible that a detective is eavesdropping on their conversation. Jones *to Smith*: "I'll be hiking Pine Ridge all day tomorrow."

Easy Counter-example #2:

Jones & Smith are criminal co-conspirators, & Jones both believes it's possible that a detective Dan is eavesdropping on their conversation (which Smith knows too), and wants to get caught (which Smith doesn't know). Jones to Smith: "I'll be hiking Pine Ridge all day tomorrow." Here, Jones knows Dan knows that Smith has an injury that precludes his doing any hiking tomorrow, and Smith doesn't know about the injury. But Smith knows that Jones won't be hiking tomorrow and will be at home.

What *is* Logic?

What *is* Logic?

- The key to becoming rational.
- "The science of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — and hence the basis for disciplines based on the formal sciences (e.g., engineering, computer science).
- The way of escape from shallow content and context to pure, immaterial, and immortal form and structure (which is why the exotic, imaginary, and seemingly non-sensical is so pedagogically useful).
- The most challenging subject there is.
- One of the chief differentiators between dogs and monkeys versus you (let alone bears and you); and mindless machines (like Deep Blue & Watson) versus you.
- A key to riches.
- The key to divining the meaning of life (and other such big questions).
- The better way to program computers; and fundamentally the *only* way to *reliably* program computers.
- One of two fundamental approaches to studying minds, and replicating/simulating minds in machines…
- The thing many creatures of fiction have mastered — have you (as a New Yorker)?…

What *is* Logic?

- The key to becoming rational.
- "The science of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — and hence the basis for disciplines based on the formal sciences (e.g., engineering, computer science).
- The way of escape from shallow content and context to pure, immaterial, and immortal form and structure (which is why the exotic, imaginary, and seemingly non-sensical is so pedagogically useful).
- The most challenging subject there is.
- One of the chief differentiators between dogs and monkeys versus you (let alone bears and you); and mindless machines (like Deep Blue & Watson) versus you.
- A key to riches.
- The key to divining the meaning of life (and other such big questions).
- The better way to program computers; and fundamentally the *only* way to *reliably* program computers.
- One of two fundamental approaches to studying minds, and replicating/simulating minds in machines…
- The thing many creatures of fiction have mastered — have you (as a New Yorker)?…

Background 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, their minds are fundamentally qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is "No." For starters, if x can't read, write, and create, x can't be rational; computing machines/robots can neither read nor write nor create; ergo, they aren't fundamentally rational.

abstract-and-valid inference schemata

Background 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, their minds are fundamentally qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is "No." For starters, if x can't read, write, and create, x can't be rational; computing machines/robots can neither read nor write nor create; ergo, they aren't fundamentally rational.

abstract-and-valid inference schemata

quantification

Background 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, their minds are fundamentally qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is "No." For starters, if x can't read, write, and create, x can't be rational; computing machines/robots can neither read nor write nor create; ergo, they aren't fundamentally rational.

abstract-and-valid inference schemata

quantification

intensional reasoning

Background 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, their minds are fundamentally qualitatively inferior to the human mind. As to whether computing machines/robots are fundamentally rational, the answer is "No." For starters, if x can't read, write, and create, x can't be rational; computing machines/robots can neither read nor write nor create; ergo, they aren't fundamentally rational.

abstract-and-valid inference schemata

quantification

intensional reasoning

recursion

abstract-and-valid inference schemata

quantification

intensional reasoning

recursion

self-reference

abstract-and-valid inference schemata

quantification

intensional reasoning

recursion

self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning

recursion

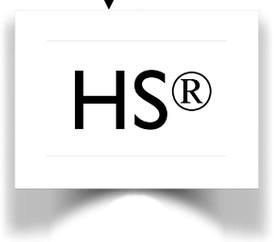
self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning



recursion

self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning

recursion

self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning

HS®

recursion

self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning

HS®

recursion

self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning

HS®

recursion

self-reference

To infinity and beyond! — routinely

abstract-and-valid inference schemata

quantification

intensional reasoning

HS®

recursion

self-reference

To infinity and beyond! — routinely

“NYS 1”

Given the statements

¬a ∨ ¬b
b
c → a

which one of the following statements can you prove?

c
¬b
¬c
h
a
none of the above

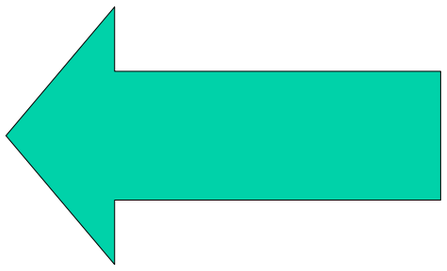
“NYS 1”

Given the statements

¬a ∨ ¬b
b
c → a

which one of the following statements can you prove?

c
¬b
¬c
h
a
none of the above



“NYS 2”

Which one of the following statements is provable from the following statement: “If you are not part of the solution, then you are part of the problem.”

If you are part of the solution, then you are not part of the problem.

If you are not part of the problem, then you are part of the solution.

If you are part of the problem, then you are not part of the solution.

If you are not part of the problem, then you are not part of the solution.

“NYS 2”

Which one of the following statements is provable from the following statement: “If you are not part of the solution, then you are part of the problem.”

If you are part of the solution, then you are not part of the problem.

If you are not part of the problem, then you are part of the solution.

If you are part of the problem, then you are not part of the solution.

If you are not part of the problem, then you are not part of the solution.

“NYS 3”

Given the statements

¬¬c

c → a

¬a ∨ b

b → d

¬(d ∨ e)

which of the following statements are provable?

¬c

e

h

¬a

all of the above

“NYS 3”

Given the statements

¬¬c

c → a

¬a ∨ b

b → d

¬(d ∨ e)

which of the following statements are provable?

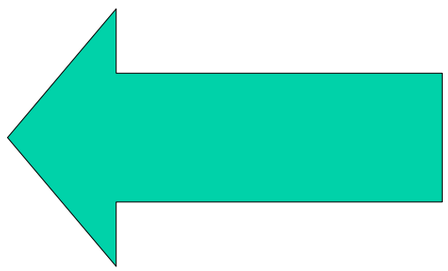
¬c

e

h

¬a

all of the above



A criminal genius nearly a match for Sherlock Holmes (Do you recognize the Dr?) has built a massive hydrogen bomb, and life on Earth is hanging in the balance, hinging on whether you make the logical prediction.

Dr M gives you a sporting chance to: make the right prediction, snip or not snip accordingly, and prove that you're right …

(With a debt to P. Johnson-Laird for the king-ace kernel used here.)

A criminal genius nearly a match for Sherlock Holmes (Do you recognize the Dr?)

(With a debt to P. Johnson-Laird for the king-ace kernel used here.)

A criminal genius nearly a match for Sherlock Holmes (Do you recognize the Dr?) has built a massive hydrogen bomb, and life on Earth is hanging in the balance, hinging on whether you make the logical prediction.

Dr M gives you a sporting chance to: make the right prediction, snip or not snip accordingly, and prove that you're right …

(With a debt to P. Johnson-Laird for the king-ace kernel used here.)

If one of the following assertions is true then so is the other:

(1) If the red wire runs to the bomb, then the blue wire runs to the bomb; and, if the blue wire runs to the bomb, then the red wire runs to the bomb.

(2) The red wire runs to the bomb.

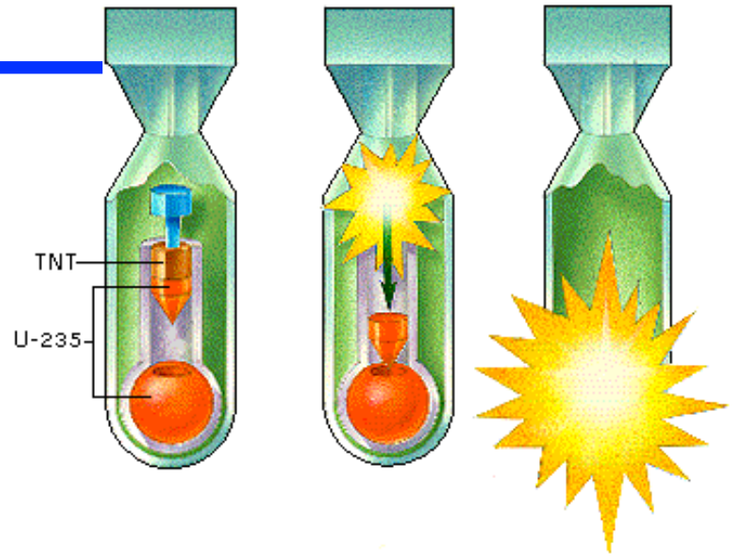
Given this perfectly reliable clue from Dr Moriarty, if either wire is more likely to run to the bomb, that wire *do*es run to the bomb, and the bomb is ticking, with only a minute left! If both are equiprobable, neither runs to the bomb, and you are powerless. Make your prediction as to what will happen when a wire is snipped, and then make your selected snip by clicking on the wire you want to snip! Or leave well enough alone!

Red more likely.

Blue more likely.

Equiprobable.

Snip



Life on Earth has ended.

advance one more slide to see a proof that you indeed made an irrational decision…

Proposition: The blue wire is more likely!

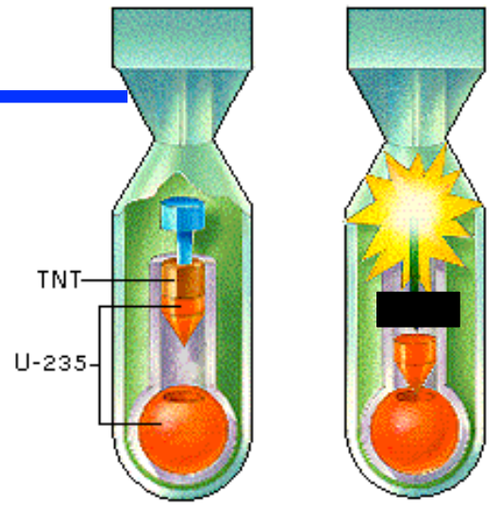
Proof: (1) can be treated as a biconditional, obviously (R <=> B).

There are two top-level cases to consider: (1) and (2) are both true; or both are false. In the case where they are both true, it's trivial to deduce both R and B. So far, then, R and B are equiprobable. What happens in the case where (1) and (2) are both false? We immediately have ~R from the denial of (2). But a biconditional is true just in case both sides are true, or both sides are false; so we have two sub-cases to consider.

Consider first the case where R is true and B is false. We have an immediate contradiction in this sub-case, so both R and B can both be deduced here, and we have not yet departed from equiprobable. So what about the case where R is false and B is true? The falsity of R is not new information (we already have that from the denial of (2)), but we can still derive B. Hence the blue wire is more likely. **QED**

STOP

Snip



Life on Earth is saved!

if you can now hand Dr M a proof that your decision was the rational one!

Advance one more slide to see a proof from Bringsjord that yours had better match up to …

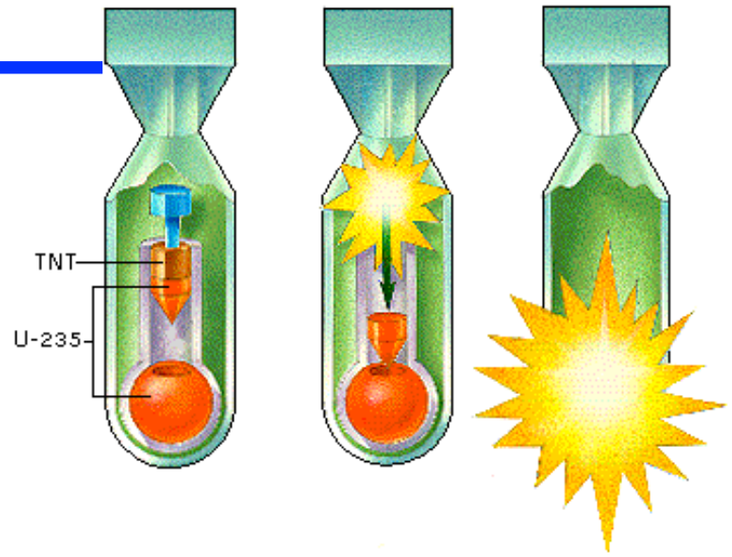
Proposition: The blue wire is more likely!

Proof: (1) can be treated as a biconditional, obviously (R <=> B).

There are two top-level cases to consider: (1) and (2) are both true; or both are false. In the case where they are both true, it's trivial to deduce both R and B. So far, then, R and B are equiprobable. What happens in the case where (1) and (2) are both false? We immediately have ~R from the denial of (2). But a biconditional is true just in case both sides are true, or both sides are false; so we have two sub-cases to consider.

Consider first the case where R is true and B is false. We have an immediate contradiction in this sub-case, so both R and B can both be deduced here, and we have not yet departed from equiprobable. So what about the case where R is false and B is true? The falsity of R is not new information (we already have that from the denial of (2)), but we can still derive B. Hence the blue wire is more likely. **QED**

STOP



Life on Earth has ended.

advance one more slide to see a proof that you indeed made an irrational decision…

Proposition: The blue wire is more likely!

Proof: (1) can be treated as a biconditional, obviously (R <=> B).

There are two top-level cases to consider: (1) and (2) are both true; or both are false. In the case where they are both true, it's trivial to deduce both R and B. So far, then, R and B are equiprobable. What happens in the case where (1) and (2) are both false? We immediately have ~R from the denial of (2). But a biconditional is true just in case both sides are true, or both sides are false; so we have two sub-cases to consider.

Consider first the case where R is true and B is false. We have an immediate contradiction in this sub-case, so both R and B can both be deduced here, and we have not yet departed from equiprobable. So what about the case where R is false and B is true? The falsity of R is not new information (we already have that from the denial of (2)), but we can still derive B. Hence the blue wire is more likely. **QED**

STOP

Logic kan redde menneskehten!