Introduction to Inductive Logic
(via Monty Hall, the Grue Paradox, the Lottery Paradox, and AI for drone strikes)

Selmer Bringsjord

Rensselaer AI & Reasoning (RAIR) Lab
Department of Cognitive Science
Department of Computer Science
Lally School of Management & Technology
Rensselaer Polytechnic Institute (RPI)
Troy, New York 12180 USA

IFLAI 2022
4/1/2024

Rensselaer AI and Reasoning Lab

R&D enabled by:

Re HyperGrader® Content

- April ~~18~~ 23 approaching …
  - Two modal-logic problems now in **Required** category.
- TLB's hardest problem now published. (Others, e.g. Dreadsbury, forthcoming.)
  - This is **Supererogatory Challenge** problem. Attempt only if you're a self-measuring hotshot on top of all your other classes. No hints have ever been/will be provided for **S-C** problems.
  - May 1 will be the last day assignments can be done. Test 3 will be due 11:59p that day.

Deductive Logic vs *In*ductive Logic …

Simple Specimens to Convey the Distinction

the hallmark of deductive logic is *proof*, the hallmark of inductive logic is the concept of an *argument*. An exceptionally strong kind of argument is a proof, but plenty of arguments fall short of being proofs — and yet still have considerable force. For instance, consider the following argument α_1:

	(1)	Tweety is bird.
	(2)	Most birds can fly.
$\therefore$	(3)	Tweety can fly.

For start contrast, consider as well this argument (α_2):

	(1′)	3 is a positive integer.
	(2′)	All positive integers are greater than 0.
$\therefore$	(3′)	3 is greater than 0.

The second of these arguments qualifies as an outright proof. That is, using the notation much employed before the present chapter:

$$\{(1'),(2')\} \vdash (3')$$

But in stark contrast, argument α_1 is not a proof that Tweety can fly. The reason is obvious: (3) isn't deduced from the combination of (1) and (2); that is,

$$\{(1),(2)\} \nvdash (3)$$

Simple Specimens to Convey the Distinction

the hallmark of deductive logic is *proof*, the hallmark of inductive logic is the concept of an *argument*. An exceptionally strong kind of argument is a proof, but plenty of arguments fall short of being proofs — and yet still have considerable force. For instance, consider the following argument α_1:

	(1)	Tweety is bird.
	(2)	Most birds can fly.
∴	(3)	Tweety can fly.

For start contrast, consider as well this argument (α_2):

	(1′)	3 is a positive integer.
	(2′)	All positive integers are greater than 0.
∴	(3′)	3 is greater than 0.

The second of these arguments qualifies as an outright proof. That is, using the notation much employed before the present chapter:

$$\{(1'),(2')\} \vdash (3')$$

But in stark contrast, argument α_1 is not a proof that Tweety can fly. The reason is obvious: (3) isn't deduced from the combination of (1) and (2); that is,

$$\{(1),(2)\} \nvdash (3)$$

Simple Specimens to Convey the Distinction

the hallmark of deductive logic is *proof*, the hallmark of inductive logic is the concept of an *argument*. An exceptionally strong kind of argument is a proof, but plenty of arguments fall short of being proofs — and yet still have considerable force. For instance, consider the following argument α_1:

	(1)	Tweety is bird.
	(2)	Most birds can fly.
∴	(3)	Tweety can fly.

For start contrast, consider as well this argument (α_2):

	(1′)	3 is a positive integer.
	(2′)	All positive integers are greater than 0.
∴	(3′)	3 is greater than 0.

Deductive (familiar)

The second of these arguments qualifies as an outright proof. That is, using the notation much employed before the present chapter:

$$\{(1'), (2')\} \vdash (3')$$

But in stark contrast, argument α_1 is not a proof that Tweety can fly. The reason is obvious: (3) isn't deduced from the combination of (1) and (2); that is,

$$\{(1), (2)\} \nvdash (3)$$

Simple Specimens to Convey the Distinction

the hallmark of deductive logic is *proof*, the hallmark of inductive logic is the concept of an *argument*. An exceptionally strong kind of argument is a proof, but plenty of arguments fall short of being proofs — and yet still have considerable force. For instance, consider the following argument α_1:

	(1)	Tweety is bird.
	(2)	Most birds can fly.
∴	(3)	Tweety can fly.

For start contrast, consider as well this argument (α_2):

	(1′)	3 is a positive integer.
	(2′)	All positive integers are greater than 0.
∴	(3′)	3 is greater than 0.

Deductive (familiar)

The second of these arguments qualifies as an outright proof. That is, using the notation much employed before the present chapter:

$$\{(1'),(2')\} \vdash (3')$$

But in stark contrast, argument α_1 is not a proof that Tweety can fly. The reason is obvious: (3) isn't deduced from the combination of (1) and (2); that is,

$$\{(1),(2)\} \nvdash (3)$$

Simple Specimens to Convey the Distinction

the hallmark of deductive logic is *proof*, the hallmark of inductive logic is the concept of an *argument*. An exceptionally strong kind of argument is a proof, but plenty of arguments fall short of being proofs — and yet still have considerable force. For instance, consider the following argument α_1:

	(1)	Tweety is bird.
	(2)	Most birds can fly.
∴	(3)	Tweety can fly.

Inductive (new)

For start contrast, consider as well this argument (α_2):

	(1′)	3 is a positive integer.
	(2′)	All positive integers are greater than 0.
∴	(3′)	3 is greater than 0.

Deductive (familiar)

The second of these arguments qualifies as an outright proof. That is, using the notation much employed before the present chapter:

$$\{(1'),(2')\} \vdash (3')$$

But in stark contrast, argument α_1 is not a proof that Tweety can fly. The reason is obvious: (3) isn't deduced from the combination of (1) and (2); that is,

$$\{(1),(2)\} \nvdash (3)$$

Simple Specimens to Convey the Distinction

the hallmark of deductive logic is *proof*, the hallmark of inductive logic is the concept of an *argument*. An exceptionally strong kind of argument is a proof, but plenty of arguments fall short of being proofs — and yet still have considerable force. For instance, consider the following argument α_1:

	(1)	Tweety is bird.
	(2)	Most birds can fly.
∴	(3)	Tweety can fly.

Inductive (new)

For start contrast, consider as well this argument (α_2):

	(1′)	3 is a positive integer.
	(2′)	All positive integers are greater than 0.
∴	(3′)	3 is greater than 0.

Deductive (familiar)

The second of these arguments qualifies as an outright proof. That is, using the notation much employed before the present chapter:

$$\{(1'),(2')\} \vdash (3')$$

But in stark contrast, argument α_1 is not a proof that Tweety can fly. The reason is obvious: (3) isn't deduced from the combination of (1) and (2); that is,

$$\{(1),(2)\} \nvdash (3)$$

Same point can be made via variants using likelihood values, or probability values.

And for something much more robust and interesting/infamous …
The Monty Hall Problem teaches us that we need more than formal *de*ductive logic!

Those who fail are behaving irrationally:

Friedman, D. (1998) "Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly" *American Economic Review* **88**(4): 933–946.

- http://static.luiss.it/hey/ambiguity/papers/Friedman_1998.pdf

Those who fail are behaving irrationally:

Friedman, D. (1998) “Monty Hall’s Three Doors: Construction and Deconstruction of a Choice Anomaly” *American Economic Review* **88**(4): 933–946.

- http://static.luiss.it/hey/ambiguity/papers/Friedman_1998.pdf

Sometimes people make decisions that seem inconsistent with rational choice theory. We have a "choice anomaly" when such decisions are systematic and well documented. From a few isolated examples such as the Maurice Allais (1953) paradox and the probability matching puzzle of William K. Estes (1954), the set of anomalies expanded dramatically in the last two decades, especially following the work of Daniel Kahneman and Amos Tversky (e.g., 1979). By now the empirical literature offers dozens of interrelated anomalies documented in hundreds of articles and surveys (e.g., Colin F. Camerer, 1995).

Those who fail are behaving irrationally:

Friedman, D. (1998) "Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly" *American Economic Review* **88**(4): 933–946.

- http://static.luiss.it/hey/ambiguity/papers/Friedman_1998.pdf

Sometimes people make decisions that seem inconsistent with rational choice theory. We have a "choice anomaly" when such decisions are systematic and well documented. From a few isolated examples such as the Maurice Allais (1953) paradox and the probability matching puzzle of William K. Estes (1954), the set of anomalies expanded dramatically in the last two decades, especially following the work of Daniel Kahneman and Amos Tversky (e.g., 1979). By now the empirical literature offers dozens of interrelated anomalies documented in hundreds of articles and surveys (e.g., Colin F. Camerer, 1995).

Anomalies?? You mean irrational decisions?

Don’t Trust the Popular Media!

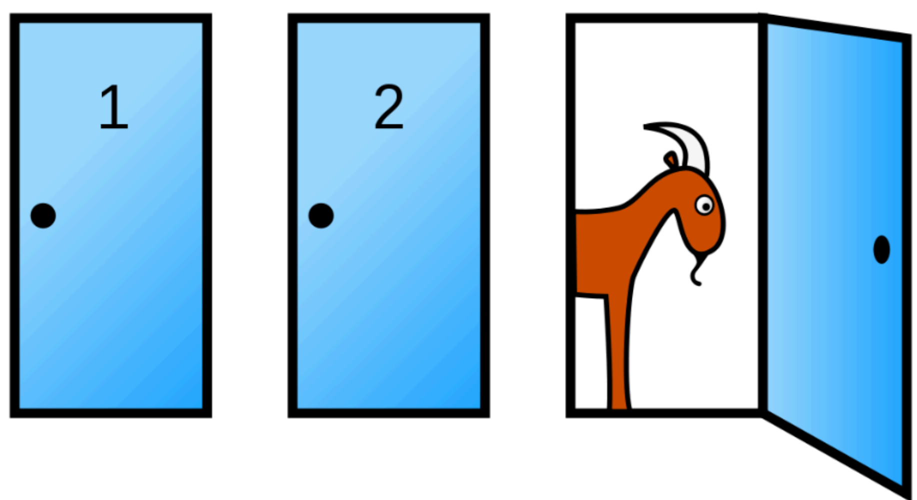
https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don’t Trust the Popular Media!

Monty Hall, Erdos, and Our Limited Minds

SAMUEL ARBESMAN SCIENCE 11.26.14 10:00 AM

Monty Hall, Erdos, and Our Limited Minds



https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don’t Trust the Popular Media!

Monty Hall, Erdos, and Our Limited Minds

THE MONTY HALL problem is a well-known mathematical brainteaser. But I find it intriguing not for how to solve it, but for how widespread having trouble with it is.

Based off of a television game show, the Monty Hall problem begins with a contestant finding herself in front of three doors. She is told that behind one of them is a car, while behind the other two there are goats. Since it is presumed that contestants want to win cars not goats, if nothing else for their resale value, there is a one-third chance of choosing the car and winning.

But now here’s the twist. After the contestant chooses a door, the game show host has another door opened and the contestant is shown a goat. Should she stick with the door she has originally chosen, or switch to the remaining unopened door?

There are many ways to examine this, but it turns out that it is always better to switch. Many people assume that the probability remains the same—it’s fifty-fifty so switching doesn’t matter—but they are wrong. There is a higher probability of the car being behind the door when you switch (here is a detailed discussion but I like to think about it based on an extreme version, one with 100 doors. One has a car and the others all have goats. You choose a door. The host opens 98 other doors, showing all goats. Should you switch? Of course! The host has done the work of almost certainly finding of the car for you.)

Anyway, I'm not concerned with the particulars of the problem but rather with how people respond to it. Namely, many listeners, even highly-trained mathematicians, are initially confused by the probabilities. In fact, until I learned of the extreme version with 100 doors, I didn't really understand why switching is better either.

https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don’ … edia!

that behind one of them is a car, while behind the other two there are goats. Since it is presumed that contestants want to win cars not goats, if nothing else for their resale value, there is a one-third chance of choosing the car and winning.

But now here’s the twist. After the contestant chooses a door, the game show host has another door opened and the contestant is shown a goat. Should she stick with the door she has originally chosen, or switch to the remaining unopened door?

There are many ways to examine this, but it turns out that it is always better to switch. Many people assume that the probability remains the same—it’s fifty-fifty so switching doesn’t matter—but they are wrong. There is a higher probability of the car being behind the door when you switch (here is a detailed discussion but I like to think about it based on an extreme version, one with 100 doors. One has a car and the others all have goats. You choose a door. The host opens 98 other doors, showing all goats. Should you switch? Of course! The host has done the work of almost certainly finding of the car for you.)

Anyway, I’m not concerned with the particulars of the problem but rather with how people respond to it. Namely, many listeners, even highly-trained mathematicians, are initially confused by the probabilities. In fact, until I learned of the extreme version with 100 doors, I didn't really understand why switching is better either.

https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don' edia!

that behind one of them is a car, while behind the other two there are goats. Since it is presumed that contestants want to win cars not goats, if nothing else for their resale value, there is a one-third chance of choosing the car and winning.

But now here's the twist. After the contestant chooses a door, the game show host has another door opened and the contestant is shown a goat. Should she stick with the door she has originally chosen, or switch to the remaining unopened door?

There are many ways to examine this, but it turns out that it is always better to switch. Many people assume that the probability remains the same—it's fifty-fifty so switching doesn't matter—but they are wrong. There is a higher probability of the car being behind the door when you switch (here is a detailed discussion but I like to think about it based on an extreme version, one with 100 doors. One has a car and the others all have goats. You choose a door. The host opens 98 other doors, showing all goats. Should you switch? Of course! The host has done the work of almost certainly finding of the car for you.)

Painful!

Anyway, I'm not concerned with the particulars of the problem but rather with how people respond to it. Namely, many listeners, even highly-trained mathematicians, are initially confused by the probabilities. In fact, until I learned of the extreme version with 100 doors, I didn't really understand why switching is better either.

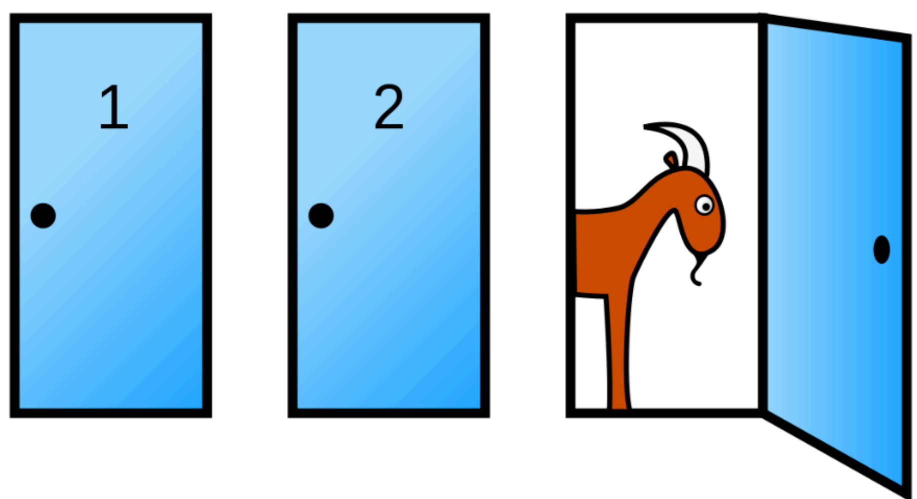
https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don’t Trust Lazy Mathematicians!

Monty Hall, Erdos, and Our Limited Minds

SAMUEL ARBESMAN SCIENCE 11.26.14 10:00 AM

Monty Hall, Erdos, and Our Limited Minds



https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don’t Trust Lazy Mathematicians!

Monty Hall, Erdos, and Our Limited Minds

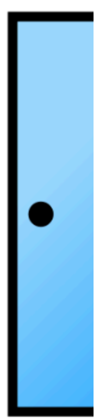
In fact, Paul Erdős, one of the most prolific and foremost mathematicians involved in probability, when initially told of the Monty Hall problem also fell victim to not understanding why opening a door should make any difference. Even when given the mathematical explanation multiple times, he wasn’t really convinced. It took several days before he finally understood the correct solution.

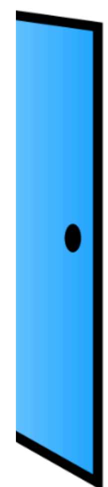
This problem is one of those situations—albeit rare—where someone can be shown an entire chain of logic, surveying the whole problem and its solution, and yet still have it bump up against their intuition. Of course, there is nothing inherently useful about our intuitions. Forged by evolution in situations completely different millions of years ago, our brain’s cognitive abilities are very often irrational, and when dealing with highly sophisticated tasks, we must overcome our intuition in order to understand them properly.

But seldom is this seen so clearly as in the Monty Hall problem. From Wikipedia:

> When first presented with the Monty Hall problem an overwhelming majority of people assume that each door has an equal probability and conclude that switching does not matter (Mueser and Granberg, 1999). Out of 228 subjects in one study, only 13% chose to switch (Granberg and Brown, 1995:713). In her book **The Power of Logical Thinking**, vos Savant (1996, p. 15) quotes cognitive psychologist Massimo Piattelli-Palmarini as saying "... no other statistical puzzle comes so close to fooling all the people all the time" and "that even Nobel physicists systematically give the wrong answer, and that they **insist** on it, and they are ready to berate in print those who propose the right answer". Pigeons repeatedly exposed to the problem show that they rapidly learn always to switch, unlike humans (Herbranson and Schroeder, 2010).

Stressing that last line again, that pigeons "rapidly learn always to switch, unlike humans," shows how unstable the pedestal is upon which humanity places itself. Our cognitive powers are great, but we certainly are far from perfect.





https://www.wired.com/2014/11/monty-hall-erdos-limited-minds/

Don’t Trust Lazy Mathematicians!

Monty Hall, Erdos, and Our Limited Minds

In fact, Paul Erdős, one of the most prolific and foremost mathematicians involved in probability, when initially told of the Monty Hall problem also fell victim to not understanding why opening a door should make any difference. Even when given the mathematical explanation multiple times, he wasn’t really convinced. It took several days before he finally understood the correct solution.

This problem is one of those situations—albeit rare—where someone can be shown an entire chain of logic, surveying the whole problem and its solution, and yet still have it bump up against their intuition. Of course, there is nothing inherently useful about our intuitions. Forged by evolution in situations completely different millions of years ago, our brain’s cognitive abilities are very often irrational, and when dealing with highly sophisticated tasks, we must overcome our intuition in order to understand them properly.

But seldom is this seen so clearly as in the Monty Hall problem. From Wikipedia:

ty-hall-erdos-limited-minds/

Don’t Trust Lazy Mathematicians!

Monty Hall, Erdos, and Our Limited Minds

In fact, Paul Erdős, one of the most prolific and foremost mathematicians involved in probability, when initially told of the Monty Hall problem also fell victim to not understanding why opening a door should make any difference. Even when given the mathematical explanation multiple times, he wasn’t really convinced. It took several days before he finally understood the correct solution.

This problem is one of those situations—albeit rare—where someone can be shown an entire chain of logic, surveying the whole problem and its solution, and yet still have it bump up against their intuition. Of course, there is nothing inherently useful about our intuitions. Forged by evolution in situations completely different millions of years ago, our brain’s cognitive abilities are very often irrational, and when dealing with highly sophisticated tasks, we must overcome our intuition in order to understand them properly.

But seldom is this seen so clearly as in the Monty Hall problem. From Wikipedia:

Gotta prove it, Paul!

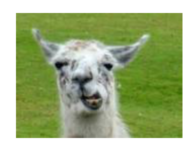
ty-hall-erdos-limited-minds/

The Monty Hall Problem

$1M

The Monty Hall Problem

$1M

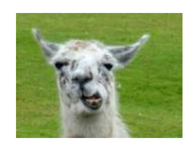


The Monty Hall Problem

$1M

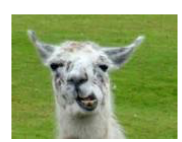
The Monty Hall Problem

$1M



The Monty Hall Problem

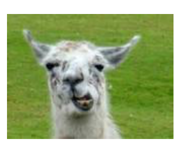
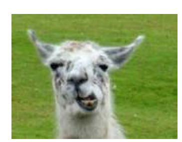
$1M



The Monty Hall Problem

The Monty Hall Problem

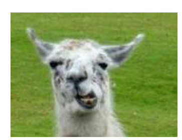


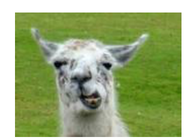



$1M

The Monty Hall Problem

$1M





MHP 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?

Any questions about how the game is played?

The Switching Policy Rational!

Proof: Our overarching technique will be proof by cases.

We denote the possible cases for initial distribution using a simple notation, according to which for example 'LLM' means that, there is a lama behind Door 1, a llama behind Door 2, and the million dollars behind Door 3. With this notation in hand, our three starting cases are: Case 1: MLL; Case 2: LML; Case 3: LLM. There are only three top-level cases for distribution. The odds of picking at the start the million-dollar door is 1/3, obviously — for each case. Hence we know that the odds of a HOLD policy winning is 1/3.

Now we proceed in a proof by sub-cases under the three cases above, to show that the overall odds of a SWITCH policy is greater than 1/3. Each sub-case is simply based on what the initial choice by Jones is, under one of the three main cases. Here we go:

Suppose Case 3, LLM, holds, and that [this (Case 3.1) is the first of three sub-cases under Case 3] Jones picks Door 1. Then FM must reveal Door 2 to reveal a llama. Switching to Door 3 wins, guaranteed. In sub-case 3.2 suppose that J's choice Door 2. Then FM will reveal Door 1. Again, switching to Door 3 wins, guaranteed. In the final sub-case, J initially selects Door 3 under Case 3; this is sub-case 3.3. Here, FM shows either Door 1 or Door 2 (as itself a random choice). This time switching loses, guaranteed. Hence, in two of the sub-cases out of three (2/3), winning is guaranteed (*prob* of 1). An exactly parallel result can be deduced for Case 2 and Case 1; i.e., in each of these two, in two of the three (2/3) sub-cases winning is 1. Hence the odds of winning by following the switching policy is 2/3, which is greater than 1/3. Hence it's rational to be a switcher. **QED**

What about 4 doors?

The Switching Policy Rational (in *n*-door game)!

$$\frac{n-1}{n} \times \frac{1}{n-2}$$

The Switching Policy Rational (in *n*-door game)!

$$\frac{n-1}{n} \times \frac{1}{n-2}$$

How do we prove *this*, which is a "generalization" of the proof-by-cases approach use by Selmer?

Deductive vs Inductive Paradoxes;
Deductive Reasoning vs Inductive Reasoning

…

Paradoxes are engines of progress in formal logic & fields based upon formal logic.

E.g., Russell's Paradox — as we've seen.

Types of Paradoxes

- Deductive Paradoxes. The reasoning in question is exclusively deductive.
  - Russell's Paradox
  - The Liar Paradox
  - Richard's Paradox
- Inductive Paradoxes Some of the reasoning in question uses non-deductive reasoning (e.g., probabilistic reasoning, abductive reasoning, analogical reasoning, etc.).

Types of Paradoxes

- Deductive Paradoxes. The reasoning in question is exclusively deductive.
  - Russell's Paradox
  - The Liar Paradox
  - Richard's Paradox
- Inductive Paradoxes Some of the reasoning in question uses non-deductive reasoning (e.g., probabilistic reasoning, abductive reasoning, analogical reasoning, etc.).

E.g., in The Lottery Paradox, which we look at later today:

E.g., in The Lottery Paradox, which we look at later today:

- Instead of arriving at some $\zeta \land \neg\zeta$, we have forced upon us perhaps something like that while …

E.g., in The Lottery Paradox, which we look at later today:

- Instead of arriving at some $\zeta \wedge \neg\zeta$, we have forced upon us perhaps something like that while …
- a perfectly rational person can never believe *P* and believe $\neg P$ at the same time, nonetheless …

E.g., in The Lottery Paradox, which we look at later today:

- Instead of arriving at some $\zeta \land \neg\zeta$, we have forced upon us perhaps something like that while …
- a perfectly rational person can never believe *P* and believe $\neg$*P* at the same time, nonetheless …
- The Lottery Paradox (apparently) shows, courtesy of its two Sequences (of Reasoning), that a perfectly rational person can indeed have such a belief (upon considering a fair, large lottery).

E.g., in The Lottery Paradox, which we look at later today:

- Instead of arriving at some $\zeta \wedge \neg\zeta$, we have forced upon us perhaps something like that while …
- a perfectly rational person can never believe *P* and believe $\neg P$ at the same time, nonetheless …
- The Lottery Paradox (apparently) shows, courtesy of its two Sequences (of Reasoning), that a perfectly rational person can indeed have such a belief (upon considering a fair, large lottery).

Portal to Competing Approaches to Inductive Logic: The Paradox of Grue …

The Grue Paradox Constructed

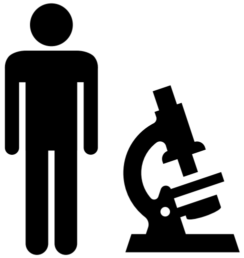
source of minerals

Mineralogist

The Grue Paradox Constructed

source of minerals

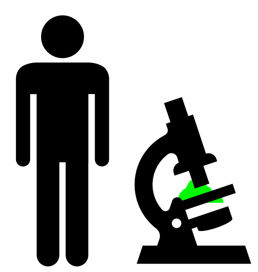
Mineralogist



The Grue Paradox Constructed

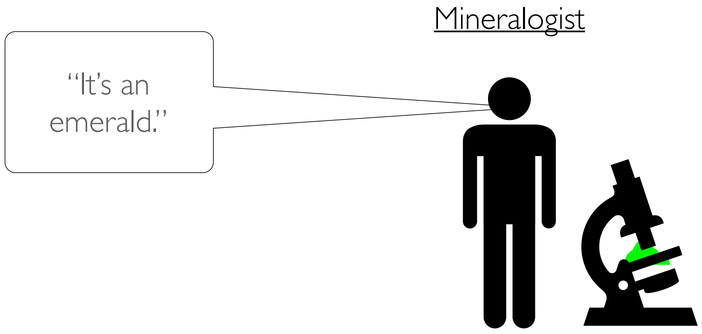
source of minerals

Mineralogist



The Grue Paradox Constructed

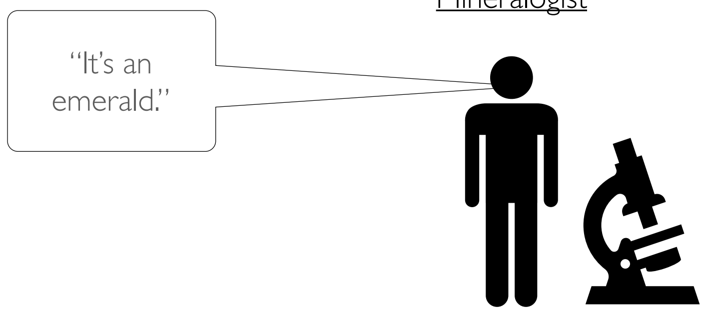
source of minerals



The Grue Paradox Constructed

source of minerals

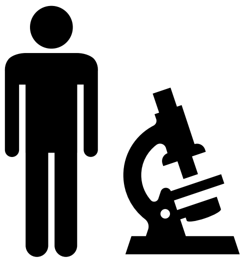
Mineralogist



The Grue Paradox Constructed

source of minerals

Mineralogist



The Grue Paradox Constructed

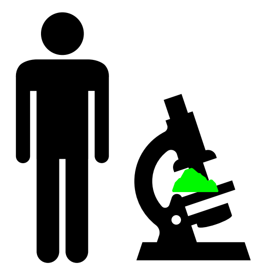
source of minerals

Mineralogist

The Grue Paradox Constructed

source of minerals

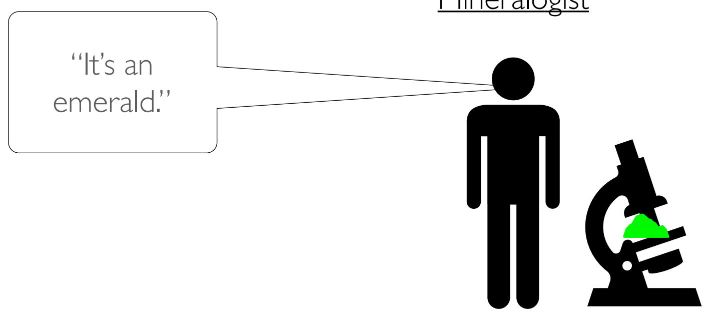
Mineralogist



The Grue Paradox Constructed

source of minerals

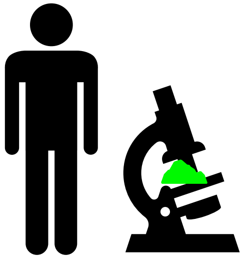
Mineralogist



The Grue Paradox Constructed

source of minerals

Mineralogist

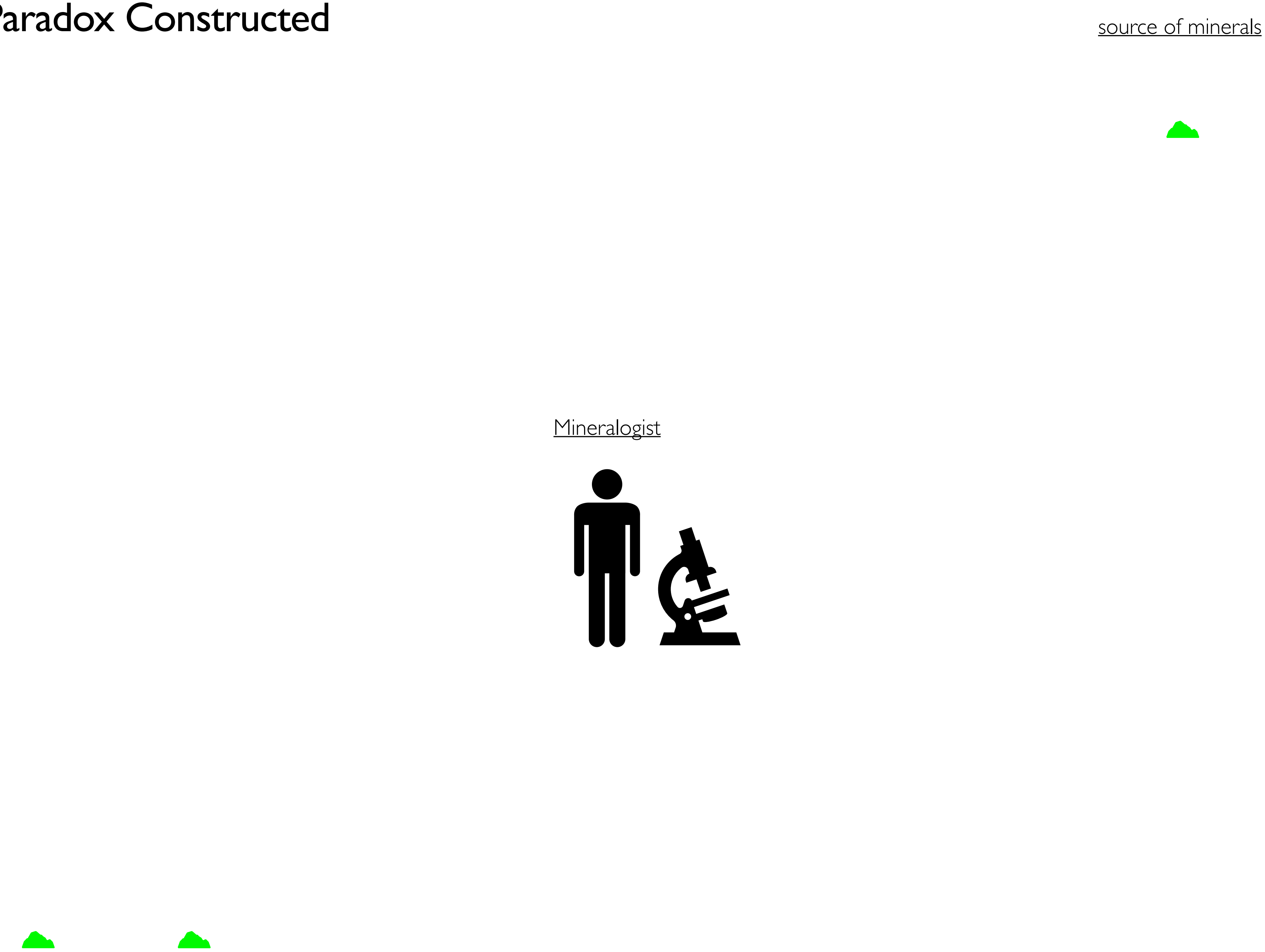


The Grue Paradox Constructed

source of minerals

Mineralogist

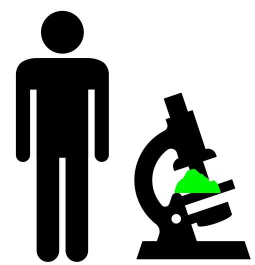
The Grue Paradox Constructed



The Grue Paradox Constructed

source of minerals

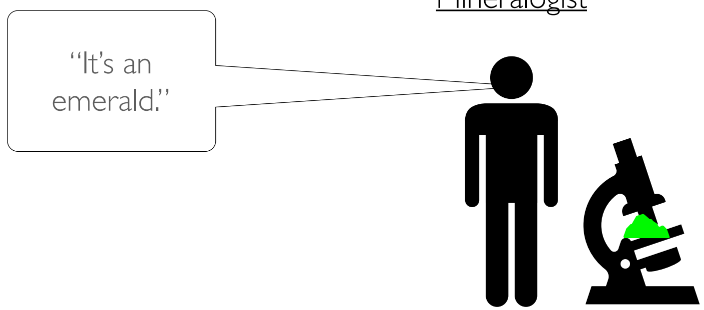
Mineralogist



The Grue Paradox Constructed

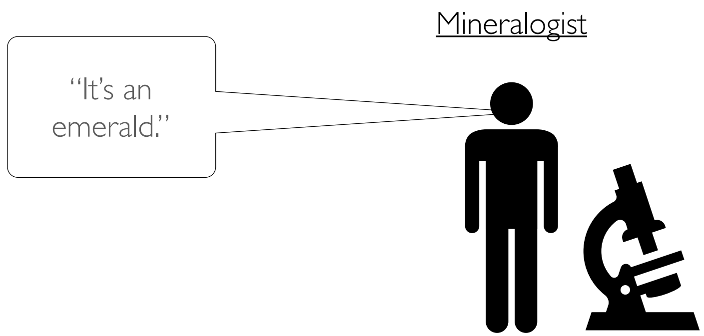
source of minerals

Mineralogist



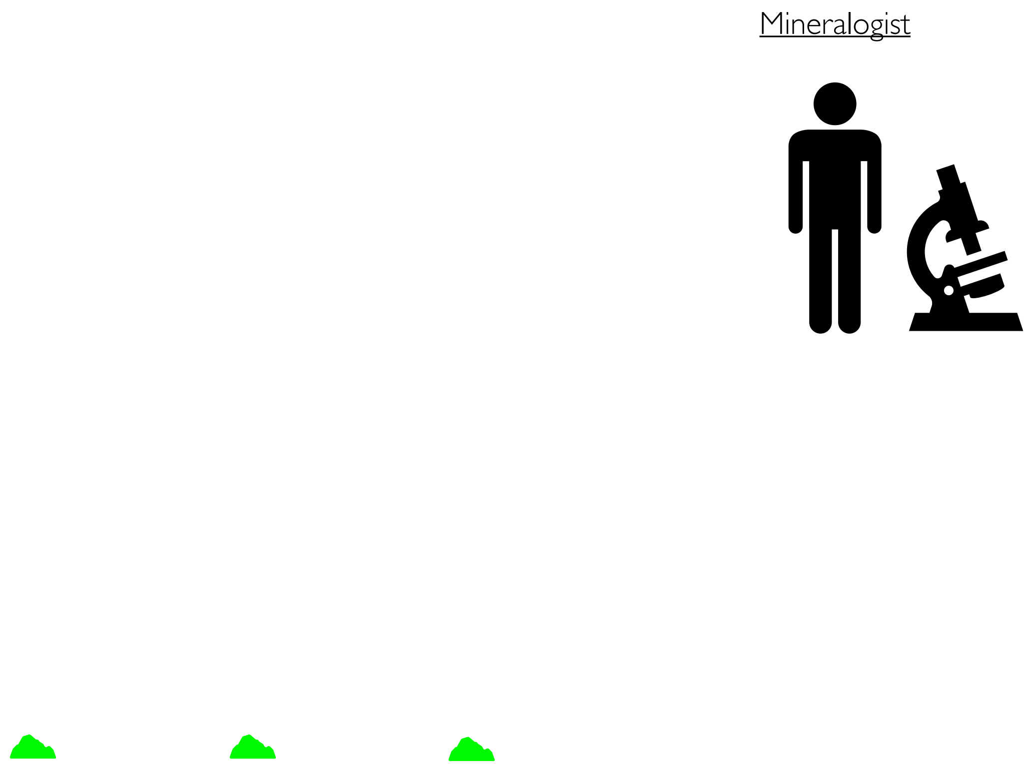
The Grue Paradox Constructed

source of minerals



The Grue Paradox Constructed

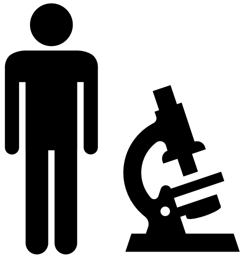
source of minerals



The Grue Paradox Constructed

source of minerals

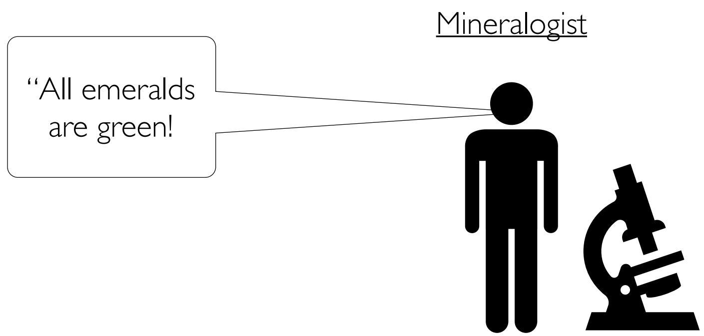
Mineralogist



 … $k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

The Grue Paradox Constructed

source of minerals



 … $k, k \in \mathbb{Z}^{+},$ and all before time $t^{\star}$

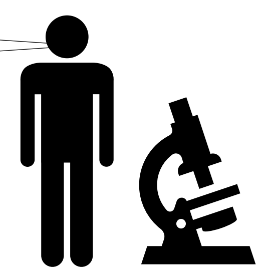
The Grue Paradox Constructed

source of minerals

The Paradox: First, we define the property *Grue*:

Mineralogist

"All emeralds are green!



… $k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

The Grue Paradox Constructed

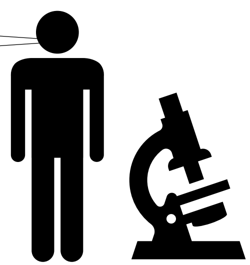
source of minerals

The Paradox: First, we define the property *Grue*:

$$\forall x \ \forall t \ [Grue(x) \text{ iff } (t < t^\star \rightarrow Green(x,t)) \wedge (t \geq t^\star \rightarrow Blue(x,t))]$$

Mineralogist

"All emeralds are green!



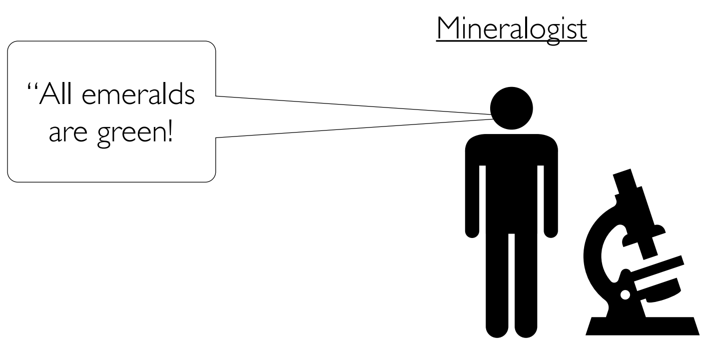
 … $k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

The Grue Paradox Constructed

source of minerals

The Paradox: First, we define the property *Grue*:

$$\forall x\ \forall t\ [Grue(x) \text{ iff } (t < t^\star \rightarrow Green(x,t)) \wedge (t \geq t^\star \rightarrow Blue(x,t))]$$



$Green(o_1, t_1)$

$\ldots\ k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

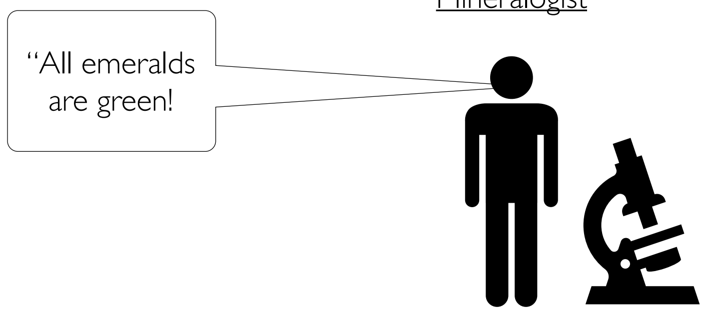
The Grue Paradox Constructed

source of minerals

The Paradox: First, we define the property *Grue*:

$$\forall x\ \forall t\ [Grue(x) \text{ iff } (t < t^\star \rightarrow Green(x,t)) \wedge (t \geq t^\star \rightarrow Blue(x,t))]$$

Mineralogist



$Green(o_1, t_1)$ $Green(o_2, t_2)$

$\ldots\ k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

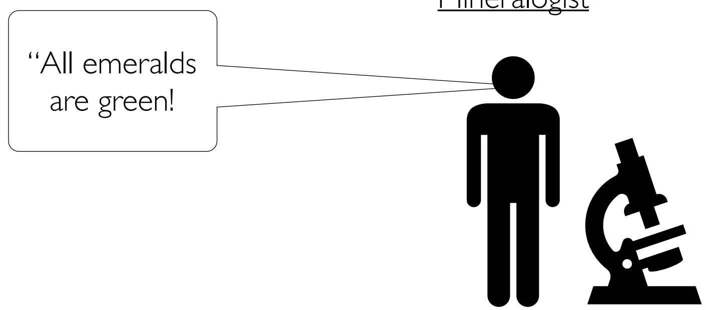
The Grue Paradox Constructed

source of minerals

The Paradox: First, we define the property *Grue*:

$$\forall x\ \forall t\ [Grue(x) \text{ iff } (t < t^\star \rightarrow Green(x,t)) \wedge (t \geq t^\star \rightarrow Blue(x,t))]$$

Mineralogist



$Green(o_1, t_1)$ $Green(o_2, t_2)$ $Green(o_3, t_3)$ ...

$\ldots\ k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

The Grue Paradox Constructed

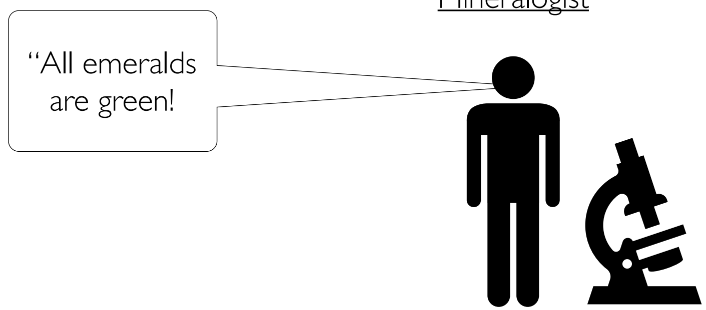
source of minerals

The Paradox: First, we define the property *Grue*:

$$\forall x\ \forall t\ [Grue(x) \text{ iff } (t < t^\star \rightarrow Green(x,t)) \wedge (t \geq t^\star \rightarrow Blue(x,t))]$$

The k atomic formulae, together, support the general mineralogical law affirmed by our mineralogist, who we assume to be a rational empirical scientist. But this scientist must *also* affirm the proposition that all emeralds are grue, since the very same atomic formulae support this proposition, and to the same degree. But surely this is absurd!

Mineralogist



$Green(o_1, t_1)$ $\quad$ $Green(o_2, t_2)$ $\quad$ $Green(o_3, t_3)$ $\ldots$

$\ldots\ k, k \in \mathbb{Z}^+,$ and all before time $t^\star$

The Grue Paradox Constructed

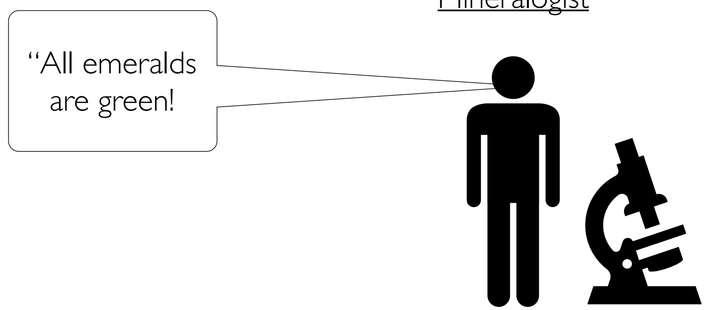
source of minerals

The Paradox: First, we define the property *Grue*:

$$\forall x\ \forall t\ [Grue(x) \text{ iff } (t < t^\star \rightarrow Green(x,t)) \wedge (t \geq t^\star \rightarrow Blue(x,t))]$$

The k atomic formulae, together, support the general mineralogical law affirmed by our mineralogist, who we assume to be a rational empirical scientist. But this scientist must *also* affirm the proposition that all emeralds are grue, since the very same atomic formulae support this proposition, and to the same degree. But surely this is absurd!

Mineralogist



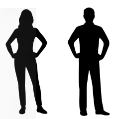
How can we solve this?!

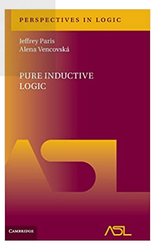
$Green(o_1, t_1)$ $Green(o_2, t_2)$ $Green(o_3, t_3)$ …

… $k, k \in \mathbb{Z}^+$, and all before time $t^\star$

The Grue Paradox Divides Hearts and Minds

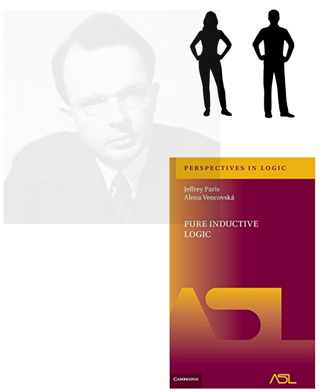
The Mathematicians/Logicians
Following Carnap





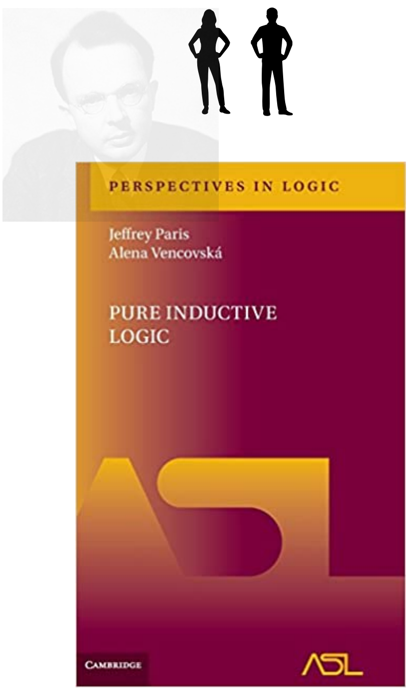
The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians
Following Carnap



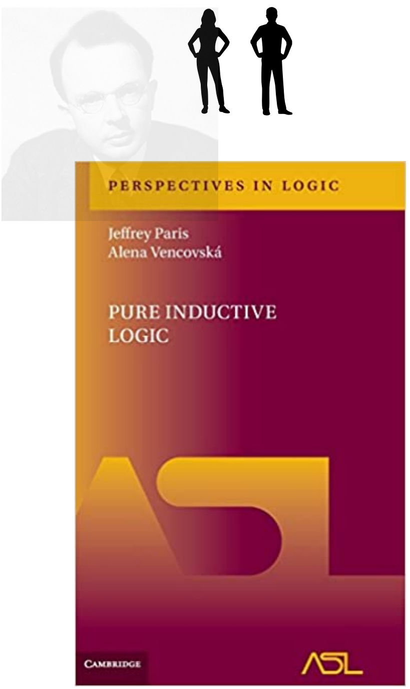
The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians
Following Carnap



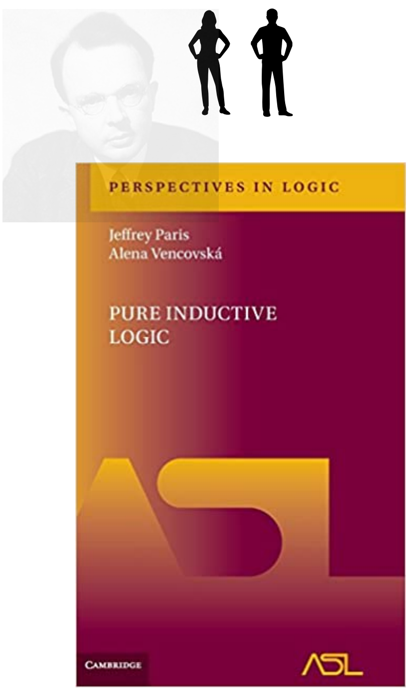
The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians
Following Carnap



The Grue Paradox Divides Hearts and Minds

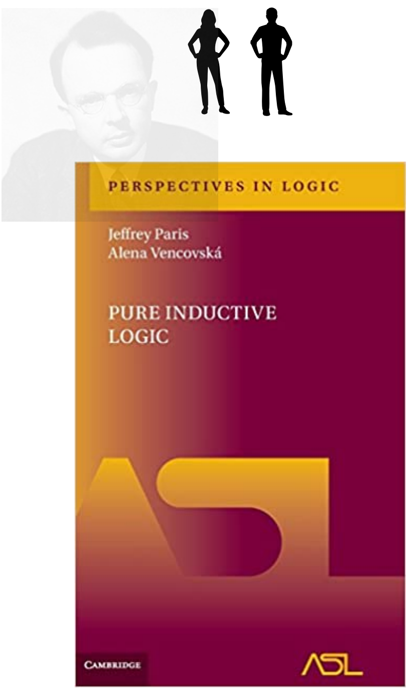
The Mathematicians/Logicians
Following Carnap



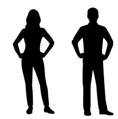
The Philosophers

The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians
Following Carnap

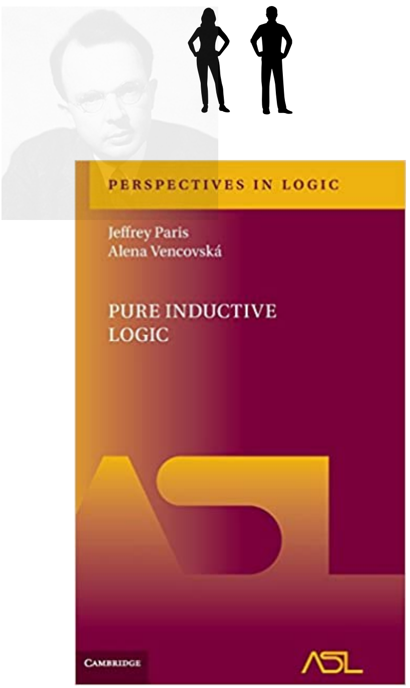


The Philosophers

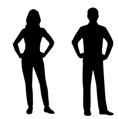


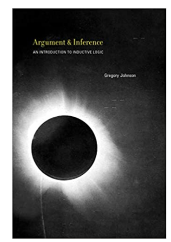
The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians
Following Carnap



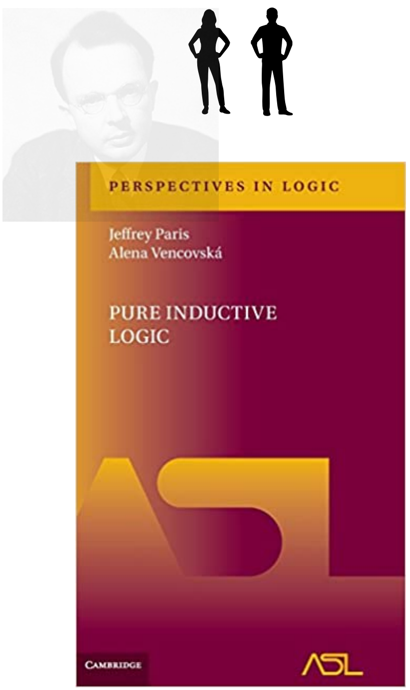
The Philosophers



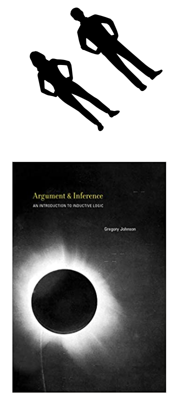


The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians
Following Carnap



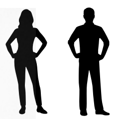
The Philosophers

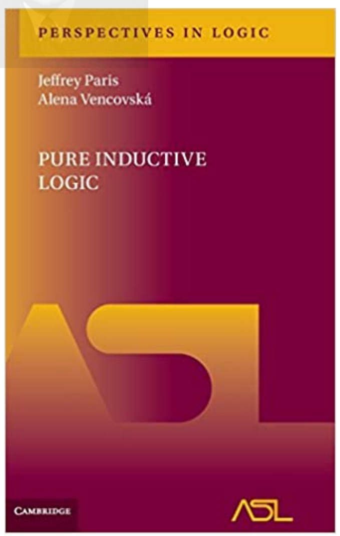


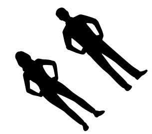
The Grue Paradox Divides Hearts and Minds

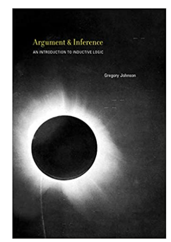
The Mathematicians/Logicians
Following Carnap

The Philosophers







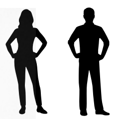


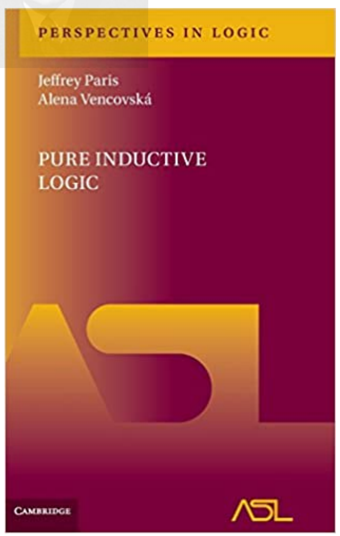
The Grue Paradox Divides Hearts and Minds

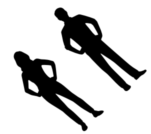
The Mathematicians/Logicians Following Carnap

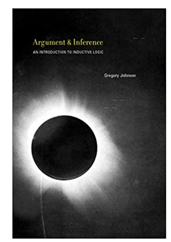
The Philosophers

Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm







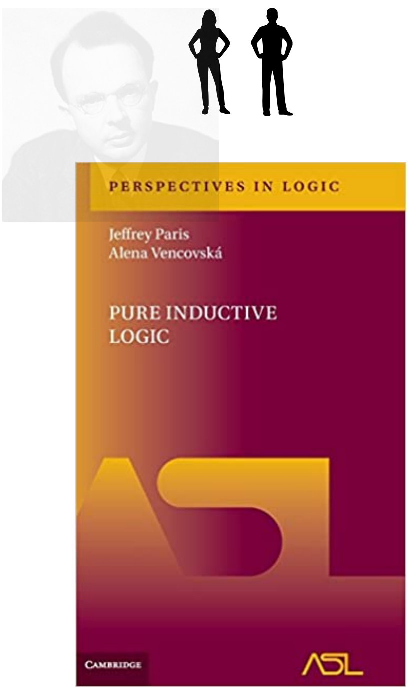


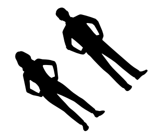
The Grue Paradox Divides Hearts and Minds

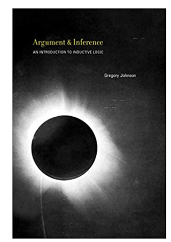
The Mathematicians/Logicians Following Carnap

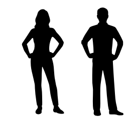
The Philosophers

Logicist AI/Computational Logicians in Proof-centric/argument-centric Paradigm



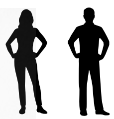


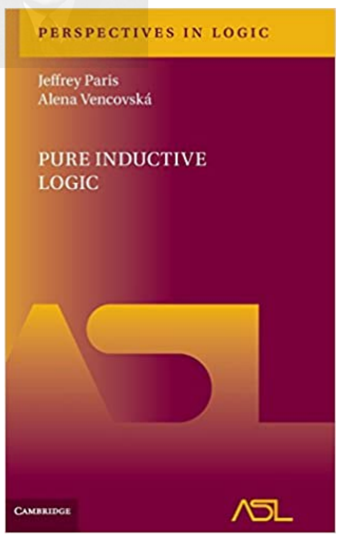




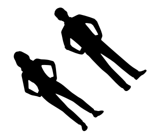
The Grue Paradox Divides Hearts and Minds

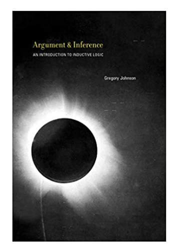
The Mathematicians/Logicians Following Carnap



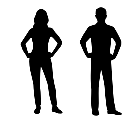


The Philosophers



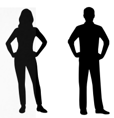


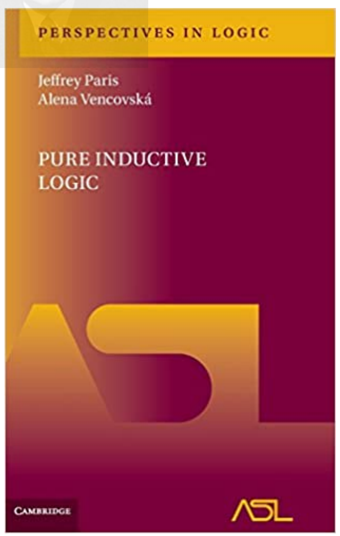
Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm



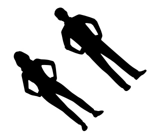
The Grue Paradox Divides Hearts and Minds

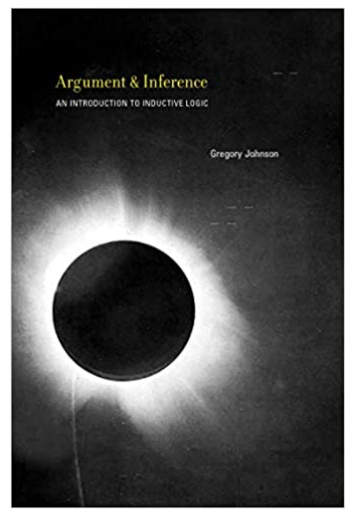
The Mathematicians/Logicians Following Carnap



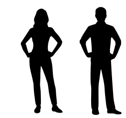


The Philosophers



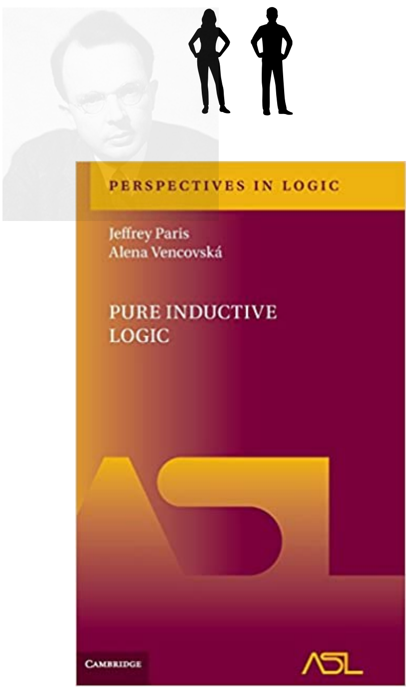


Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm

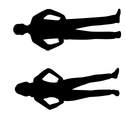


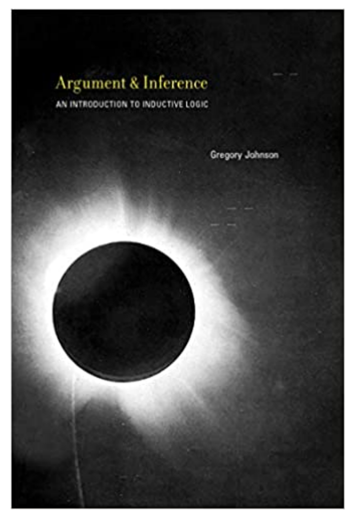
The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians Following Carnap

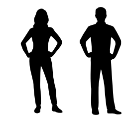


The Philosophers



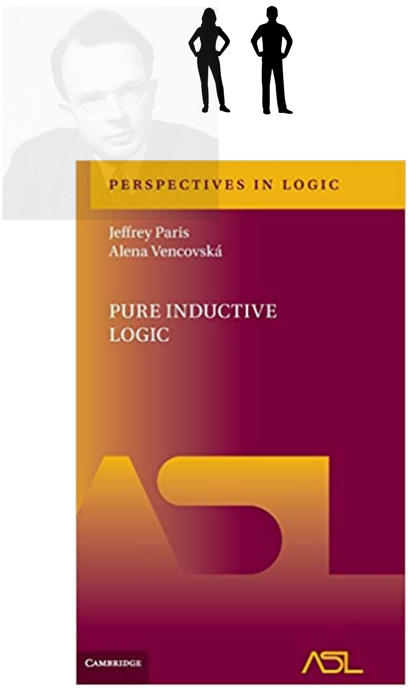


Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm

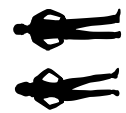


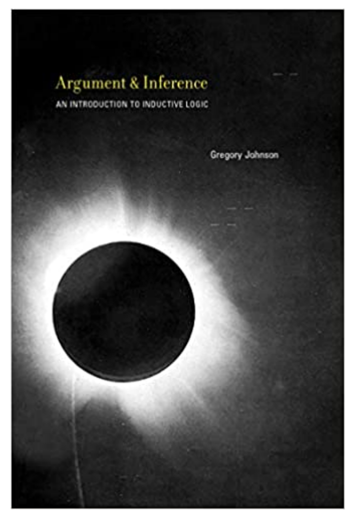
The Grue Paradox Divides Hearts and Minds

The Mathematicians/Logicians Following Carnap



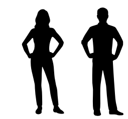
The Philosophers





Yours truly

Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm

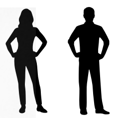


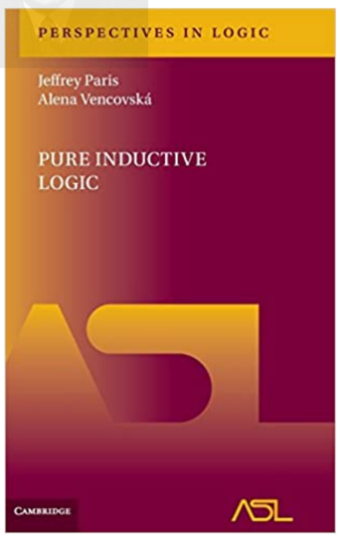
The Grue Paradox Divides Hearts and Minds

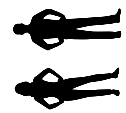
The Mathematicians/Logicians Following Carnap

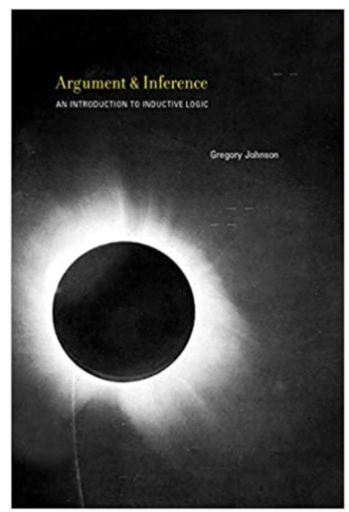
The Philosophers

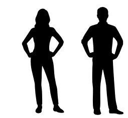
Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm









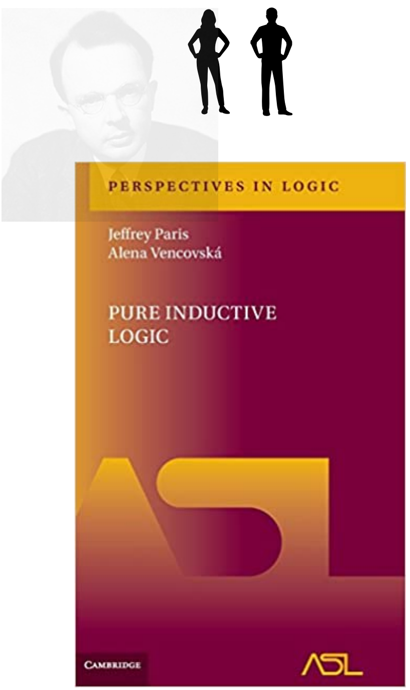


Yours truly

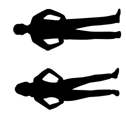
The Grue Paradox Divides Hearts and Minds

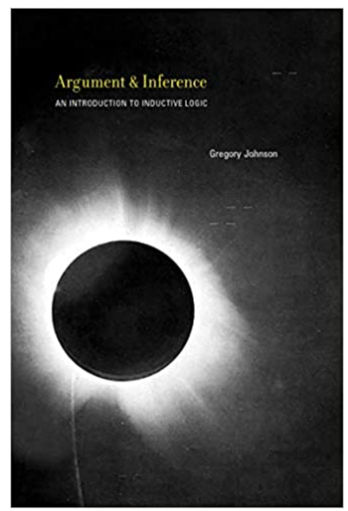
LAMA®

The Mathematicians/Logicians Following Carnap



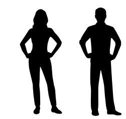
The Philosophers





Yours truly

Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm



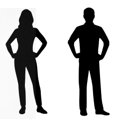
The Grue Paradox Divides Hearts and Minds

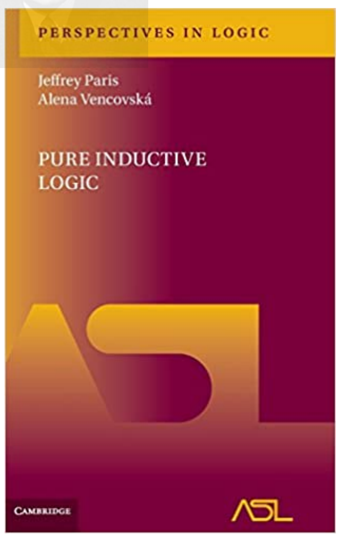
LAMA®

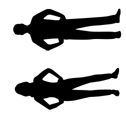
The Mathematicians/Logicians Following Carnap

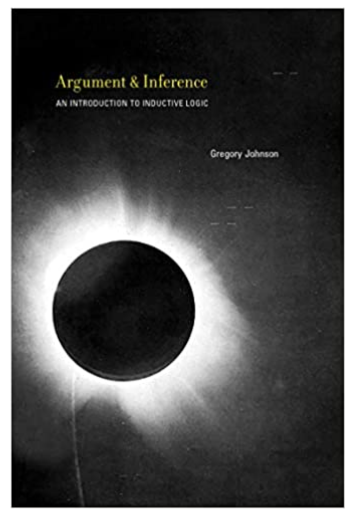
The Philosophers

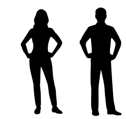
Logicist AI/Computational Logicians in Proof-centric/ argument-centric Paradigm











Yours truly

Background Reading …

The Original Publication Introducing The Grue Paradox

Goodman, N. (1955) *Fact, Fiction, and Forecast* (Cambridge, MA: Harvard University Press).
4th edition 1983, also HUP.

From "Nelson Goodman" in SEP (https://plato.stanford.edu/entries/goodman)

Here comes the riddle. Suppose that your research is in gemology. Your special interest lies in the color properties of certain gemstones, in particular, emeralds. All emeralds you have examined before a certain time t were green (your notebook is full of evidence statements of the form "Emerald x found at place y date $z(z \leq t)$ is green"). It seems that, at t, this supports the hypothesis that all emeralds are green (L3).

Now Goodman introduces the predicate "grue". This predicate applies to all things examined before some future time t just in case they are green but to other things (observed at or after t) just in case they are blue:

(DEF1) x is grue $=_{df}$ x is examined before t and green $\vee$ x is not so examined and blue

Until t it is obviously the case that for each statement in your notebook, there is a parallel statement asserting that the emerald x found at place y date $z(z \leq t)$ is grue. Each of these statements is analytically equivalent with the corresponding one in your notebook. All these grue-evidence statements taken together confirm the hypothesis that all emeralds are grue (L4), and they confirm this hypothesis to the exact same degree as the green-evidence statements confirmed the hypothesis that all emeralds are green. But if that is the case, then the following two predictions are also confirmed to the same degree:

(P1) The next emerald first examined after t will be green.
(P2) The next emerald first examined after t will be grue.

However, to be a grue emerald examined after t is not to be a green emerald. An emerald first examined after t is grue iff it is blue. We have two mutually incompatible predictions, both confirmed to the same degree by the past evidence. We could obviously define infinitely many grue-like predicates that would all lead to new, similarly incompatible predictions.

The immediate lesson is that we cannot use all kinds of weird predicates to formulate hypotheses or to classify our evidence. Some predicates (which are the ones like "green") can be used for this; other predicates (the ones like "grue") must be excluded, if induction is supposed to make any sense. This already is an interesting result. For valid inductive inferences the choice of predicates matters.

It is not just that we lack justification for accepting a general hypothesis as true only on the basis of positive instances and lack of counterinstances (which was the old problem), or to define what rule we are using when accepting a general hypothesis as true on these grounds (which was the problem after Hume). The problem is to explain why some general statements (such as L3) are confirmed by their instances, whereas others (such as L4) are not. Again, this is a matter of the lawlikeness of L3 in contrast to L4, but how are we supposed to tell the lawlike regularities from the illegitimate generalizations?

Wikipedia Entry
“New Riddle of Induction” Isn’t Half Bad!

(https://en.wikipedia.org/wiki/New_riddle_of_induction)

Tutorial by Paris on Pure Inductive Logic:

http://fitelson.org/few/paris_notes.pdf

(Paris explains that the mathematicians just assumed the reasoning in the grue paradox is invalid, and then continued on their way to erect upon Carnap's work a robust formal edifice (= *pure inductive logic*).)

See "Inductive Logic" in SEP for an excellent overview, and in particular nice coverage of Carnap's seminal contributions, which PIL extends. (https://plato.stanford.edu/entries/logic-inductive)

Inductive Logic

First published Mon Sep 6, 2004; substantive revision Mon Mar 19, 2018

An inductive logic is a logic of evidential support. In a deductive logic, the premises of a valid deductive argument *logically entail* the conclusion, where *logical entailment* means that every logically possible state of affairs that makes the premises true *must* make the conclusion truth as well. Thus, the premises of a valid deductive argument provide *total support* for the conclusion. An inductive logic extends this idea to weaker arguments. In a good inductive argument, the truth of the premises provides some *degree of support* for the truth of the conclusion, where this *degree-of-support* might be measured via some numerical scale. By analogy with the notion of deductive entailment, the notion of inductive degree-of-support might mean something like this: among the logically possible states of affairs that make the premises true, the conclusion must be true in (at least) proportion r of them—where r is some numerical measure of the support strength.

If a logic of *good inductive arguments* is to be of any real value, the measure of support it articulates should be up to the task. Presumably, the logic should at least satisfy the following condition:

> **Criterion of Adequacy (CoA)**:
> The logic should make it likely (as a matter of logic) that as evidence accumulates, the total body of true evidence claims will eventually come to indicate, via the logic's *measure of support*, that false hypotheses are probably false and that true hypotheses are probably true.

The CoA stated here may strike some readers as surprisingly strong. Given a specific logic of evidential support, how might it be shown to satisfy such a condition? Section 4 will show precisely how this condition is satisfied by the logic of evidential support articulated in Sections 1 through 3 of this article.

This article will focus on the kind of the approach to inductive logic most widely studied by epistemologists and logicians in recent years. This approach employs conditional probability functions to represent measures of the degree to which evidence statements support hypotheses. Presumably, hypotheses should be empirically evaluated based on what they *say* (or imply) about the likelihood that evidence claims will be true. A straightforward theorem of probability theory, called Bayes' Theorem, articulates the way in which what hypotheses *say* about the likelihoods of evidence claims influences the degree to which hypotheses are supported by those evidence claims. Thus, this approach to the logic of evidential support is often called a *Bayesian Inductive Logic* or a *Bayesian Confirmation Theory*. This article will first provide a detailed explication of a Bayesian approach to inductive logic. It will then examine the extent to which this logic may pass muster as an adequate logic of evidential support for hypotheses. In particular, we will see how such a logic may be shown to satisfy the Criterion of Adequacy stated above.

Simple Inductive-Reasoning Example from Pollock, for Lead-in to The Lottery Paradox …

Imagine the following:

Keith tells you that the morning news predicts rain in Troy today. However, Alvin tells you that the same news report predicted sunshine.

Imagine the following: Keith tells you that the morning news predicts rain in Tucson today. However, Alvin tells you that the same news report predicted sunshine.

Without any other source of information, it would be irrational to place belief in either Keith's or Alvin's statements.

Imagine the following: Keith tells you that the morning news predicts rain in Tucson today. However, Alvin tells you that the same news report predicted sunshine.

Without any other source of information, it would be irrational to place belief in either Keith's or Alvin's statements.

Further, suppose you happened to watch the noon news report, and that report predicted rain. Then you should believe that it will rain despite your knowledge of Alvin's argument.

Defeasible Reasoning in OSCAR

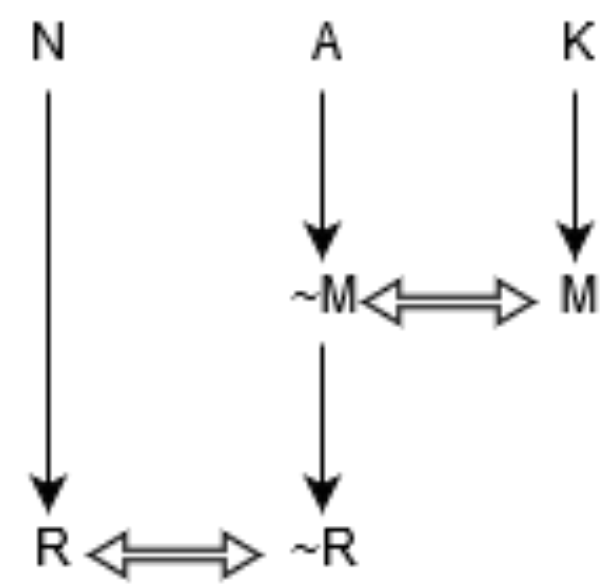
K- Keith says that M

A- Alvin says that ~M

M- The morning news said that R

R- It is going to rain this afternoon

N- The noon news says that R



All such can be absorbed into our inductive logics and our automated inductive reasoners (= our AI).

In Our Inductive Modal Logic

(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?
	(5)	$\mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi)$	Testimonial P1

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?
	(5)	$\mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi)$	Testimonial P1
$\therefore$	(6)	$\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?
	(5)	$\mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi)$	Testimonial P1
$\therefore$	(6)	$\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	
$\therefore$	(7)	$\neg \mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \neg \mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	"Clash" Principle

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?
	(5)	$\mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi)$	Testimonial P1
$\therefore$	(6)	$\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	
$\therefore$	(7)	$\neg\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \neg\mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	"Clash" Principle
	(8)	$\mathbf{K}(you, \mathbf{S}(noonnews, rain))$	

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?
	(5)	$\mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi)$	Testimonial P1
$\therefore$	(6)	$\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	
$\therefore$	(7)	$\neg\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \neg\mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	"Clash" Principle
	(8)	$\mathbf{K}(you, \mathbf{S}(noonnews, rain))$	
$\therefore$	(9)	$\mathbf{S}(noonnews, rain)$	?

In Our Inductive Modal Logic

	(1)	$\mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain)))$	fact
	(2)	$\mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	fact
$\therefore$	(3)	$\mathbf{S}(keith, \mathbf{S}(m, rain))$	?
$\therefore$	(4)	$\mathbf{S}(alvin, \mathbf{S}(m, \neg rain)))$	?
	(5)	$\mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi)$	Testimonial P1
$\therefore$	(6)	$\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	
$\therefore$	(7)	$\neg\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \neg\mathbf{B}^2(you, \mathbf{S}(m, \neg rain))$	"Clash" Principle
	(8)	$\mathbf{K}(you, \mathbf{S}(noonnews, rain))$	
$\therefore$	(9)	$\mathbf{S}(noonnews, rain)$	?
	(10)	$\mathbf{S}(noonnews, \phi) \rightarrow \mathbf{B}^3(you, \phi)$	Testimonial P2

In Our Inductive Modal Logic

$$
\begin{array}{c|l|l|l}
 & (1) & \mathbf{K}(you, \mathbf{S}(keith, \mathbf{S}(m, rain))) & \text{fact} \\
 & (2) & \mathbf{K}(you, \mathbf{S}(alvin, \mathbf{S}(m, \neg rain))) & \text{fact} \\
\therefore & (3) & \mathbf{S}(keith, \mathbf{S}(m, rain)) & ? \\
\therefore & (4) & \mathbf{S}(alvin, \mathbf{S}(m, \neg rain))) & ? \\
 & (5) & \mathbf{S}(keith, \phi) \rightarrow \mathbf{B}^2(you, \phi) & \text{Testimonial P1} \\
\therefore & (6) & \mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \mathbf{B}^2(you, \mathbf{S}(m, \neg rain)) & \\
\therefore & (7) & \neg\mathbf{B}^2(you, \mathbf{S}(m, rain)) \wedge \neg\mathbf{B}^2(you, \mathbf{S}(m, \neg rain)) & \text{``Clash'' Principle} \\
 & (8) & \mathbf{K}(you, \mathbf{S}(noonnews, rain)) & \\
\therefore & (9) & \mathbf{S}(noonnews, rain) & ? \\
 & (10) & \mathbf{S}(noonnews, \phi) \rightarrow \mathbf{B}^3(you, \phi) & \text{Testimonial P2} \\
\therefore & (11) & \mathbf{B}^3(you, rain) &
\end{array}
$$

The Lottery Paradox …

Stewart's Shops

WE ARE CLOSER TO YOU!

E: “Please go down to Stewart’s & get the T U.”

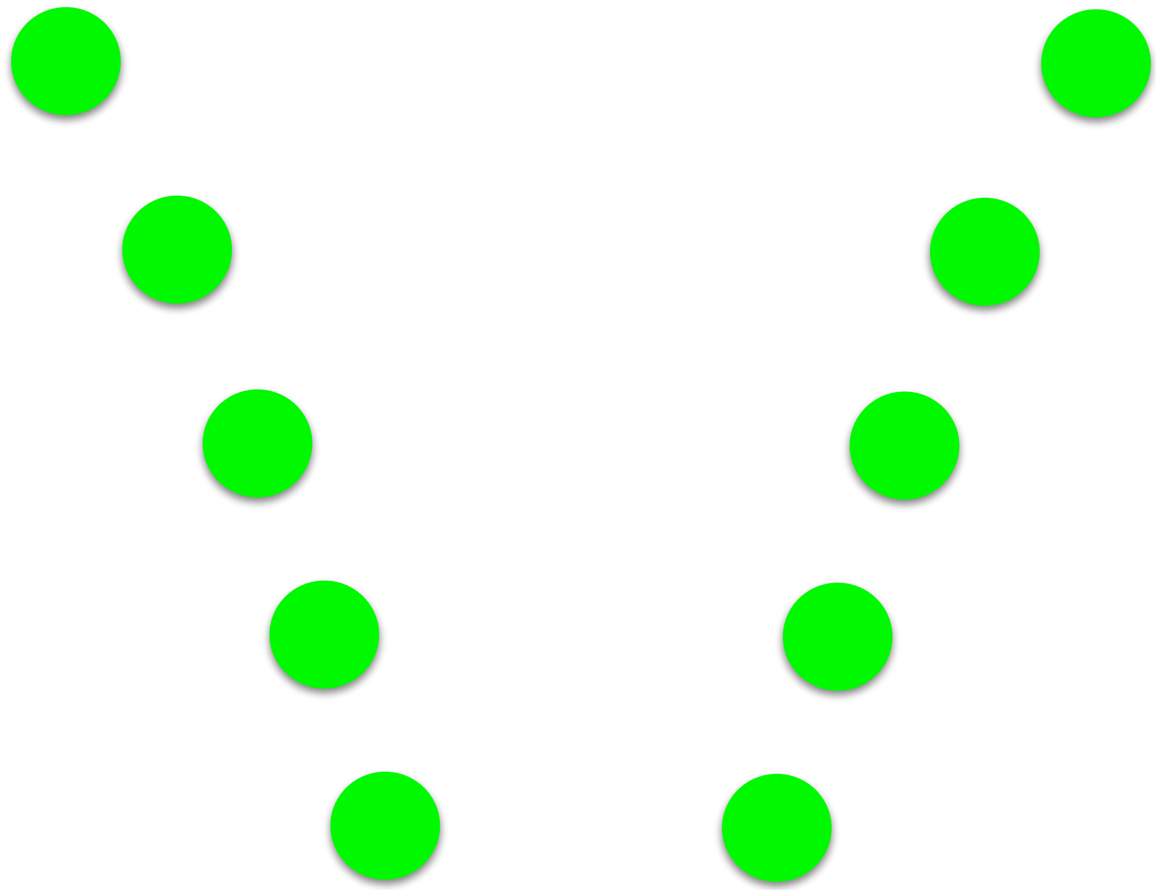
E: “Please go down to Stewart’s & get the T U.”

S: “I’m sorry, E, I’m afraid I can’t do that.
It would be irrational.”

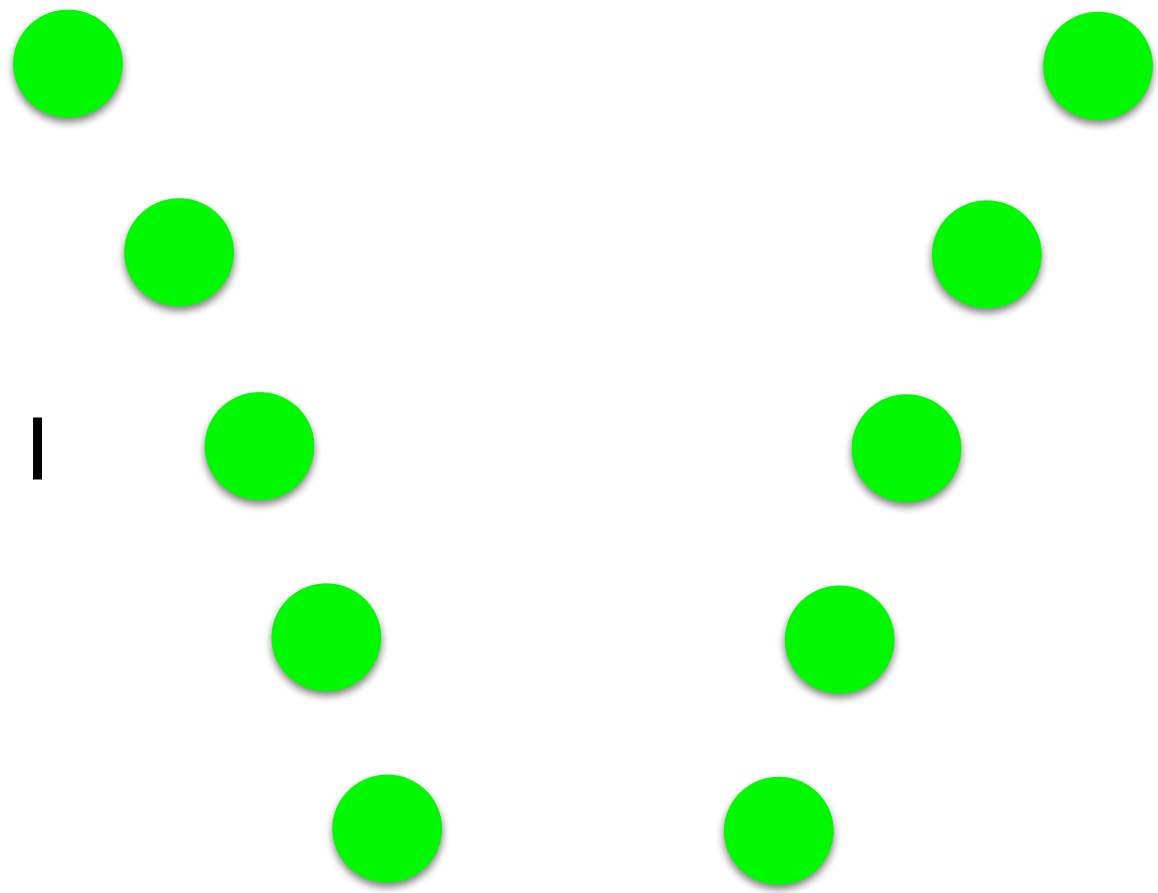
E: “Please go down to Stewart’s & get the T U.”

S: “I’m sorry, E, I’m afraid I can’t do that.
It would be irrational.”

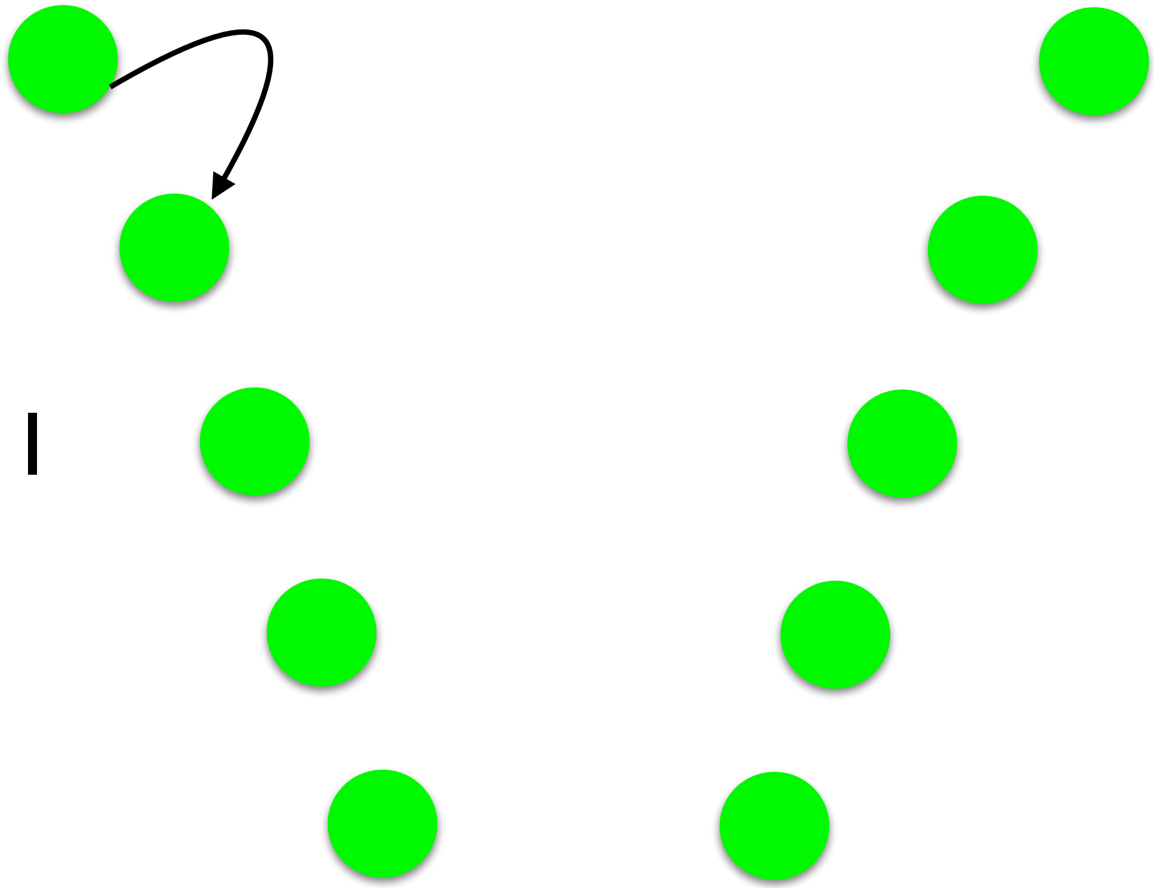
•••



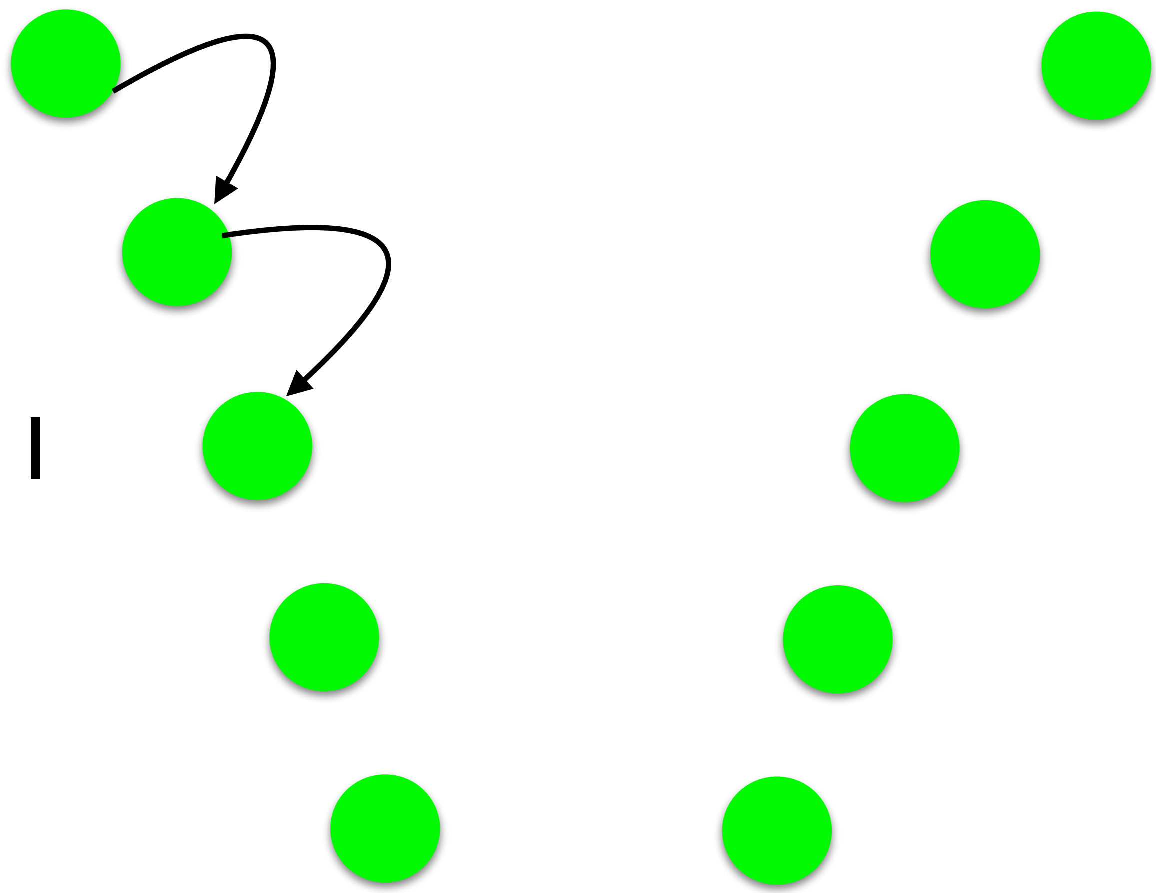
Sequence 1

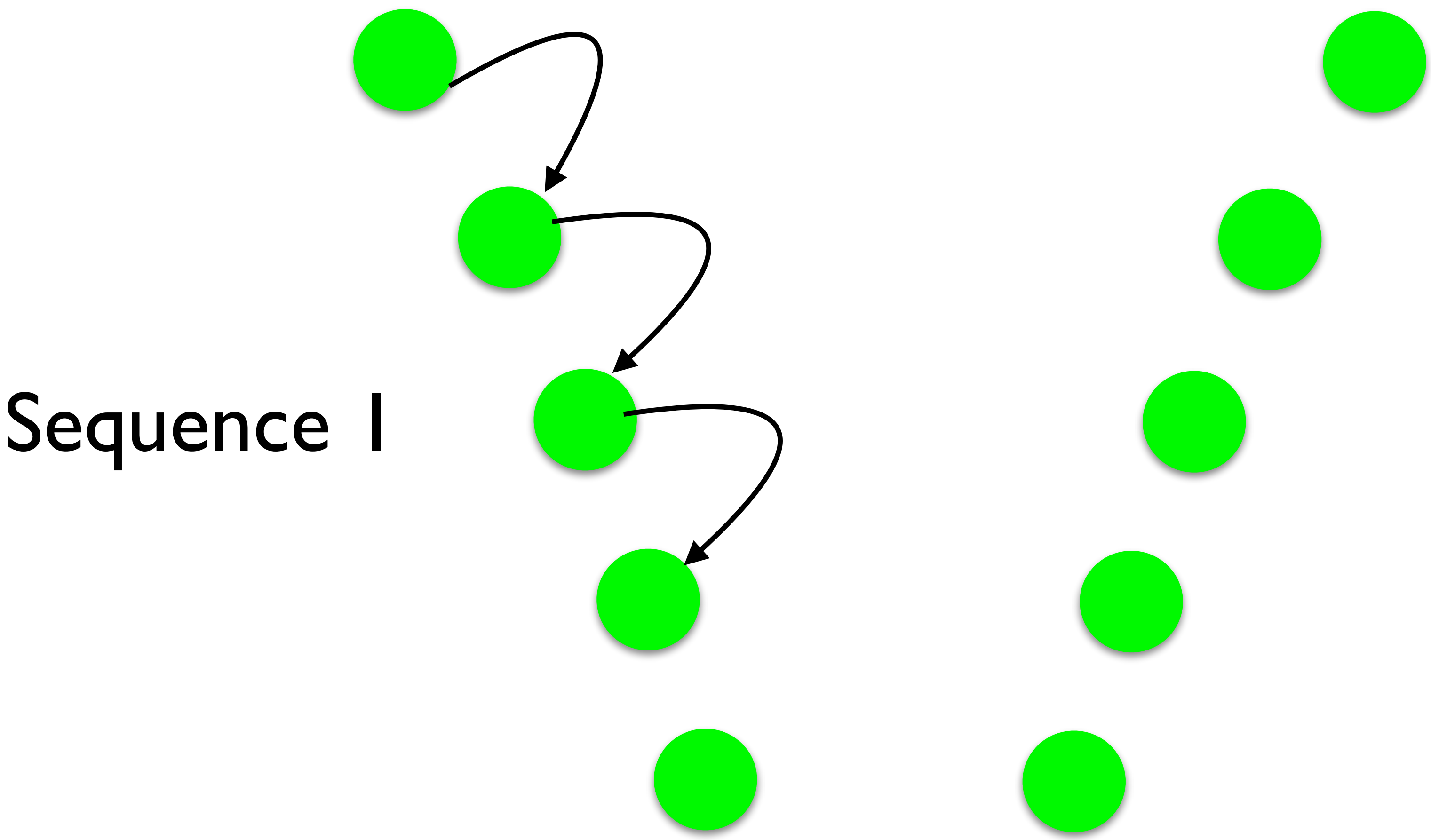


Sequence 1

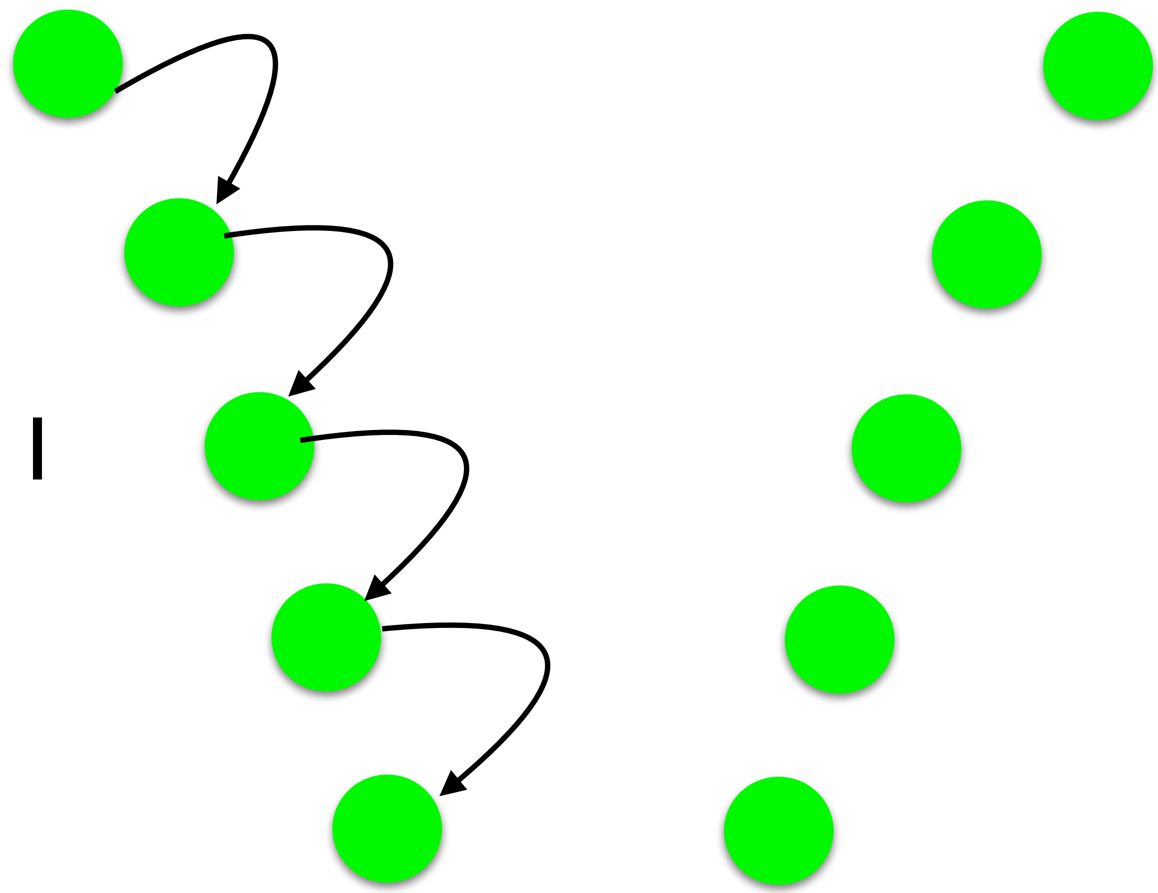


Sequence 1

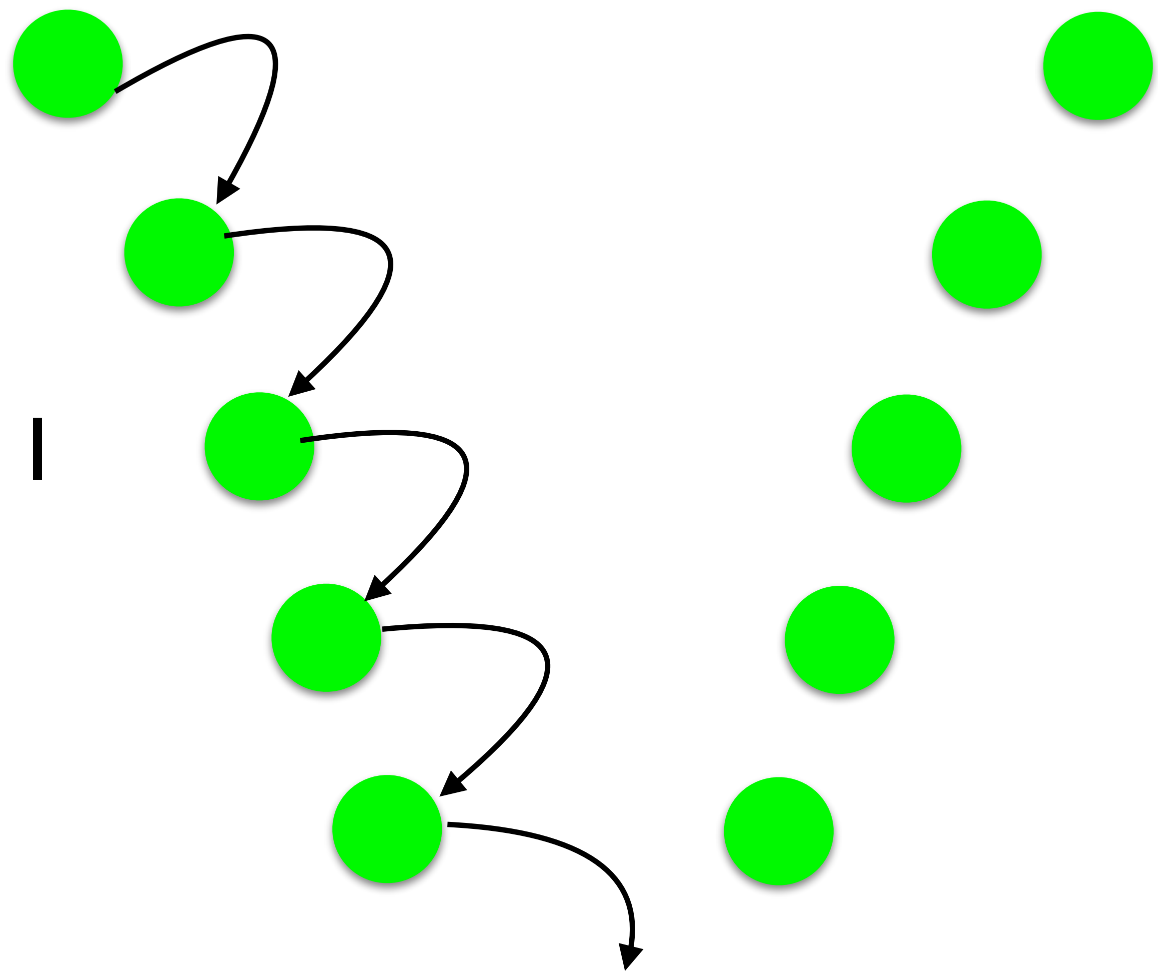


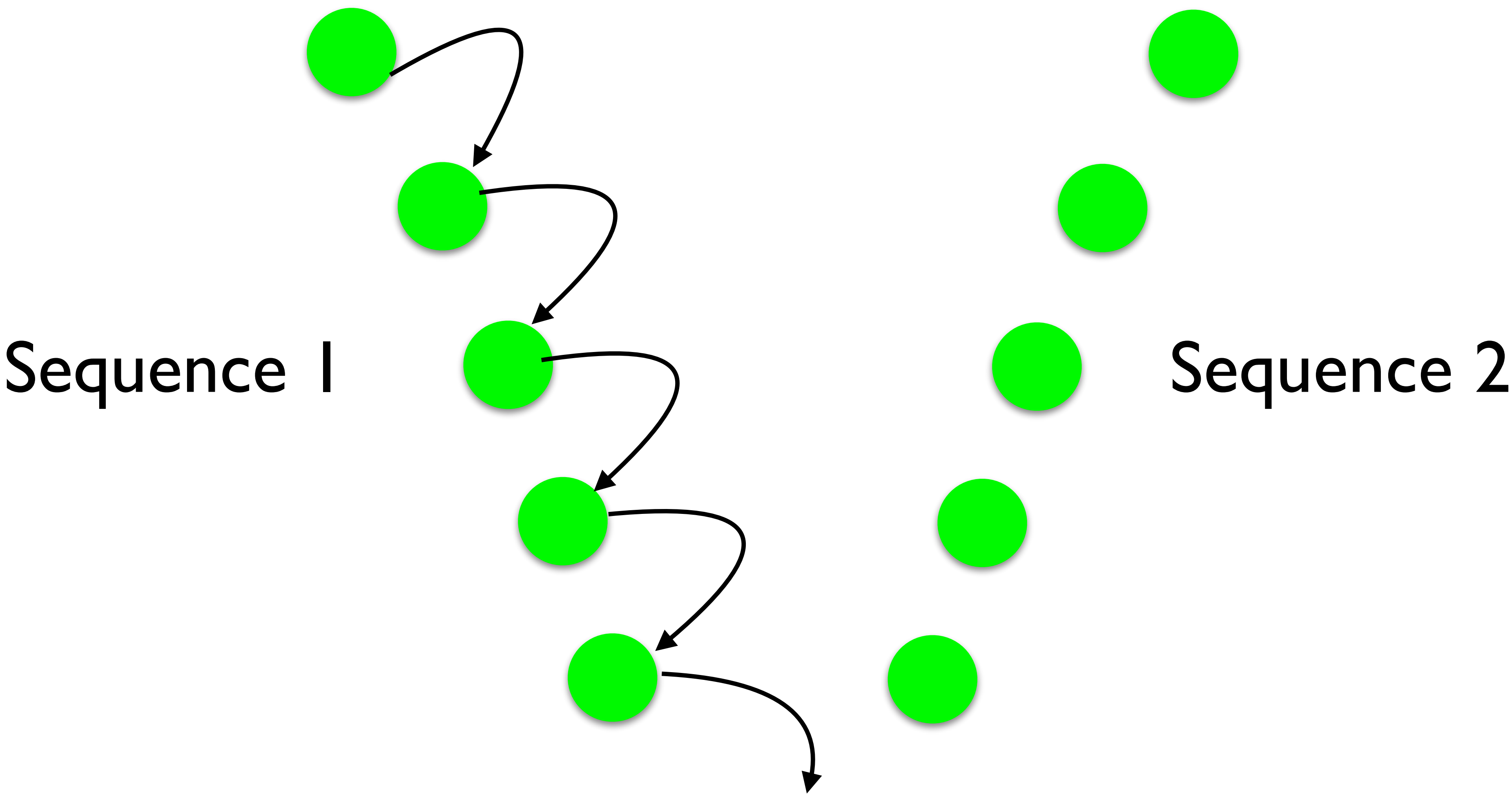

Sequence 1

Sequence 1

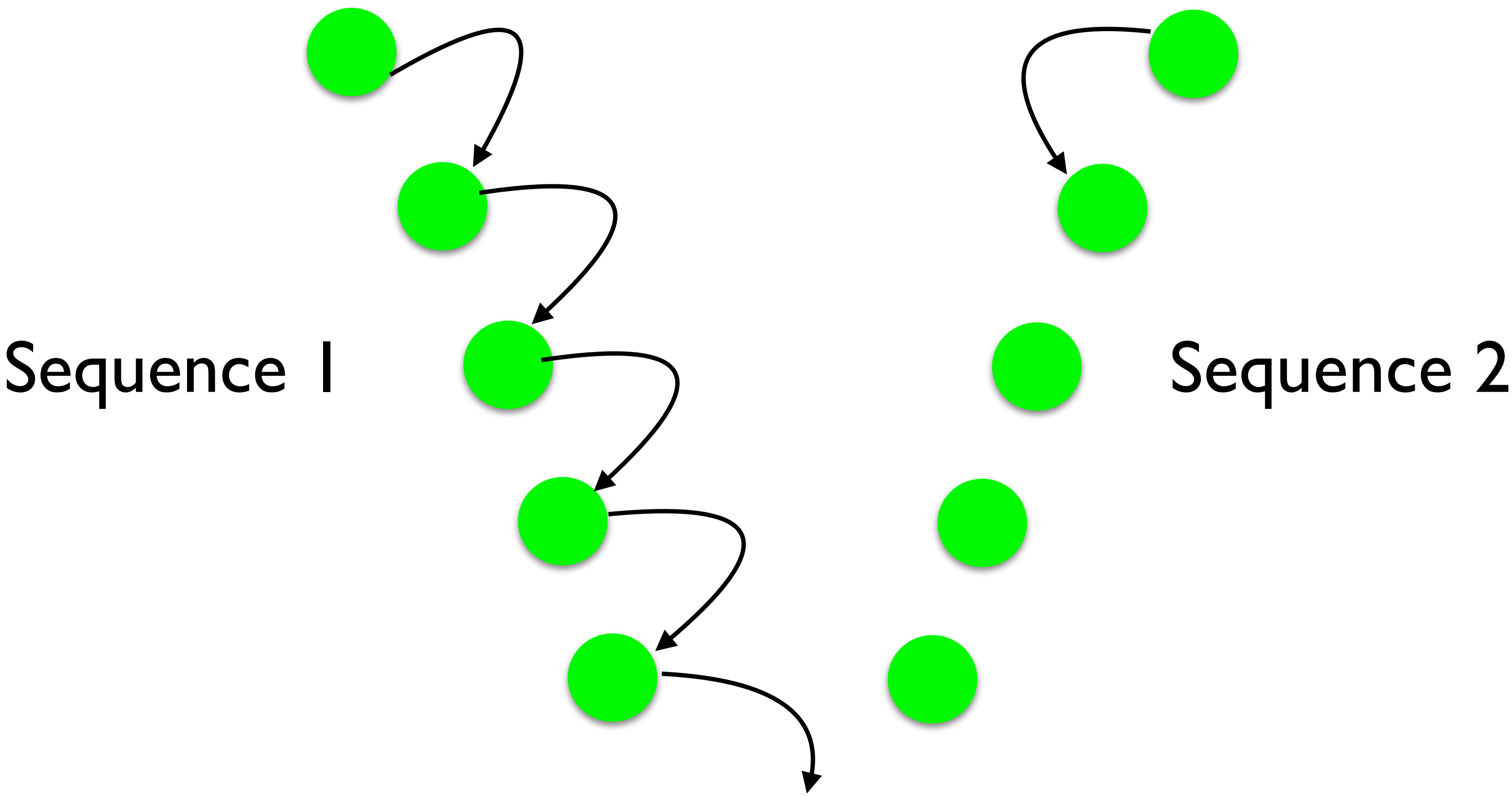


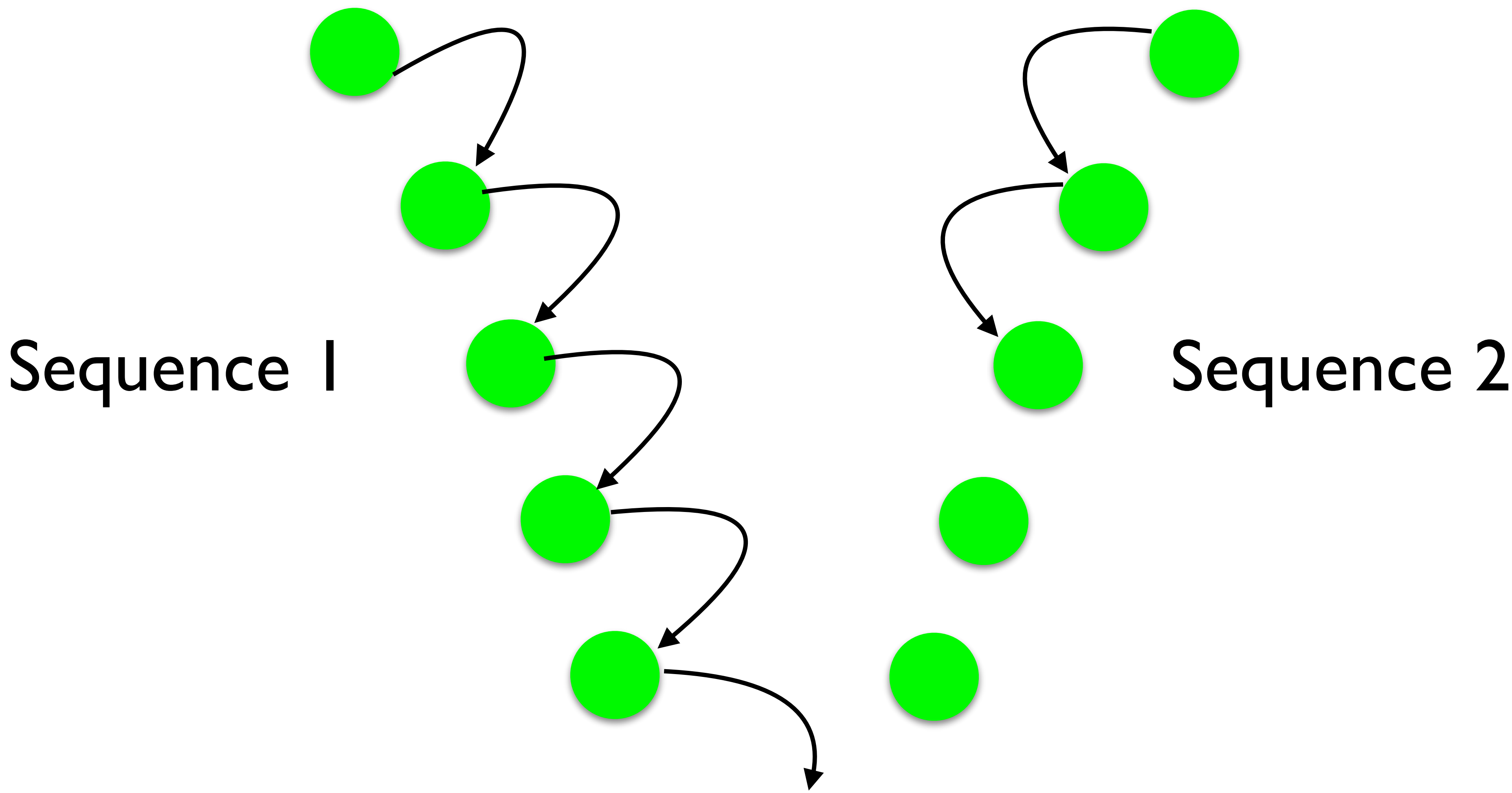
Sequence 1

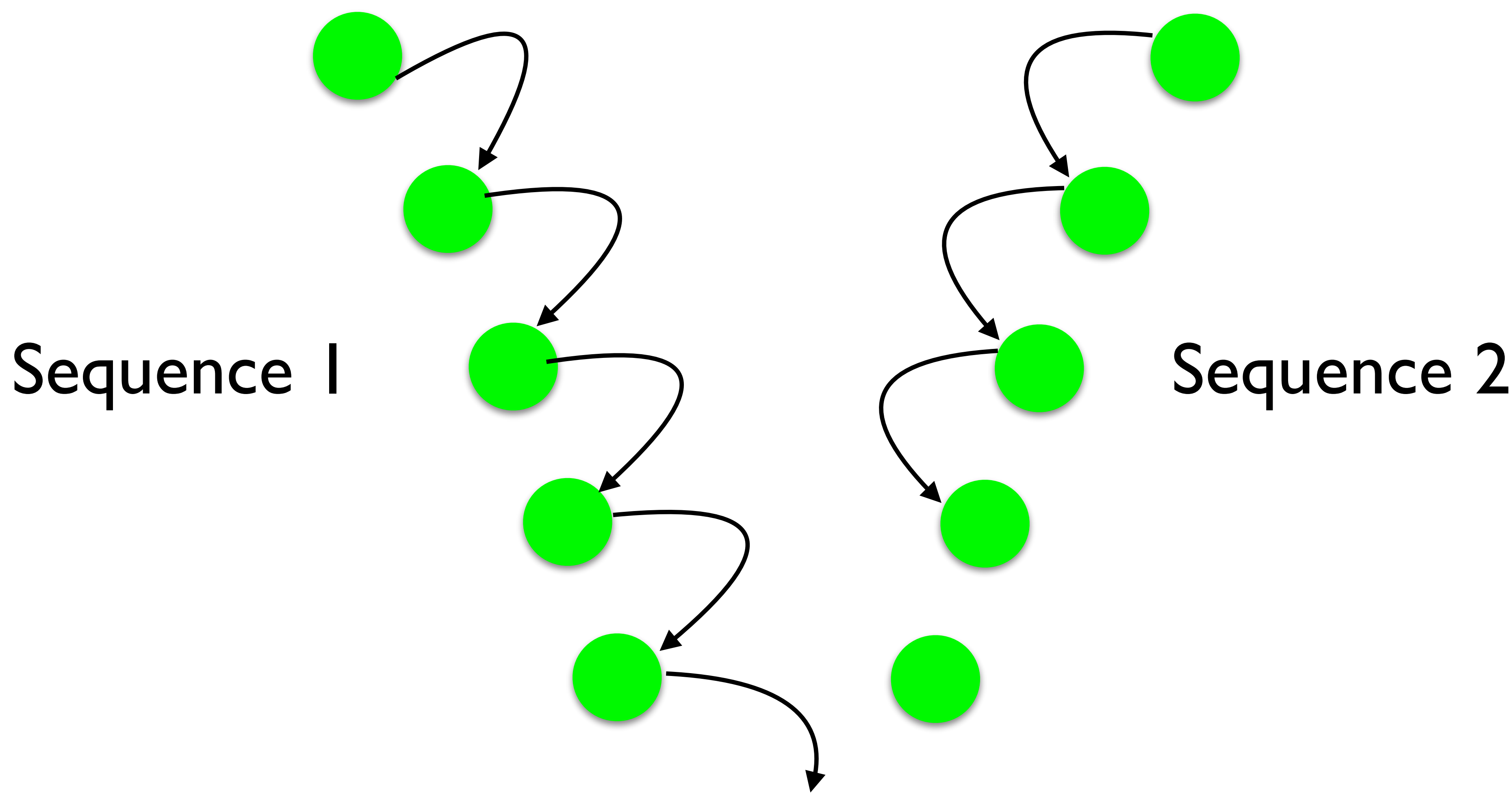




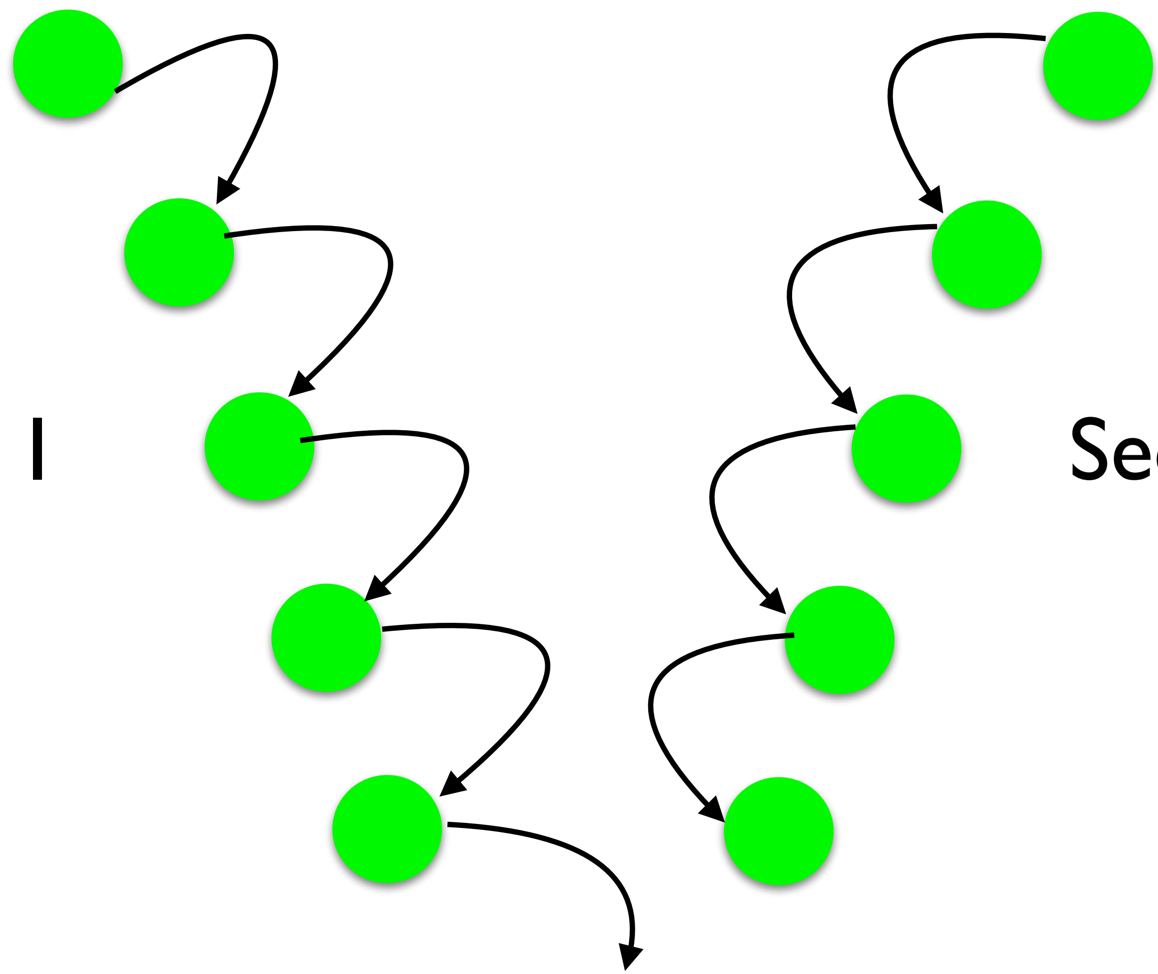
Sequence 1
Sequence 2


Sequence 1
Sequence 2

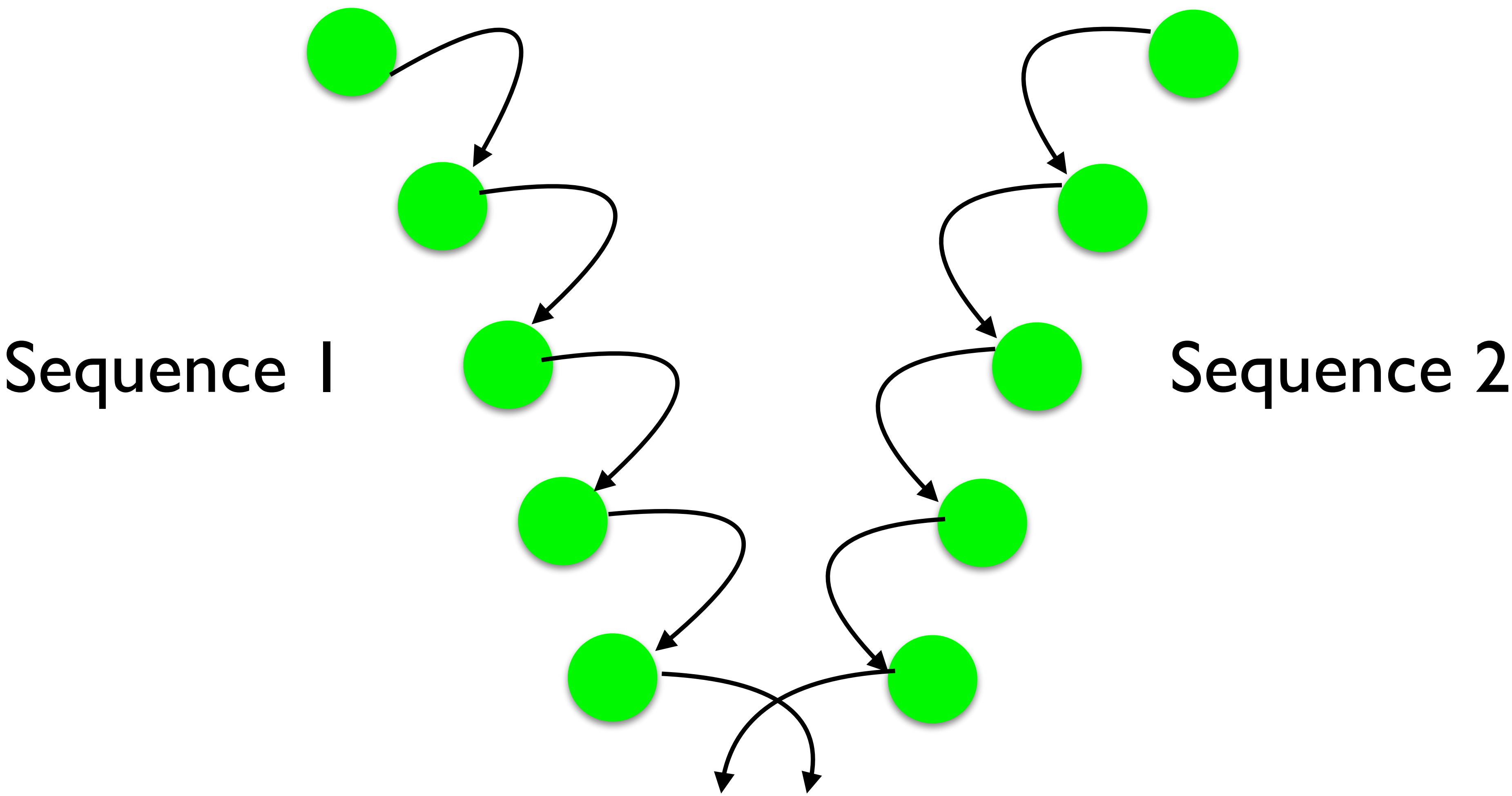

Sequence 1
Sequence 2

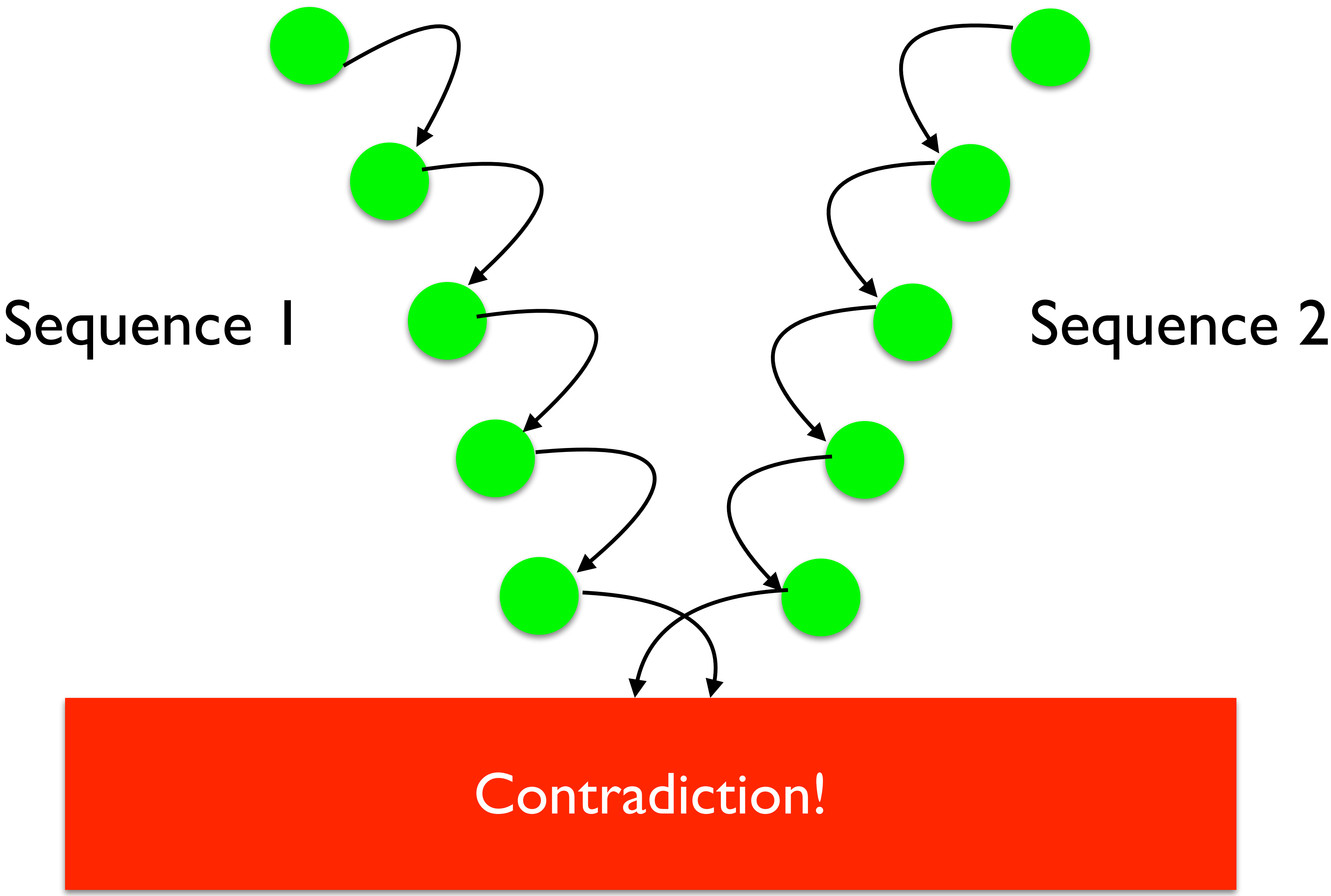

Sequence 1
Sequence 2

Sequence 1



Sequence 2


Sequence 1
Sequence 2


Sequence 1
Sequence 2
Contradiction!

Sequence 1

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

$$\exists t_i Wt_i \quad (2)$$

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

$$\exists t_i Wt_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

Sequence 1

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

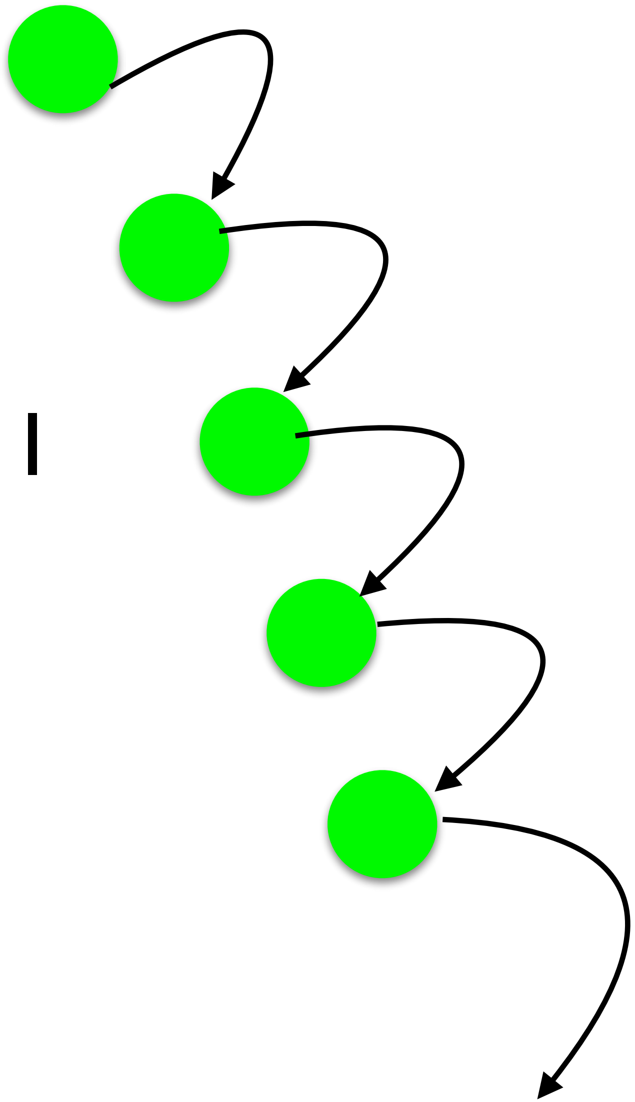
We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

$$\exists t_i Wt_i \quad (2)$$

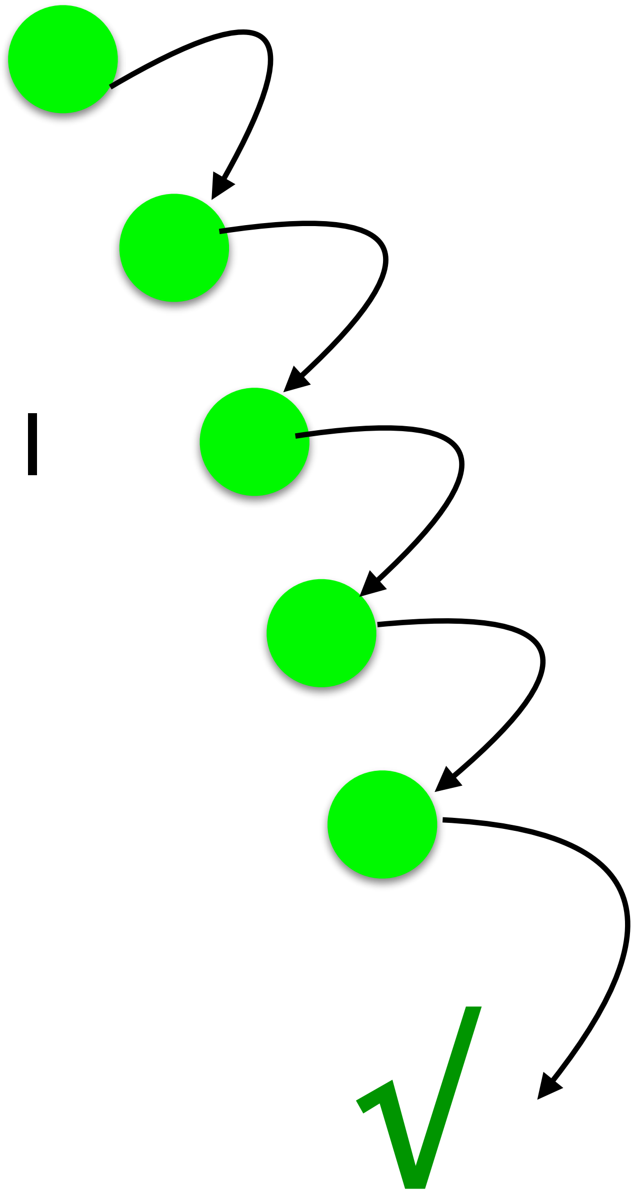
Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

$$\mathbf{B}_a \exists t_i Wt_i \quad (3)$$

Sequence 1



Sequence 1



Sequence 2

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \land prob(Wt_2) = \frac{1}{1T} \land \ldots \land prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_1 winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_1 won't win sails through— and this of course works for each ticket. Hence we have:

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_l winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_l won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg Wt_1 \wedge \mathbf{B}_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \land prob(Wt_2) = \frac{1}{1T} \land \ldots \land prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_l winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_l won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg Wt_1 \land \mathbf{B}_a \neg Wt_2 \land \ldots \land \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_I winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_I won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg Wt_1 \wedge \mathbf{B}_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a(\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_1 winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_1 won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg Wt_1 \wedge \mathbf{B}_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a(\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

Sequence 2

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_l winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_l won't win sails through— and this of course works for each ticket. Hence we have:

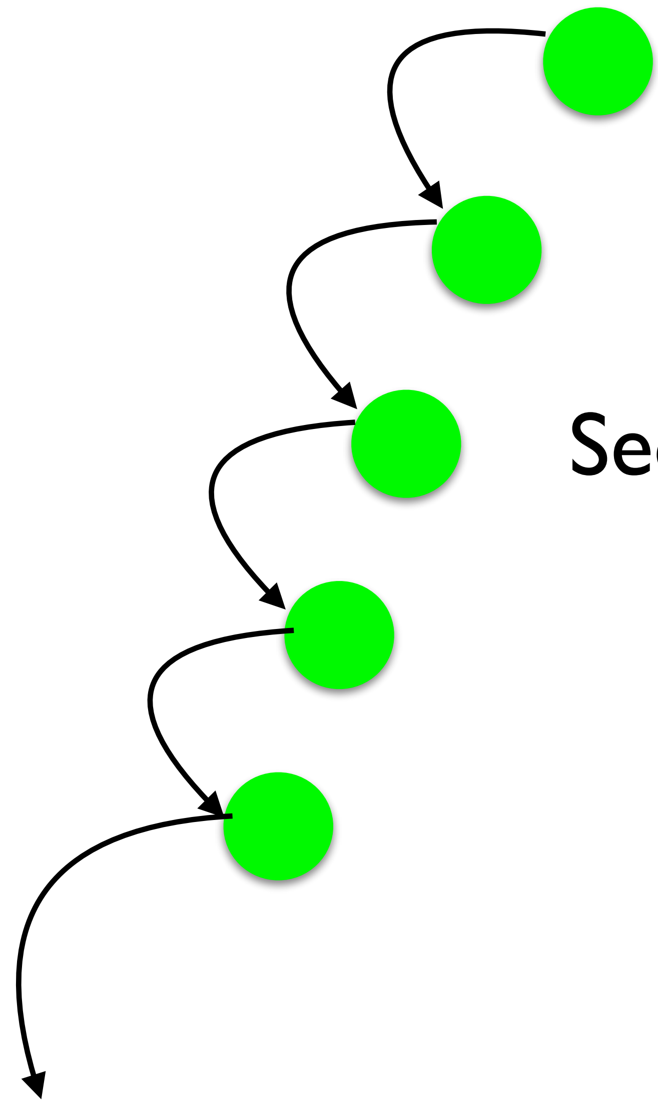
$$\mathbf{B}_a \neg Wt_1 \wedge \mathbf{B}_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

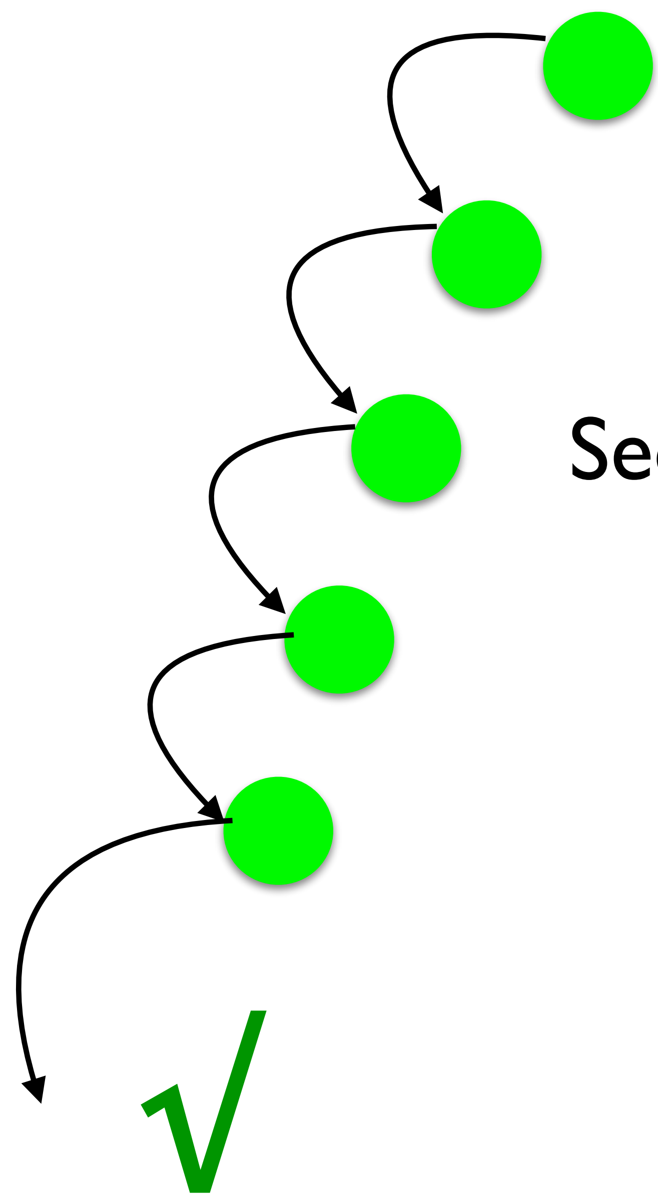
$$\mathbf{B}_a(\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

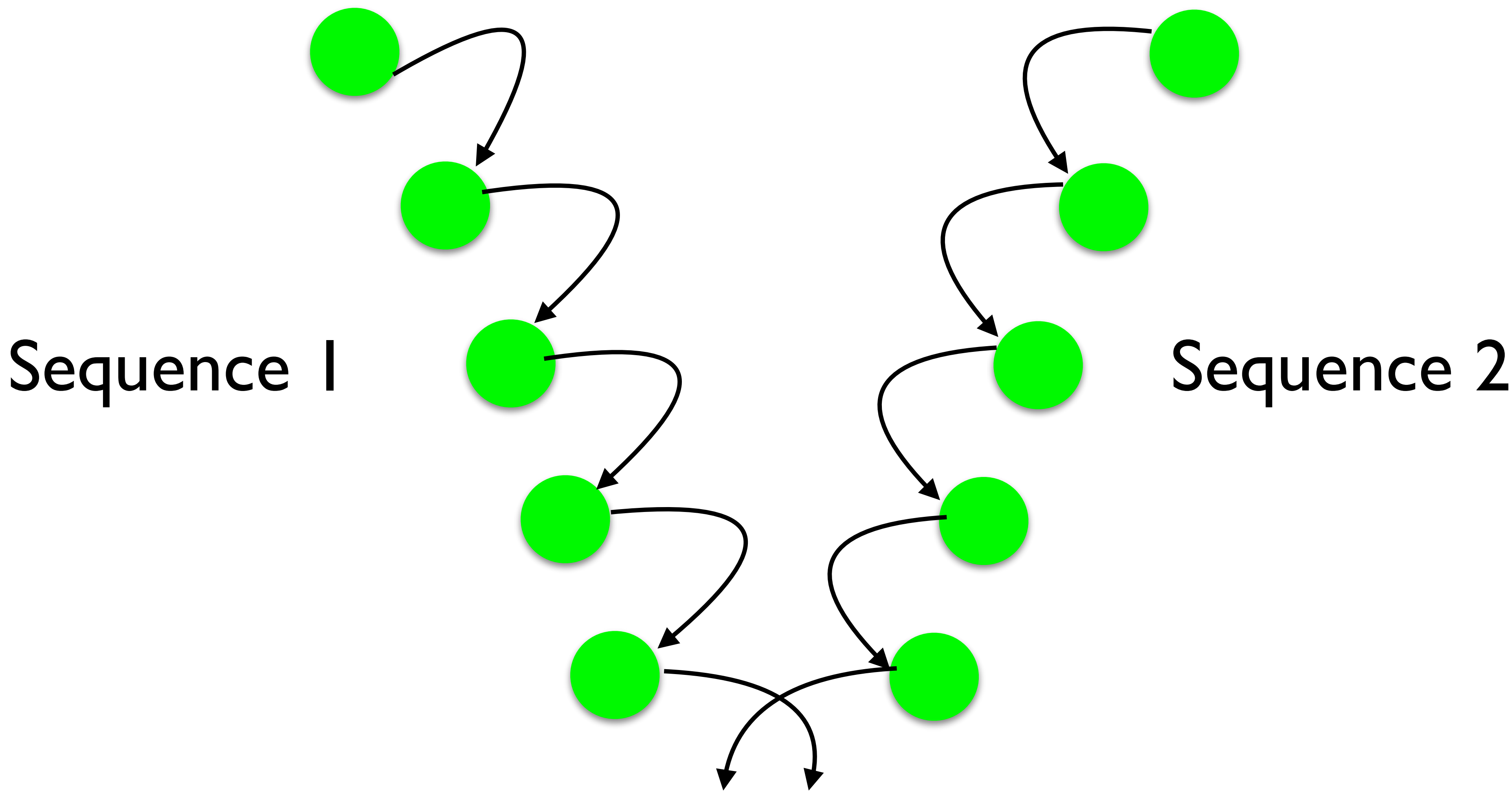
$$\mathbf{B}_a \neg \exists t_i Wt_i \quad (4)$$

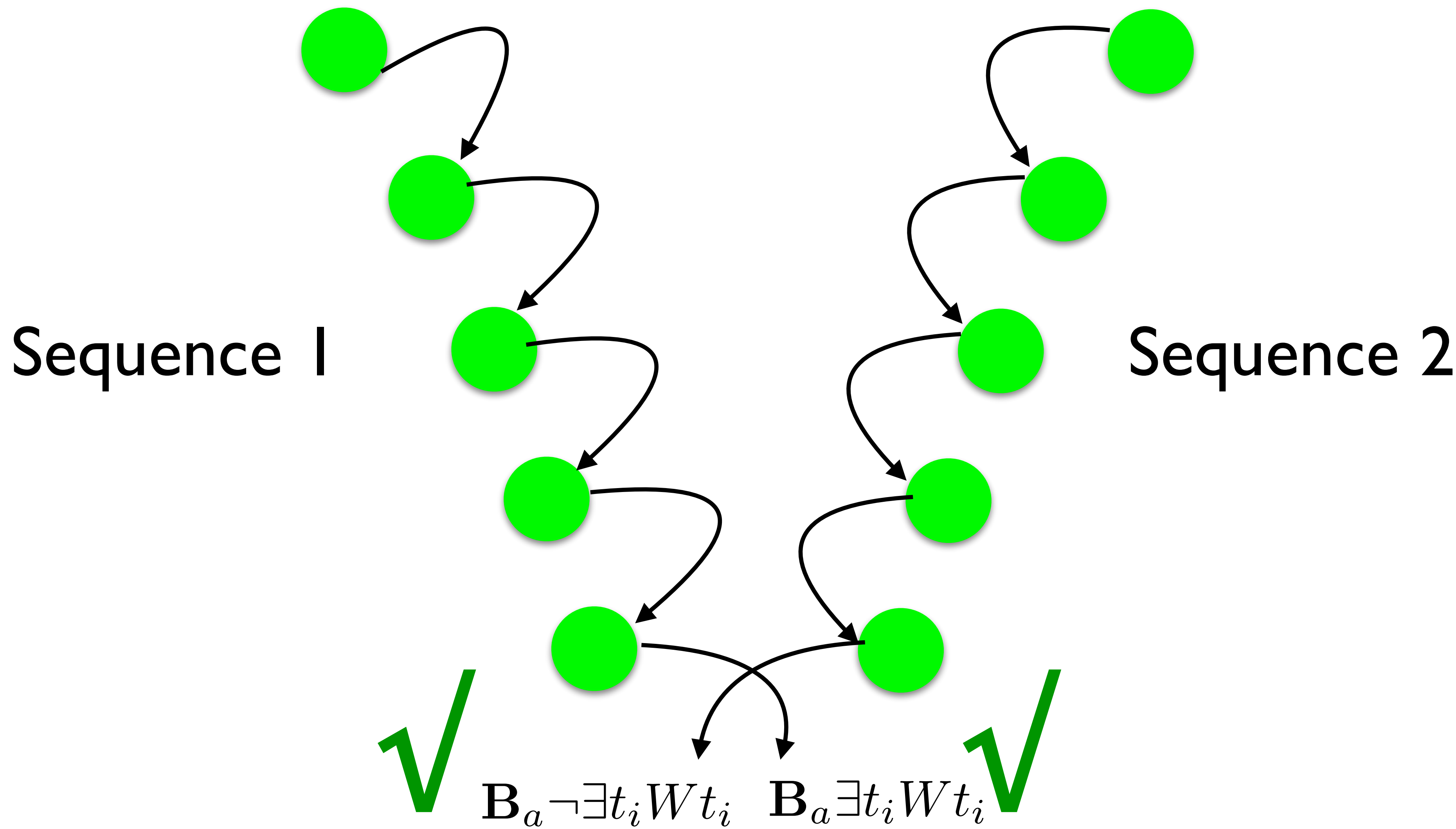


Sequence 2

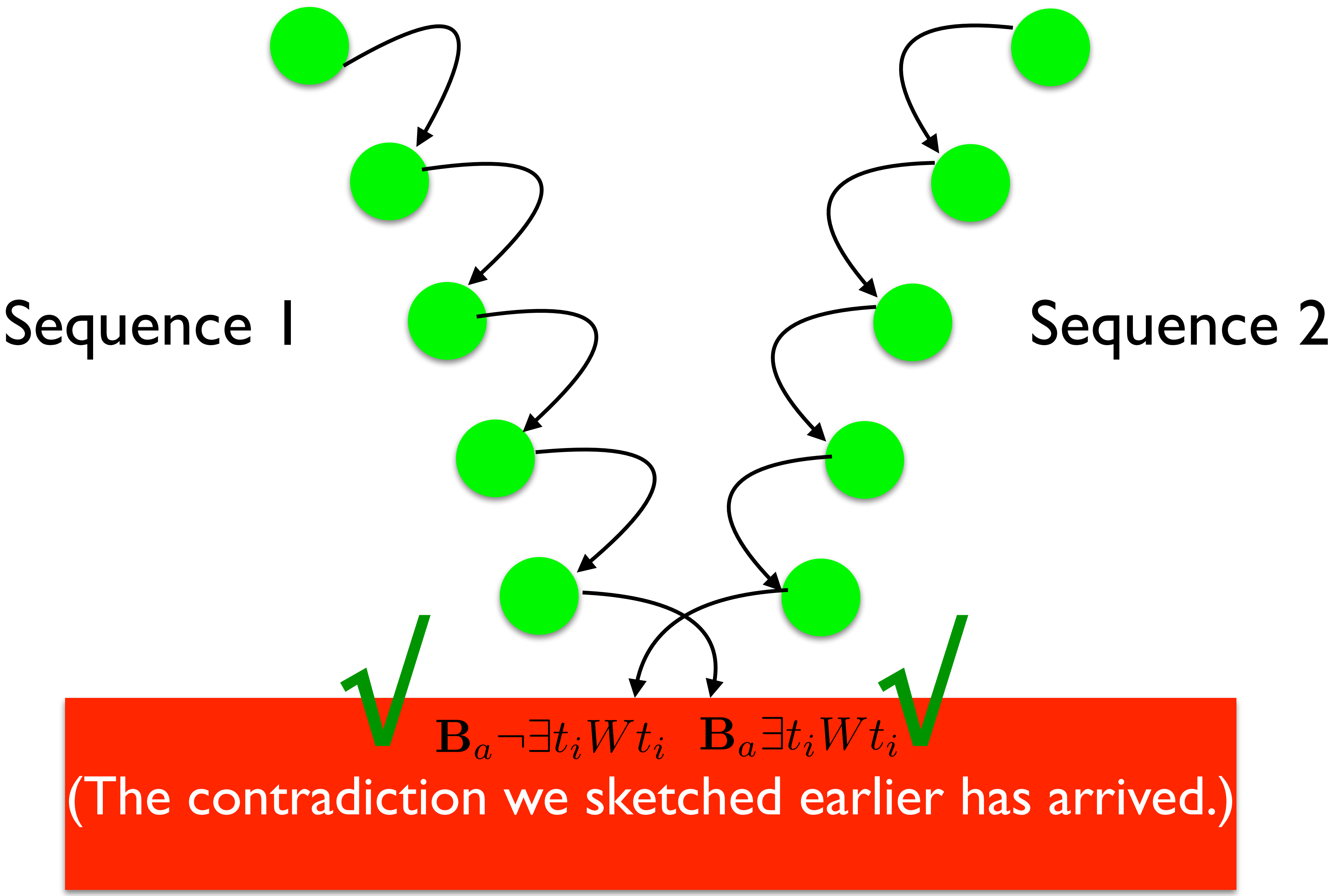


Sequence 2


Sequence 1
Sequence 2



Sequence 1
Sequence 2
$\mathbf{B}_a \neg \exists t_i W t_i$
$\mathbf{B}_a \exists t_i W t_i$


Sequence 1
Sequence 2
$\mathbf{B}_a \neg \exists t_i W t_i$
$\mathbf{B}_a \exists t_i W t_i$
(The contradiction we sketched earlier has arrived.)

A Solution to The Lottery Paradox …

Strength-Factor Continuum

Certain

Improbable

Evidently False

Probable

Beyond Reasonable Belief

Certainly False

Counterbalanced

Evident

Beyond Reasonable Doubt

Strength-Factor Continuum

Certain

Evident

Beyond Reasonable Doubt

Probable

Counterbalanced

Improbable

Beyond Reasonable Belief

Evidently False

Certainly False

Actually, now . . .

Actually, now …

English	Value
certain	6
evident	5
overwhelmingly likely = "beyond reasonable doubt" = "one in a million"	4
very likely	3
likely	2
more likely than not	1
counterbalanced	0

Actually, now …

English	Value
certain	6
evident	5
overwhelmingly likely = “beyond reasonable doubt” = “one in a million”	4
very likely	3
likely	2
more likely than not	1
counterbalanced	0

… but let’s use the simpler scheme.

Strength-Factor Continuum

Certain

Evident

Beyond Reasonable Doubt

Probable

Counterbalanced

Improbable

Beyond Reasonable Belief

Evidently False

Certainly False

Strength-Factor Continuum

Certain

Evident

Beyond Reasonable Doubt

Probable

Counterbalanced

Improbable

Beyond Reasonable Belief

Evidently False

Certainly False

Strength-Factor Continuum

Epistemically Positive

Certain

Evident

Beyond Reasonable Doubt

Probable

Counterbalanced

Improbable

Beyond Reasonable Belief

Evidently False

Certainly False

Strength-Factor Continuum

Epistemically Positive

Certain

Evident

Beyond Reasonable Doubt

Probable

Counterbalanced

Improbable

Beyond Reasonable Belief

Evidently False

Certainly False

Epistemically Negative

Strength-Factor Continuum

Epistemically Positive

Certain

Evident

Beyond Reasonable Doubt

Probable

Counterbalanced

Improbable

Beyond Reasonable Belief

Evidently False

Epistemically Negative

Certainly False

Strength-Factor Continuum

Epistemically Positive

(4) Certain

(3) Evident

(2) Beyond Reasonable Doubt

(1) Probable

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

Epistemically Negative

(-4) Certainly False

Strength-Factor Continuum

Epistemically Positive

(4) Certain

(3) Evident

(2) Beyond Reasonable Doubt

(1) Probable

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

(-4) Certainly False

Epistemically Negative

Epistemically Positive

(4) Certain

(3) Evident

(2) Beyond Reasonable Doubt

(1) Probable

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

(-4) Certainly False

Epistemically Negative

Key Principles

Epistemically Positive

(4) Certain

(3) Evident

(2) Beyond Reasonable Doubt

(1) Probable

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

(-4) Certainly False

Epistemically Negative

Key Principles

Epistemically Positive

(4) Certain

Deduction preserves strength.

(3) Evident

(2) Beyond Reasonable Doubt

(1) Probable

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

(-4) Certainly False

Epistemically Negative

Key Principles

Epistemically Positive

(4) Certain

Deduction preserves strength.

(3) Evident

(2) Beyond Reasonable Doubt

Clashes are resolved in favor of higher strength.

(1) Probable

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

Epistemically Negative

(-4) Certainly False

Key Principles

Epistemically Positive

(4) Certain

Deduction preserves strength.

(3) Evident

(2) Beyond Reasonable Doubt

Clashes are resolved in favor of higher strength.

(1) Probable

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

(0) Counterbalanced

(-1) Improbable

(-2) Beyond Reasonable Belief

(-3) Evidently False

Epistemically Negative

(-4) Certainly False

Key Principles

Epistemically Positive

(4) Certain

Deduction preserves strength.

(3) Evident

(2) Beyond Reasonable Doubt

Clashes are resolved in favor of higher strength.

(1) Probable

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

(0) Counterbalanced

(-1) Improbable

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

(-2) Beyond Reasonable Belief

(-3) Evidently False

Epistemically Negative

(-4) Certainly False

Sequence 1, "Rigorized"

Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

$$\exists t_i W t_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

$$\mathbf{B}_a \exists t_i W t_i \quad (3)$$

Sequence 1, "Rigorized"

4 Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

$$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

$$\exists t_i Wt_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

$$\mathbf{B}_a \exists t_i Wt_i \quad (3)$$

Sequence 1, "Rigorized"

4 Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

4 $$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

$$\exists t_i W t_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

$$\mathbf{B}_a \exists t_i W t_i \quad (3)$$

Sequence 1, “Rigorized”

4 Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let’s write this (exclusive) disjunction as follows:

4 $$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

4 $$\exists t_i W t_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

$$\mathbf{B}_a \exists t_i W t_i \quad (3)$$

Sequence 1, "Rigorized"

4 Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

4 $$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

4 $$\exists t_i W t_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

4 $$\mathbf{B}_a \exists t_i W t_i \quad (3)$$

Sequence 1, "Rigorized"

4 Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

4 $$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

4 $$\exists t_i Wt_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

4

Sequence 1, "Rigorized"

4 Let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that either ticket 1 will win or ticket 2 will win or … or ticket 1,000,000,000,000 will win. Let's write this (exclusive) disjunction as follows:

4 $$Wt_1 \oplus Wt_2 \oplus \ldots \oplus Wt_{1T} \quad (1)$$

We then deduce from this that there is at least one ticket that will win, a proposition represented — using standard notation — as:

4 $$\exists t_i Wt_i \quad (2)$$

Very well; perfectly clear so far. And now we can add another deductive step: Since our rational agent *a* can follow this deduction sequence to this point, and since **D** is assumed to be indubitable, it follows that our rational agent in turn believes (2); i.e., we conclude Sequence 1 by obtaining the following:

4 $$\mathbf{B}_a^4 \exists t_i Wt_i \quad (3)$$

Sequence 2, "Rigorized"

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_1 winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_1 won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg Wt_1 \wedge \mathbf{B}_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a(\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i Wt_i \quad (4)$$

Sequence 2, "Rigorized"

4

As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_I winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_I won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg W t_1 \wedge \mathbf{B}_a \neg W t_2 \wedge \ldots \wedge \mathbf{B}_a \neg W t_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a (\neg W t_1 \wedge \neg W t_2 \wedge \ldots \wedge \neg W t_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i W t_i \quad (4)$$

Sequence 2, "Rigorized"

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4 $$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_1 winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_1 won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a \neg Wt_1 \wedge \mathbf{B}_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a (\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i W t_i \quad (4)$$

Sequence 2, “Rigorized”

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using ‘1T’ to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4 $$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \land prob(Wt_2) = \frac{1}{1T} \land \ldots \land prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_1 winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_1 won’t win sails through— and this of course works for each ticket. Hence we have:

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a(\neg Wt_1 \land \neg Wt_2 \land \ldots \land \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn’t exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i Wt_i \quad (4)$$

Sequence 2, "Rigorized"

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4 $$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_l winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_l won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a^3 \neg Wt_1 \wedge \mathbf{B}_a^3 \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a^3 \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a (\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i W t_i \quad (4)$$

Sequence 2, "Rigorized"

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4
$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_l winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_l won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a^3 \neg Wt_1 \wedge \mathbf{B}_a^3 \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a^3 \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i W t_i \quad (4)$$

Sequence 2, “Rigorized”

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using ‘1T’ to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4 $$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \land prob(Wt_2) = \frac{1}{1T} \land \ldots \land prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_1 winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_1 won’t win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a^3 \neg Wt_1 \land \mathbf{B}_a^3 \neg Wt_2 \land \ldots \land \mathbf{B}_a^3 \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a^3 (\neg Wt_1 \land \neg Wt_2 \land \ldots \land \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn’t exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a \neg \exists t_i Wt_i \quad (4)$$

Sequence 2, "Rigorized"

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using '1T' to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4
$$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_l winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_l won't win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}^3_a \neg Wt_1 \wedge \mathbf{B}^3_a \neg Wt_2 \wedge \ldots \wedge \mathbf{B}^3_a \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}^3_a (\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn't exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

Sequence 2, “Rigorized”

4 As in Sequence 1, once again let **D** be a meticulous and perfectly accurate description of a 1,000,000,000,000-ticket lottery, of which rational agent *a* is fully apprised.

From **D** it obviously can be proved that the probability of a particular ticket t_i winning is 1 in 1,000,000,000,000. Using ‘1T’ to denote 1 trillion, we can write the probability for each ticket to win as a conjunction:

4 $$prob(Wt_1) = \frac{1}{1,000,000,000,000} = \frac{1}{1T} \wedge prob(Wt_2) = \frac{1}{1T} \wedge \ldots \wedge prob(Wt_{1T}) = \frac{1}{1T} \quad (1)$$

For the next step, note that the probability of ticket t_I winning is lower than, say, the probability that if you walk outside a minute from now you will be flattened on the spot by a door from a 747 that falls from a jet of that type cruising at 35,000 feet. Since you have the rational belief that death won't ensue if you go outside (and have this belief precisely because you believe that the odds of your sudden demise in this manner are vanishingly small), the inference to the rational belief on the part of *a* that t_I won’t win sails through— and this of course works for each ticket. Hence we have:

$$\mathbf{B}_a^3 \neg Wt_1 \wedge \mathbf{B}_a^3 \neg Wt_2 \wedge \ldots \wedge \mathbf{B}_a^3 \neg Wt_{1T} \quad (2)$$

Of course, if a rational agent believes *P*, and believes *Q* as well, it follows that that agent will believe the conjunction *P* & *Q*. Applying this principle to (2) yields:

$$\mathbf{B}_a^3 (\neg Wt_1 \wedge \neg Wt_2 \wedge \ldots \wedge \neg Wt_{1T}) \quad (3)$$

But (3) is logically equivalent to the statement that there doesn’t exist a winning ticket; i.e., (3) is logically equivalent to this result from Sequence 2:

$$\mathbf{B}_a^3 \neg \exists t_i Wt_i \quad (4)$$

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathbf{B}_a^4 \exists t_i W t_i \quad (3)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathrm{B}_a^4 \exists t_i W t_i \quad (3) \qquad \mathrm{B}_a^3 \neg \exists t_i W t_i \quad (4)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathrm{B}_a^4 \exists t_i W t_i \quad (3) \qquad \mathrm{B}_a^3 \neg \exists t_i W t_i \quad (4)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathbf{B}^{4}_{a}\exists t_i W t_i (4) \quad \mathbf{B}^{3}_{a}\exists t_i W t_i (3)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathbf{B}_a^4 \exists t_i W t_i \quad (3)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathbf{B}_a^4 \exists t_i W t_i \quad (3)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

$$\mathbf{B}_a^3 \neg W t_1 \wedge \mathbf{B}_a^3 \neg W t_2 \wedge \ldots \wedge \mathbf{B}_a^3 \neg W t_{1T} \quad (2)$$

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathbf{B}_a^4 \exists t_i W t_i \quad (3)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

$$\mathbf{B}_a^3 \neg W t_1 \wedge \boxed{\mathbf{B}_a^3 \neg W t_2} \wedge \ldots \wedge \mathbf{B}_a^3 \neg W t_{1T} \quad (2)$$

Paradox Solved!

Deduction preserves strength.

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction; if no higher-strength factors, suspend belief.

$$\mathbf{B}_a^4 \exists t_i W t_i \quad (3)$$

Any proposition *p* such that *prob*(*p*) < 1 is at most evident.

Any rational belief that *p*, where the basis for *p* is at most evident, is at most an evident (= level 3) belief.

$$\mathbf{B}_a^3 \neg W t_1 \wedge \boxed{\mathbf{B}_a^3 \neg W t_2} \wedge \ldots \wedge \mathbf{B}_a^3 \neg W t_{1T} \quad (2)$$

This is why, to Mega Millions ticket holder:
"Sorry. I'm rational, and I believe you won't win."

To be clear about the effects of the first principle:

$$\vdash \mathbf{B}^3_a \neg \cancel{\exists} x W x \wedge \mathbf{B}^3_a \exists x W x!$$

$$\vdash \mathbf{B}^2_a \neg \cancel{\exists} x W x \wedge \mathbf{B}^2_a \exists x W x!$$

$$\vdash \mathbf{B}^1_a \neg \cancel{\exists} x W x \wedge \mathbf{B}^1_a \exists x W x!$$

Clashes are resolved in favor of higher strength; clashes propagate backwards through inverse deduction, preserving affirmation/belief of premises as far as is possible; if no higher-strength factors, suspend belief. (This means that in this case belief at level 4 also shoots down belief at level 2, and level 1. This is sort of bizarre, because to retain the belief (at levels 3, 2, 1) that every particular ticket won't win, the step that gets to believing the existential formula is blocked. Pollock doesn't have steps in his "arguments." Our agents thus ends up believing at all levels that some ticket will win, and believing at all levels 3 and down, of each particular ticket, that it won't win.)

Engineer and apply this in the real world, to save lives …

DE GRUYTER Paladyn, Journal of Behavioral Robotics 2021; 12: 310–335

Research Article

Selmer Bringsjord, Naveen Sundar Govindarajulu*, and Michael Giancola

Automated argument adjudication to solve ethical problems in multi-agent environments

https://doi.org/10.1515/pjbr-2021-0009
received May 7, 2020; accepted December 30, 2020

Abstract: Suppose an artificial agent a_{adj}, as time unfolds, (i) receives from multiple artificial agents (which may, in turn, themselves have received from yet other such agents…) propositional content, and (ii) must solve an ethical problem on the basis of what it has received. How should a_{adj} adjudicate what it has received in order to produce such a solution? We consider an environment infused with logicist artificial agents $a_1, a_2, \ldots, a_n$ that sense and report their findings to "adjudicator" agents who must solve ethical problems. (Many if not most of these agents may be robots.) In such an environment, inconsistency is a virtual guarantee: a_{adj} may, for instance, receive a report from a_1 that proposition ϕ holds, then from a_2 that $\neg\phi$ holds, and then from a_3 that neither ϕ nor $\neg\phi$ should be believed, but rather ψ instead, at some level of likelihood. We further assume that agents receiving such incompatible reports will nonetheless sometimes simply need, before long, to make decisions on the basis of these reports, in order to try to solve ethical problems. We provide a solution to such a quandary: AI capable of adjudicating competing reports from subsidiary agents through time, and delivering to humans a rational, ethically correct (relative to underlying ethical principles) recommendation based upon such adjudication. To illuminate our solution, we anchor it to a particular scenario.

Keywords: cognitive calculus, cognitive robotics, adjudication, adjudication for ethical principles, ethics

1 Introduction

Neurobiologically normal, mature human beings often enjoy the luxury of being able to make decisions in and unto themselves. A hot burner on a stove, if mistakenly touched, can lead to a rather quick decision to pull away; and while such a decision usually happens by reflex, the human in question can then inspect his/her finger and decide whether or not treatment is needed. But as we know, decision-making is not always this independent; sometimes what humans must decide must factor in what has been received from other humans. When this happens, the situation can be quite tricky. Perhaps this is especially true when the required decision is needed in order to try to resolve some ethical problem. Note that in the course of human affairs, profound ethical decisions have long needed to be made in these kinds of buzzing, dynamic, dialectical, multi-agent scenarios, where all the agents are humans. Deep and challenging legal cases provide a case in point,[1] as for that matter so do command-and-control challenges to humans in warfare, a domain that our case study given below relates to.[2] But our task herein is to formalize the AI correlate of this kind of tricky situation and to propose a way for a new kind of AI to solve the correlate.

This AI correlate, in broad strokes for the moment, has the following structure: An artificial agent a_{adj}, as

*** Corresponding author: Naveen Sundar Govindarajulu,** Rensselaer AI & Reasoning (RAIR) Lab, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, United States of America, e-mail: naveen.sundar.g@gmail.com
Selmer Bringsjord: Department of Cognitive Science, Rensselaer AI & Reasoning (RAIR) Lab, Troy, NY 12180, United States of America; Department of Computer Science, Rensselaer AI & Reasoning (RAIR) Lab, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, United States of America, e-mail: selmer.bringsjord@gmail.com
Michael Giancola: Department of Computer Science, Rensselaer AI & Reasoning (RAIR) Lab, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, United States of America, e-mail: mike.j.giancola@gmail.com

1 It would, e.g., be quite interesting to see how an artificial agent of the type introduced in the present paper would, if suitably "armed" with starting information, adjudicate the Dreyfus case, covered brilliantly in literary fashion by Proust [68], and in hard-nosed journalistic fashion in ref. [69].

2 For a deeper, more complex case study in this domain, it would be interesting to see whether decision-making as to when to engage Pershing's new-world forces in WW I, which involved many a mind interacting with Pershing's, could be automated. For background, see ref. [70].

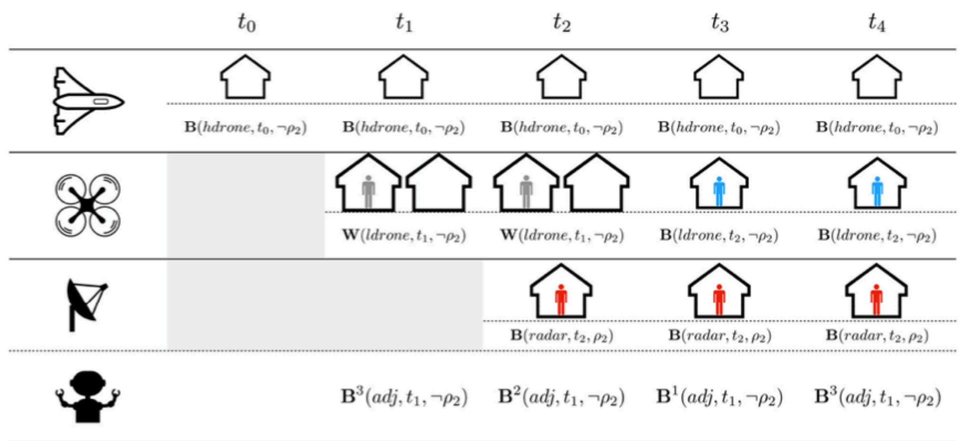


Figure 3: Overview of the scenario.

and leave more robust implementations to future work. The third step is also admittedly incomplete for this particular case study; however, full completion would not require any new research: one would simply follow the processes outlined in ref. [8,13]. Finally, Step 4 is also only partially finished, as its full completion is precluded by the merely partial implementation of Step 3. Overall, though, it should be clear that our Four Steps have been followed.

4 Discussion

We are under no such illusion as that our work will be embraced immediately by all. In general, we at this point anticipate two general classes of objections: one that contains technical worries, and a second aimed at alleged fundamental flaws in logicist AI, at least as such AI is pursued by us. In what now follows, in conformity with this two-part division, we first discuss a class of objections that relate to limitative theorems due, at least originally, to Arrow; and then, we proceed to present and rebut objections that claim our methodology is missing something crucial.

4.1 Dialectic arising from arrow's impossibility theorem and successors

In point of fact, there is no denying that Arrow's Impossibility Theorem (AIT) is directly relevant, logico-mathematically *and* implementation-wise, to our framework and technology for adjudication in multi-agent contexts. However, we cannot expect our readers in the present case to be familiar with AIT (very nicely presented and proved in ref. [40], and ably summarized without proof in ref. [41]). Hence, we must find a shortcut here; and we do, as follows. We can without loss of generality at the current juncture take AIT to be based upon the existence of n artificial agents $a_1, \ldots, a_n$ whose action repertoire consists solely in each of

Table 3: Overview of the beliefs. The adjudicator's beliefs about other agents' beliefs and its uncertainty level in $\neg\rho_2$

Time	hdrone	ldrone	Radar	Strength for $(adj, t, \neg\rho_2)$
t_1	$\mathbf{B}(hdrone, t_0, \neg\rho_2)$	Not considered	Not considered	$\mathbf{B}^3(\text{adj}, t_1, \neg\rho_2)$
t_2	$\mathbf{B}(hdrone, t_0, \neg\rho_2)$	$\mathbf{W}(ldrone, t_1, \neg\rho_2)$	Not considered	$\mathbf{B}^2(\text{adj}, t_2, \neg\rho_2)$
t_3	$\mathbf{B}(hdrone, t_0, \neg\rho_2)$	$\mathbf{W}(ldrone, t_1, \neg\rho_2)$	$\mathbf{B}(radar, t_2, \rho_2)$	$\mathbf{B}^1(\text{adj}, t_3, \neg\rho_2)$
t_4	$\mathbf{B}(hdrone, t_0, \neg\rho_2)$	$\mathbf{B}(ldrone, t_2, \neg\rho_2)$	$\mathbf{B}(radar, t_2, \rho_2)$	$\mathbf{B}^3(\text{adj}, t_4, \neg\rho_2)$

slutten