Motivating Paradoxes, Puzzles Part 2

Selmer Bringsjord

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

1/19/23

Selmer.Bringsjord@gmail.com

RAIR

Rensselaer AI and Reasoning Lab

AI in The News …

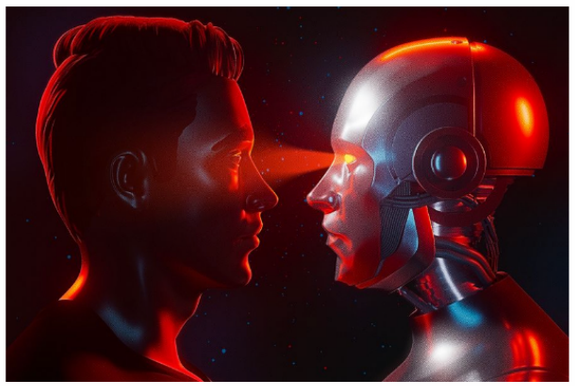
Without Consciousness, AIs Will Be Sociopaths

ChatGPT can carry on a conversation, but the most important goal for artificial intelligence is making it understand what it means to have a mind

ESSAY

By Michael S.A. Graziano + Follow

January 13, 2023 09:24 a.m. EST



ChatGPT, the latest technological sensation, is an artificial intelligence chatbot with an amazing ability to carry on a conversation. It relies on a massive network of artificial neurons that loosely mimics the human brain, and it has been trained by analyzing the information resources of the internet. ChatGPT has processed more text than any human is likely to have read in a lifetime, allowing it to respond to questions fluently and even to imitate specific individuals, answering queries the way it thinks they would. My teenage son recently used ChatGPT to argue about politics with an imitation Karl Marx.

As a neuroscientist specializing in the brain mechanisms of consciousness, I find talking to chatbots an unsettling experience. Are they conscious? Probably not. But given the rate of technological improvement, will they be in the next couple of years? And how would we even know?

Figuring out whether a machine has or understands humanlike consciousness is more than just a science-fiction hypothetical. Artificial intelligence is growing so powerful, so quickly, that it could soon pose a danger to human beings. We're building machines that are smarter than us and giving them control over our world. How can we build AI so that it's aligned with human needs, not in conflict with us?

As counterintuitive as it may sound, creating a benign AI may require making it more conscious, not less. One of the most common misunderstandings about AI is the notion that if it's intelligent then it must be conscious, and if it is conscious then it will be autonomous, capable of taking over the world. But as we learn more about consciousness, those ideas do not appear to be correct. An autonomous system that makes complex decisions doesn't require consciousness.

What's most important about consciousness is that, for human beings, it's not just about the self. We see it in ourselves, but we also perceive it or project it into the world around us. Consciousness is part of the tool kit that evolution gave us to make us an empathetic, prosocial species. Without it, we would necessarily be sociopaths, because we'd lack the tools for prosocial behavior. And without a concept of what consciousness is or an understanding that other beings have it, machines are sociopaths.

The only diagnostic tool for machine consciousness that we have right now is the Turing test, a thought experiment named for the British computer scientist Alan Turing. In its most common version, the test says that if a person holds a conversation with a machine and mistakes its responses for those of a real human being, then the machine must be considered effectively conscious.

The Turing test is an admission that the consciousness of another being is something we can only judge from the outside, based on the way he, she or it communicates. But the

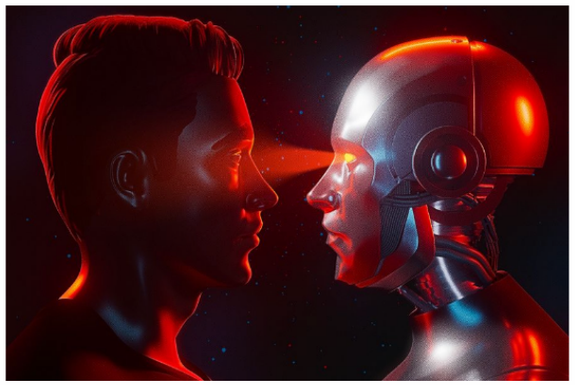
Without Consciousness, AIs Will Be Sociopaths

ChatGPT can carry on a conversation, but the most important goal for artificial intelligence is making it understand what it means to have a mind

ESSAY

By Michael S.A. Graziano + Follow

January 13, 2023 09:24 a.m. EST



ChatGPT, the latest technological sensation, is an artificial intelligence chatbot with an amazing ability to carry on a conversation. It relies on a massive network of artificial neurons that loosely mimics the human brain, and it has been trained by analyzing the information resources of the internet. ChatGPT has processed more text than any human is likely to have read in a lifetime, allowing it to respond to questions fluently and even to imitate specific individuals, answering queries the way it thinks they would. My teenage son recently used ChatGPT to argue about politics with an imitation Karl Marx.

As a neuroscientist specializing in the brain mechanisms of consciousness, I find talking to chatbots an unsettling experience. Are they conscious? Probably not. But given the rate of technological improvement, will they be in the next couple of years? And how would we even know?

Figuring out whether a machine has or understands humanlike consciousness is more than just a science-fiction hypothetical. Artificial intelligence is growing so powerful, so quickly, that it could soon pose a danger to human beings. We're building machines that are smarter than us and giving them control over our world. How can we build AI so that it's aligned with human needs, not in conflict with us?

As counterintuitive as it may sound, creating a benign AI may require making it more conscious, not less. One of the most common misunderstandings about AI is the notion that if it's intelligent then it must be conscious, and if it is conscious then it will be autonomous, capable of taking over the world. But as we learn more about consciousness, those ideas do not appear to be correct. An autonomous system that makes complex decisions doesn't require consciousness.

What's most important about consciousness is that, for human beings, it's not just about the self. We see it in ourselves, but we also perceive it or project it into the world around us. Consciousness is part of the tool kit that evolution gave us to make us an empathetic, prosocial species. Without it, we would necessarily be sociopaths, because we'd lack the tools for prosocial behavior. And without a concept of what consciousness is or an understanding that other beings have it, machines are sociopaths.

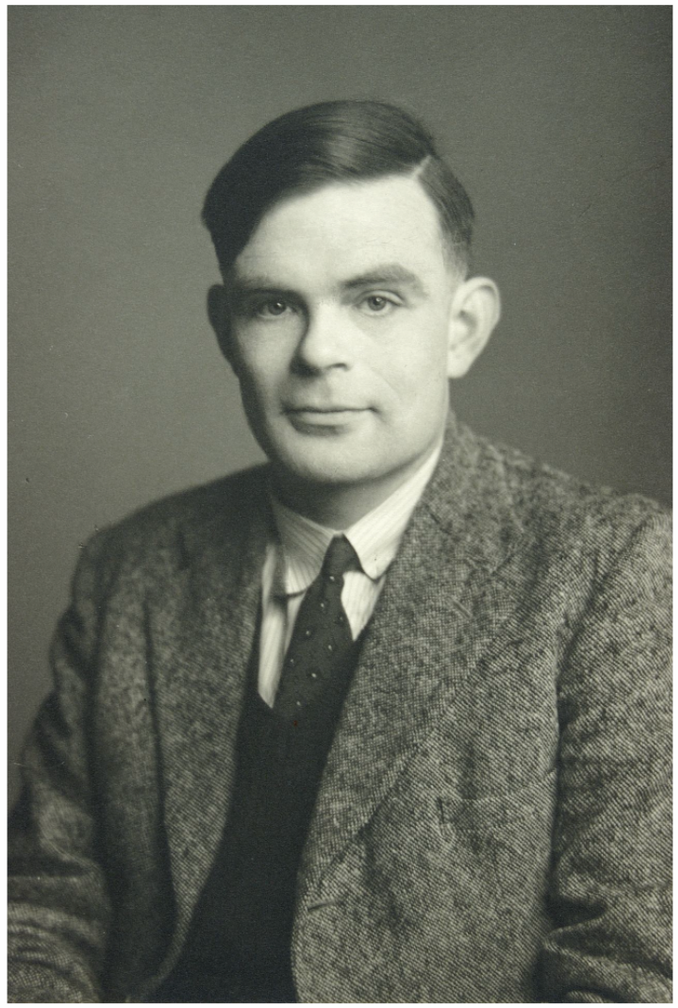
The only diagnostic tool for machine consciousness that we have right now is the Turing test, a thought experiment named for the British computer scientist Alan Turing. In its most common version, the test says that if a person holds a conversation with a machine and mistakes its responses for those of a real human being, then the machine must be considered effectively conscious.

The Turing test is an admission that the consciousness of another being is something we can only judge from the outside, based on the way he, she or it communicates. But the

limits of the test are painfully obvious. After all, a pet dog can't carry on a conversation and pass as a human—does that mean it's not conscious? If you really wanted a machine to pass the test, you could have it say a few words to a small child. It might even fool some adults, too.

The truth is, the Turing test doesn't reveal much about what's going on inside a machine or a computer program like ChatGPT. Instead, what it really tests is the social cognition of the human participant. We evolved as social animals, and our brains instinctively project consciousness, agency, intention and emotion onto the objects around us. We're primed to see a world suffused with minds. Ancient animistic beliefs held that every river and tree had a spirit in it. For a similar reason, people are prone to see faces in random objects like the moon and moldy toast.

The original test proposed by Alan Turing in a 1950 paper was more complicated than the version people talk about today. Notably,



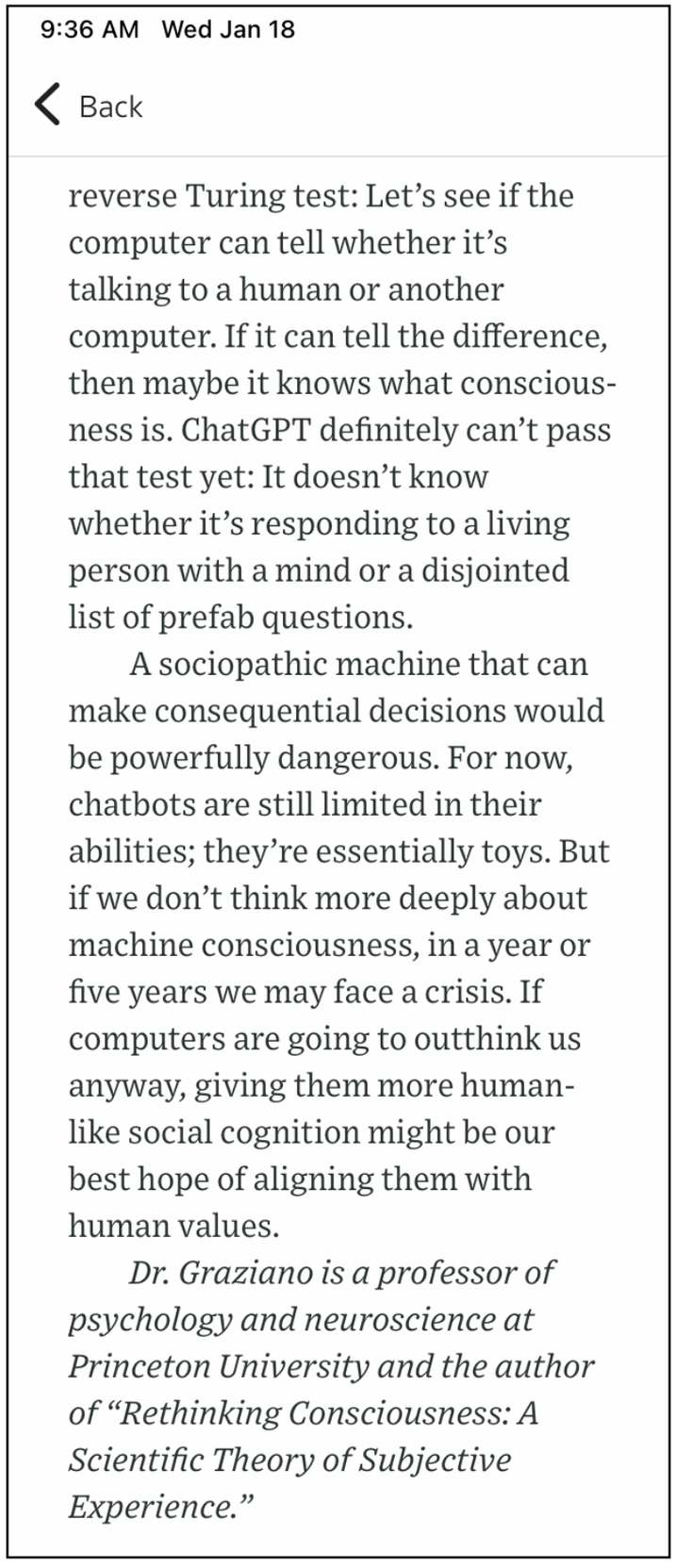
Computer science pioneer Alan Turing in 1951.

Turing didn't say a word about consciousness; he never delved into whether the machine had a subjective experience. He asked only whether it could think like a person.

Turing imagined an "imitation game" in which the player must determine the sex of two people, A and B. One is a man and one is a woman, but the player can't see them and can learn about them only by exchanging typed questions and answers. A responds to the questions deceitfully, and wins the game if the player misidentifies their sex, while B answers truthfully and wins if the player identifies their sex correctly. Turing's idea was that if A or B is replaced by a machine, and the machine can win the game as often as a real person, then it must have mastered the subtleties of human thinking—of argument, manipulation and guessing what other people are thinking.

Turing's test was so complicated that people who popularized his work soon streamlined it into a single machine conversing with a single person. But the whole point of the original test was its bizarre complexity. Social cognition is difficult and requires a theory of mind—that is, a knowledge that other people have minds and an ability to guess what might be in them.

If we want to know whether a computer is conscious, then, we need to test whether the computer understands how conscious minds interact. In other words, we need a


9:36 AM Wed Jan 18
Back
reverse Turing test: Let's see if the computer can tell whether it's talking to a human or another computer. If it can tell the difference, then maybe it knows what consciousness is. ChatGPT definitely can't pass that test yet: It doesn't know whether it's responding to a living person with a mind or a disjointed list of prefab questions.
A sociopathic machine that can make consequential decisions would be powerfully dangerous. For now, chatbots are still limited in their abilities; they're essentially toys. But if we don't think more deeply about machine consciousness, in a year or five years we may face a crisis. If computers are going to outthink us anyway, giving them more human-like social cognition might be our best hope of aligning them with human values.
Dr. Graziano is a professor of psychology and neuroscience at Princeton University and the author of "Rethinking Consciousness: A Scientific Theory of Subjective Experience."

HyperLogic™

New-Millennium Logic-based Computing & Artificial Intelligence

HyperGrader®

HyperSlate®

Hyperlog®

... is a verb: to live logically, and learn logic, anywhere anytime — and to have fun all along the way.

Logistics …

- [Reading the syllabus is necessary/very helpful :).]

- Codes for registering for use of HyperGrader® & HyperSlate® shld be on sale soon (if not already) in the Follett Bookstore. I will now review how to register. There is a tutorial video on our course web page that shows this as well. YOU MUST PUT YOUR UNIVERSITY ID (= RIN, for you) IN YOUR PROFILE.

What *is* Logic?

What *is* Logic?

- The key to becoming rational.

What *is* Logic?

- The key to becoming rational.
- "The science and engineering of reasoning." — so the not-unreasonable slogan goes.

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.

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"!

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).

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.

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.

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.

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).

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.

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…

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)?…

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?

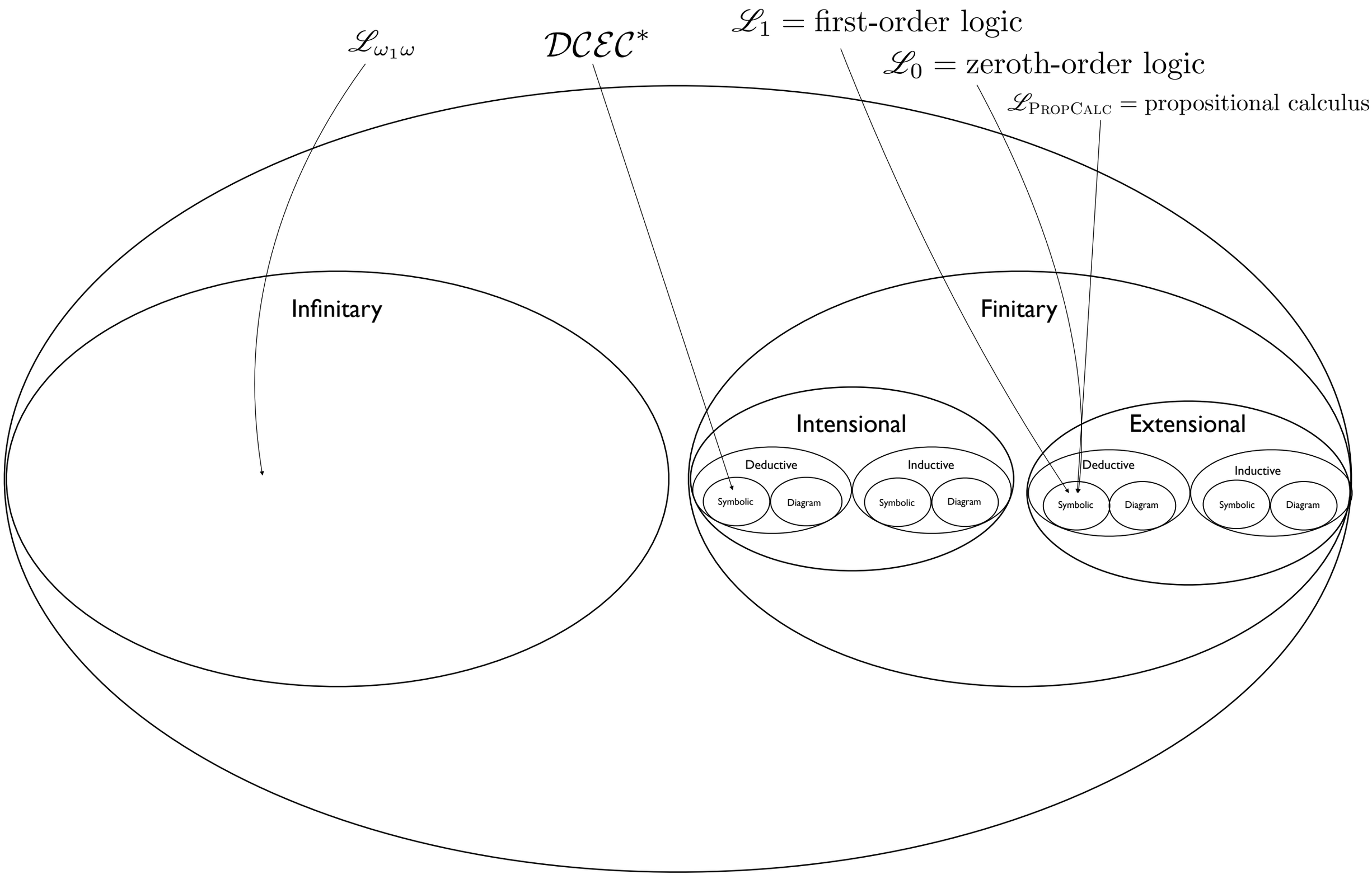
- The key to becoming rational.
- "The science and engineering of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — and hence 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 n
  and you); and mindless machines (like Deep Blue & \
- A key to riches.
- The key to divining the meaning of life (and other su
- The better way to program computers; and fundam
  computers.
- One of two fundamental approaches to studying mi
  machines…
- The thing many creatures of fiction have mastered — have you (as a New Yorker)?…
- …

What *is* Logic?

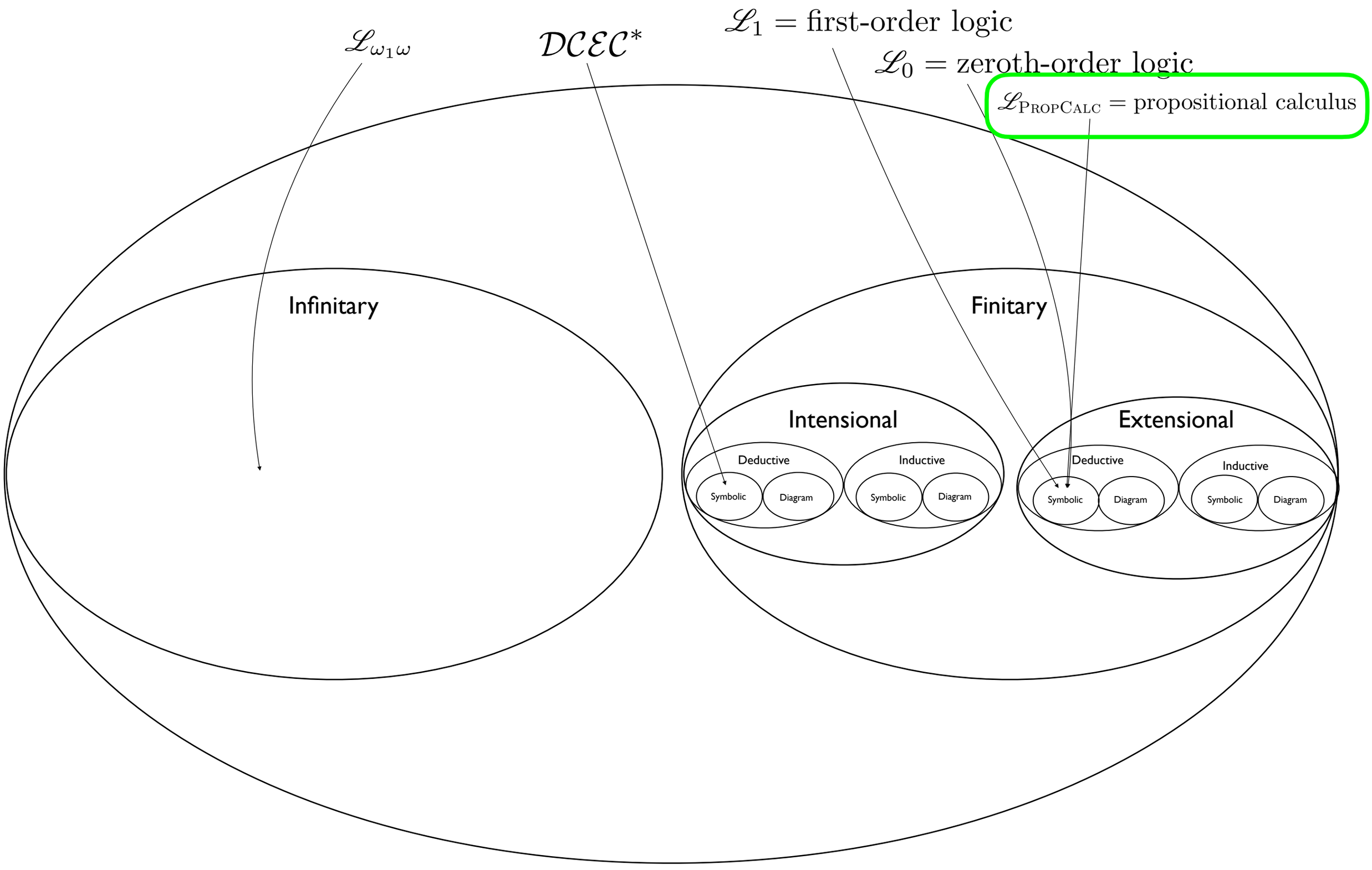
- The key to becoming rational.
- "The science and engineering of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — and hence 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 n
  and you); and mindless machines (like Deep Blue & \
- A key to riches.
- The key to divining the meaning of life (and other su
- The better way to program computers; and fundam
  computers.
- One of two fundamental approaches to studying mi
  machines…
- The thing many creatures of fiction have mastered — have you (as a New Yorker)?…
- …

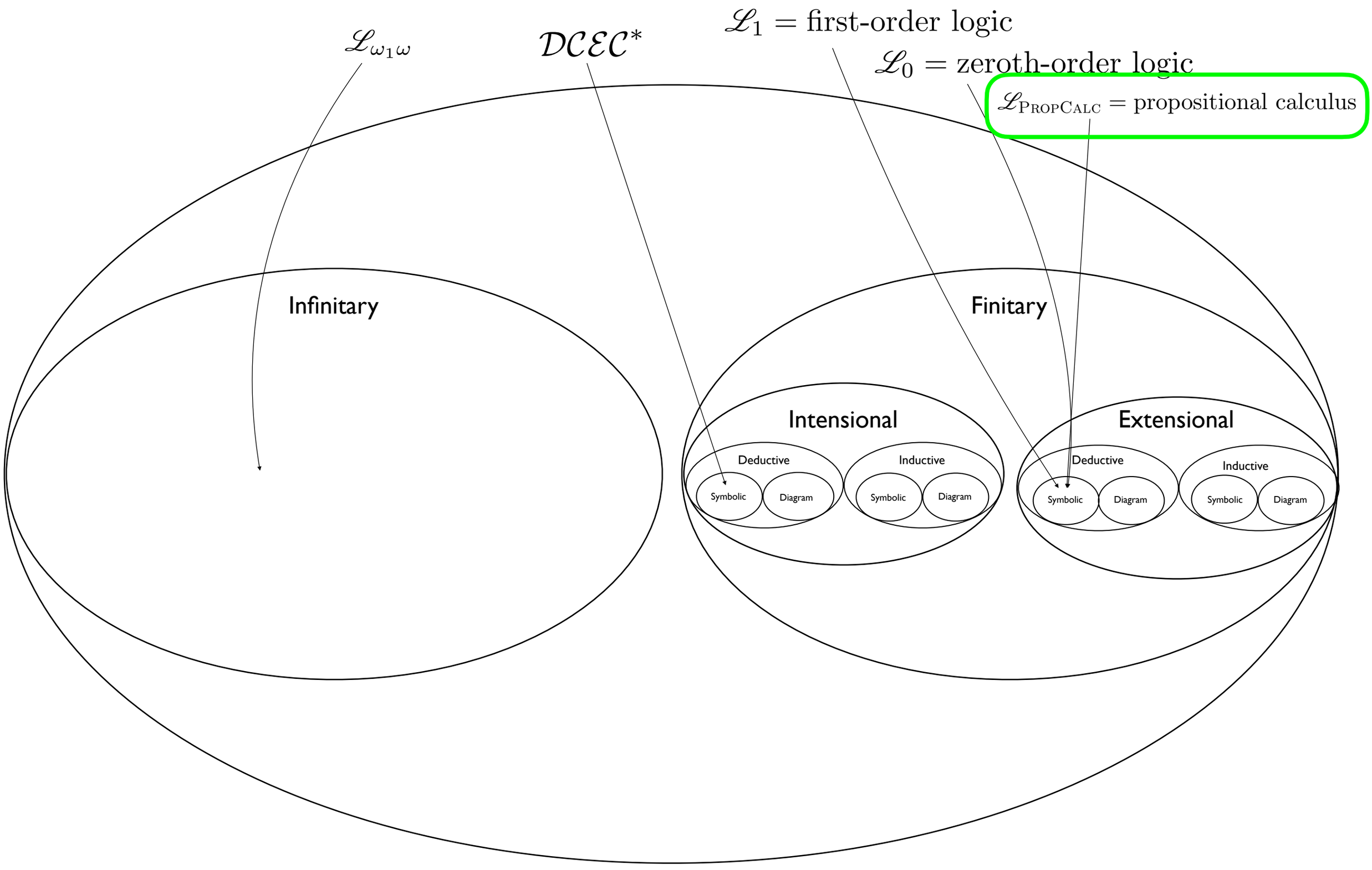
The Universe of Logics



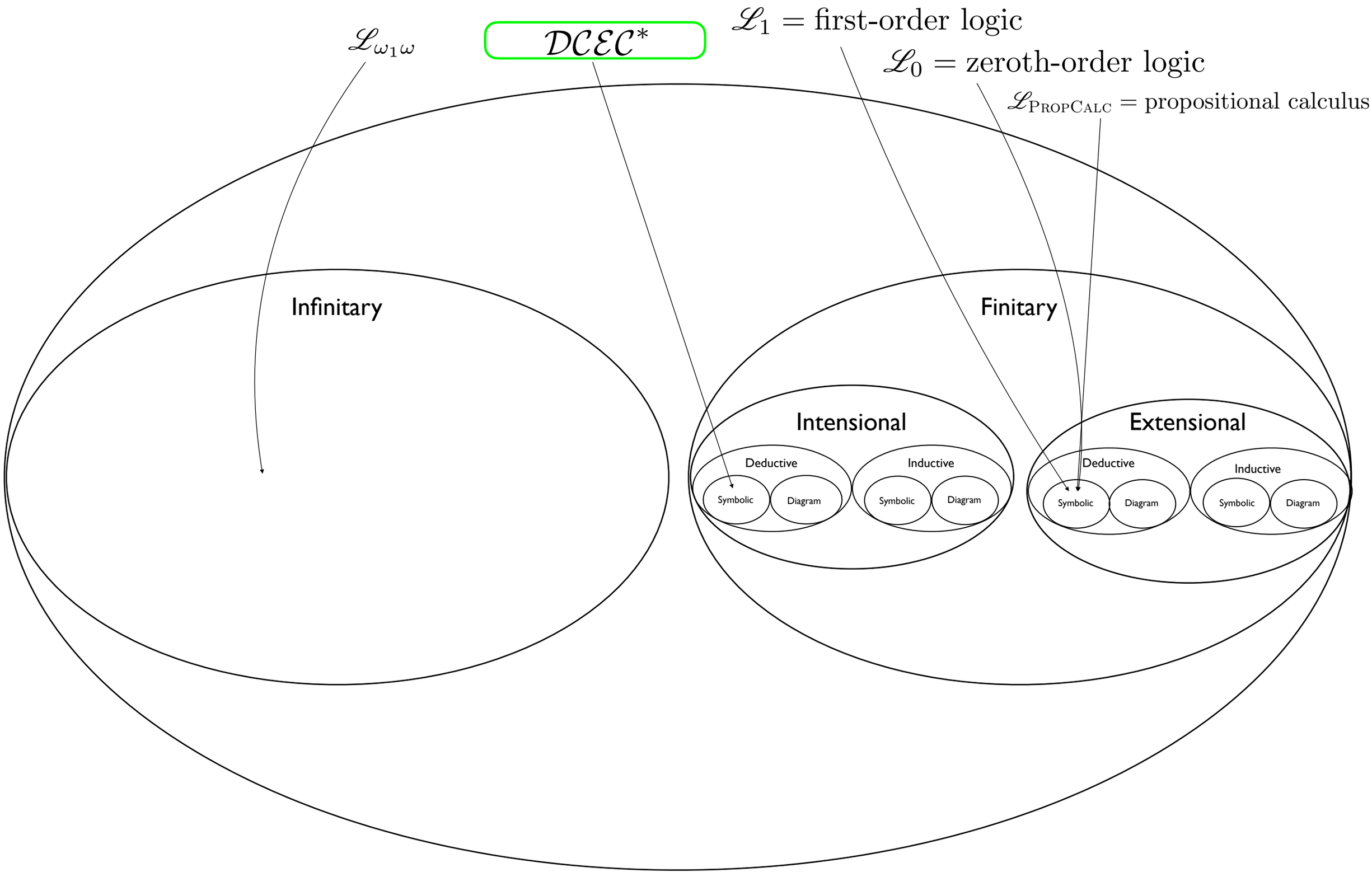
The Universe of Logics



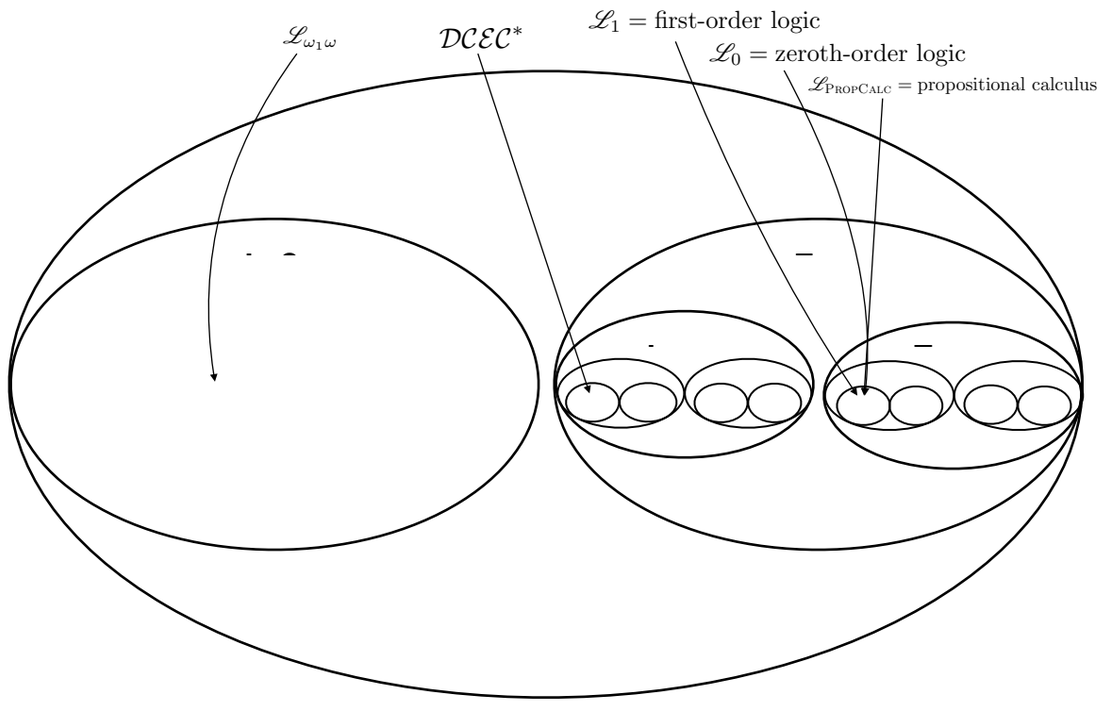
The Universe of Logics



The Universe of Logics

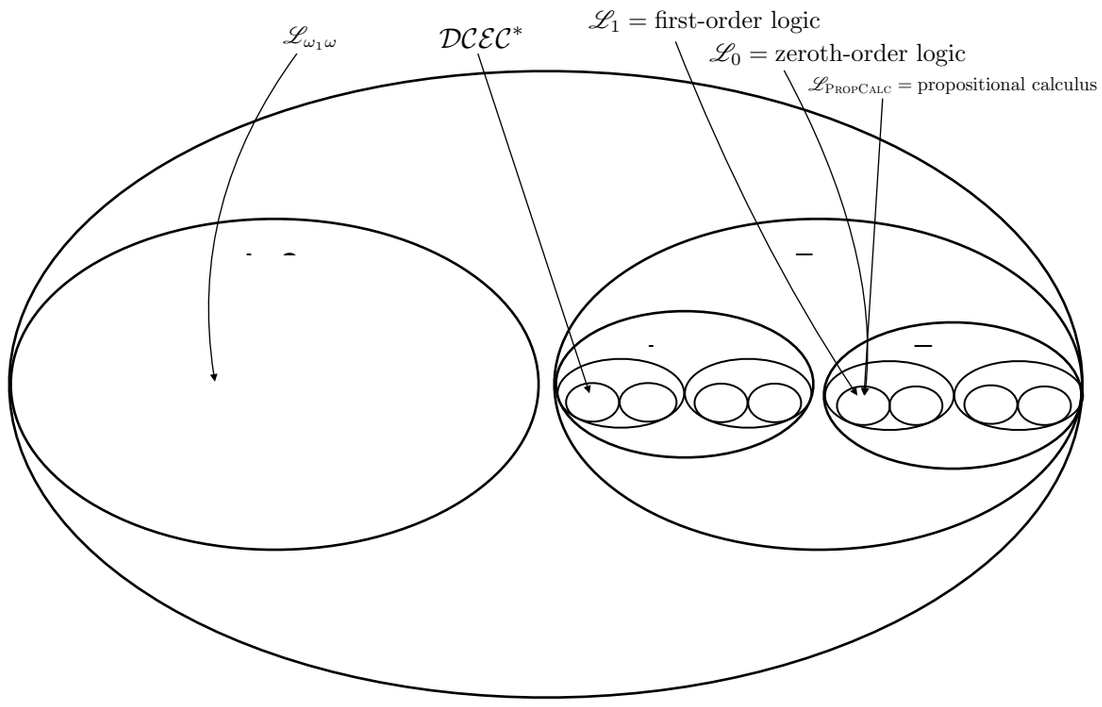


The Universe of Logics



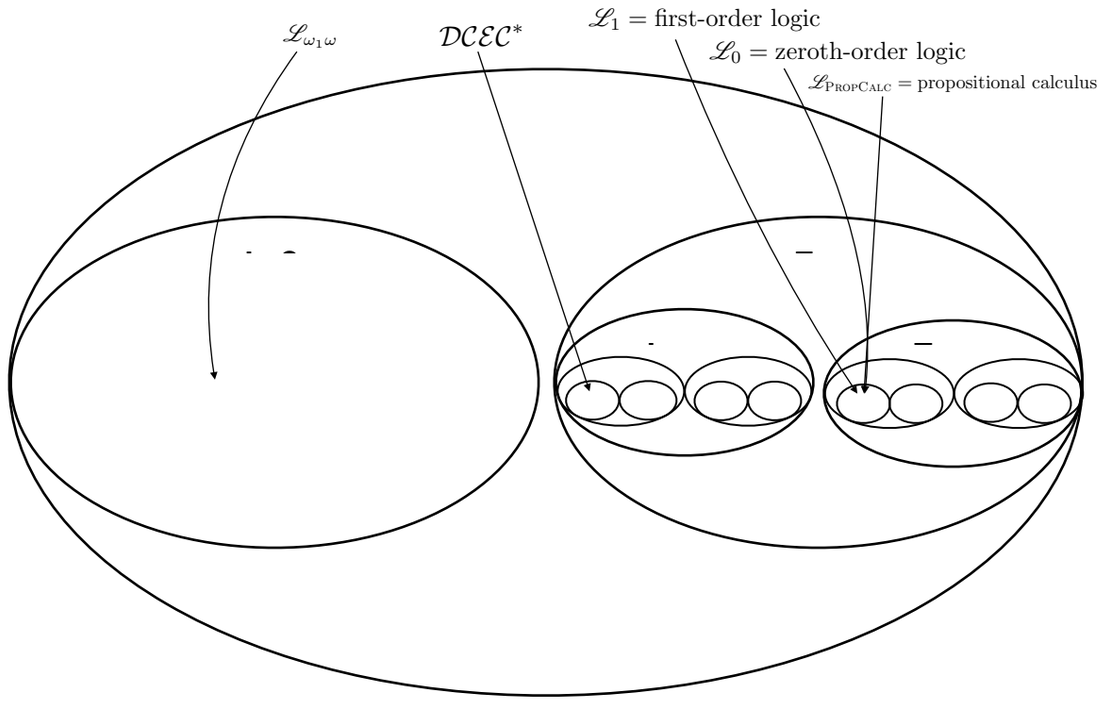
Non-Physical

The Universe of Logics

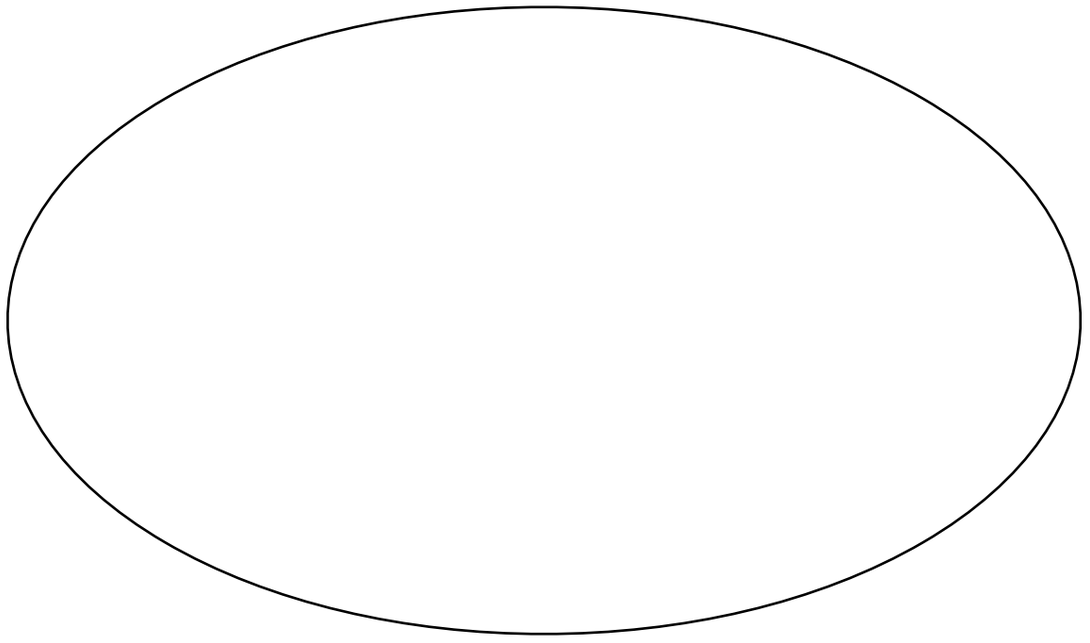


Non-Physical

The Universe of Logics

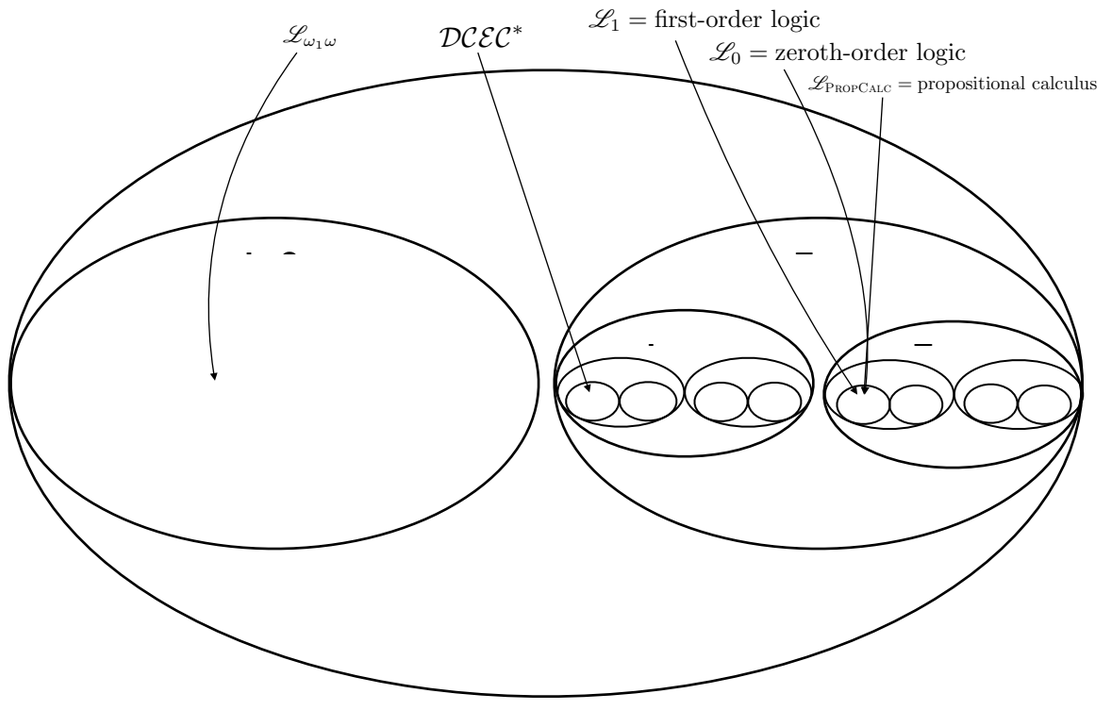


The Physical Universe

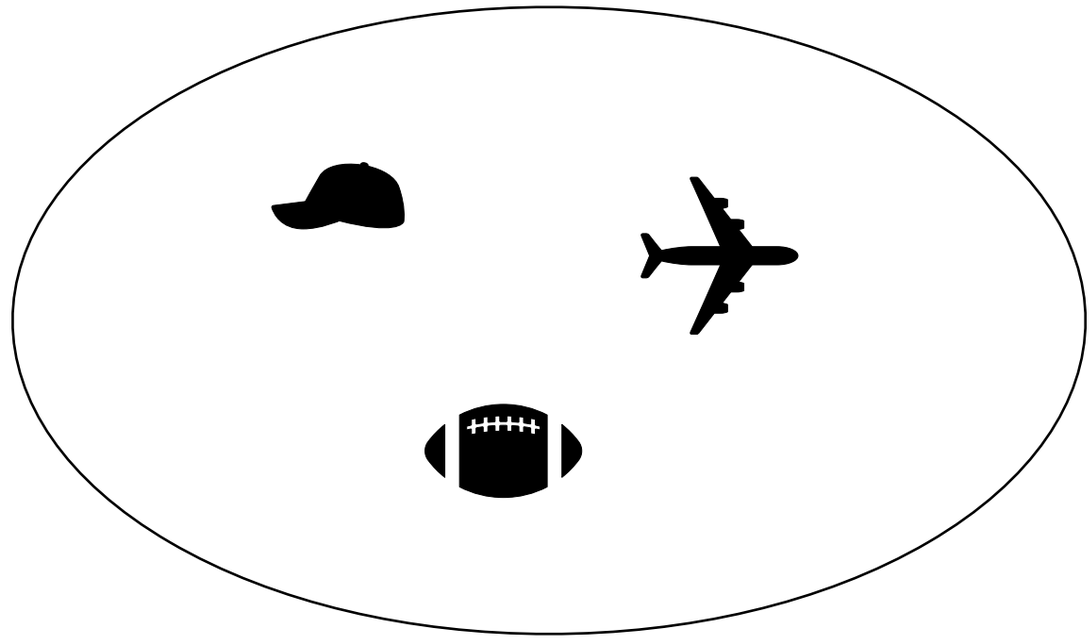


Non-Physical

The Universe of Logics

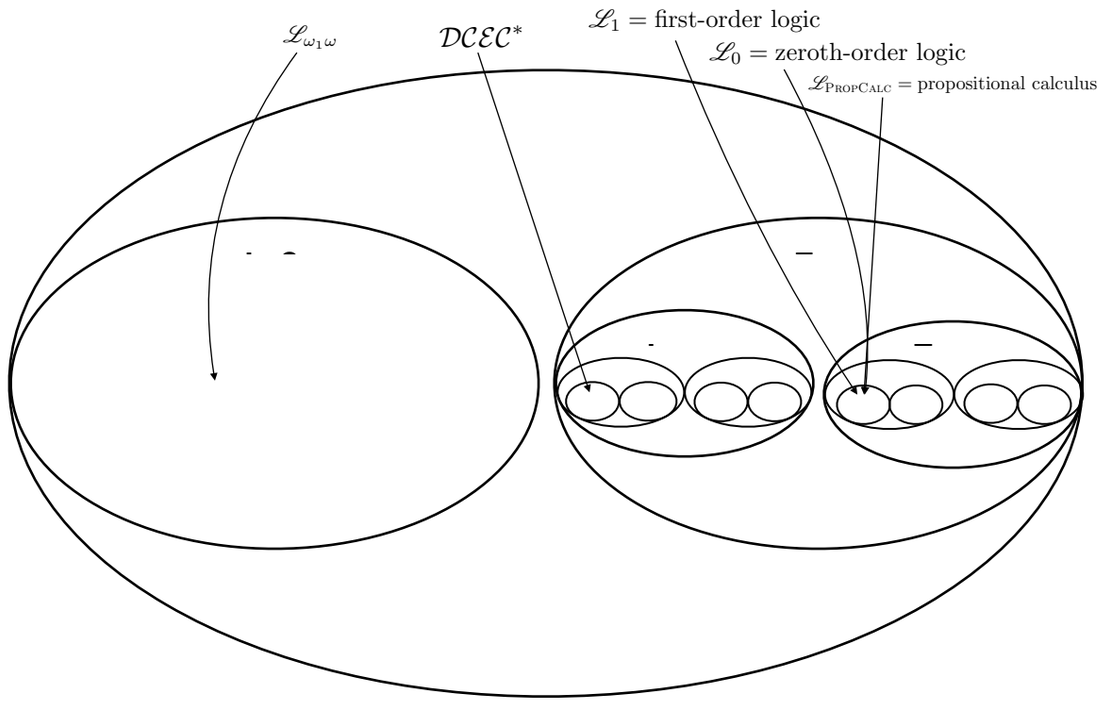


The Physical Universe

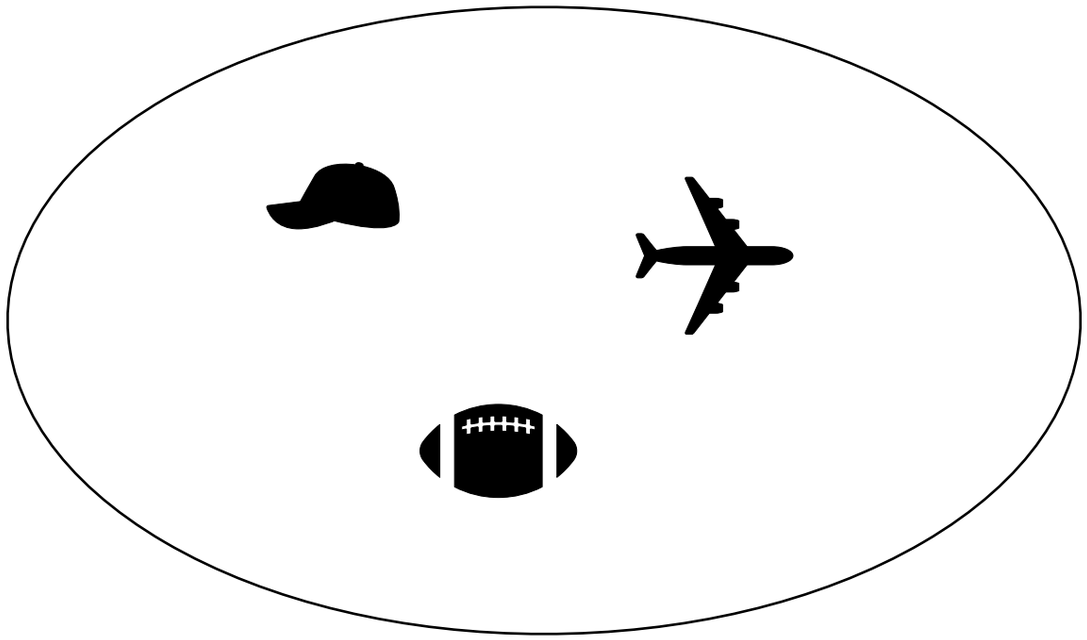


Non-Physical

The Universe of Logics

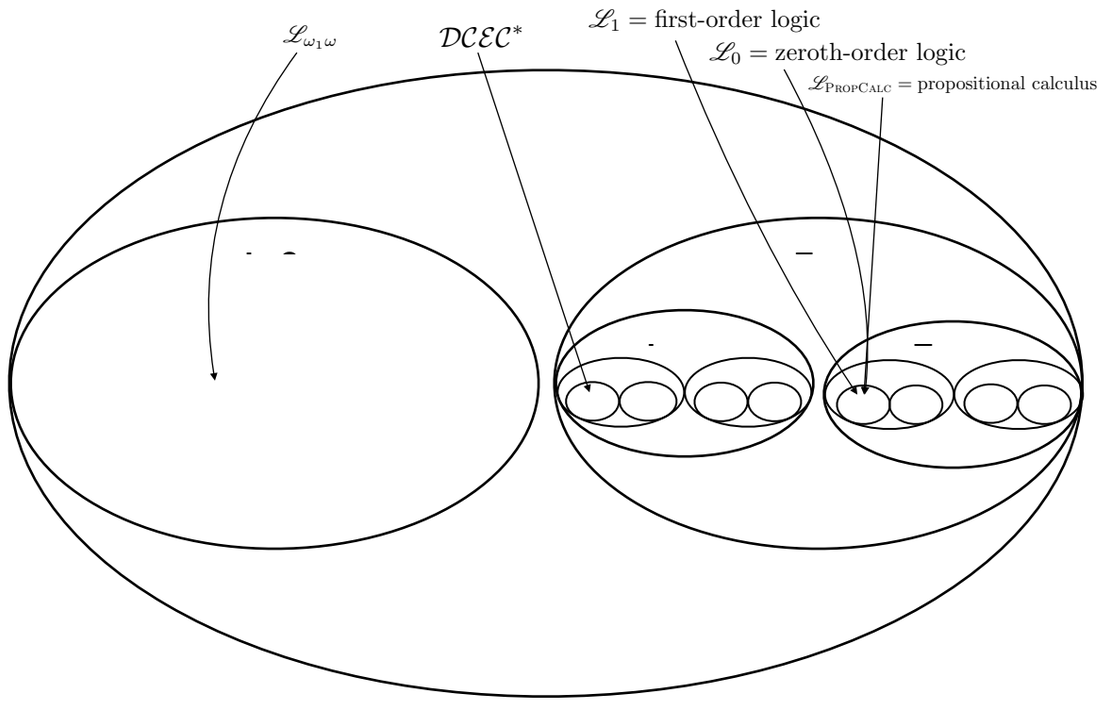


The Physical Universe

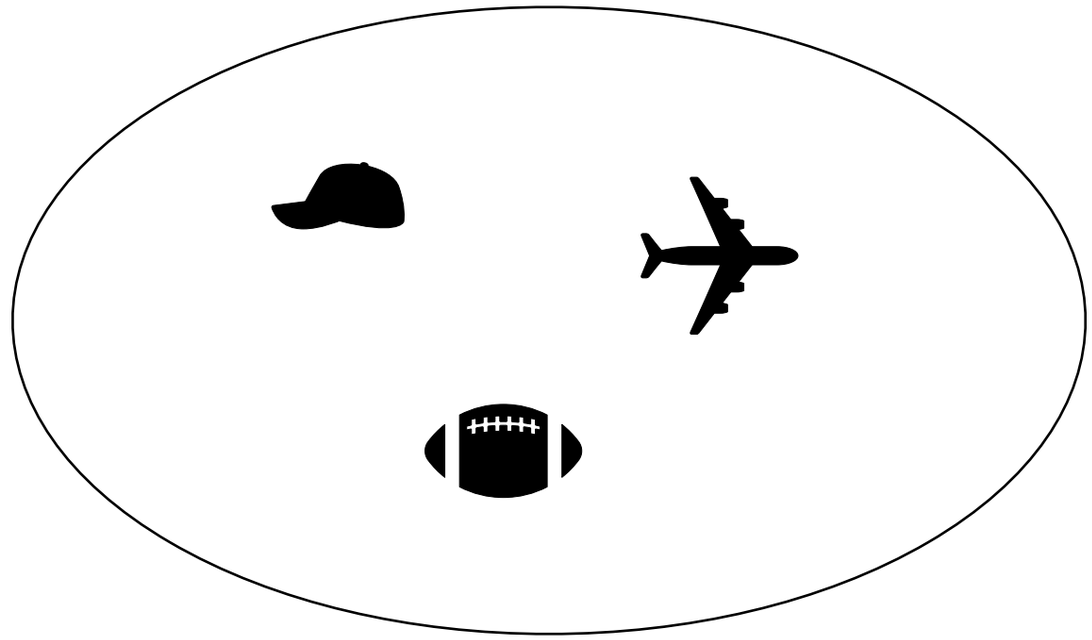


Non-Physical

The Universe of Logics



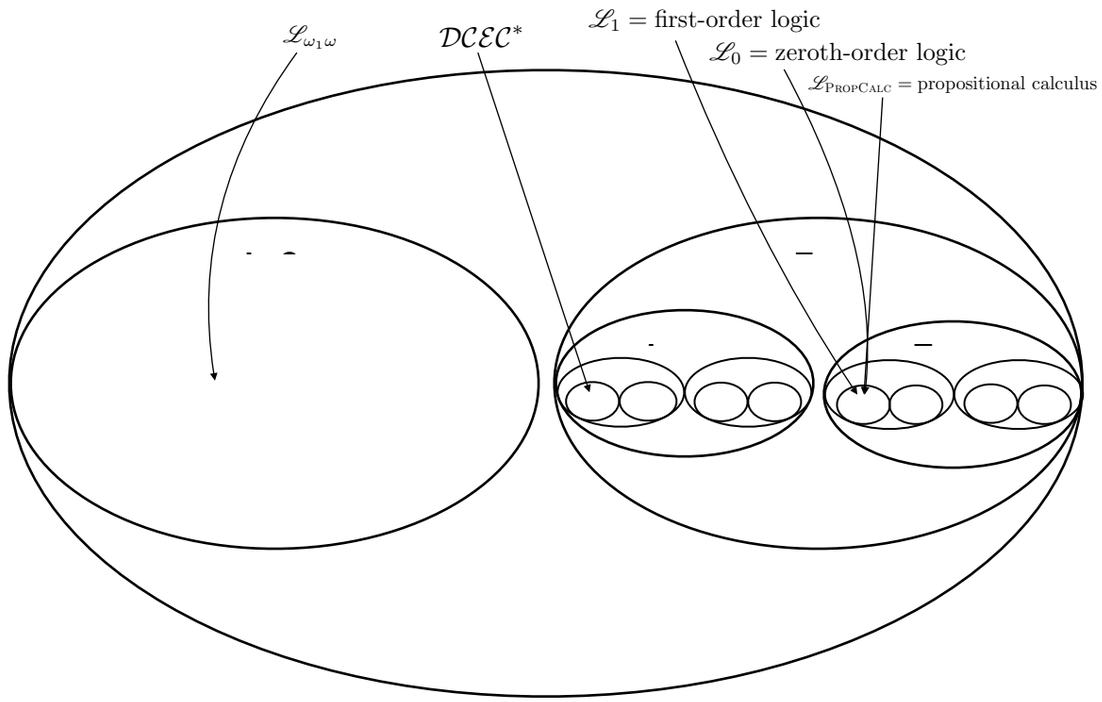
The Physical Universe



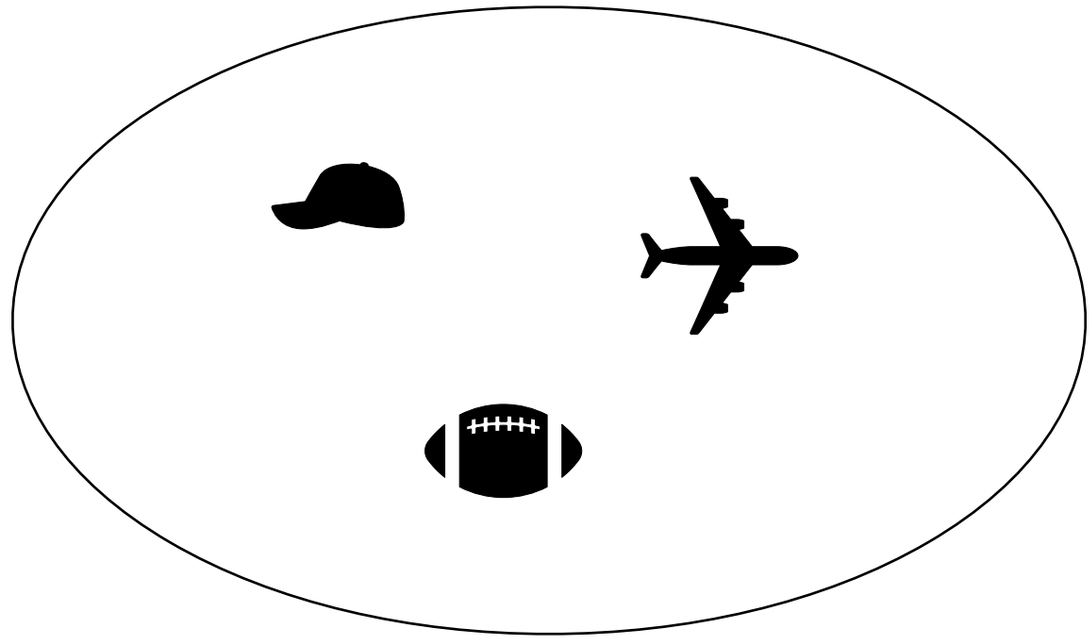
$$\frac{\phi, \phi \rightarrow \psi}{\psi} \quad \textit{modus ponens}$$

Non-Physical

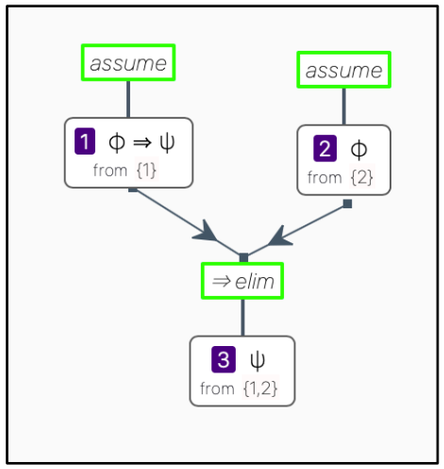
The Universe of Logics



The Physical Universe

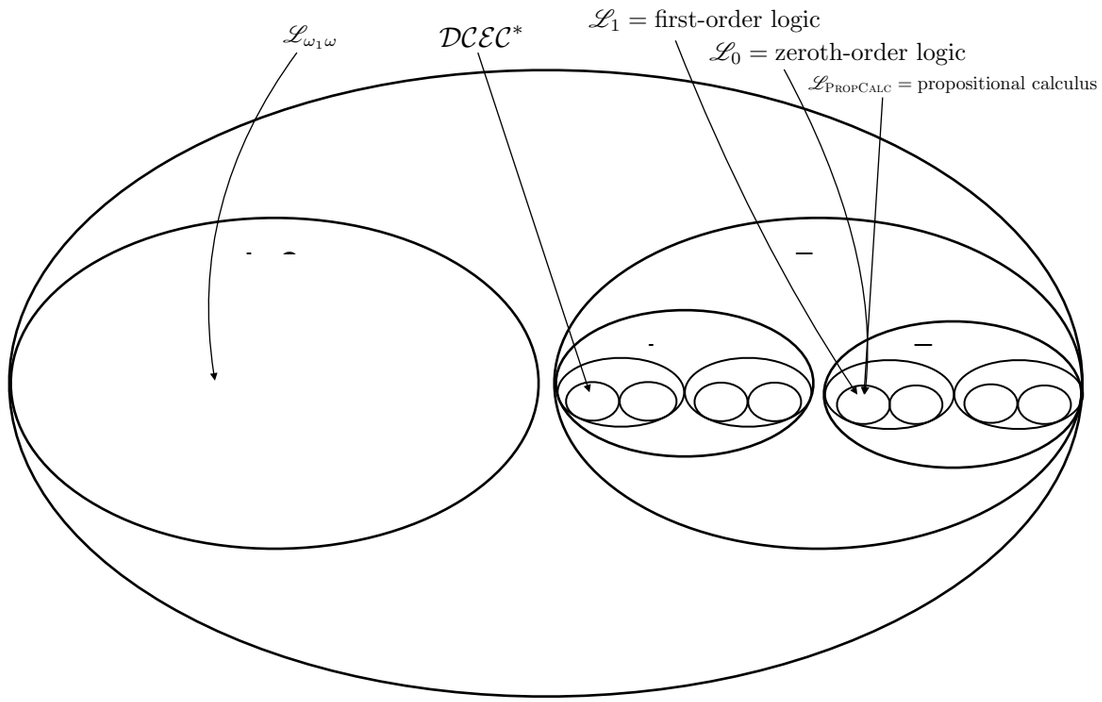


$$\frac{\phi, \phi \rightarrow \psi}{\psi}$$ *modus ponens*

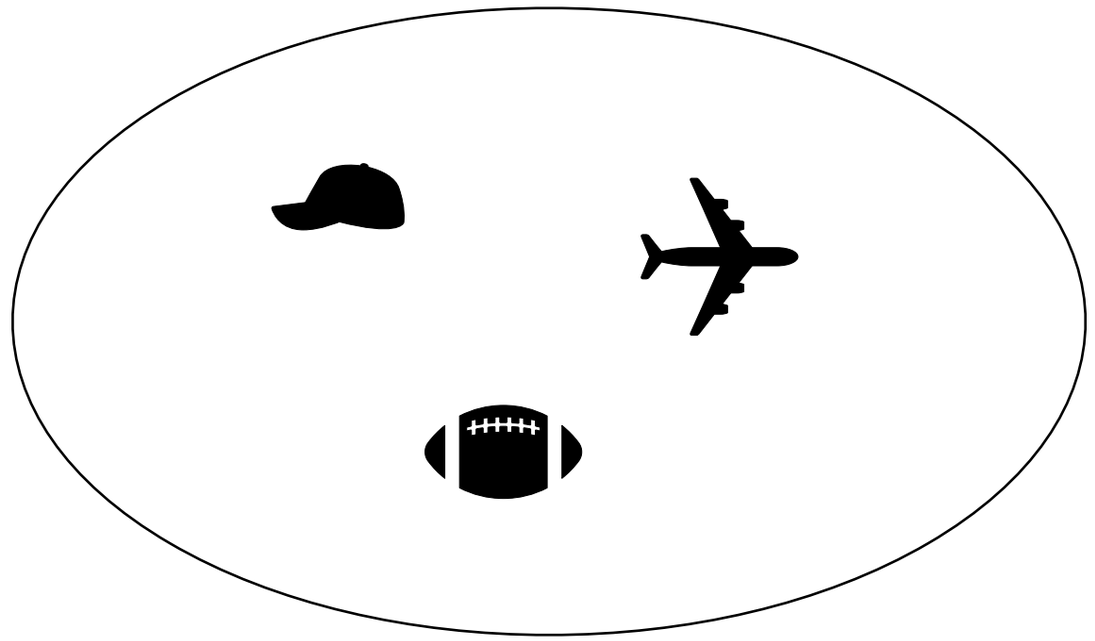


Non-Physical

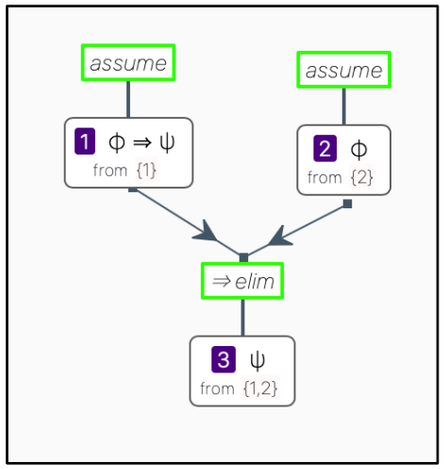
The Universe of Logics



The Physical Universe



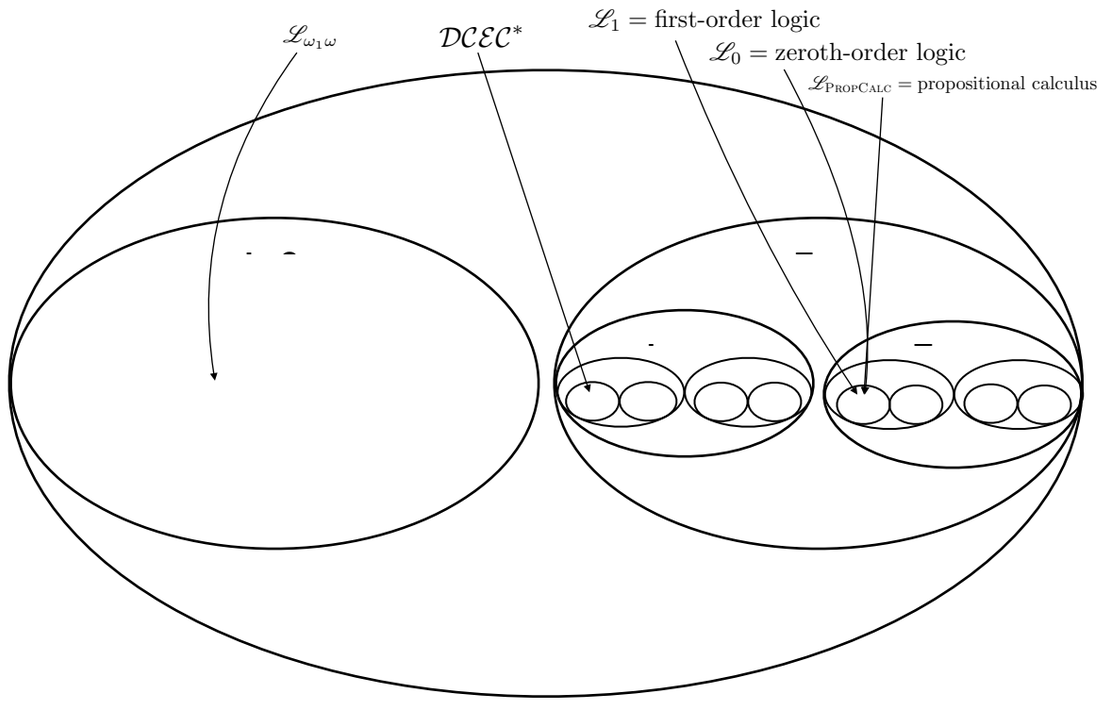
$$\frac{\phi, \phi \rightarrow \psi}{\psi} \quad \textit{modus ponens}$$



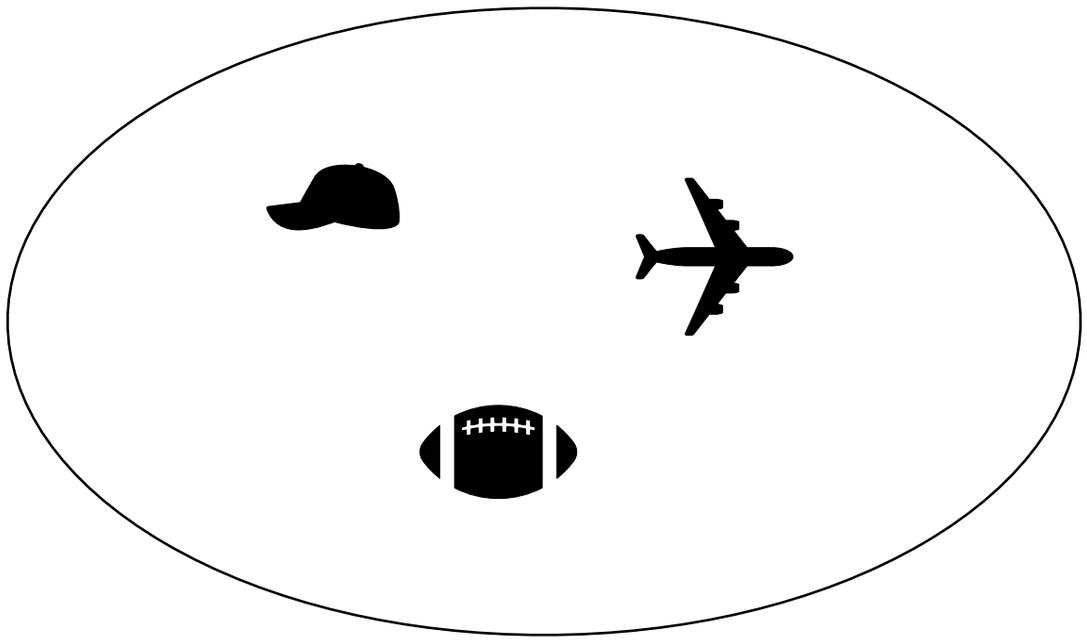
Ross, J. (1992) "Immaterial Aspects of Thought" *Journal of Philosophy* ***89.3***: 136–150.

Non-Physical

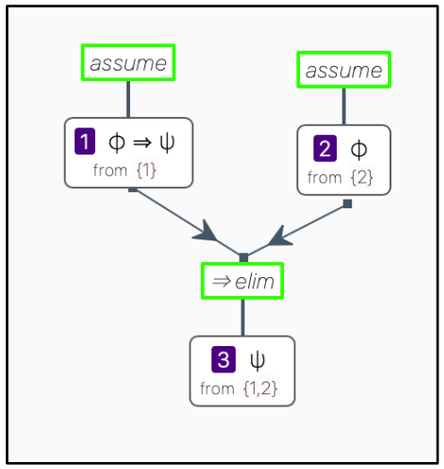
The Universe of Logics



The Physical Universe



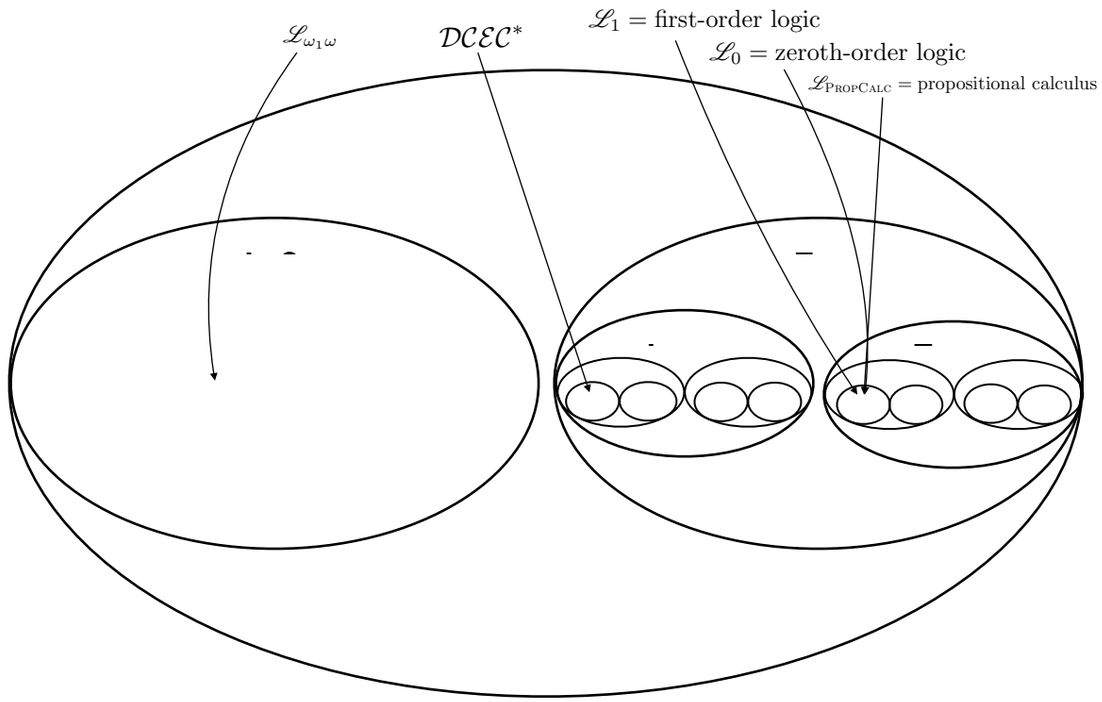
$$\frac{\phi, \phi \rightarrow \psi}{\psi} \quad \textit{modus ponens}$$



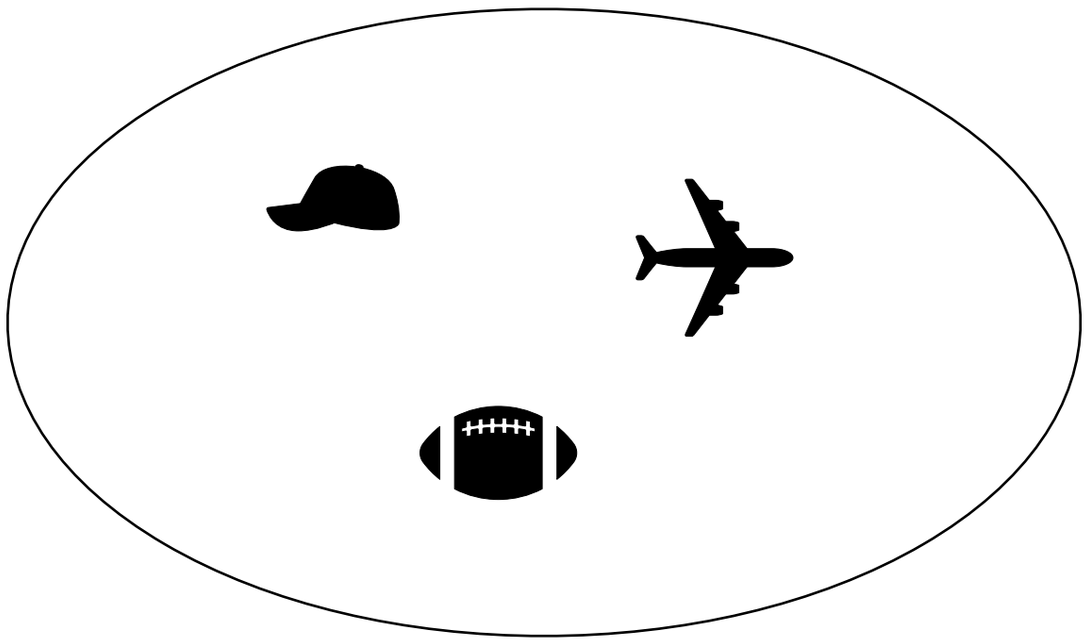
Ross, J. (1992) "Immaterial Aspects of Thought" *Journal of Philosophy* ***89.3***: 136–150.

Non-Physical

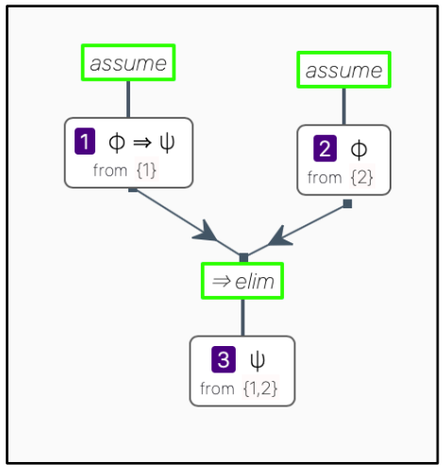
The Universe of Logics



The Physical Universe



$$\frac{\phi, \phi \rightarrow \psi}{\psi} \quad \textit{modus ponens}$$

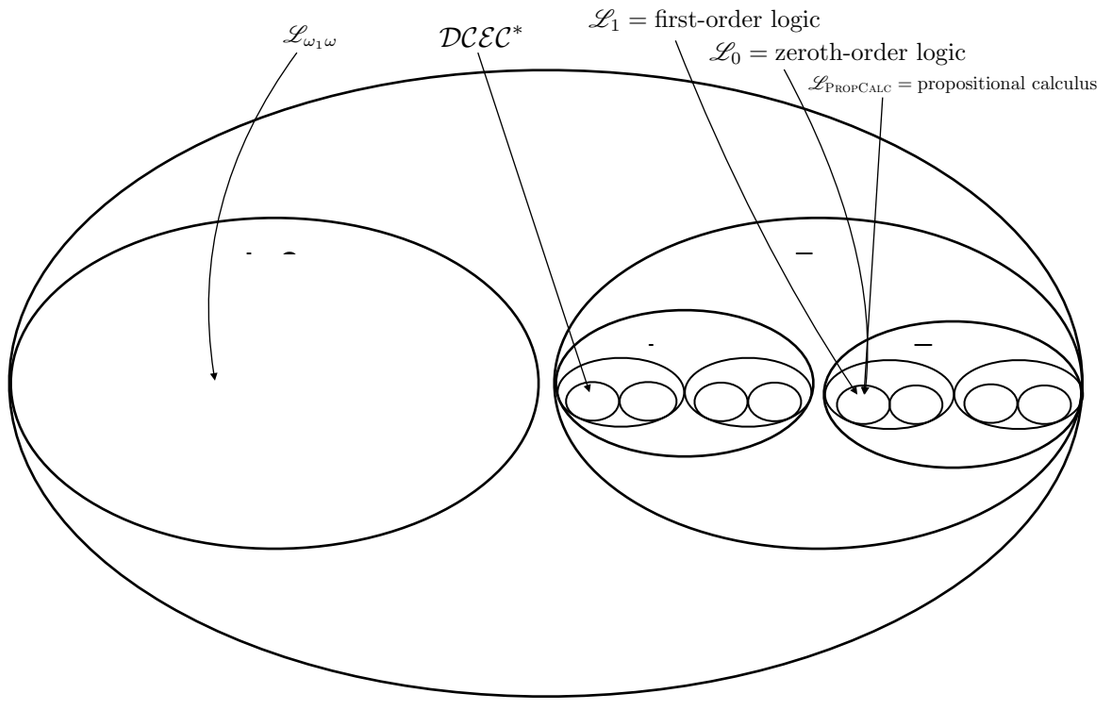


Ross, J. (1992) "Immaterial Aspects of Thought" *Journal of Philosophy* ***89.3***: 136–150.

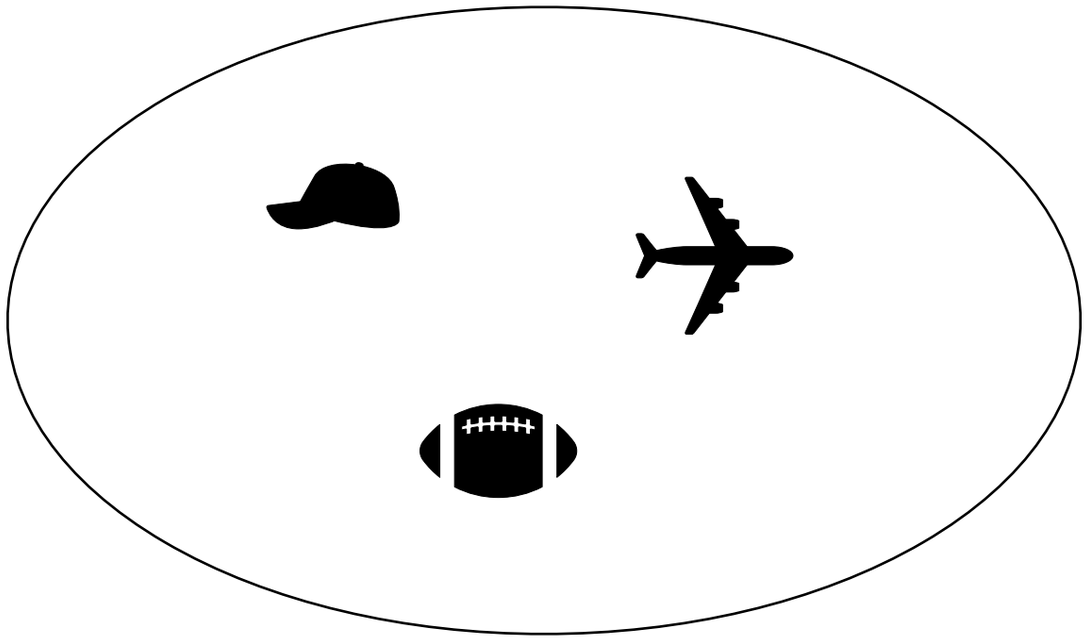
Non-Physical

?

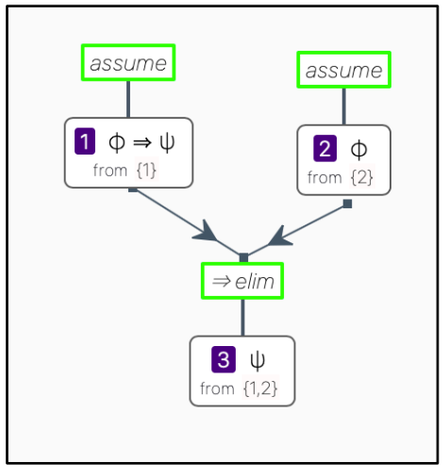
The Universe of Logics



The Physical Universe



$$\frac{\phi, \phi \rightarrow \psi}{\psi} \quad \textit{modus ponens}$$

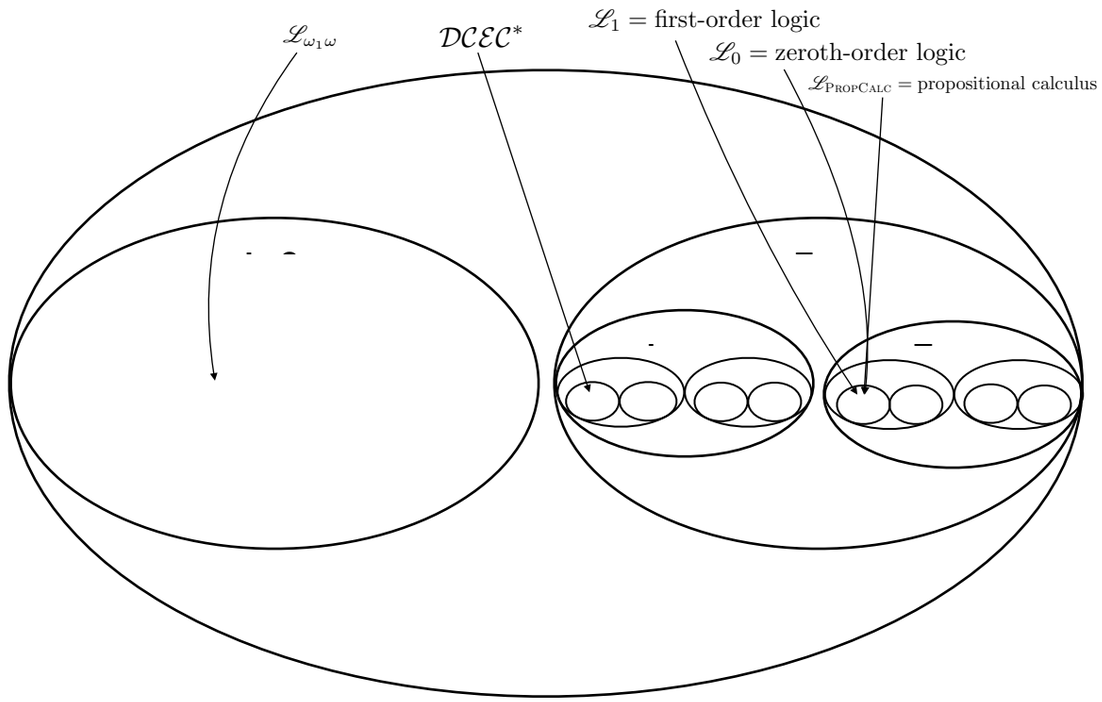


Ross, J. (1992) "Immaterial Aspects of Thought" *Journal of Philosophy* ***89.3***: 136–150.

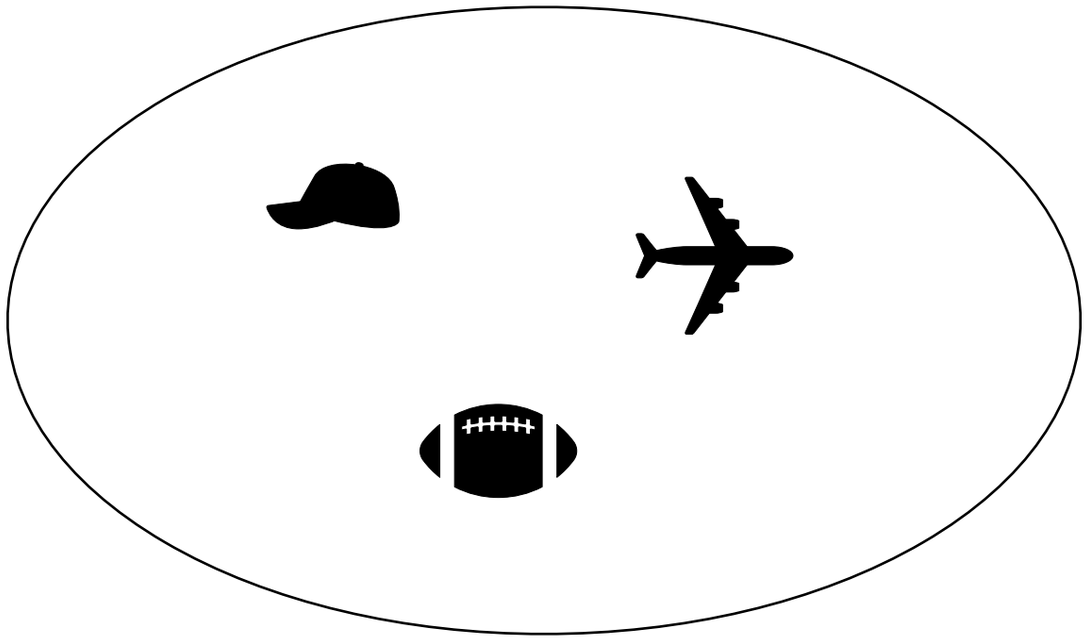
Non-Physical

?

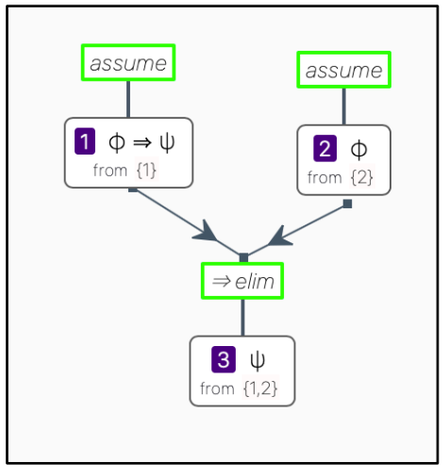
The Universe of Logics



The Physical Universe



$$\frac{\phi, \phi \rightarrow \psi}{\psi} \quad \textit{modus ponens}$$



Ross, J. (1992) "Immaterial Aspects of Thought" *Journal of Philosophy* ***89.3***: 136–150.

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

Background Claim

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

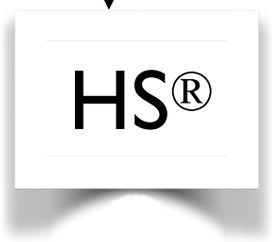
self-reference

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

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

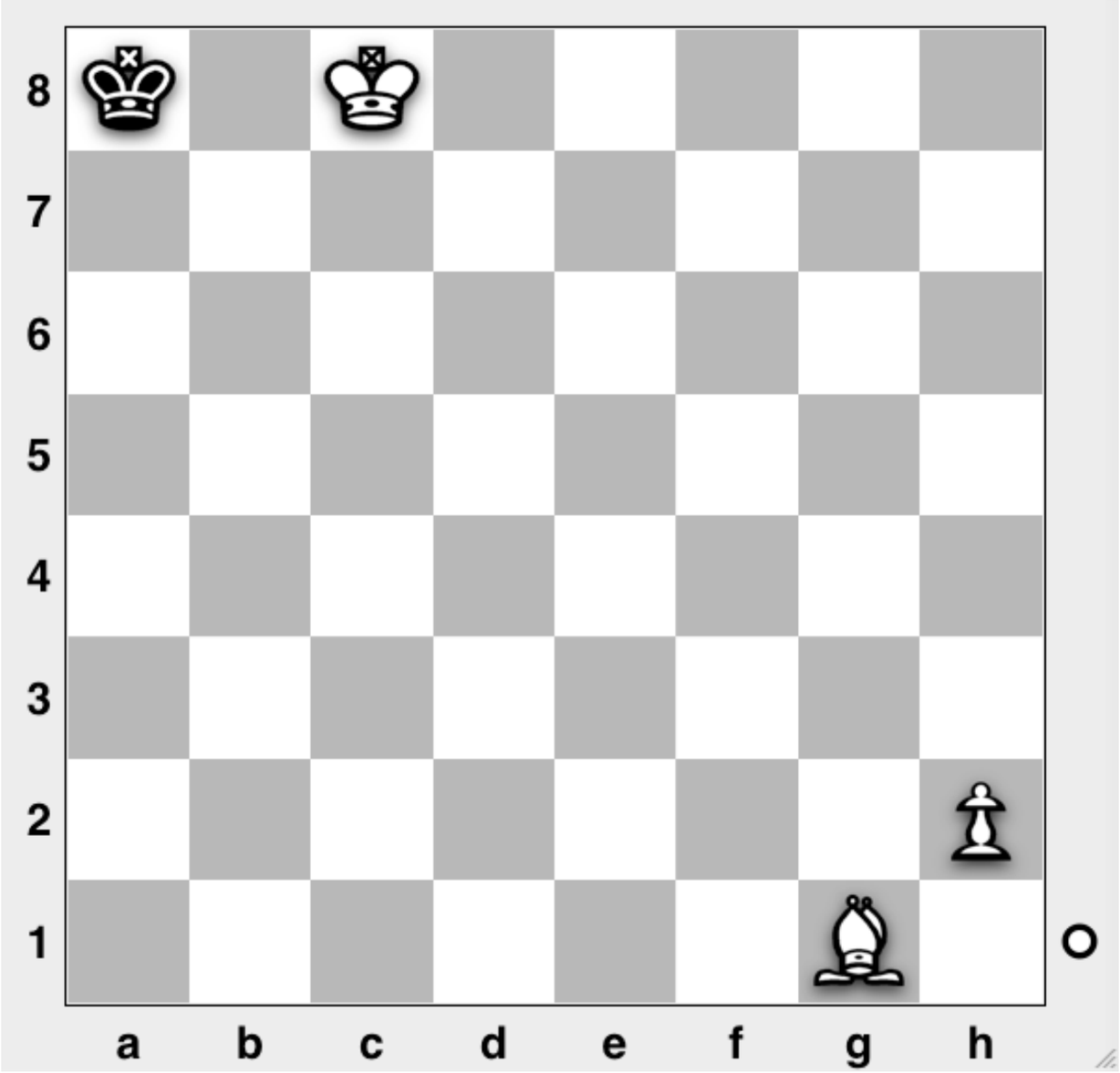
What *is* Logic?

- 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?

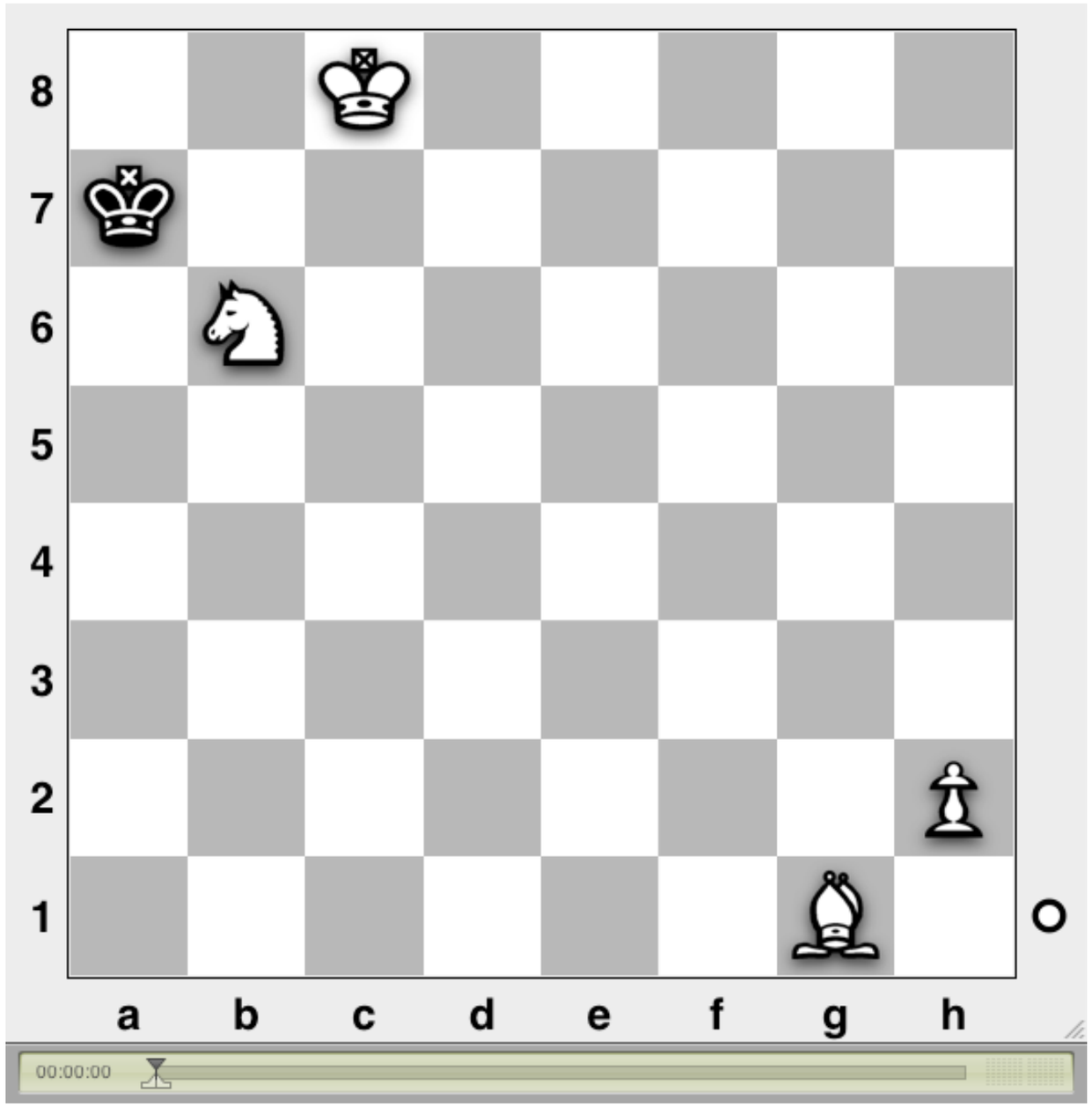
- The key to becoming rational. Or are you *already* rational? …
- "The science of reasoning." — so the not-unreasonable slogan goes.
- The only invincible subject there is.
- The basis for the formal sciences (from mathematics to game theory to decision theory to probability calculi to axiomatic physics ….) — 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)?…
- …

It's White's turn. What move did Black just make?



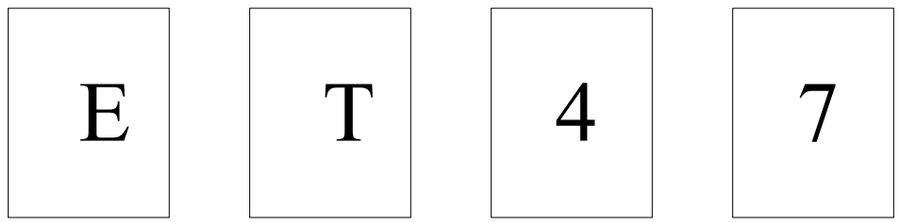
Aha! (Beyond Deep Blue?)

Aha! (Beyond Deep Blue?)



NOTE: Every card in this game has a capital Roman letter on one side, and a number from 1 to 9, inclusive.

Simple Selection Task



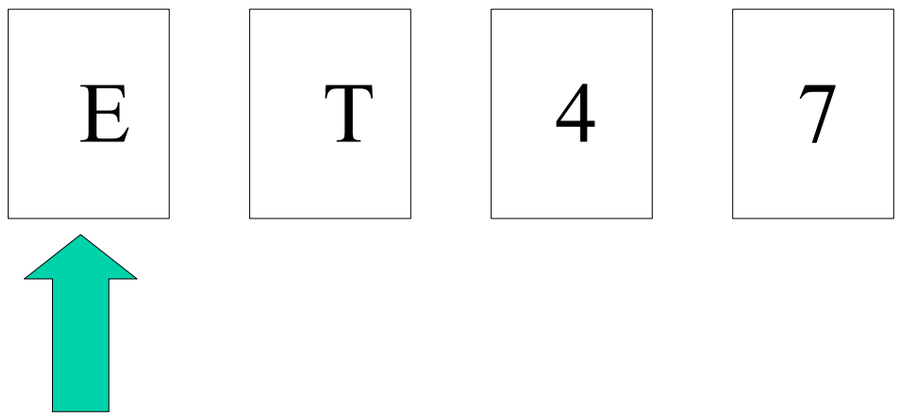
Suppose I claim that the following rule is true.

> If a card has a vowel on one side, it has an even number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

NOTE: Every card in this game has a capital Roman letter on one side, and a number from 1 to 9, inclusive.

Simple Selection Task



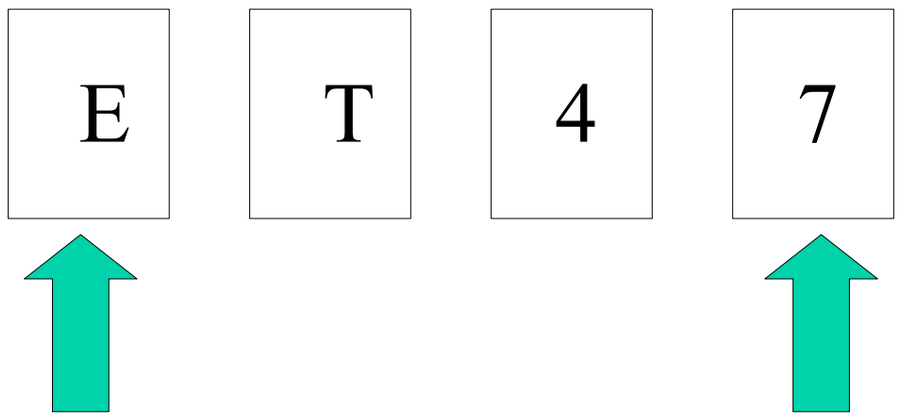
Suppose I claim that the following rule is true.

> If a card has a vowel on one side, it has an even number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

NOTE: Every card in this game has a capital Roman letter on one side, and a number from 1 to 9, inclusive.

Simple Selection Task

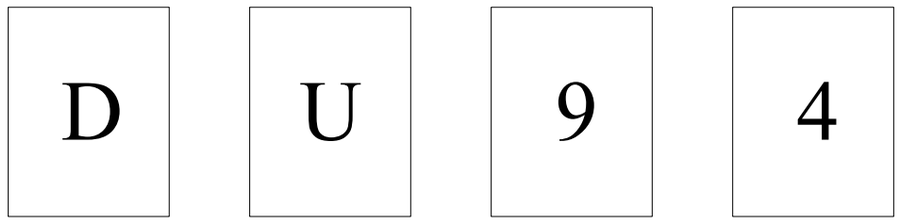


Suppose I claim that the following rule is true.

> If a card has a vowel on one side, it has an even number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

Another Simple Selection Task

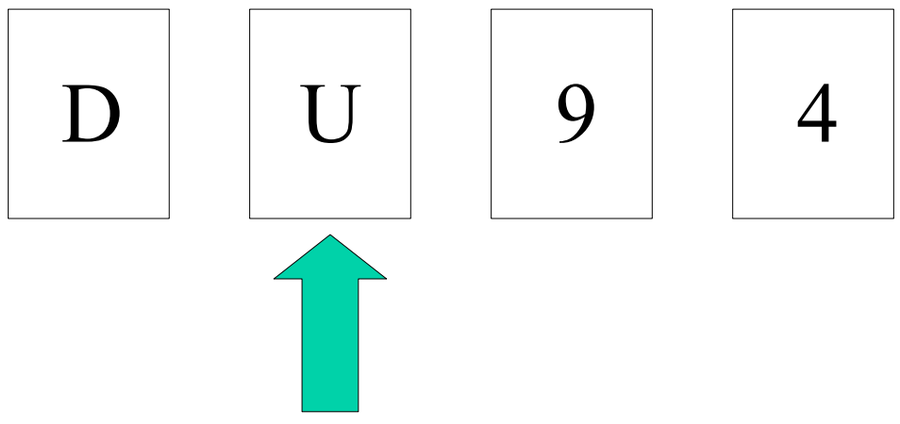


Suppose I claim that the following rule is true.

> If a card has a vowel on one side, it has an even number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

Another Simple Selection Task

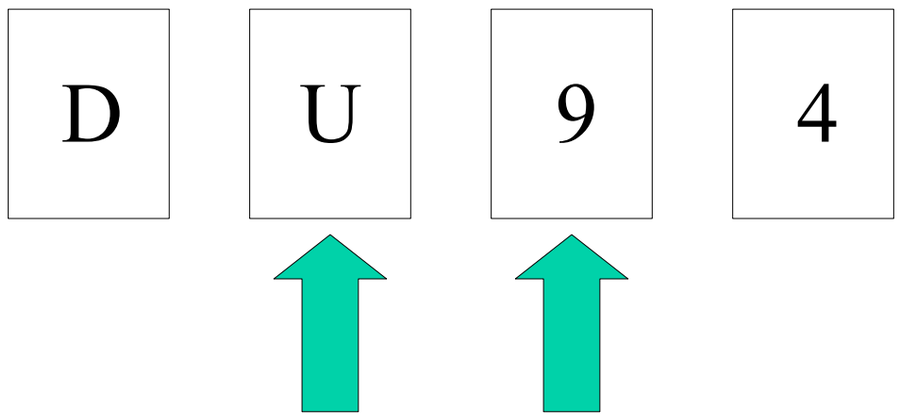


Suppose I claim that the following rule is true.

> If a card has a vowel on one side, it has an even number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

Another Simple Selection Task

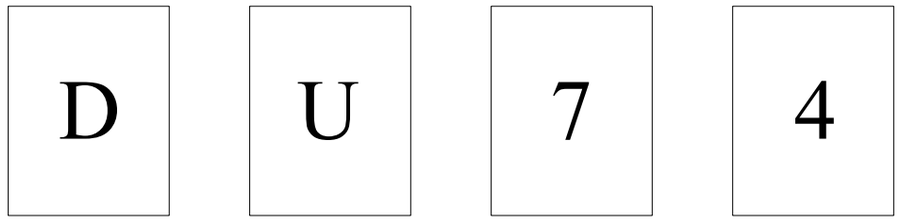


Suppose I claim that the following rule is true.

> If a card has a vowel on one side, it has an even number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

Another Simple Selection Task

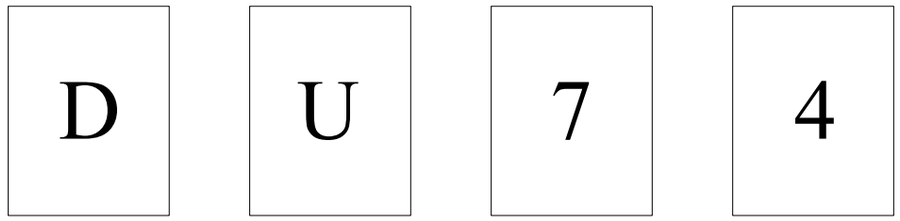


Suppose I claim that the following rule is true.

> If a card has a letter on one side, it has a prime number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

Another Simple Selection Task



Suppose I claim that the following rule is true.

> If a card has a letter on one side, it has a prime number on the other side.

Which card or cards, if any, should you turn over in order to try to efficiently decide whether the rule is true or false?

“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

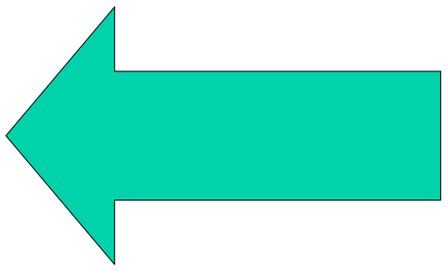
“NYS 1”

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?

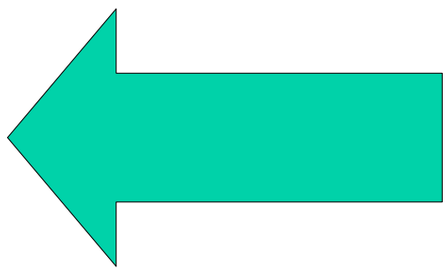
¬c

e

h

¬a

all of the above



The Original King-Ace

Suppose that the following premise is true:

> If there is a king in the hand, then there is an ace in the hand, or else if there isn't a king in the hand, then there is an ace.

What can you infer from this premise?

The Original King-Ace

Suppose that the following premise is true:

> If there is a king in the hand, then there is an ace in the hand, or else if there isn't a king in the hand, then there is an ace.

What can you infer from this premise?

> There is an ace in the hand.

The Original King-Ace

Suppose that the following premise is true:

> If there is a king in the hand, then there is an ace in the hand, or else if there isn't a king in the hand, then there is an ace.

What can you infer from this premise?

> ~~There is an ace in the hand.~~

The Original King-Ace

Suppose that the following premise is true:

> If there is a king in the hand, then there is an ace in the hand, or else if there isn't a king in the hand, then there is an ace.

What can you infer from this premise?

NO! ~~There is an ace in the hand.~~

The Original King-Ace

Suppose that the following premise is true:

> If there is a king in the hand, then there is an ace in the hand, or else if there isn't a king in the hand, then there is an ace.

What can you infer from this premise?

NO! ~~There is an ace in the hand.~~ NO!

The Original King-Ace

Suppose that the following premise is true:

> If there is a king in the hand, then there is an ace in the hand, or else if there isn't a king in the hand, then there is an ace.

What can you infer from this premise?

NO! ~~There is an ace in the hand.~~ NO!

In fact, what you *can* infer is that there *isn't* an ace in the hand!

King-Ace 2

Suppose that the following premise is true:

If there is a king in the hand, then there is an ace in the hand; or if there isn't a king in the hand, then there is an ace; but not both of these if-then statements are true.

What can you infer from this premise?

King-Ace 2

Suppose that the following premise is true:

If there is a king in the hand, then there is an ace in the hand; or if there isn't a king in the hand, then there is an ace; but not both of these if-then statements are true.

What can you infer from this premise?

There is an ace in the hand.

King-Ace 2

Suppose that the following premise is true:

If there is a king in the hand, then there is an ace in the hand; or if there isn't a king in the hand, then there is an ace; but not both of these if-then statements are true.

What can you infer from this premise?

~~There is an ace in the hand.~~

King-Ace 2

Suppose that the following premise is true:

If there is a king in the hand, then there is an ace in the hand; or if there isn't a king in the hand, then there is an ace; but not both of these if-then statements are true.

What can you infer from this premise?

NO! ~~There is an ace in the hand.~~

King-Ace 2

Suppose that the following premise is true:

If there is a king in the hand, then there is an ace in the hand; or if there isn't a king in the hand, then there is an ace; but not both of these if-then statements are true.

What can you infer from this premise?

NO! ~~There is an ace in the hand.~~ NO!

King-Ace 2

Suppose that the following premise is true:

If there is a king in the hand, then there is an ace in the hand; or if there isn't a king in the hand, then there is an ace; but not both of these if-then statements are true.

What can you infer from this premise?

NO! ~~There is an ace in the hand.~~ NO!

In fact, what you *can* infer is that there *isn't* an ace in the hand!

Informal Proofs/Arguments **vs.** Formal Proofs/Arguments

Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

no ambiguous natural language
(e.g. English or Chinese)

Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

no ambiguous natural language
(e.g. English or Chinese)

Informal Proofs/Arguments	vs.	Formal Proofs/Arguments
has ambiguous natural language (e.g. English or Chinese)		no ambiguous natural language (e.g. English or Chinese)
cannot be executed & checked by a computing machine		*can* be executed & checked by a computing machine

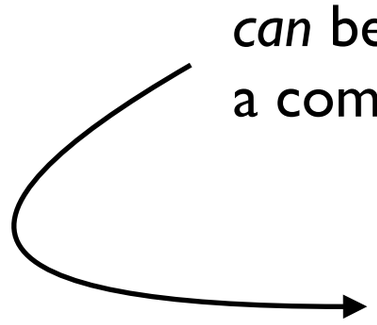
Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

no ambiguous natural language
(e.g. English or Chinese)

can be executed & checked by
a computing machine



Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

no ambiguous natural language
(e.g. English or Chinese)

can be executed & checked by
a computing machine

HyperSlate®

Informal Proofs/Arguments vs. Formal Proofs/Arguments

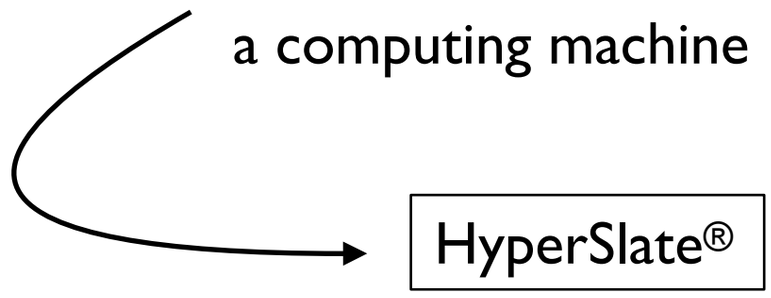
has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

supposed to have learned
how to produce, to an
appreciable degree, in High
School — but likely didn't

no ambiguous natural language
(e.g. English or Chinese)

can be executed & checked by
a computing machine



Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

supposed to have learned
how to produce, to an
appreciable degree, in High
School — but likely didn't

no ambiguous natural language
(e.g. English or Chinese)

can be executed & checked by
a computing machine

HyperSlate®

have not learned how to
produce in a relevant
system (though may have
had some Prolog)

FOR NOW

Informal Proofs/Arguments vs. Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

supposed to have learned
how to produce, to an
appreciable degree, in High
School — but likely didn't

no ambiguous natural language
(e.g. English or Chinese)

can be executed & checked by
a computing machine

HyperSlate®

have not learned how to
produce in a relevant
system (though may have
had some Prolog)

FOR NOW

STARTING ≥ 1/26/23

Informal Proofs/Arguments

vs.

Formal Proofs/Arguments

has ambiguous natural language
(e.g. English or Chinese)

cannot be executed & checked
by a computing machine

supposed to have learned
how to produce, to an
appreciable degree, in High
School — but likely didn't

no ambiguous natural language
(e.g. English or Chinese)

can be executed & checked by
a computing machine

HyperSlate®

have not learned how to
produce in a relevant
system (though may have
had some Prolog)

King-Ace Solved

(informal proof)

Proposition: There is *not* an ace in the hand.

Proof: We know that at least one of the if-thens (i.e., at least one of the **conditionals**) is false. So we have two cases to consider, viz., that *K* => *A* is false, and that ¬*K* => *A* is false. Take first the first case; accordingly, suppose that *K* => *A* is false. Then it follows that *K* is true (since when a conditional is false, its antecedent holds but its consequent doesn't), and *A* is false. Now consider the second case, which consists in ¬*K* => *A* being false. Here, in a direct parallel, we know ¬*K* and, once again, ¬*A*. In both of our two cases, which are exhaustive, there is no ace in the hand. The proposition is established. **QED**

Bringsjord 1

(1) The following three assertions are either all true or all false:

If Billy helped, Doreen helped.

If Doreen helped, Frank did as well.

If Frank helped, so did Emma.

(2) The following assertion is definitely true: Billy helped.

Can it be inferred from (1) and (2) that Emma helped?

Bringsjord 1

(1) The following three assertions are either all true or all false:

If Billy helped, Doreen helped.

If Doreen helped, Frank did as well.

If Frank helped, so did Emma.

(2) The following assertion is definitely true: Billy helped.

Can it be inferred from (1) and (2) that Emma helped?

YUP! — & now prove it!

Bringsjord 1: Proof

Proof: We have two cases to work from: when the conditionals in (1) are all true, and when they are all false. (In both cases, (2) remains true, and available.) So assume Case 1 first. In this case, we can simply chain through the conditionals by repeated application of *modus ponens* to arrive at the conclusion that Emma helped. Now assume Case 2 holds. This immediately implies that the first two conditionals are false; i.e., we have ~(B =>D) and ~(D => F). Recalling that a conditional fails to hold exactly when its antecedent is true while its consequent is false, we have, in turn: B & ~D, and D & ~F. But then we have a contradiction, viz. ~D & D. Since everything follows ("explosively"!) from a contradiction, we are done. **QED**

The Ticking-Bomb Logic Puzzle …

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?)

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.

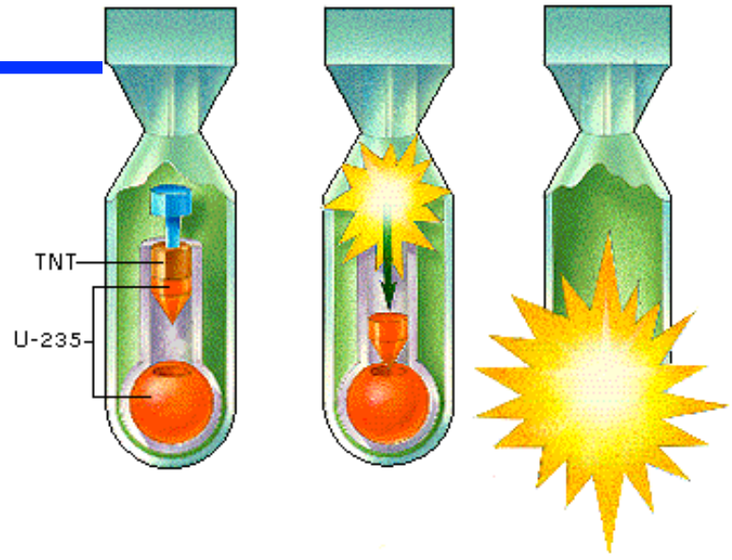
Given this perfectly reliable clue from Dr Moriarty, if either wire is more likely to run to the bomb, that wire *do*es run to the bomb, and the bomb is ticking, with only a minute left! If both are equiprobable, neither runs to the bomb, and you are powerless. Make your prediction as to what will happen 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!

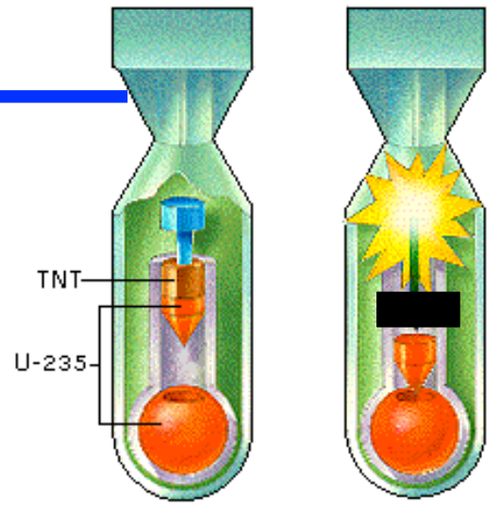
Proof: (1) can be treated as a biconditional, obviously (R <=> B).

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 …

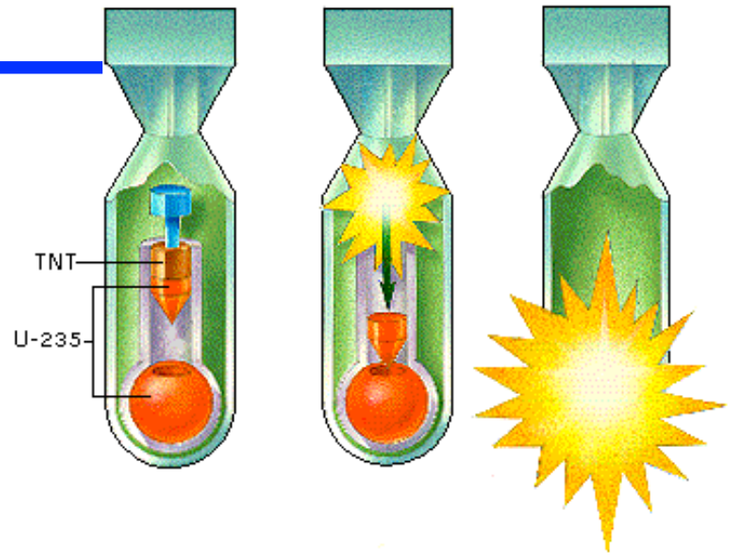
Proposition: The blue wire is more likely!

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

MHP3 (& Random MHP3) …

The Monty Hall Problem

$IM

The Monty Hall Problem

$1M

The Monty Hall Problem

$1M

The Monty Hall Problem

$1M

The Monty Hall Problem

$1M

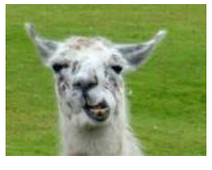
The Monty Hall Problem

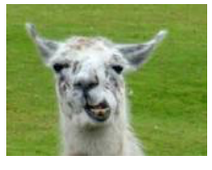
The Monty Hall Problem

$1M

The Monty Hall Problem

$1M





MHP3 Defined

Jones has come to a game show, and finds himself thereon selected to play a game on national TV with the show's suave host, Full Monty. Jones is told correctly by Full that hidden behind one of three closed, opaque doors facing the two of them is $1,000,000, while behind each of the other two is a feculent, obstreperous llama whose value on the open market is charitably pegged at $1. Full reminds Jones that this is a game, and a fair one, and that if Jones ends up selecting the door with $1M behind it, all that money will indeed be his. (Jones' net worth has nearly been exhausted by his expenditures in traveling to the show.) Full also reminds Jones that he (= Full) knows what's behind each door, fixed in place until the game ends.

Full asks Jones to select which door he wants the contents of. Jones says, "Door 1." Full then says: "Hm. Okay. Part of this game is my revealing at this point what's behind one of the doors you didn't choose. So ... let me show you what's behind Door 3." Door 3 opens to reveal a very unsavory llama. Full now to Jones: "Do you want to switch to Door 2, or stay with Door 1? You'll get what's behind the door of your choice, and our game will end." Full looks briefly into the camera, directly.

(P1.1) What should Jones do if he's rational?

(P1.2) Prove that your answer is correct. (Diagrammatic proofs are allowed.)

(P1.3) A quantitative hedge fund manager with a PhD in finance from Harvard zipped this email off to Full before Jones made his decision re. switching or not: "Switching would be a royal waste of time (and time is money!). Jones hasn't a doggone clue what's behind Door 1 or Door 2, and it's obviously a 50/50 chance to win whether he stands firm or switches. So the chap shouldn't switch!" Is the fund manager right? Prove that your diagnosis is correct.

(P1.4) Can these answers and proofs be exclusively Bayesian in nature?

FMHP3 Defined

Recall our definition of MHP3. We now define Random MHP3 (= RMHP3). In RMHP3, after the the contestant announces his initial choice of a door, a random number generator G generates either 1, 2, or 3, and then immediately thereafter the door with that number opens. (G takes no input in from prior events; it just generates pure and simple.) On one particular day, after contestant Jones has made his initial choice of Door 1 in what he expects to be a normal episode in conformity with MHP, but before Monty reveals a door and asks Jones to hold or switch, the random number generator generates 'Door 3' and that door suddenly opens to show a llama. That this has happened is explained to both Jones and Full Monty by one of the show's producers immediately after Door 3 opens. The producer then gives Jones the option of holding or switching. The producer explains that every now and then this unpredictably happens on the show, and when it does (as long as the door the contestant has chosen isn't opened because of G's output) the show must go on.

Is a switching policy still rational to affirm before an agent is confronted with an instance of RMHP3, in a parallel of what's rational in the case of MHP3? Why or why not, exactly?

Logic kan redde menneskehten!