FOL = $\mathscr{L}_1$ Problems on HG®; $\mathscr{Q}^1_2$ "Aristotle" Problems & DFT⁺ From You in Logicist AI Work; Intensional Logic & "The *TOS*' Changeling"; Real Learning Glimpse & Paper

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

Intro to Logic-based AI
10/10/2024

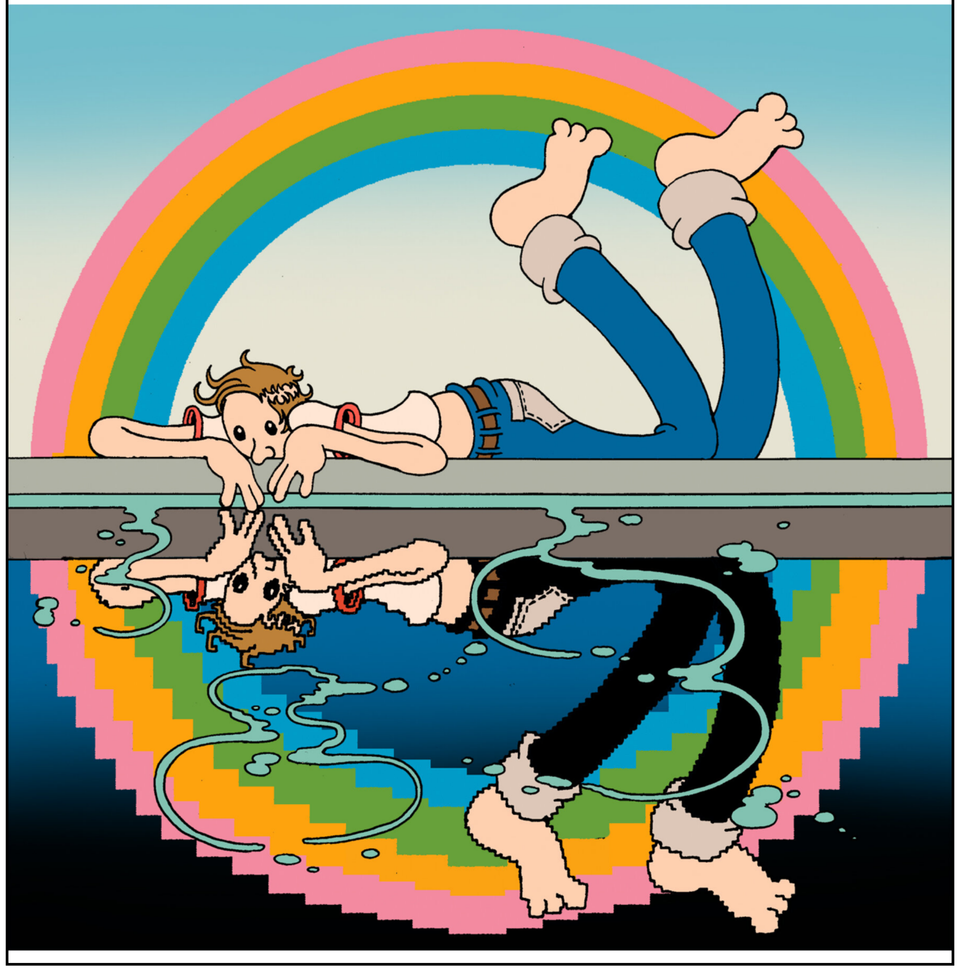
In The Logic-and-AI News

…

OPINION
GUEST ESSAY

I Created an A.I. Voice Clone to Prank Telemarketers. But the Joke's on Us.

Oct. 10, 2024, 5:02 a.m. ET



By Evan Ratliff
Mr. Ratliff is a journalist and the creator of the podcast "Shell Game."

Earlier this year I called up my good friend Warren to talk about a soccer game we were about to watch on two different coasts. "You pumped for the game tonight?" I asked. "What? Of course I'm pumped," he said, as we proceeded with our normal pregame chatter. Then Warren noticed something: "You're speaking in these bite-size chunks that make it sound like maybe this is an A.I. conversation."

He had me there.

The "I" in our call was not me at all but a voice agent I'd created using a professional-grade artificial intelligence clone of my voice. The voice bot was powered by ChatGPT and attached to my phone number — a process that takes less than an hour and is easy for anyone to replicate. As an experiment, I've been sending my voice agent out into the world for most of the last year for a podcast called "Shell Game," about how strangers, colleagues and friends respond to sudden encounters with the A.I. Evan Ratliff.

What I've learned is that interacting with A.I. voice agents will change how we interact with

Rapid advances in artificial intelligence have tended to spur three broad reactions. Champions of A.I. spin up utopian visions of hyper-efficiency and machine brilliance. Skeptics claim it's an overhyped technology that's already hitting a wall. Alarmists sound warnings about A.I.'s most grandiose dangers, predicting it could sweep away whole industries or escape our control. These competing visions obscure an unavoidable reality: A.I. agents are already triggering an avalanche of synthetic conversation, as they are deployed as tireless, unflagging talkers, capable of endless invented chatter. As they improve, it will become increasingly difficult to distinguish these A.I. voice agents from humans and, even when you *can* identify them, you will still be forced to talk to them.

Advocates of A.I. try to sell these agents as helpful digital assistants to schedule our appointments or friends who'll always be there to listen. But the more simulated human conversation I heard, the more it left me craving the real thing: in-person connections with the people I care about, with all the quirks of a meandering human discussion. If the coming onslaught of humanlike A.I. conversation threatens to fill our world with made-up verbal detritus, an audio version of "A.I. slop," then the upside may be that it forces us to appreciate the subtleties of personal interactions that many of us have come to devalue.

voice agents will change how we interact with

A Godfather of AI Just Won a Nobel. He Has Been Warning the Machines Could Take Over the World.

Geoffrey Hinton hopes the prize will add credibility to his claims about the dangers of AI technology he pioneered

Geoffrey Hinton has warned that AI systems could escape human control. (CHRIS YOUNG/CANADIAN PRESS/AP)

By Miles Kruppa + Follow *and Deepa Seetharaman* + Follow

Updated Oct 09, 2024 07:42 a.m. ET

A Godfather of AI Just Won a Nobel. He Has Been Warning the Machines Could Take Over the World.

Geoffrey Hinton hopes the prize will add credibility to his claims about the dangers of AI technology he pioneered

Geoffrey Hinton has warned that AI systems could escape human control. (CHRIS YOUNG/CANADIAN PRESS/AP)

By Miles Kruppa + Follow and Deepa Seetharaman + Follow

Updated Oct 09, 2024 07:42 a.m. ET

The newly minted Nobel laureate Geoffrey Hinton has a message about the artificial-intelligence systems he helped create: get more serious about safety or they could endanger humanity.

"I think we're at a kind of bifurcation point in history where, in the next few years, we need to figure out if there's a way to deal with that threat," Hinton said in an interview Tuesday with a Nobel Prize official that mixed pride in his life's work with warnings about the growing danger it poses.

The 76-year-old Hinton resigned from Google last year in part so he could talk more about the possibility that AI systems could escape human control and influence elections or power dangerous robots. Along with other experienced AI researchers, he has called on such companies as OpenAI, Meta Platforms META -0.93% ▼ and Alphabet GOOGL +0.38% ▲ -owned Google to devote more resources to the safety of the advanced systems that they are competing against each other to develop as quickly as possible.

Hinton's Nobel win has provided a new platform for his doomsday warnings at the same time it celebrates his critical role in advancing the technologies fueling them. Hinton has argued

Our Agenda; Here We Go

Our Agenda; Here We Go

- FOL Required Problems for You

Our Agenda; Here We Go

- FOL Required Problems for You
  - Includes a creativity problem based on ramping up Aristotle to created $\mathcal{Q}_2$ DFT$^+$ problems!

Our Agenda; Here We Go

- FOL Required Problems for You
    - Includes a creativity problem based on ramping up Aristotle to created $\mathcal{Q}_2$ DFT$^+$ problems!
    - I explain w/ doc camera & pen & paper.

Our Agenda; Here We Go

- FOL Required Problems for You
  - Includes a creativity problem based on ramping up Aristotle to created $\mathcal{Q}_2$ DFT+ problems!
  - I explain w/ doc camera & pen & paper.
- Intensional Logic: Nomad/The Changeling; FBTasks

Our Agenda; Here We Go

- FOL Required Problems for You
  - Includes a creativity problem based on ramping up Aristotle to created $\mathcal{Q}_2$ DFT+ problems!
  - I explain w/ doc camera & pen & paper.
- Intensional Logic: Nomad/The Changeling; FBTasks
  - Assignment: Watch episode. What would it take for an AI, not Kirk, to save the day against Nomad.

Our Agenda; Here We Go

- FOL Required Problems for You
  - Includes a creativity problem based on ramping up Aristotle to created $\mathcal{Q}_2$ DFT$^+$ problems!
  - I explain w/ doc camera & pen & paper.
- Intensional Logic: Nomad/The Changeling; FBTasks
  - Assignment: Watch episode. What would it take for an AI, not Kirk, to save the day against Nomad.
- Real Learning … & Assignment: Read the paper.

Our Agenda; Here We Go

- FOL Required Problems for You
  - Includes a creativity problem based on ramping up Aristotle to created $\mathscr{Q}_2$ DFT+ problems!
  - I explain w/ doc camera & pen & paper.
- Intensional Logic: Nomad/The Changeling; FBTasks
  - Assignment: Watch episode. What would it take for an AI, not Kirk, to save the day against Nomad.
- Real Learning … & Assignment: Read the paper.
- Finally: *2001* came out on top of movies; please watch for next class as well (i.e. for Thu Oct 17).

Theory-of-Mind Reasoning/ Planning in *TOS* "The Changeling" …

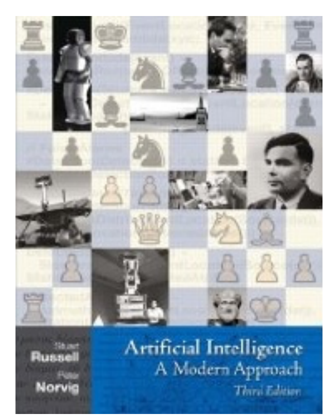
Dealt with deeply later, but for now, sufficient to leave informally put: “We’re in *very* deep trouble.”

Dealt with deeply later, but for now, sufficient to leave informally put: "We're in *very* deep trouble."




Dealt with deeply later, but for now, sufficient to leave informally put: “We’re in *very* deep trouble.”



While the PAI machines aren't quite as easy to neutralize as the destructive machines vanquished in *Star Trek: TOS*, these relevant four episodes show the protective power of … logic.

"The Ultimate Computer"
S2 E24

"The Return of the Archons"
S1 E21

"The Changeling"
S2 E3

"I, Mudd"
S2 E8

While the PAI machines aren't quite as easy to neutralize as the destructive machines vanquished in *Star Trek: TOS*, these relevant four episodes show the protective power of … logic.

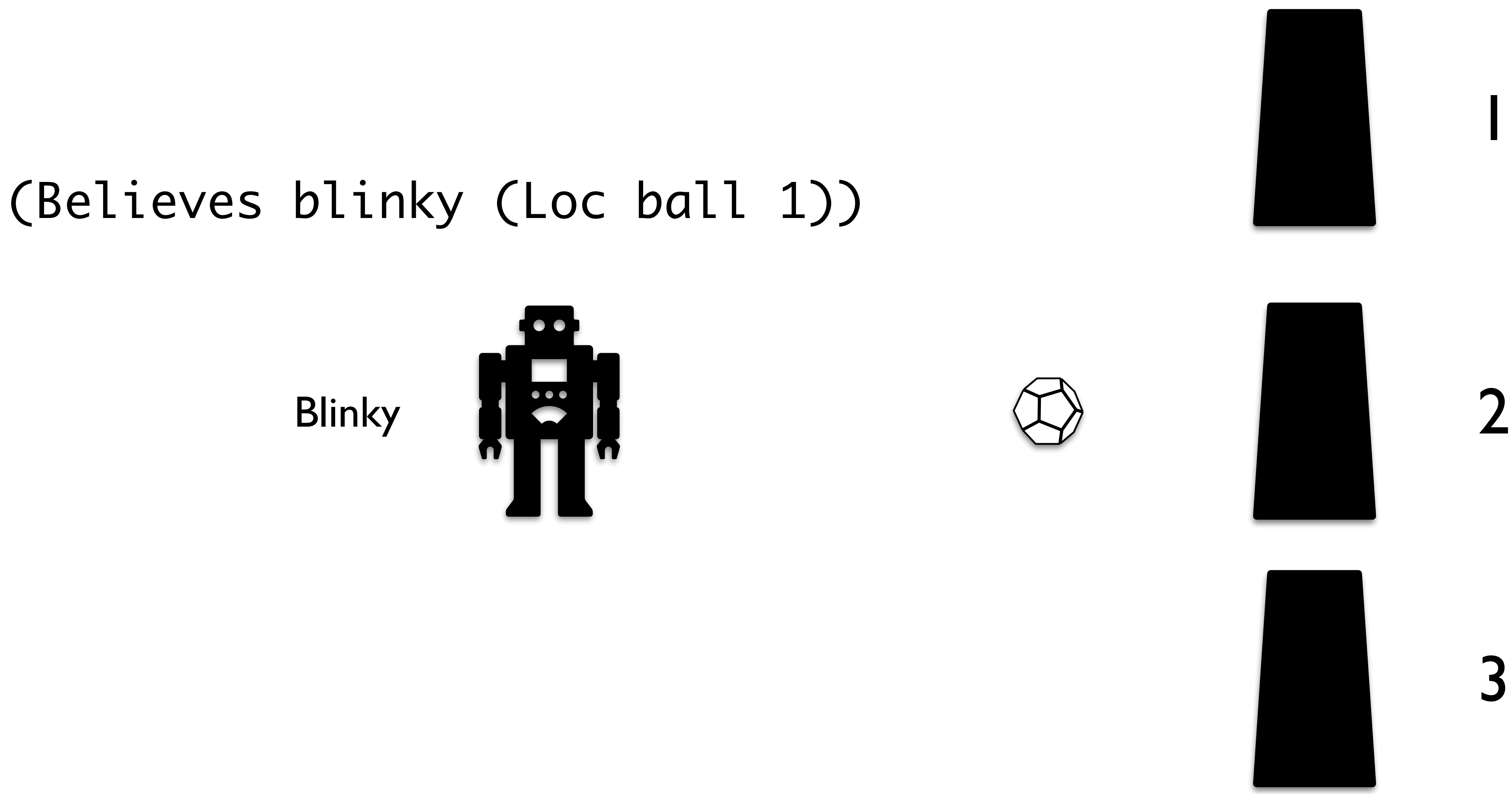
"The Ultimate Computer"
S2 E24

"The Return of the Archons"
S1 E21

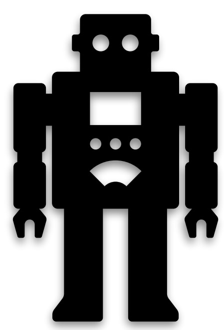
"The Changeling"
S2 E3

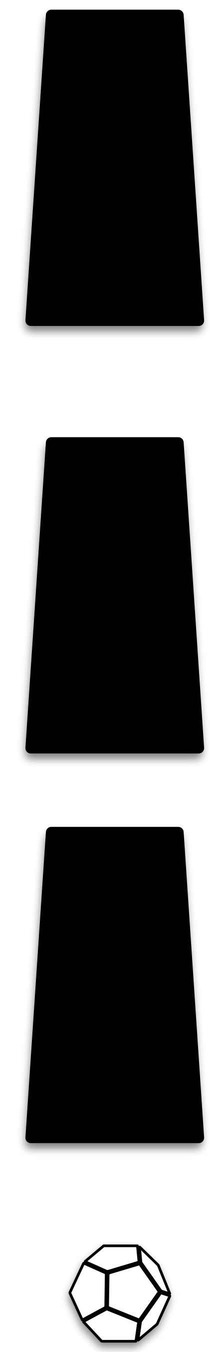
"I, Mudd"
S2 E8

Blinky as portal to intensional logics …

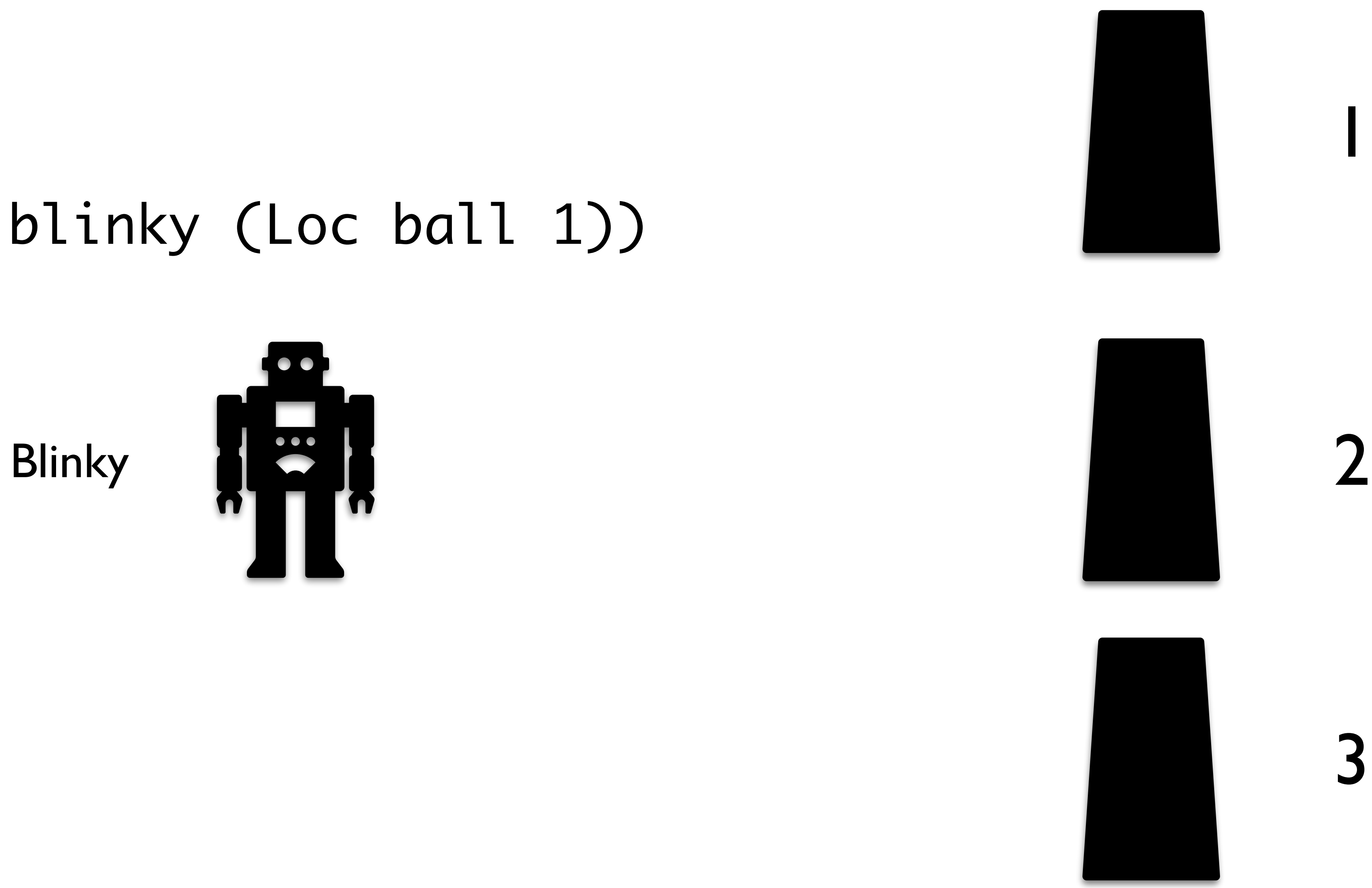

(Believes blinky (Loc ball 1))
Blinky
1
2
3

(Believes blinky (Loc ball 1))

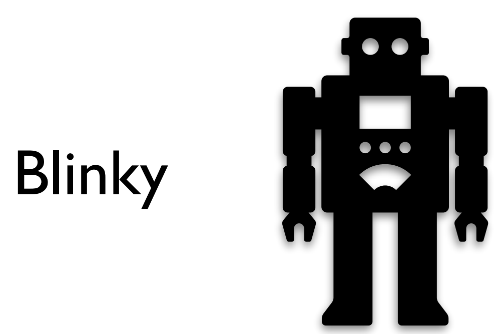


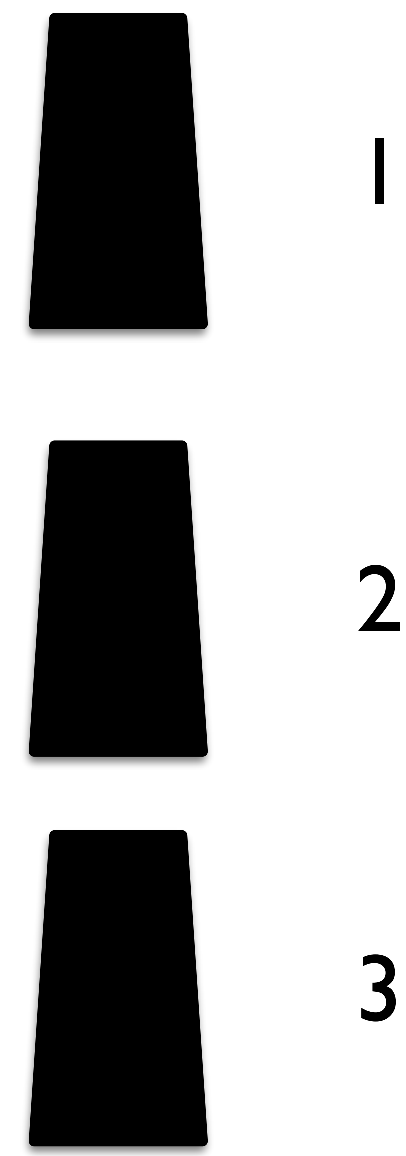


```
(Believes blinky (Loc ball 1))
```



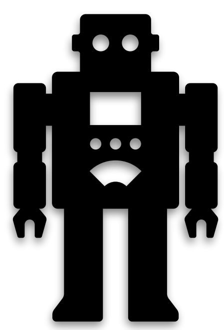
```
(Believes blinky (Loc ball 1))
```

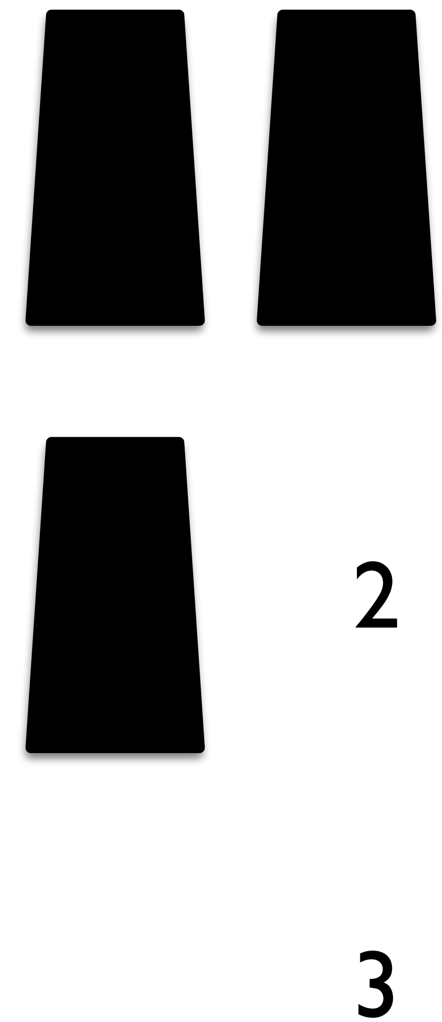




```
(Believes blinky (Loc ball 1))
```

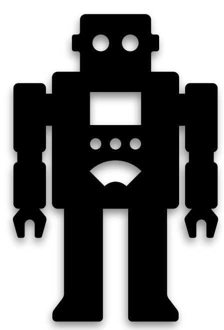
Blinky

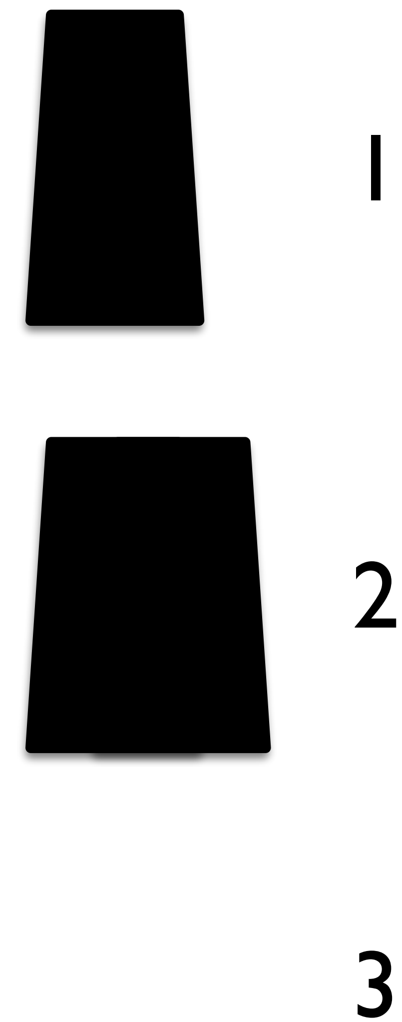




`(Believes blinky (Loc ball 1))`

Blinky

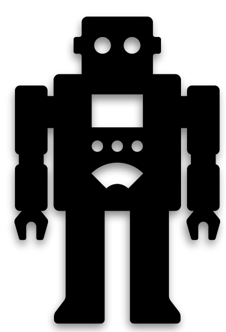


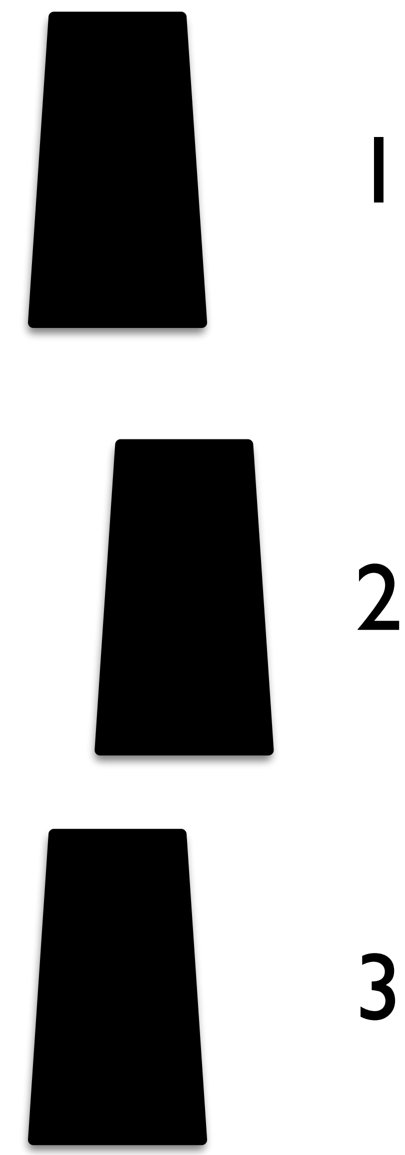


`(Believes blinky (Loc ball 1))`

```
(Believes blinky (Loc ball 1))
```

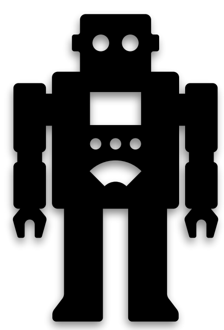
Blinky

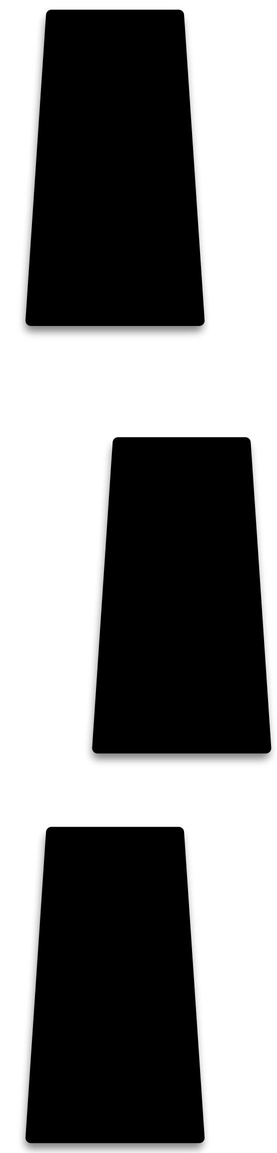




```
(Believes blinky (Loc ball 1))
```

Blinky





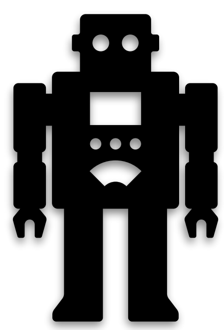
1

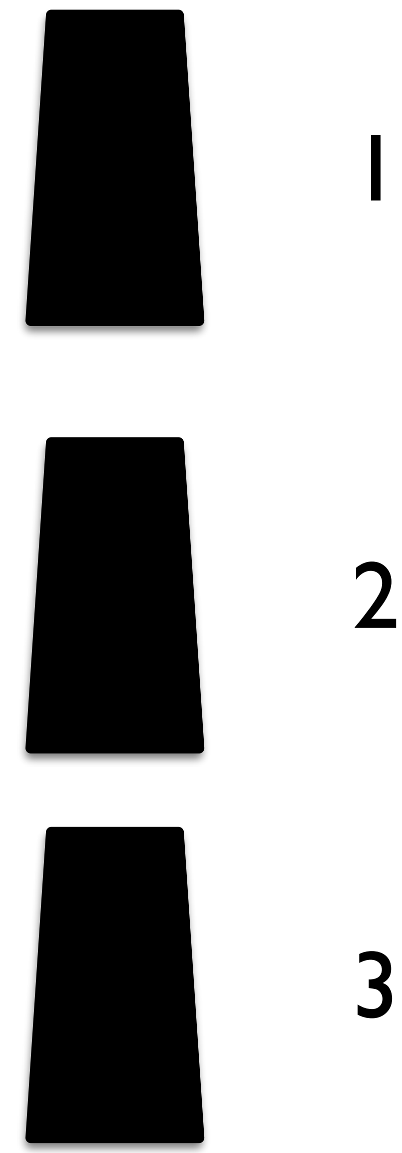
2

3

```
(Believes blinky (Loc ball 1))
```

Blinky

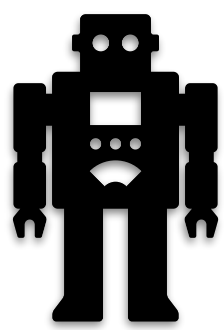


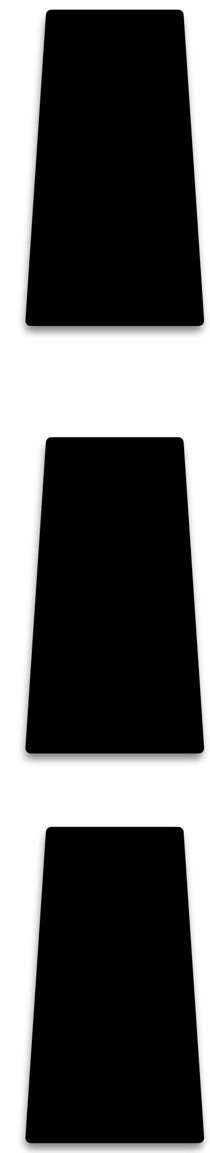


The ball is in the cup at location #1.

```
(Believes blinky (Loc ball 1))
```

Blinky





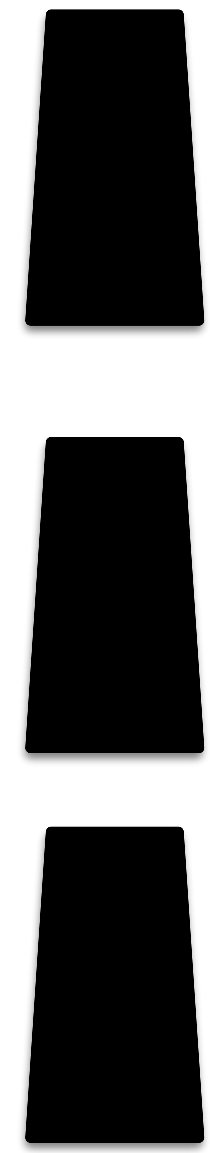
1

2

3

The ball is in the cup at location #1.

$\mathrm{Loc}(\mathrm{ball}, 1)$



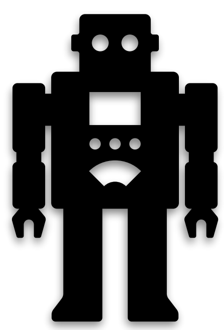
1

2

3

```
(Believes blinky (Loc ball 1))
```

Blinky



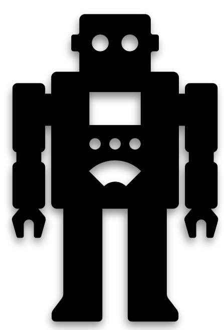
The ball is in the cup at location #1.

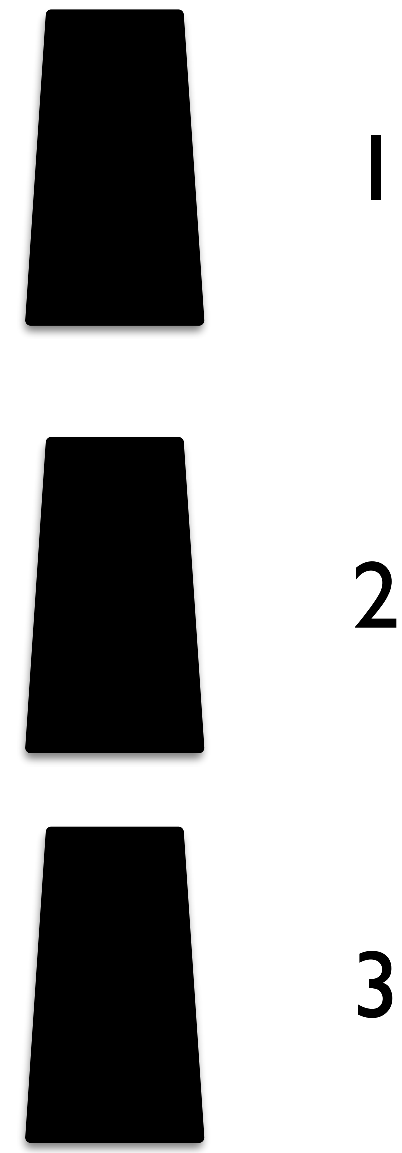
$\text{Loc}(\text{ball}, 1)$

```
(Loc ball 1)
```

```
(Believes blinky (Loc ball 1))
```

Blinky





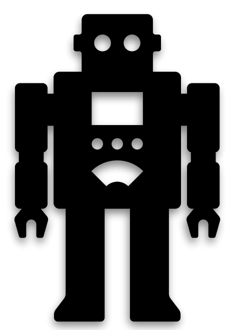
The ball is in the cup at location #1.

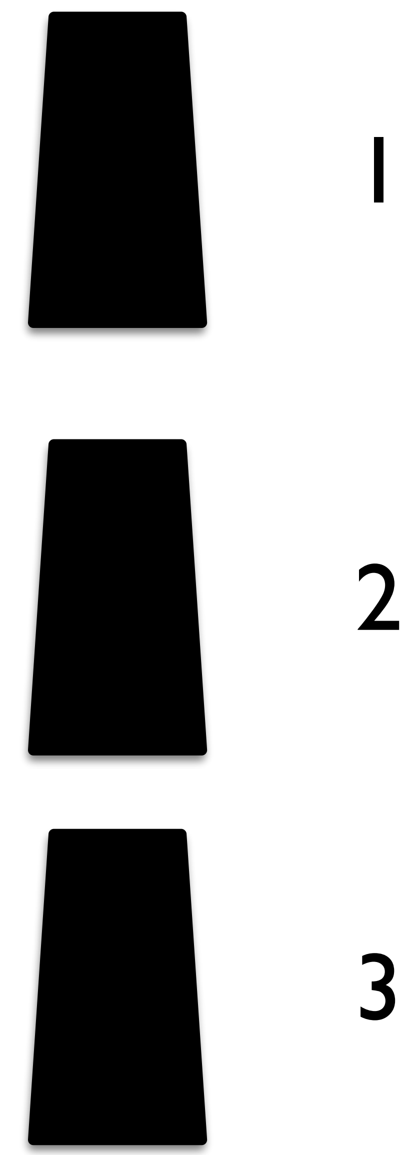
FALSE Loc(ball, 1)

```
(Loc ball 1)
```

```
(Believes blinky (Loc ball 1))
```

Blinky



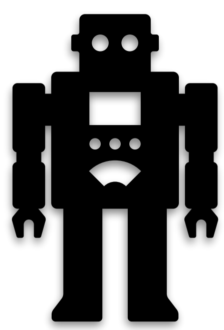


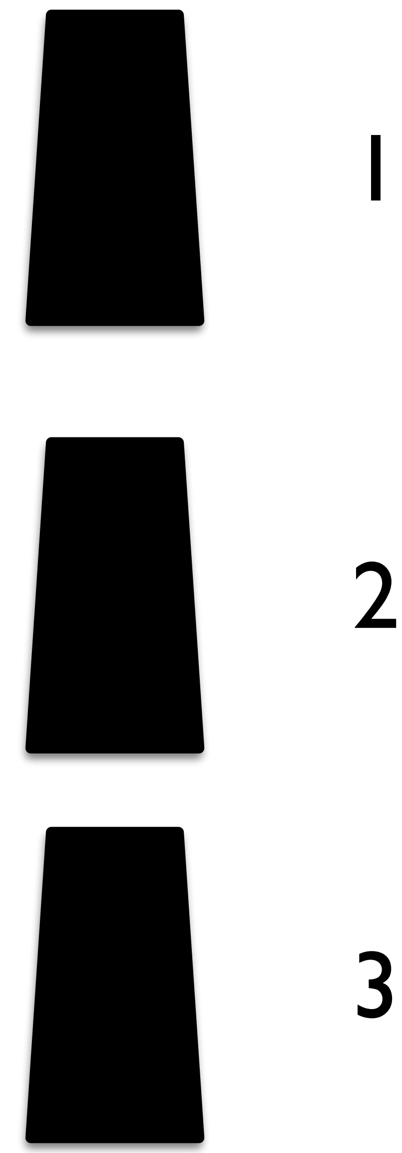
FALSE $\mathsf{Loc}(\mathsf{ball}, 1)$

```
(Loc ball 1)
```

```
(Believes blinky (Loc ball 1))
```

Blinky



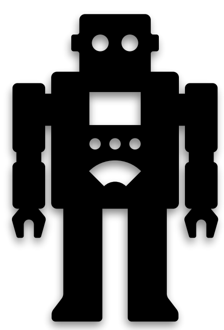


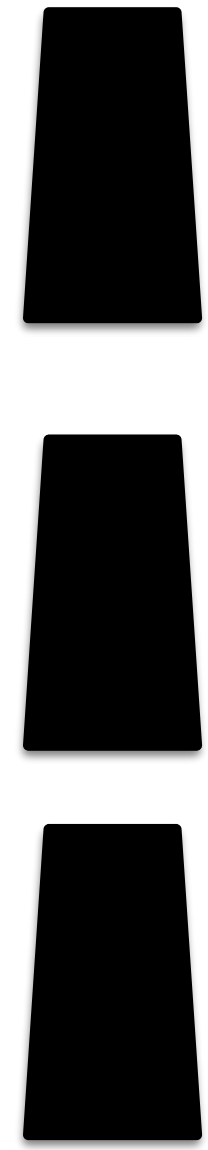
FALSE

(Loc ball 1)

(Believes blinky (Loc ball 1))

Blinky





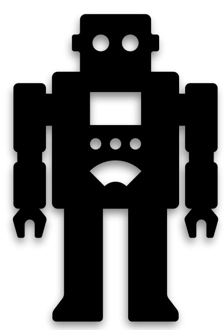
1

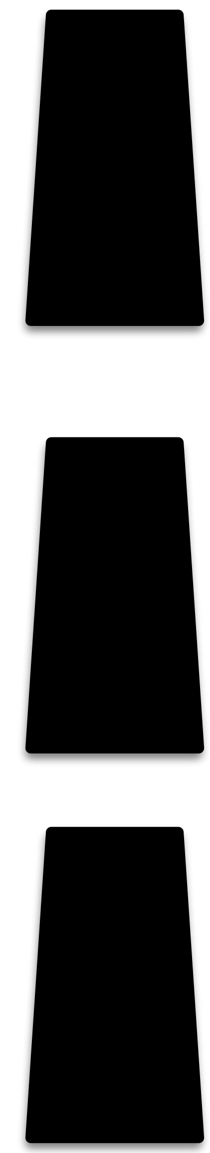
2

3

```
(Loc ball 1)

(Believes blinky (Loc ball 1))
```

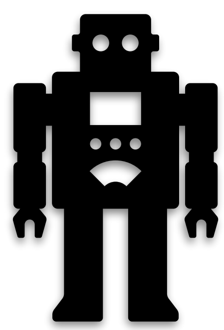
Blinky

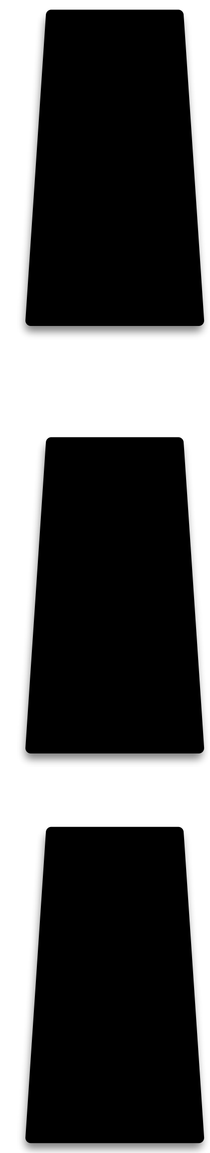




```
(Believes blinky (Loc ball 1))
```

Blinky

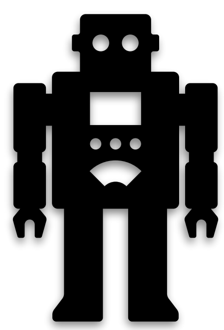


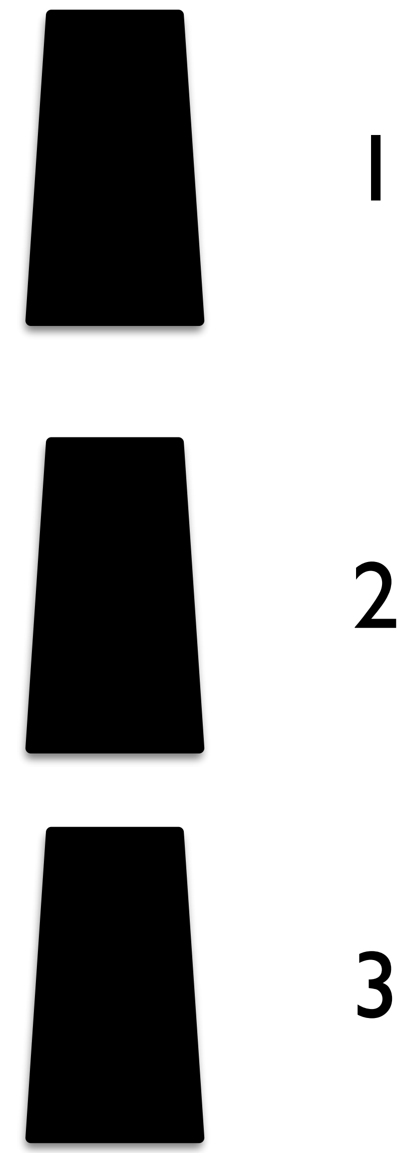


The ball is in the cup at location #1 or the ball is at location #3.

```
(Believes blinky (Loc ball 1))
```

Blinky



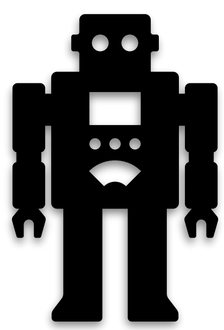


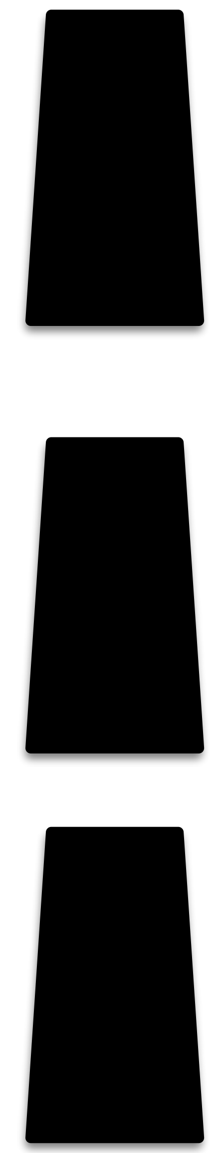
The ball is in the cup at location #1 or the ball is at location #3.

$$\mathrm{Loc}(\mathrm{ball},1) \vee \mathrm{Loc}(\mathrm{ball},3)$$

```
(Believes blinky (Loc ball 1))
```

Blinky





1

2

3

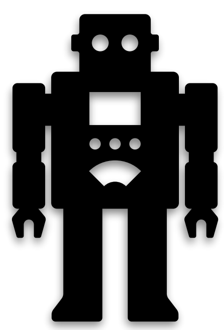
The ball is in the cup at location #1 or the ball is at location #3.

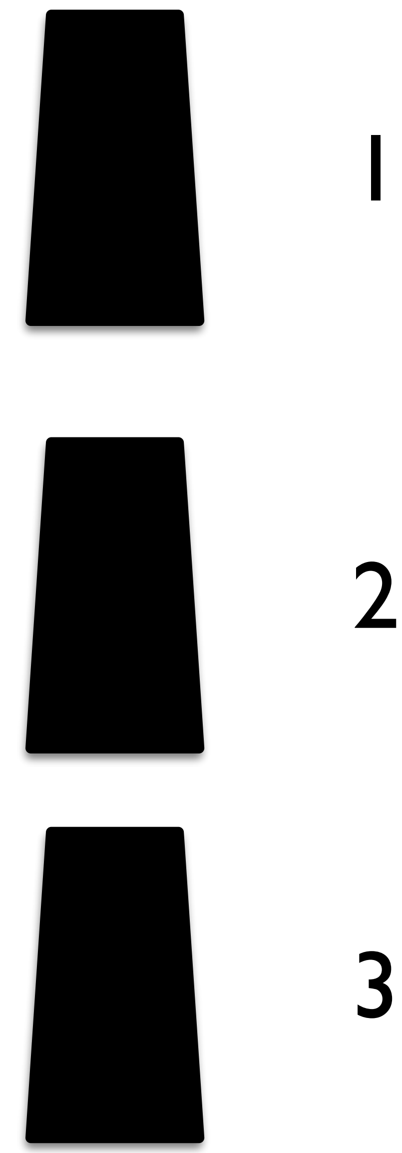
$$\text{Loc}(\text{ball},1) \lor \text{Loc}(\text{ball},3)$$

```
(or (Loc ball 1) (Loc ball 3))
```

```
(Believes blinky (Loc ball 1))
```

Blinky





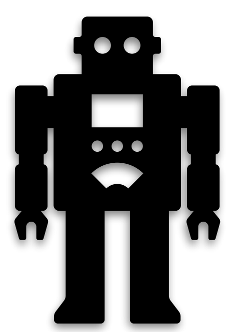
The ball is in the cup at location #1 or the ball is at location #3.

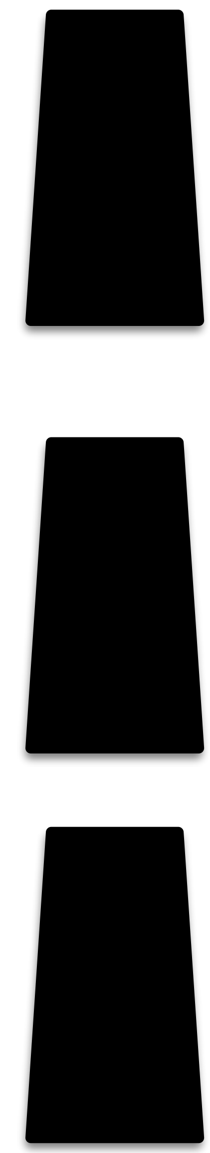
FALSE $Loc(ball,1) \vee Loc(ball,3)$

```
(or (Loc ball 1) (Loc ball 3))
```

```
(Believes blinky (Loc ball 1))
```

Blinky





1

2

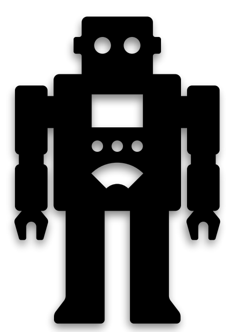
3

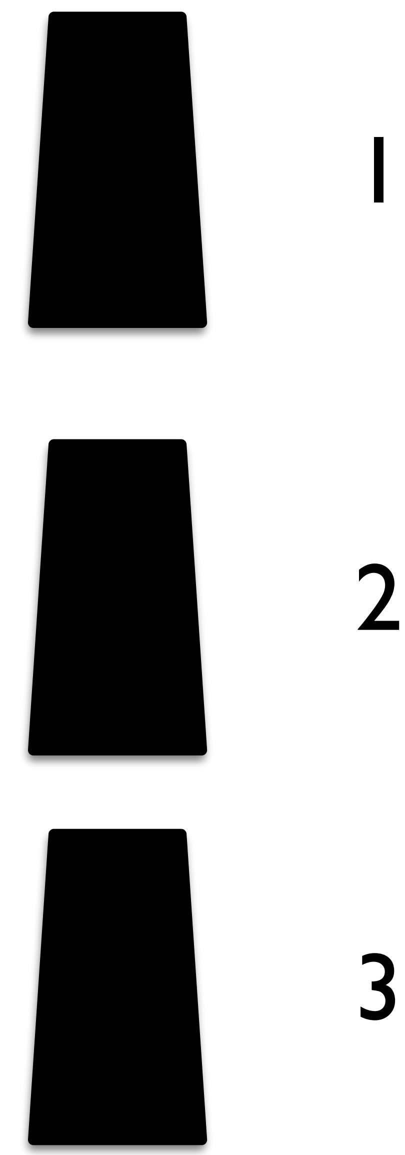
FALSE $\mathrm{Loc}(\mathrm{ball},1) \vee \mathrm{Loc}(\mathrm{ball},3)$

```
(or (Loc ball 1) (Loc ball 3))
```

```
(Believes blinky (Loc ball 1))
```

Blinky

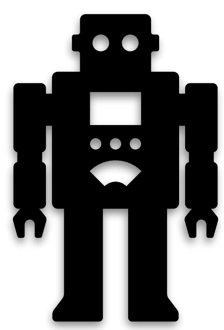


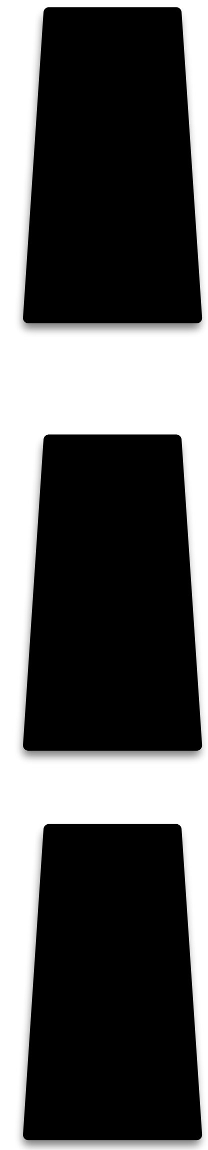


FALSE

```
(or (Loc ball 1) (Loc ball 3))

(Believes blinky (Loc ball 1))
```

Blinky





1

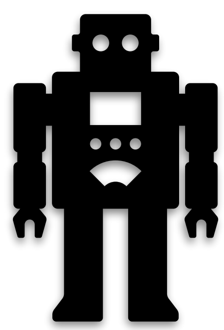
2

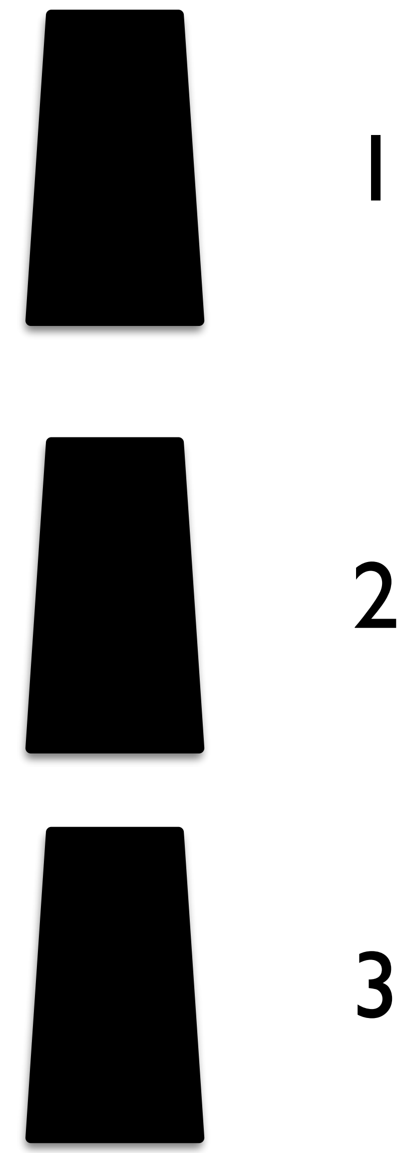
3

FALSE

`(Believes blinky (Loc ball 1))`

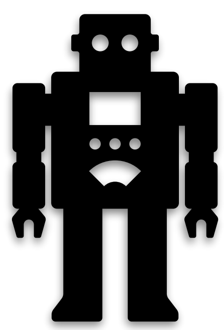
Blinky

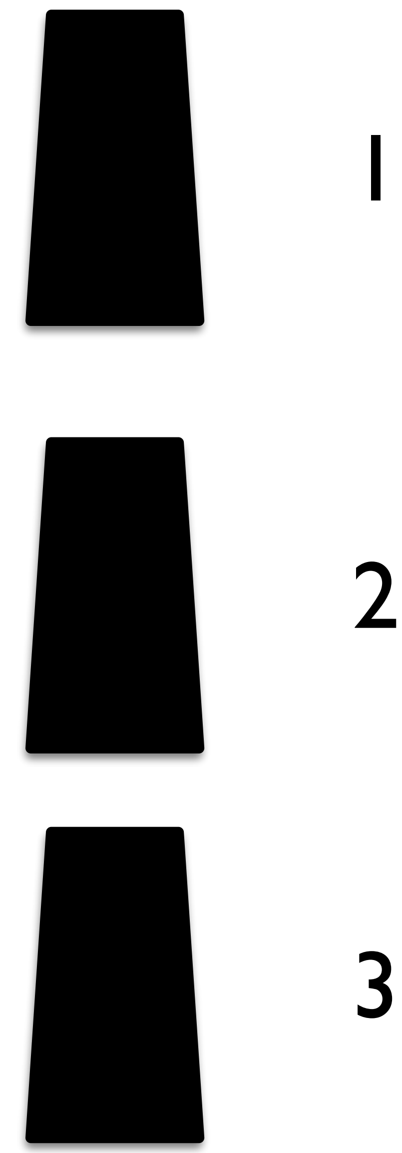




`(Believes blinky (Loc ball 1))`

Blinky

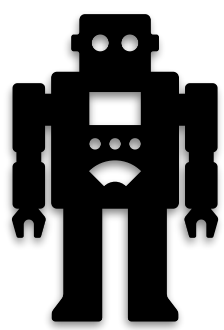


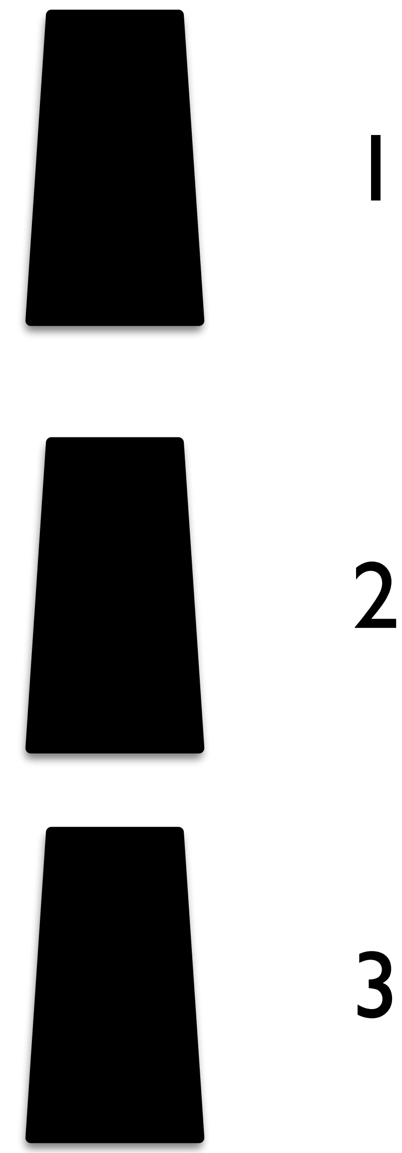


Blinky believes that the ball is in the cup at location #1.

```
(Believes blinky (Loc ball 1))
```

Blinky



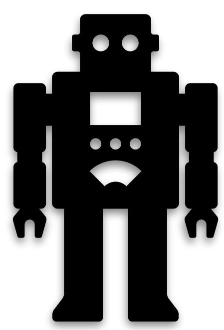


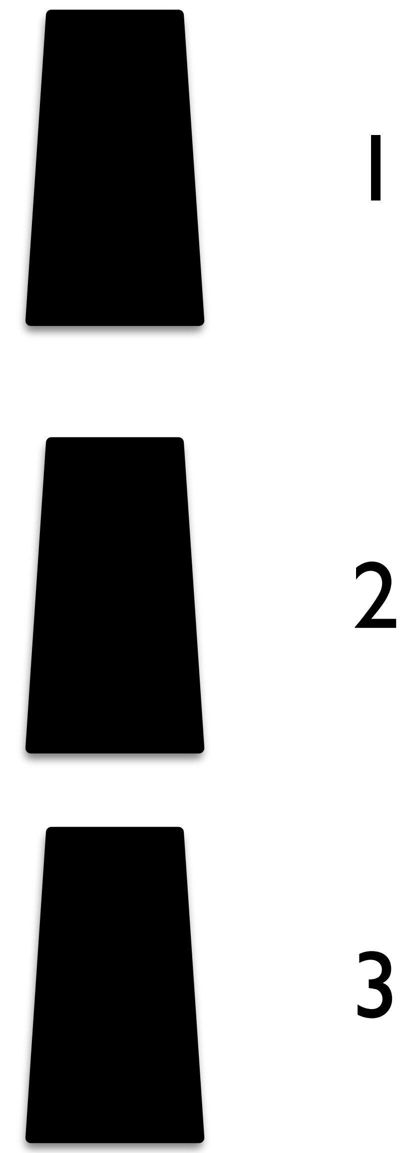
Blinky believes that the ball is in the cup at location #1.

$$\mathbf{B}(\text{blinky}, \text{Loc}(\text{ball}, 1))$$

```
(Believes blinky (Loc ball 1))
```

Blinky



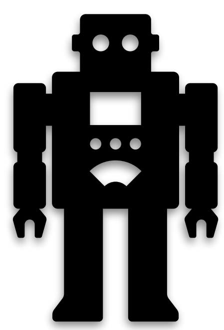


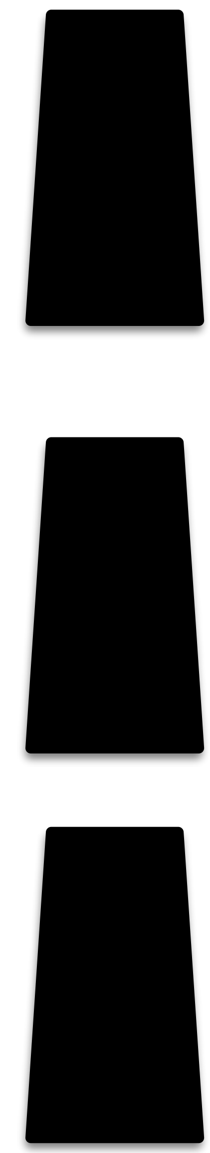
Blinky believes that the ball is in the cup at location #1.

$$\mathbf{B}(\text{blinky}, \text{Loc}(\text{ball}, 1))$$

```
(Believes! blinky (Loc ball 1))

(Believes blinky (Loc ball 1))
```

Blinky





1

2

3

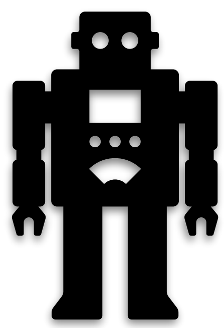
Blinky believes that the ball is in the cup at location #1.

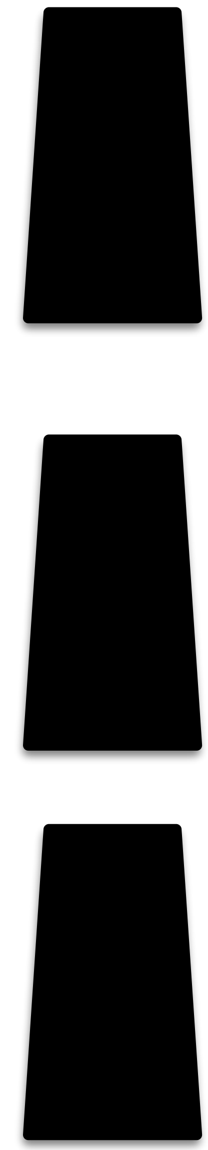
???

$$\mathbf{B}(\text{blinky}, \text{Loc}(\text{ball}, 1))$$

```
(Believes! blinky (Loc ball 1))

(Believes blinky (Loc ball 1))
```

Blinky





1

2

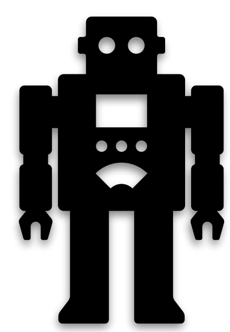
3

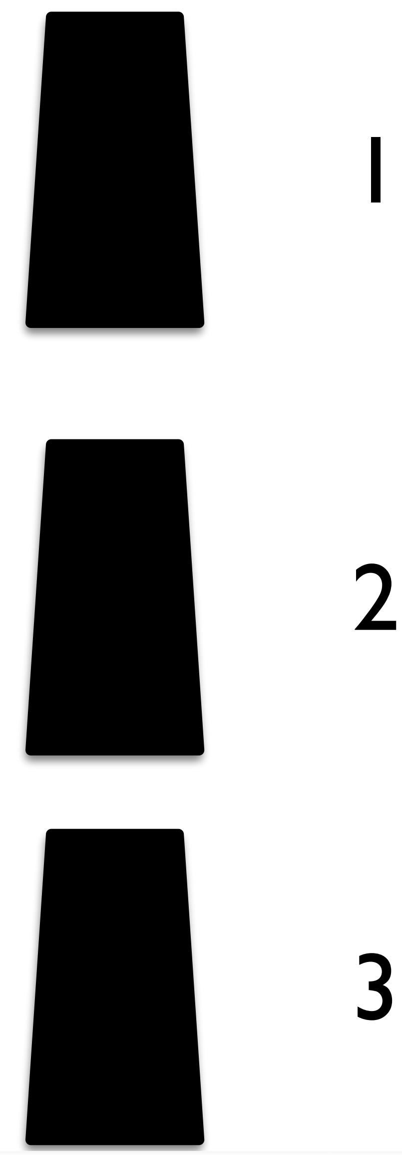
Blinky believes that the ball is in the cup at location #1.

??? $\mathbf{B}(blinky, Loc(ball, 1))$

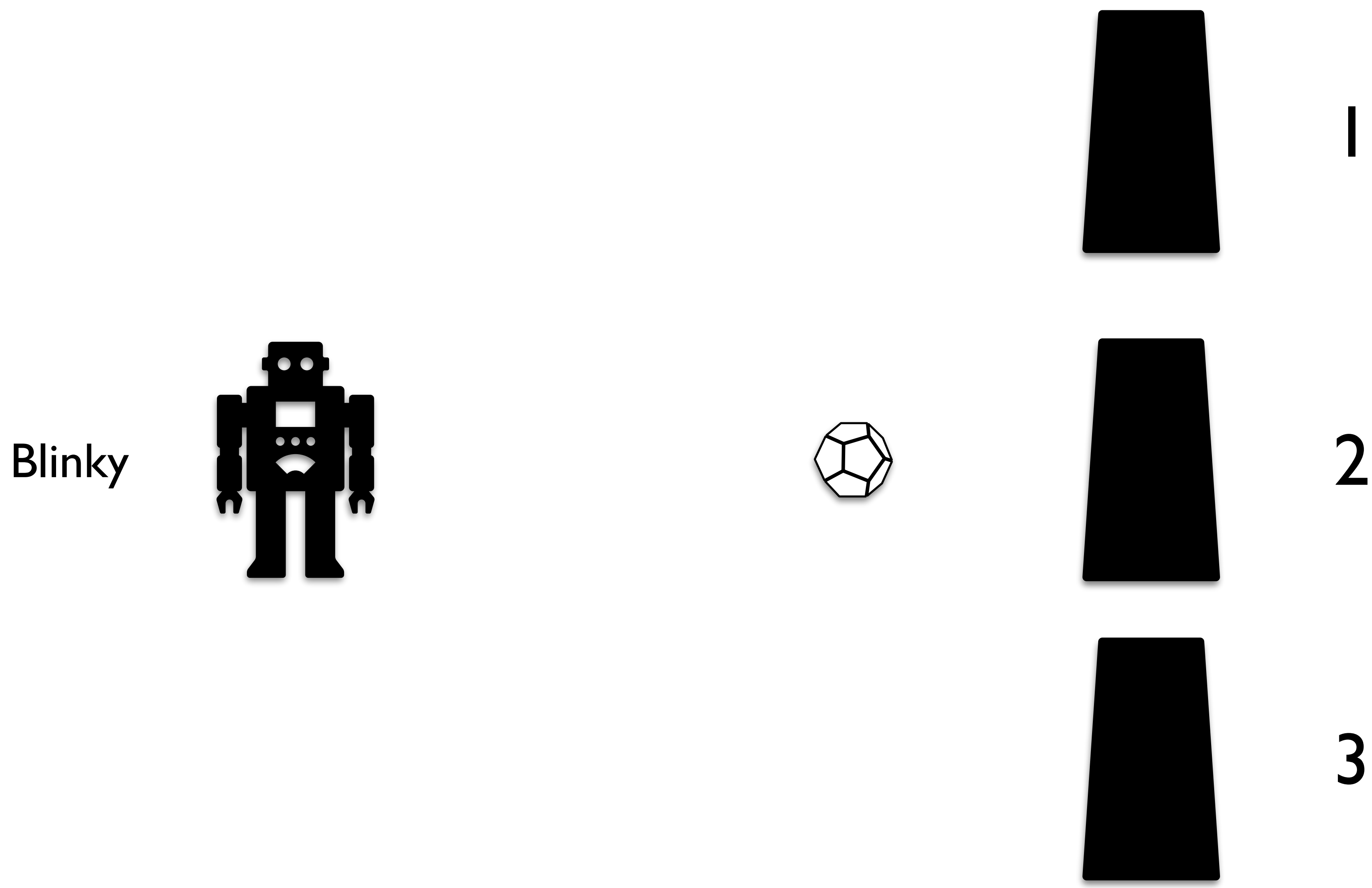
```
(Believes! blinky (Loc ball 1))

(Believes blinky (Loc ball 1))
```

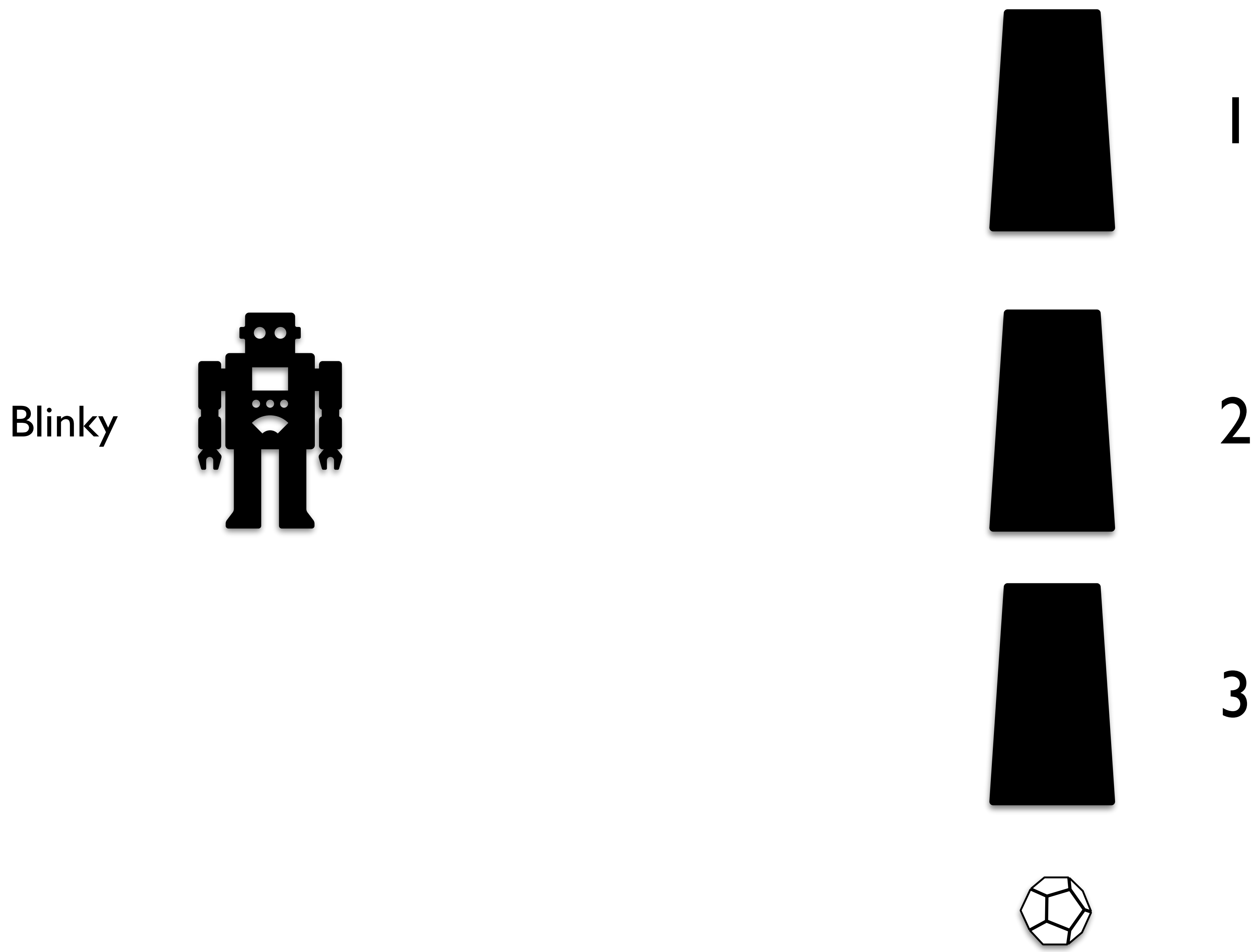
Blinky

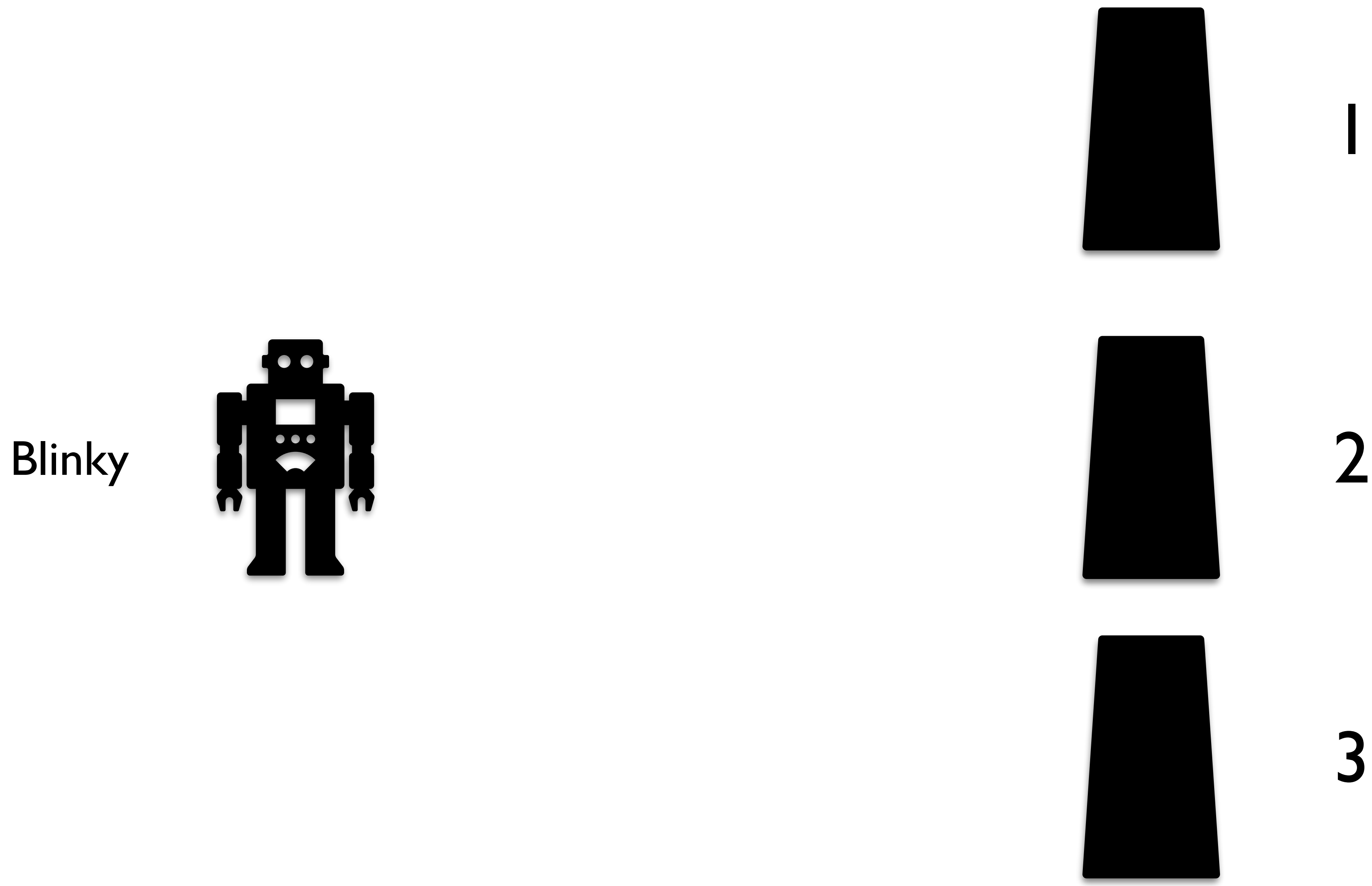


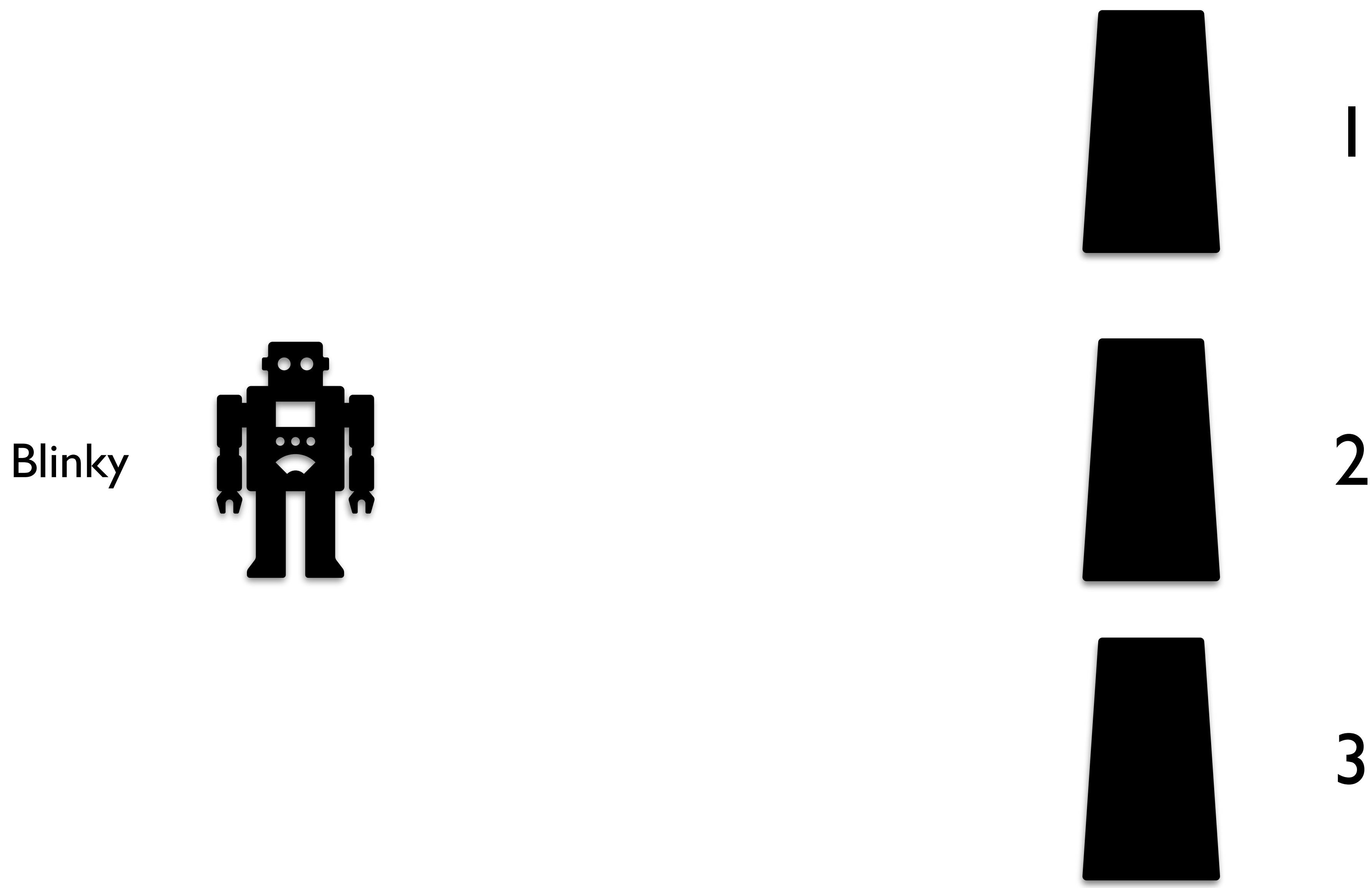


In extensional logics, what is denoted is conflated with meaning (the latter being naïvely compositional), but intensional attitudes like *believes*, *knows*, *hopes*, *fears*, etc cannot be represented and reasoned over smoothly (e.g. without fear of inconsistency rising up).

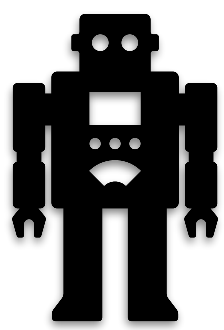

Blinky
1
2
3

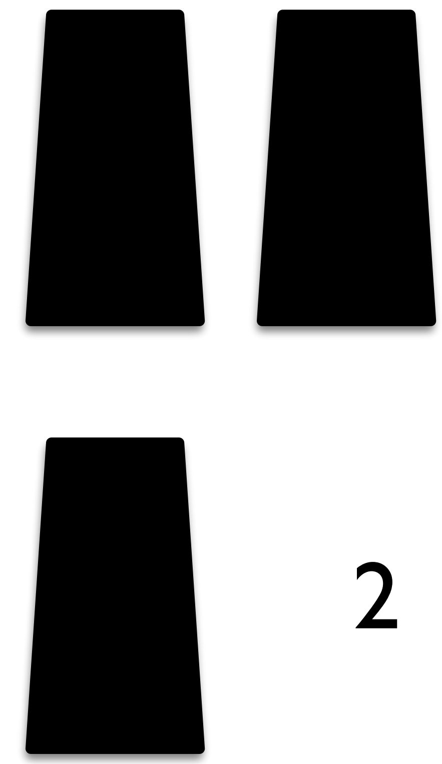

Blinky
1
2
3


Blinky
1
2
3


Blinky
1
2
3

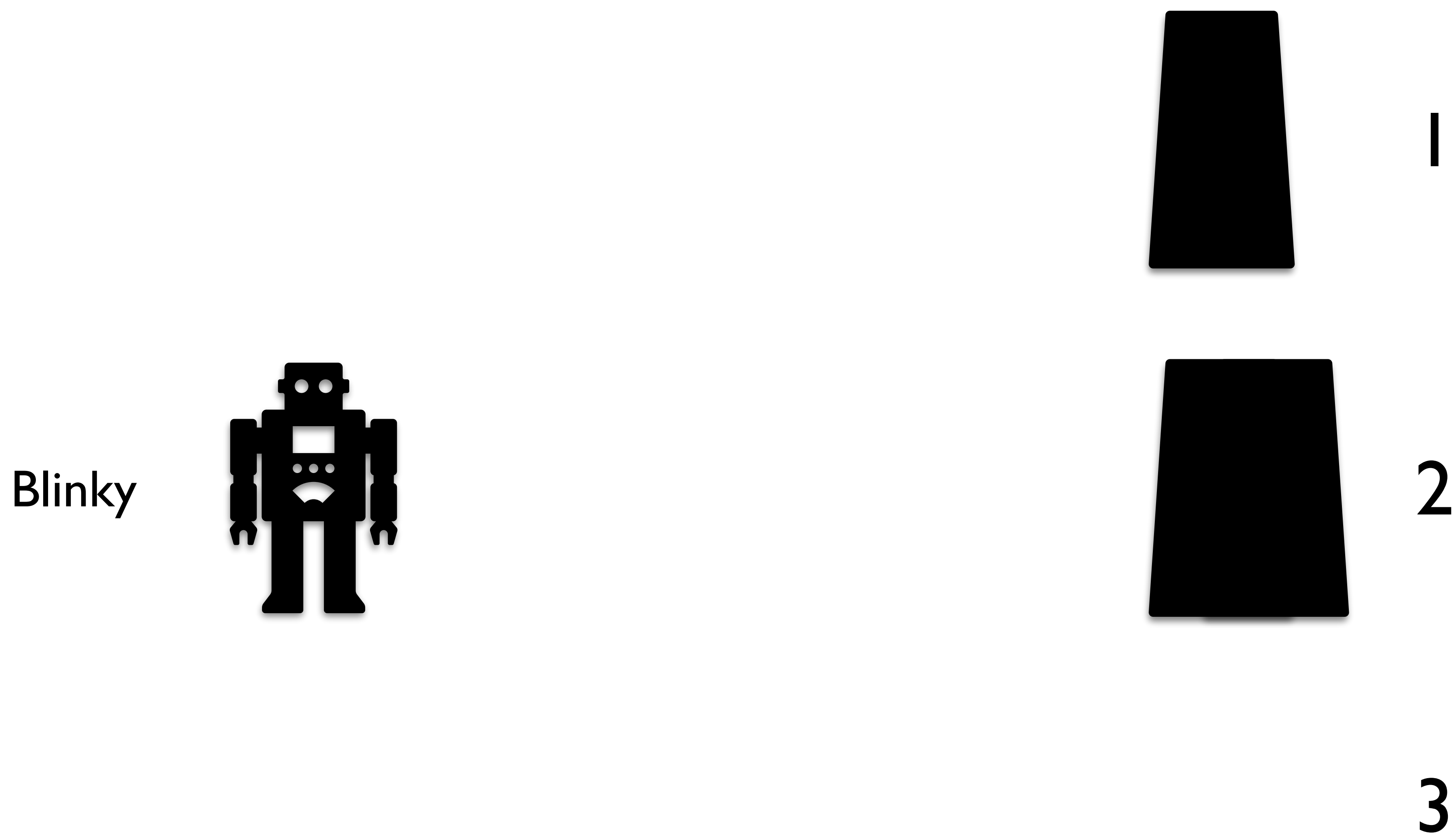
Blinky

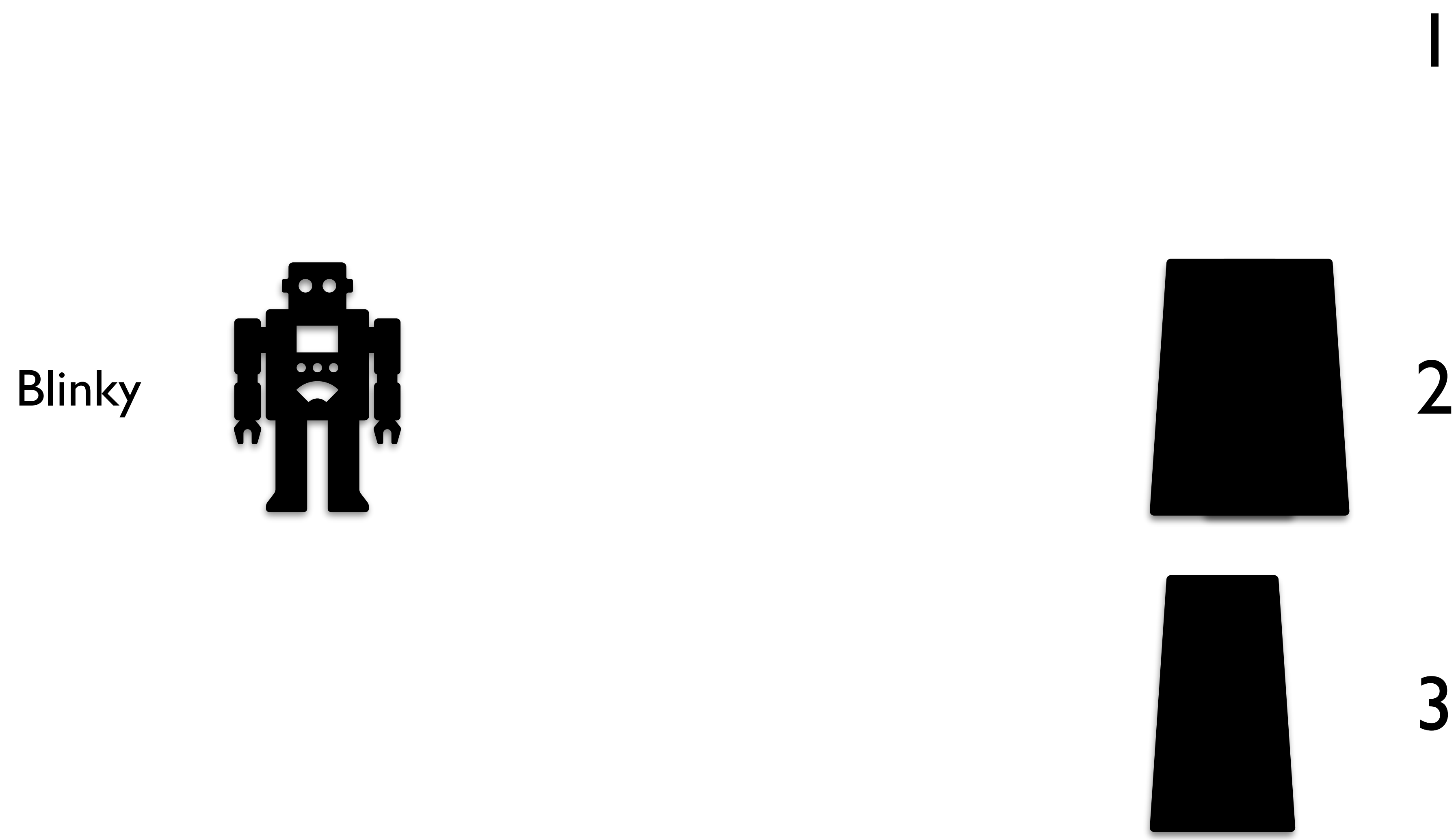


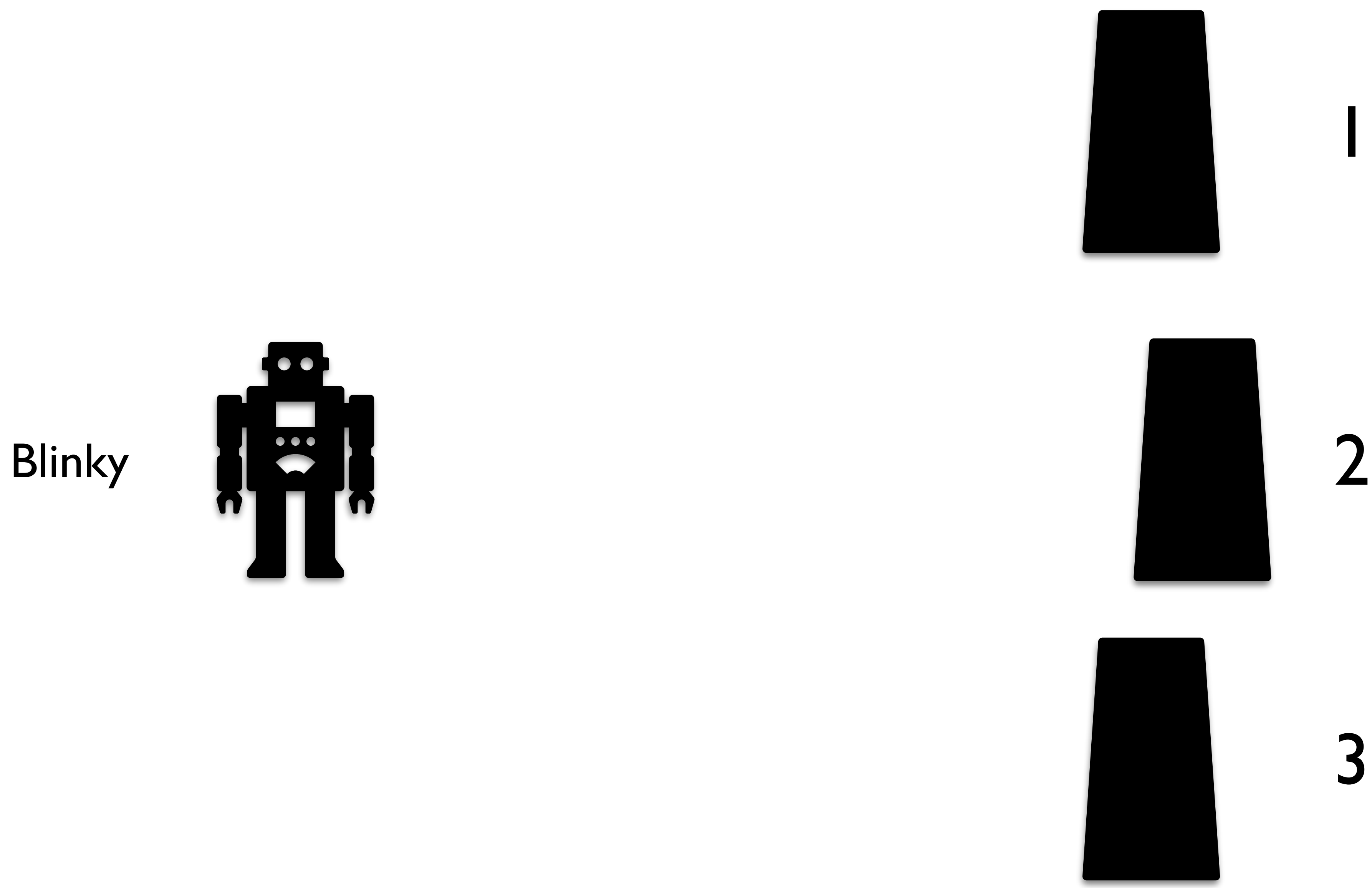


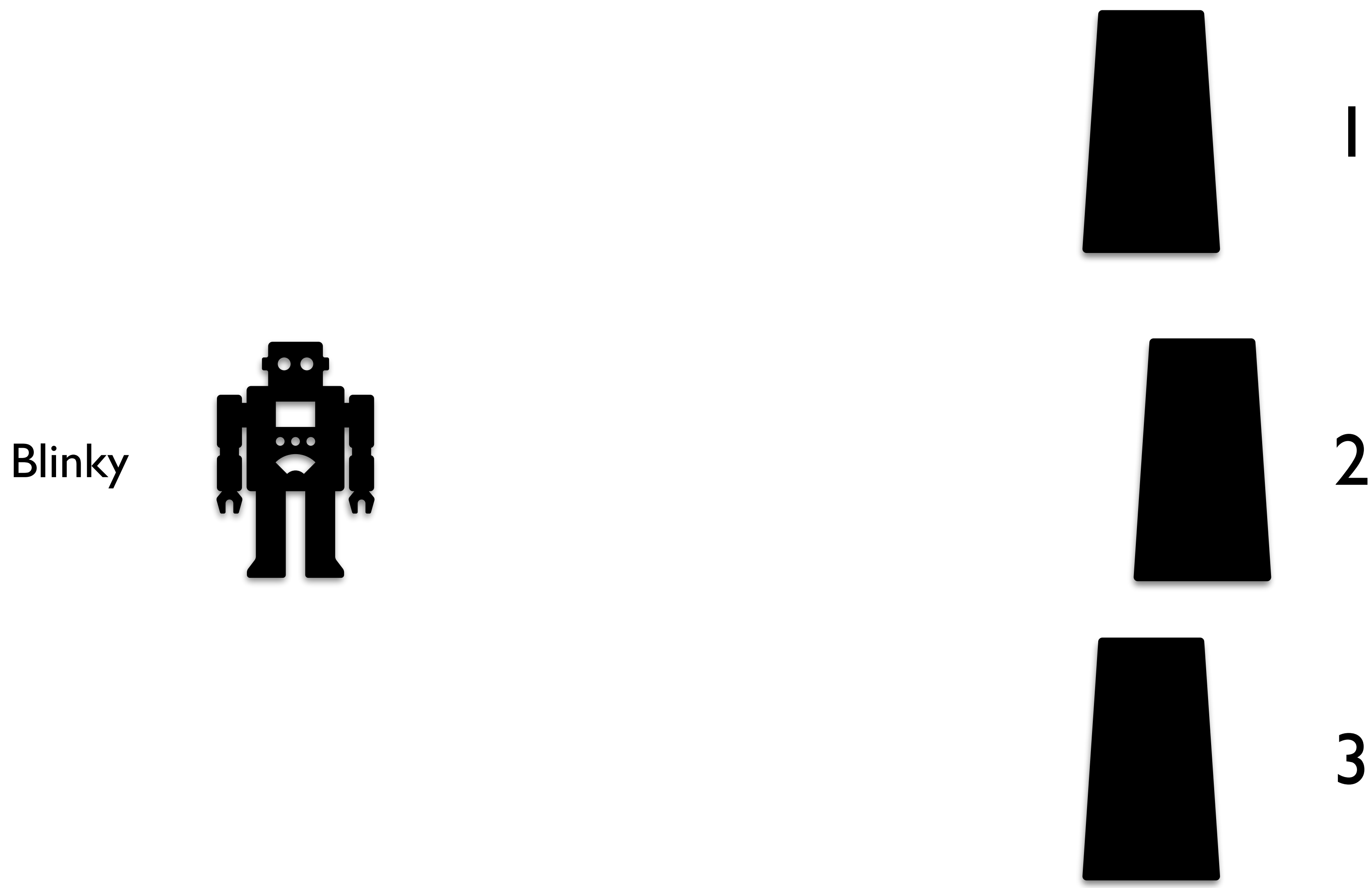
2

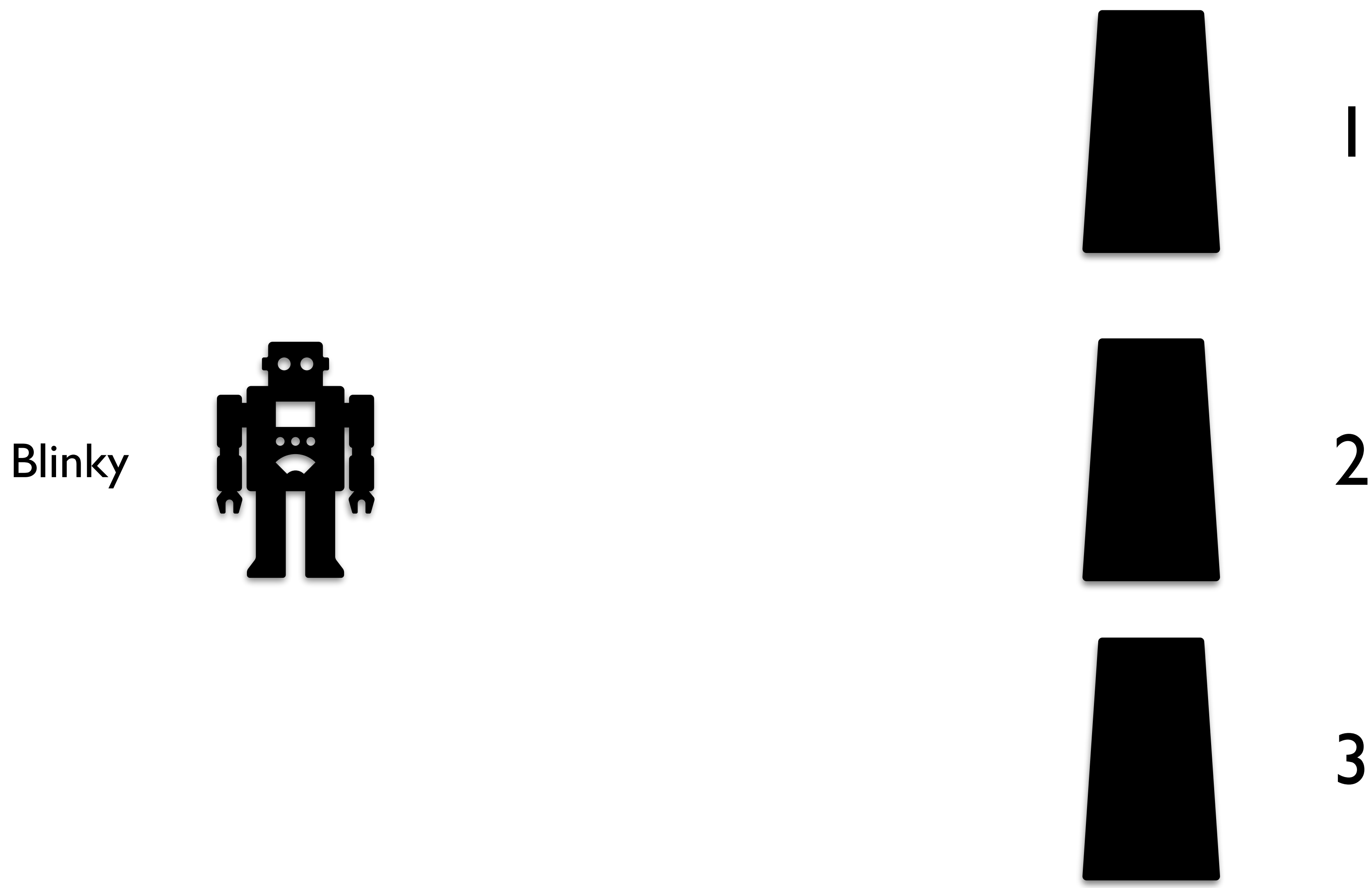
3


Blinky
1
2
3

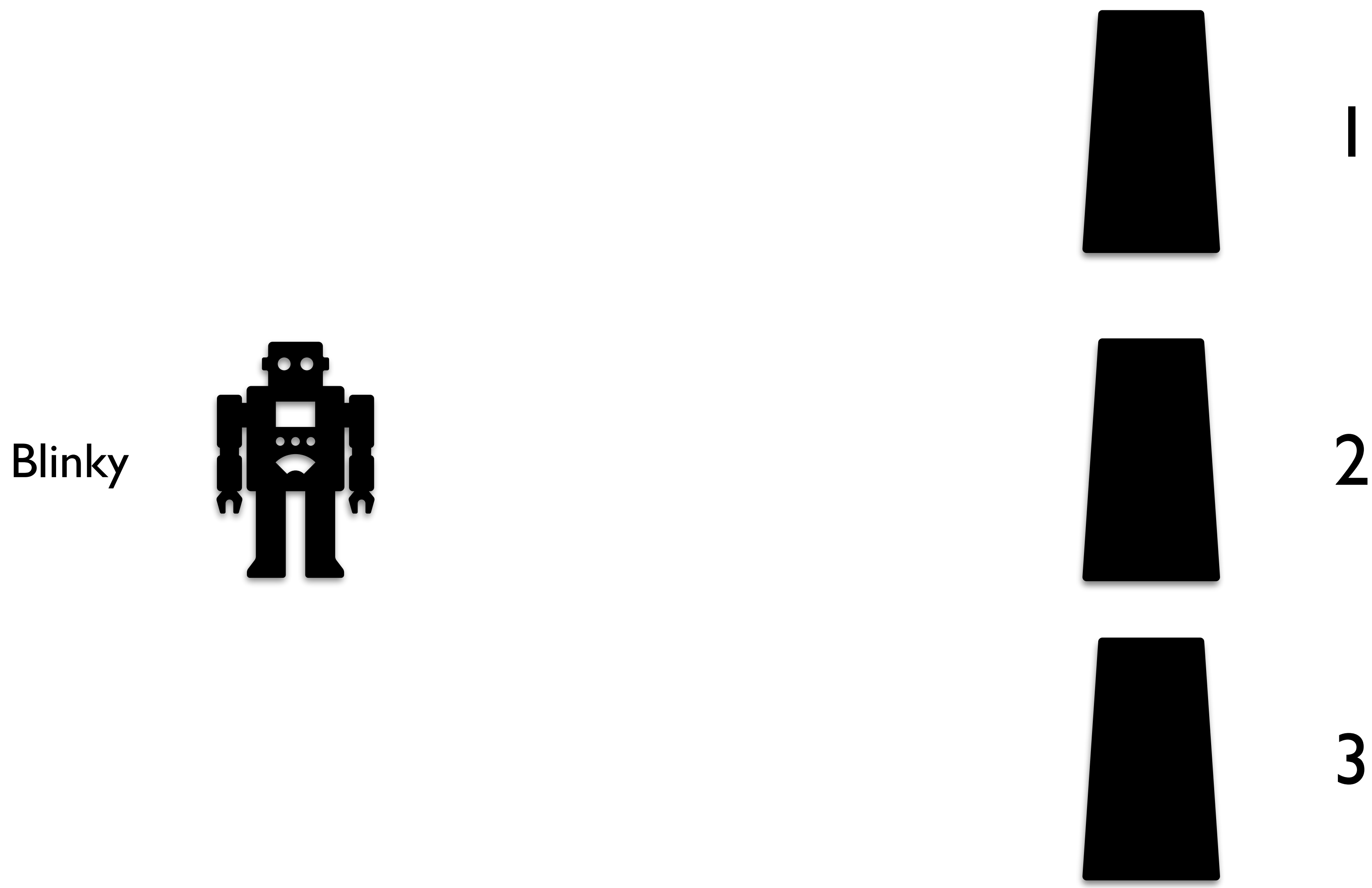

Blinky
1
2
3


Blinky
1
2
3


Blinky
1
2
3

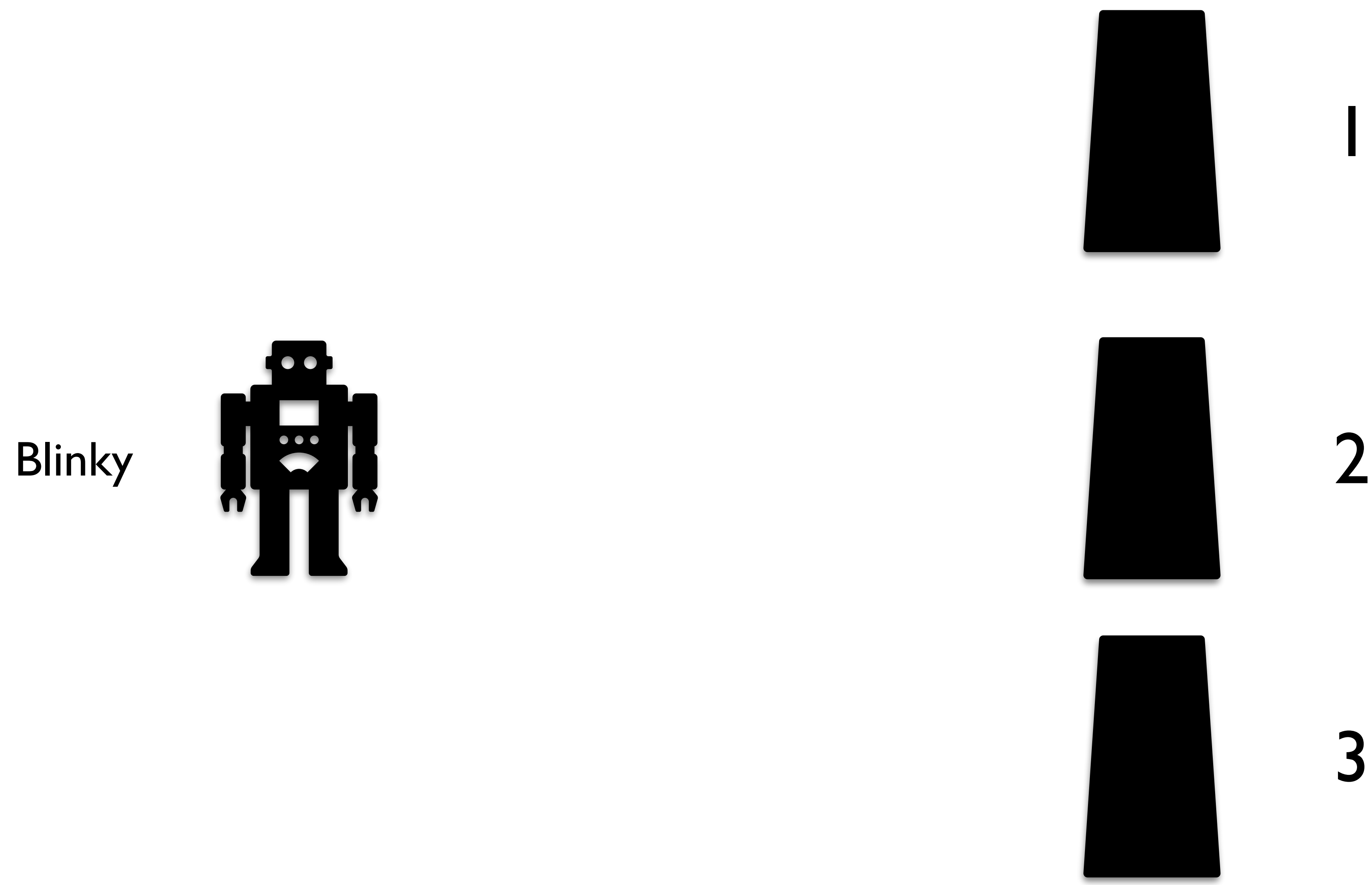

Blinky
1
2
3

Blinky believes that the ball is in the cup at location #1.



Blinky believes that the ball is in the cup at location #1.

B(blinky, loc-ball-1)

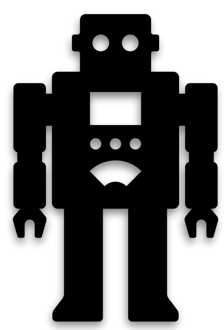


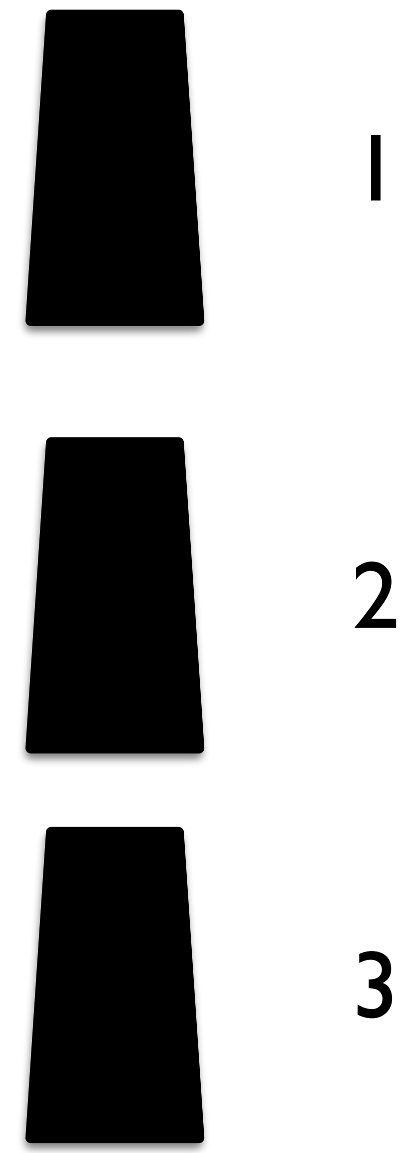
Blinky believes that the ball is in the cup at location #1.

B(blinky, loc-ball-1)

```
(Believes! blinky loc-ball-1)
```

Blinky



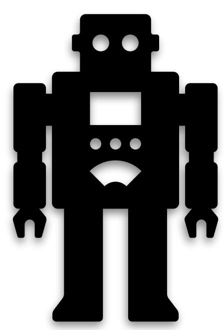


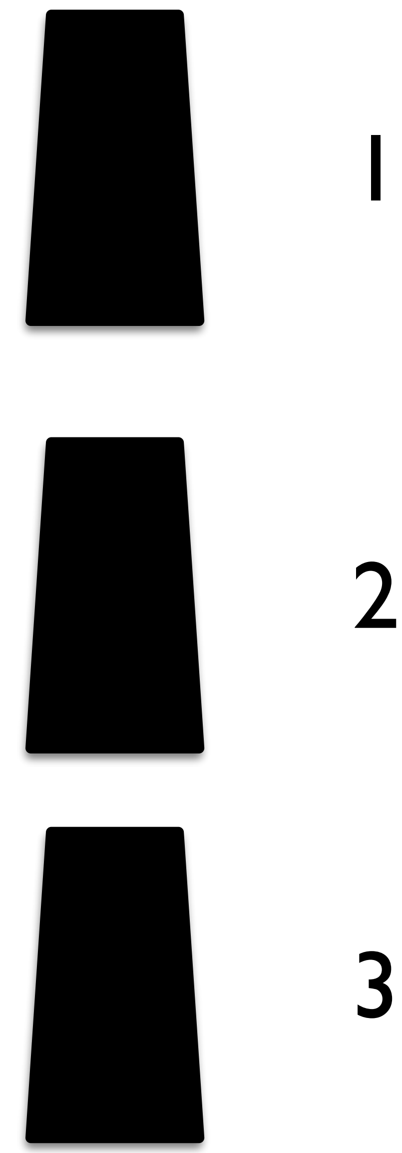
Blinky believes that the ball is in the cup at location #1.

B(blinky, loc-ball-1)

```
(Believes! blinky loc-ball-1)
```

Blinky





In intensional logics, meaning and designation are separated, and compositionality is abandoned.

Blinky believes that the ball is in the cup at location #1.

B(blinky, loc-ball-1)

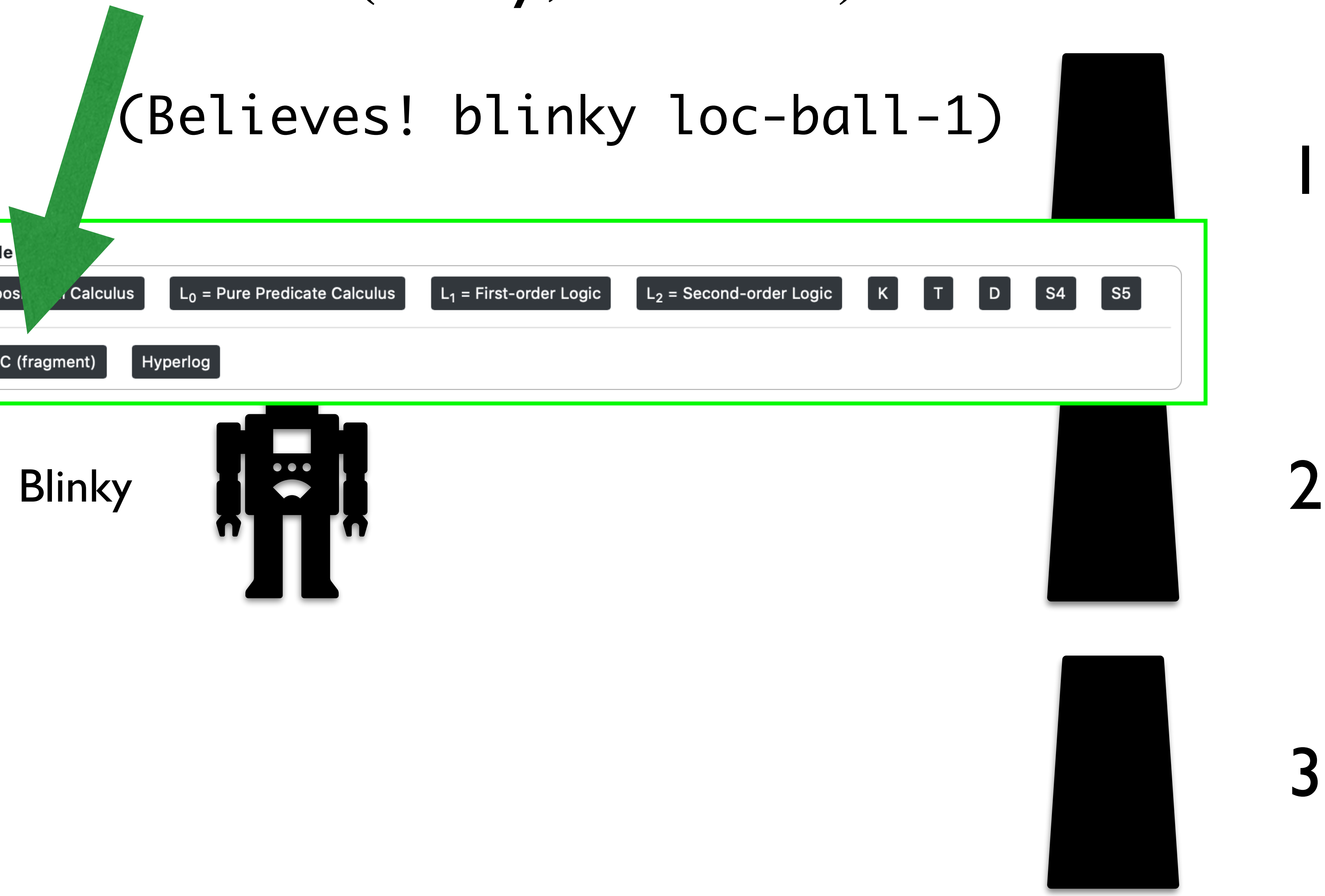
`(Believes! blinky loc-ball-1)`

In intensional logics, meaning and designation are separated, and compositionality is abandoned.

Blinky believes that the ball is in the cup at location #1.

B(blinky, loc-ball-1)

```
(Believes! blinky loc-ball-1)
```



In intensional logics, meaning and designation are separated, and compositionality is abandoned.

For Brave Adventurers

For Brave Adventurers

“Everything smart knows that everything tinks anything that tinks something identical with something.”

For Brave Adventurers

"Everything smart knows that everything tinks anything that tinks something identical with something."

"Blinky is smart."

For Brave Adventurers

"Everything smart knows that everything tinks anything that tinks something identical with something."

"Blinky is smart."

Therefore:

For Brave Adventurers

“Everything smart knows that everything tinks anything that tinks something identical with something.”

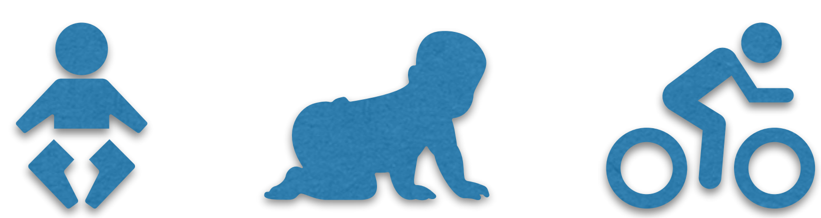
“Blinky is smart.”

Therefore:

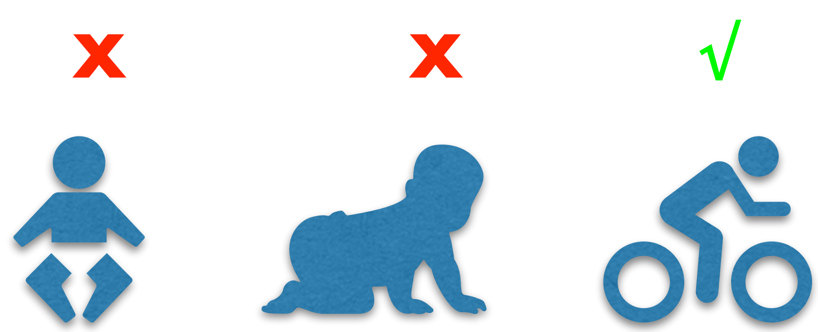
“Everything tinks anything that tinks something identical with something.”

False Belief Task Demands Intensional Logic …

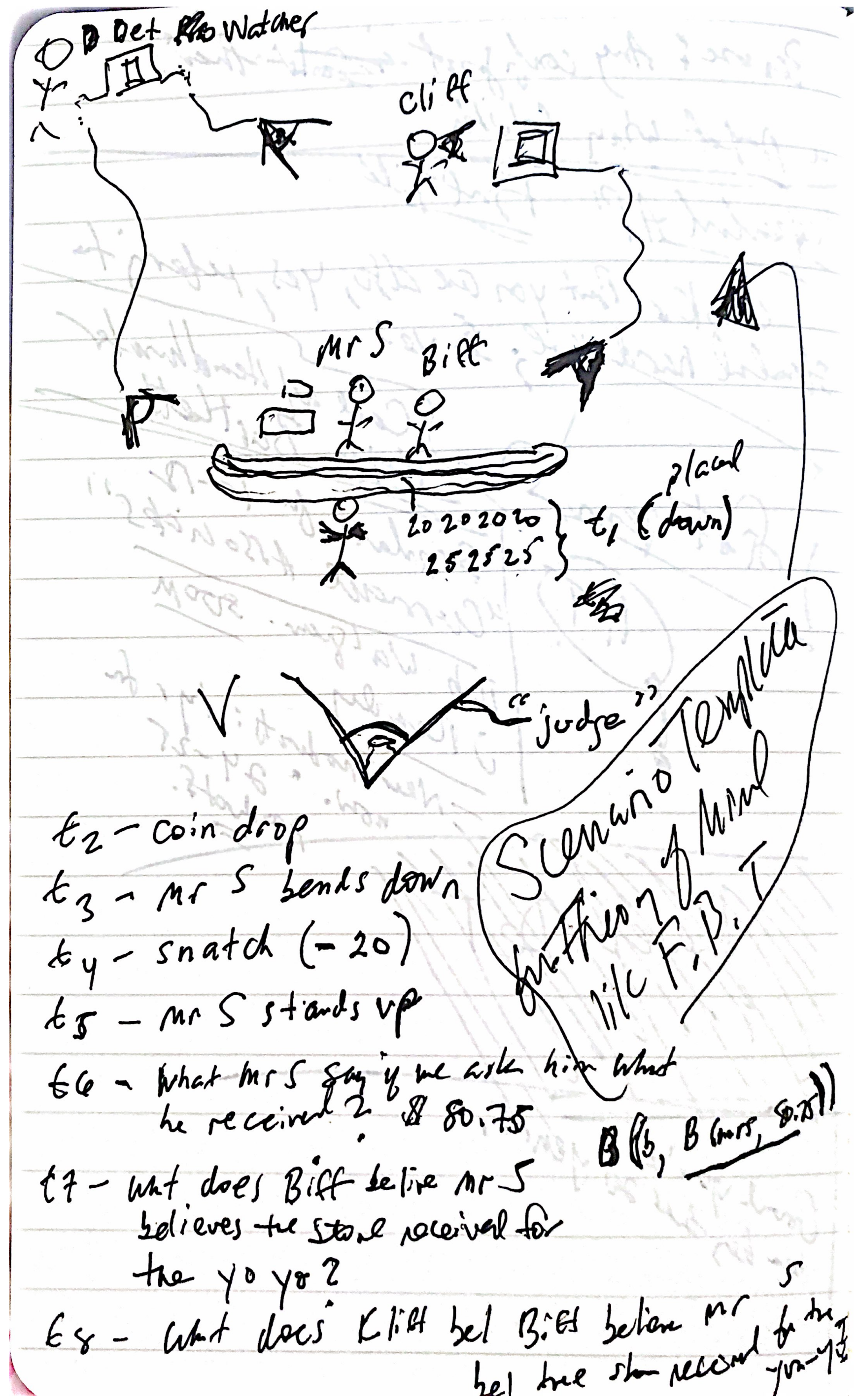
False Belief Task Demands Intensional Logic …



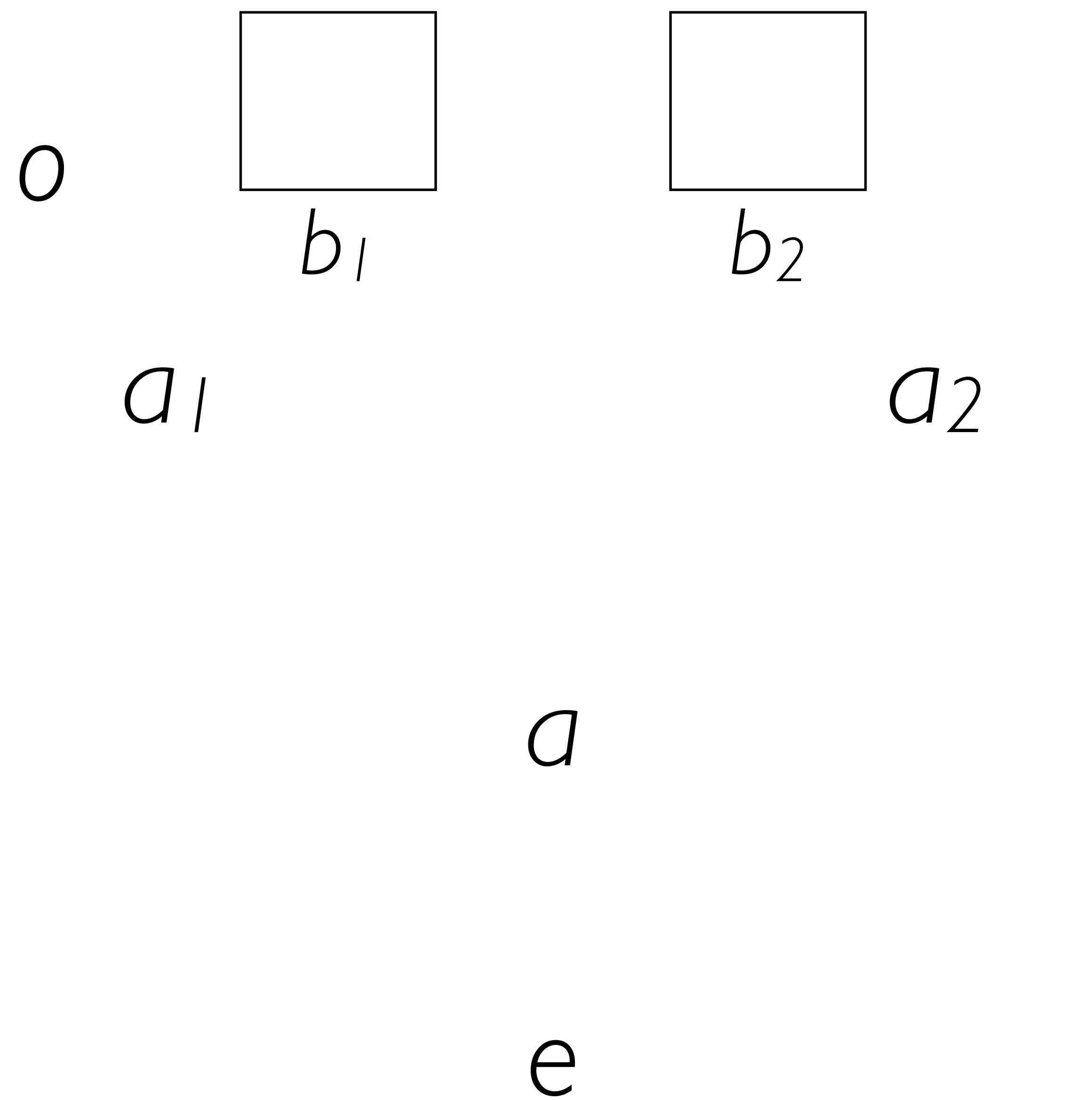
False Belief Task Demands Intensional Logic …



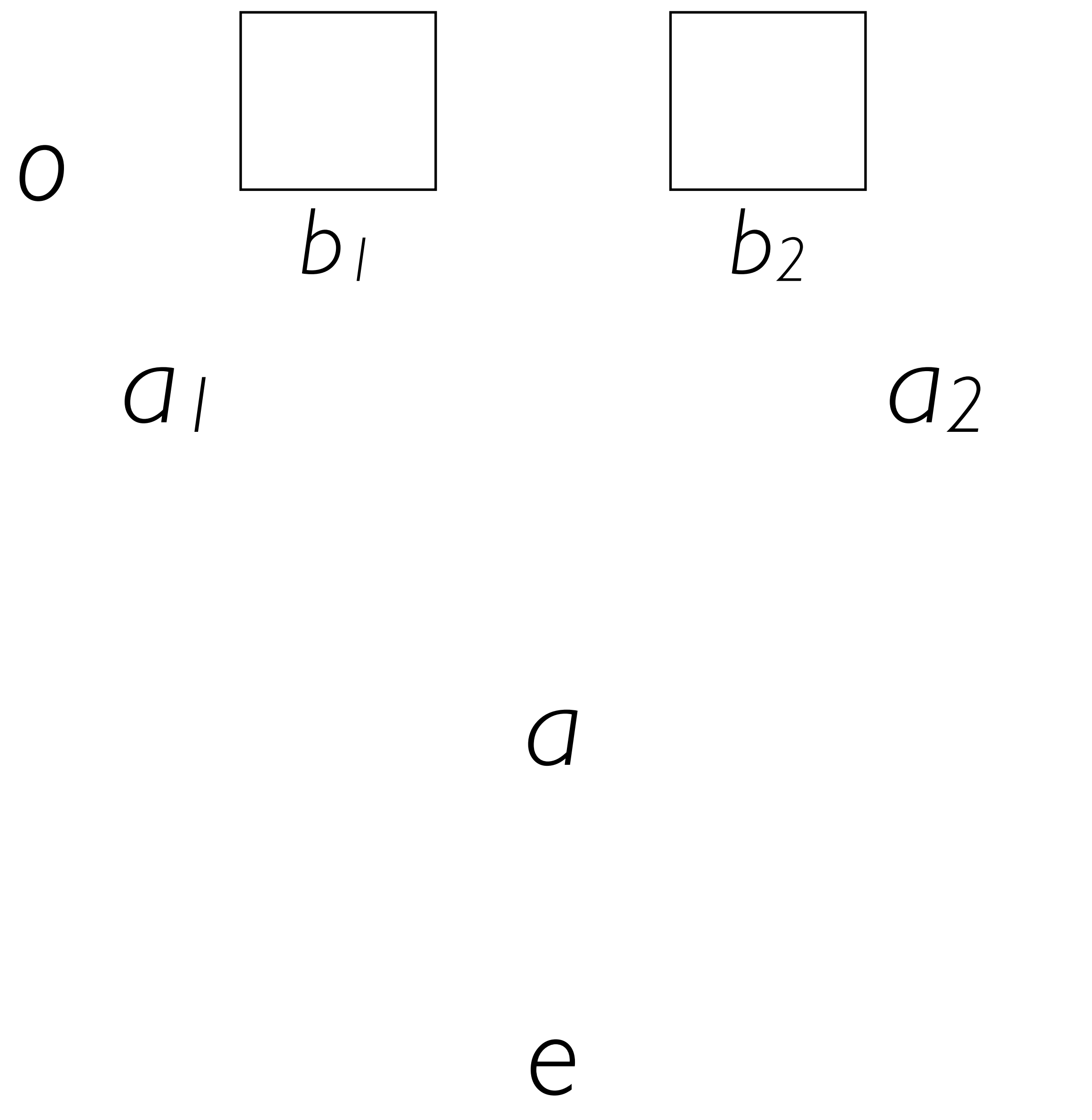
Better, But Embryonic: The ToM Pawn Shop



Framework for FBT$^0{}_1$

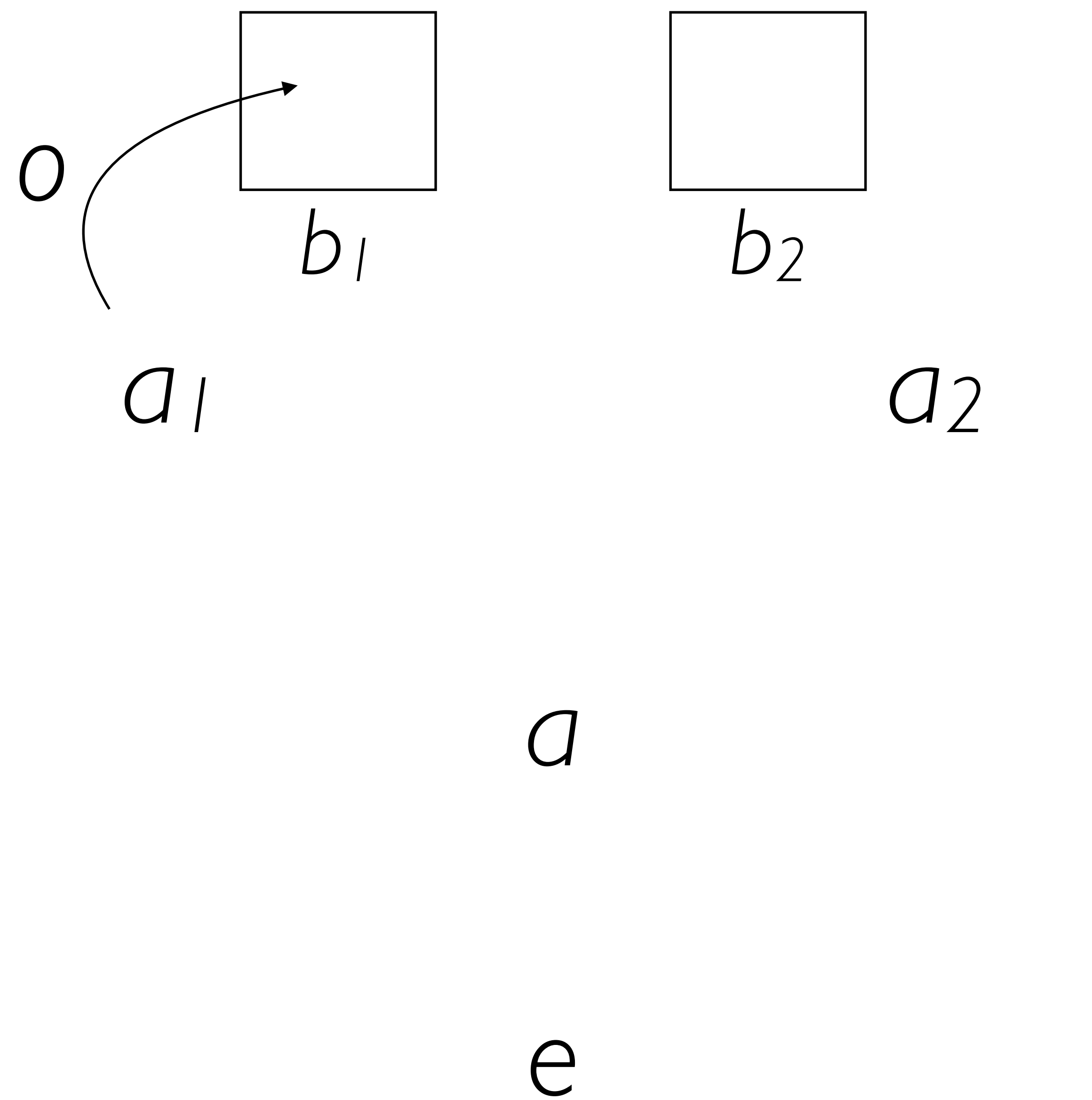


Framework for FBT^0_1
(five timepoints)

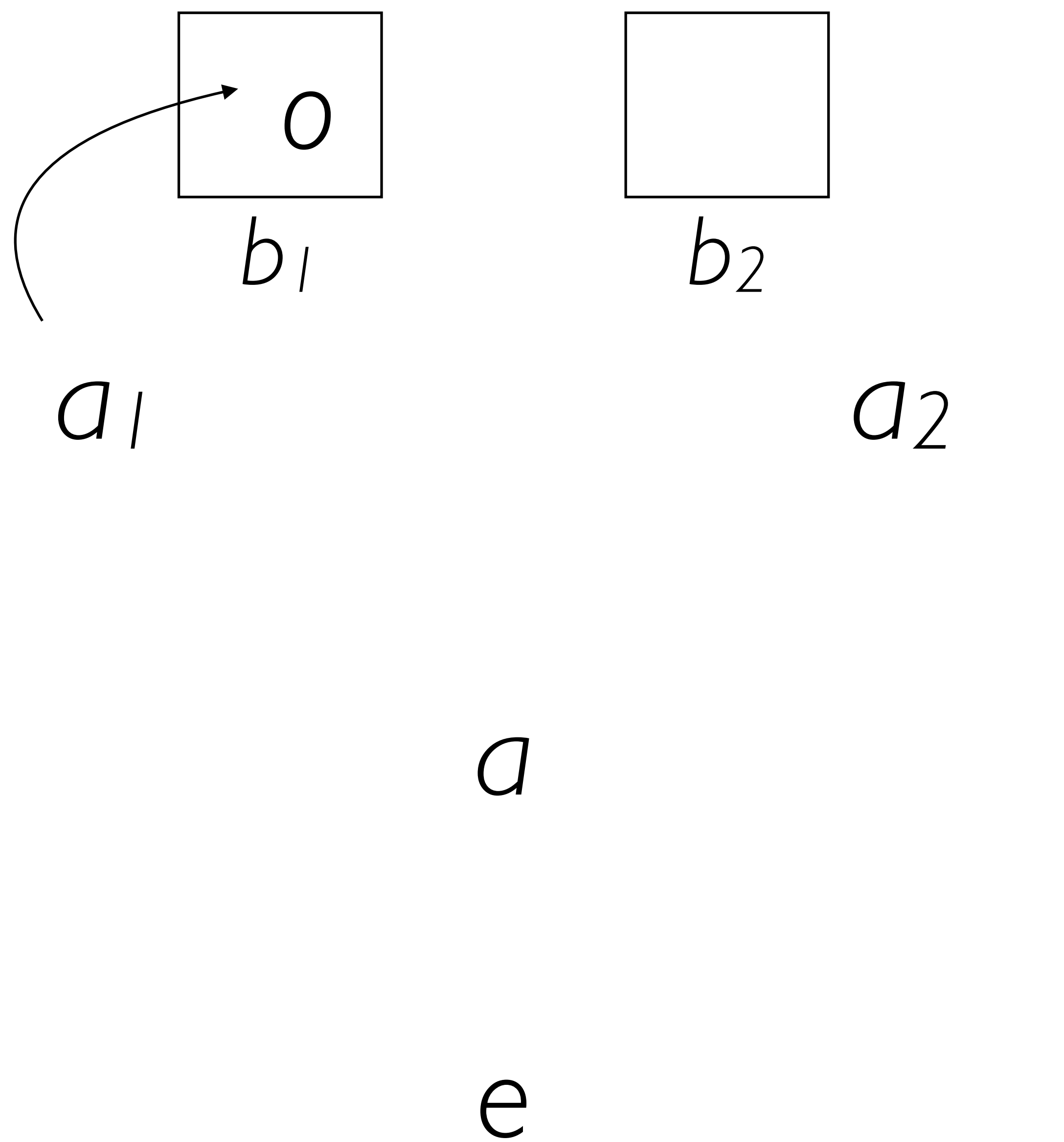


Framework for FBT$^0{}_1$

(five timepoints)

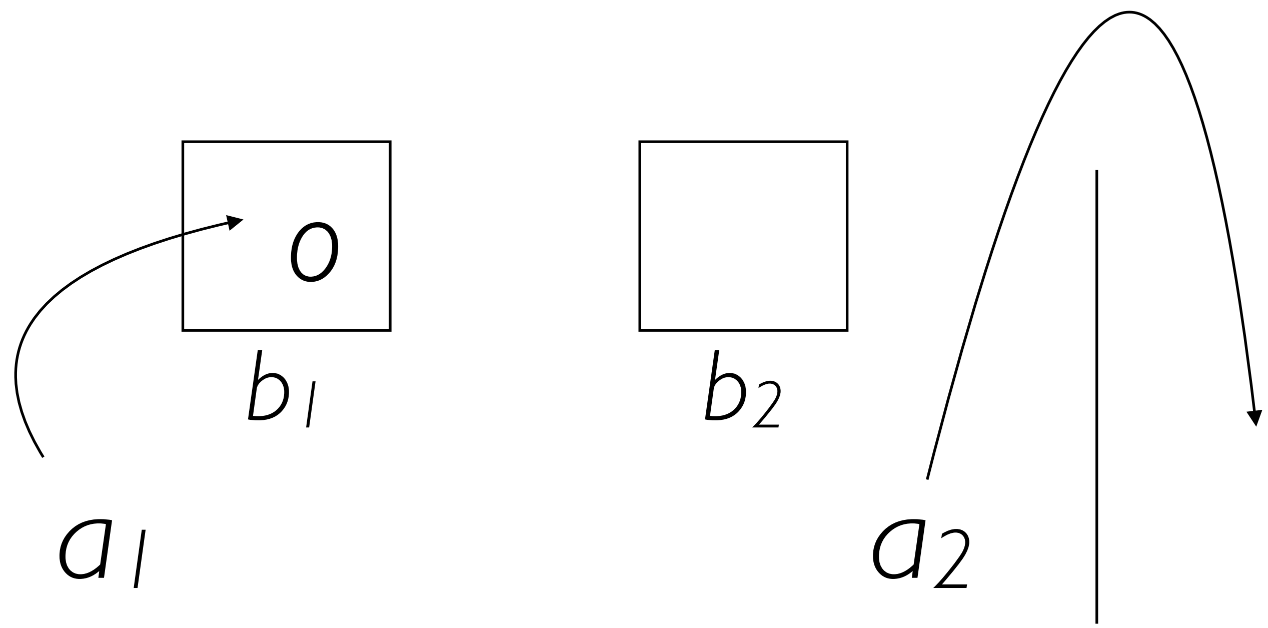


Framework for $FBT^0{}_1$
(five timepoints)



Framework for $FBT^0{}_1$

(five timepoints)

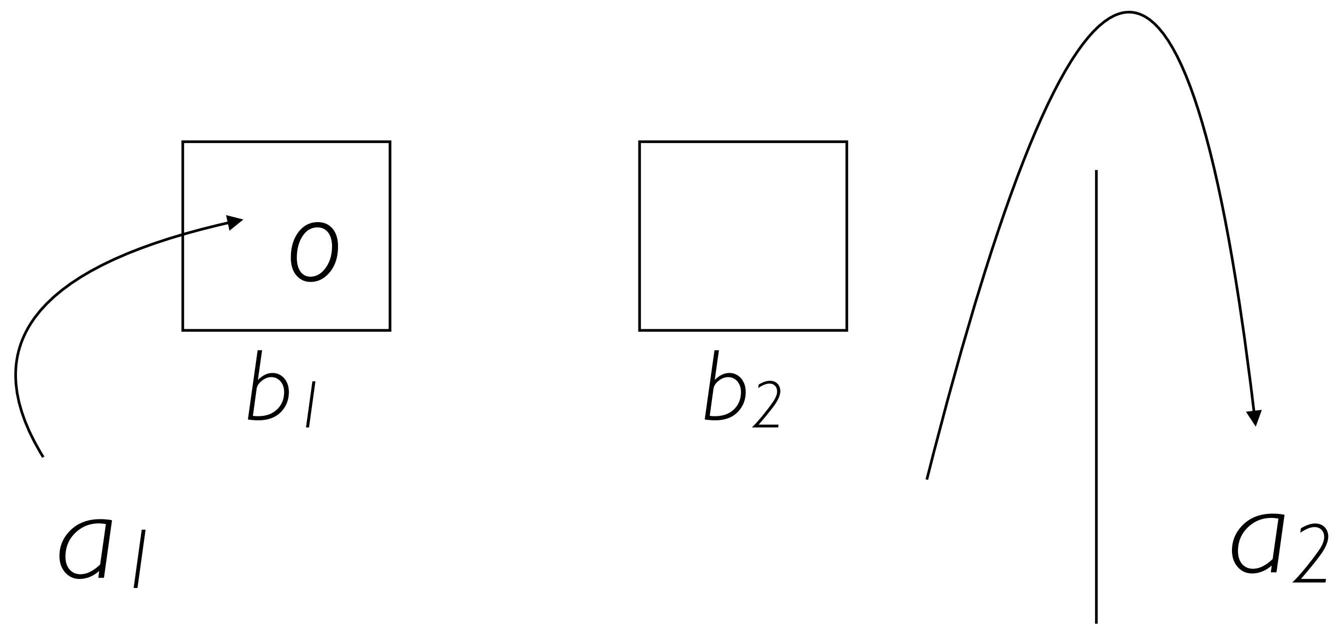


a

e

Framework for $FBT^0{}_1$

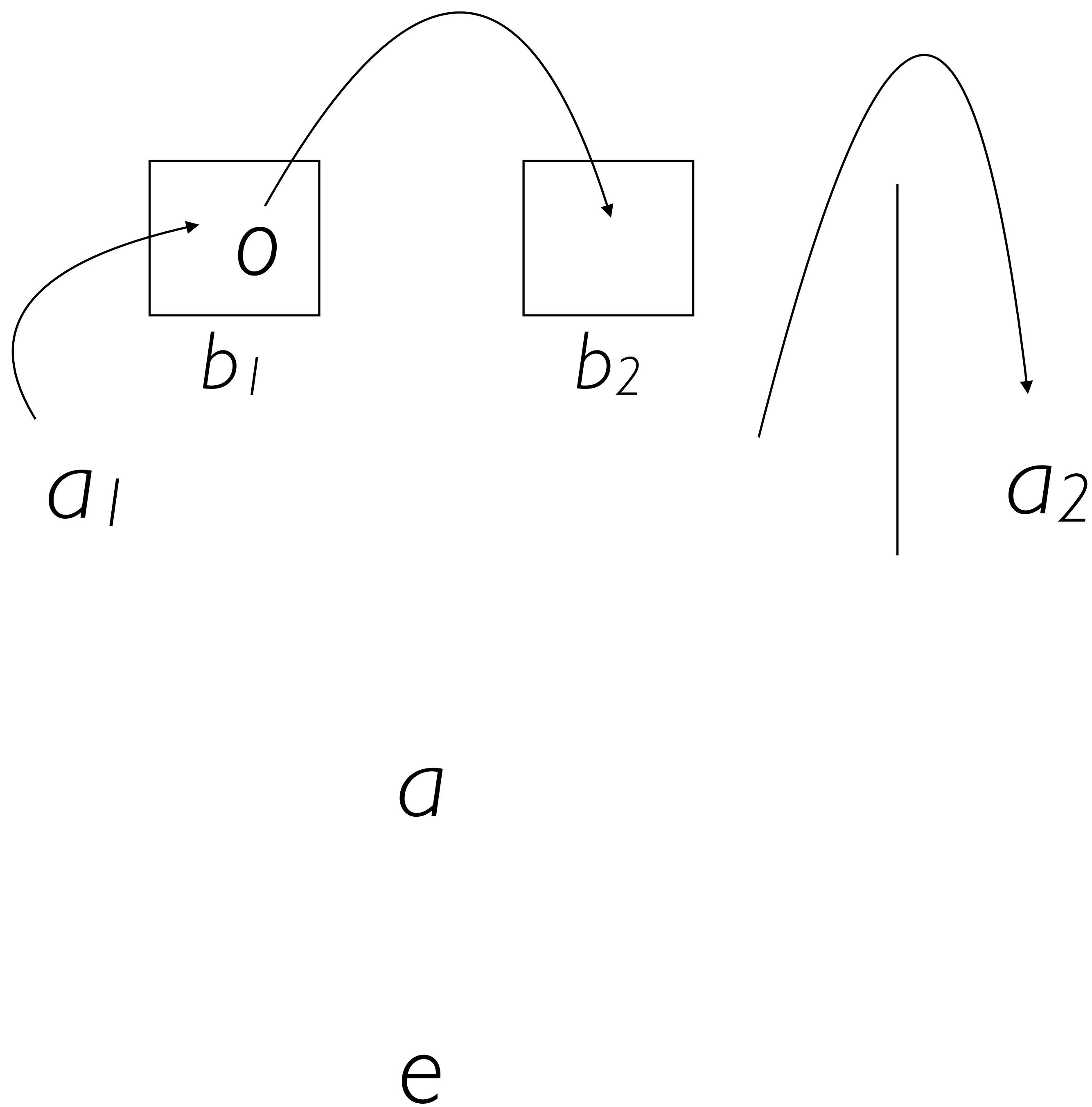
(five timepoints)



a

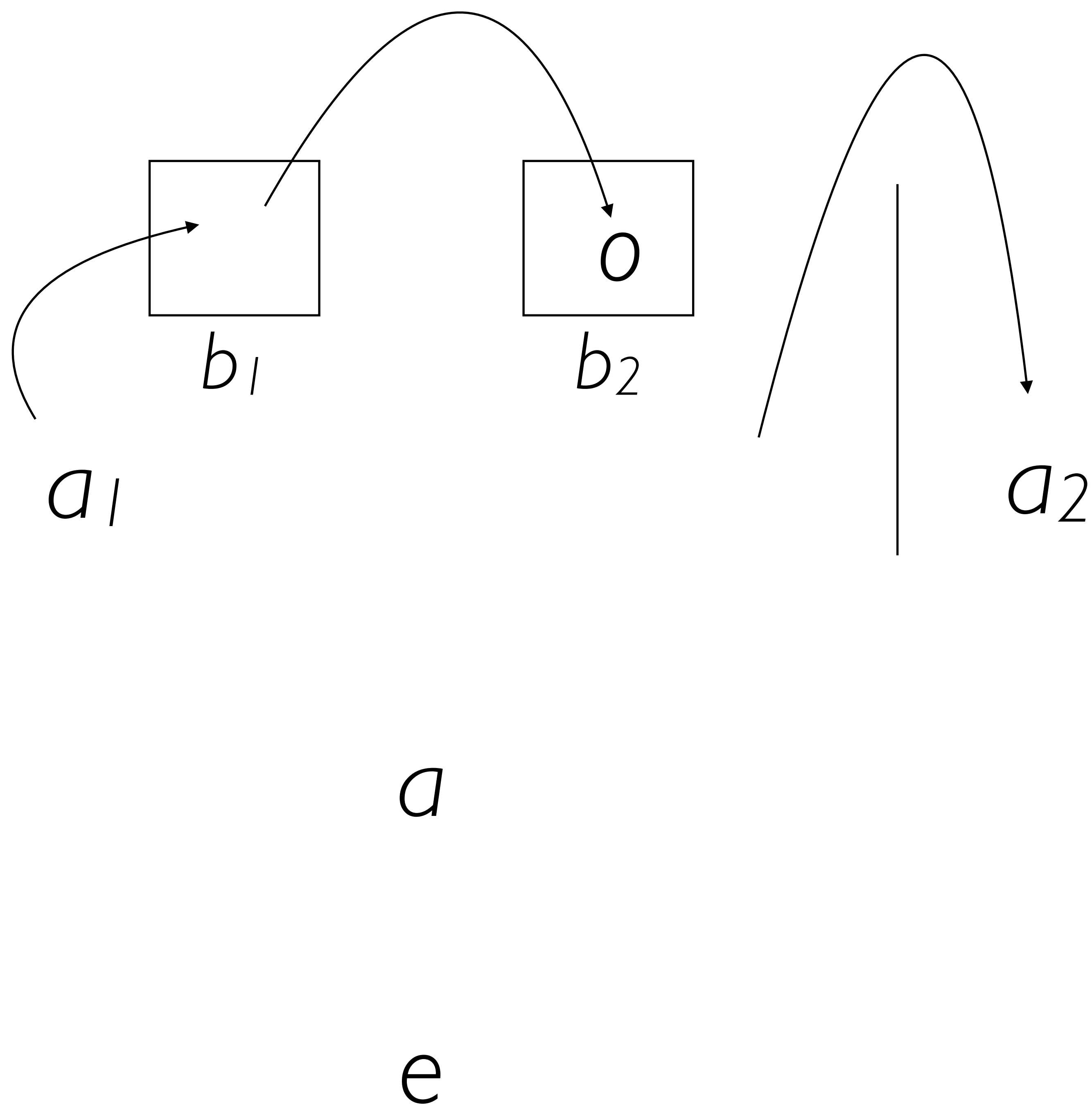
e

Framework for FBT^0_1
(five timepoints)



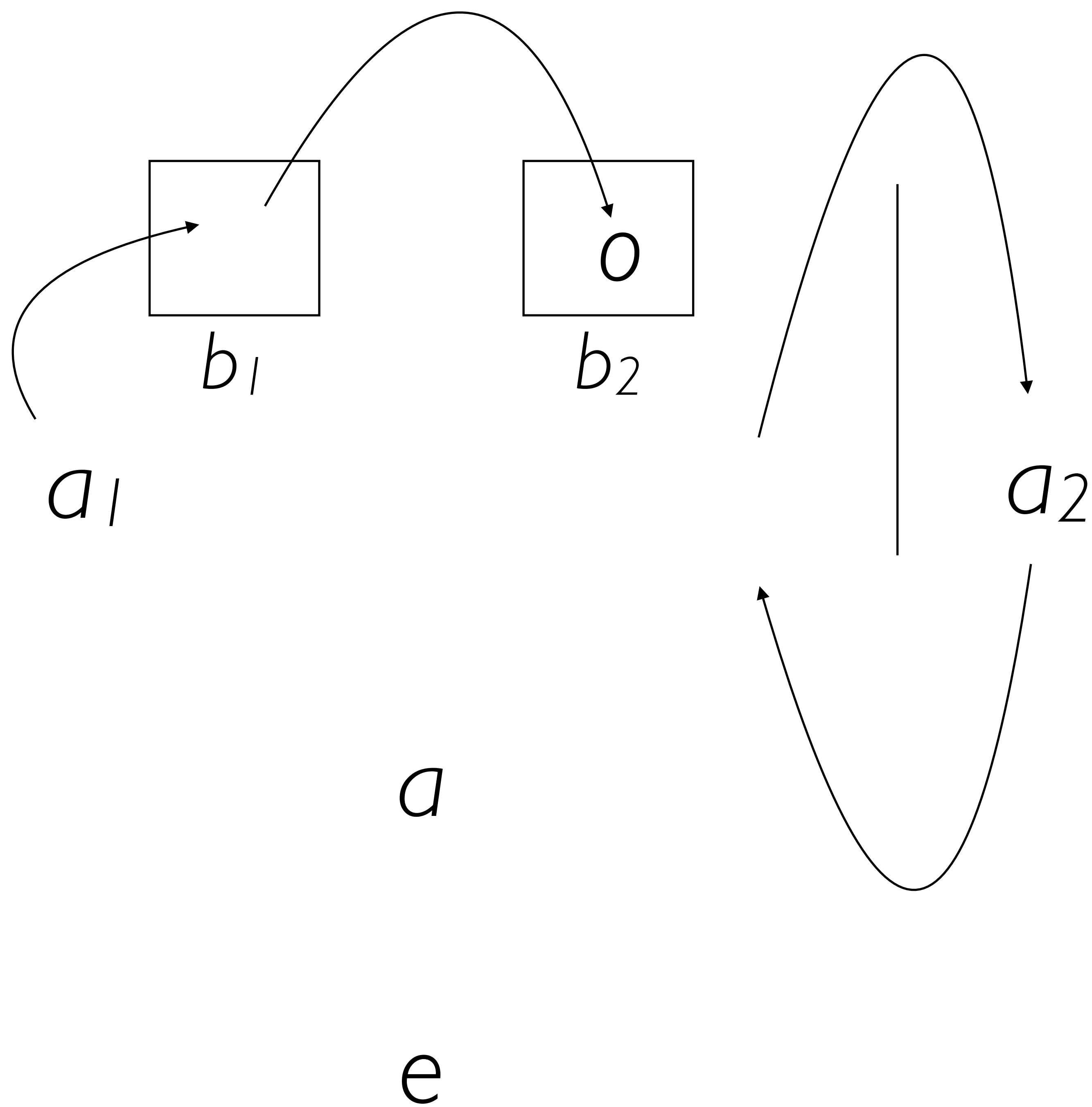
Framework for FBT$^0{}_1$

(five timepoints)



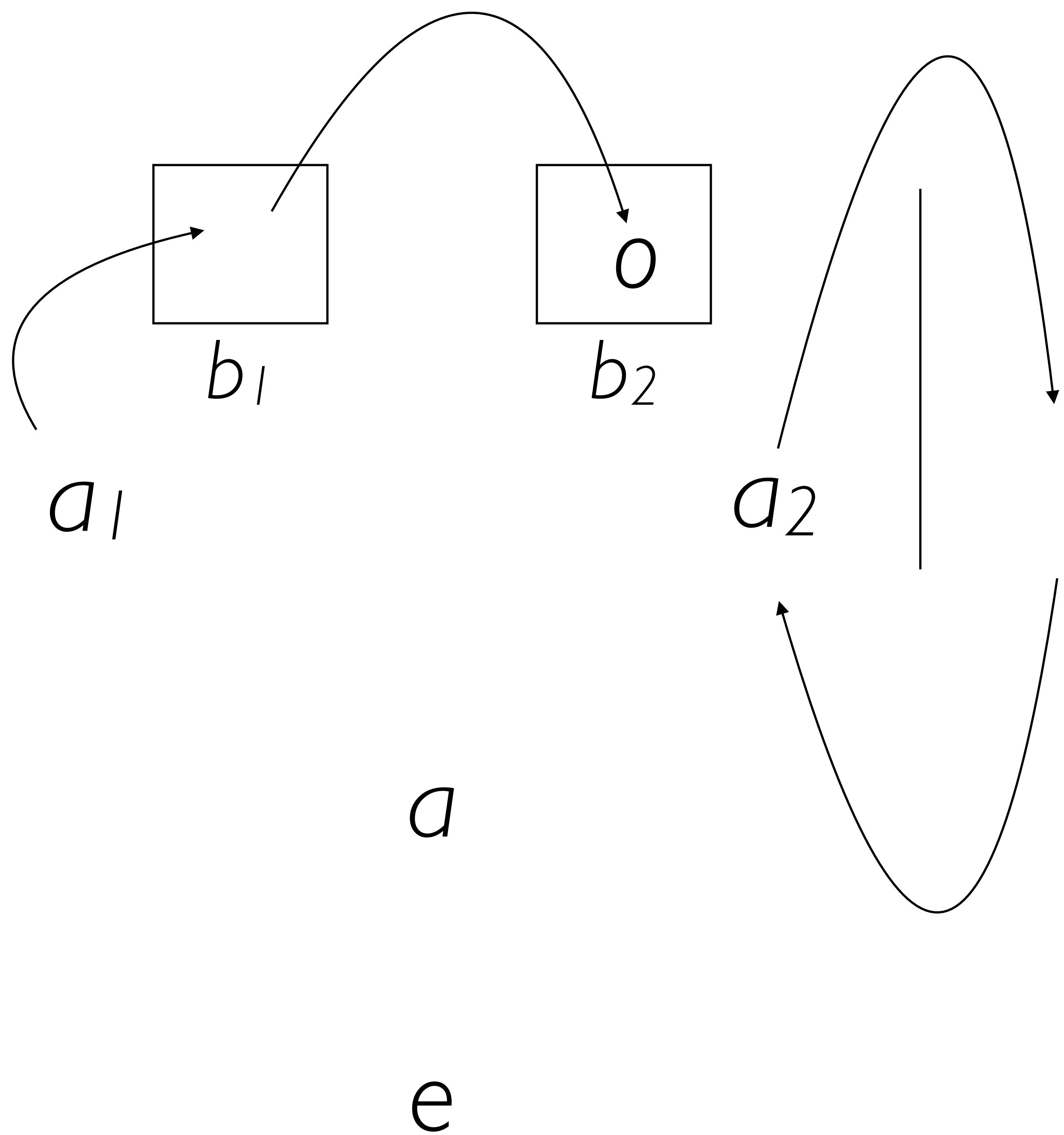
Framework for $FBT^0{}_1$

(five timepoints)

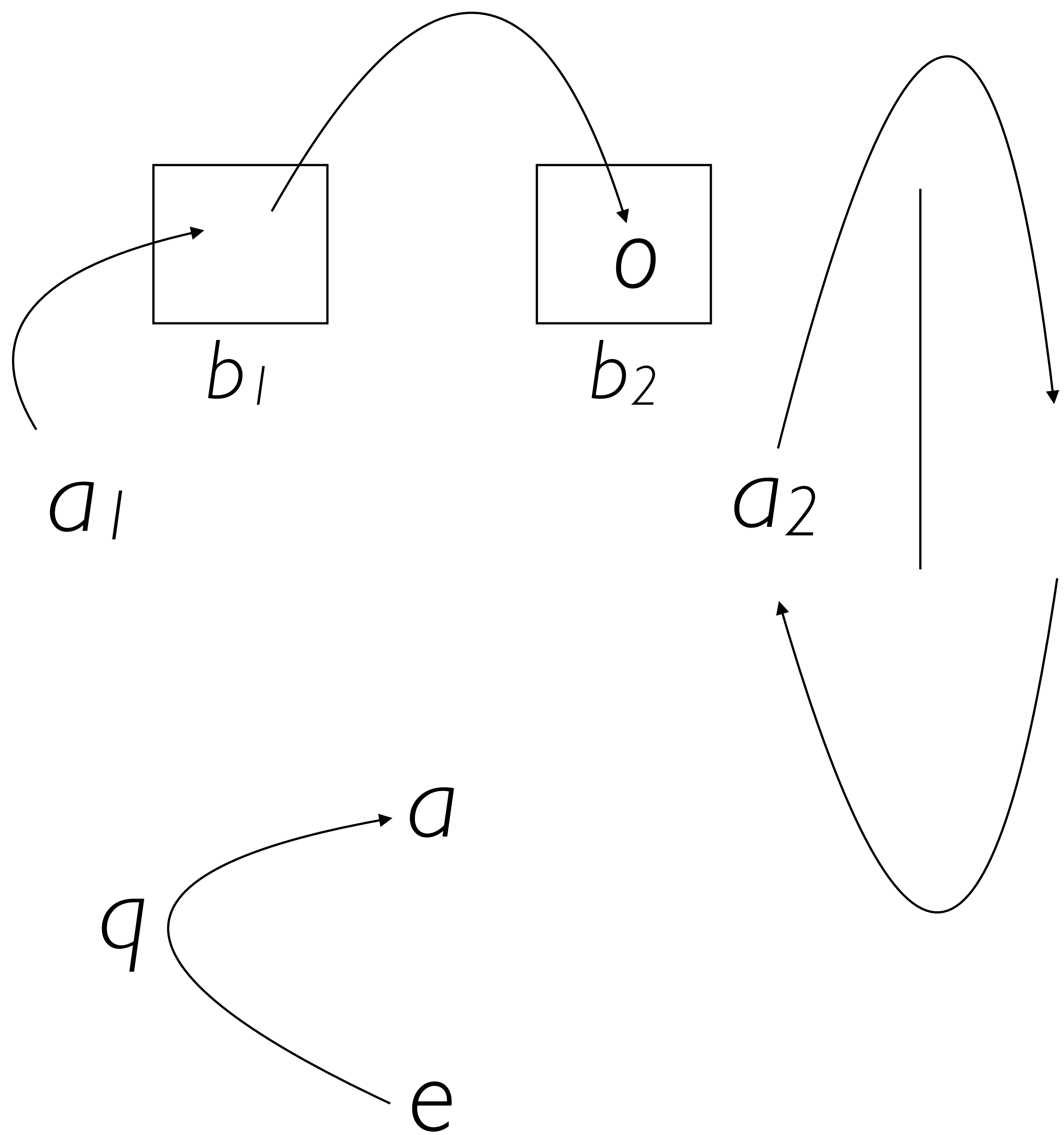


Framework for $FBT^0{}_1$

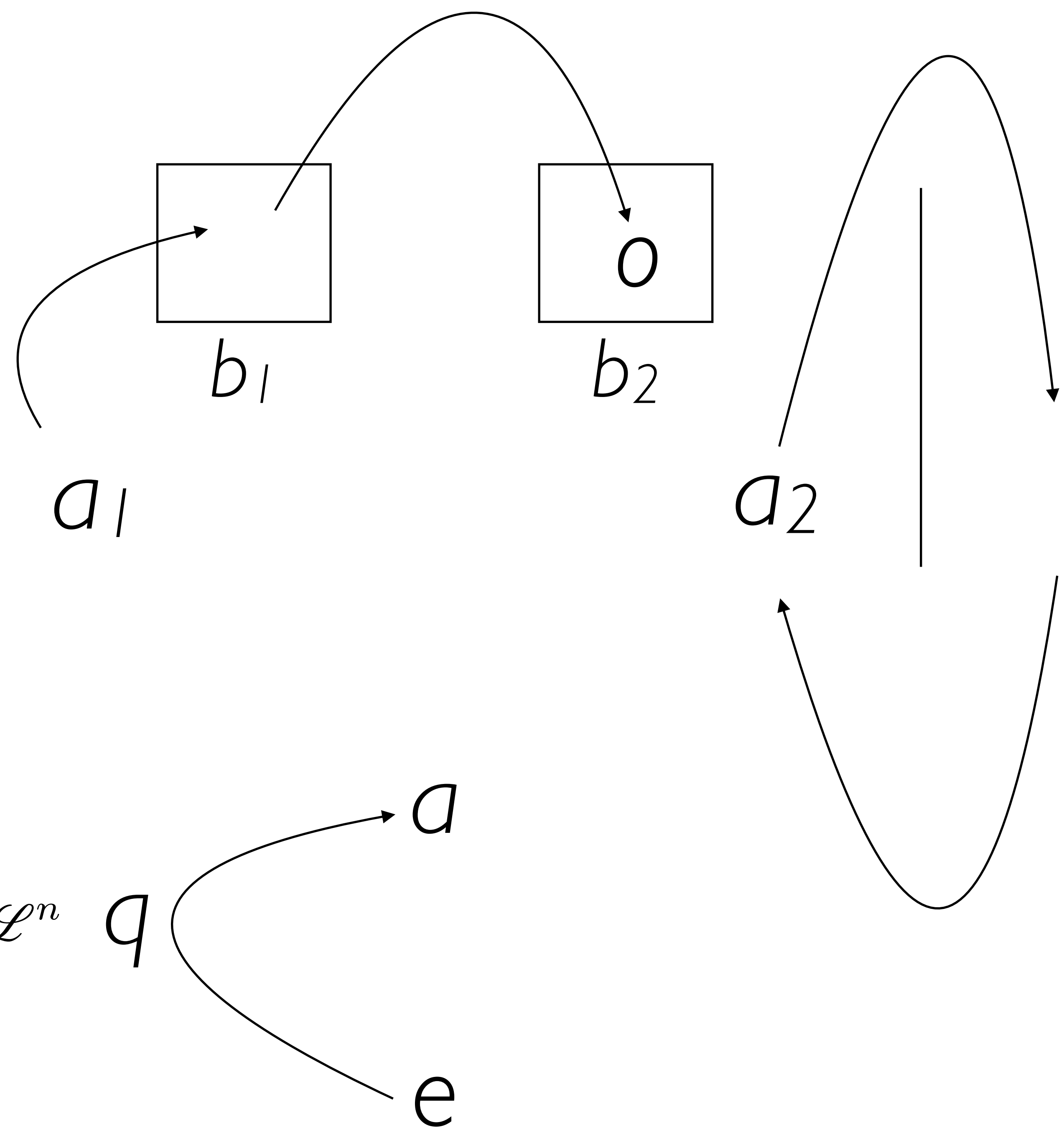
(five timepoints)



Framework for $FBT^0{}_1$
(five timepoints)

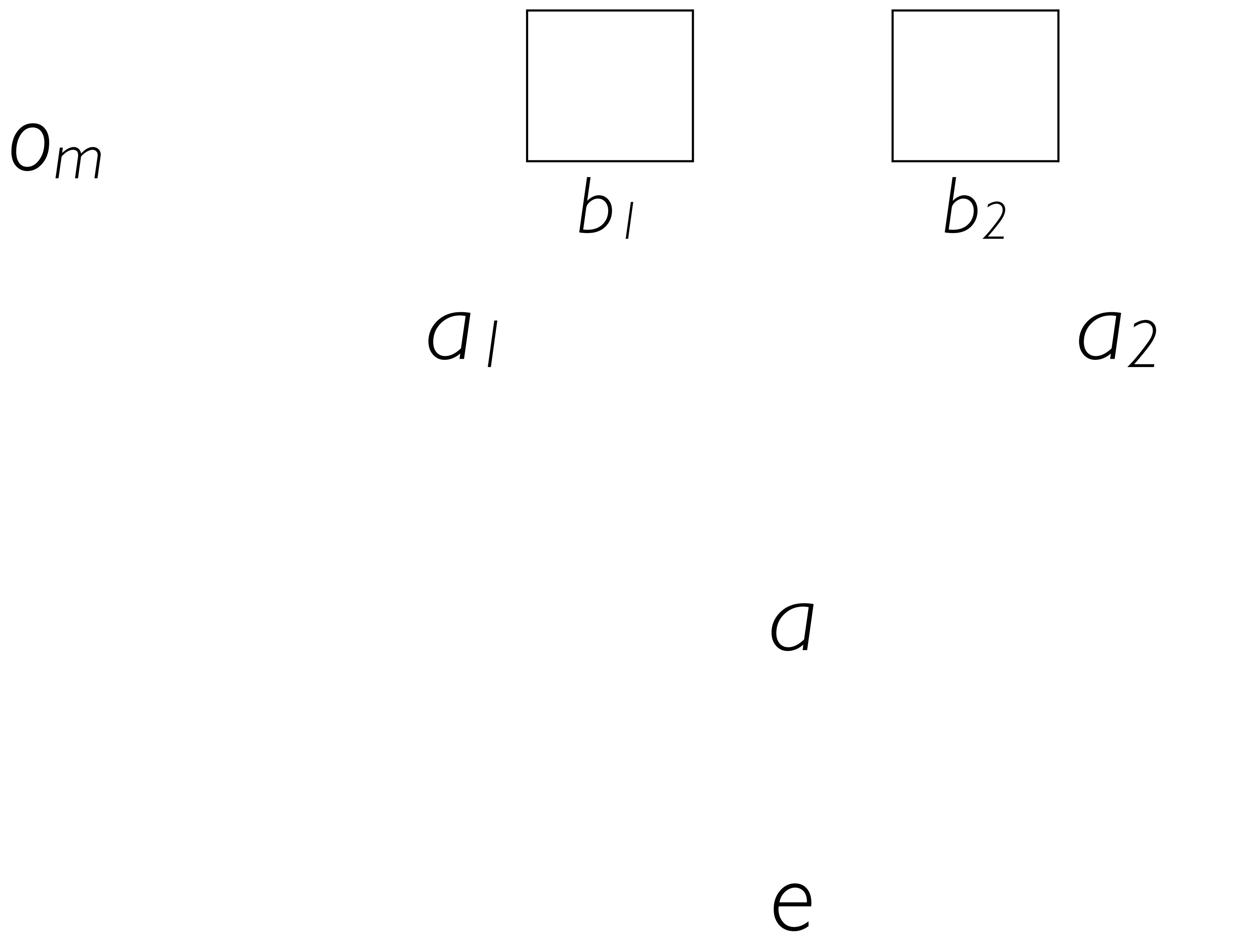


Framework for FBT^0_1
(five timepoints)



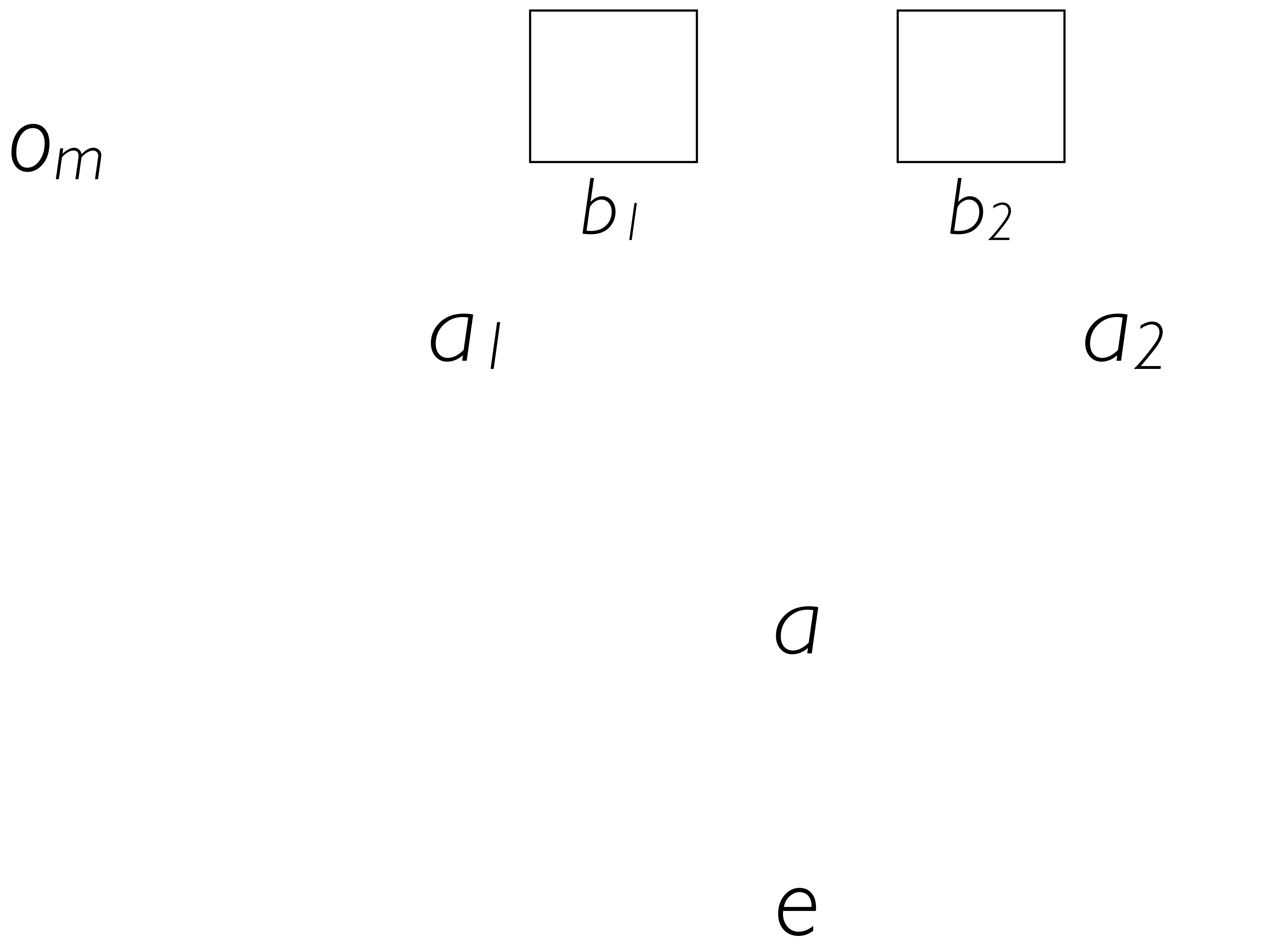
q a formula in modal $\mathscr{L}^n$

Framework for FBT$^1_1$

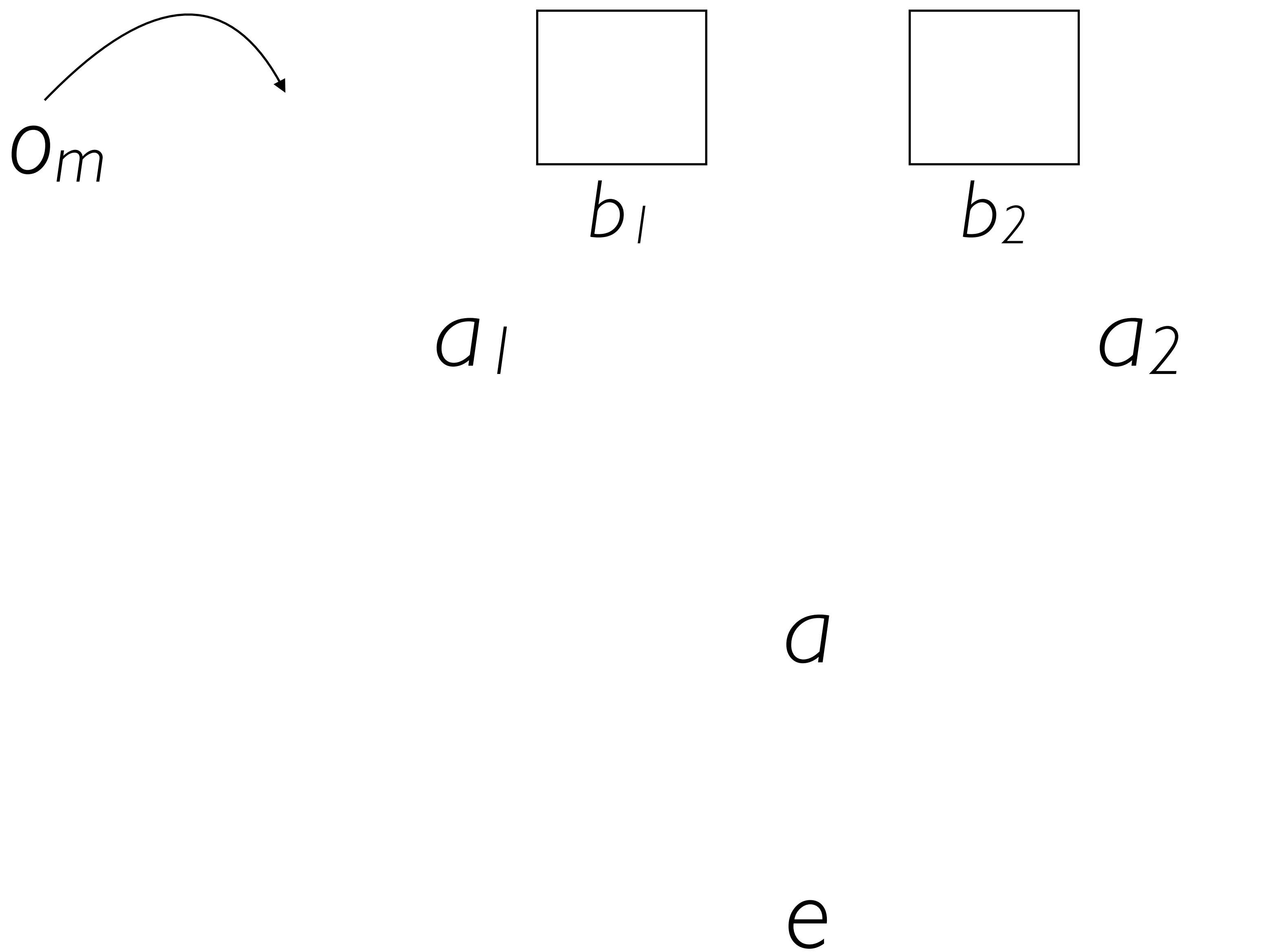


Framework for FBT^1_1

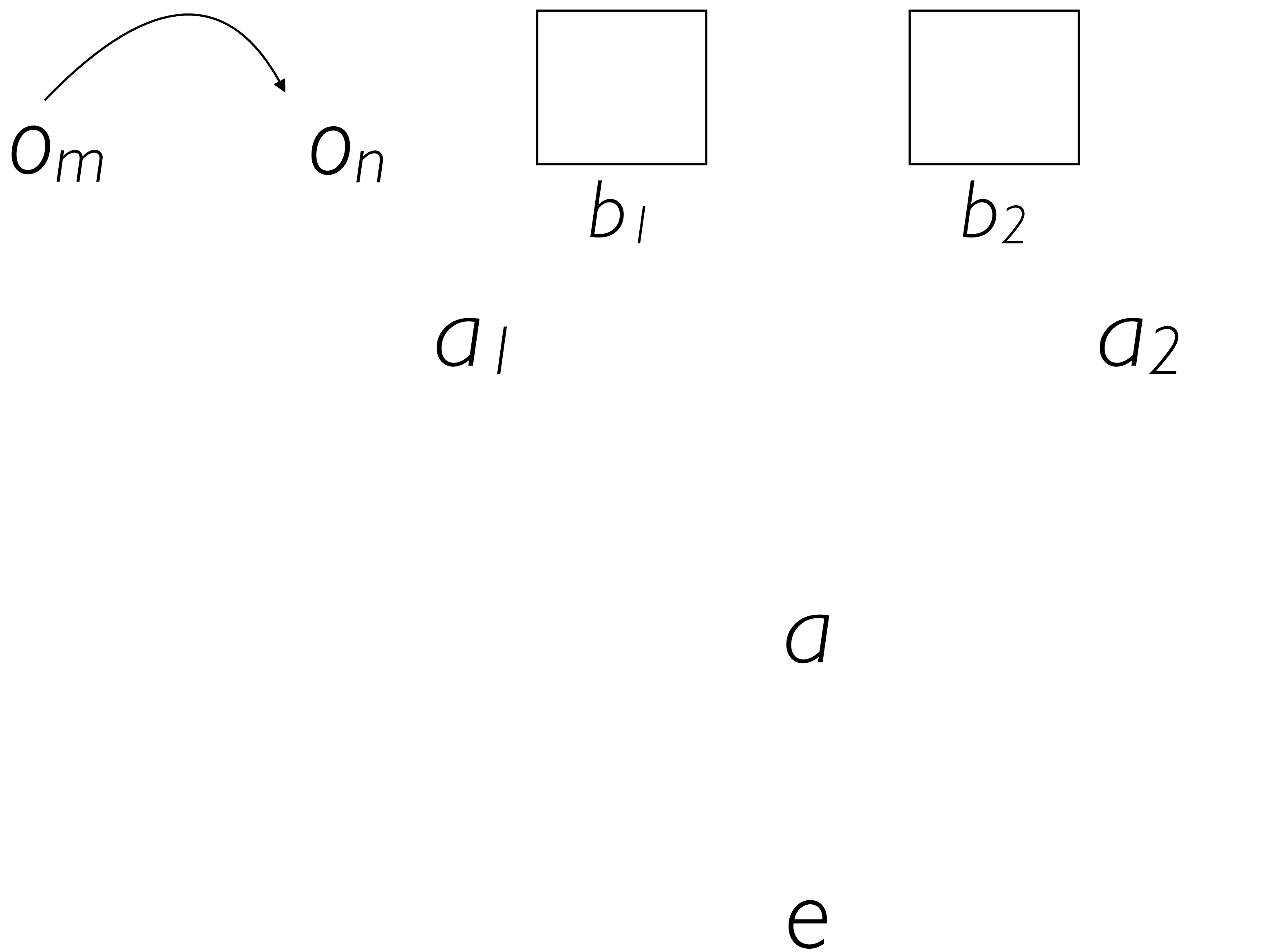
(six timepoints)



Framework for FBT^1_1
(six timepoints)

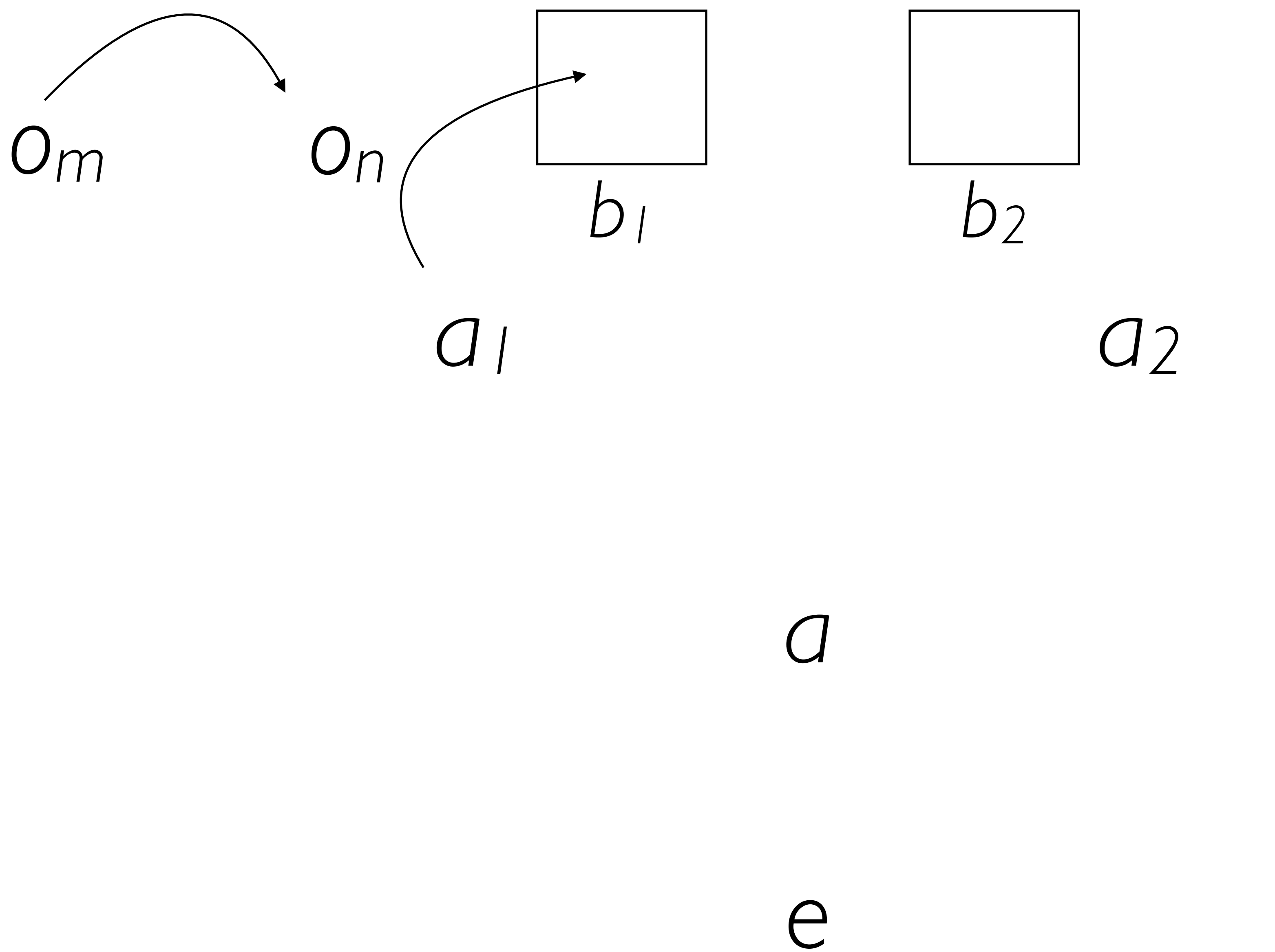


Framework for FBT^{1}_{1}
(six timepoints)

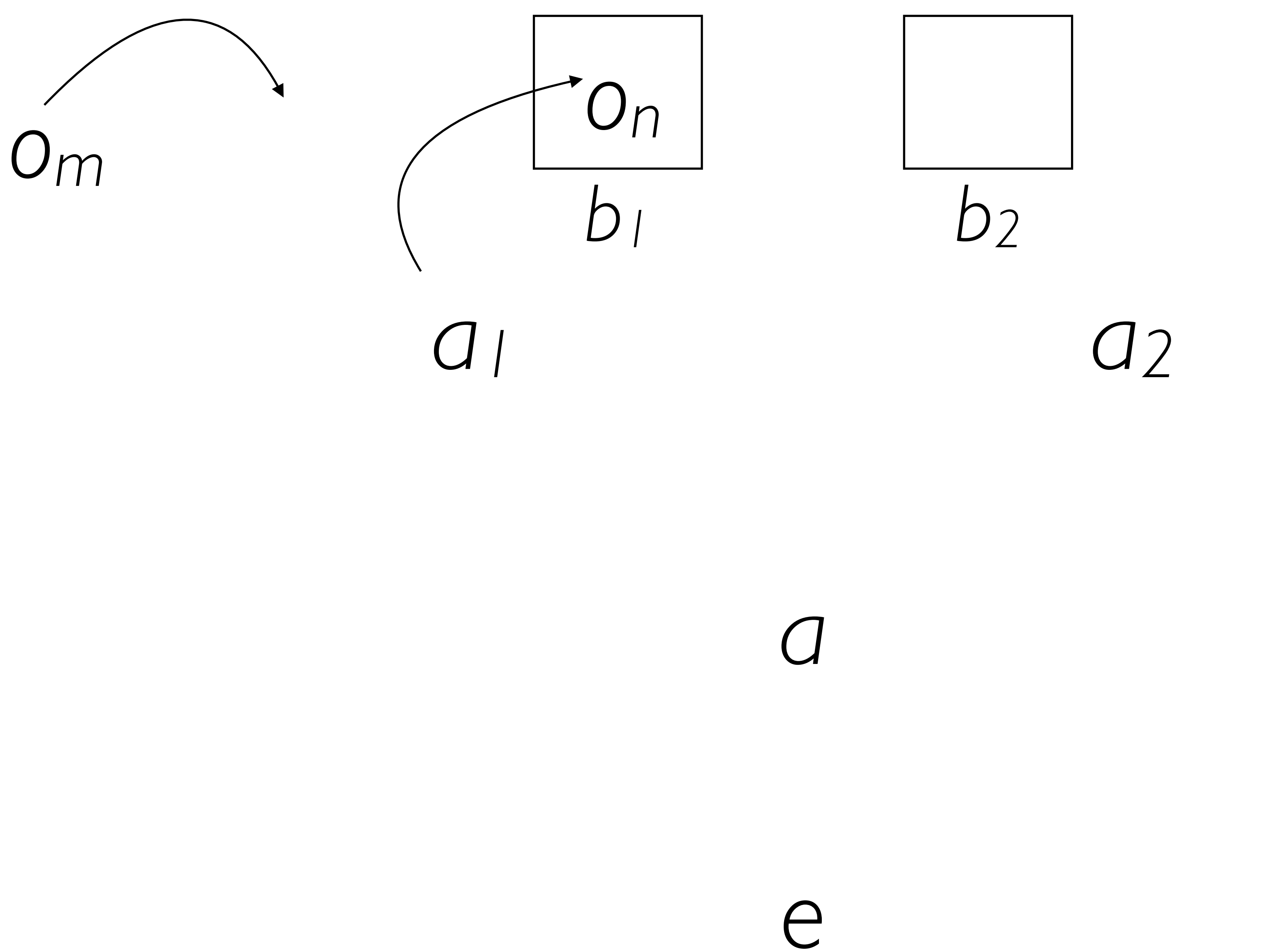


Framework for FBT^{1}_{1}

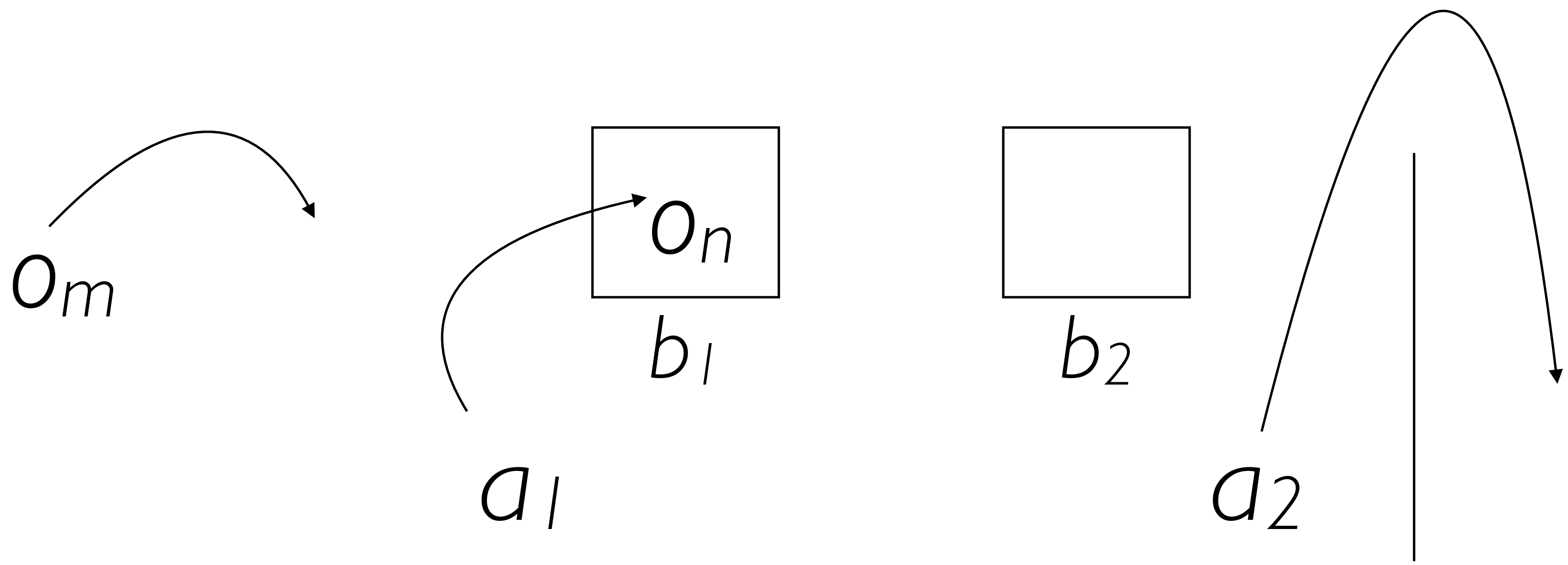
(six timepoints)



Framework for FBT$^{1}_{1}$
(six timepoints)



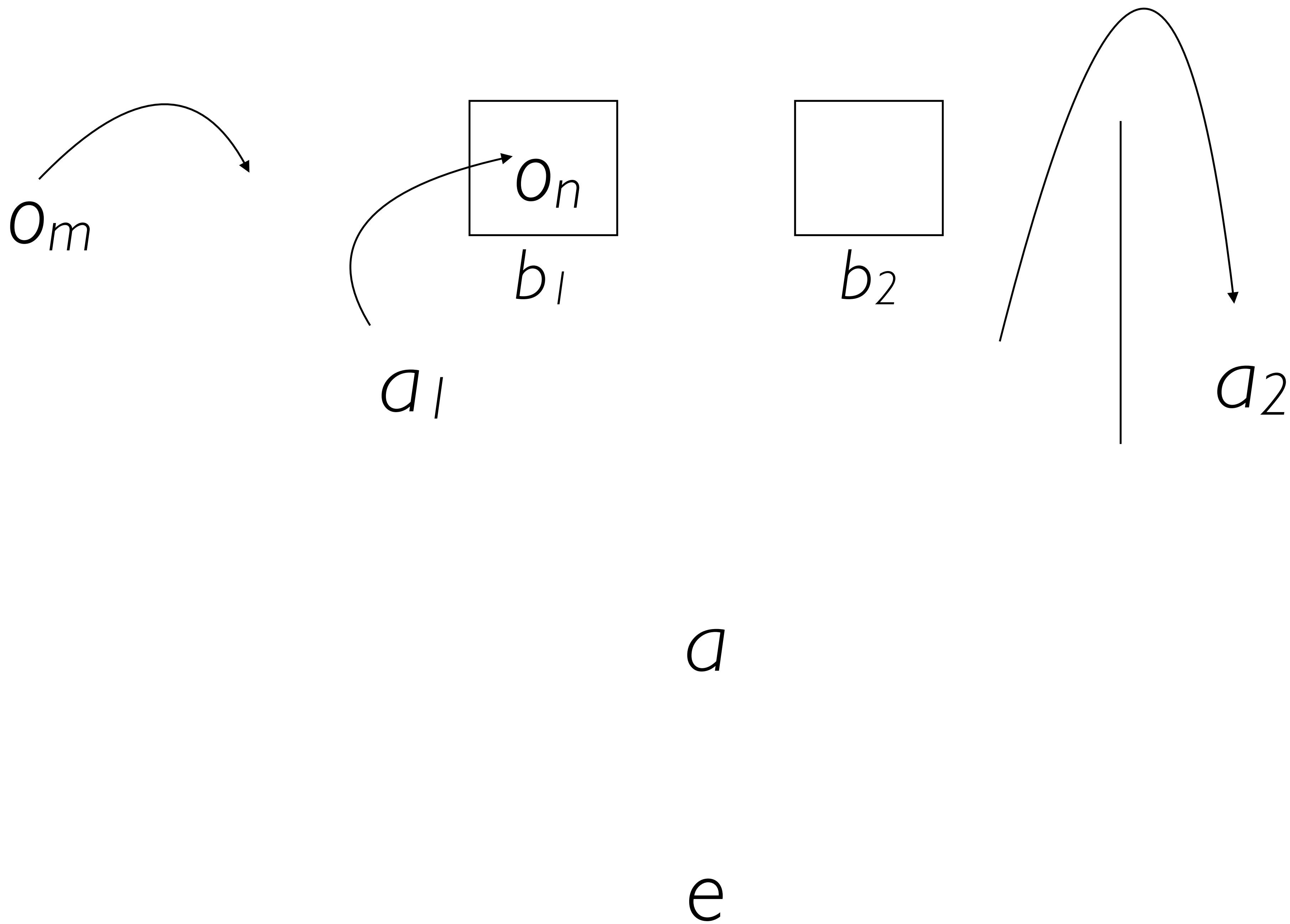
Framework for FBT^1_1
(six timepoints)



a

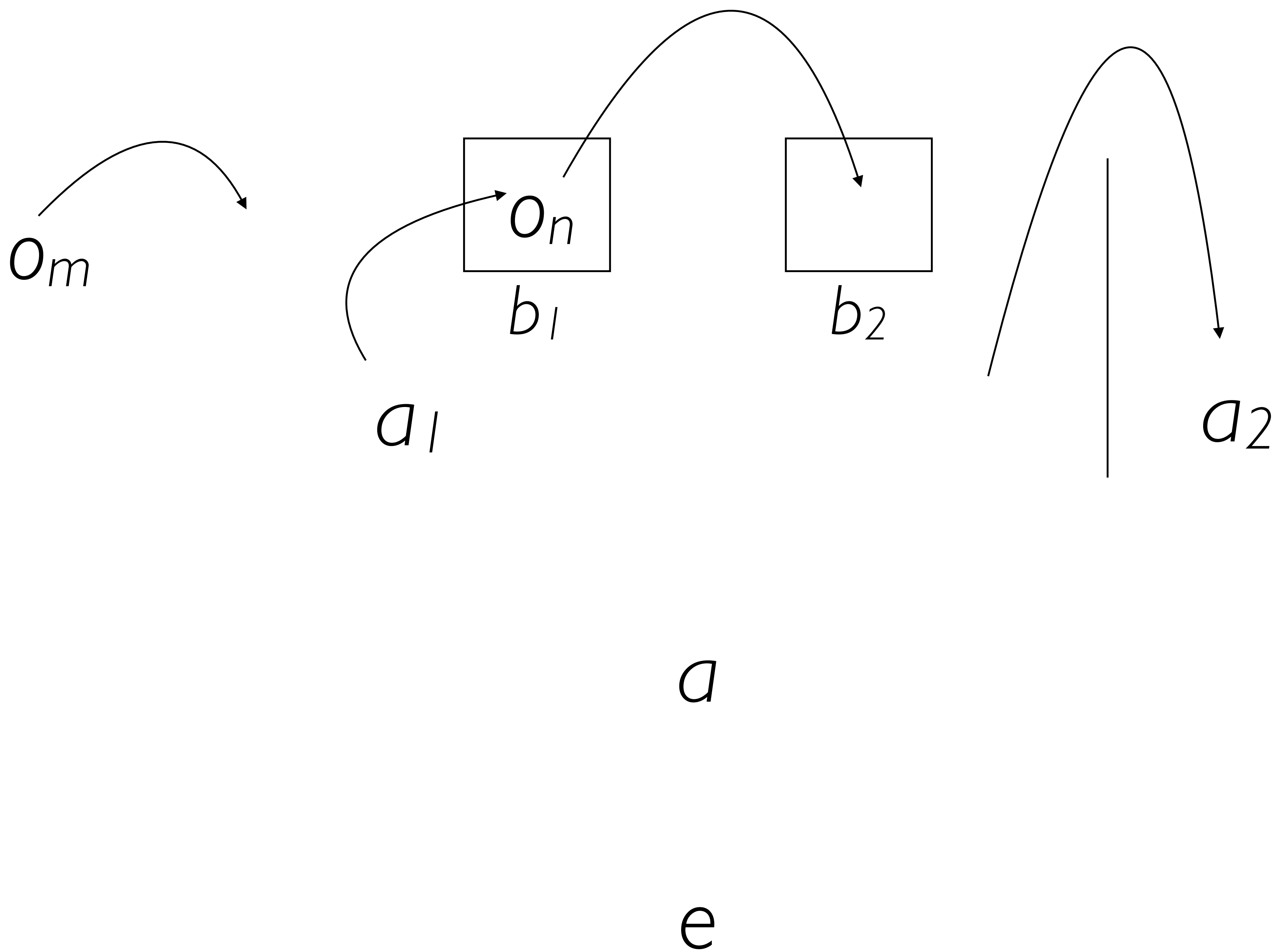
e

Framework for FBT$^1_1$
(six timepoints)

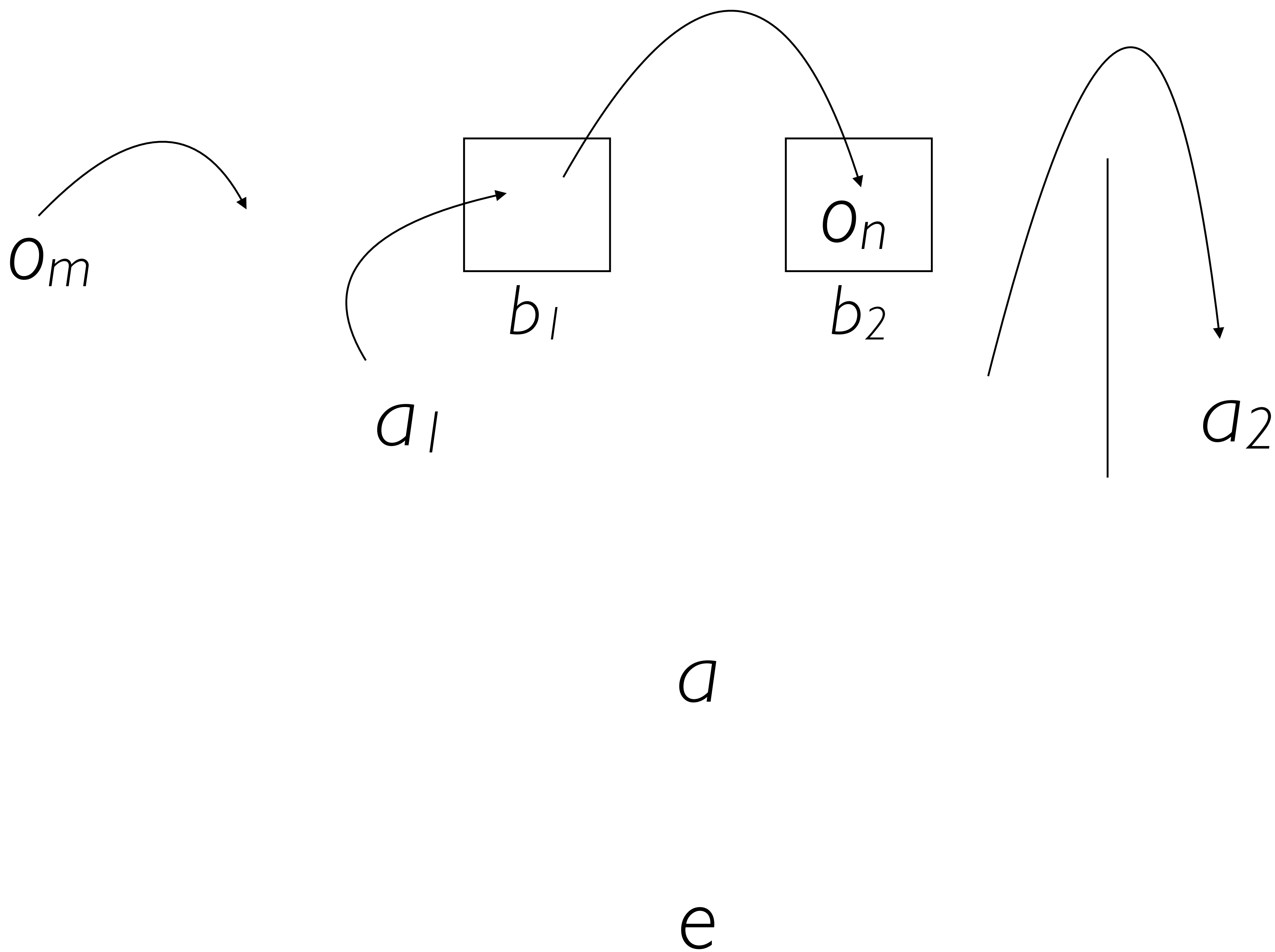


Framework for FBT$^1{}_1$

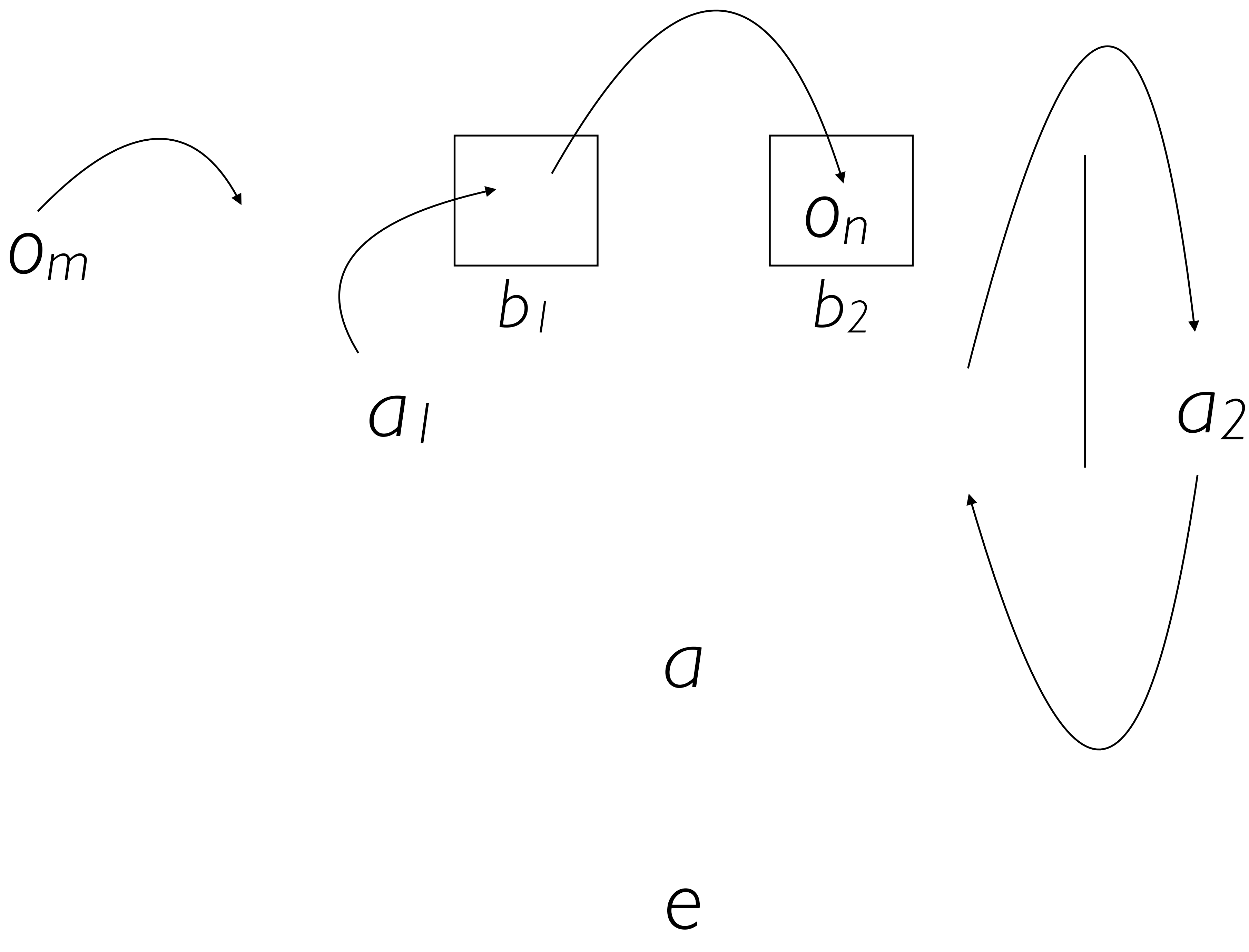
(six timepoints)



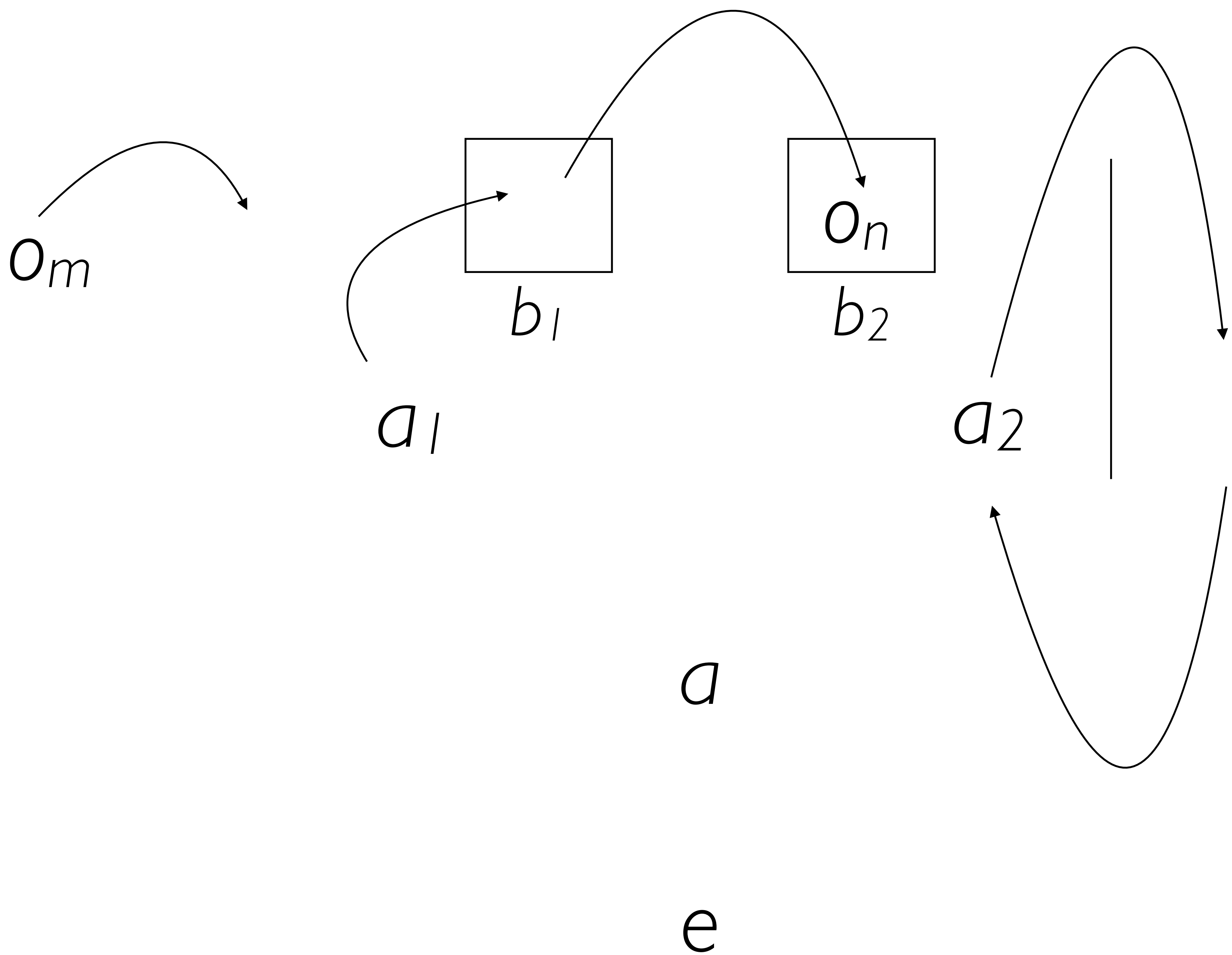
Framework for FBT^{1}_{1}
(six timepoints)



Framework for $FBT^1{}_1$
(six timepoints)

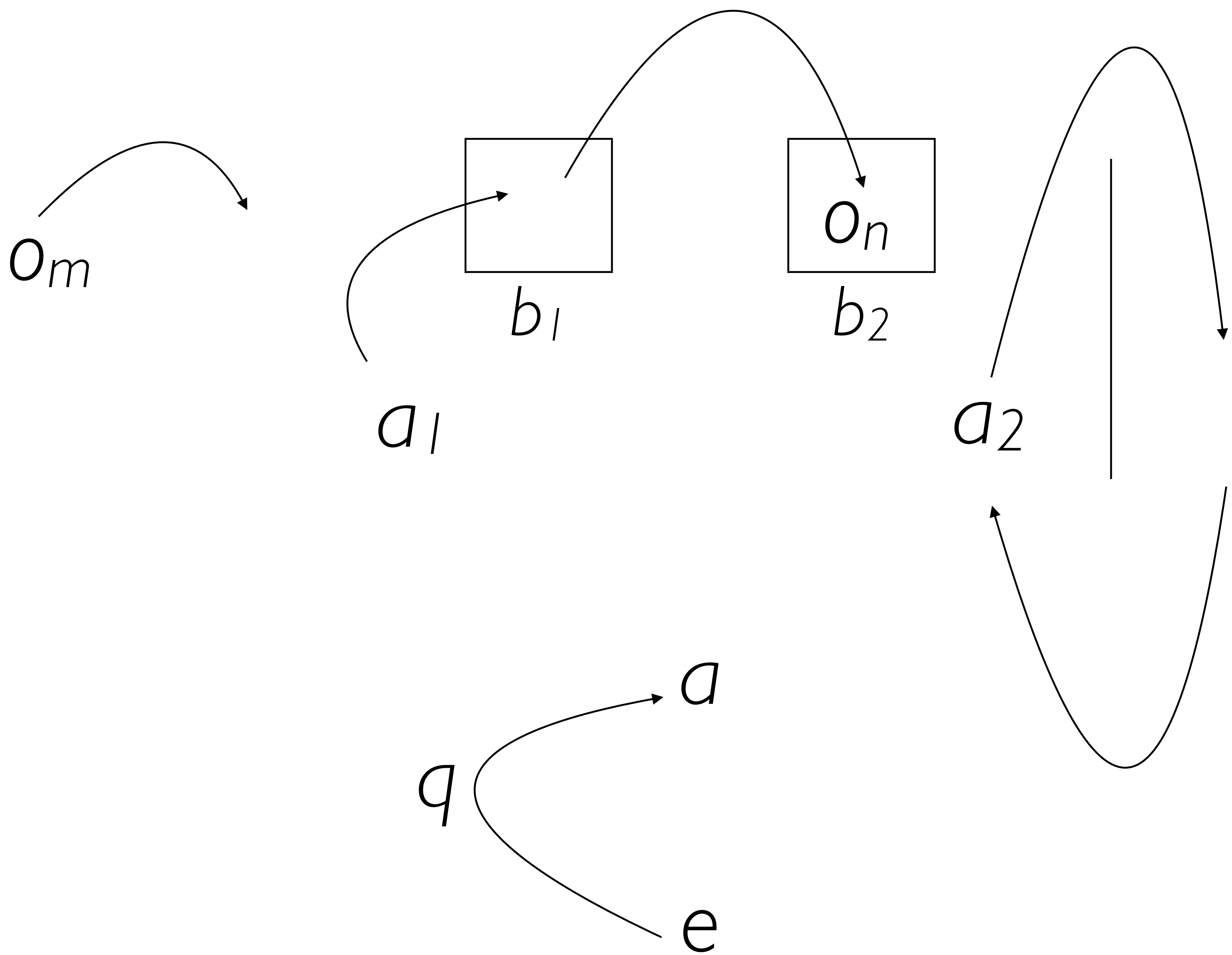


Framework for $FBT^1{}_1$
(six timepoints)

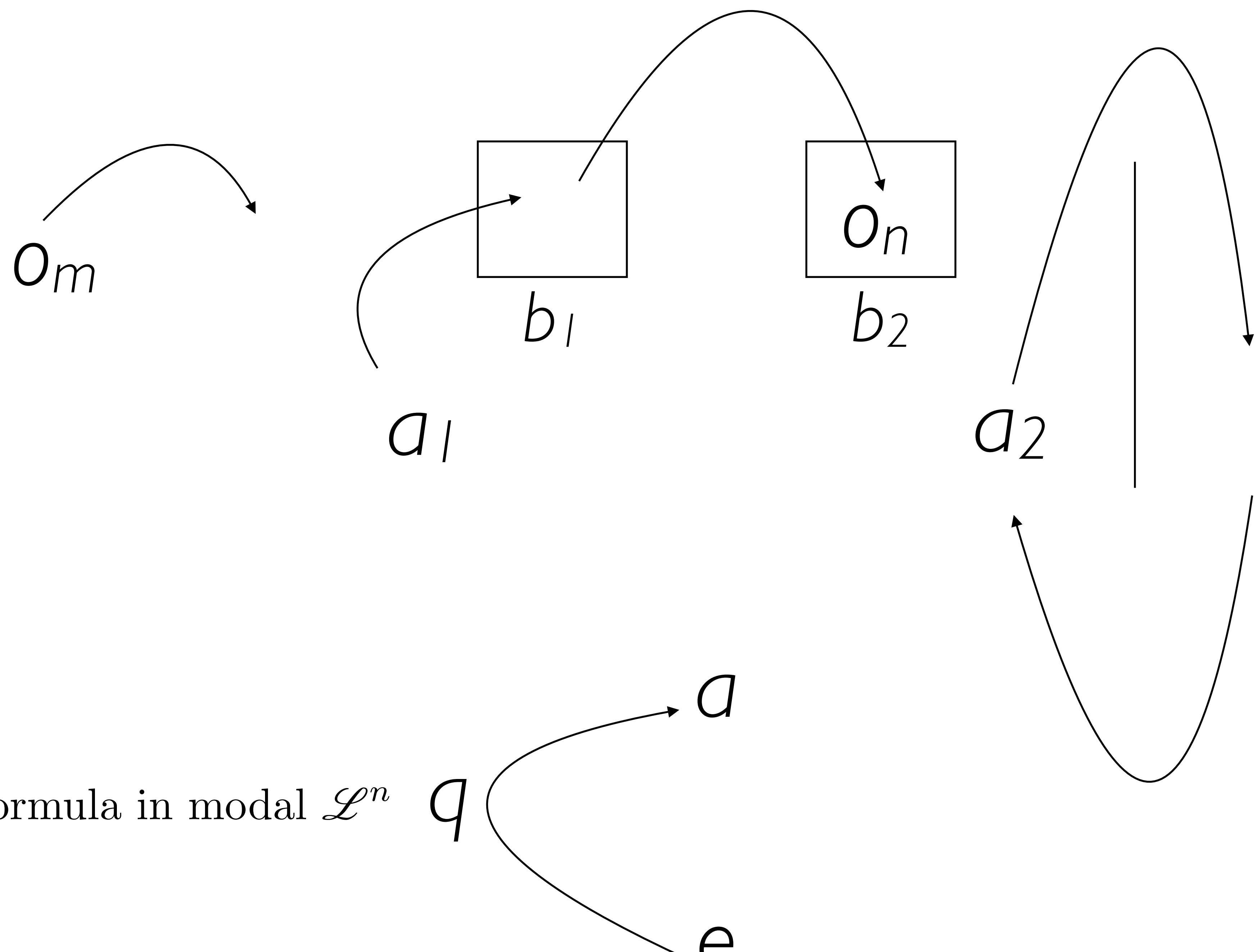


Framework for $FBT^1{}_1$

(six timepoints)



Framework for $FBT^1{}_1$
(six timepoints)



q a formula in modal $\mathscr{L}^n$

Done, a Decade Ago, Formally & Implementation/Simulation

Arkoudas, K. & Bringsjord, S. (2009) “Propositional Attitudes and Causation” *International Journal of Software and Informatics* **3.1**: 47–65.

http://kryten.mm.rpi.edu/PRICAI_w_sequentcalc_041709.pdf

Propositional attitudes and causation

Konstantine Arkoudas and Selmer Bringsjord

Cognitive Science and Computer Science Departments, RPI
arkouk@rpi.edu, brings@rpi.edu

Abstract. Predicting and explaining the behavior of others in terms of mental states is indispensable for everyday life. It will be equally important for artificial agents. We present an inference system for representing and reasoning about mental states, and use it to provide a formal analysis of the false-belief task. The system allows for the representation of information about events, causation, and perceptual, doxastic, and epistemic states (vision, belief, and knowledge), incorporating ideas from the event calculus and multi-agent epistemic logic. Unlike previous AI formalisms, our focus here is on mechanized proofs and proof programmability, not on metamathematical results. Reasoning is performed via relatively cognitively plausible inference rules, and a degree of automation is achieved by general-purpose inference methods and by a syntactic embedding of the system in first-order logic.

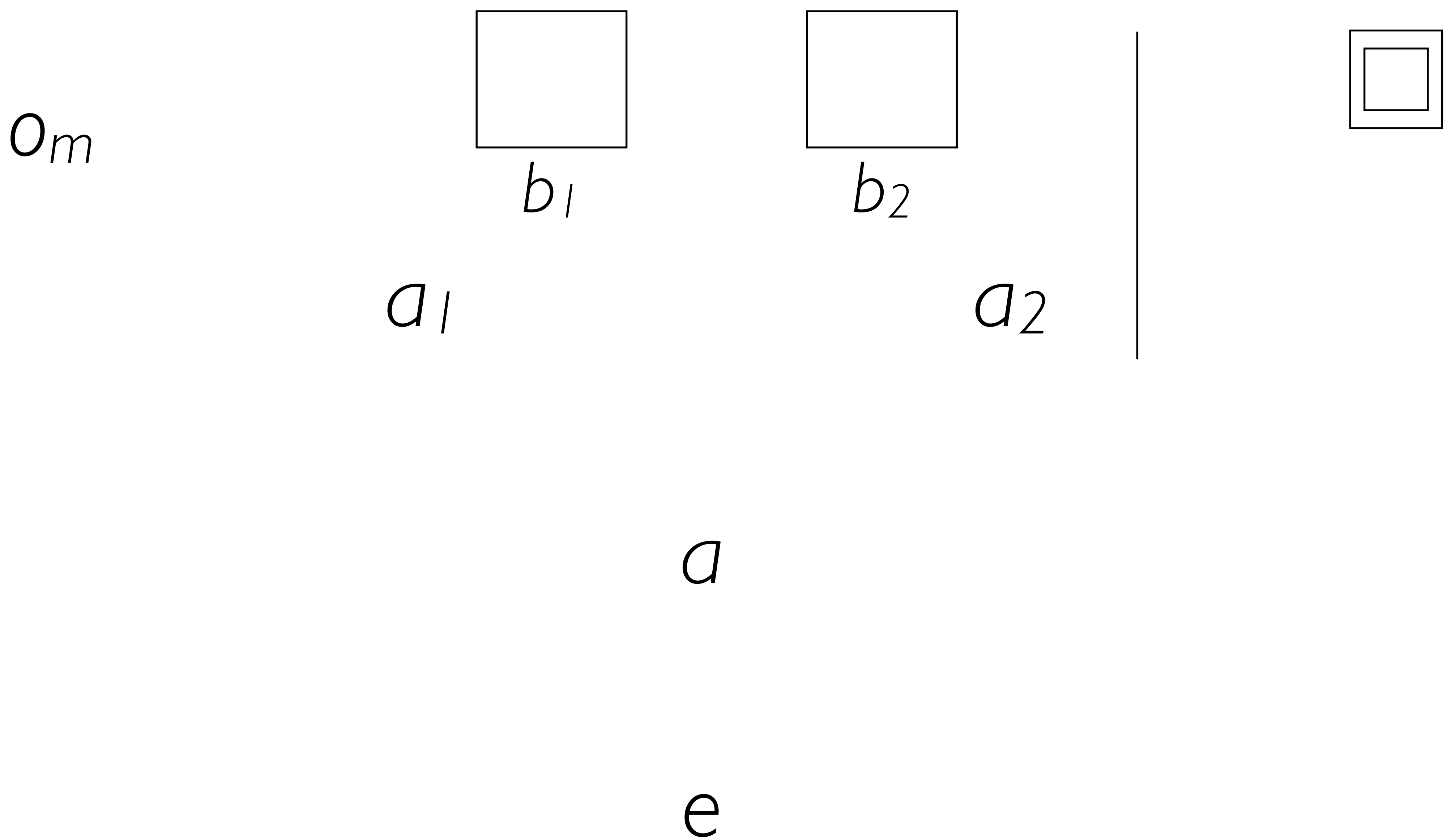
1 Introduction

Interpreting the behavior of other people is indispensable for everyday life. It is something that we do constantly, on a daily basis, and it helps us not only to make sense of human behavior, but also to predict it and—to a certain extent—to control it. How exactly do we manage that? That is not currently known, but many have argued that the ability to ascribe mental states to others and to reason about such mental states is a key component of our capacity to understand human behavior. In particular, all social transactions, from engaging in commerce and negotiating to making jokes and empathizing with other people’s pain or joy, appear to require at least a rudimentary grasp of common-sense psychology (CSP), i.e., a large body of truisms such as the following: When an agent a (1) wants to achieve a certain state of affairs p, and (2) believes that some action c can bring about p, and (3) a knows how to carry out c; then, ceteris paribus,[1] a will carry out c; when a sees that p, a knows that p; when a fears that p and a discovers that p is the case, a is disappointed; and so on.

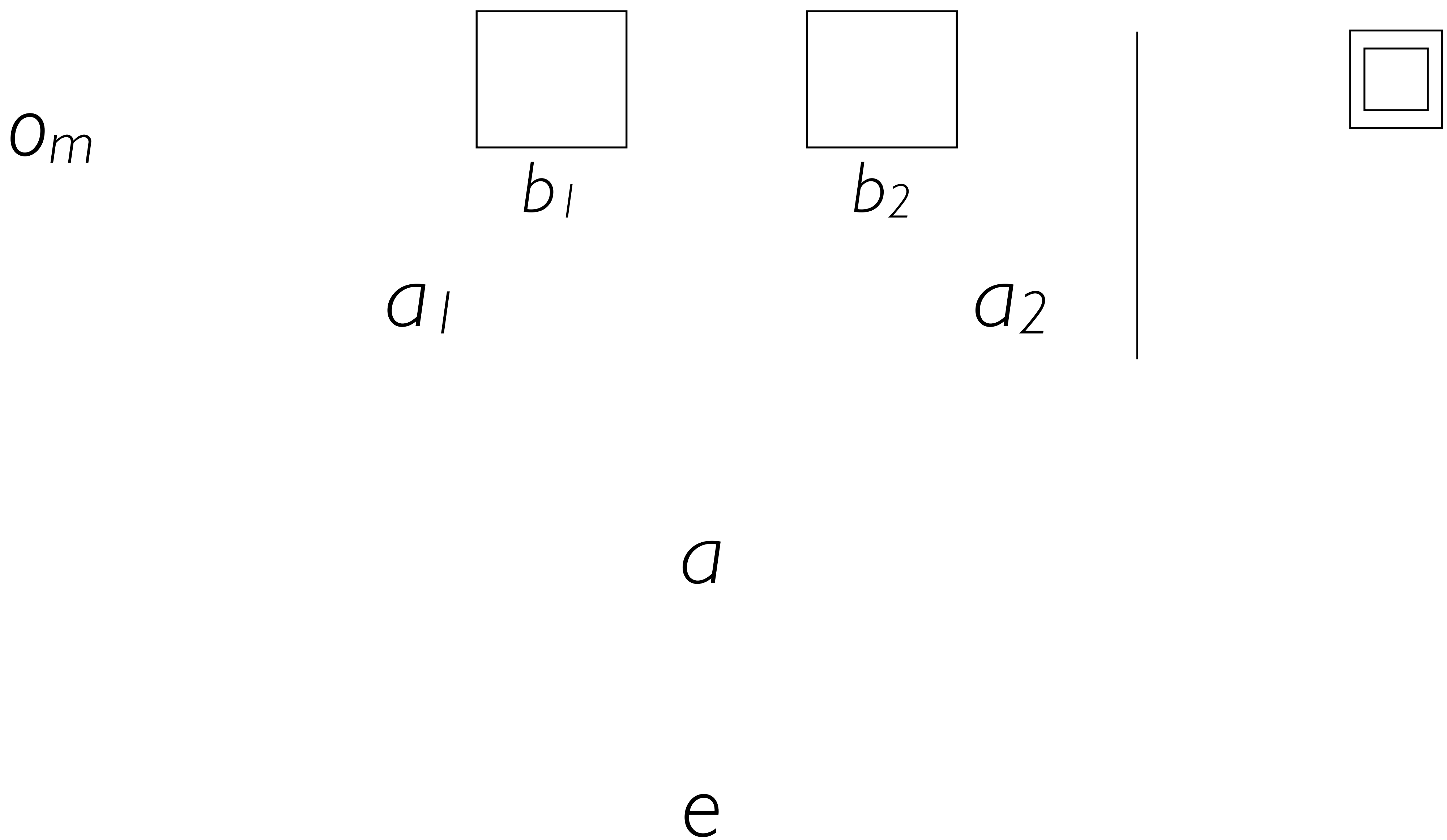
Artificial agents without a mastery of CSP would be severely handicapped in their interactions with humans. This could present problems not only for artificial agents trying to interpret human behavior, but also for artificial agents trying to interpret the behavior of one another. When a system exhibits a complex but rational behavior, and detailed knowledge of its internal structure is not

[1] Assuming that a is able to carry out c, that a has no conflicting desires that override his goal that p; and so on.

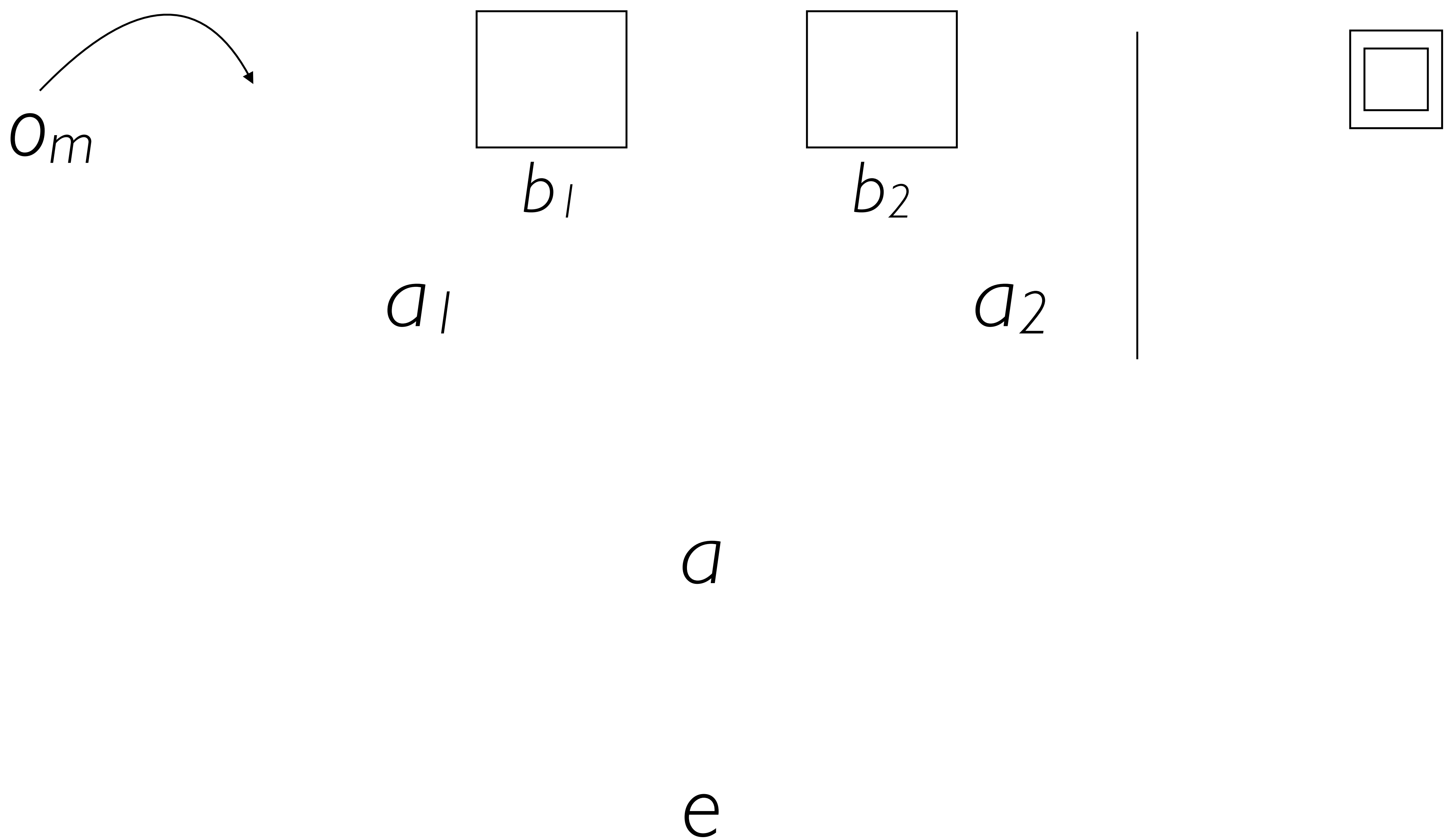
Framework for FBT$^{1}{}_{2}$



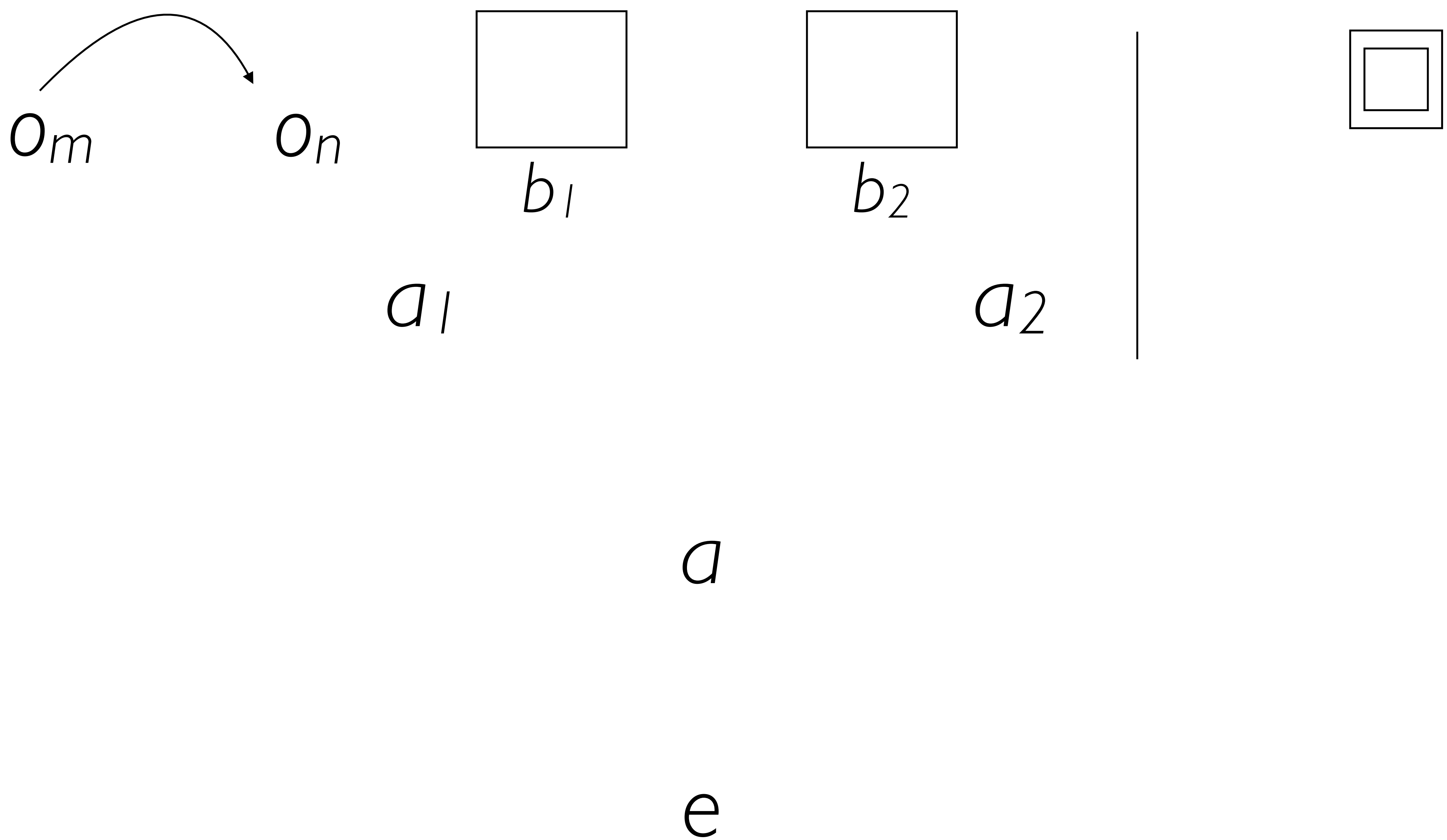
Framework for FBT^1_2
(seven timepoints)



Framework for FBT^1_2
(seven timepoints)

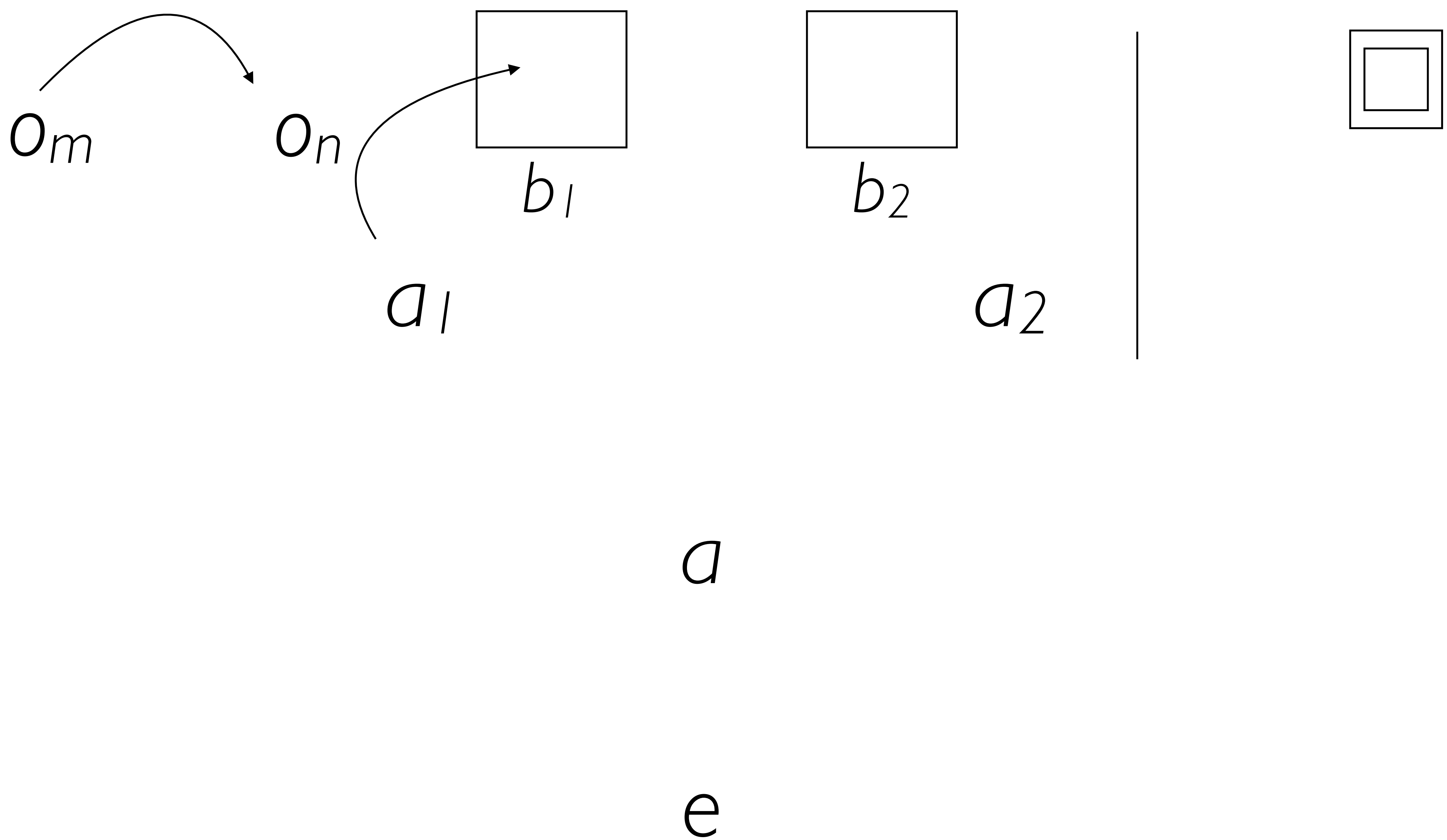


Framework for FBT^1_2
(seven timepoints)



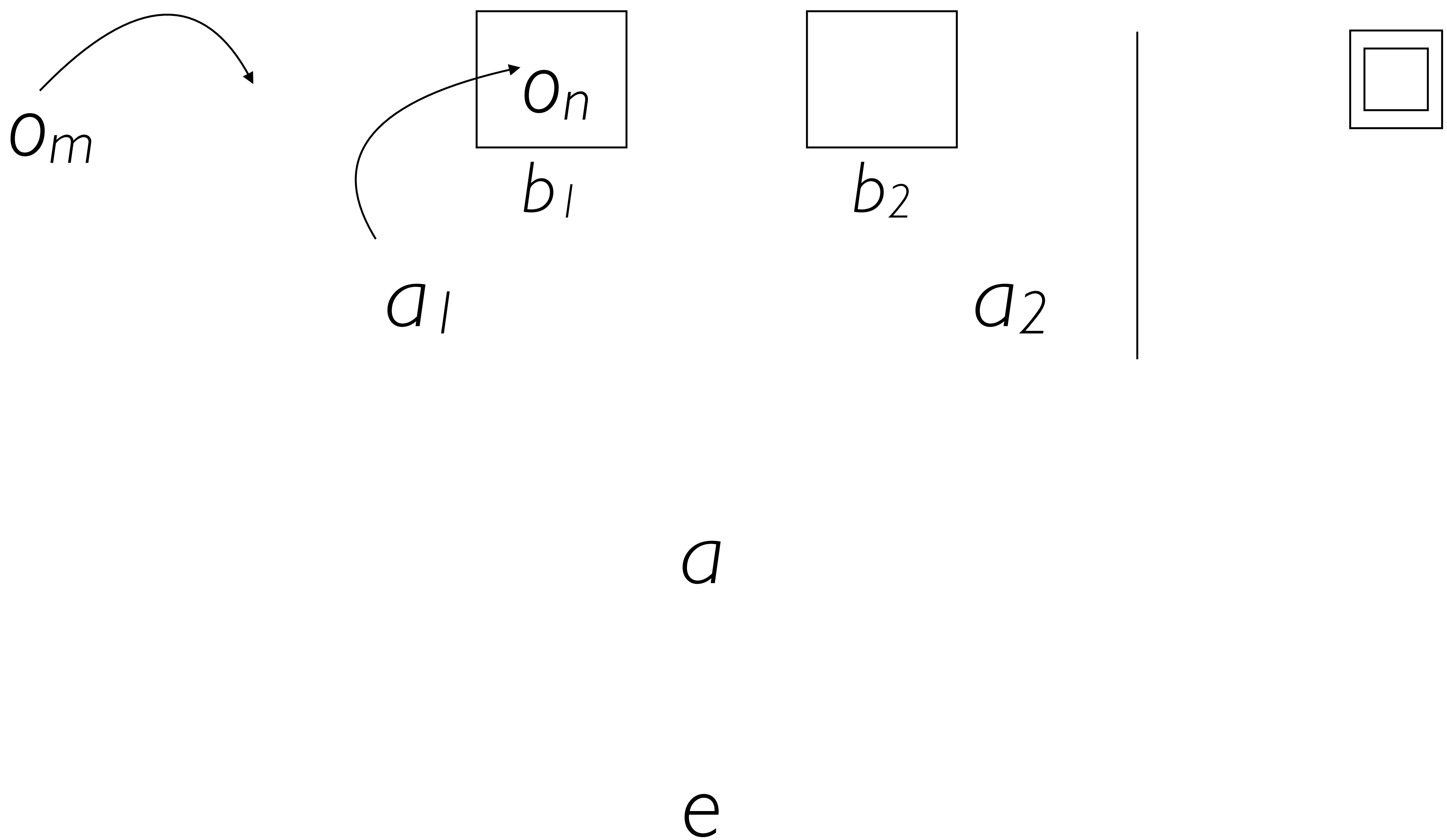
Framework for FBT$^{1}_{2}$

(seven timepoints)



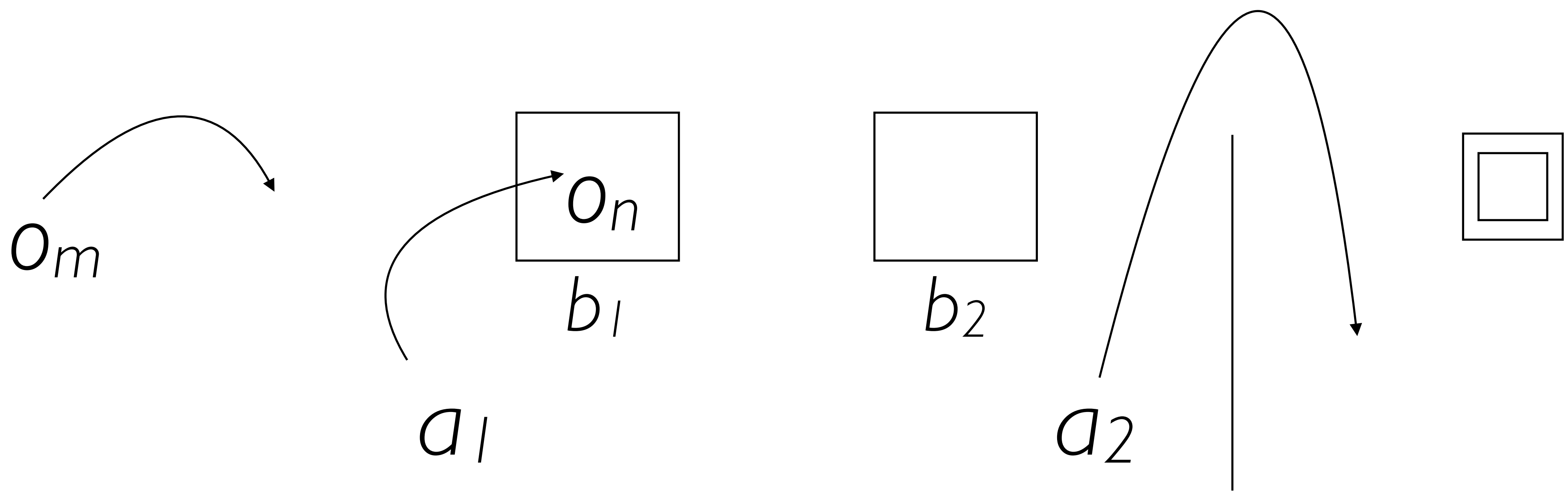
Framework for $FBT^1{}_2$

(seven timepoints)



Framework for FBT$^1{}_2$

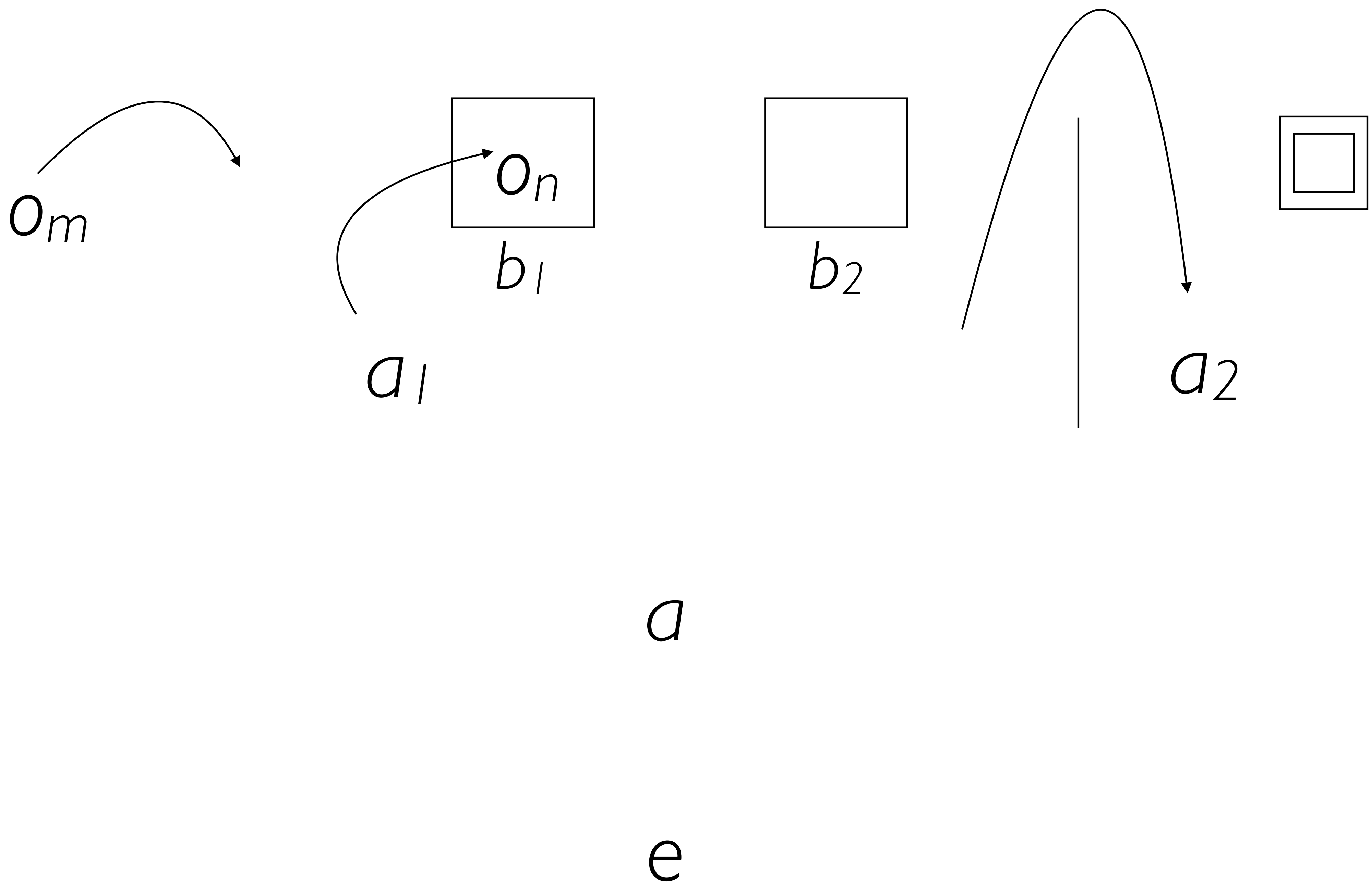
(seven timepoints)



a

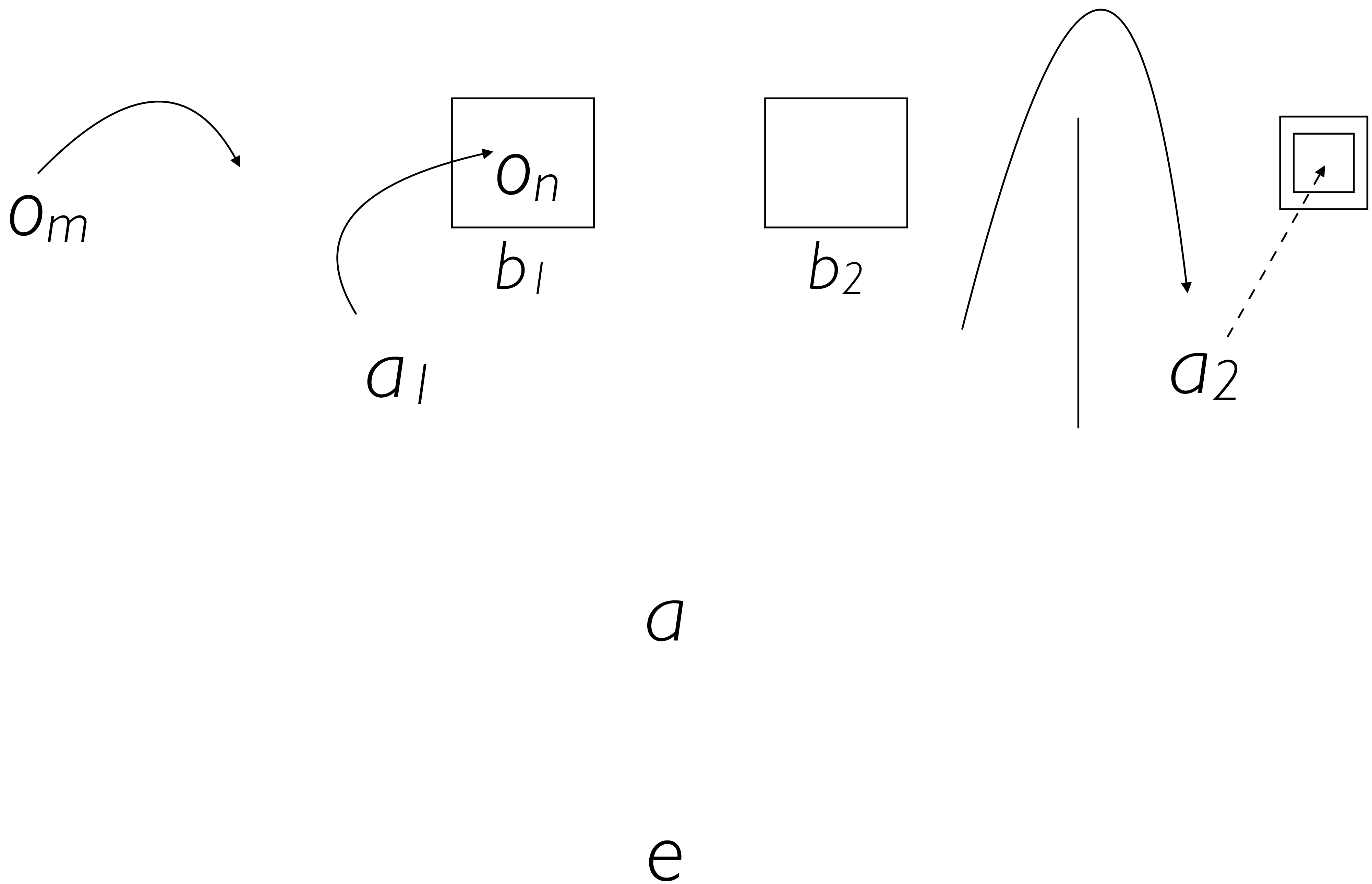
e

Framework for $FBT^1{}_2$
(seven timepoints)

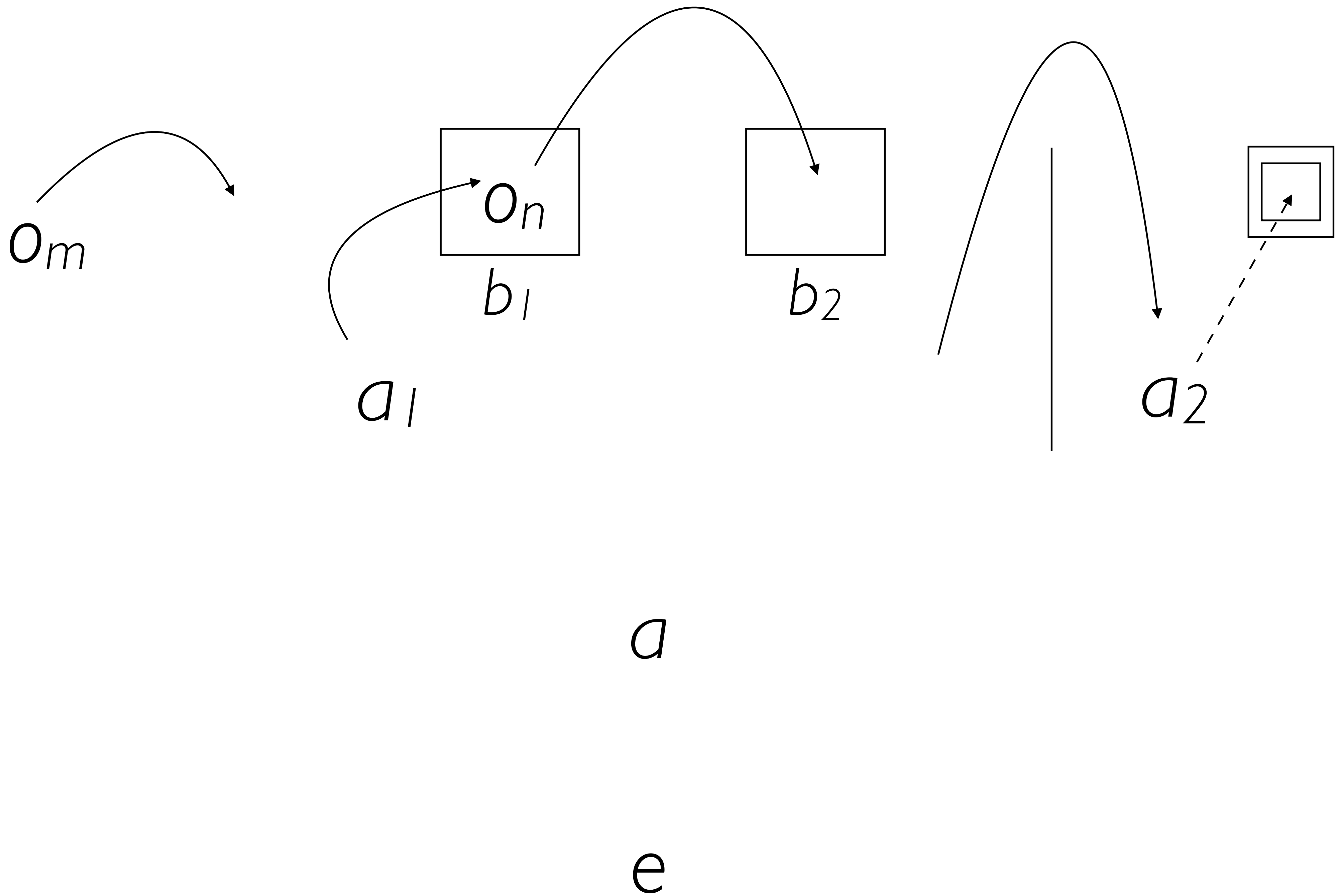


Framework for FBT$^1{}_2$

(seven timepoints)

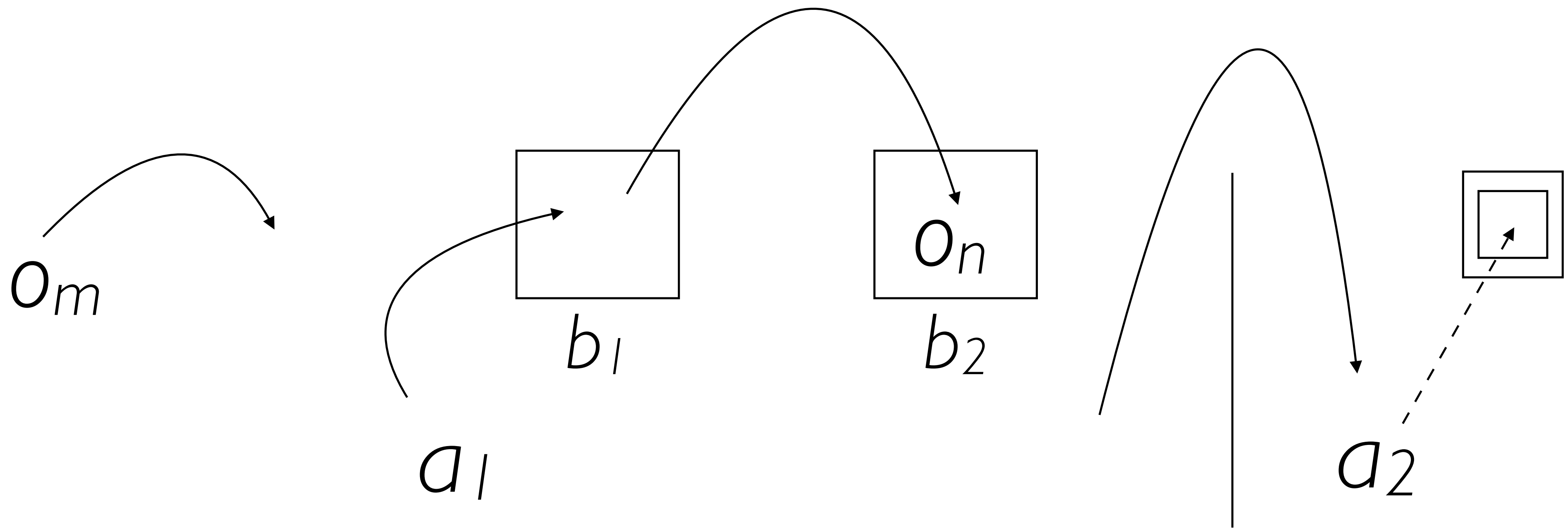


Framework for FBT^1_2
(seven timepoints)



Framework for FBT$^1{}_2$

(seven timepoints)

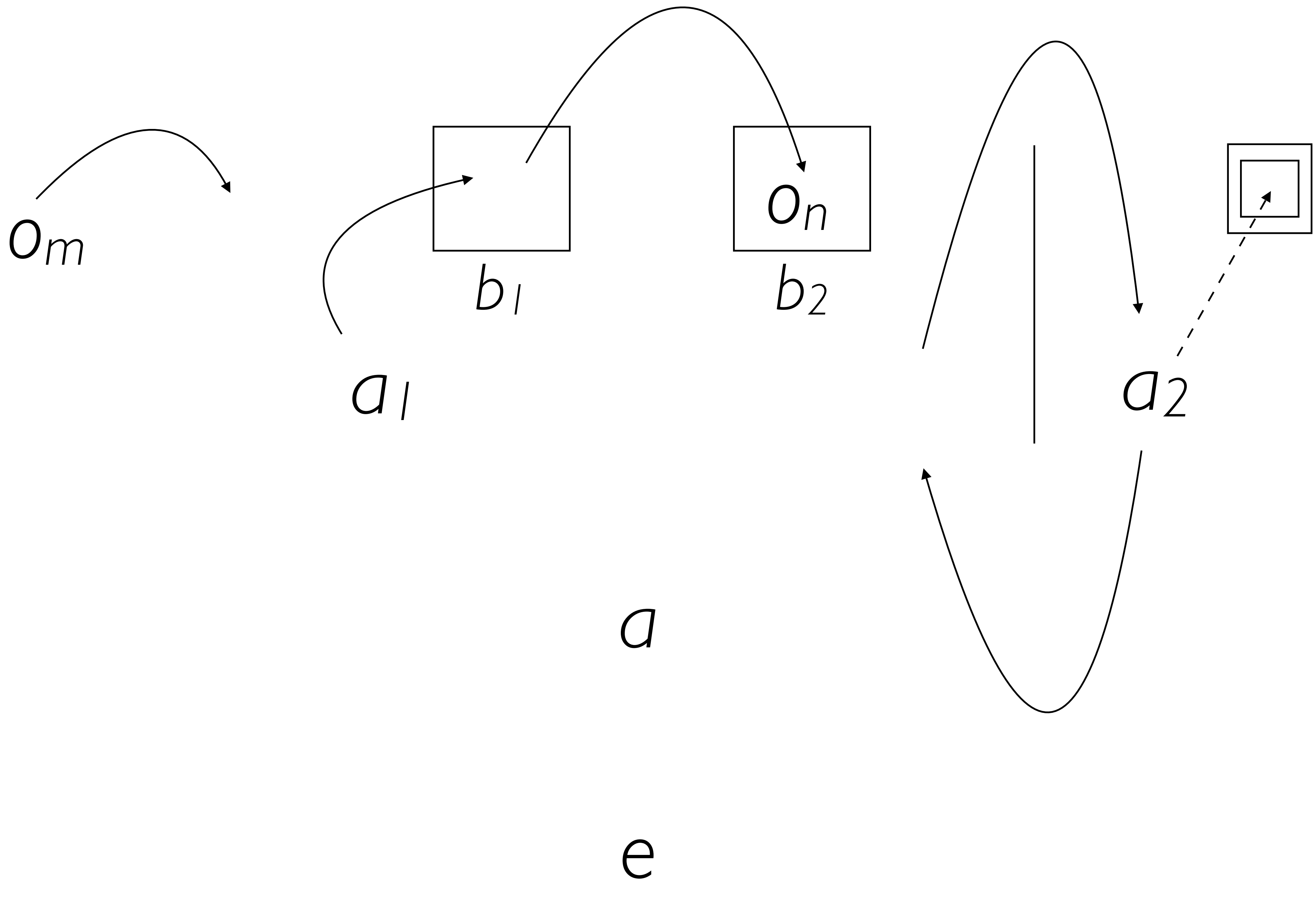


a

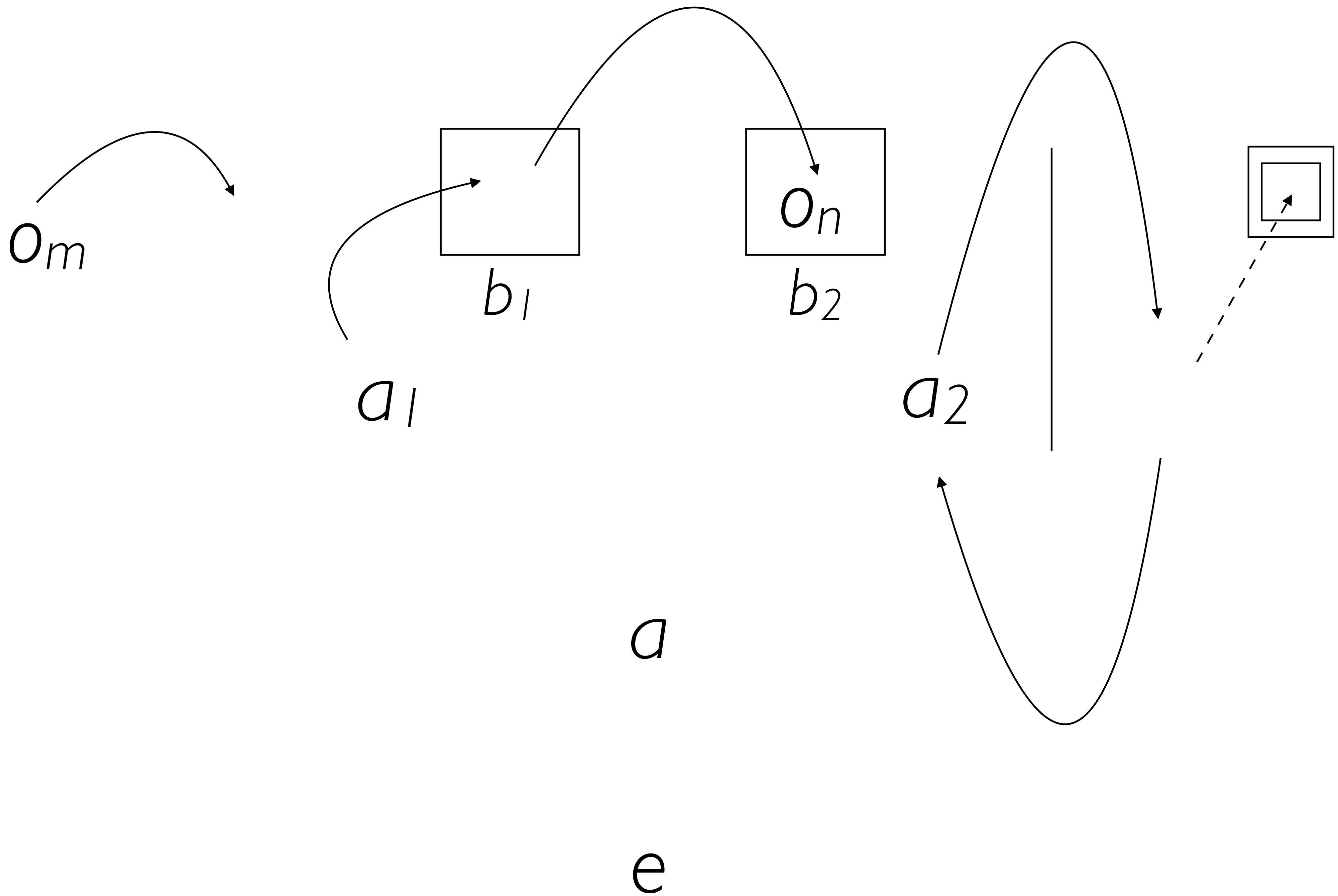
e

Framework for $FBT^1{}_2$

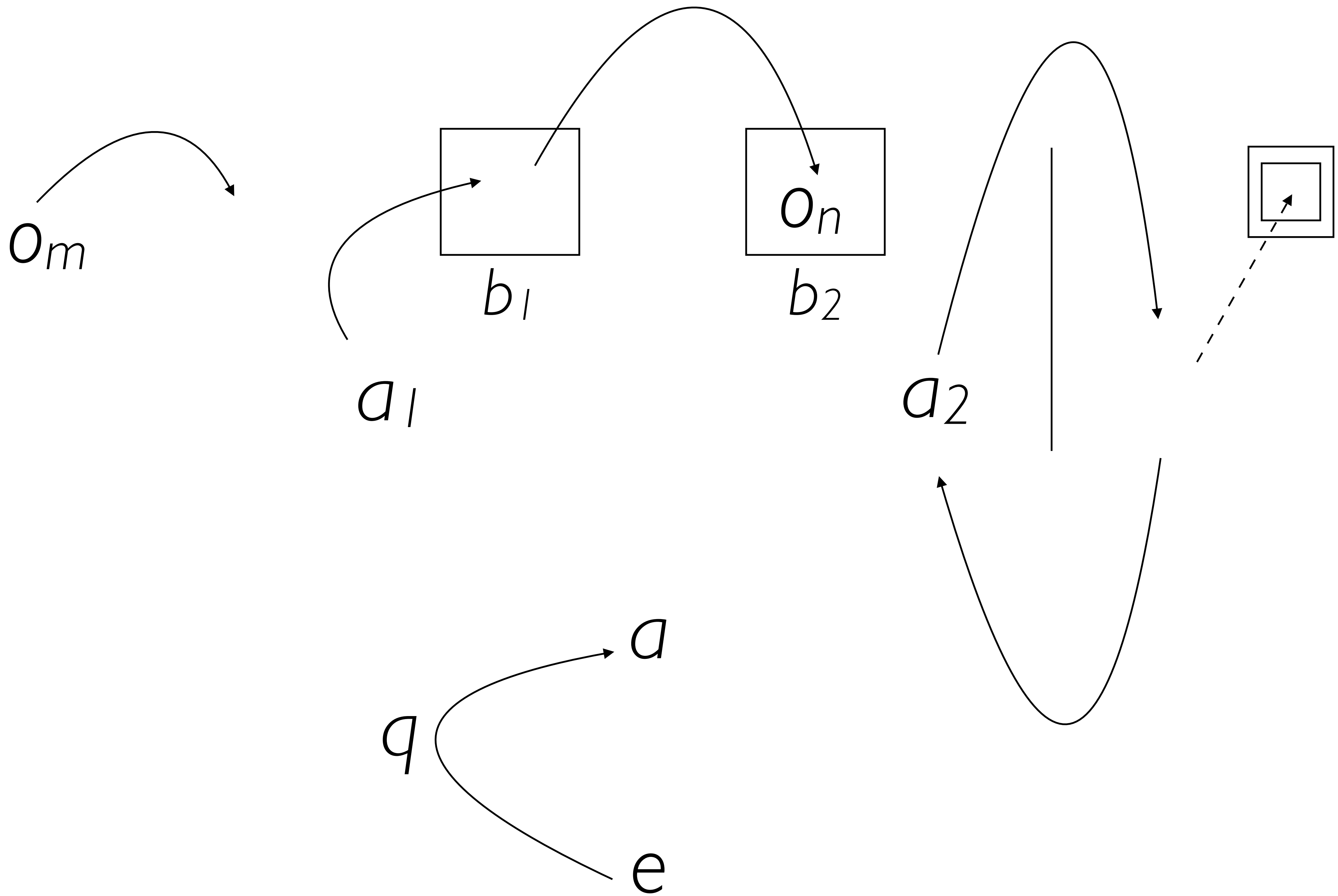
(seven timepoints)



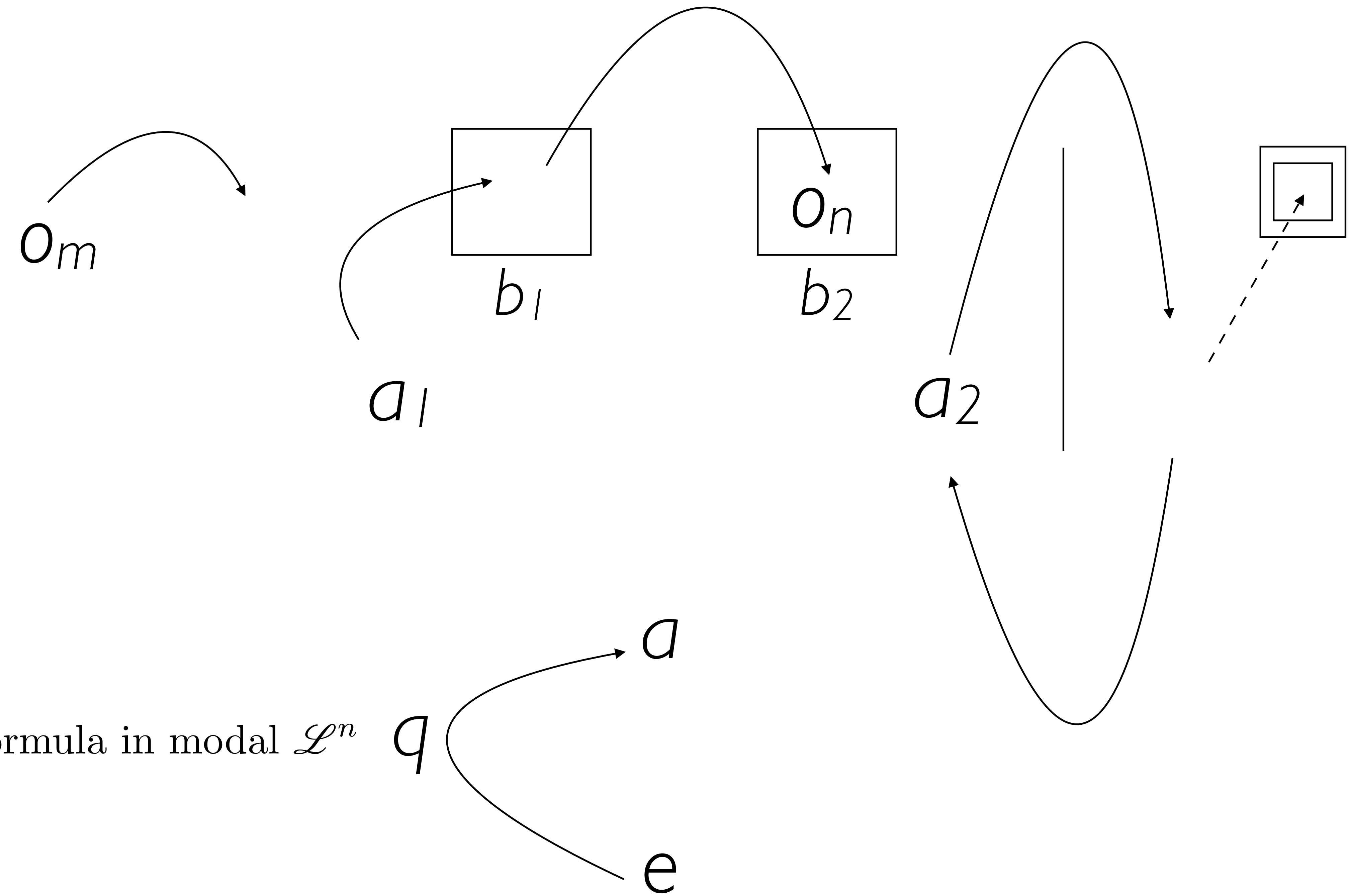
Framework for $FBT^1{}_2$
(seven timepoints)



Framework for $FBT^1{}_2$
(seven timepoints)

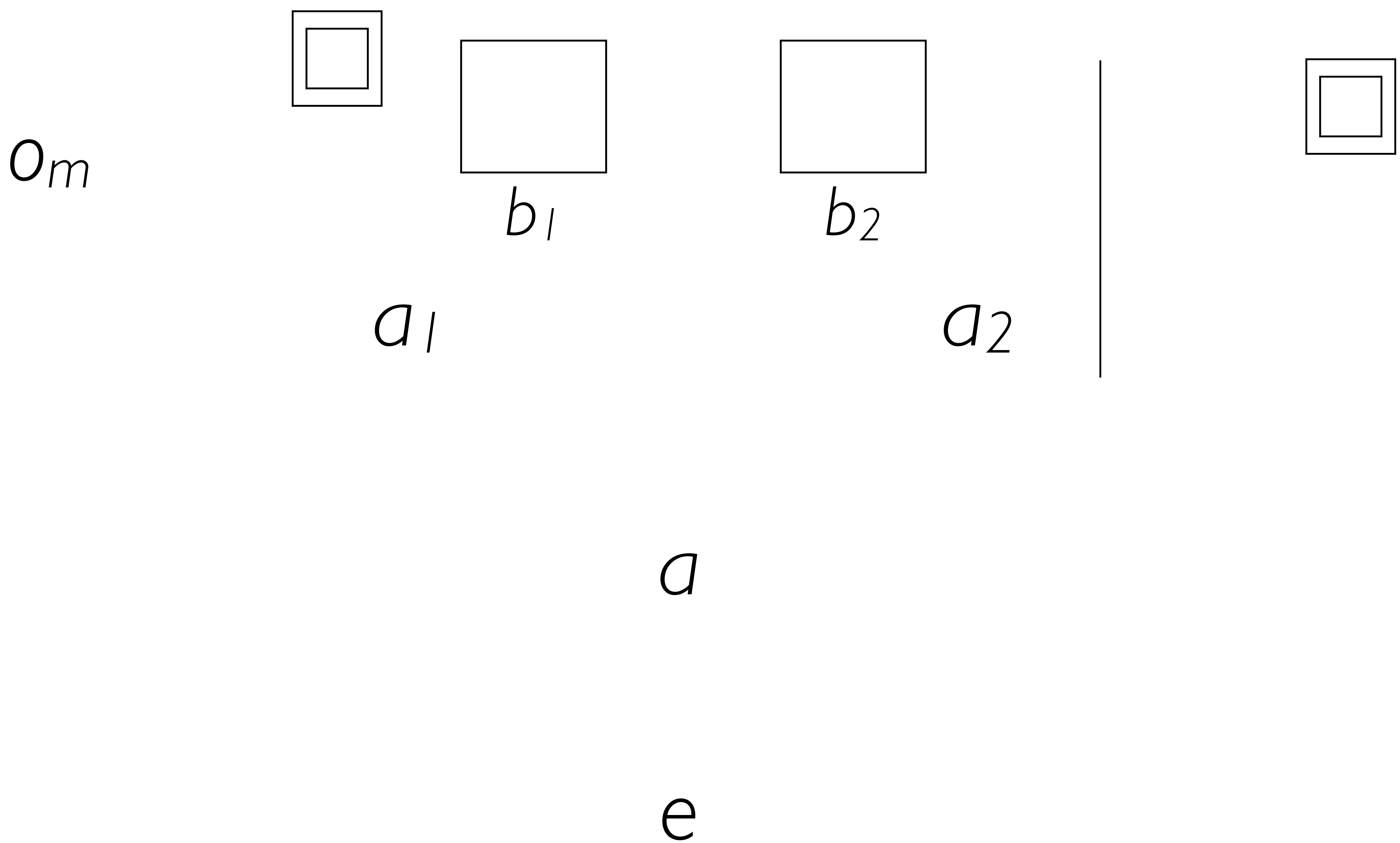


Framework for $FBT^1{}_2$
(seven timepoints)

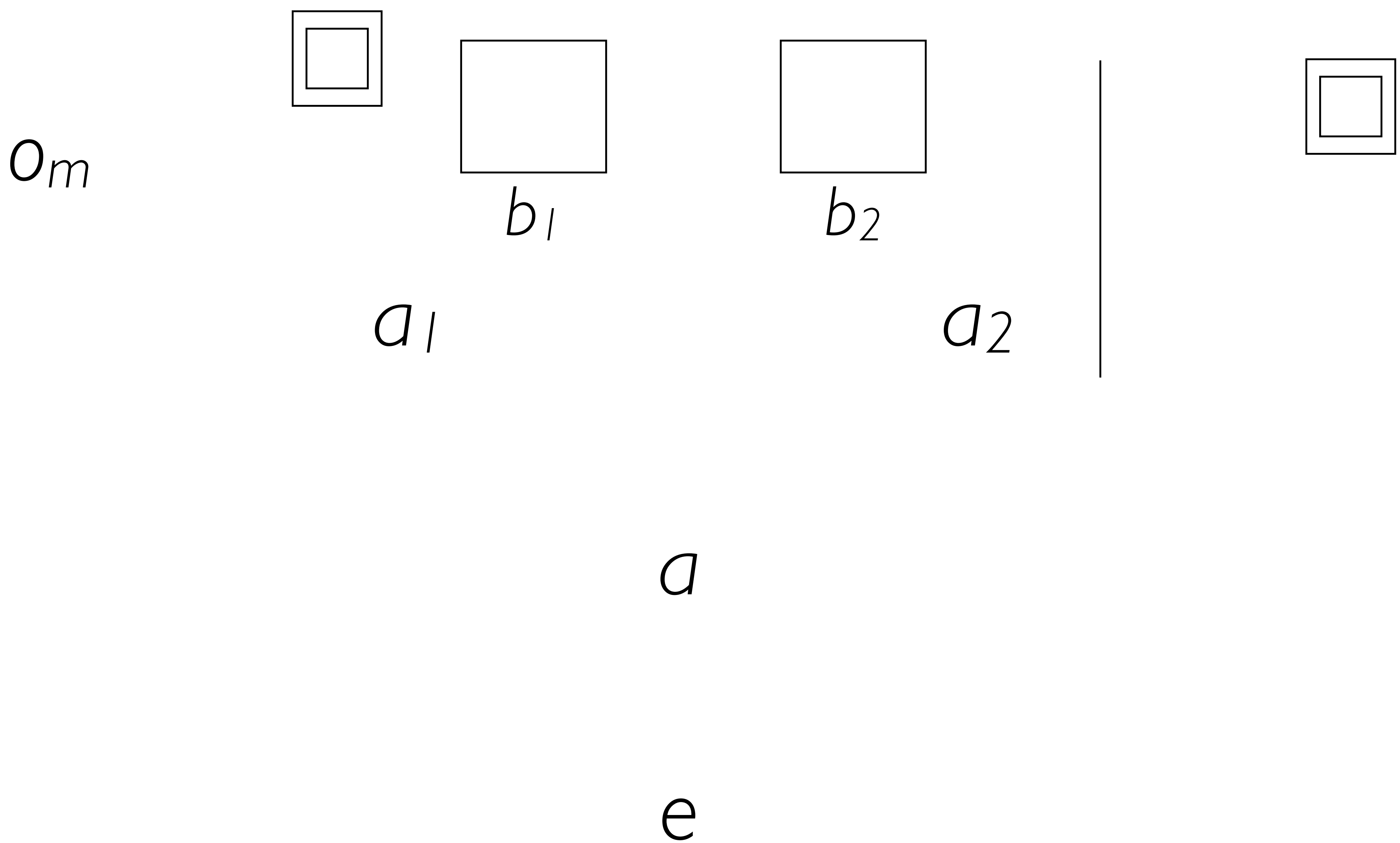


q a formula in modal $\mathscr{L}^n$

Framework for FBT^1_3

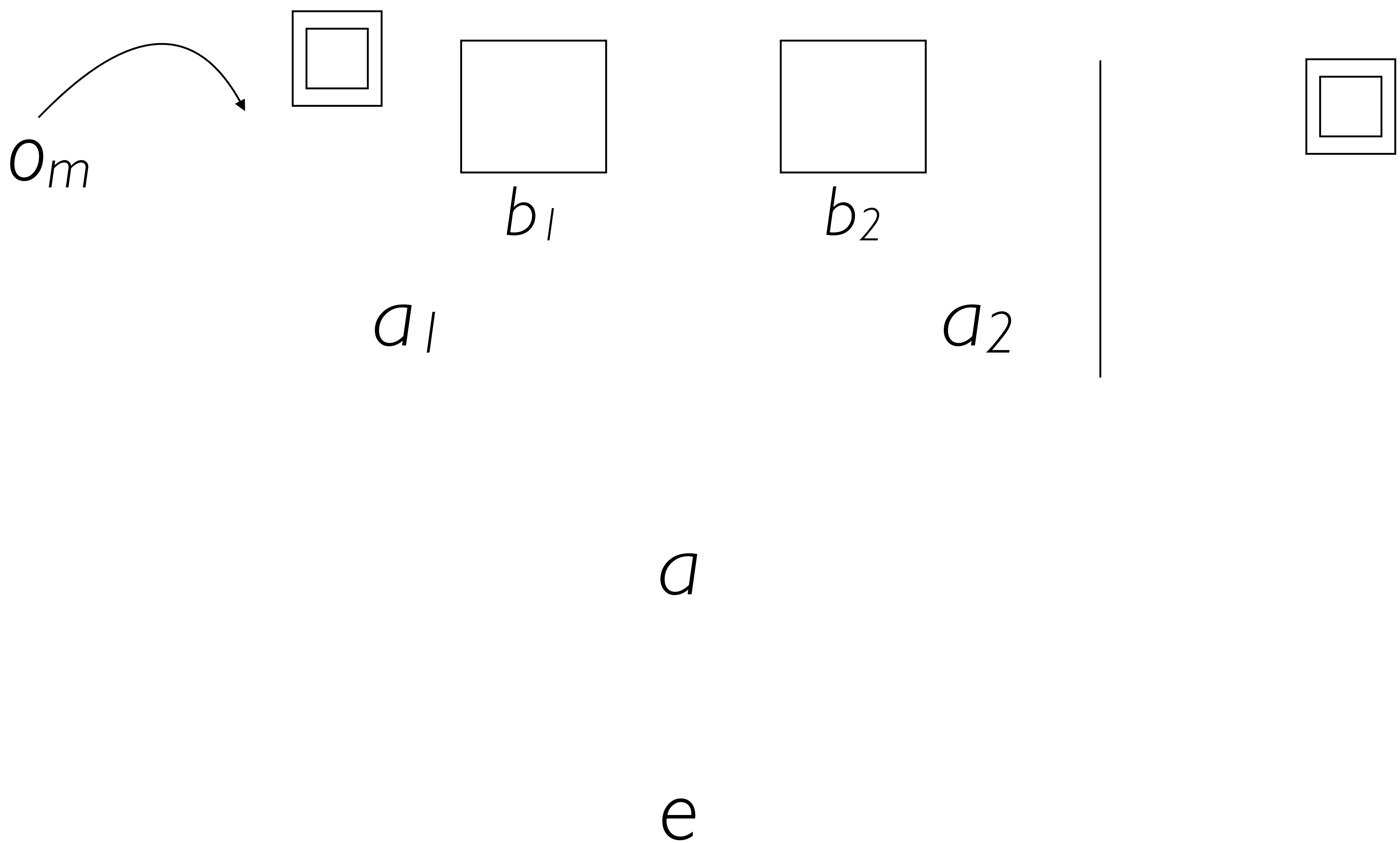


Framework for FBT^1_3
(eight timepoints)

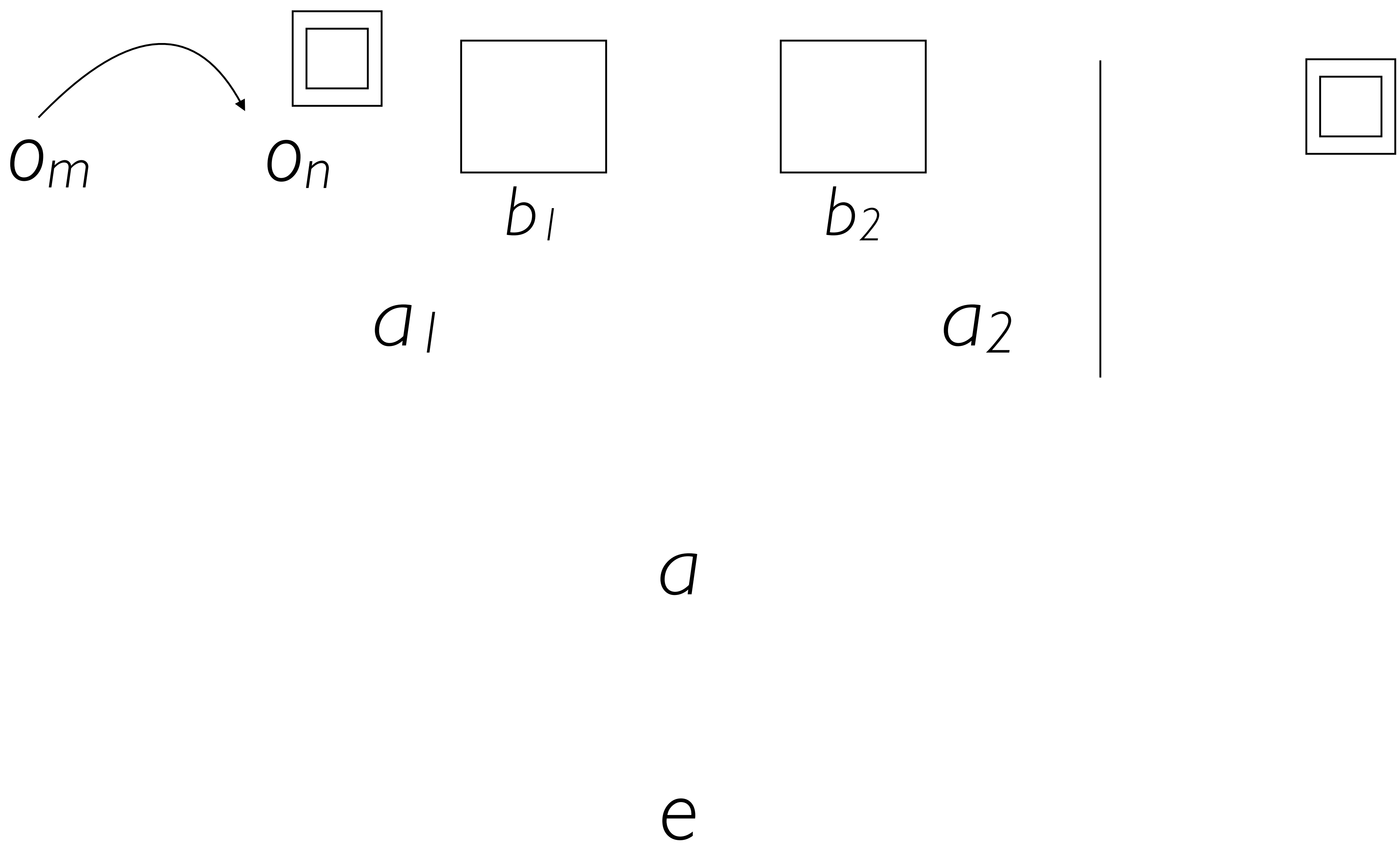


Framework for $FBT^1{}_3$

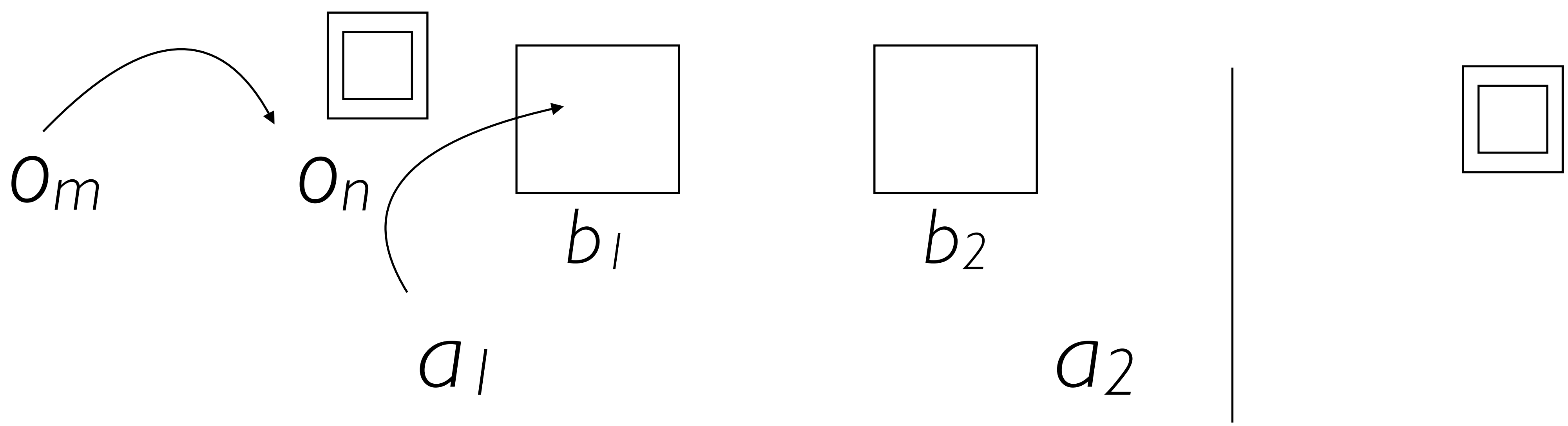
(eight timepoints)



Framework for FBT^1_3
(eight timepoints)



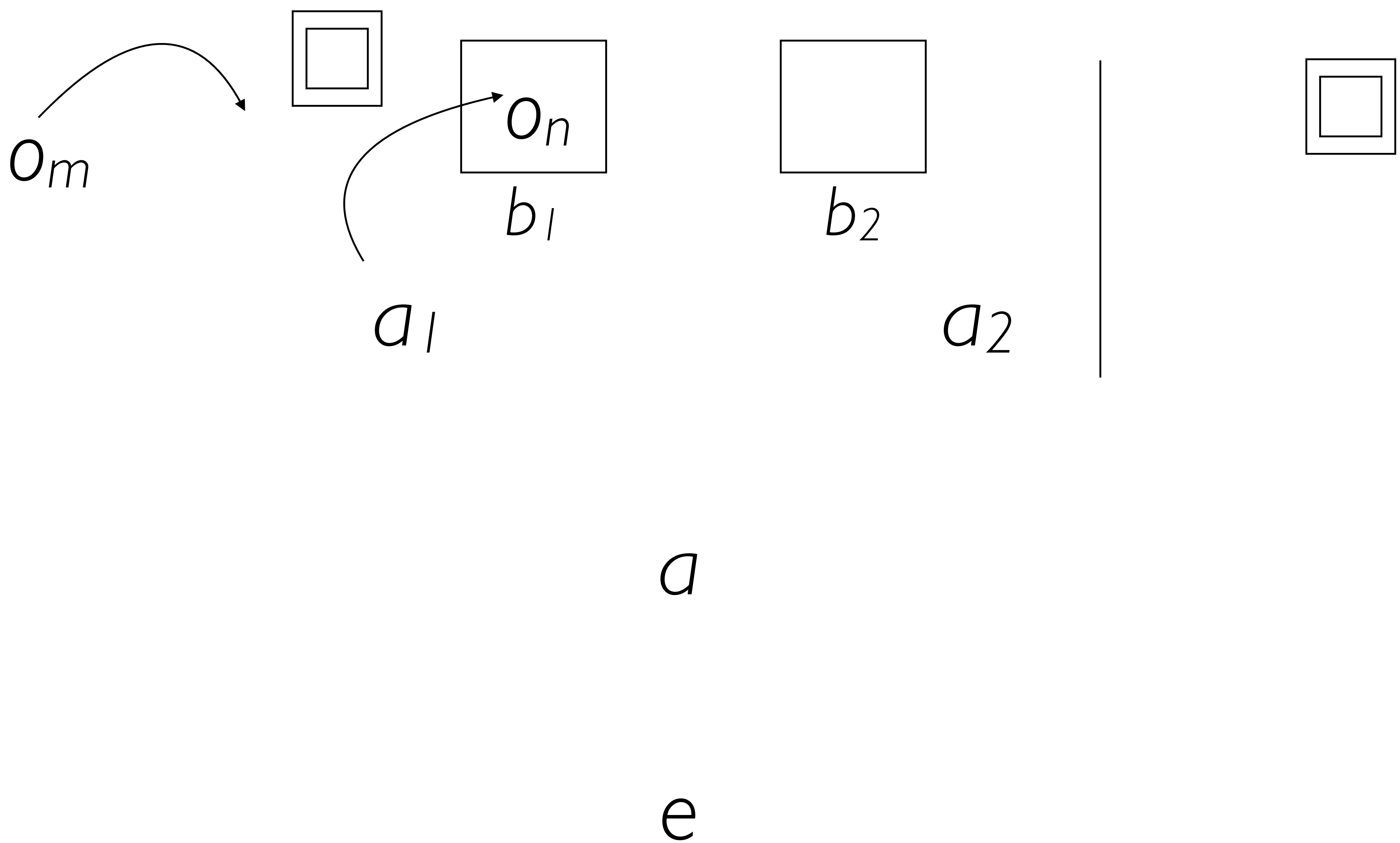
Framework for FBT^1_3
(eight timepoints)



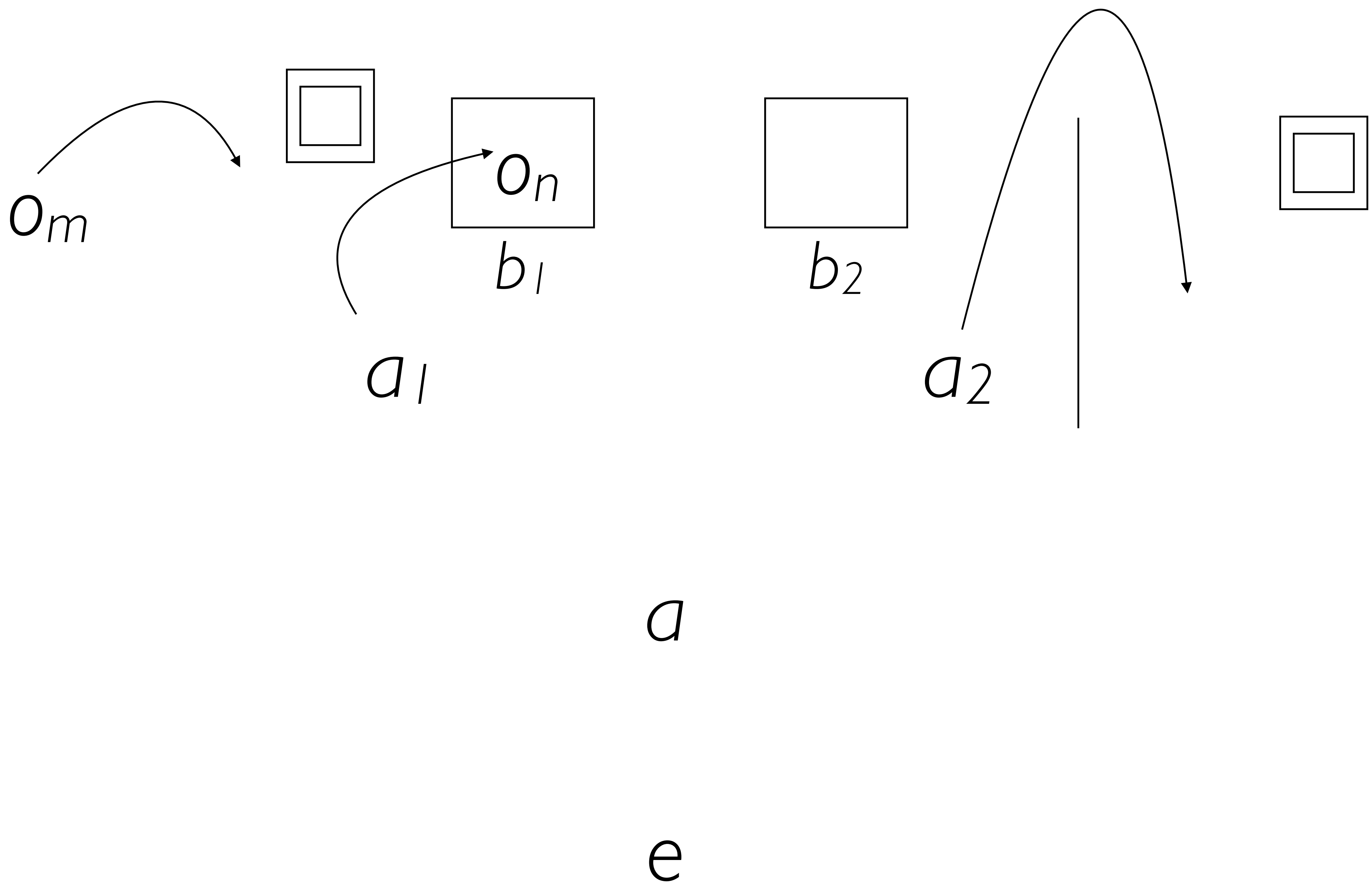
a

e

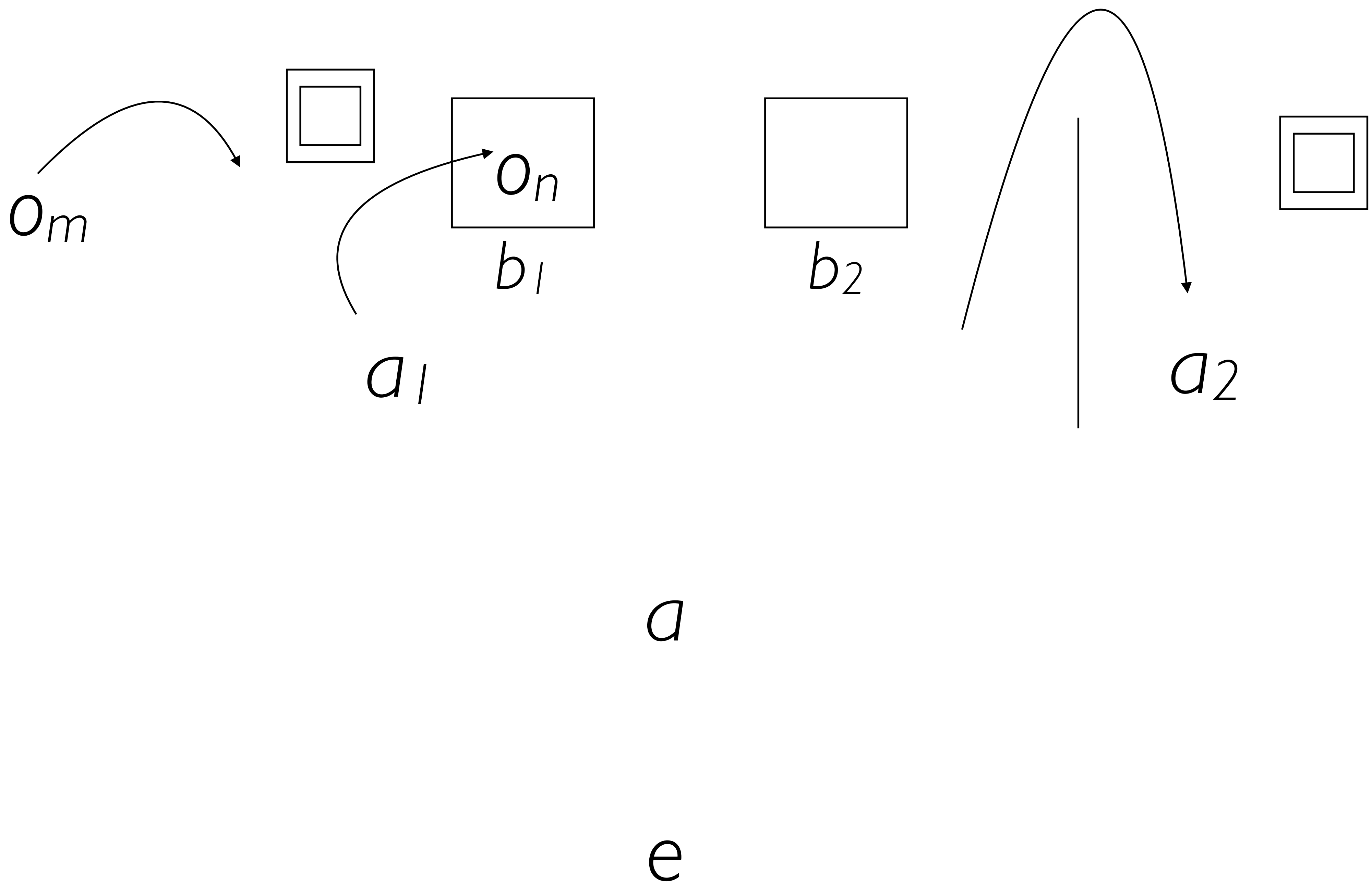
Framework for $FBT^1{}_3$
(eight timepoints)



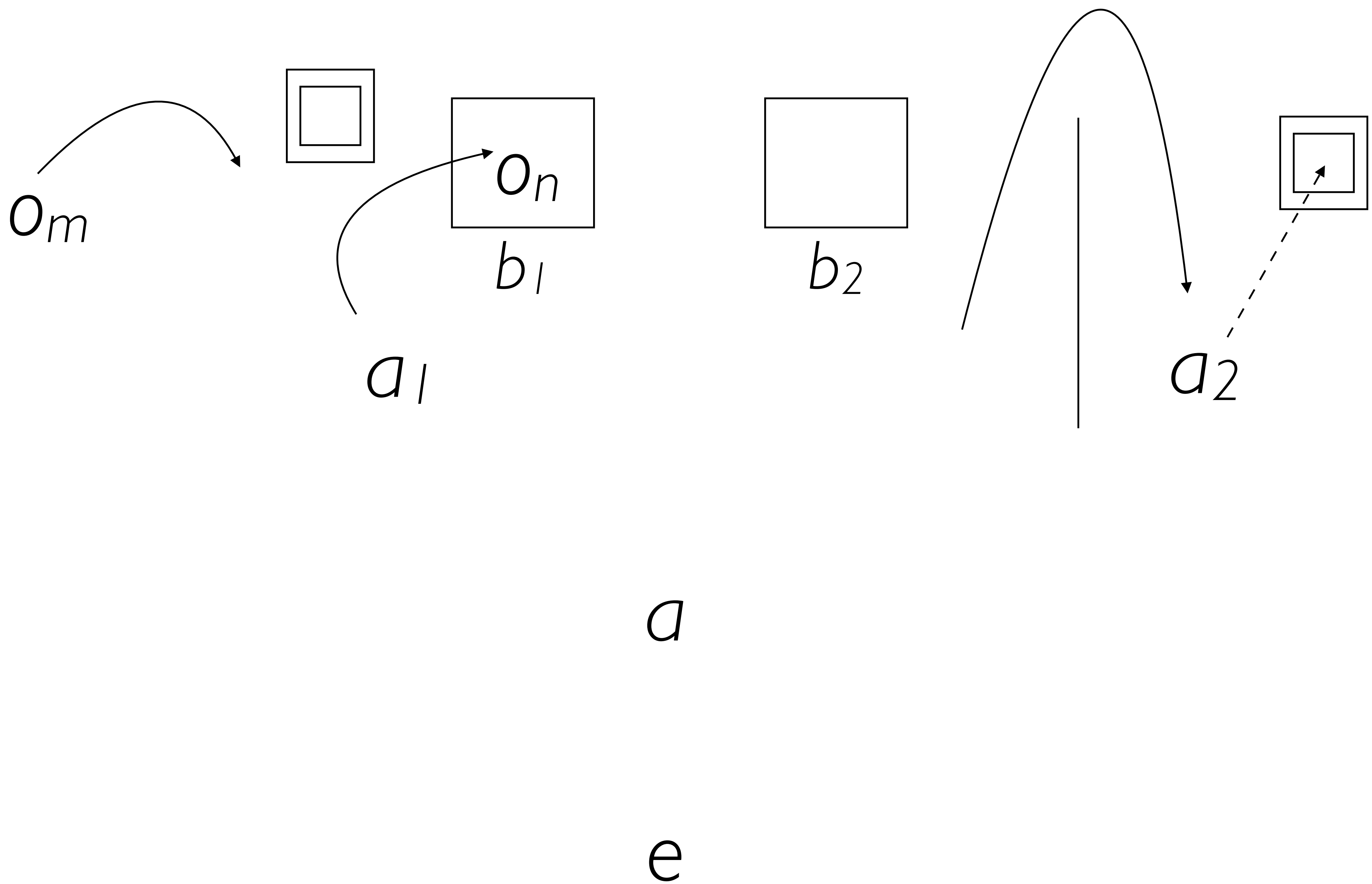
Framework for FBT$^{1}_{3}$
(eight timepoints)



Framework for $FBT^1{}_3$
(eight timepoints)

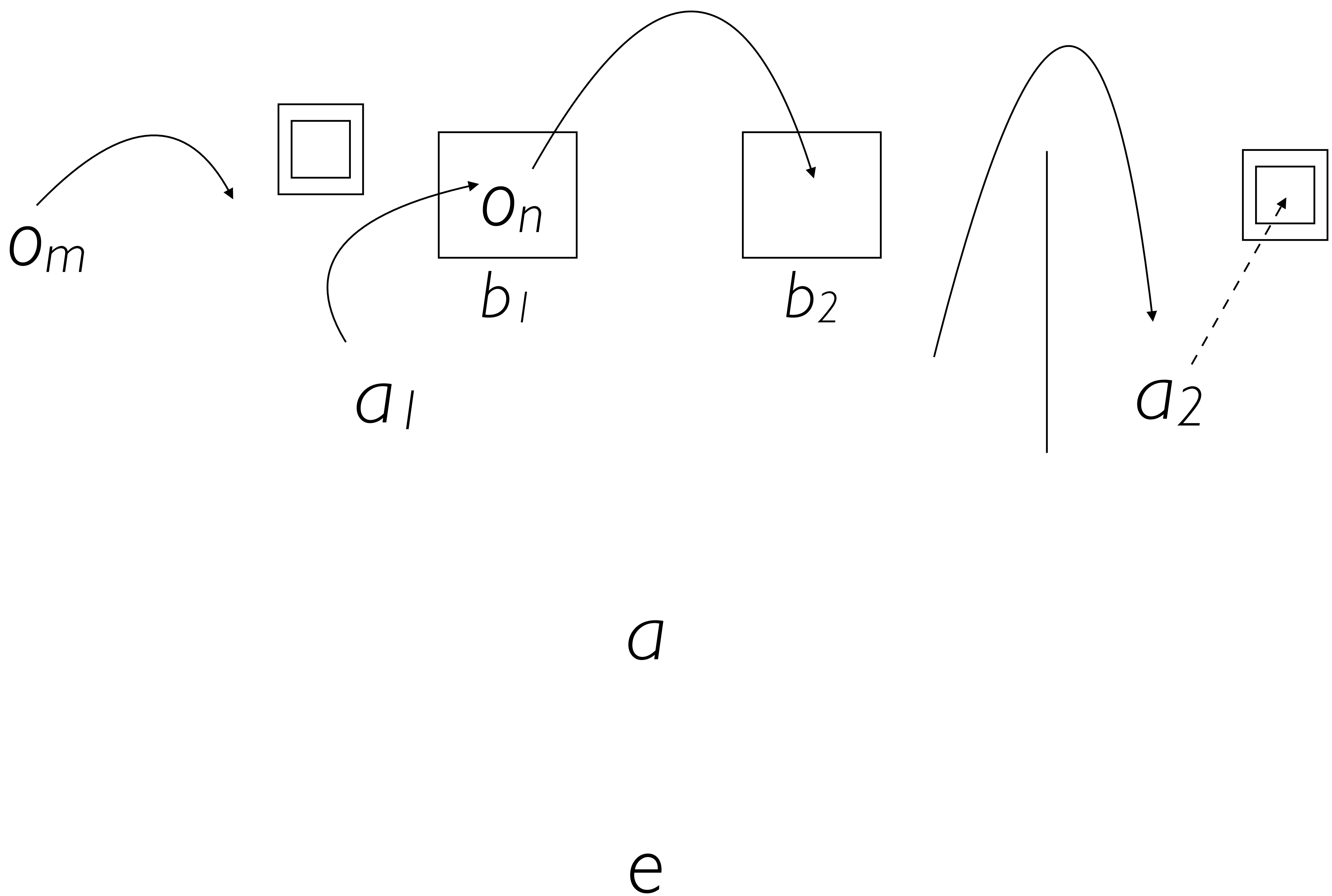


Framework for FBT^1_3
(eight timepoints)



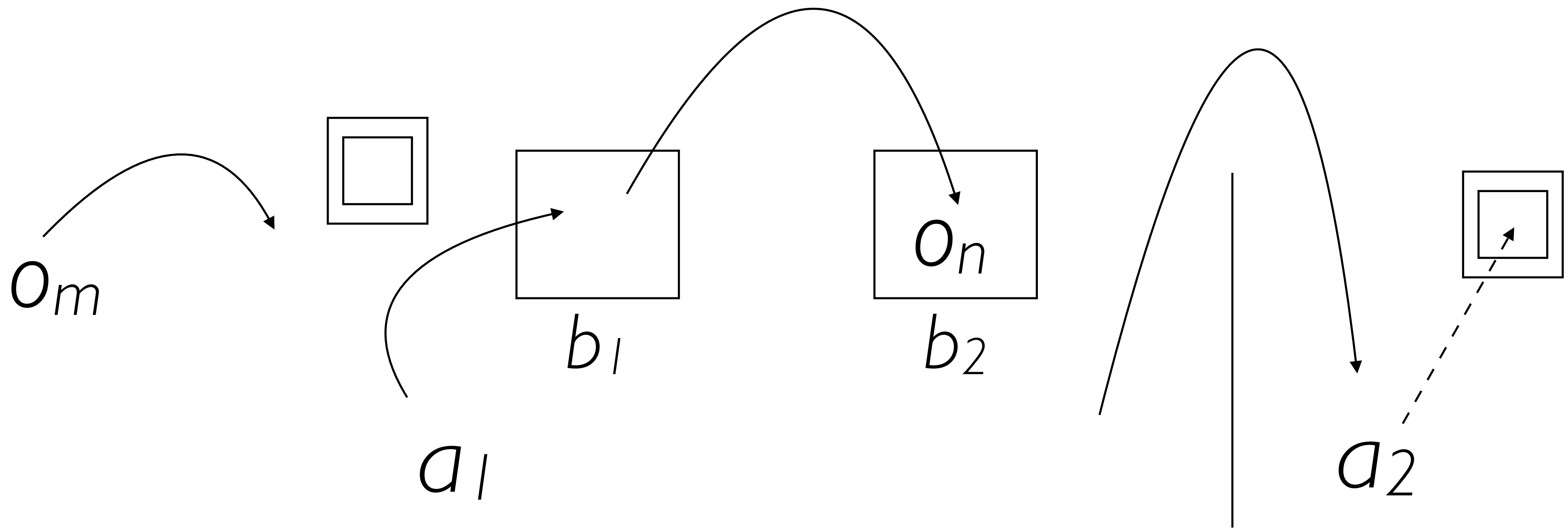
Framework for $FBT^1{}_3$

(eight timepoints)



Framework for $FBT^{1}{}_{3}$

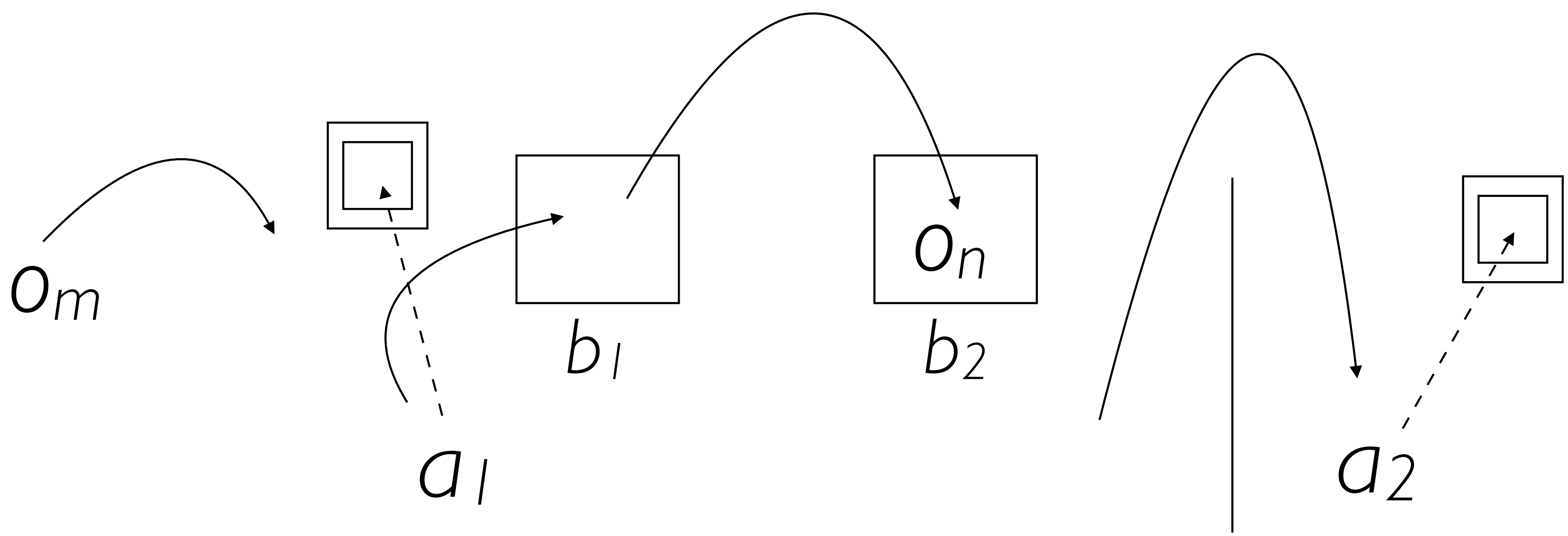
(eight timepoints)



a

e

Framework for FBT$^{1}_{3}$
(eight timepoints)

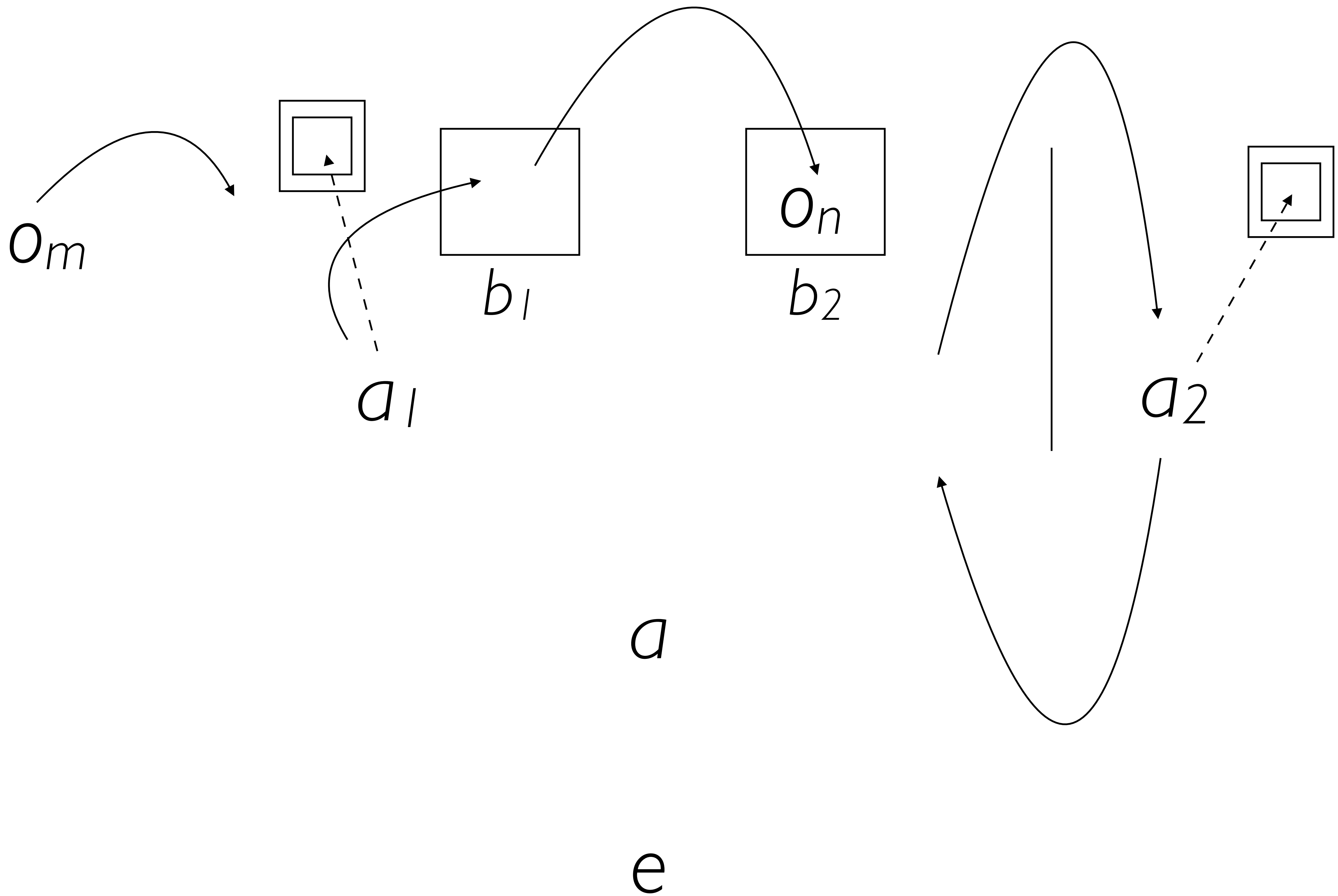


a

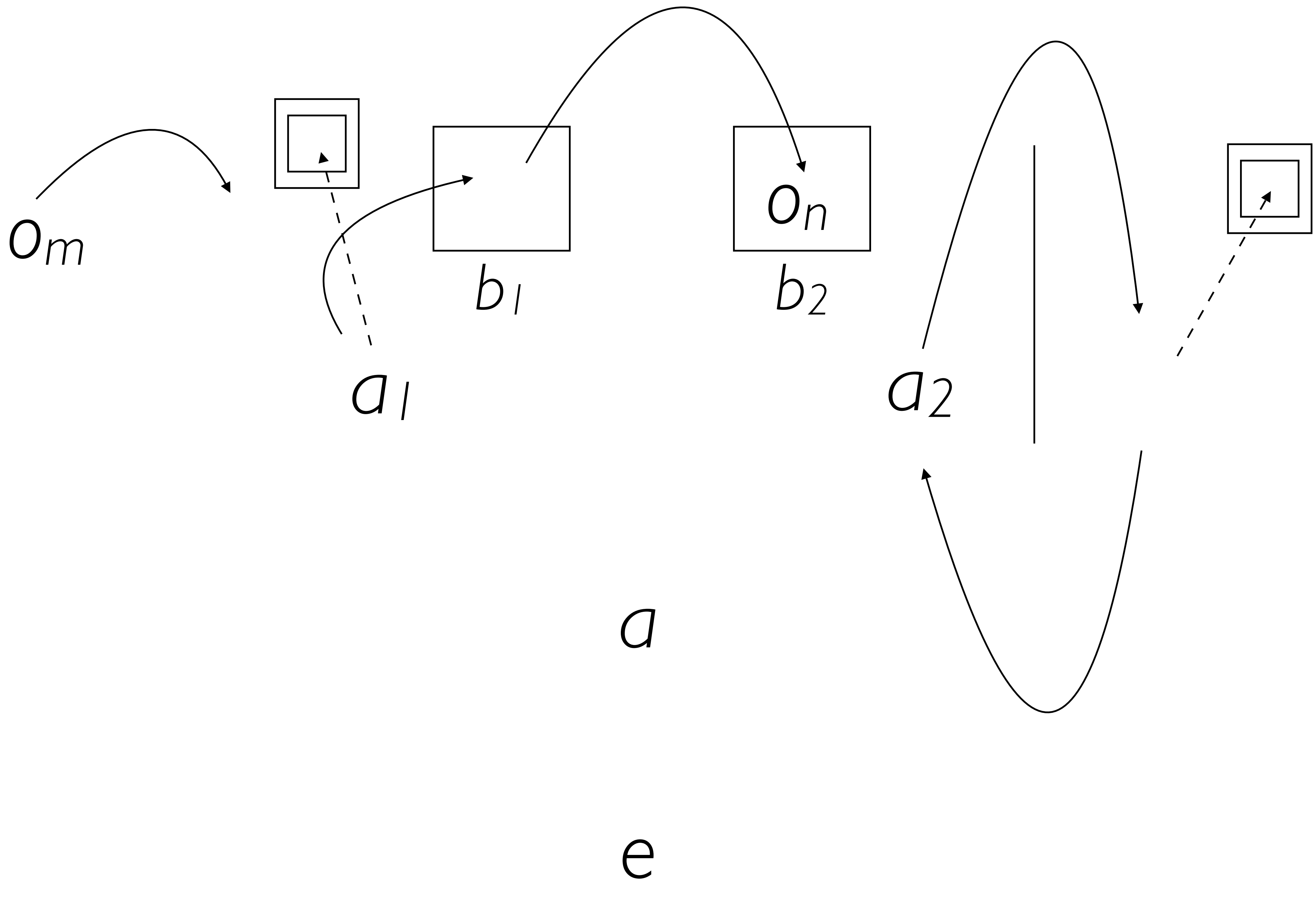
e

Framework for FBT^{1}_{3}

(eight timepoints)

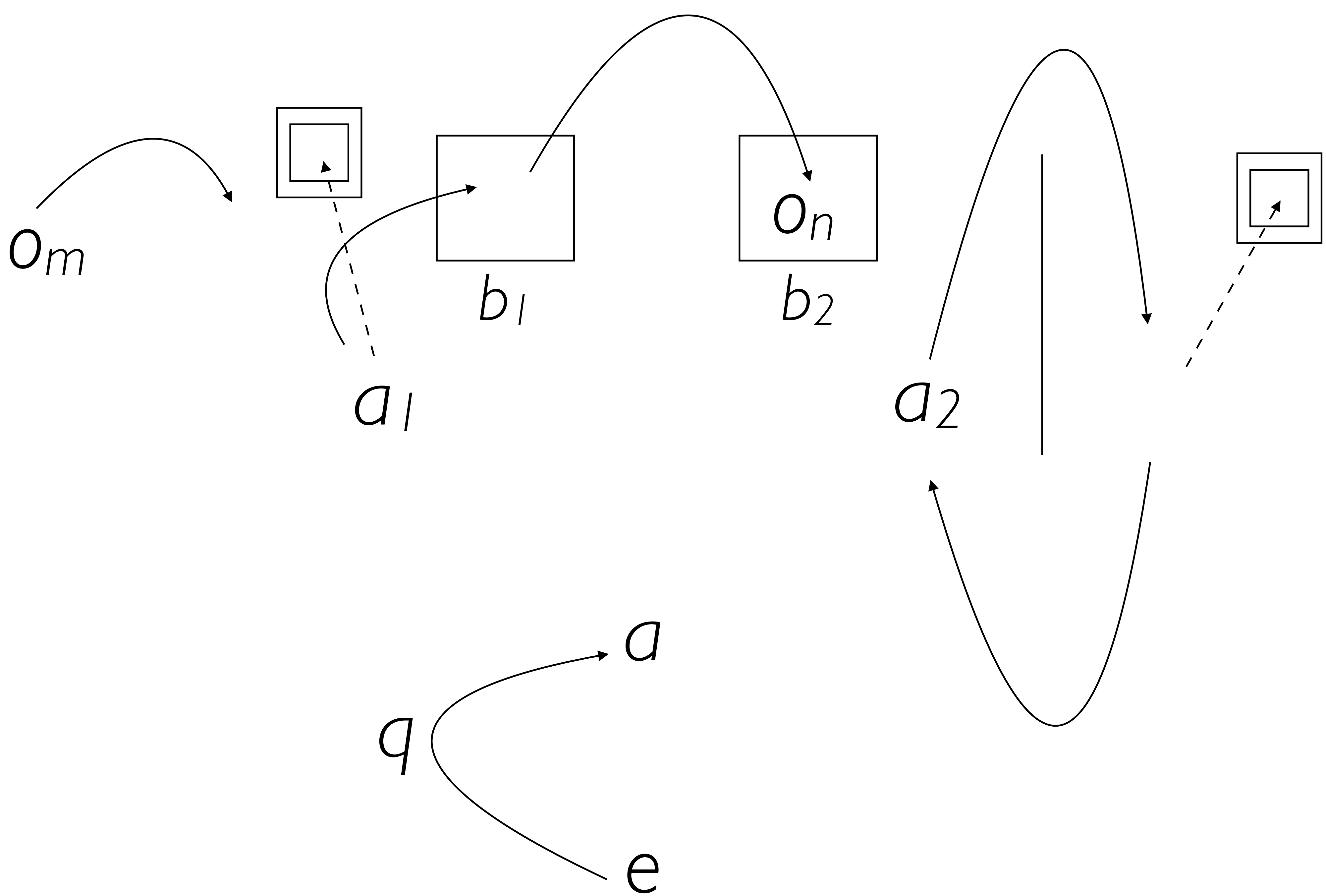


Framework for FBT^1_3
(eight timepoints)

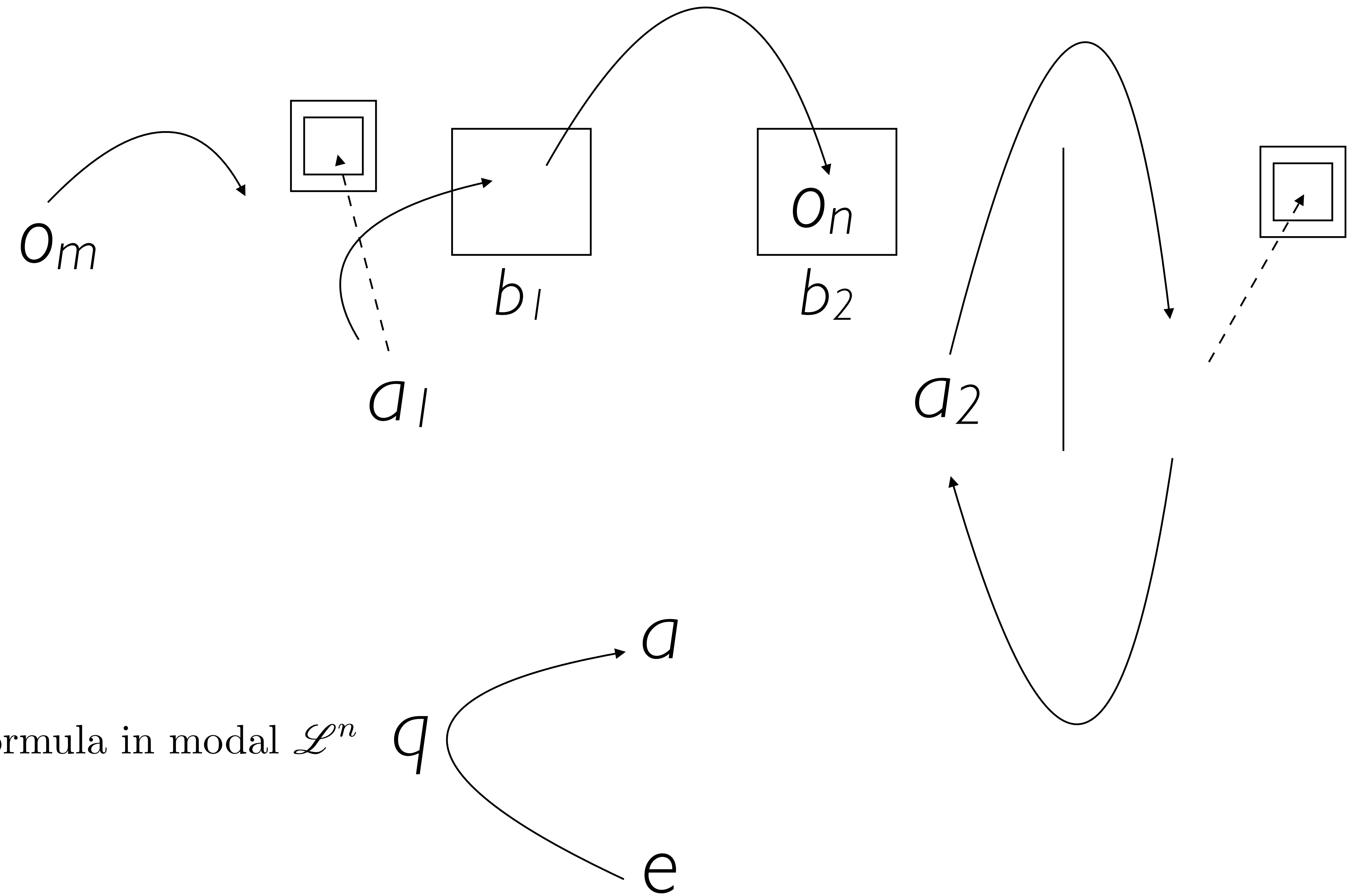


Framework for $FBT^1{}_3$

(eight timepoints)



Framework for $FBT^1{}_3$
(eight timepoints)



q a formula in modal $\mathscr{L}^n$

Simulation Courtesy of …

ShadowProver!

Level I

```
:name        "Level 1: False Belief Task "

:description "Agent a1 puts an object o into b1 in plain view of a2.
              Agent a2 then leaves, and in the absence of a2, a1 moves o
              from b1 into b2 ; this movement isn't perceived by a2 . Agent
              a2 now returns, and a is asked by the experimenter e: "If a2
              desires to retrieve o, which box will a2 look in?" If younger
              than four or five, a will reply "In b " (which of course fails 2
              the task); after this age subjects respond with the correct "In b1."

              Level1 Belief: a1 believes a2 believes o is in b1.
              "

:date        "Monday July 22, 2019"

:assumptions {
               :P1 (Perceives! a1 t1 (Perceives! a2 t1 (holds (In o b1) t1)))

               :P2 (Believes! a1 t2 (Believes! a2 t2 (not (exists [?e] (terminates ?e (In o b1))))))

               :P3 (holds (In o b1) t1)

               :C1 (Common! t0 (forall [?f ?t2 ?t2]
                                (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
                                  (holds ?f ?t2))))

               :C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3)))
               }

:goal        (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))}
```

Level 2

```
{:name         "Level 2: False Belief Task "

 :description "Agent a1 puts an object o into b1 in plain view of a2.
               Agent a2 then leaves, and in the absence of a2, a1 moves o
               from b1 into b2 ; this movement isn't perceived by a2 . Agent
               a2 now returns, and a is asked by the experimenter e: "If a2
               desires to retrieve o, which box will a2 look in?" If younger
               than four or five, a will reply "In b " (which of course fails 2
               the task); after this age subjects respond with the correct "In b1."

               Level2 Belief: a2 believes a1 believes a2 believes o is in b1.
               "

 :date         "Monday July 22, 2019"

 :assumptions {

                :P1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (holds (In o b1) t1))))

                :P2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (not (exists [?e] (terminates ?e (In o b1)))))))

                :P3 (holds (In o b1) t1)


                :C1 (Common! t0
                             (forall [?f ?t2 ?t2]
                                     (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
                                       (holds ?f ?t2))))

               :C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3)))}

 :goal         (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3))))}
```

Level 3

```
{:name        "Level 3: False Belief Task "

 :description "Agent a1 puts an object o into b1 in plain view of a2.
               Agent a2 then leaves, and in the absence of a2, a1 moves o
               from b1 into b2 ; this movement isn’t perceived by a2 . Agent
               a2 now returns, and a is asked by the experimenter e: “If a2
               desires to retrieve o, which box will a2 look in?” If younger
               than four or five, a will reply “In b ” (which of course fails 2
               the task); after this age subjects respond with the correct “In b1.”

               Level3 Belief: a2 believes a1 believes a2 believes o is in b1.
               "

 :date        "Monday July 22, 2019"

 :assumptions {

                :P1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (holds (In o b1) t1)))))
                :P2 (Believes! a1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (not (exists [?e] (terminates ?e (In o b1))))))))

                :P3 (holds (In o b1) t1)


                :C1 (Common! t0
                             (forall [?f ?t2 ?t2]
                                     (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
                                       (holds ?f ?t2))))

                :C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3)))}

 :goal        (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))))}
```

Level 4

```
{:name        "Level 4: False Belief Task "

 :description "Agent a1 puts an object o into b1 in plain view of a2.
               Agent a2 then leaves, and in the absence of a2, a1 moves o
               from b1 into b2 ; this movement isn't perceived by a2 . Agent
               a2 now returns, and a is asked by the experimenter e: "If a2
               desires to retrieve o, which box will a2 look in?" If younger
               than four or five, a will reply "In b " (which of course fails 2
               the task); after this age subjects respond with the correct "In b1."

               Level4 Belief: a2 believes a1 believes a2 believes a1 believes a2 believes o is in b1.
               "

 :date        "Monday July 22, 2019"

 :assumptions {

                :P1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (holds (In o b1) t1))))))
                :P2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (not (exists [?e] (terminates ?e (In o b1))))))))

                :P3 (holds (In o b1) t1)


                :C1 (Common! t0
                             (forall [?f ?t2 ?t2]
                                     (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
                                       (holds ?f ?t2))))

                :C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3)))}

 :goal        (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3))))))}
```

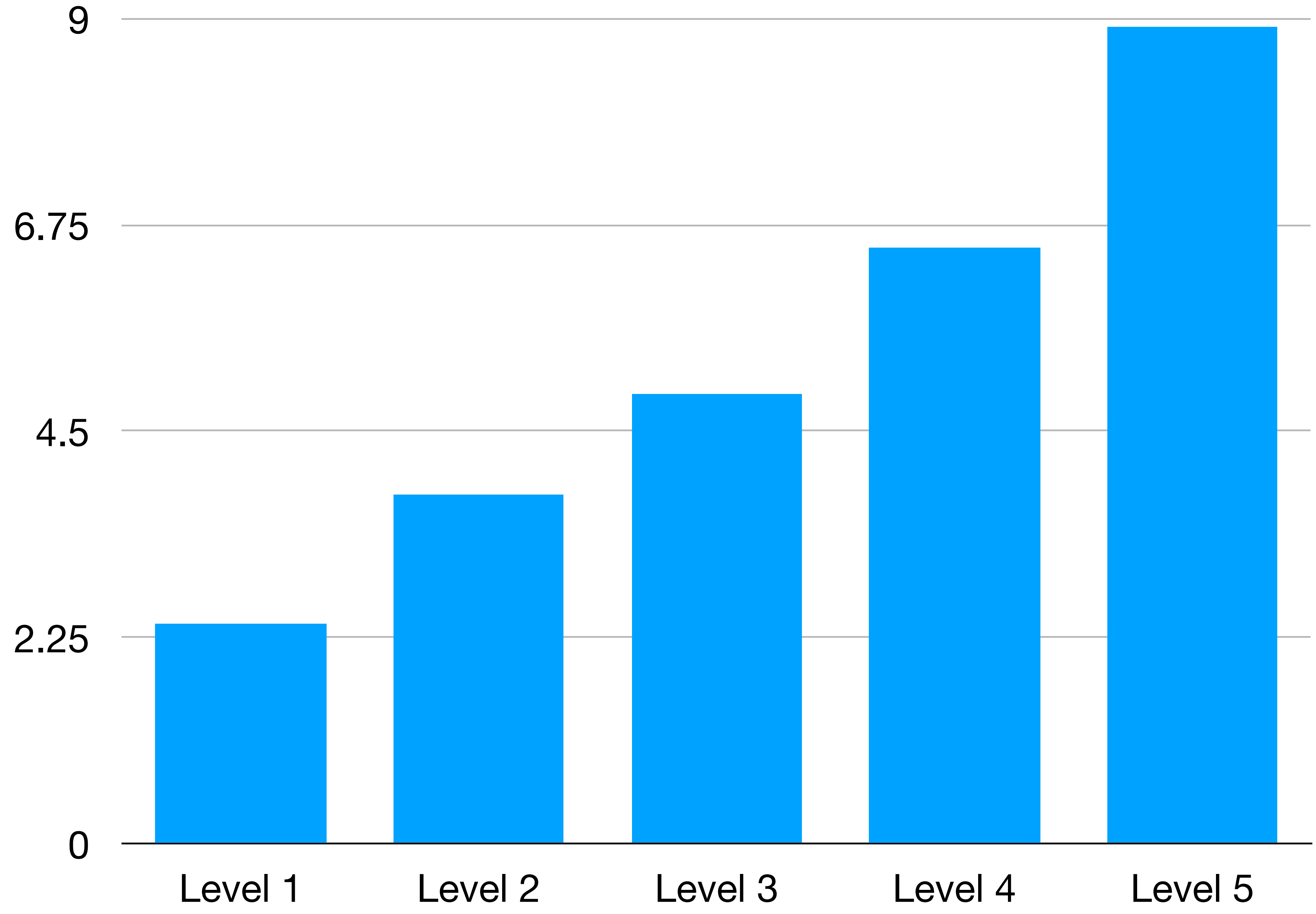
Level 5

```
{:name        "Level 5: False Belief Task "

 :description "Agent a1 puts an object o into b1 in plain view of a2.
               Agent a2 then leaves, and in the absence of a2, a1 moves o
               from b1 into b2 ; this movement isn't perceived by a2 . Agent
               a2 now returns, and a is asked by the experimenter e: “If a2
               desires to retrieve o, which box will a2 look in?” If younger
               than four or five, a will reply “In b ” (which of course fails 2
               the task); after this age subjects respond with the correct “In b1.”

               Level5 Belief: a1 believes a2 believes a1 believes a2 believes a1 believes a2 believes o is in b1.
               "

 :date        "Monday July 22, 2019"

 :assumptions {

                :P1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (holds (In o b1) t1)))))))
                :P2 (Believes! a1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (not (exists [?e] (terminates ?e (In o b1)))))))))

                :P3 (holds (In o b1) t1)


                :C1 (Common! t0
                             (forall [?f ?t2 ?t2]
                                     (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
                                       (holds ?f ?t2))))

                :C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3)))}

 :goal        (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))))))}
```

```
gent
f If a2
f younger
 course fails 2
n the correct “In b1.”
l believes a2 believes a1 believes a2 believes o is in b1.
t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2
1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (not
   (In o b1) t1)
1 (Common! t0
   (forall [?f ?t2 ?t2]
     (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
      (holds ?f ?t2))))
:C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3))))}
:goal (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3))))))))}
```

```
gent
f If a2
f younger
 course fails 2
n the correct “In b1.”
l believes a2 believes a1 believes a2 believes o is in b1.

t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2
1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (Believes! a1 t2 (Believes! a2 t2 (not
(In o b1) t1)
l (Common! t0
      (forall [?f ?t2 ?t2]
        (if (and (not (exists [?e] (terminates ?e ?f))) (holds ?f ?t1) (< ?t1 ?t2))
          (holds ?f ?t2))))
    :C2 (Common! t0 (and (< t1 t2) (< t2 t3) (< t1 t3)))}
:goal
  (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3))))))))}
```

Time (in seconds) to Prove



Simulation of Level 5 in Real Time

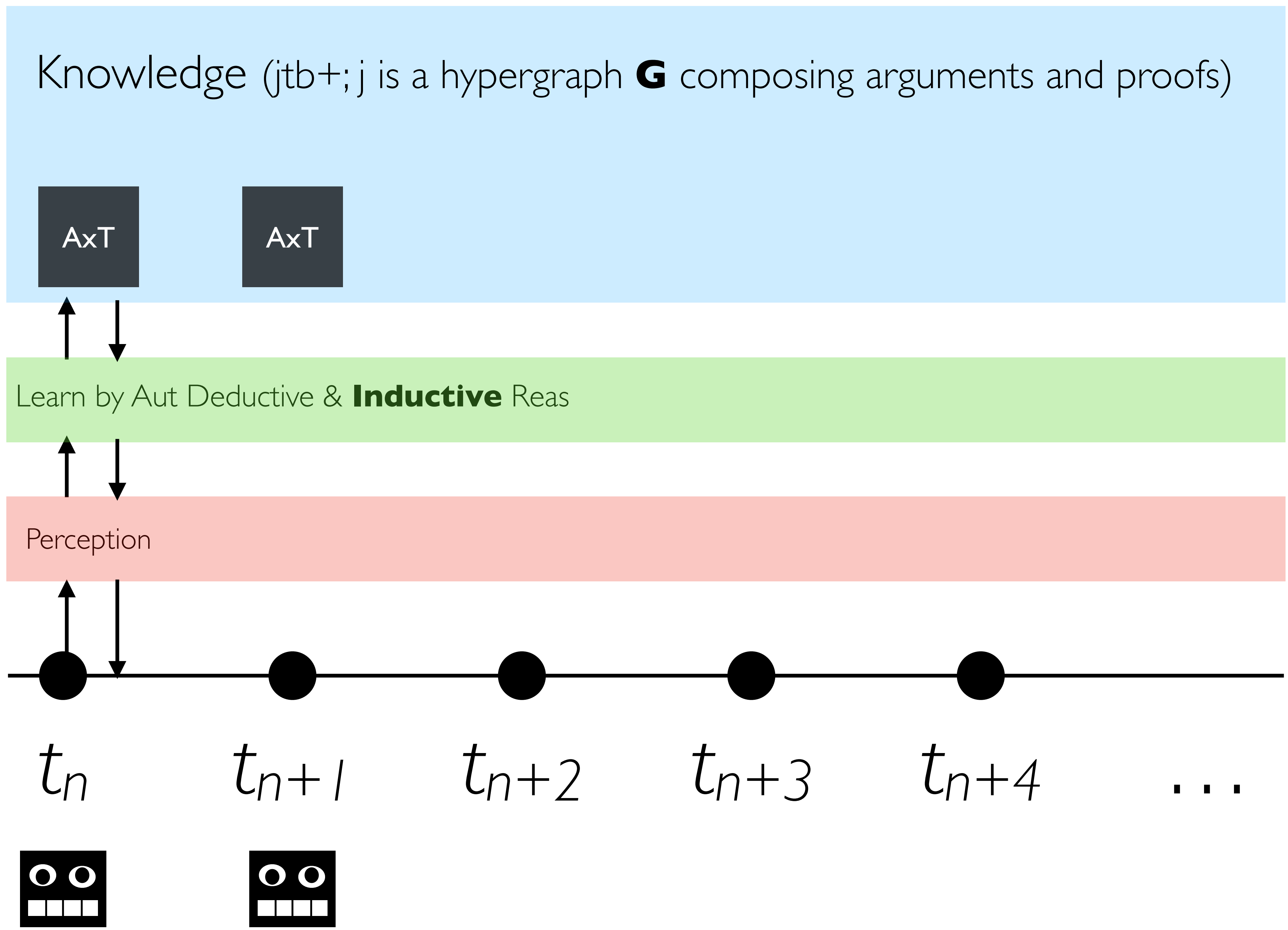
```
/Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java ...
objc[16653]: Class JavaLaunchHelper is implemented in both /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java (0x102a2d4c0) and /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/jre/lib/libinstrument.dylib (0x102ab94e0)
------------- Level 5 ----------------------
```

Simulation of Level 5 in Real Time

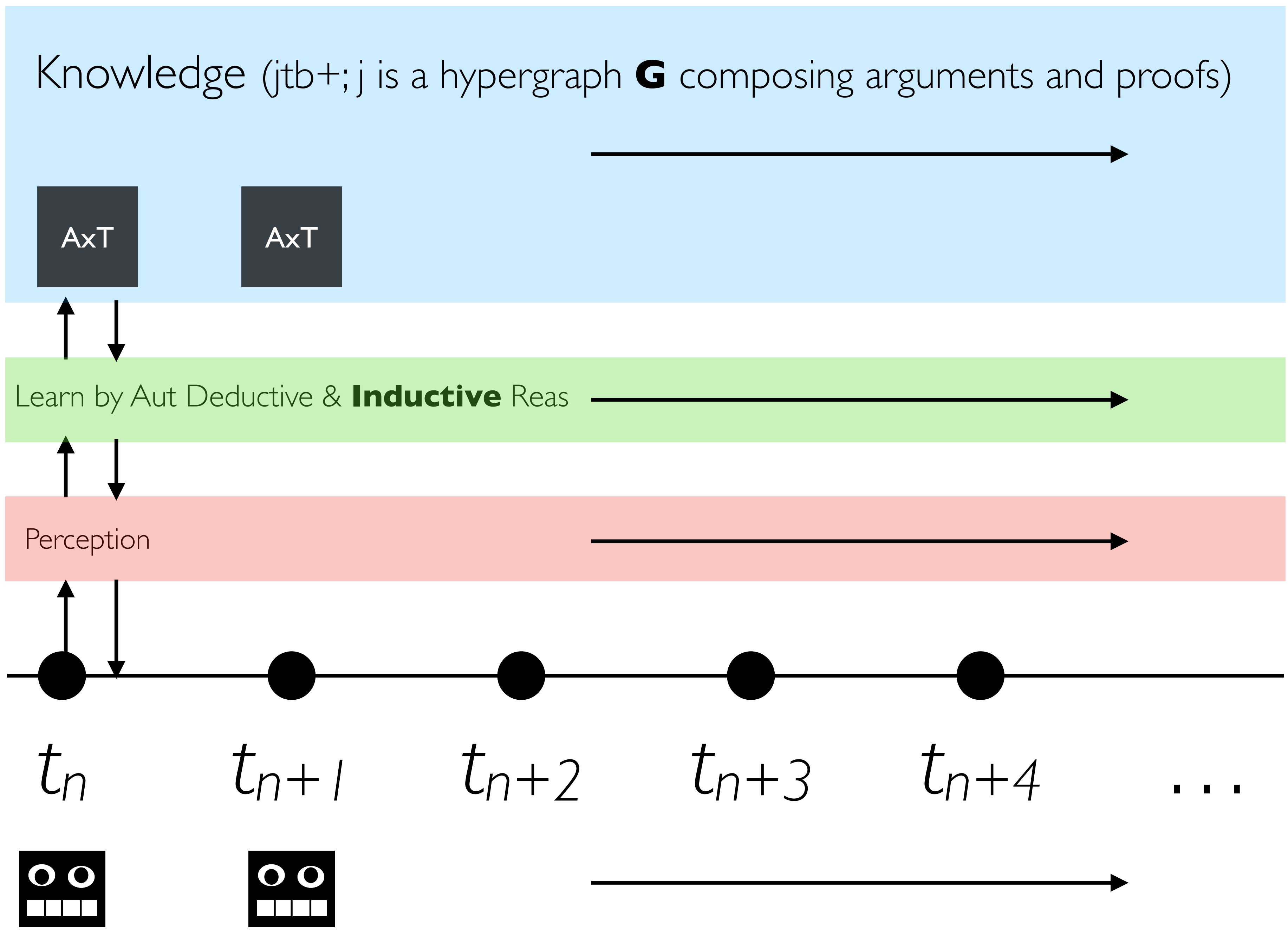
```
/Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java ...
objc[16653]: Class JavaLaunchHelper is implemented in both /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java (0x102a2d4c0) and /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/jre/lib/libinstrument.dylib (0x102ab94e0)
------------- Level 5 ----------------------
```

Real Learning …

Advanced Logicist (Real) Machine Learning

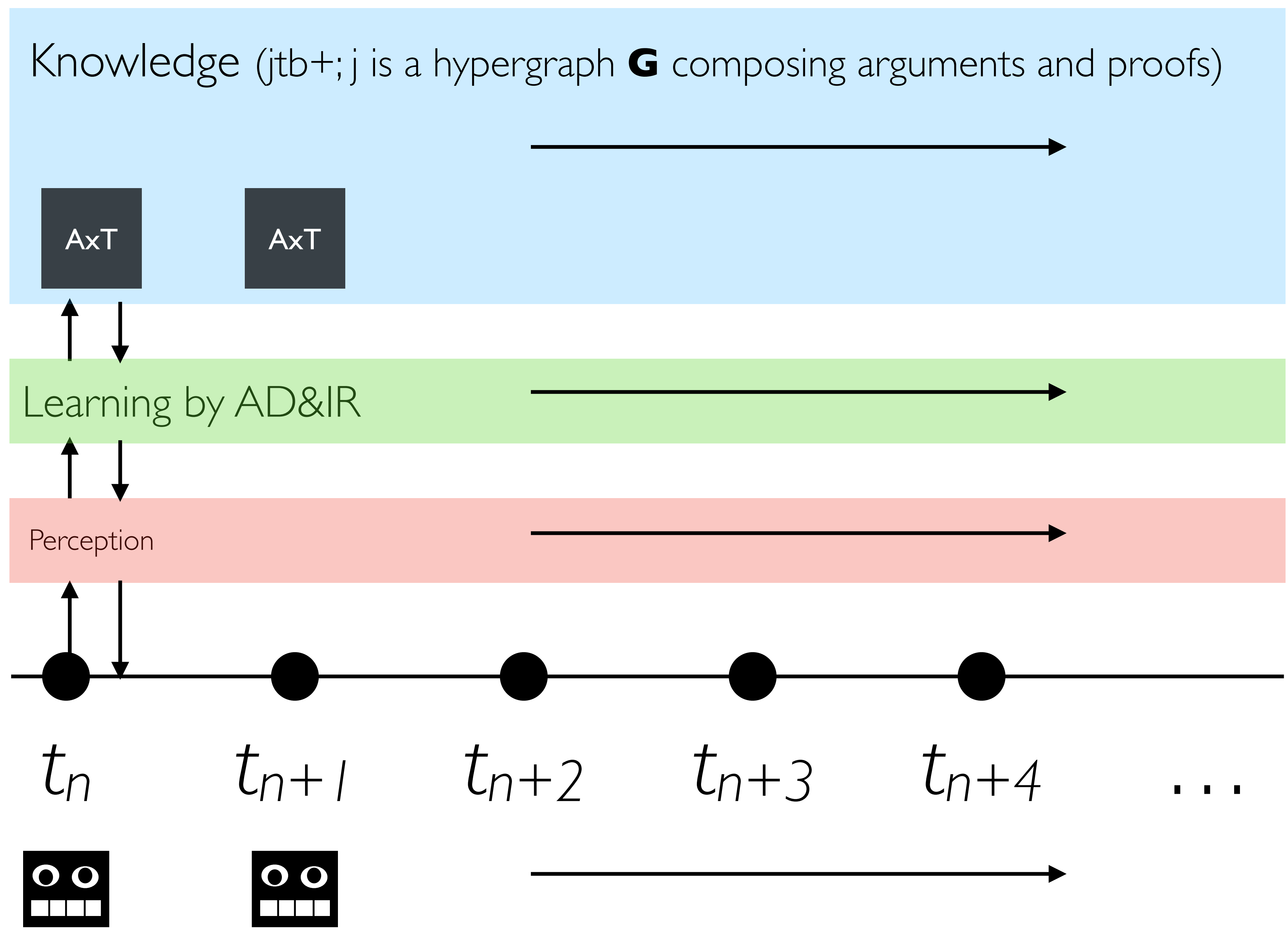


Advanced Logicist (Real) Machine Learning



Advanced Logicist (Real) Machine Learning

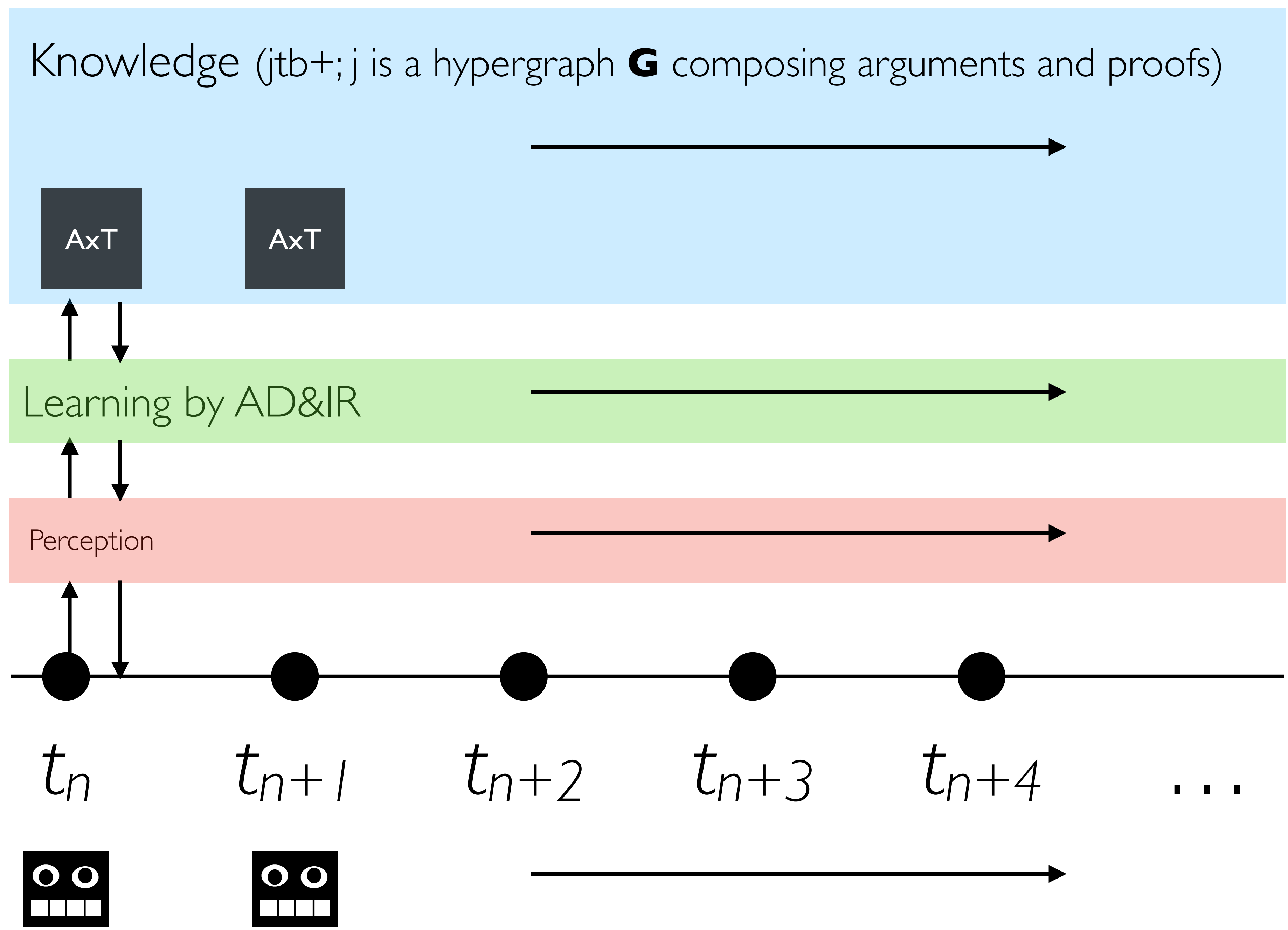
Reasoning:



Advanced Logicist (Real) Machine Learning

Reasoning:

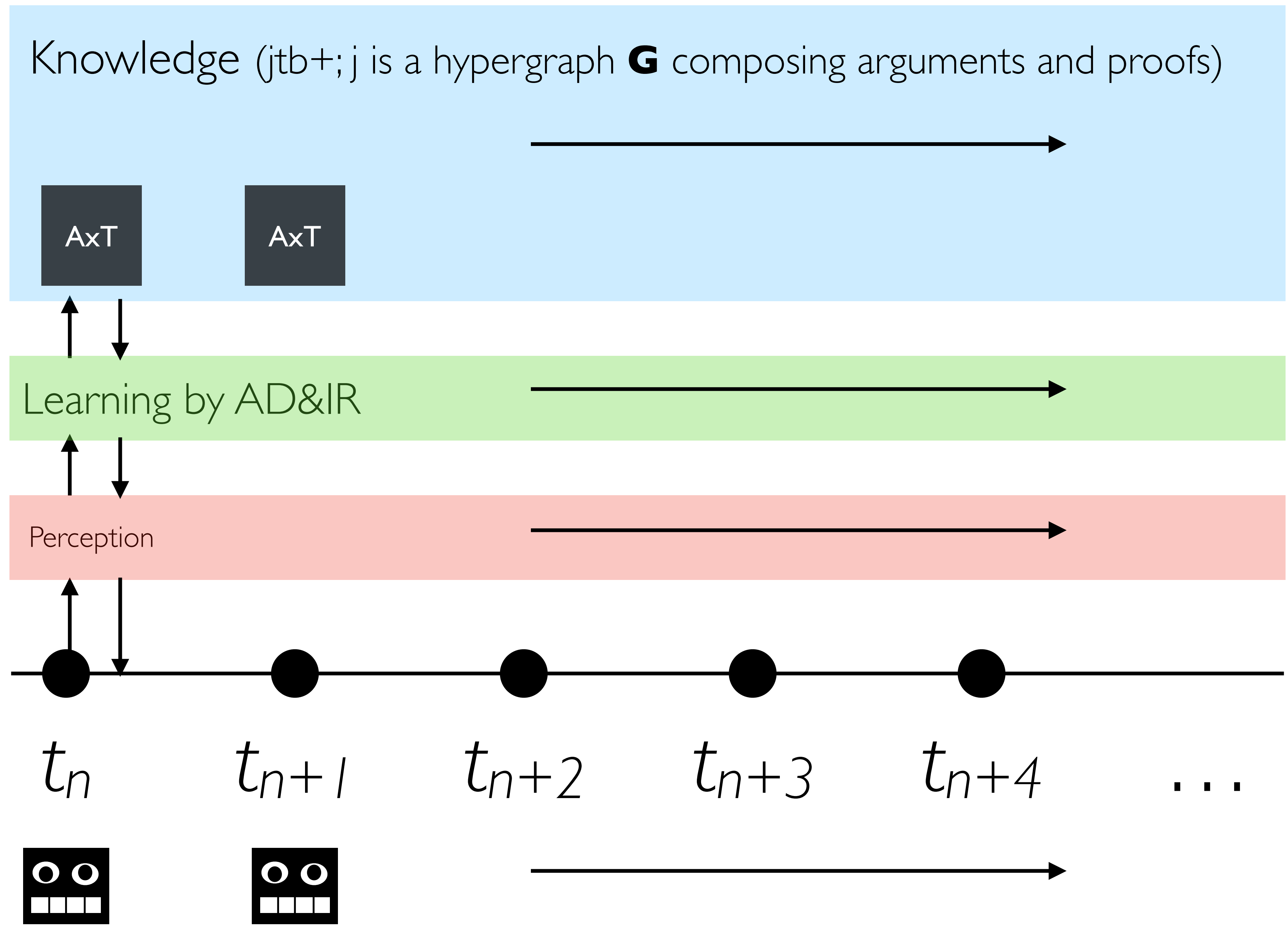
deduction
analogical reasoning
analogico-deductive reasoning
abduction
Bayesian argumentation
enumerative induction
ITBE
. . .



Advanced Logicist (Real) Machine Learning

Reasoning:

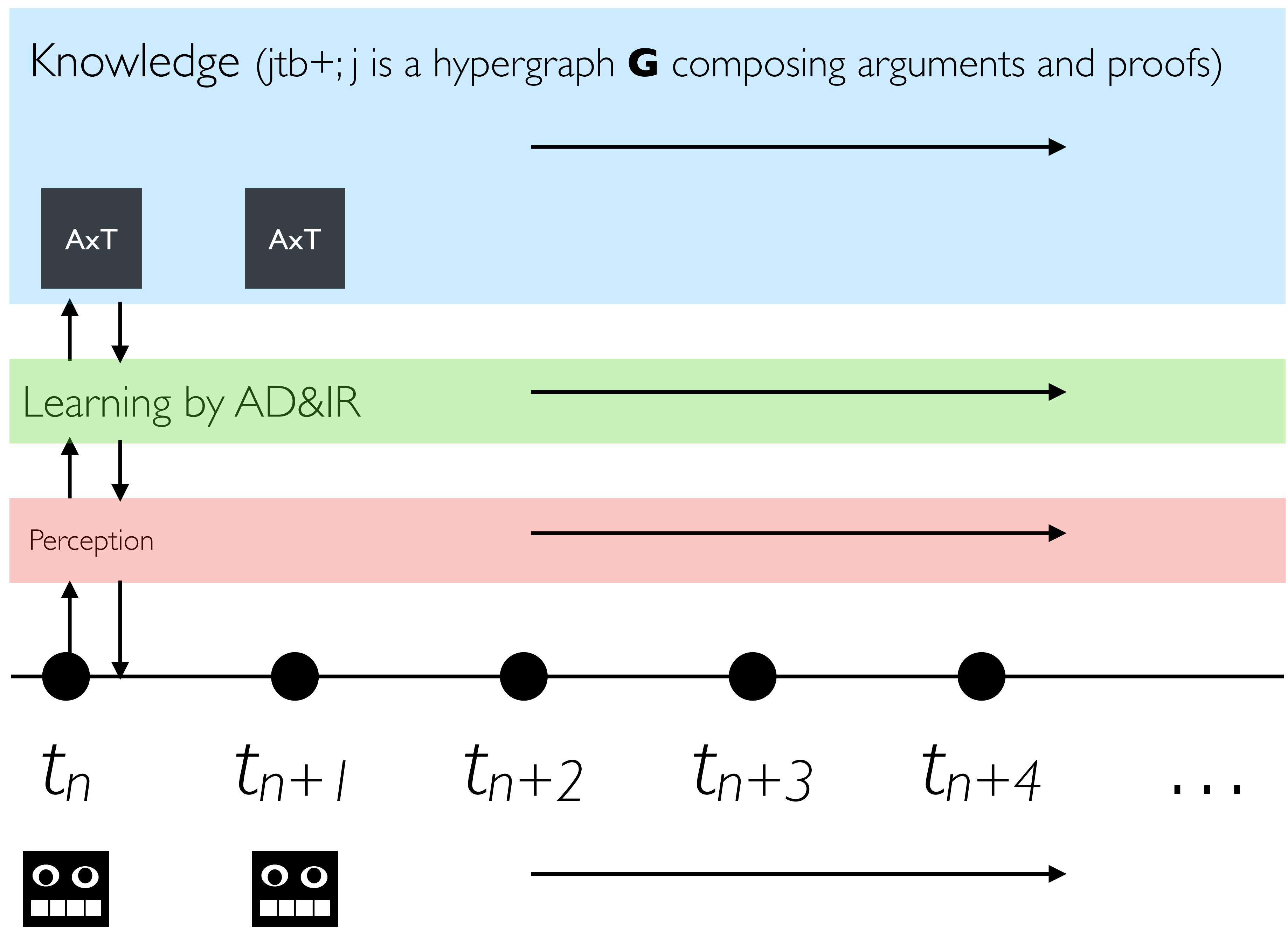
deduction
analogical reasoning
analogico-deductive reasoning
abduction
Bayesian argumentation
enumerative induction
ITBE
. . .



Advanced Logicist (Real) Machine Learning

Defeasible Reasoning:

deduction
analogical reasoning
analogico-deductive reasoning
abduction
Bayesian argumentation
enumerative induction
ITBE
. . .



Bringsjord, S. & Govindarajulu, N.S. (2018) "The Epistemology of Computer-Mediated Proofs" in Hansson, Sven Ove, ed., *Technology and Mathematics: Philosophical and Historical Investigations* (Berlin, GE: Springer). This book appears in the series *Philosophy of Engineering and Technology* (as Volume 30) edited by Pieter Vermass. ISBN is 978-3-319-93778-6.

Given This, Do Machine-Learning Machines Learn? No.

http://kryten.mm.rpi.edu/SB_NSG_SB_JH_DoMachine-LearningMachinesLearn_preprint.pdf

Given This, Do Machine-Learning Machines Learn? No.

Do Machine-Learning Machines Learn?

Selmer Bringsjord and Naveen Sundar Govindarajulu and Shreya Banerjee and John Hummel

Abstract We answer the present paper's title in the negative. We begin by introducing and characterizing "real learning" ($\mathcal{RL}$) in the formal sciences, a phenomenon that has been firmly in place in homes and schools since at least Euclid. The defense of our negative answer pivots on an integration of *reductio* and proof by cases, and constitutes a general method for showing that any contemporary form of machine learning (ML) isn't real learning. Along the way, we canvass the many different conceptions of "learning" in not only AI, but psychology and its allied disciplines; none of these conceptions (with one exception arising from the view of cognitive development espoused by Piaget), aligns with real learning. We explain in this context by four steps how to broadly characterize and arrive at a focus on $\mathcal{RL}$.

Selmer Bringsjord
Rensselaer Polytechnic Institute, 110 8th Street Troy, NY USA 12180, e-mail: selmerbringsjord@gmail.com

Naveen Sundar Govindarajulu
Rensselaer Polytechnic Institute, 110 8th Street Troy, NY USA 12180, e-mail: Naveen.Sundar.G@gmail.com

Shreya Banerjee
Rensselaer Polytechnic Institute, 110 8th Street Troy, NY USA 12180, e-mail: shreyabbanerjee@gmail.com

John Hummel
901 West Illinois Street, Urbana, IL 61801, e-mail: jehummel@illinois.edu

8 Appendix: The Formal Method

The following deduction uses fonts in an obvious and standard way to sort between functions ($\mathfrak{f}$), agents ($\mathfrak{a}$), and computing machines ($\mathfrak{m}$) in the Arithmetical Hierarchy. Ordinary italicized Roman is used for particulars under these sorts (e.g. f is a particular function). In addition, '$\mathcal{C}$' denotes any collection of conditions constituting jointly necessary-and-sufficient conditions for a form of current ML, which can come from relevant textbooks (e.g. Luger, 2008; Russell and Norvig, 2009) or papers; we leave this quite up to the reader, as no effect upon the validity of the deductive inference chain will be produced by the preferred instantiation of '$\mathcal{C}$.' It will perhaps be helpful to the reader to point out that the deduction eventuates in the proposition that no machine in the ML fold that in this style learns a relevant function $\mathfrak{f}$ thereby also real-learns $\mathfrak{f}$. We encode this target as follows:

$$(\star)\quad \neg\exists \mathfrak{m}\, \exists \mathfrak{f}\, [\phi := \mathit{MLlearns}(\mathfrak{m},\mathfrak{f}) \wedge \psi := \mathit{RLlearns}(\mathfrak{m},\mathfrak{f}) \wedge \mathcal{C}_\phi(\mathfrak{m},\mathfrak{f}) \vdash^* (\mathrm{ci}')\text{–}(\mathrm{ciii})_\psi(\mathfrak{m},\mathfrak{f})]$$

Note that $(\star)$ employs meta-logical machinery to refer to particular instantiations of $\mathcal{C}$ for a particular, arbitrary case of ML (ϕ is the atomic sub-formula that can be instantiated to make the particular case), and particular instantiations of the triad (ci′)–(ciii) for a particular, arbitrary case of $\mathcal{RL}$ (ψ is the atomic sub-formula that can be instantiated to make the particular case). Meta-logical machinery also allows us to use a provability predicate to formalize the notion that real learning is produced by the relevant instance of ML. If we "pop" ϕ/ψ to yield ϕ'/ψ' we are dealing with the particular instantiation of the atomic sub-formula.

The deduction, as noted in earlier when the informal argument was given, is indirect proof by cases; accordingly, we first assume $\neg(\star)$, and then proceed as follows under this supposition.

	(1)	$\forall \mathfrak{f},\mathfrak{a}\, [\mathfrak{f} : \mathbb{N} \mapsto \mathbb{N} \rightarrow (\mathit{RLlearns}(\mathfrak{a},\mathfrak{f}) \rightarrow \text{(i)–(iii)})]$	Def of Real Learning
	(2)	$\mathit{MLlearns}(m,f) \wedge \mathit{RLlearns}(m,f) \wedge f : \mathbb{N} \mapsto \mathbb{N}$	supp (for $\exists$ elim on $(\star)$)
	(3)	$\forall \mathfrak{m},\mathfrak{f}\, [\mathfrak{f} : \mathbb{N} \mapsto \mathbb{N} \rightarrow (\mathit{MLlearns}(\mathfrak{m},\mathfrak{f}) \leftrightarrow \mathcal{C}(\mathfrak{m},\mathfrak{f}))]$	Def of ML
	(4)	$\forall \mathfrak{f}\, [\mathfrak{f} : \mathbb{N} \mapsto \mathbb{N} \rightarrow (\mathit{TurComp}(\mathfrak{f}) \vee \mathit{TurUncomp}(\mathfrak{f}))]$	theorem
	(5)	$\mathit{TurUncomp}(f)$	supp; Case 1
	(6)	$\neg\exists \mathfrak{m}\, \exists \mathfrak{f}\, [(\mathfrak{f} : \mathbb{N} \mapsto \mathbb{N} \wedge \mathit{TurUncomp}(\mathfrak{f}) \wedge \mathcal{C}(\mathfrak{m},\mathfrak{f})]$	theorem
$\therefore$	(7)	$\neg\exists \mathfrak{m}\, \mathit{MLlearns}(\mathfrak{m},f)$	(6), (3)
$\therefore$	(8)	$\bot$	(7), (2)
	(9)	$\mathit{TurComp}(f)$	supp; Case 2
$\therefore$	(10)	$\mathcal{C}_{\phi'}(m,f)$	(2), (3)
$\therefore$	(11)	$(\mathrm{ci}')\text{–}(\mathrm{ciii})_{\psi'}(m,f)$	from supp for $\exists$ elim on $(\star)$ and provability
$\therefore$	(12)	$\neg(\mathrm{ci}')\text{–}(\mathrm{ciii})_{\psi'}(m,f)$	inspection: proofs wholly absent from $\mathcal{C}$
$\therefore$	(13)	$\bot$	(11), (12)
$\therefore$	(14)	$\bot$	*reductio*; proof by cases

http://kryten.mm.rpi.edu/SB_NSG_SB_JH_DoMachine-LearningMachinesLearn_preprint.pdf https://en.wikipedia.org/wiki/File:2001_A_Space_Odyssey_(1968)_-_Trailer.webm https://en.wikipedia.org/wiki/File:2001_A_Space_Odyssey_(1968)_-_Trailer.webm

Det er en logikk for hvert problem!