On to Intensional Logics

(In this edition of IFLAI2, S5 & DCEC only.)

Selmer Bringsjord

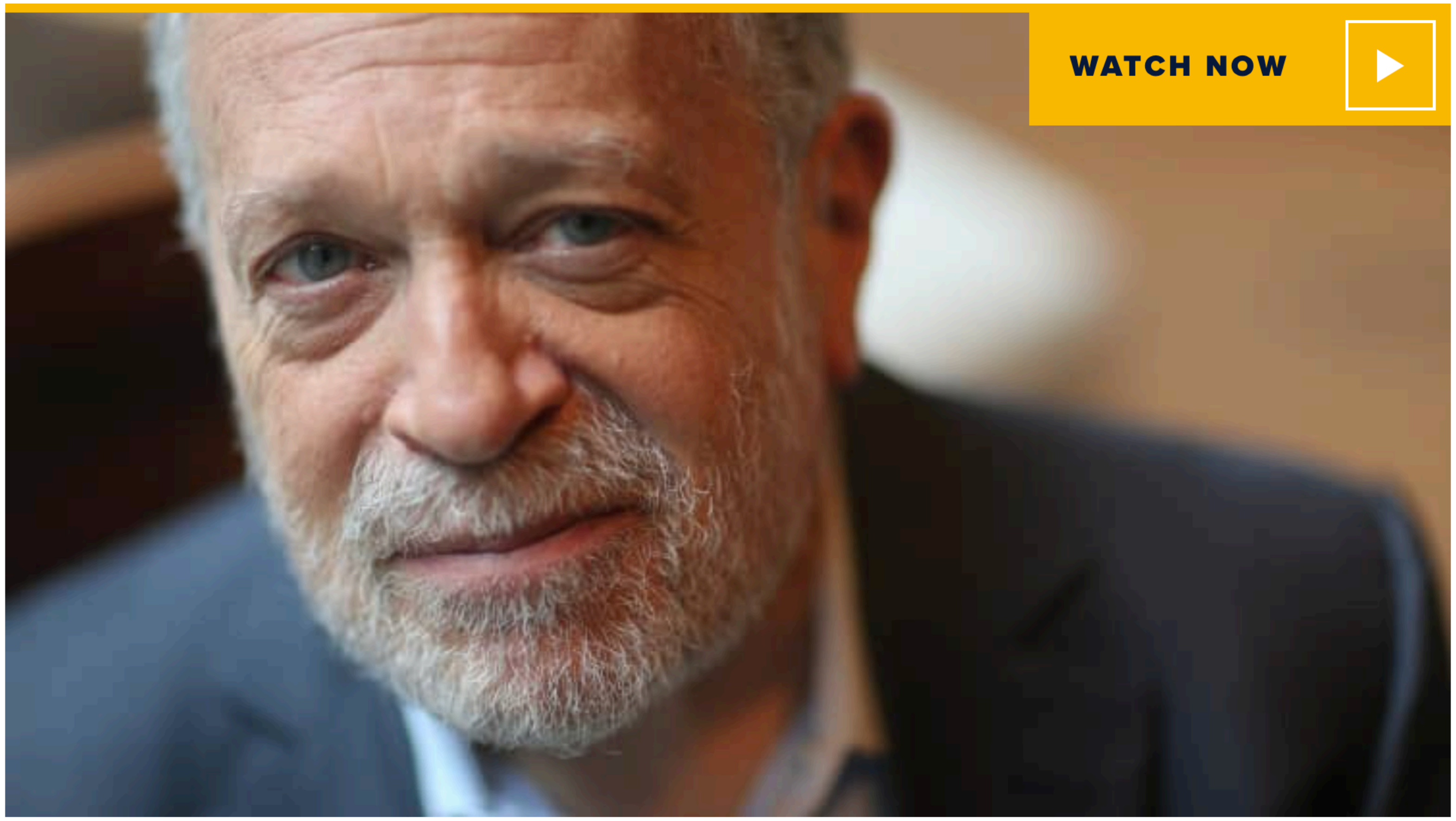
Rensselaer AI & Reasoning (RAIR) Lab
Department of Cognitive Science
Department of Computer Science
Lally School of Management & Technology
Rensselaer Polytechnic Institute (RPI)
Troy, New York 12180 USA

Intermediate Formal Logic & AI (IFLAI2)
9/7/2023

R A I R
Rensselaer AI and Reasoning Lab

In The Logic-and-AI News

…



THE BOTTOM LINE

SHARE

Why AI may end labor unions and become your new employer: Robert Reich

Sep 3 2023

Mechanics …

Mechanics …

Mechanics …

Access codes available if not registered!

Mechanics …

Access codes available if not registered!

Might find helpful the "Getting Off The Ground …" tutorial, available by link in "Tutorials" section of our course web page.

Mechanics …

Access codes available if not registered!

Might find helpful the "Getting Off The Ground …" tutorial, available by link in "Tutorials" section of our course web page.

Important: Your "University ID" as a student @ RPI is you RIN #.

Mechanics …

Access codes available if not registered!

Might find helpful the "Getting Off The Ground …" tutorial, available by link in "Tutorials" section of our course web page.

Important: Your "University ID" as a student @ RPI is you RIN #.

Important: 1 browser, 1 window w/ tabs, & immaculate housekeeping (versioning up-to-date eg).

Mechanics …

Access codes available if not registered!

Might find helpful the "Getting Off The Ground …" tutorial, available by link in "Tutorials" section of our course web page.

Important: Your "University ID" as a student @ RPI is you RIN #.

Important: 1 browser, 1 window w/ tabs, & immaculate housekeeping (versioning up-to-date eg).

As part of review, a personalized problem should now be in your HG account & due by COC today. Try it! In fact let's look @ mine now, together and have a tutorial …

On the esemplastic extensional-logic ladder …
questions?

Climbing the *k*-order Ladder

Climbing the *k*-order Ladder

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

$$Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

ZOL $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

There's some thing which is a llama and likes b (which is also a llama), and whose father is a llama too.

ZOL $Llama(a) \wedge Llama(b) \wedge Likes(a, b) \wedge Llama(fatherOf(a))$

a is a llama, as is b, a likes b, and the father of a is a llama as well.

Climbing the *k*-order Ladder

$$\exists x[Llama(x) \land Llama(b) \land Likes(x, b) \land Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $$Llama(a) \land Llama(b) \land Likes(a, b) \land Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

FOL $\exists x[Llama(x) \wedge Llama(b) \wedge Likes(x,b) \wedge Llama(fatherOf(x))]$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\exists x[Llama(x) \land Llama(b) \land Likes(x, b) \land Llama(fatherOf(x))]$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $Llama(a) \land Llama(b) \land Likes(a, b) \land Llama(fatherOf(a))$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

$$\exists x \exists y \exists R[R(x) \land R(y) \land Likes(x, y) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $$\exists x[Llama(x) \land Llama(b) \land Likes(x, b) \land Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $$Llama(a) \land Llama(b) \land Likes(a, b) \land Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

SOL $\exists x \exists y \exists R[R(x) \wedge R(y) \wedge Likes(x, y) \wedge R(fatherOf(x))]$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\exists x[Llama(x) \wedge Llama(b) \wedge Likes(x, b) \wedge Llama(fatherOf(x))]$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $Llama(a) \wedge Llama(b) \wedge Likes(a, b) \wedge Llama(fatherOf(a))$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $\exists x \exists y \exists R[R(x) \wedge R(y) \wedge \mathit{Likes}(x, y) \wedge R(\mathit{fatherOf}(x))]$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\exists x[\mathit{Llama}(x) \wedge \mathit{Llama}(b) \wedge \mathit{Likes}(x, b) \wedge \mathit{Llama}(\mathit{fatherOf}(x))]$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $\mathit{Llama}(a) \wedge \mathit{Llama}(b) \wedge \mathit{Likes}(a, b) \wedge \mathit{Llama}(\mathit{fatherOf}(a))$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

$$\exists x, y\ \exists R, R^2 [R(x) \land R(y) \land R^2(x,y) \land Positive(R^2) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $$\exists x \exists y \exists R[R(x) \land R(y) \land Likes(x,y) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $$\exists x[Llama(x) \land Llama(b) \land Likes(x,b) \land Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $$Llama(a) \land Llama(b) \land Likes(a,b) \land Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

TOL $\exists x, y \; \exists R, R^2 [R(x) \wedge R(y) \wedge R^2(x,y) \wedge Positive(R^2) \wedge R(fatherOf(x))]$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $\exists x \exists y \exists R [R(x) \wedge R(y) \wedge Likes(x,y) \wedge R(fatherOf(x))]$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\exists x [Llama(x) \wedge Llama(b) \wedge Likes(x,b) \wedge Llama(fatherOf(x))]$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

TOL
$\mathscr{L}_3$

$$\exists x, y\ \exists R, R^2[R(x) \wedge R(y) \wedge R^2(x, y) \wedge Positive(R^2) \wedge R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL
$\mathscr{L}_2$

$$\exists x \exists y \exists R[R(x) \wedge R(y) \wedge Likes(x, y) \wedge R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL
$\mathscr{L}_1$

$$\exists x[Llama(x) \wedge Llama(b) \wedge Likes(x, b) \wedge Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL
$\mathscr{L}_0$

$$Llama(a) \wedge Llama(b) \wedge Likes(a, b) \wedge Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the *k*-order Ladder

⋮

TOL $\mathscr{L}_3$

$$\exists x, y\ \exists R, R^2[R(x) \land R(y) \land R^2(x,y) \land Positive(R^2) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $\mathscr{L}_2$

$$\exists x \exists y \exists R[R(x) \land R(y) \land Likes(x,y) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\mathscr{L}_1$

$$\exists x[Llama(x) \land Llama(b) \land Likes(x,b) \land Llama(fatherOf(x))]$$

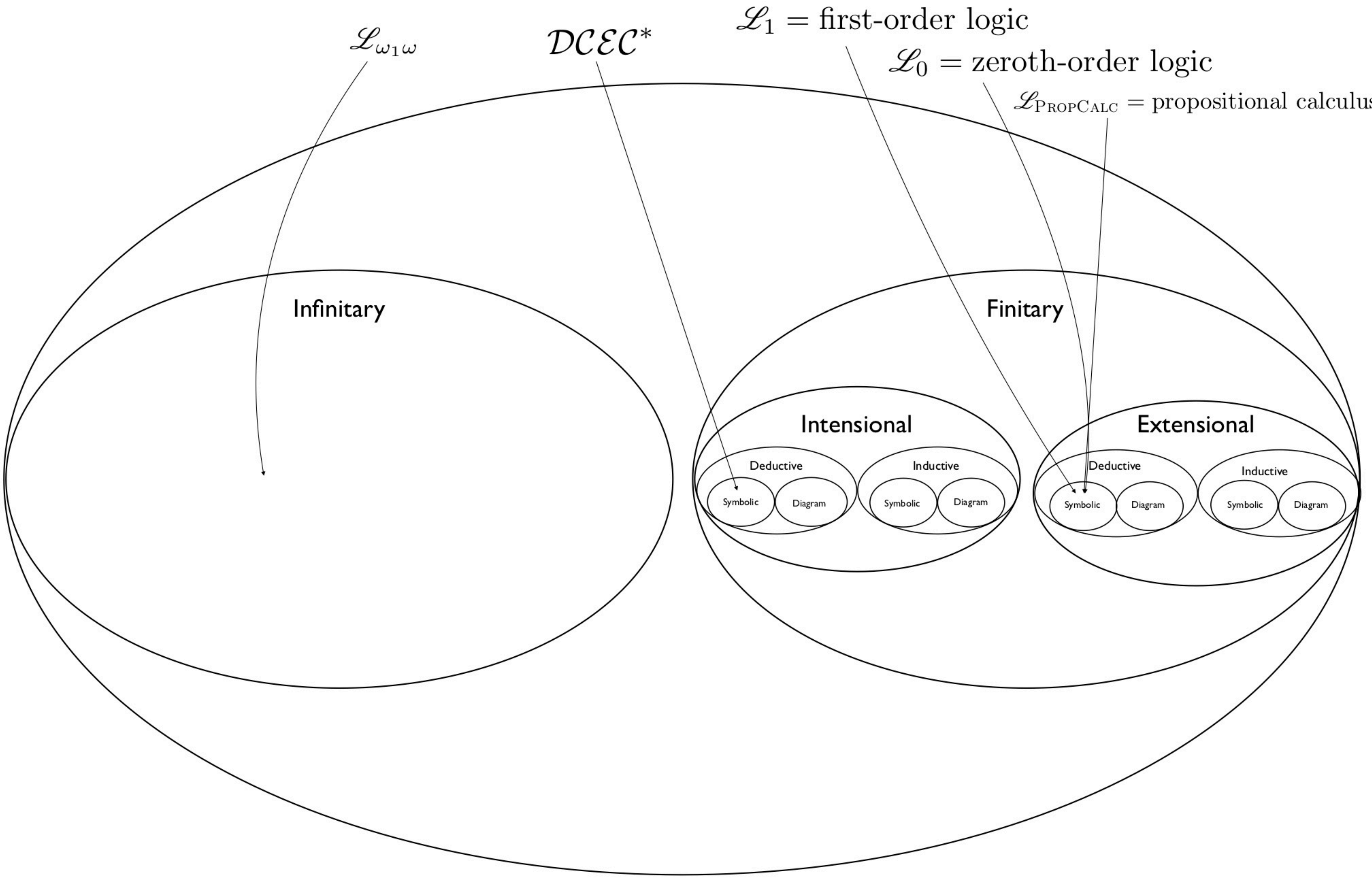
There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $\mathscr{L}_0$

$$Llama(a) \land Llama(b) \land Likes(a,b) \land Llama(fatherOf(a))$$

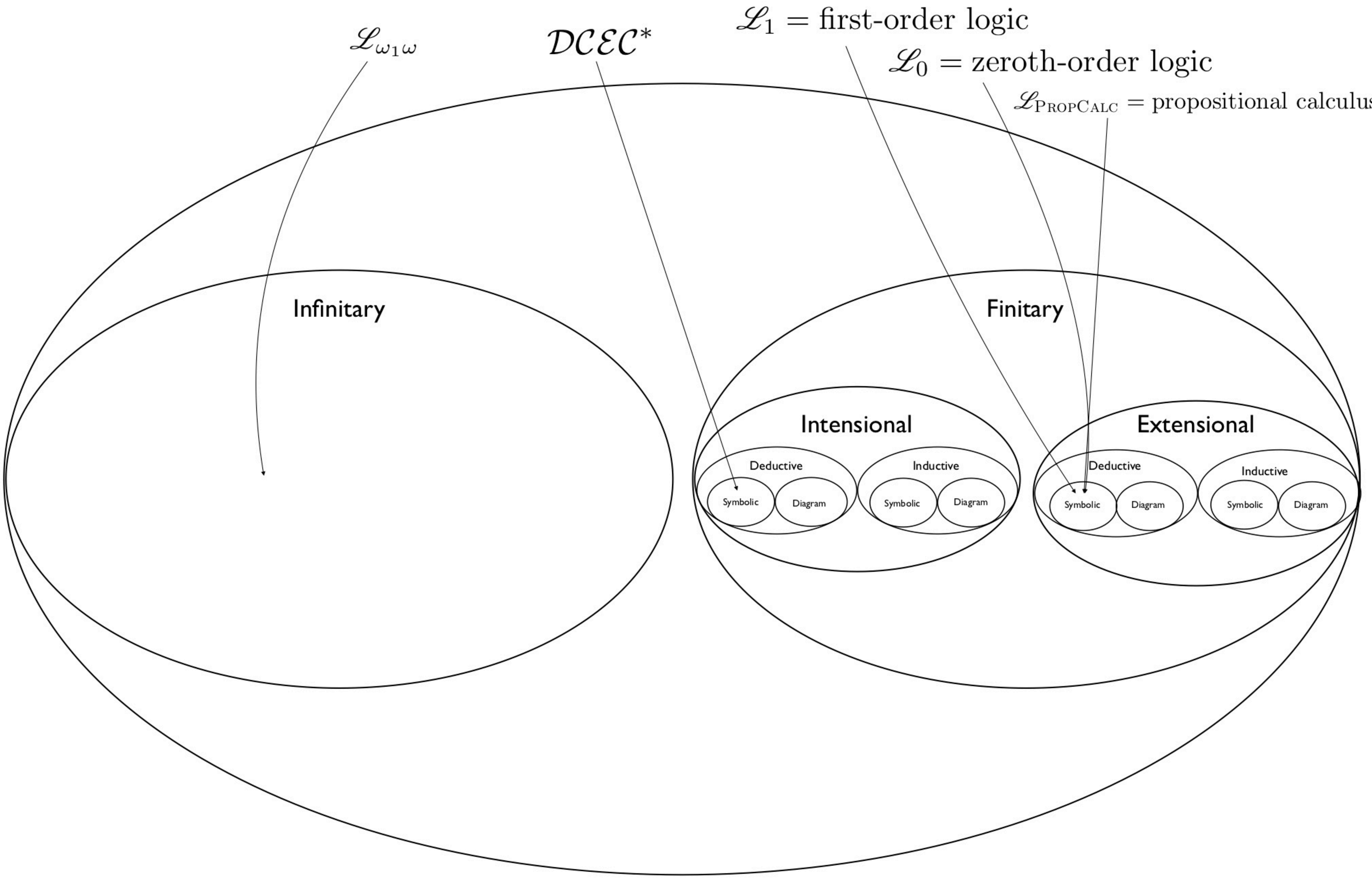
a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

The Universe of Logics

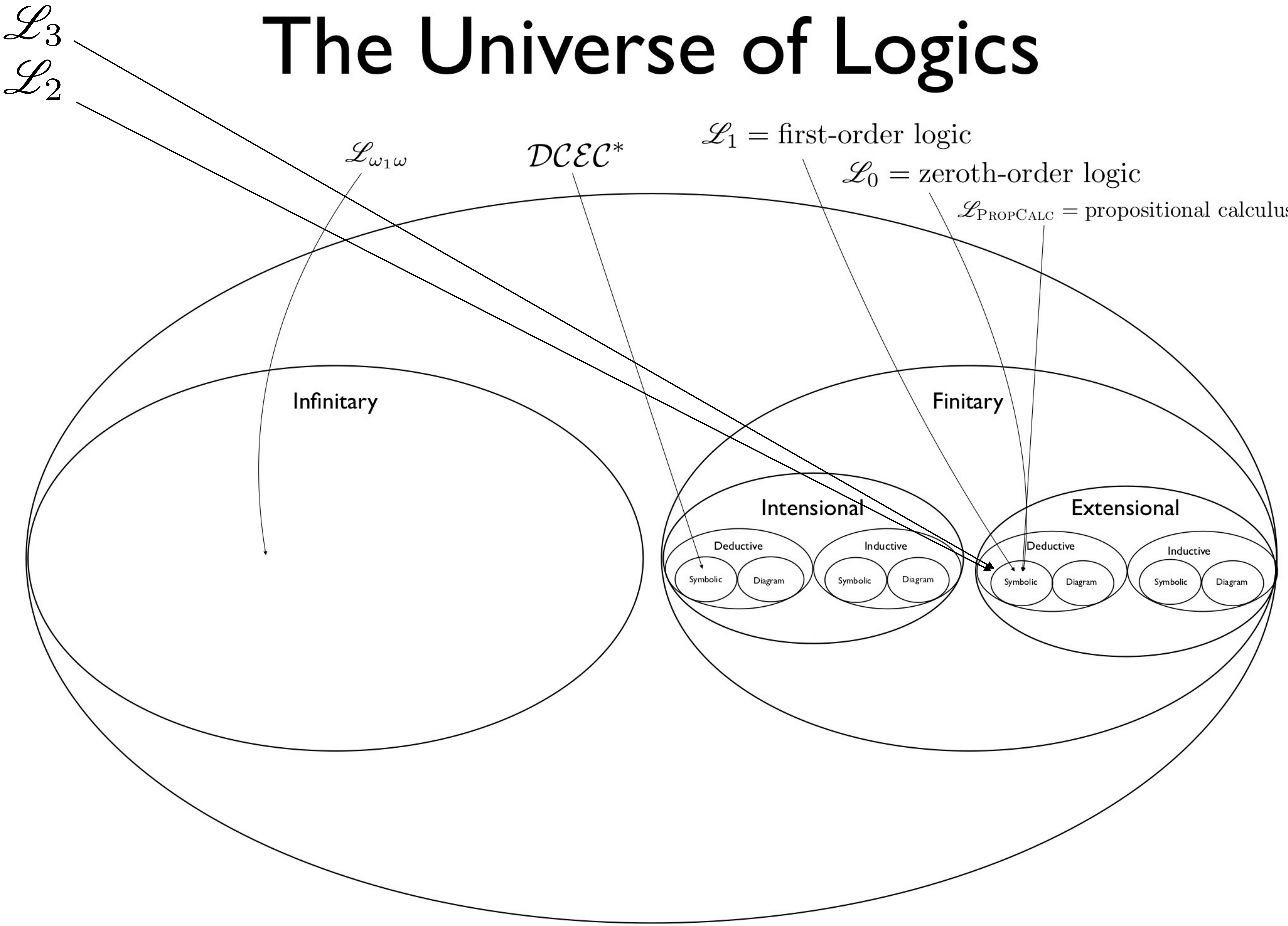


$\mathscr{L}_3$
$\mathscr{L}_2$

The Universe of Logics



The Universe of Logics



Climbing the *k*-order Ladder

$\vdots$

TOL $\mathscr{L}_3$

$$\exists x, y\ \exists R, R^2[R(x) \land R(y) \land R^2(x,y) \land Positive(R^2) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $\mathscr{L}_2$

$$\exists x \exists y \exists R[R(x) \land R(y) \land Likes(x,y) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\mathscr{L}_1$

$$\exists x[Llama(x) \land Llama(b) \land Likes(x,b) \land Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $\mathscr{L}_0$

$$Llama(a) \land Llama(b) \land Likes(a,b) \land Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the k-order Ladder

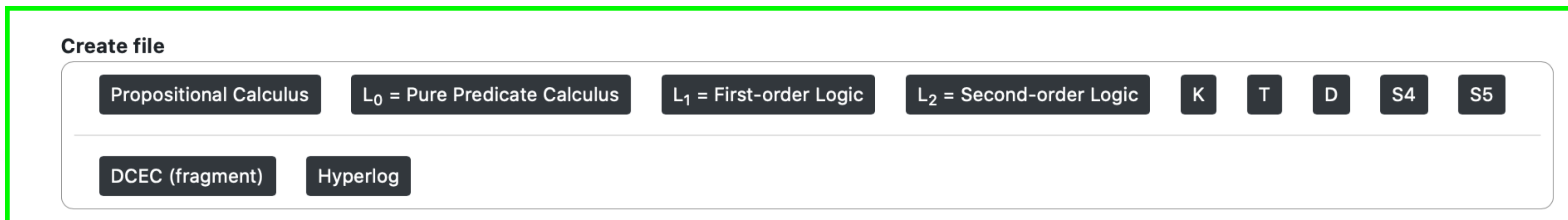
$\vdots$

TOL $\mathscr{L}_3$

$$\exists x, y\ \exists R, R^2[R(x) \land R(y) \land R^2(x,y) \land Positive(R^2) \land R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $\mathscr{L}_2$



$$herOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\mathscr{L}_1$

$$\exists x[Llama(x) \land Llama(b) \land Likes(x,b) \land Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

ZOL $\mathscr{L}_0$

$$Llama(a) \land Llama(b) \land Likes(a,b) \land Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

Climbing the k-order Ladder

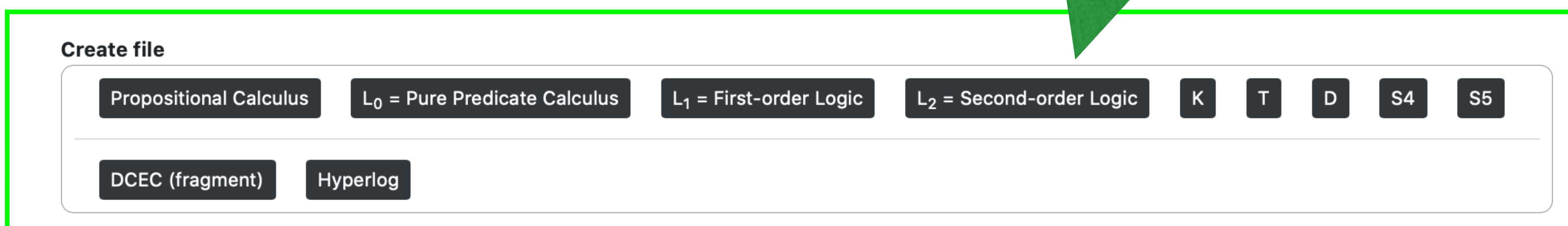
$\vdots$

TOL $\mathscr{L}_3$

$$\exists x, y\ \exists R, R^2[R(x) \wedge R(y) \wedge R^2(x,y) \wedge Positive(R^2) \wedge R(fatherOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property; and, *x* R^2s *y*, where R^2 is a positive property.

SOL $\mathscr{L}_2$



$$\ldots herOf(x))]$$

Things *x* and *y*, along with the father of *x*, share a certain property (and *x* likes *y*).

FOL $\mathscr{L}_1$

$$\exists x[Llama(x) \wedge Llama(b) \wedge Likes(x,b) \wedge Llama(fatherOf(x))]$$

There's some thing which is a llama and likes *b* (which is also a llama), and whose father is a llama too.

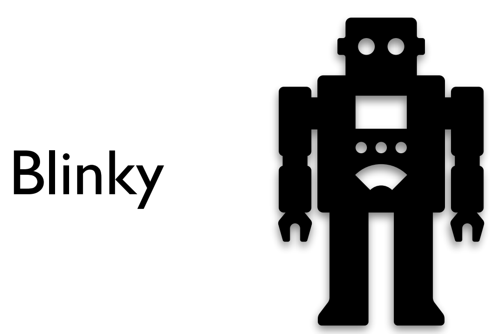
ZOL $\mathscr{L}_0$

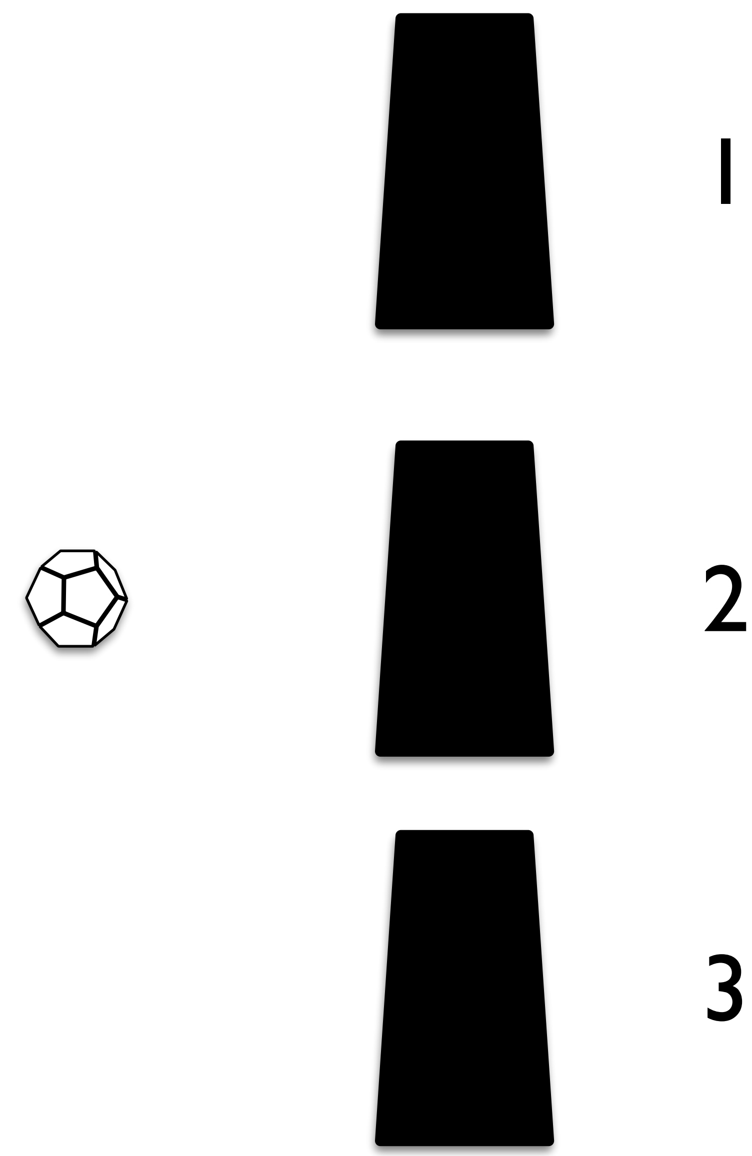
$$Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$$

a is a llama, as is *b*, *a* likes *b*, and the father of *a* is a llama as well.

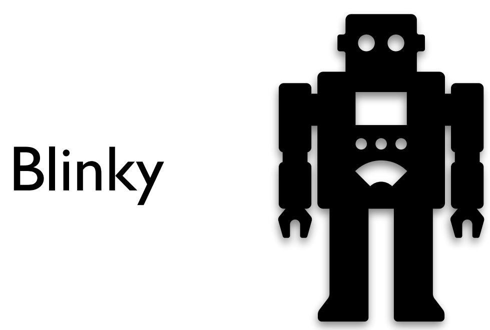
Blinky as portal to intensional logics …

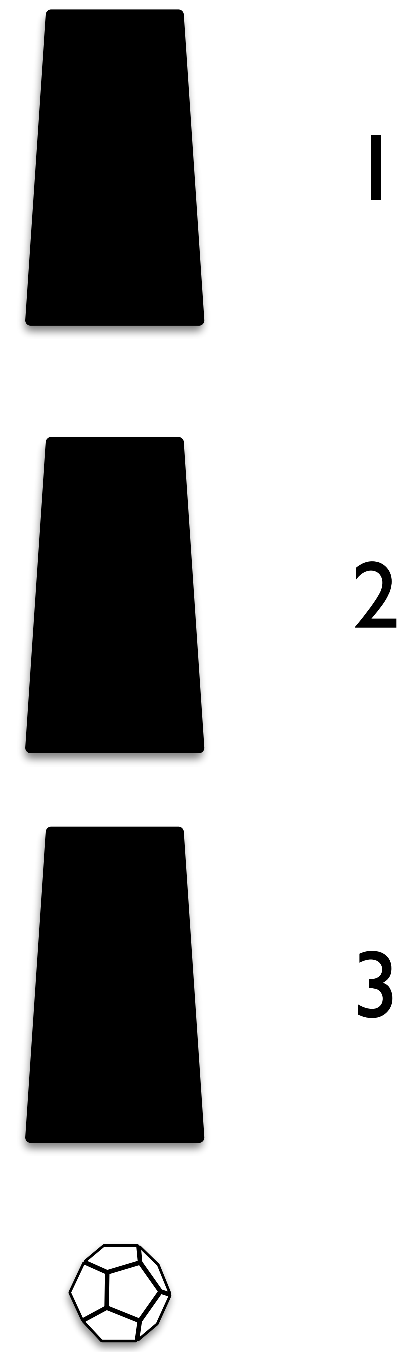
```
(Believes blinky (Loc ball 1))
```



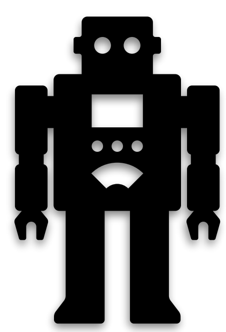


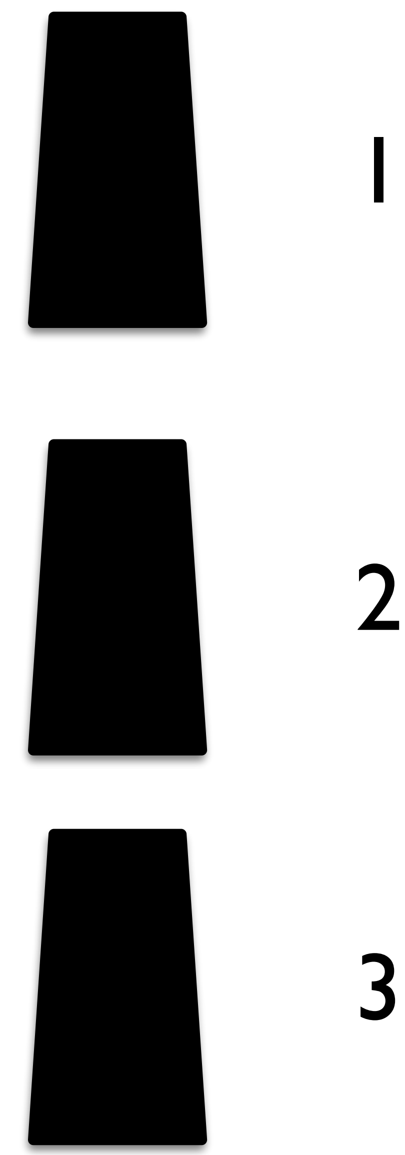
```
(Believes blinky (Loc ball 1))
```



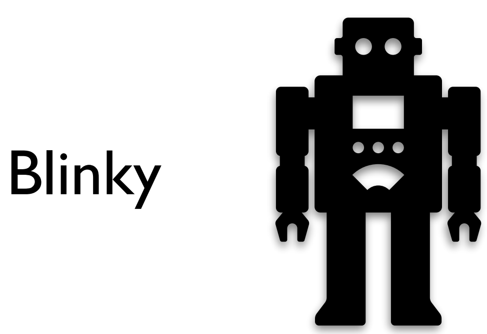


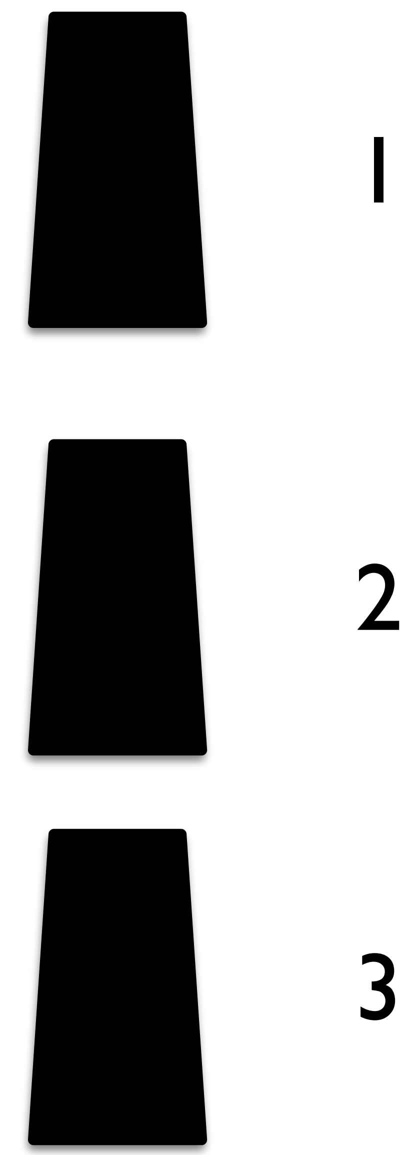
```
(Believes blinky (Loc ball 1))
```





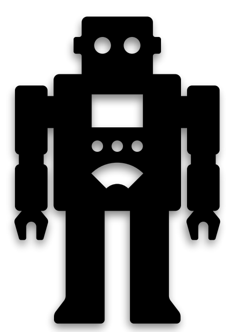
```
(Believes blinky (Loc ball 1))
```

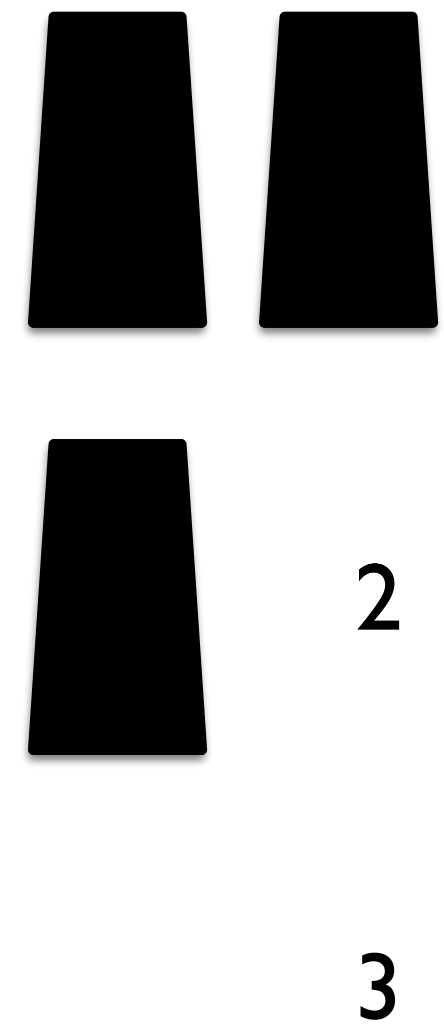




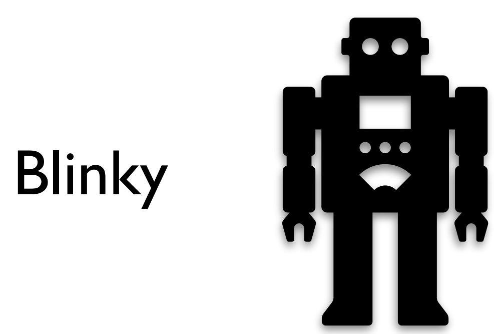
`(Believes blinky (Loc ball 1))`

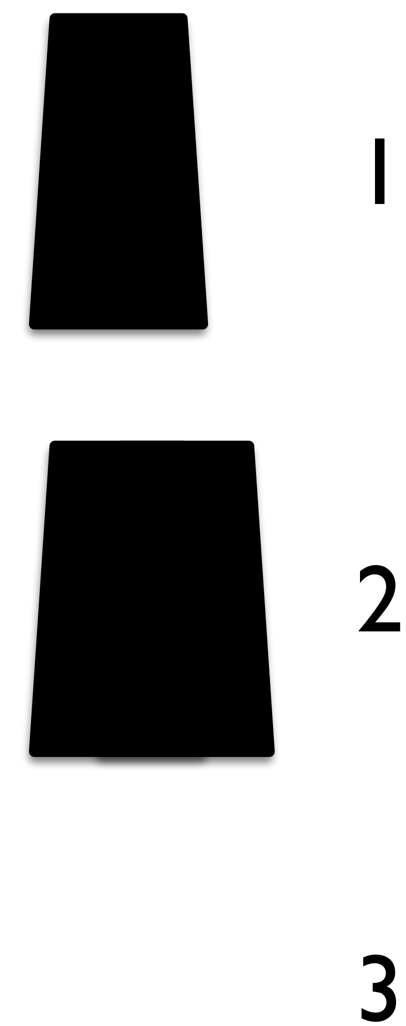
Blinky





(Believes blinky (Loc ball 1))

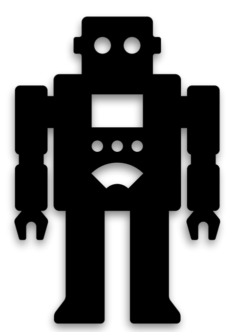


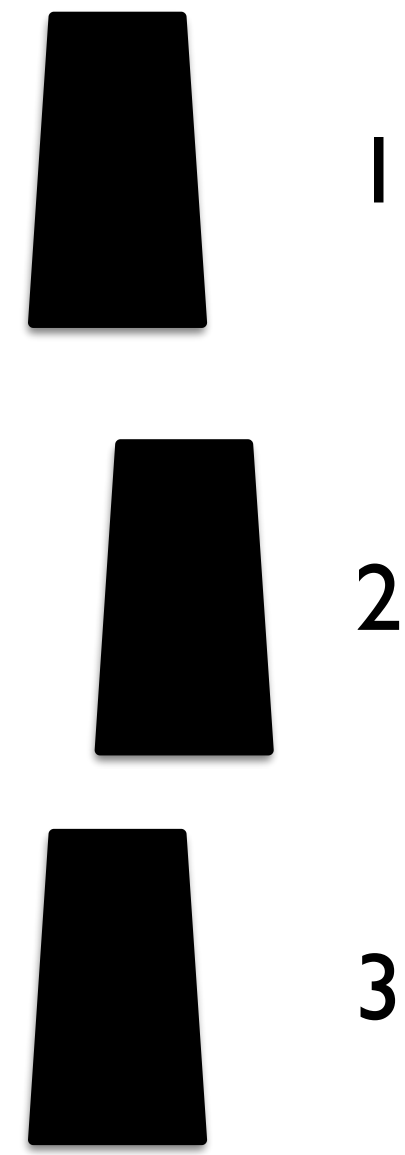


```
(Believes blinky (Loc ball 1))
```

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

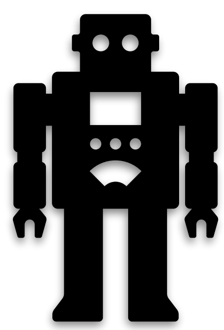
Blinky

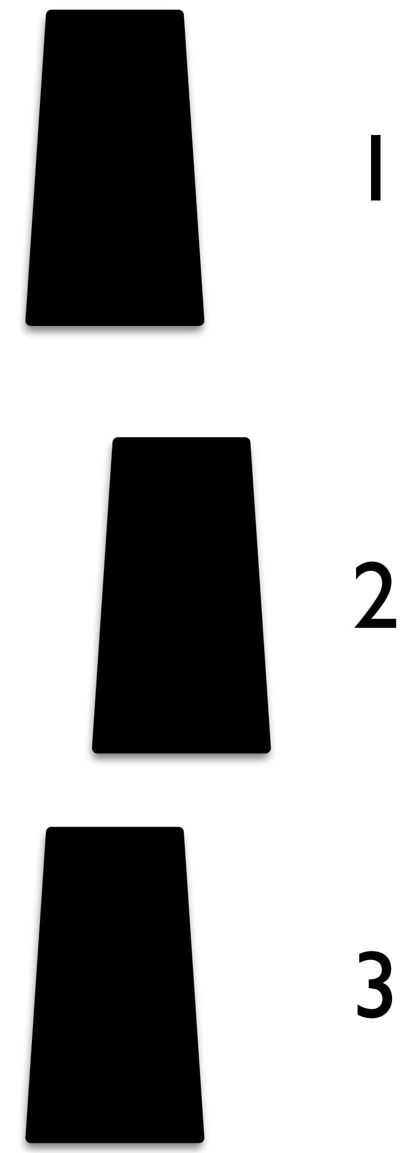




```
(Believes blinky (Loc ball 1))
```

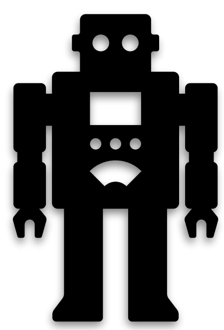
Blinky

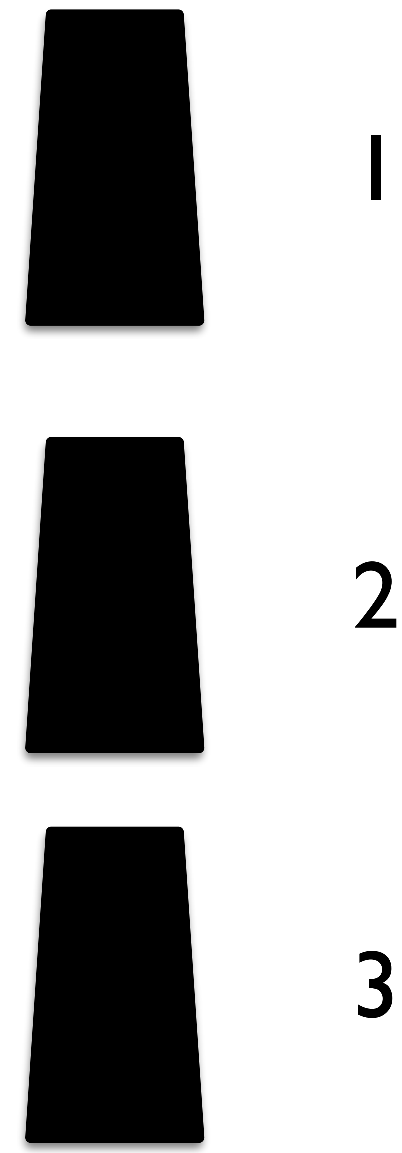




`(Believes blinky (Loc ball 1))`

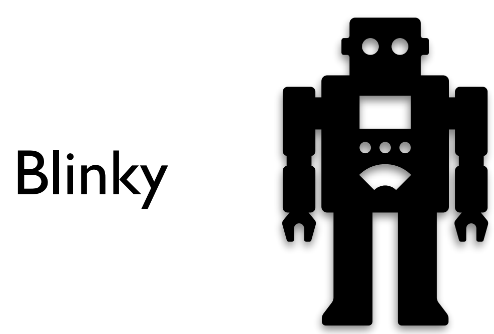
Blinky

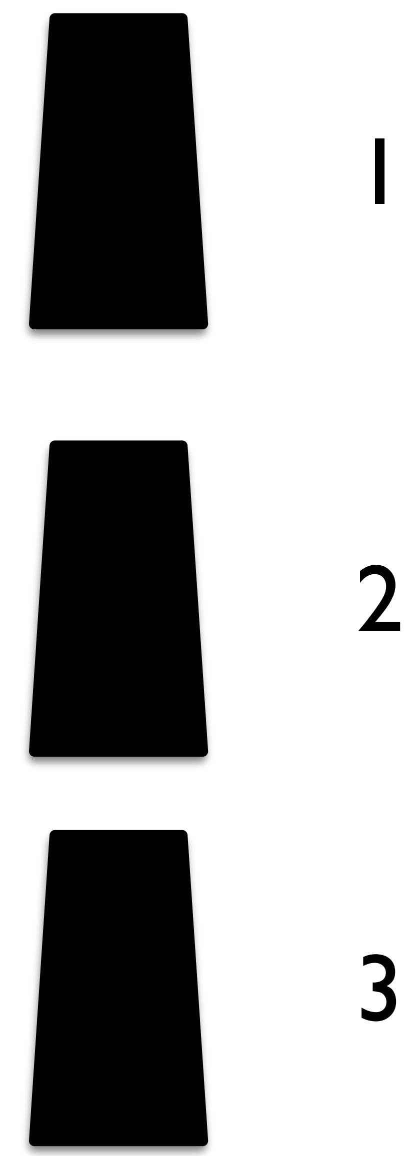




The ball is in the cup at location #1.

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



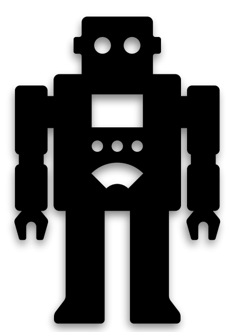


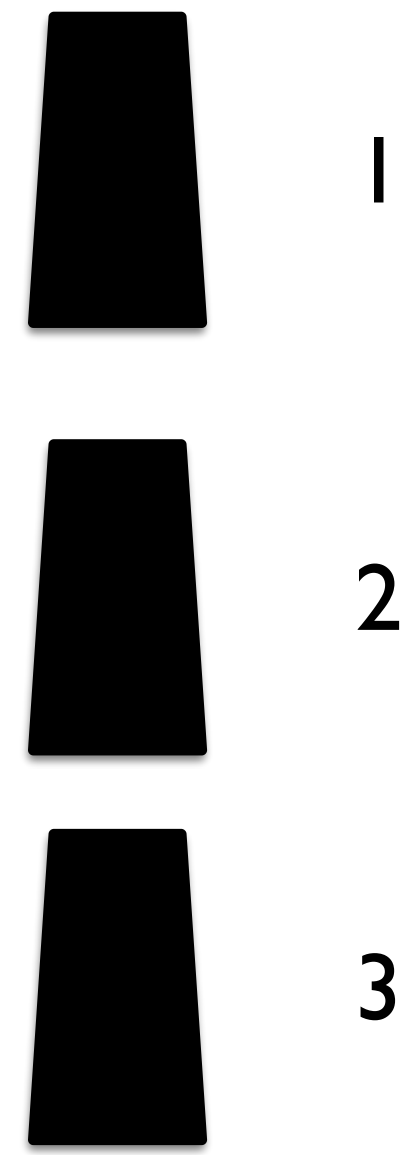
The ball is in the cup at location #1.

Loc(ball,1)

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

Blinky





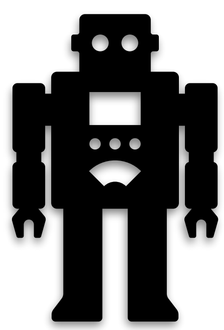
The ball is in the cup at location #1.

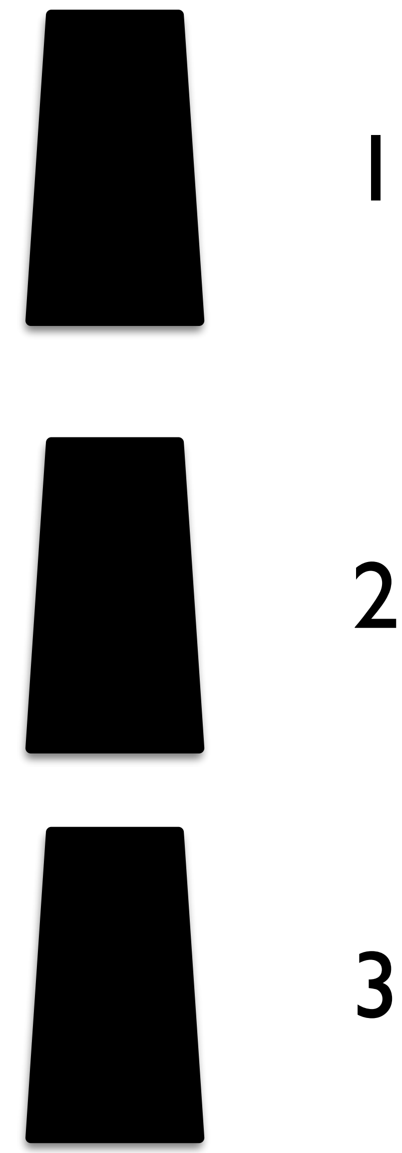
Loc(ball,1)

(Loc ball 1)

(Believes blinky (Loc ball 1))

Blinky





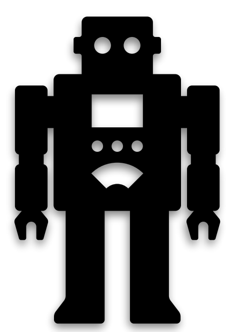
The ball is in the cup at location #1.

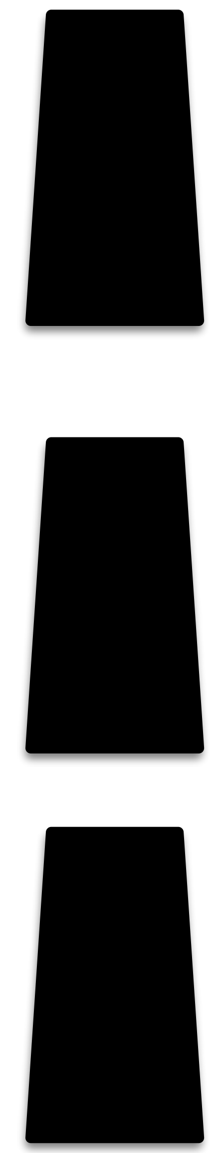
FALSE Loc(ball,1)

```
(Loc ball 1)
```

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

Blinky





1

2

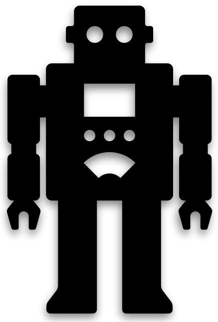
3

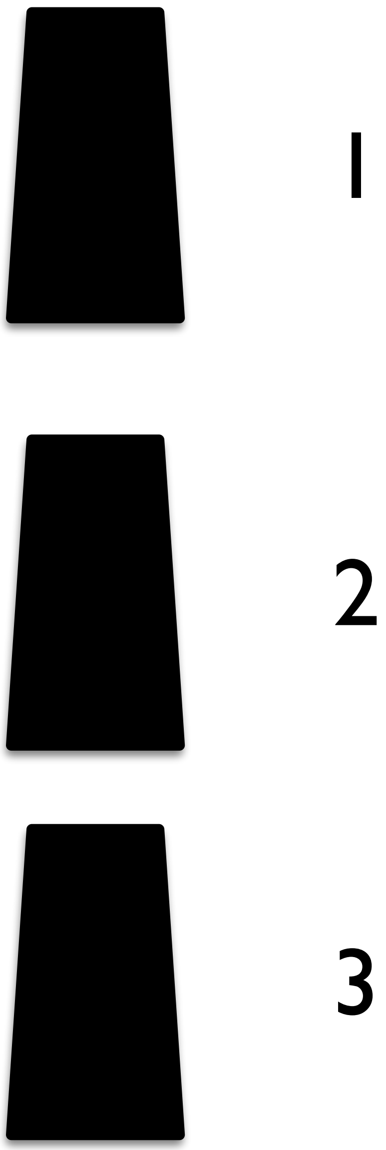
FALSE Loc(ball, 1)

`(Loc ball 1)`

`(Believes blinky (Loc ball 1))`

Blinky

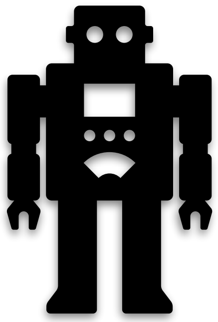


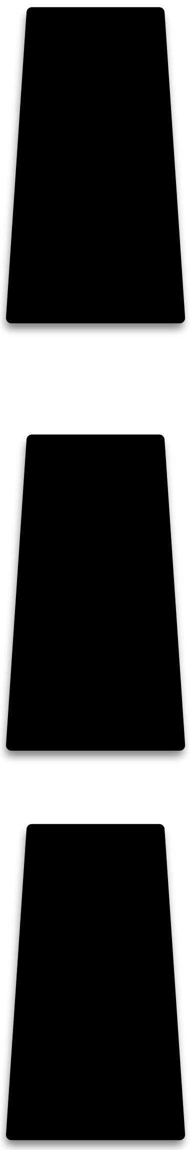


FALSE

```
(Loc ball 1)

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

Blinky





1

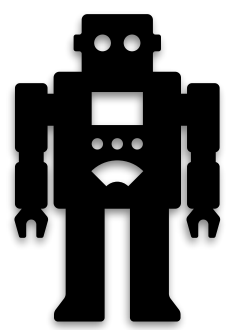
2

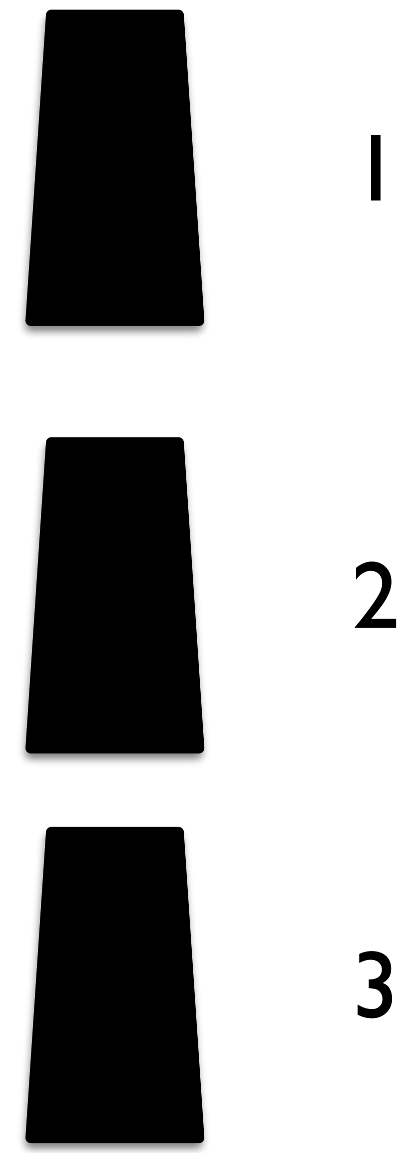
3

(Loc ball 1)

(Believes blinky (Loc ball 1))

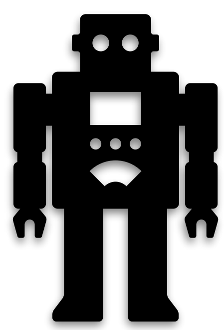
Blinky

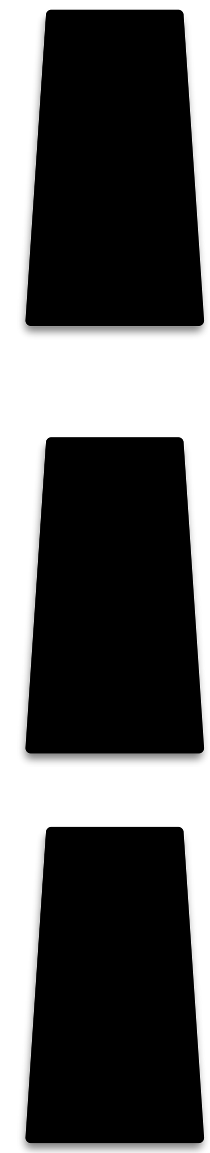




```
(Believes blinky (Loc ball 1))
```

Blinky





1

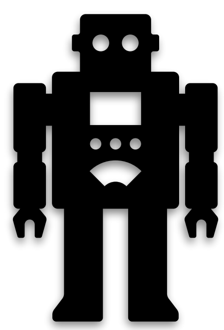
2

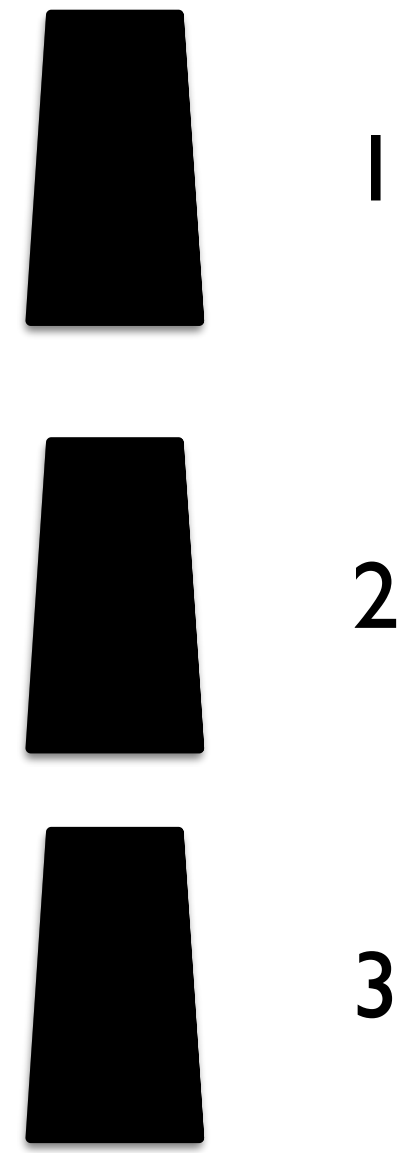
3

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

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

Blinky



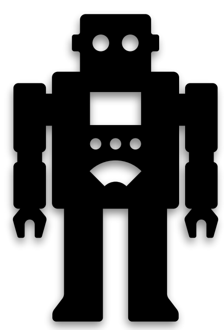


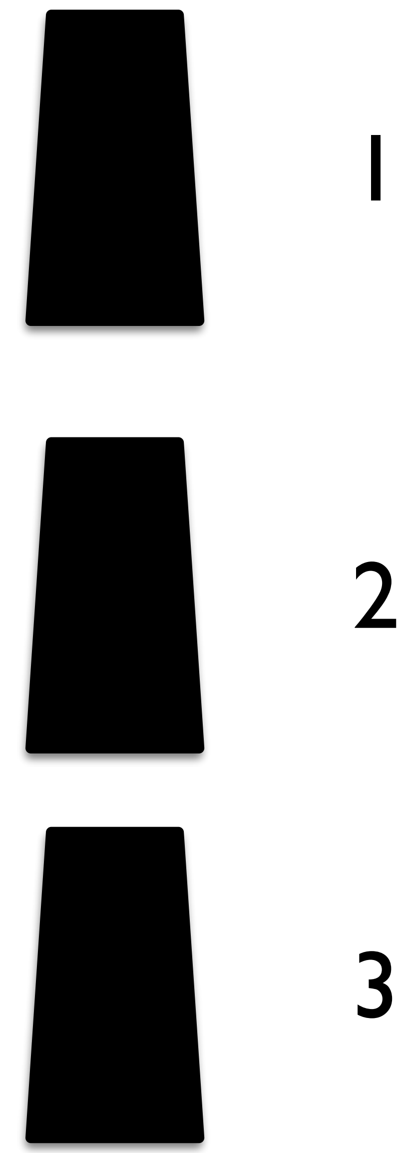
The ball is in the cup at location #1 or the ball is at location #3.

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

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

Blinky





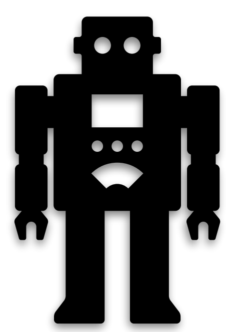
The ball is in the cup at location #1 or the ball is at location #3.

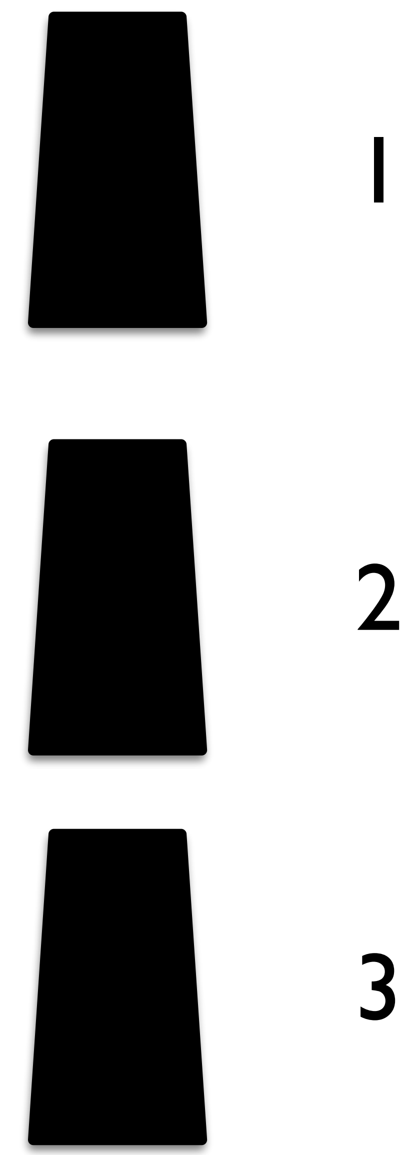
$$\text{Loc}(\text{ball}, 1) \lor \text{Loc}(\text{ball}, 3)$$

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

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

Blinky





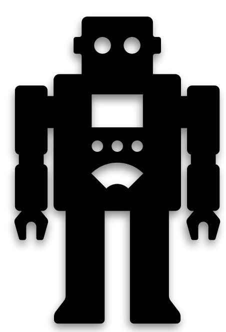
The ball is in the cup at location #1 or the ball is at location #3.

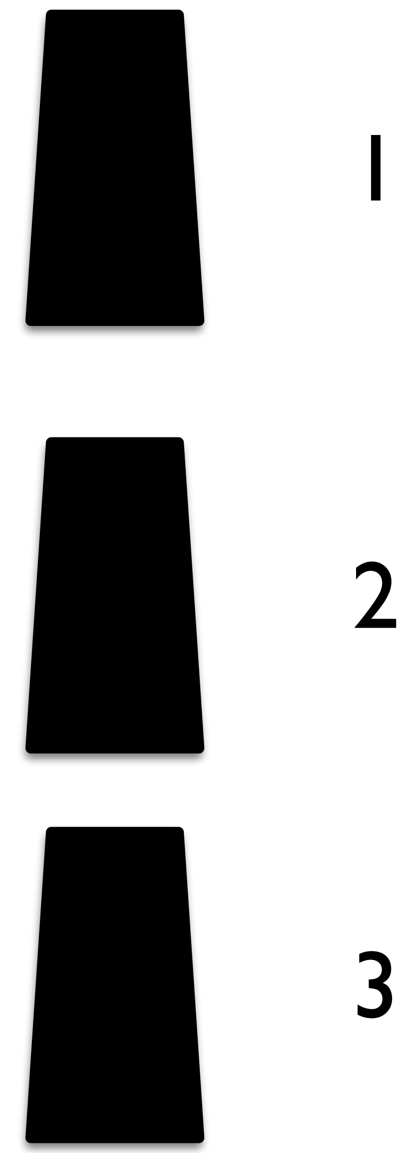
FALSE $Loc(ball,1) \vee Loc(ball,3)$

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

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

Blinky

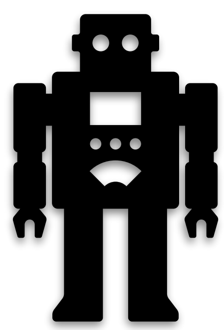


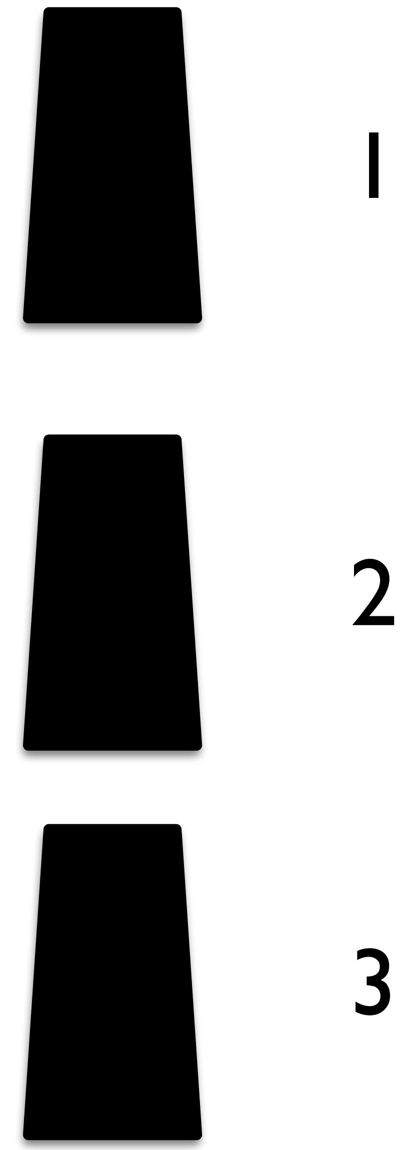


FALSE $\mathrm{Loc}(\mathrm{ball},1) \vee \mathrm{Loc}(\mathrm{ball},3)$

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

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

Blinky



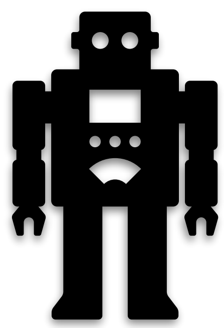


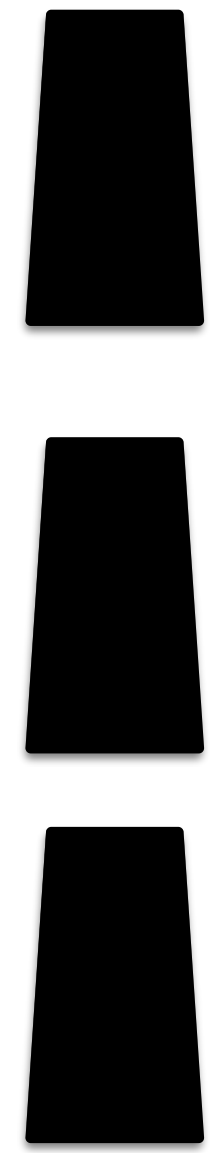
FALSE

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

`(Believes blinky (Loc ball 1))`

Blinky





1

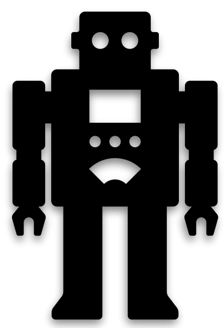
2

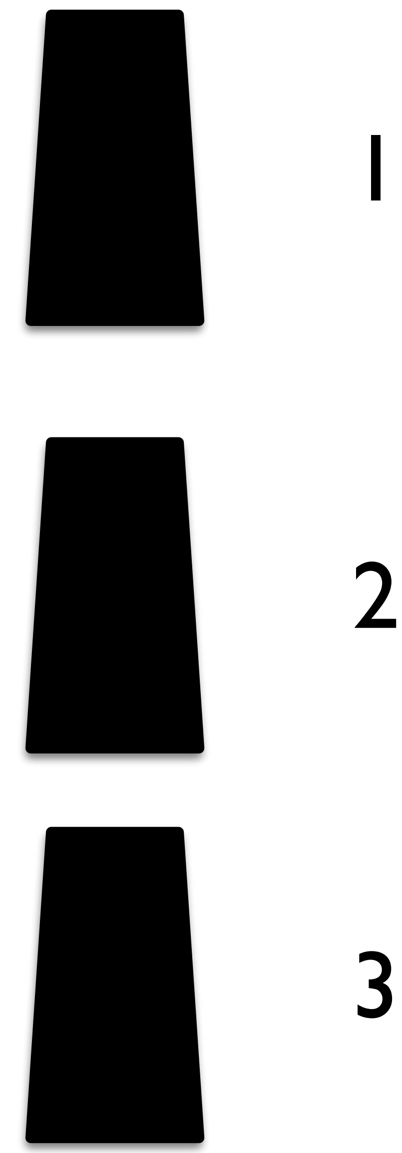
3

FALSE

(Believes blinky (Loc ball 1))

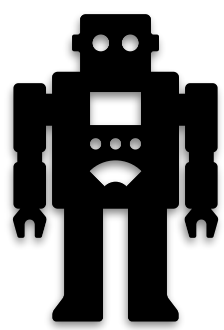
Blinky

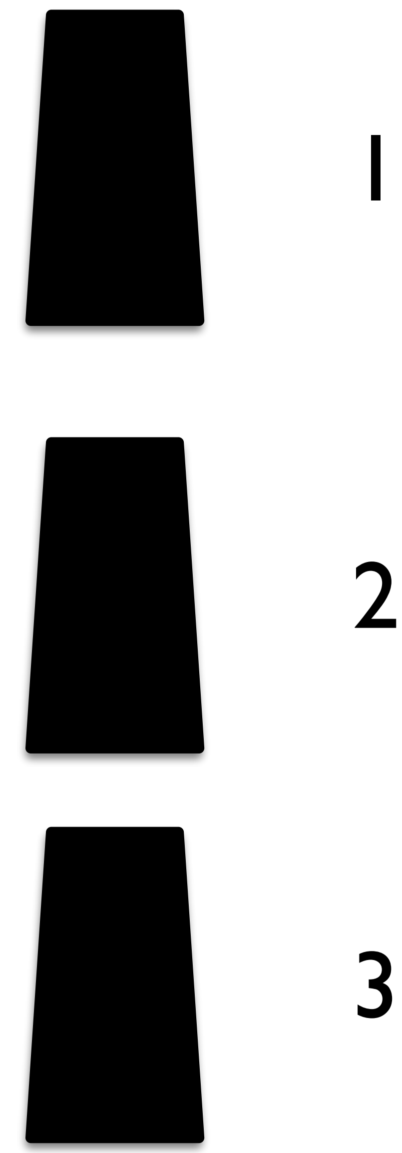




```
(Believes blinky (Loc ball 1))
```

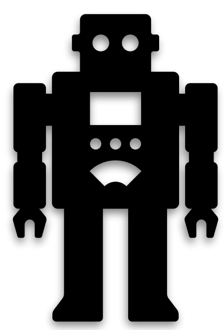
Blinky

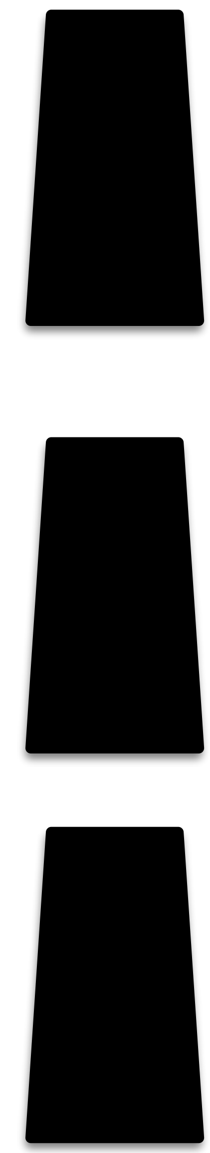




Blinky believes that the ball is in the cup at location #1.

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

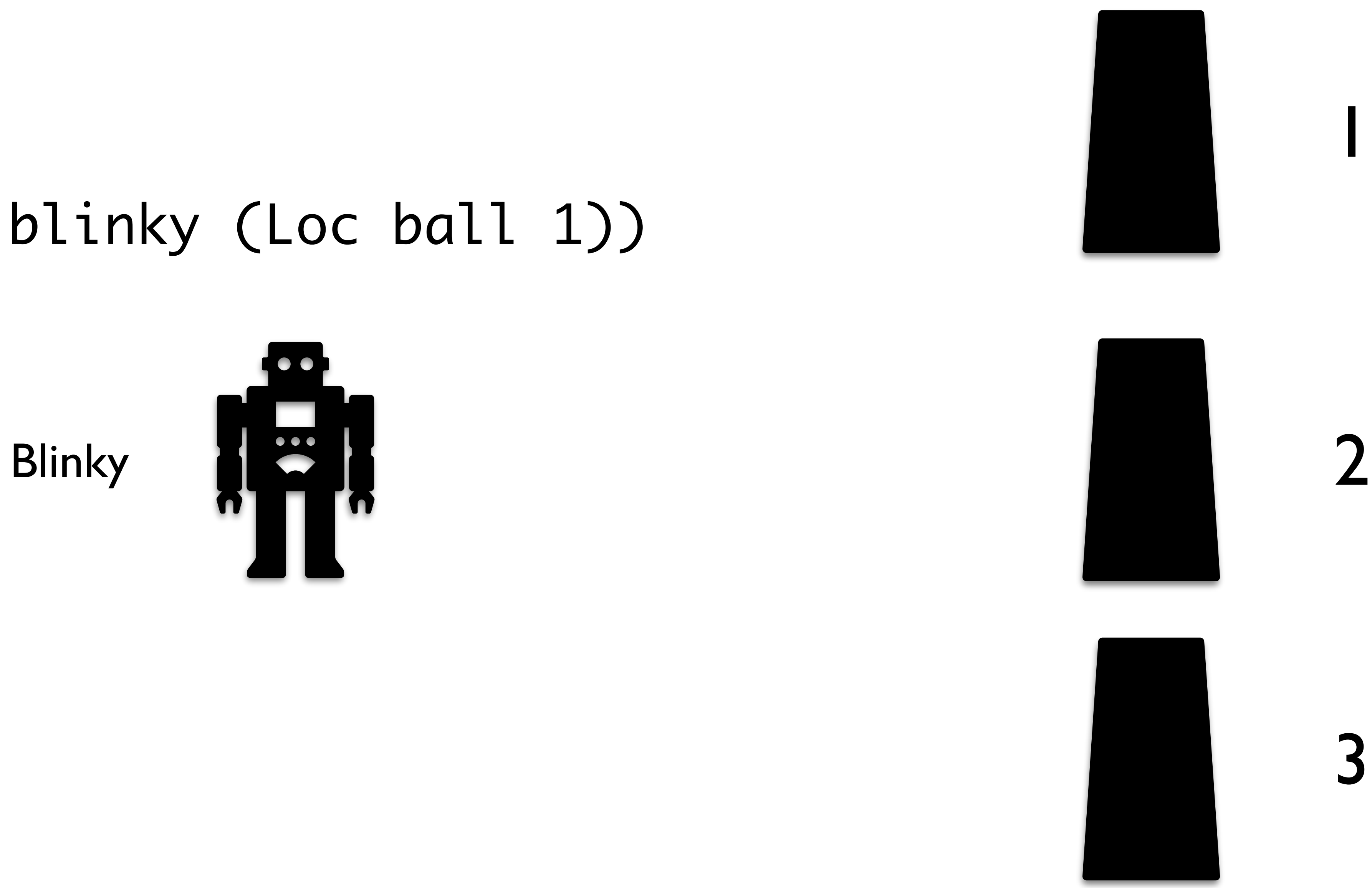




Blinky believes that the ball is in the cup at location #1.

$$B(blinky, Loc(ball, 1))$$

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



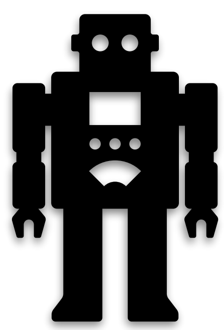
Blinky believes that the ball is in the cup at location #1.

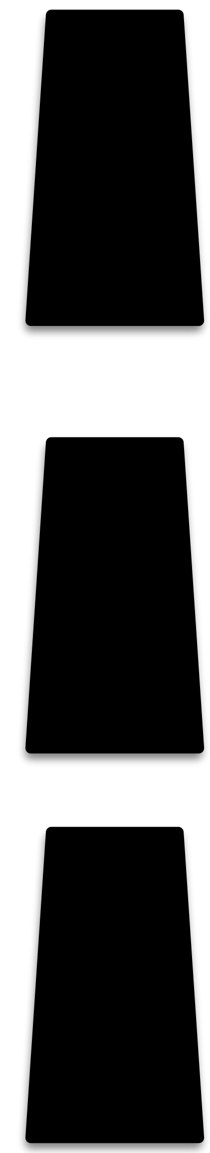
$$B(blinky, Loc(ball, 1))$$

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

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

Blinky





1

2

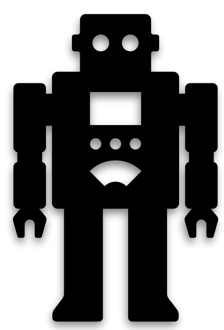
3

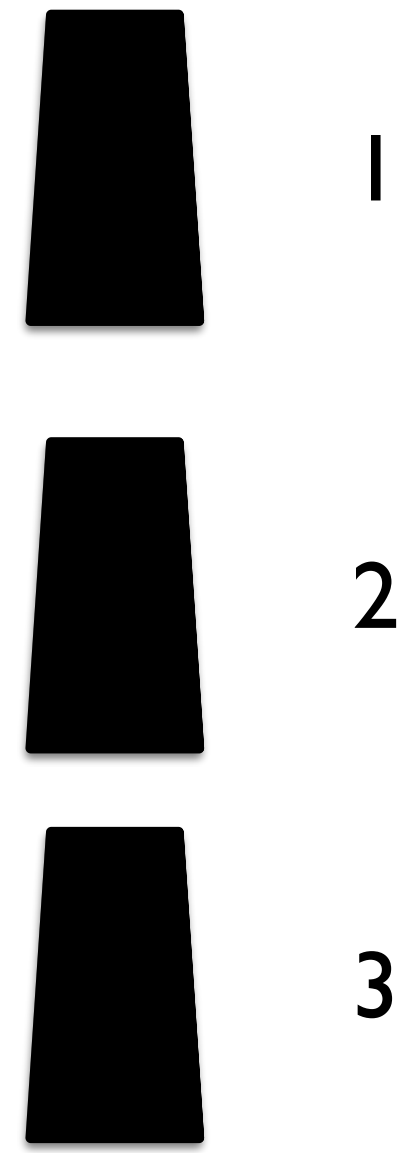
Blinky believes that the ball is in the cup at location #1.

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

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

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

Blinky



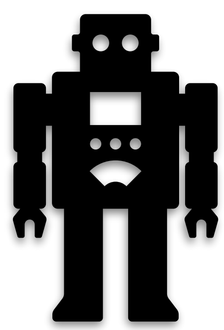


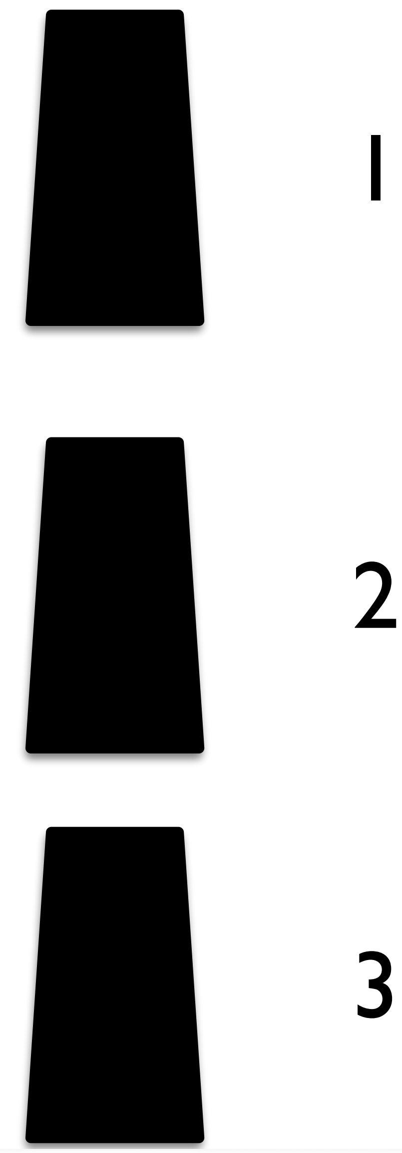
Blinky believes that the ball is in the cup at location #1.

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

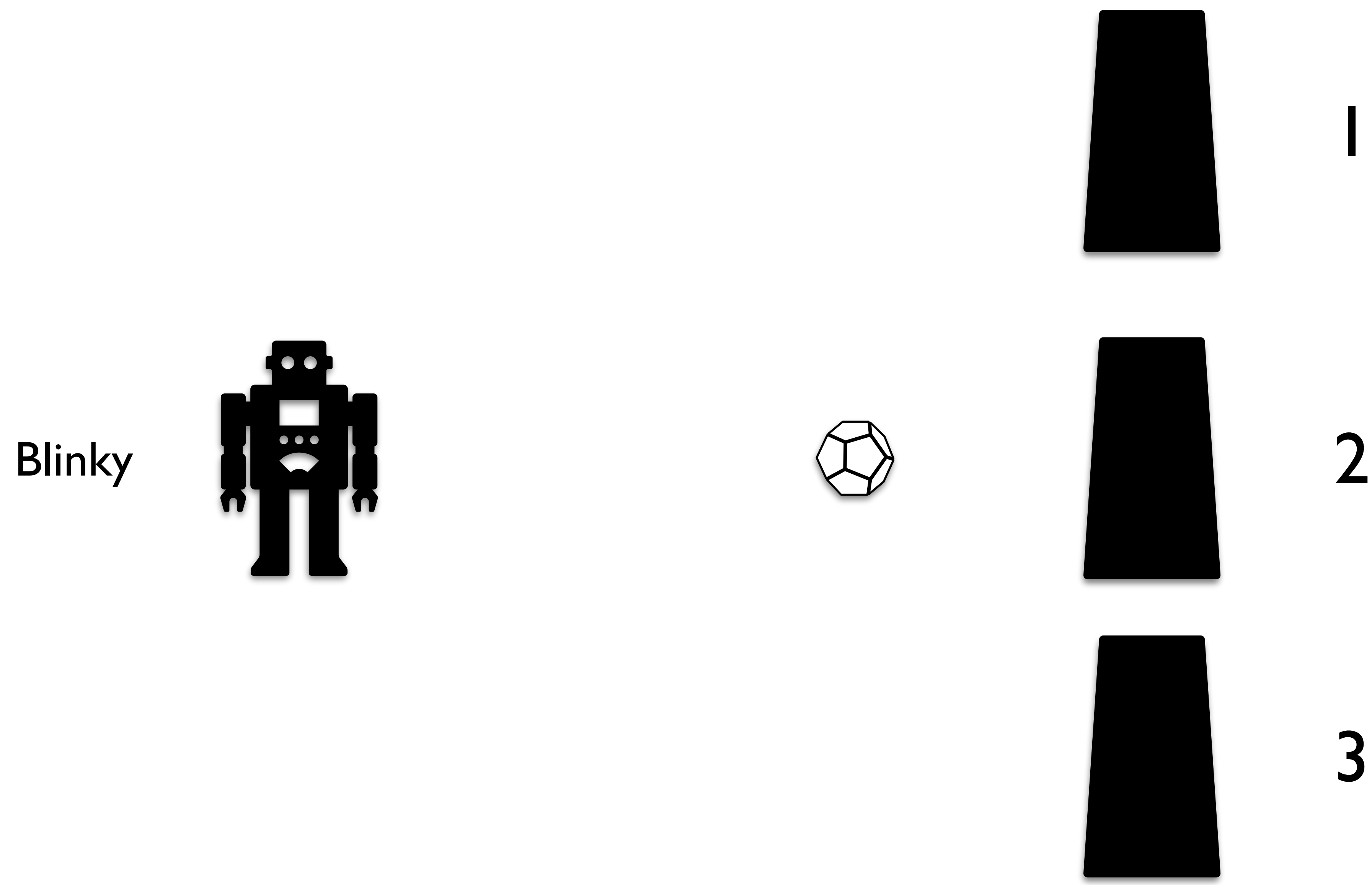
```
(Believes! blinky (Loc ball 1))

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

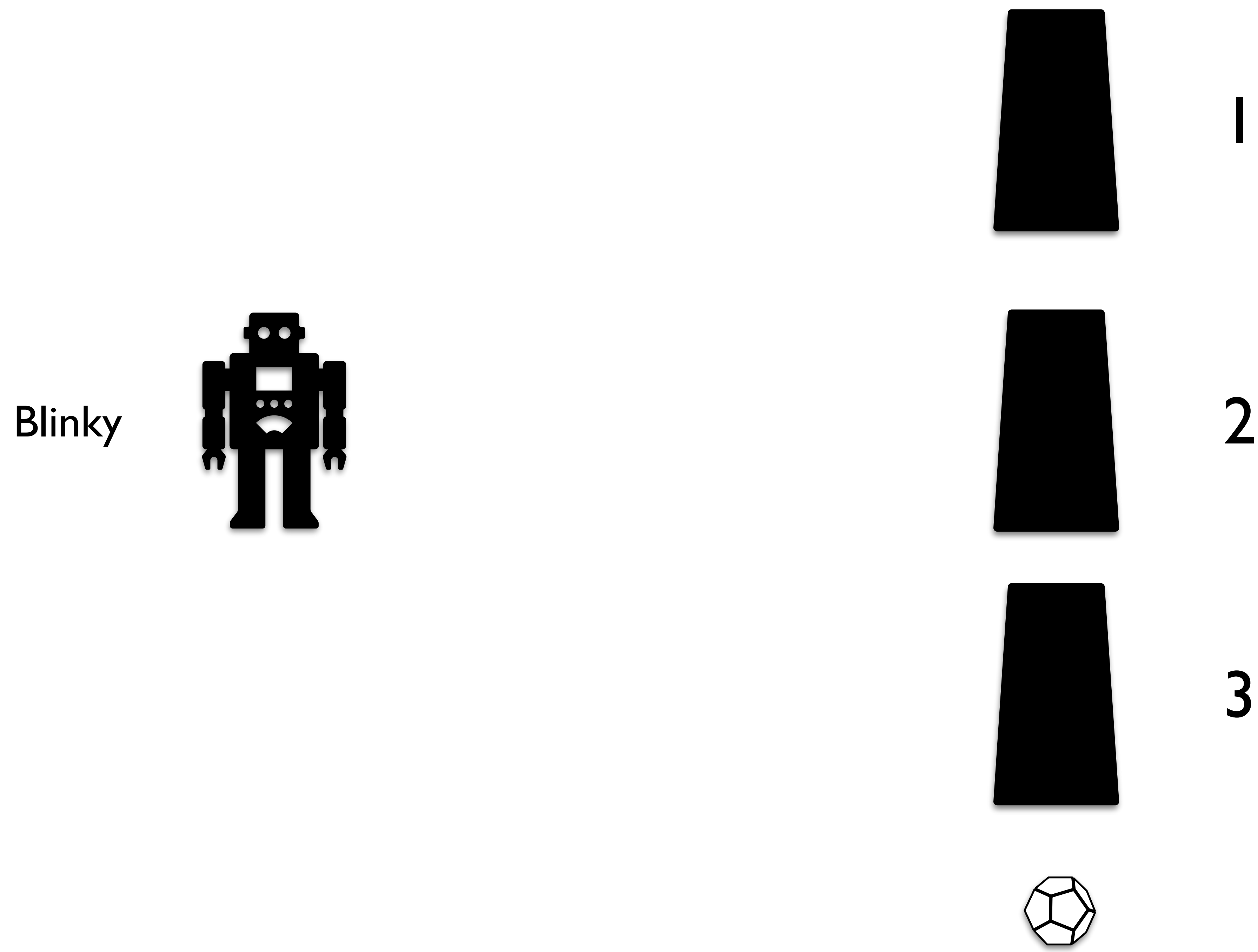
Blinky

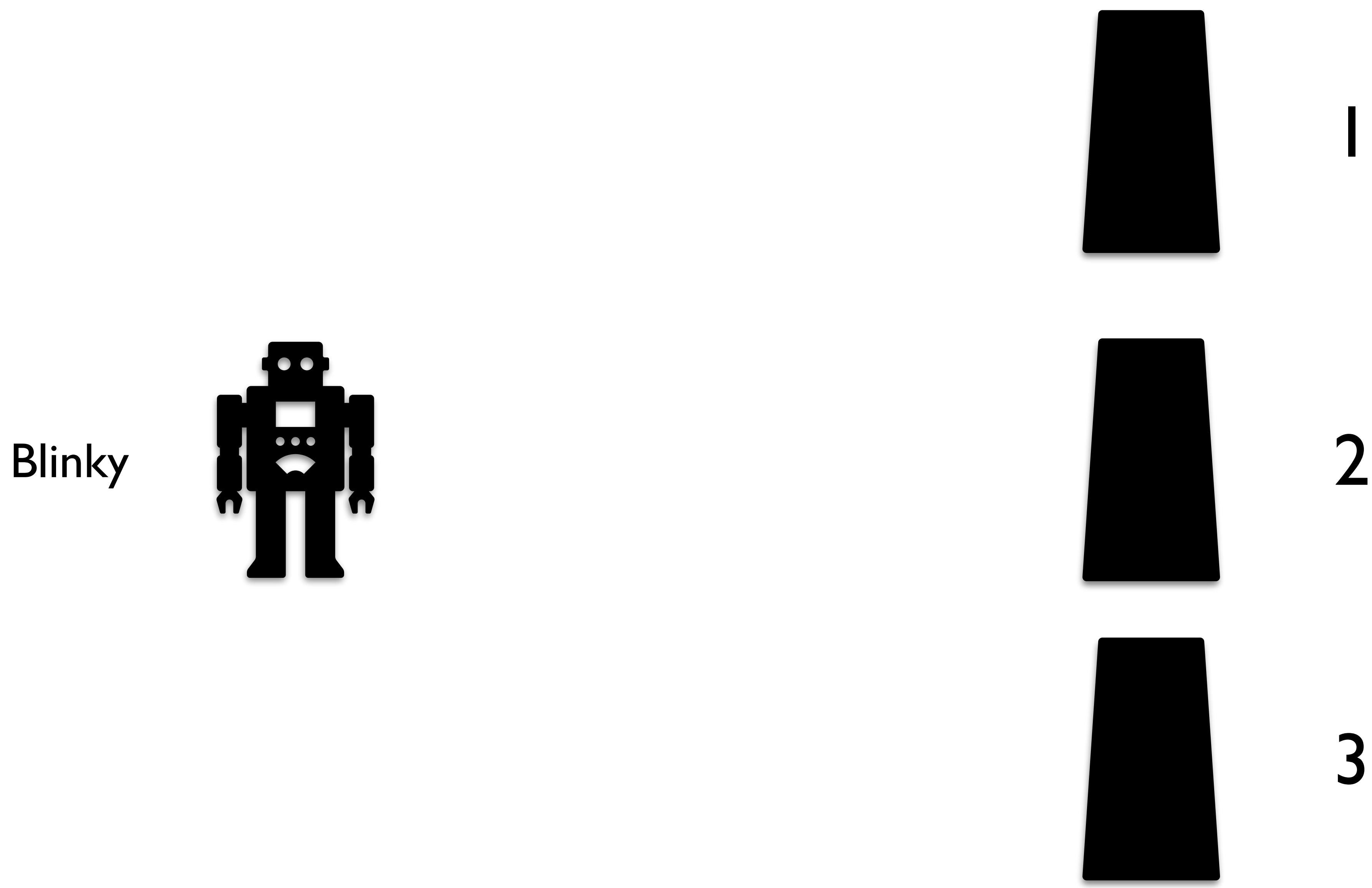


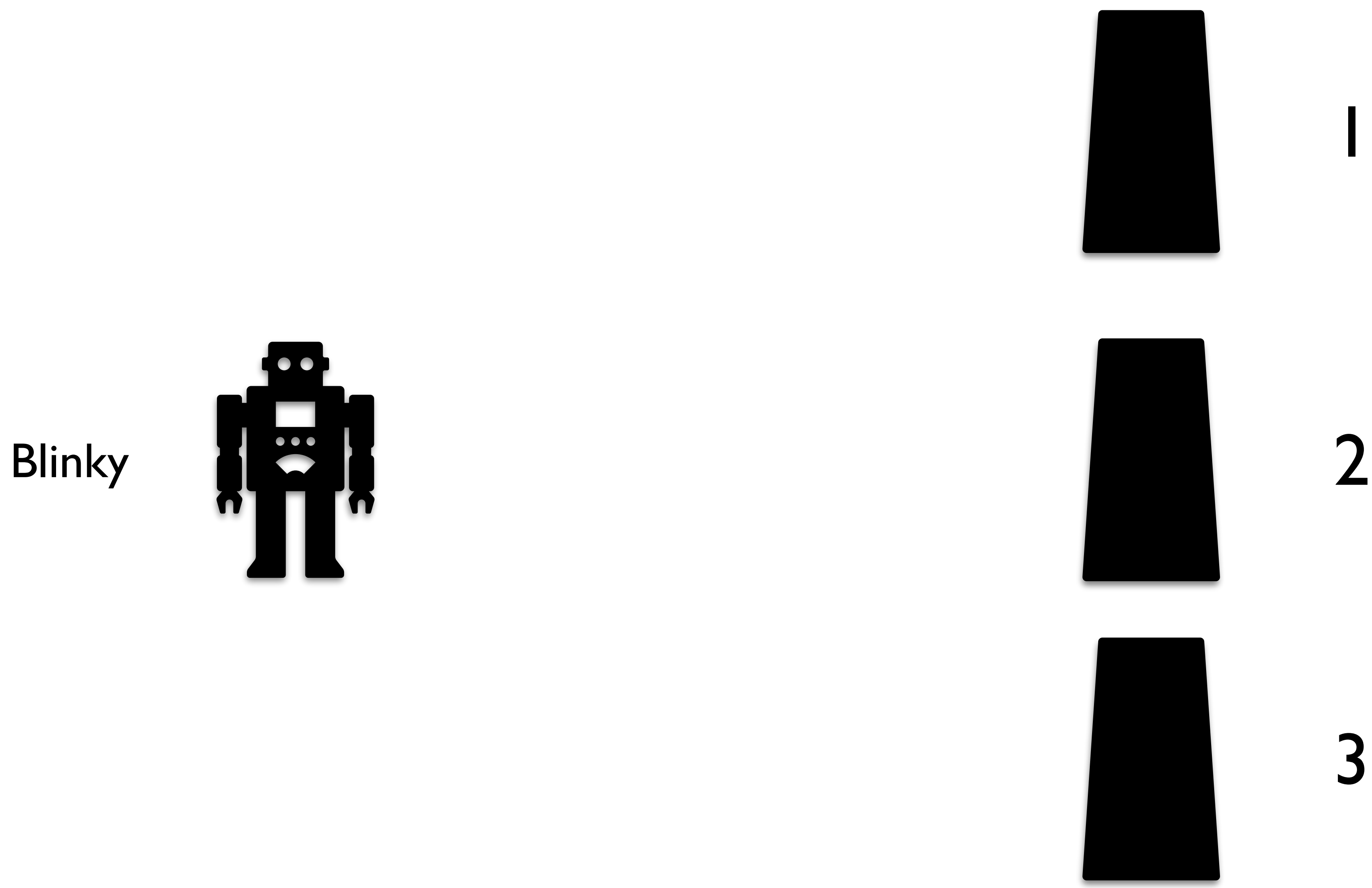


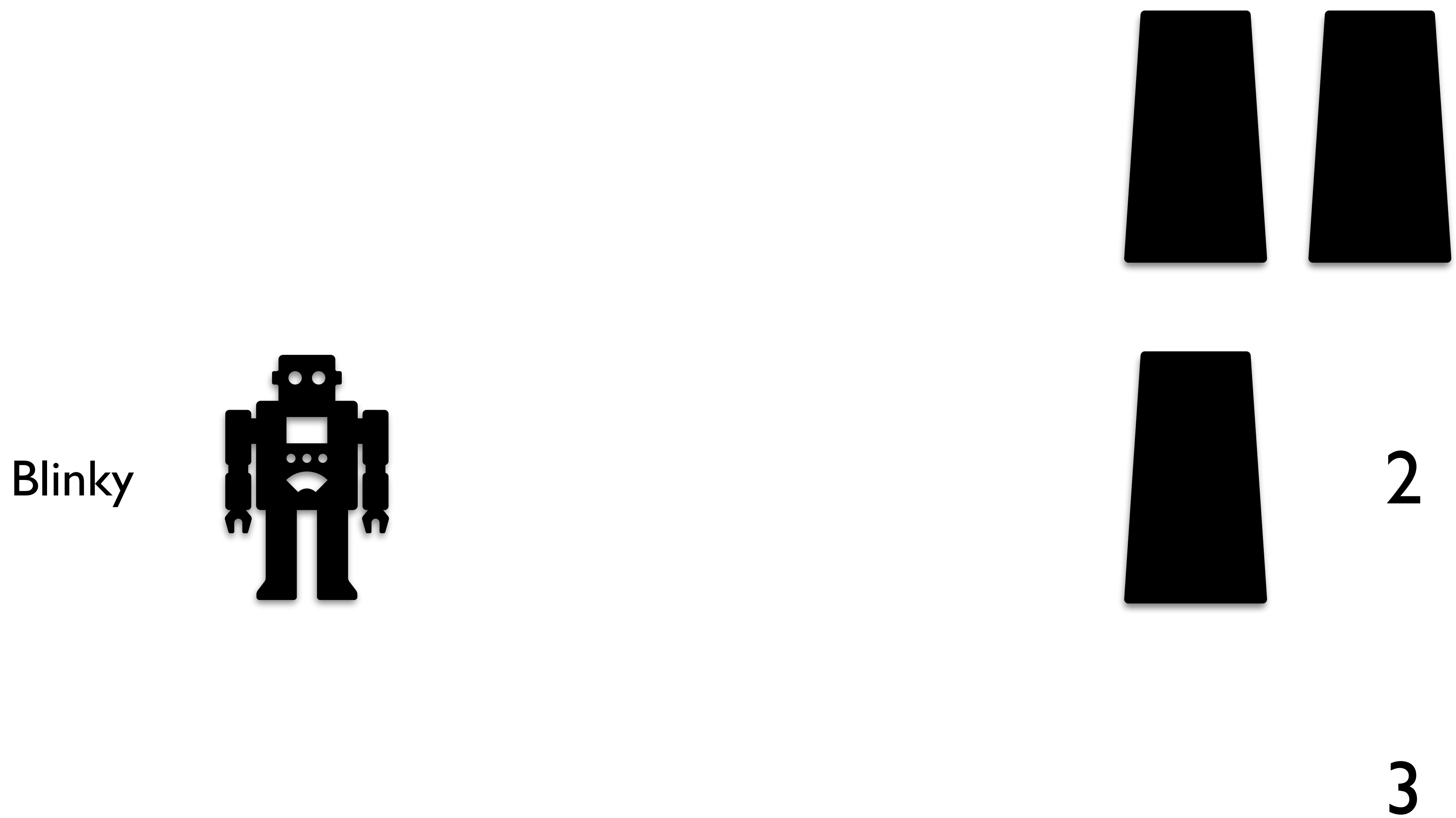
In extensional logics, what is denoted is conflated with meaning (the latter being naïvely compositional), but intensional attitudes like *believes*, *knows*, *hopes*, *fears*, etc cannot be represented and reasoned over smoothly (e.g. without fear of inconsistency rising up).

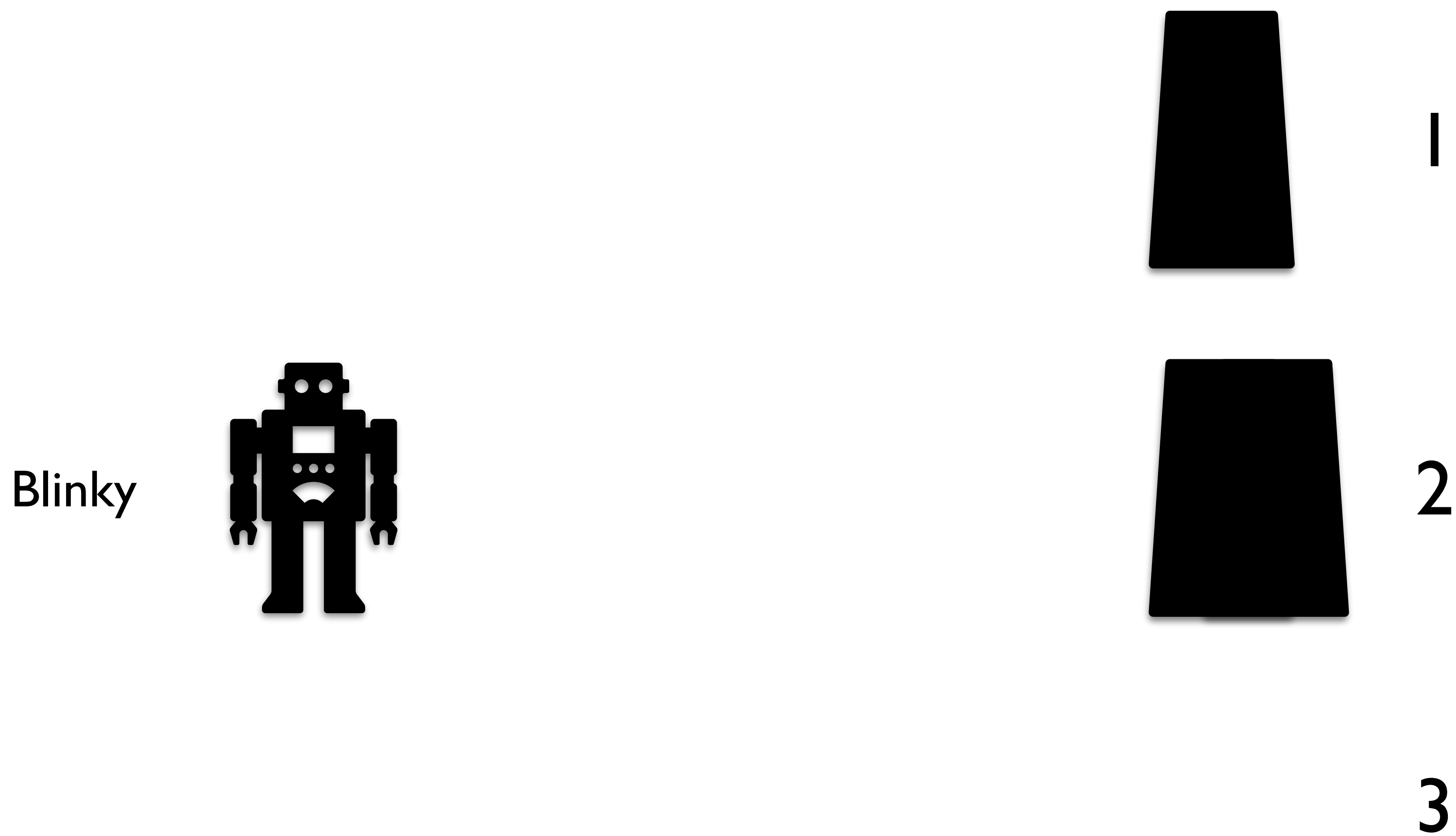

Blinky
1
2
3

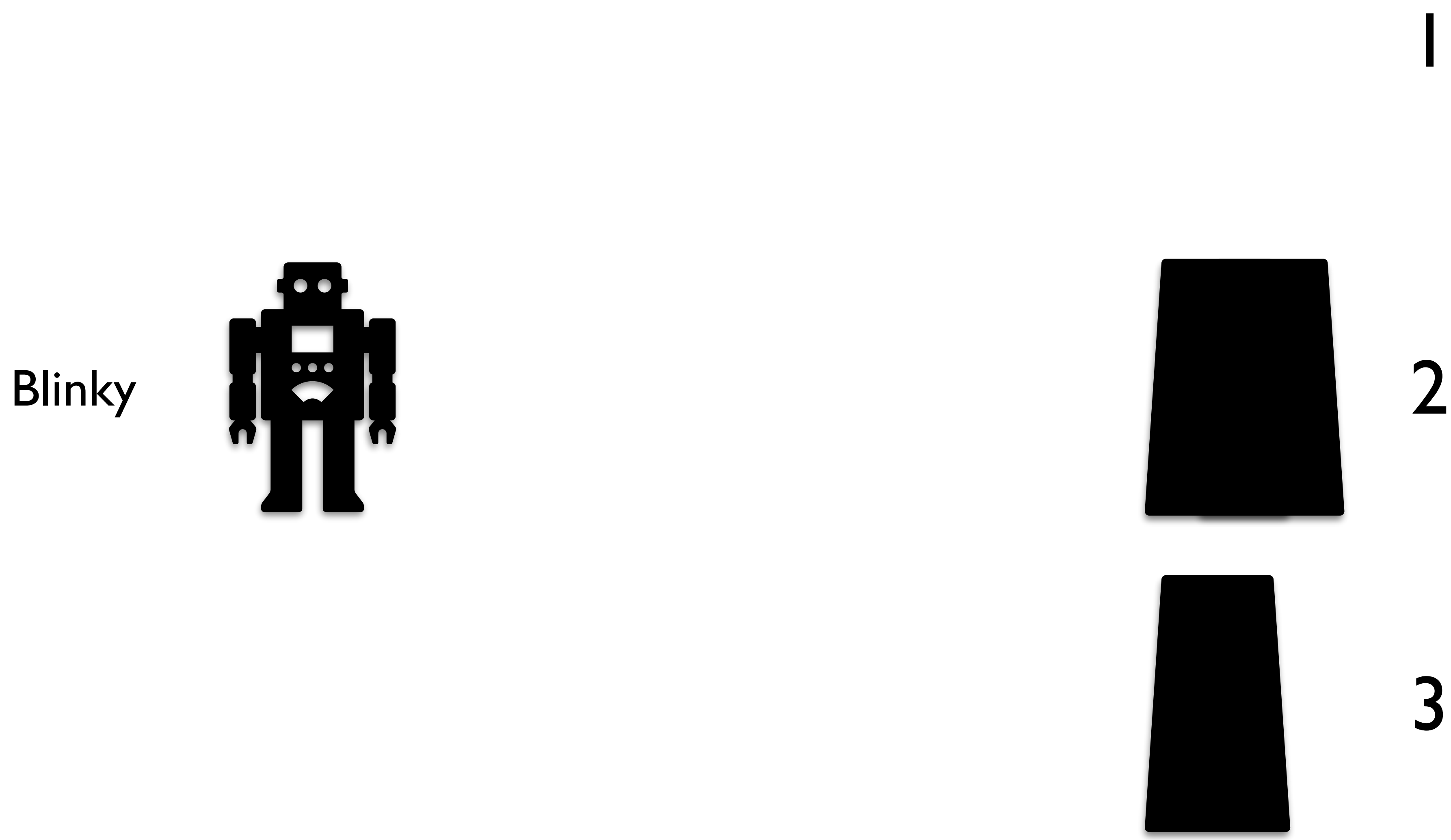

Blinky
1
2
3

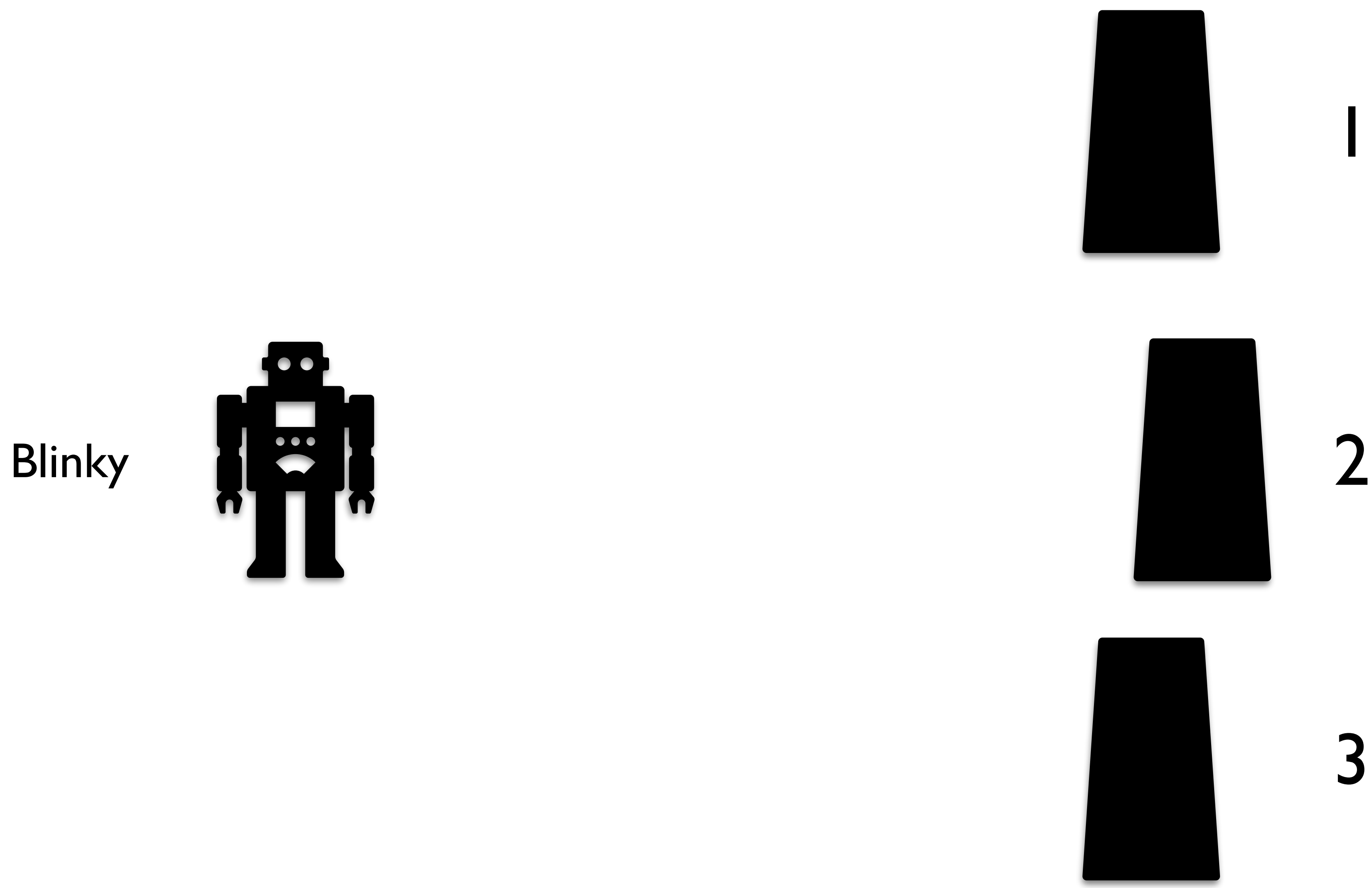

Blinky
1
2
3

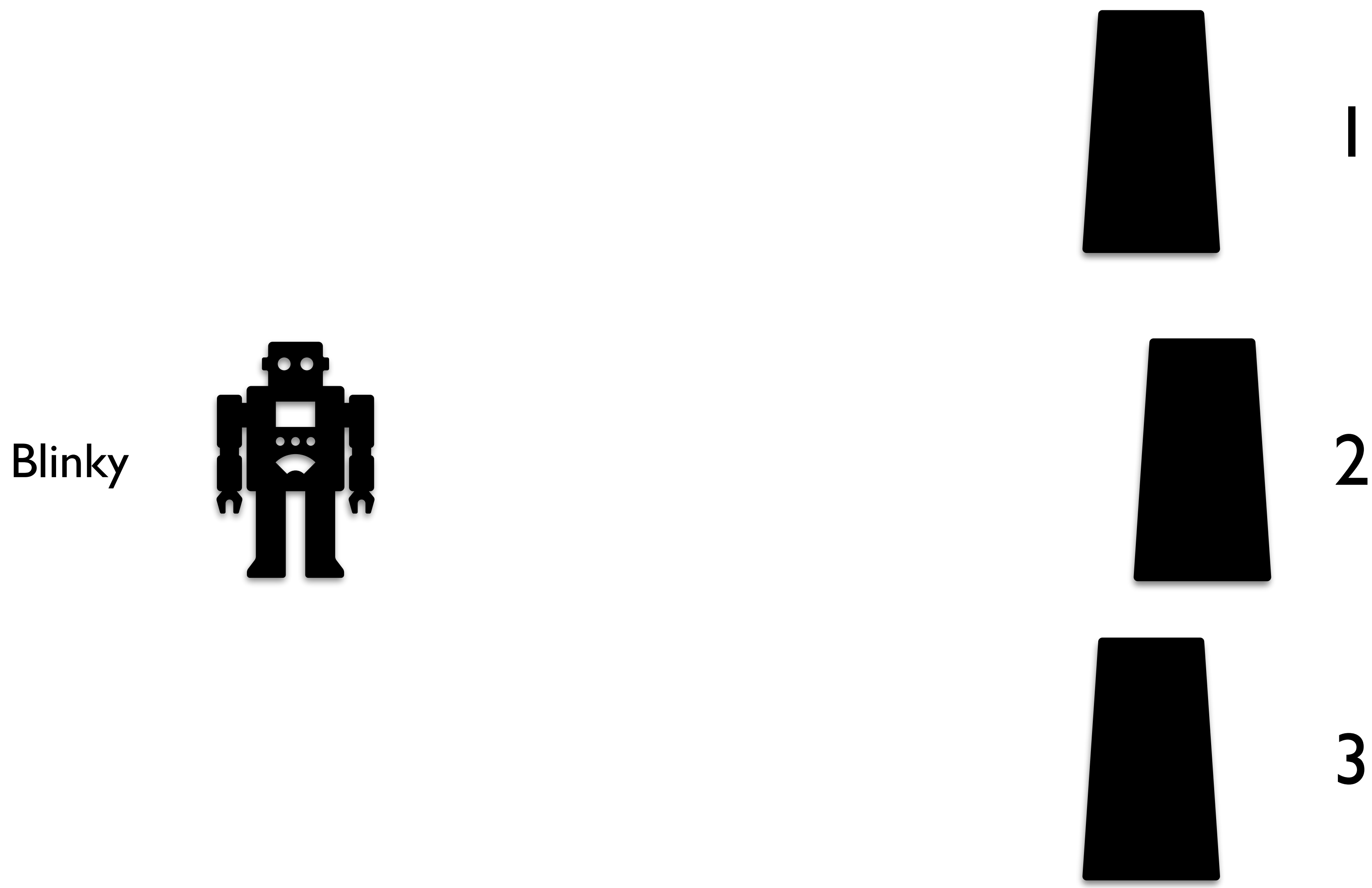

Blinky
1
2
3

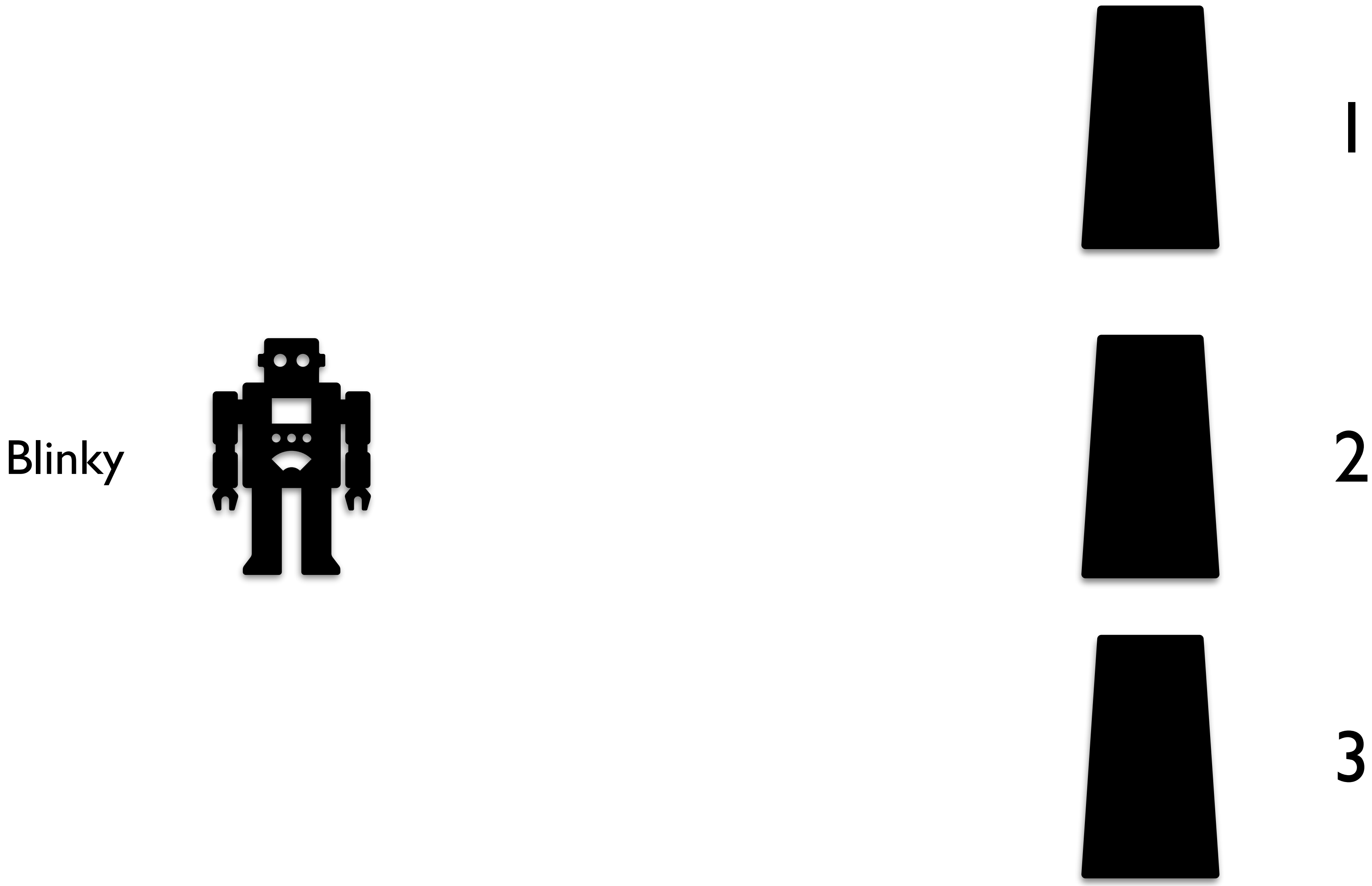

Blinky
2
3


Blinky
1
2
3


Blinky
1
2
3

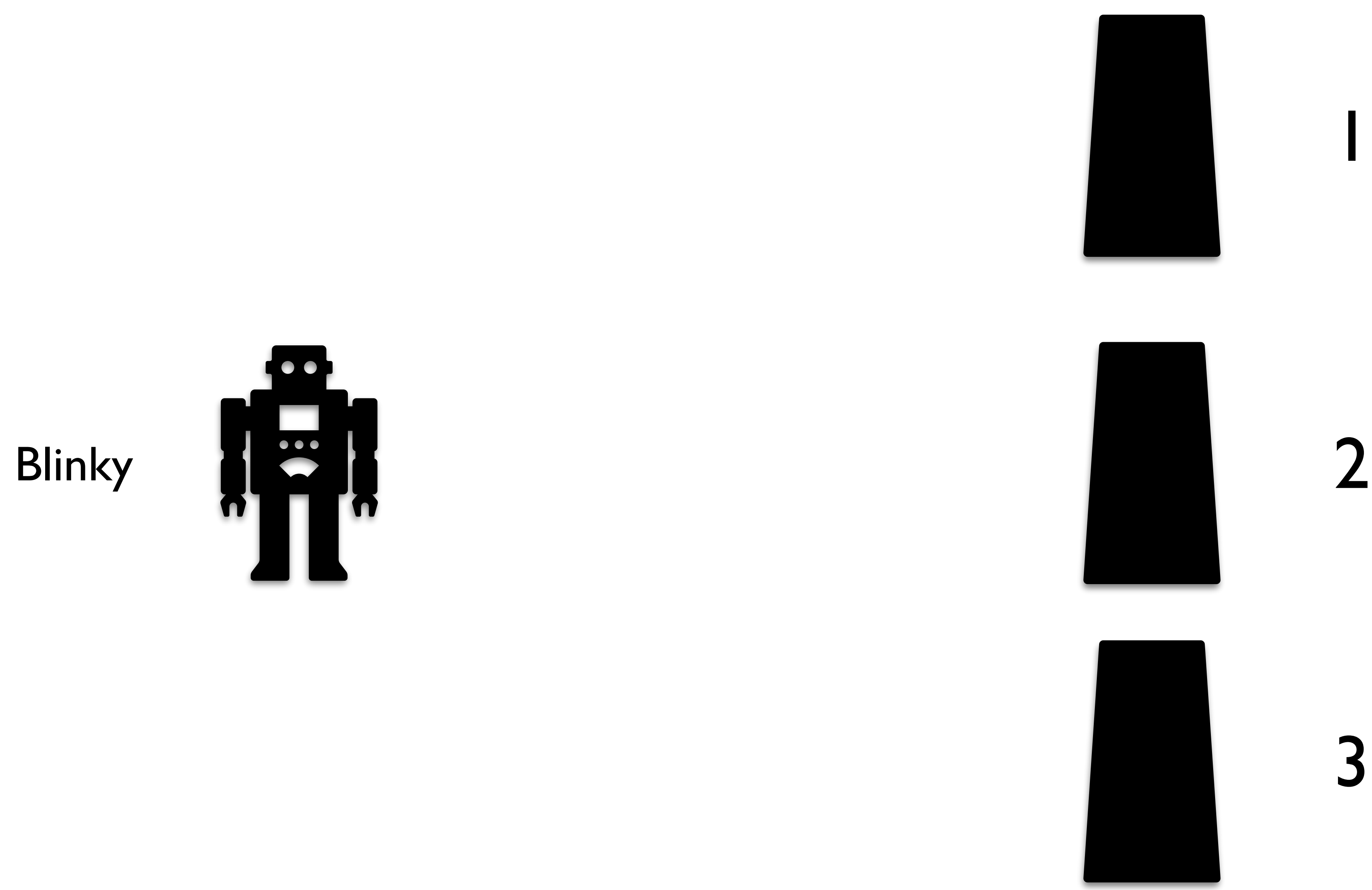

Blinky
1
2
3


Blinky
1
2
3



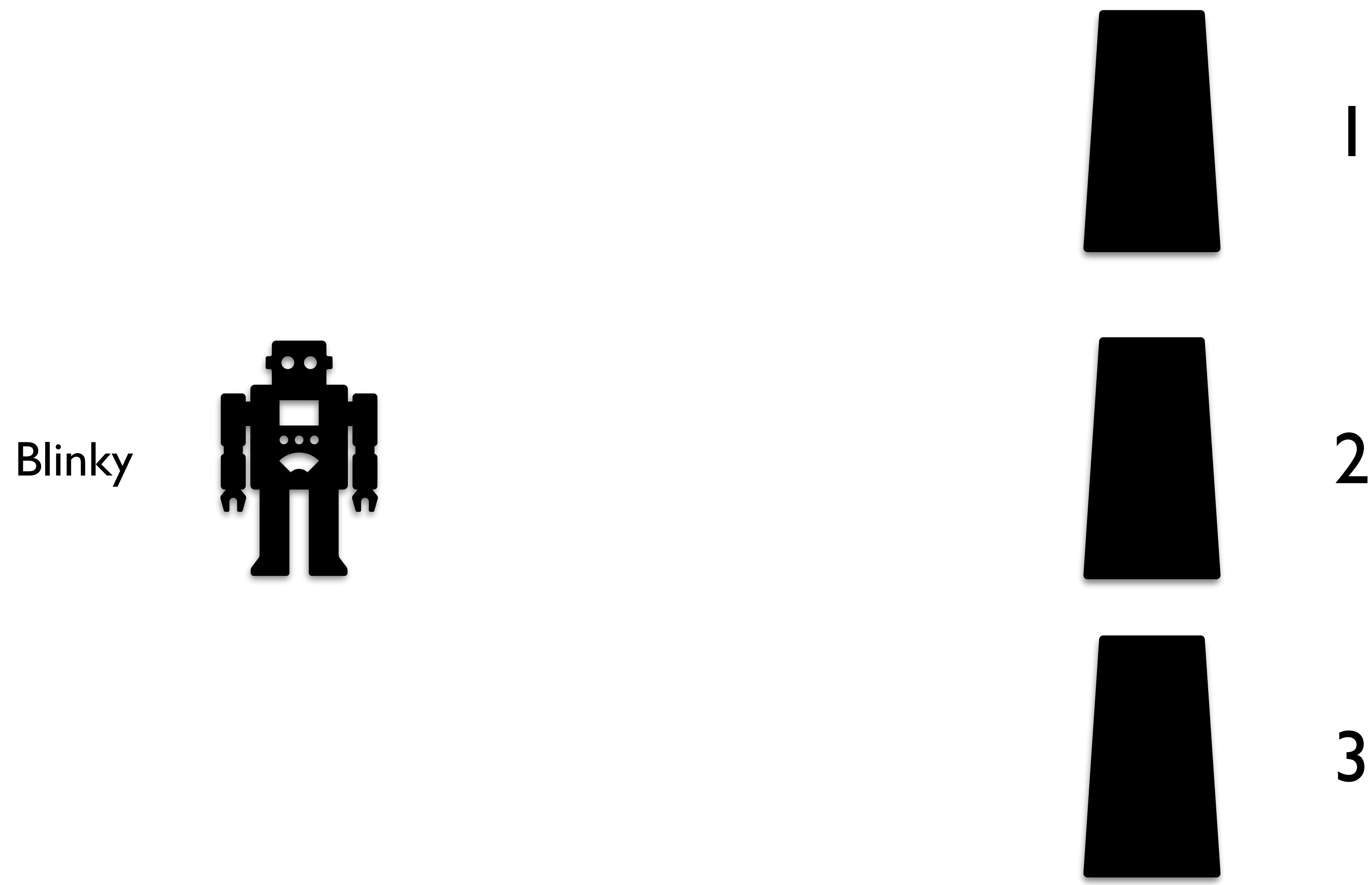
Blinky
1
2
3

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



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

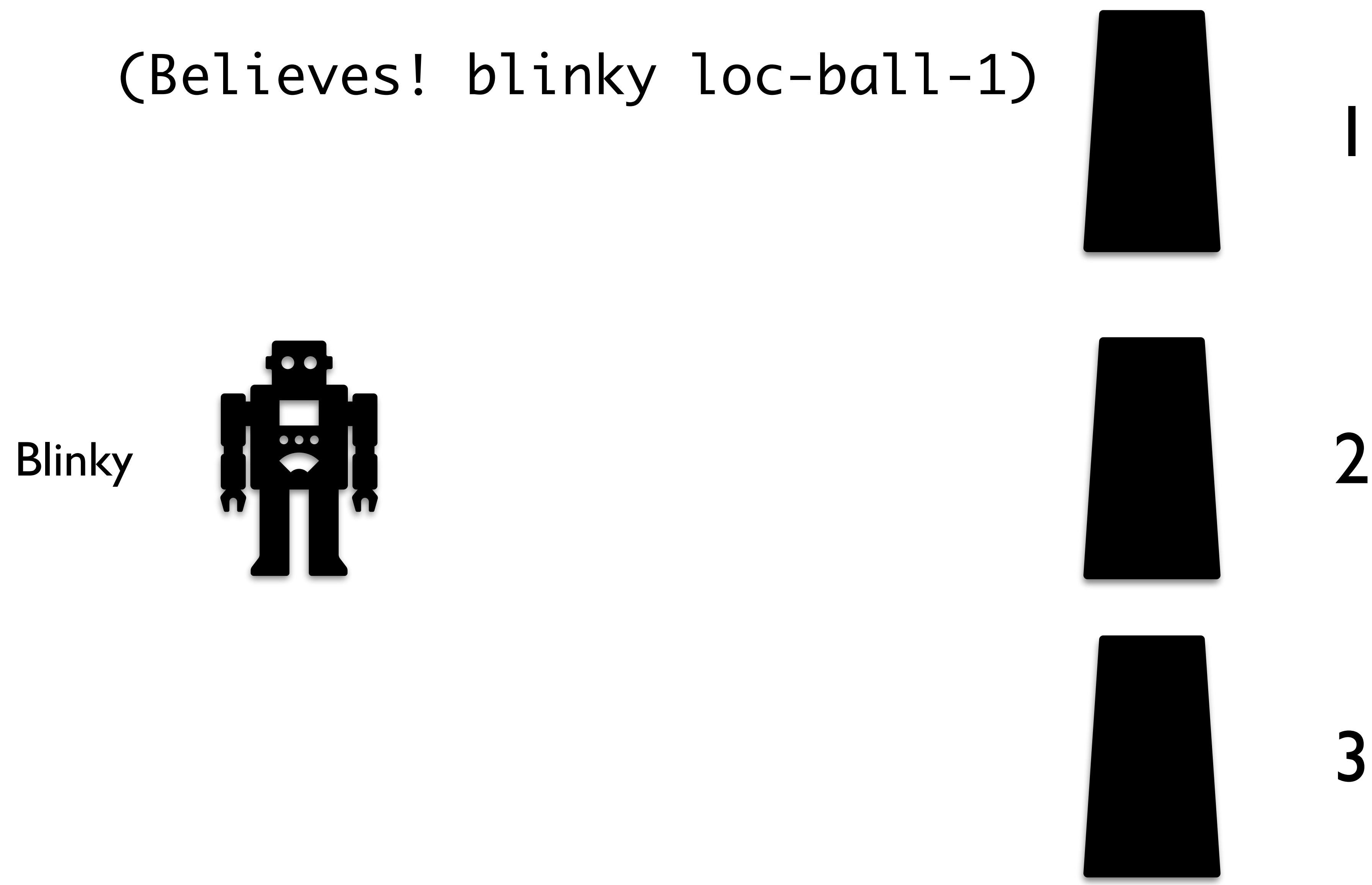
B(blinky, loc-ball-1)



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

B(blinky, loc-ball-1)

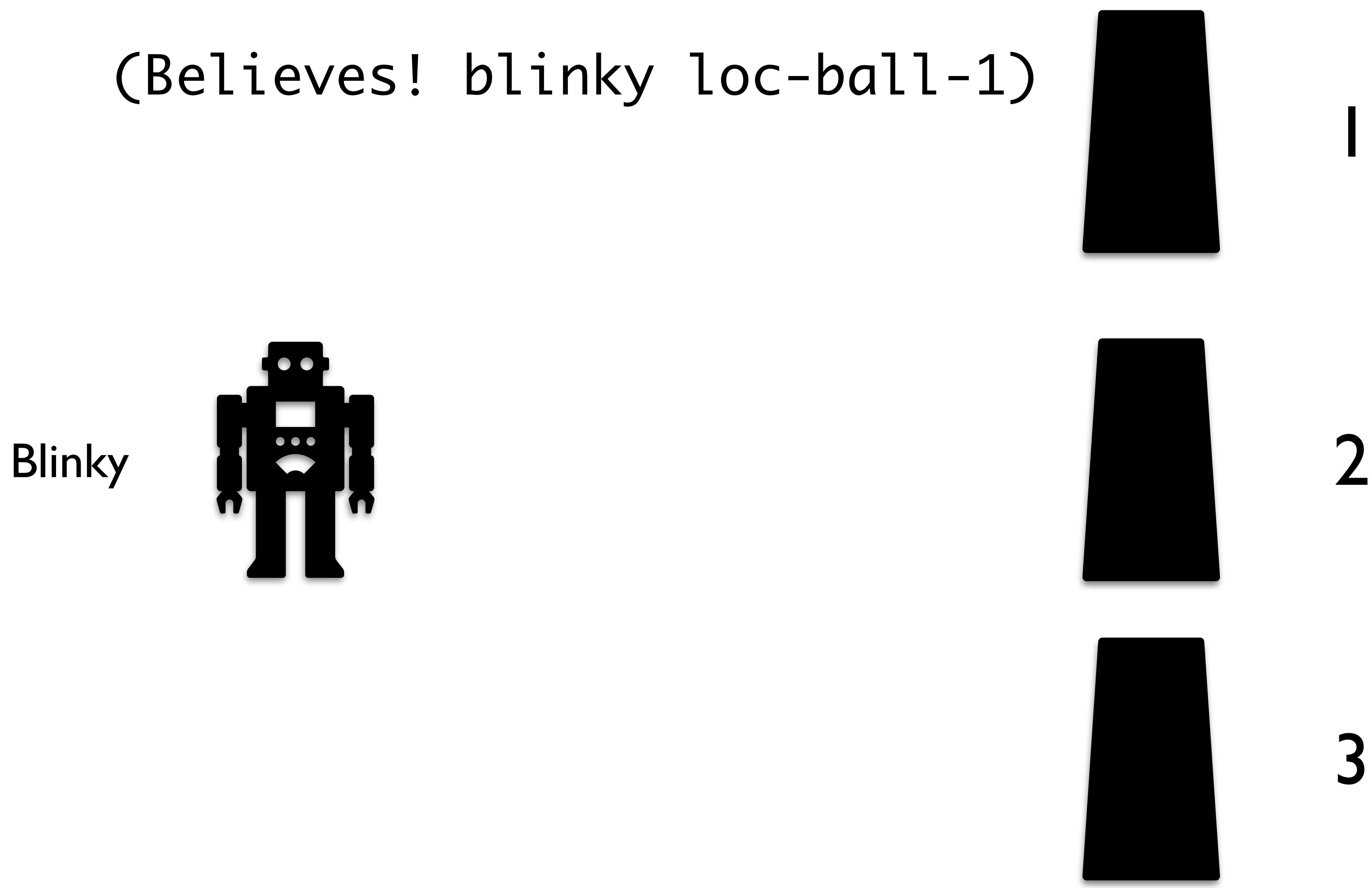
```
(Believes! blinky loc-ball-1)
```



Blinky 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)

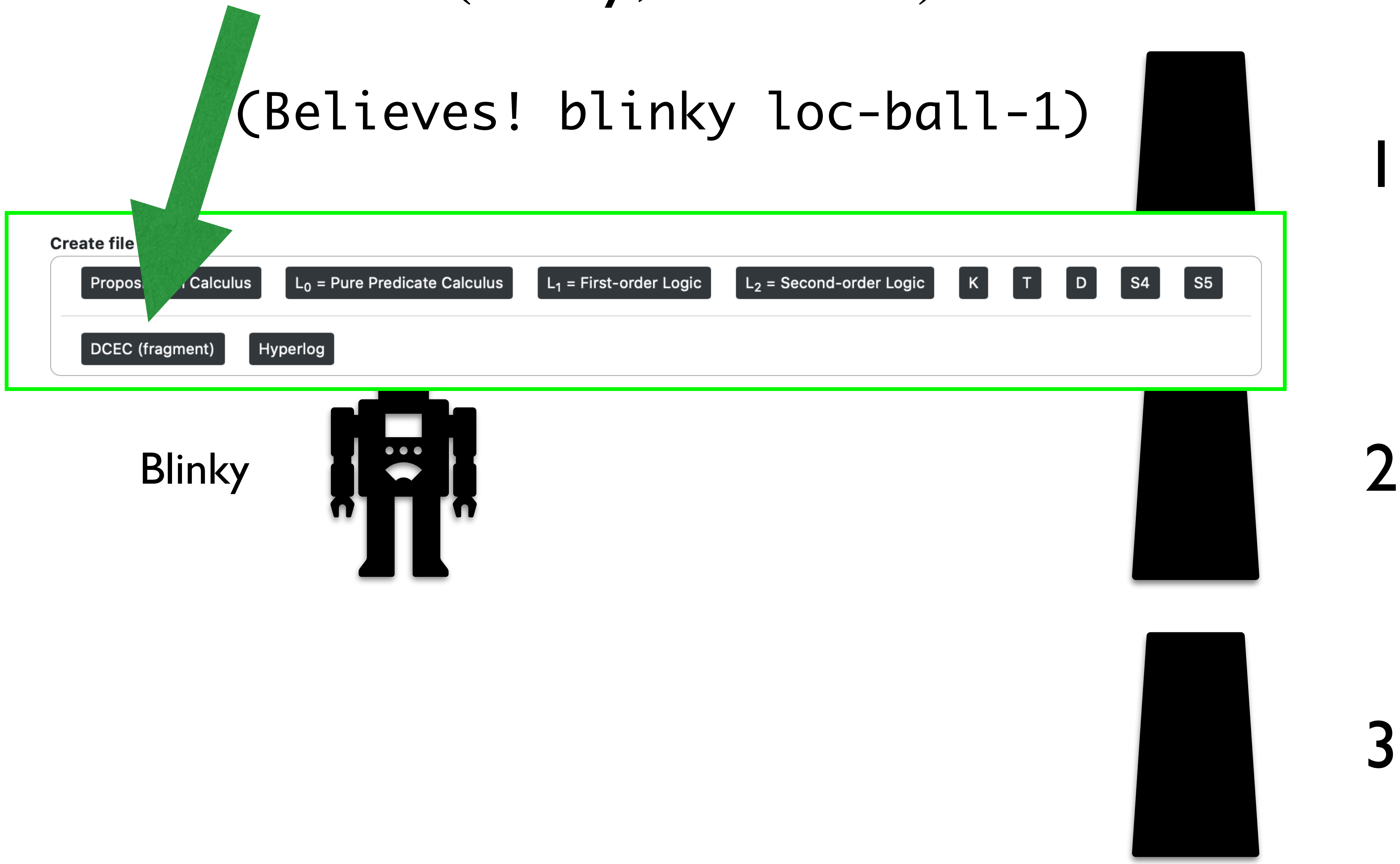
```
(Believes! blinky loc-ball-1)
```

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

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."

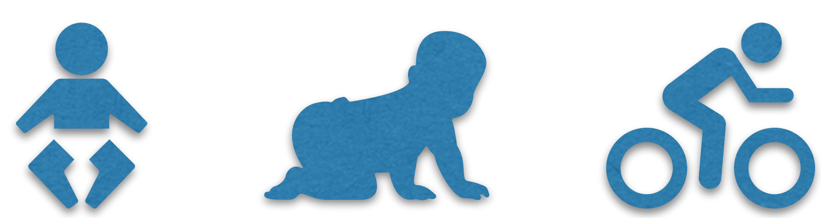
"Blinky is smart."

Therefore:

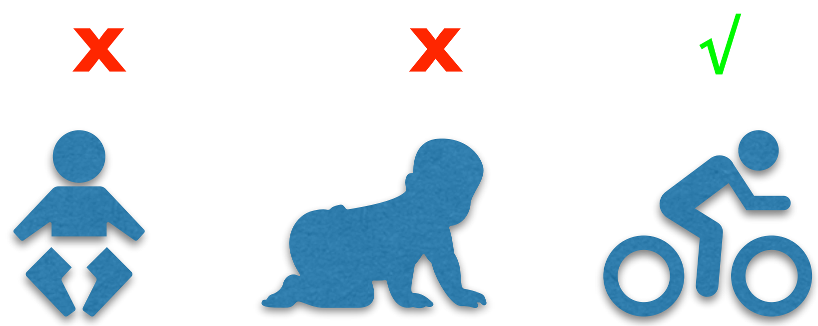
"Everything tinks anything that tinks something identical with something."

False Belief Task Demands Intensional Logic …

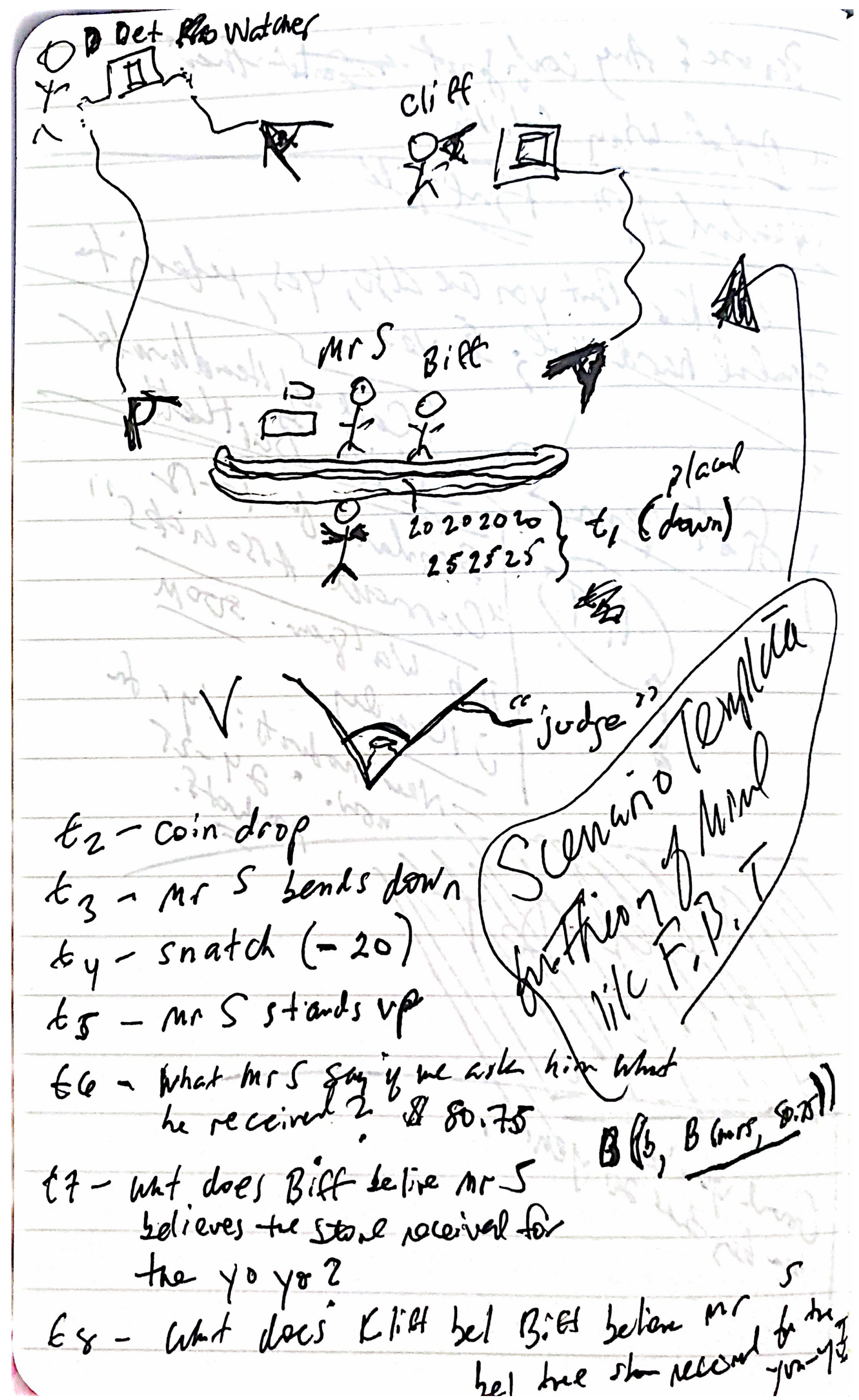
False Belief Task Demands Intensional Logic …



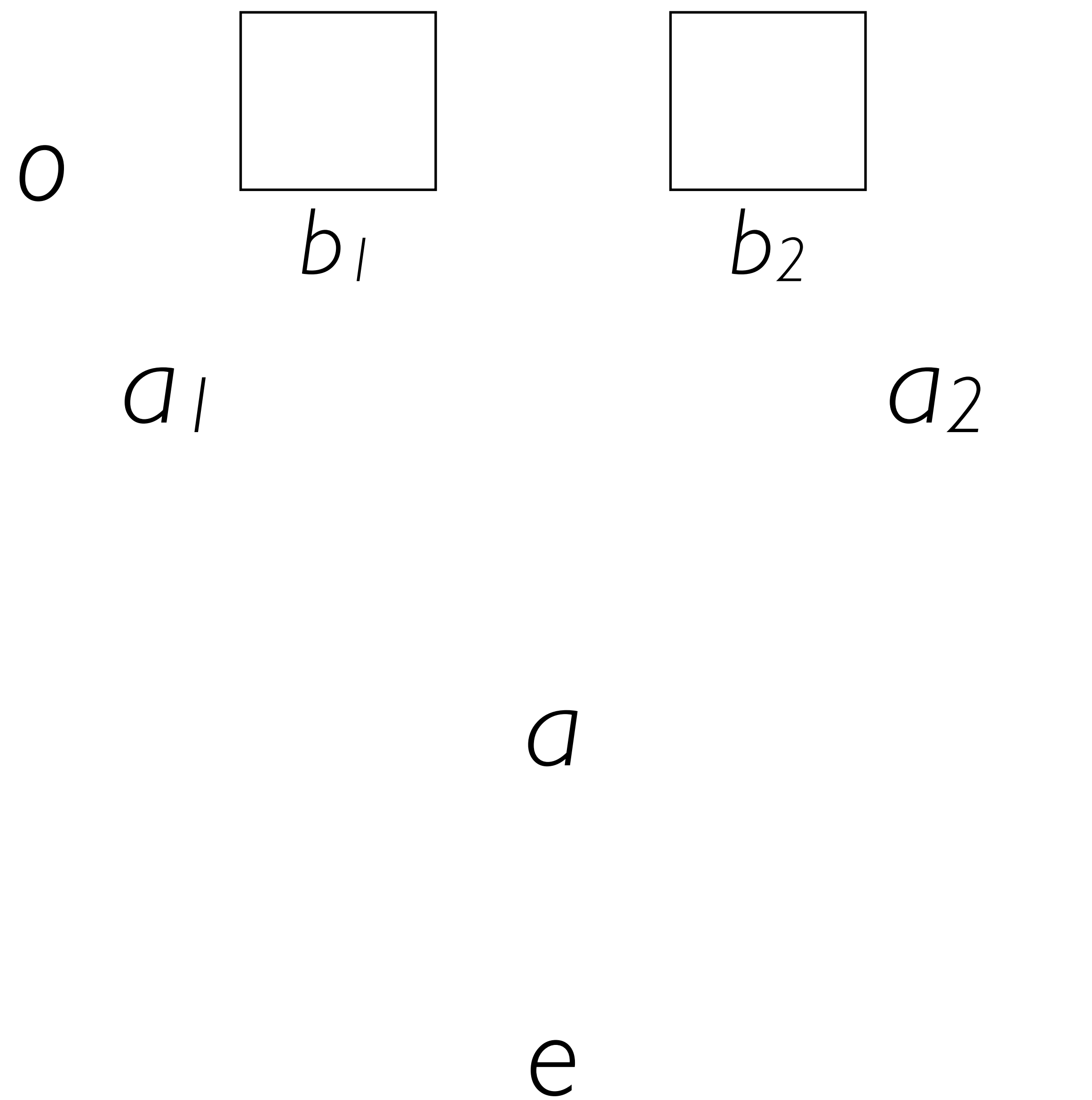
False Belief Task Demands Intensional Logic …



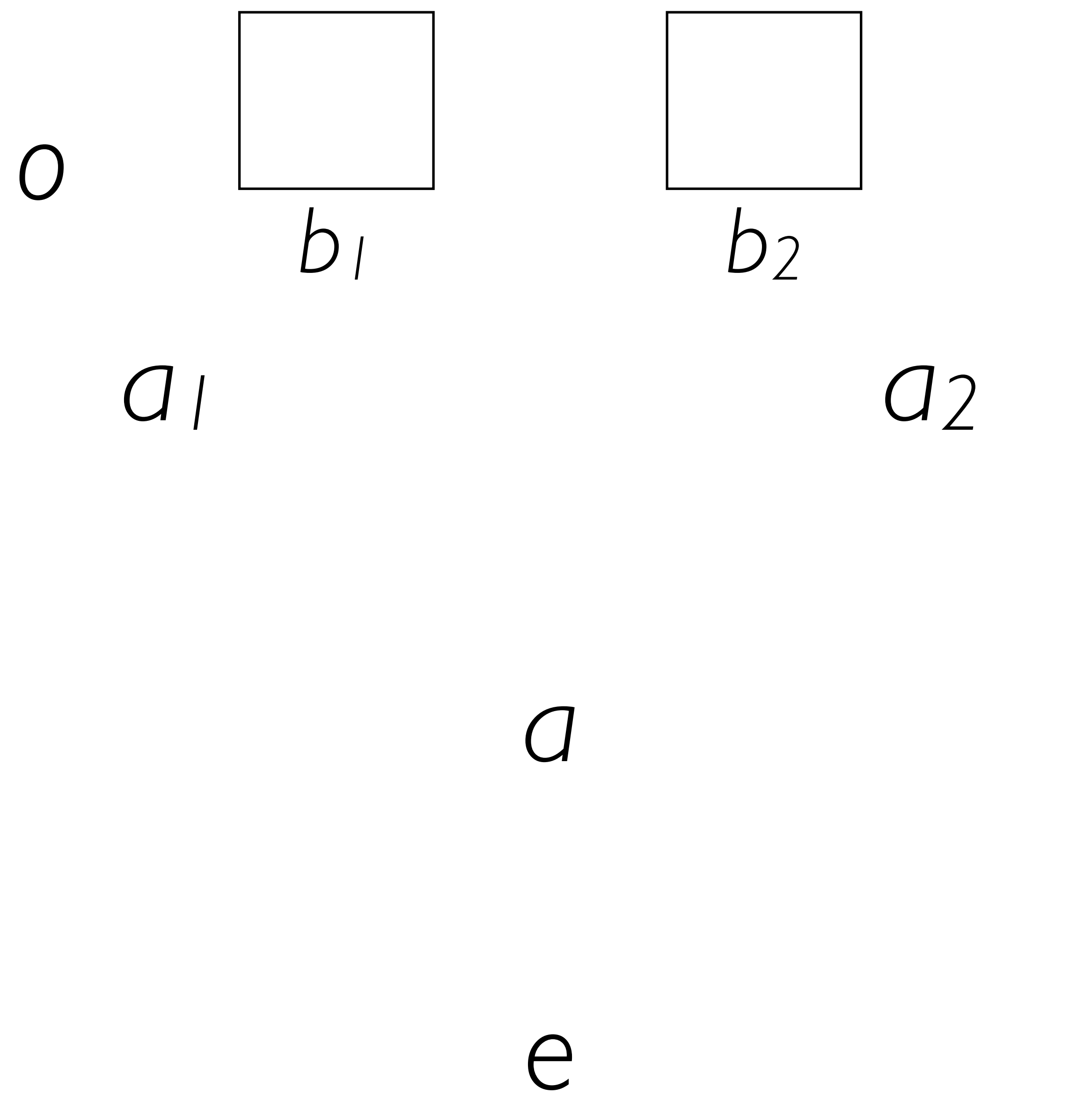
Better, But Embryonic: The ToM Pawn Shop



Framework for $FBT^0{}_1$

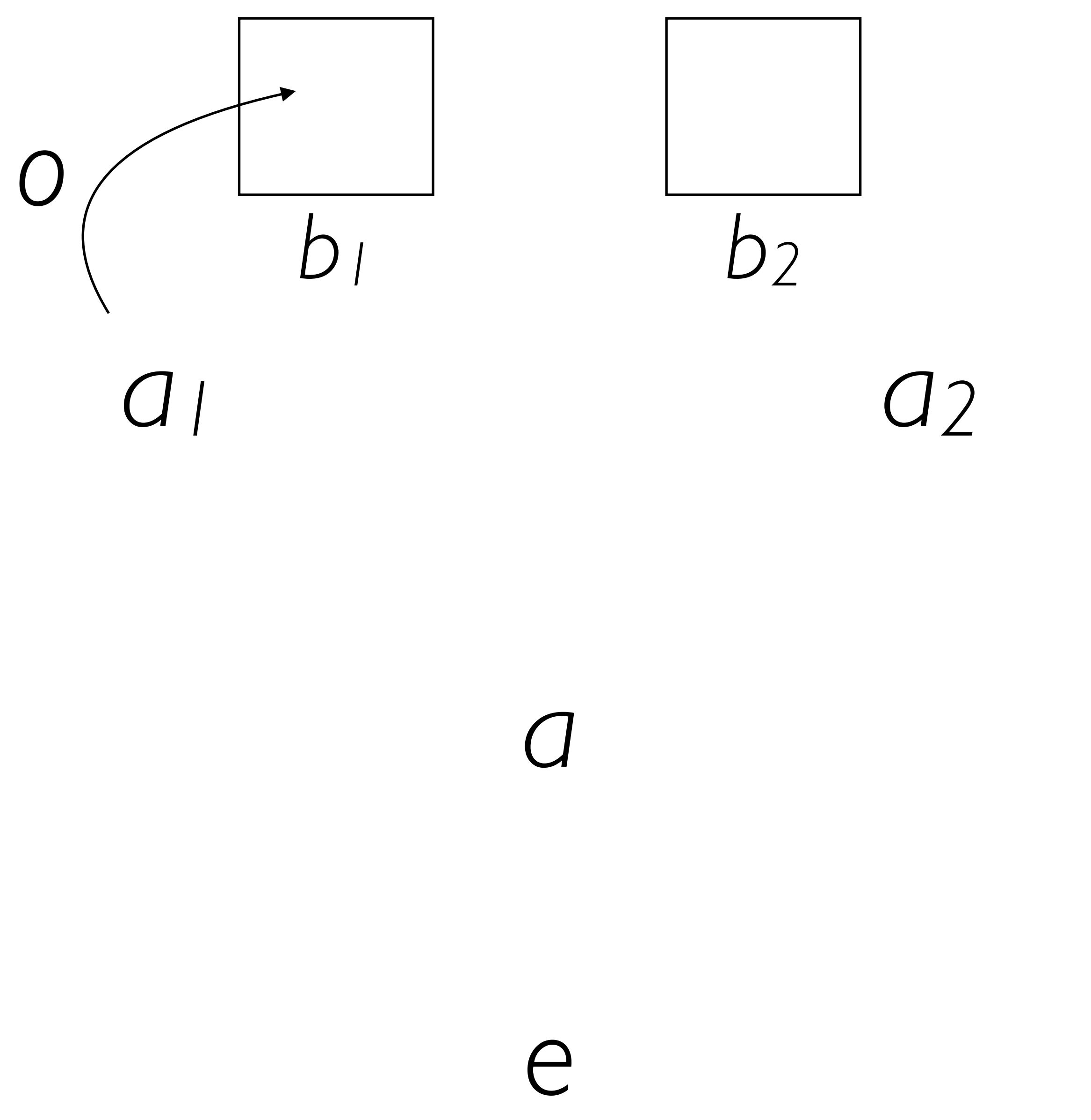


Framework for FBT^0_1
(five timepoints)

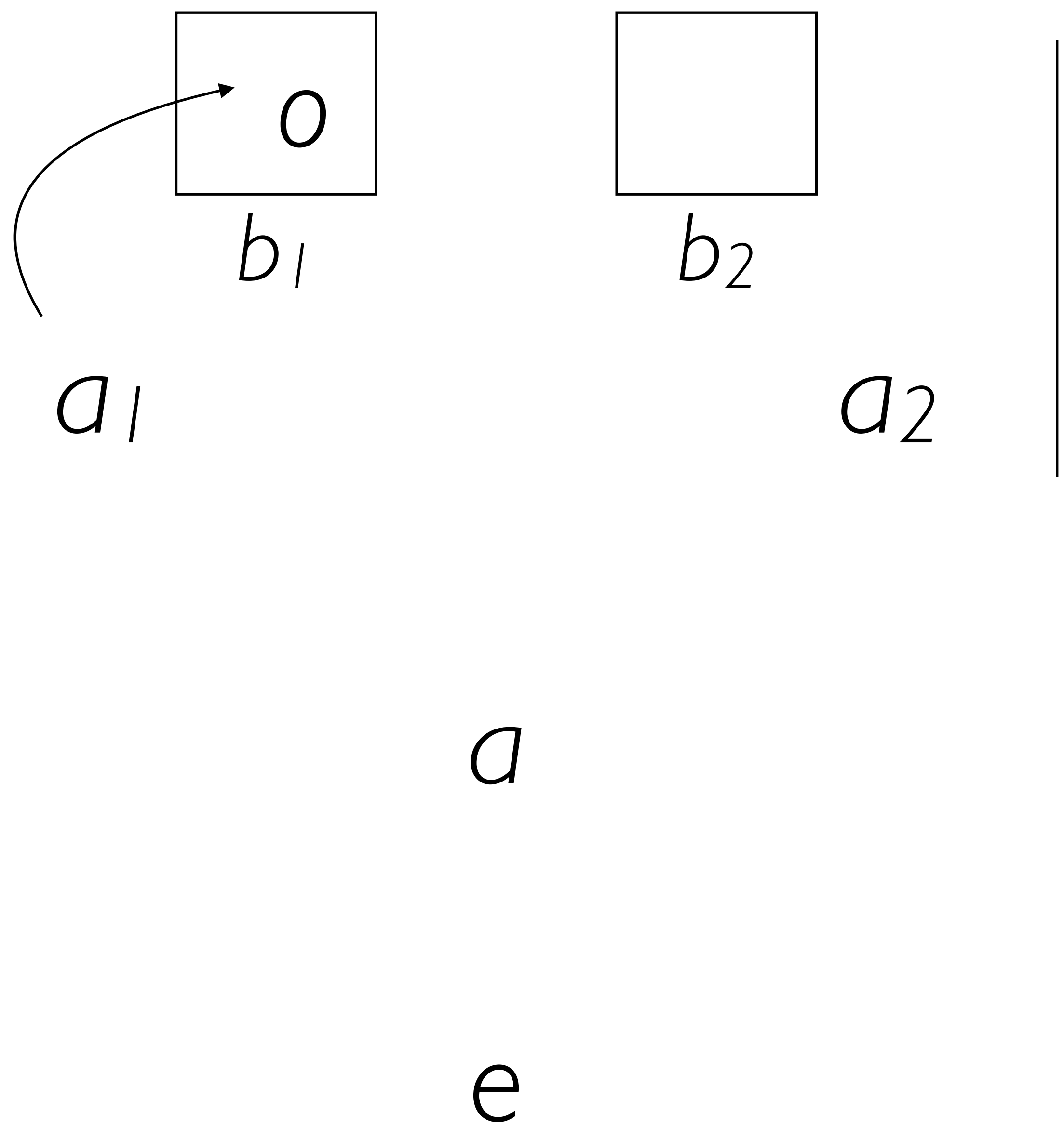


Framework for FBT^0_1

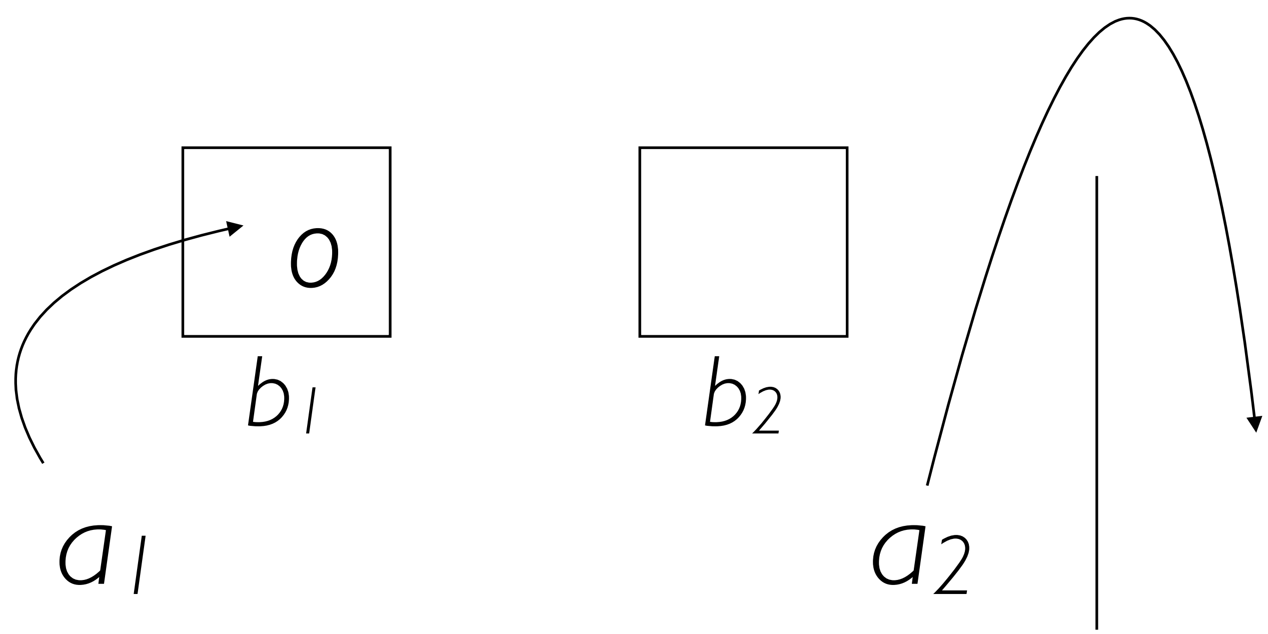
(five timepoints)



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



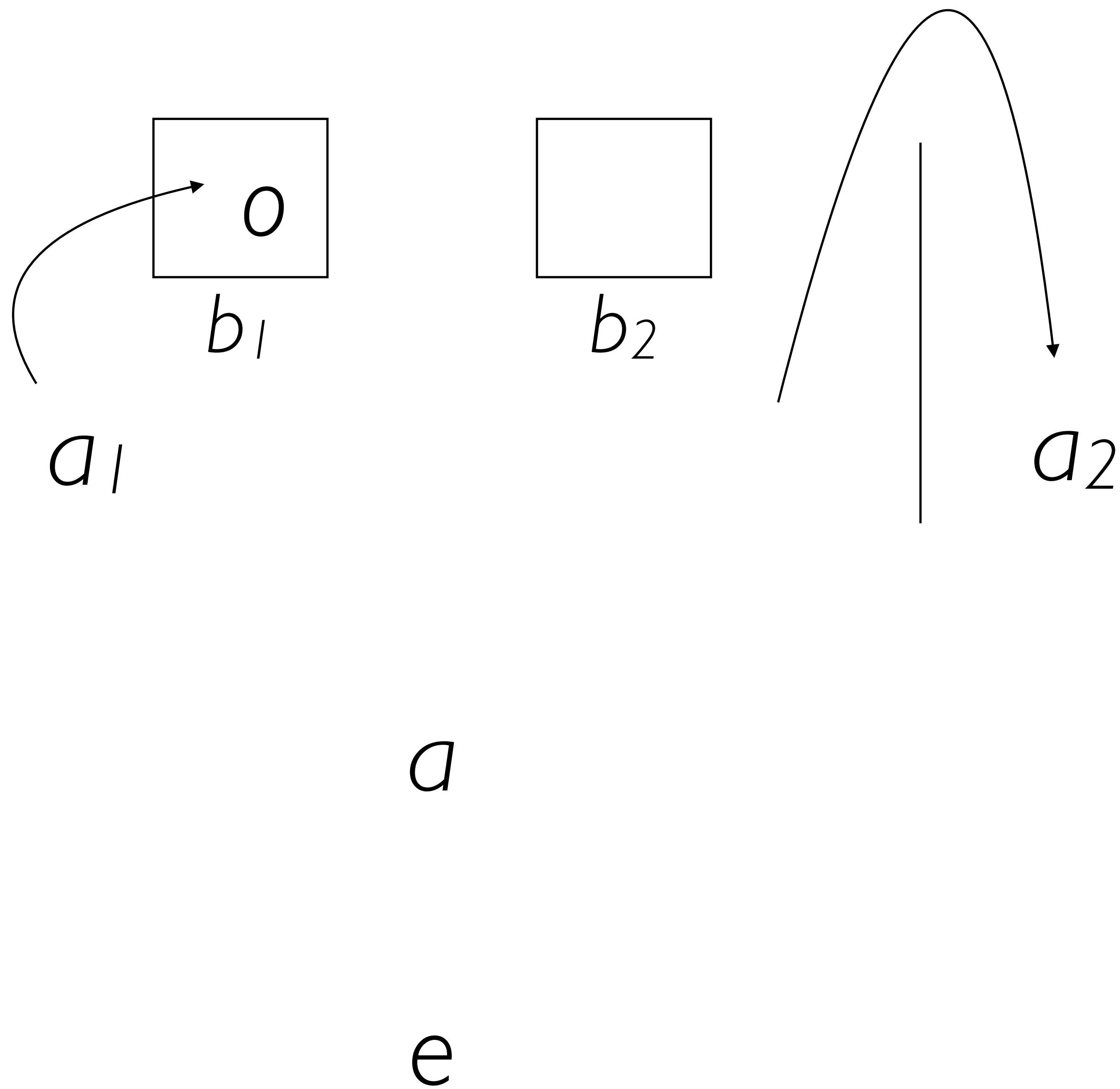
Framework for $FBT^0{}_1$
(five timepoints)



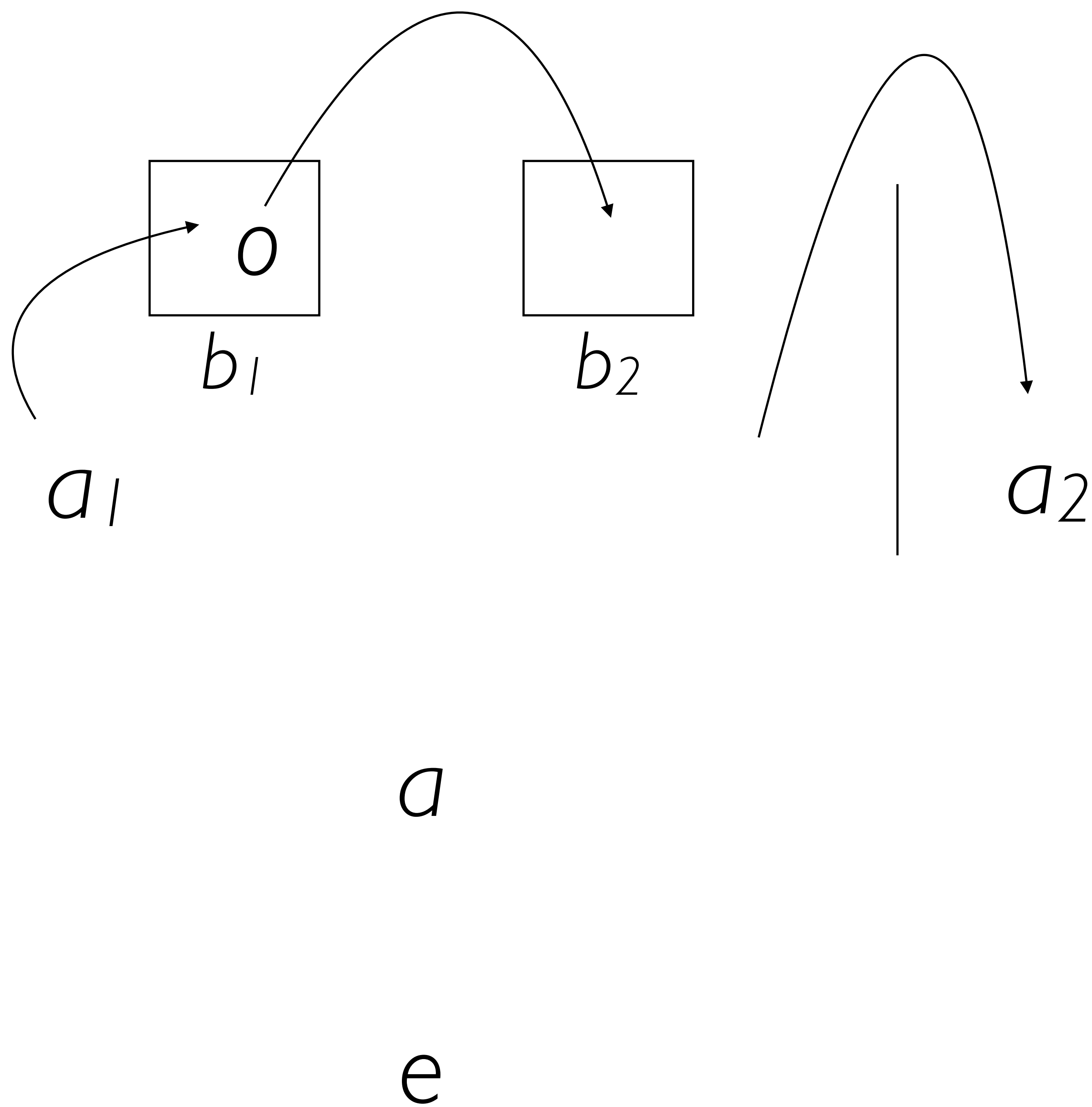
a

e

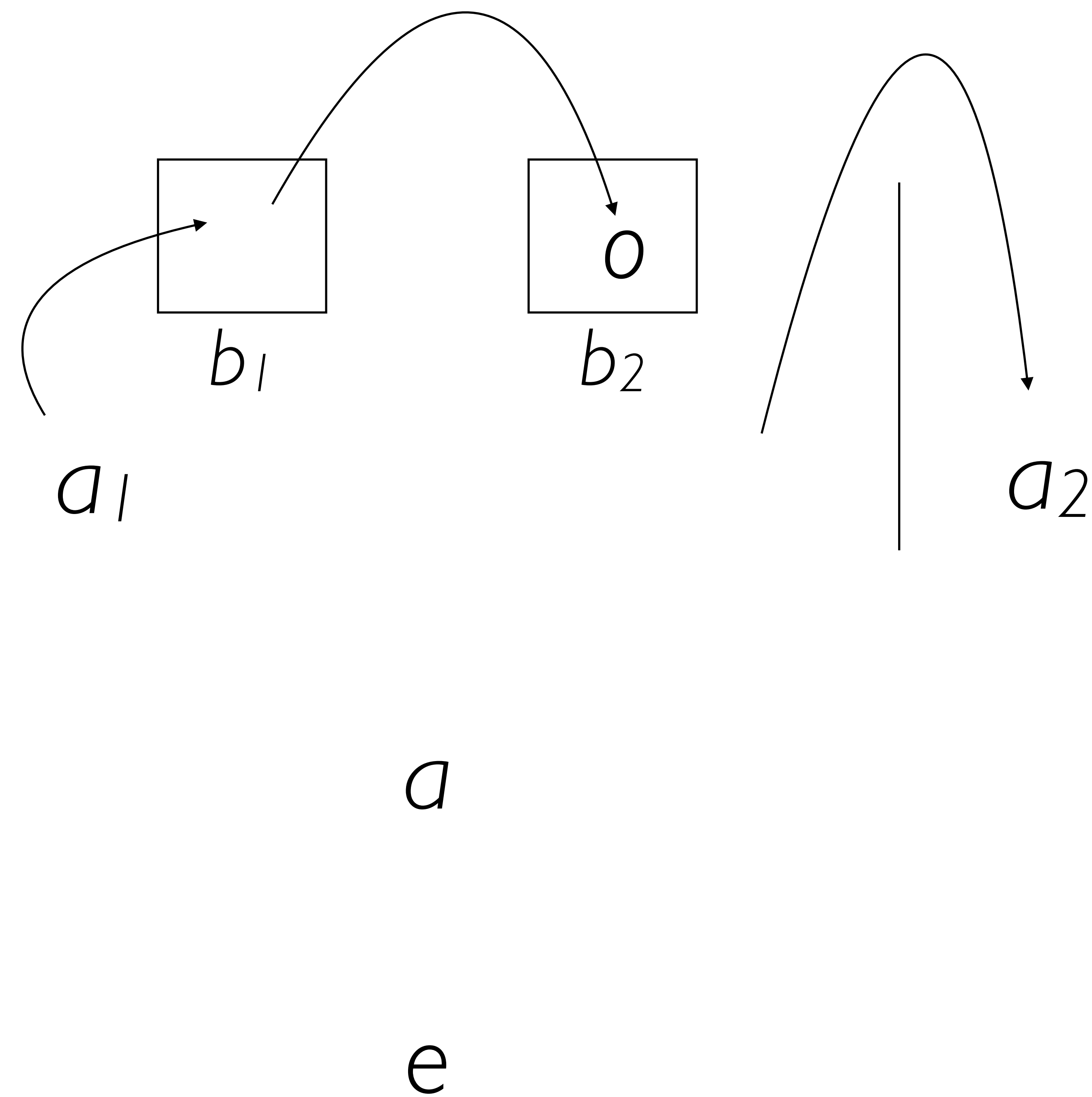
Framework for FBT^0_1
(five timepoints)



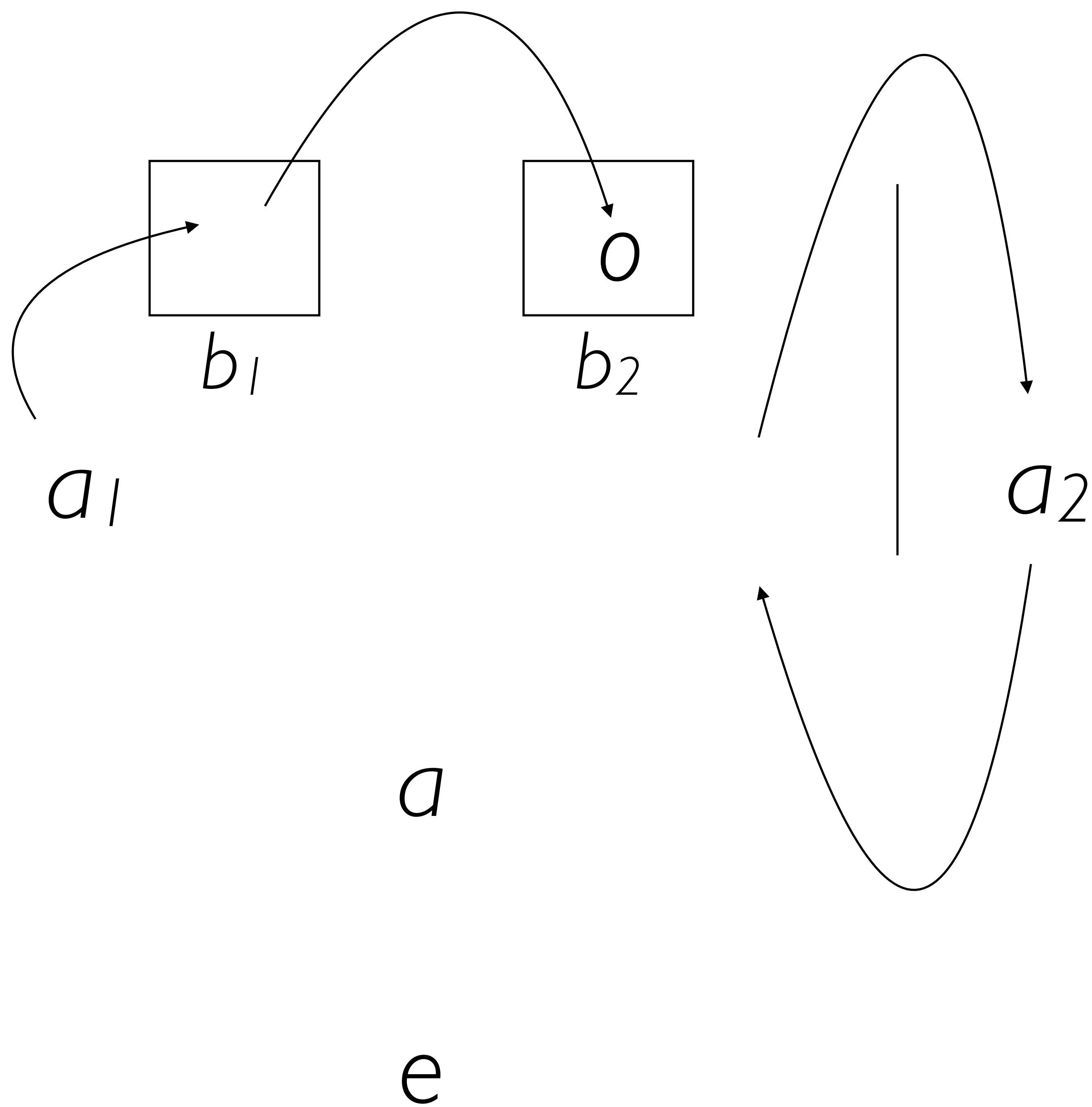
Framework for $FBT^0{}_1$
(five timepoints)



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

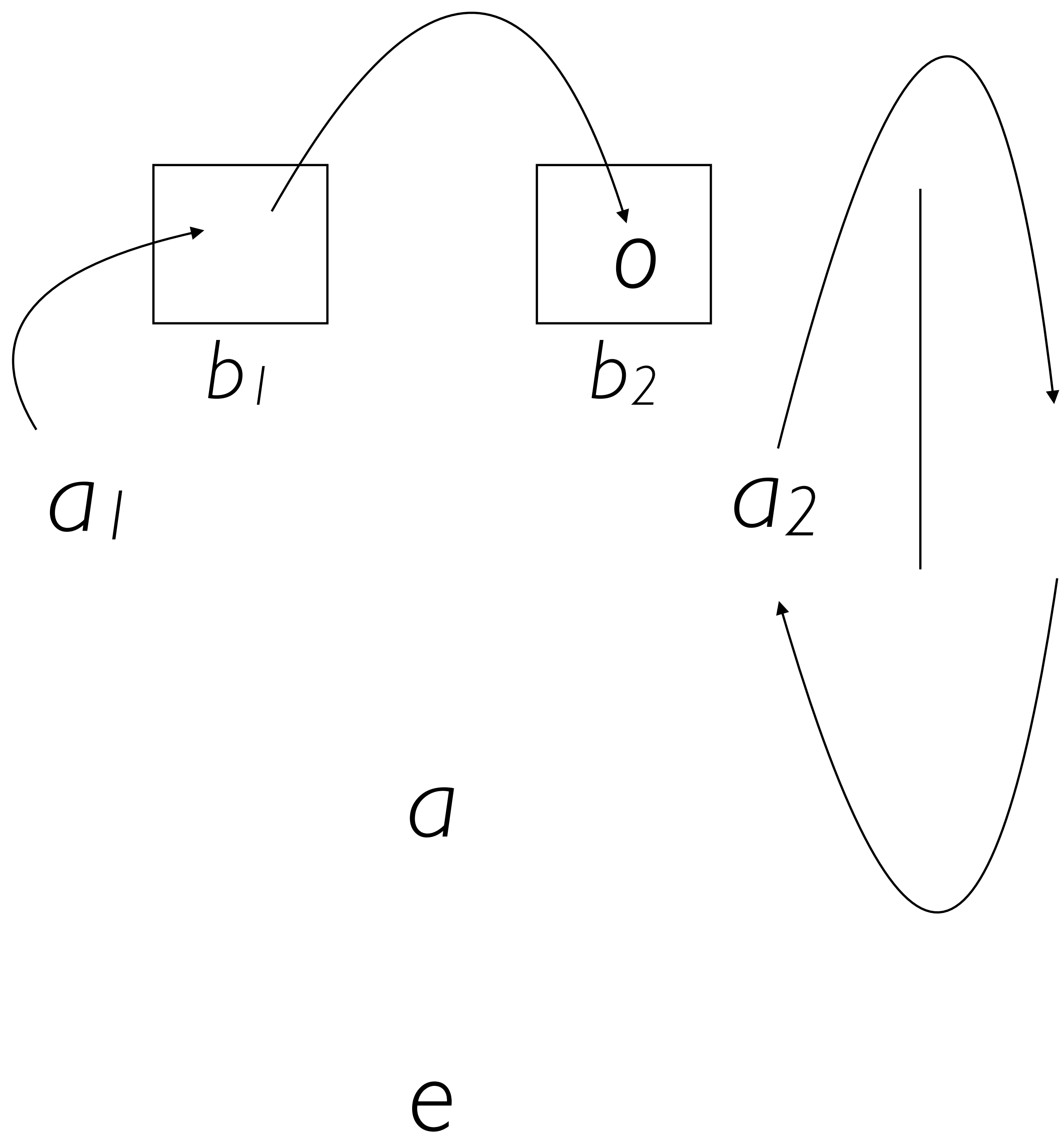


Framework for $FBT^0{}_1$
(five timepoints)

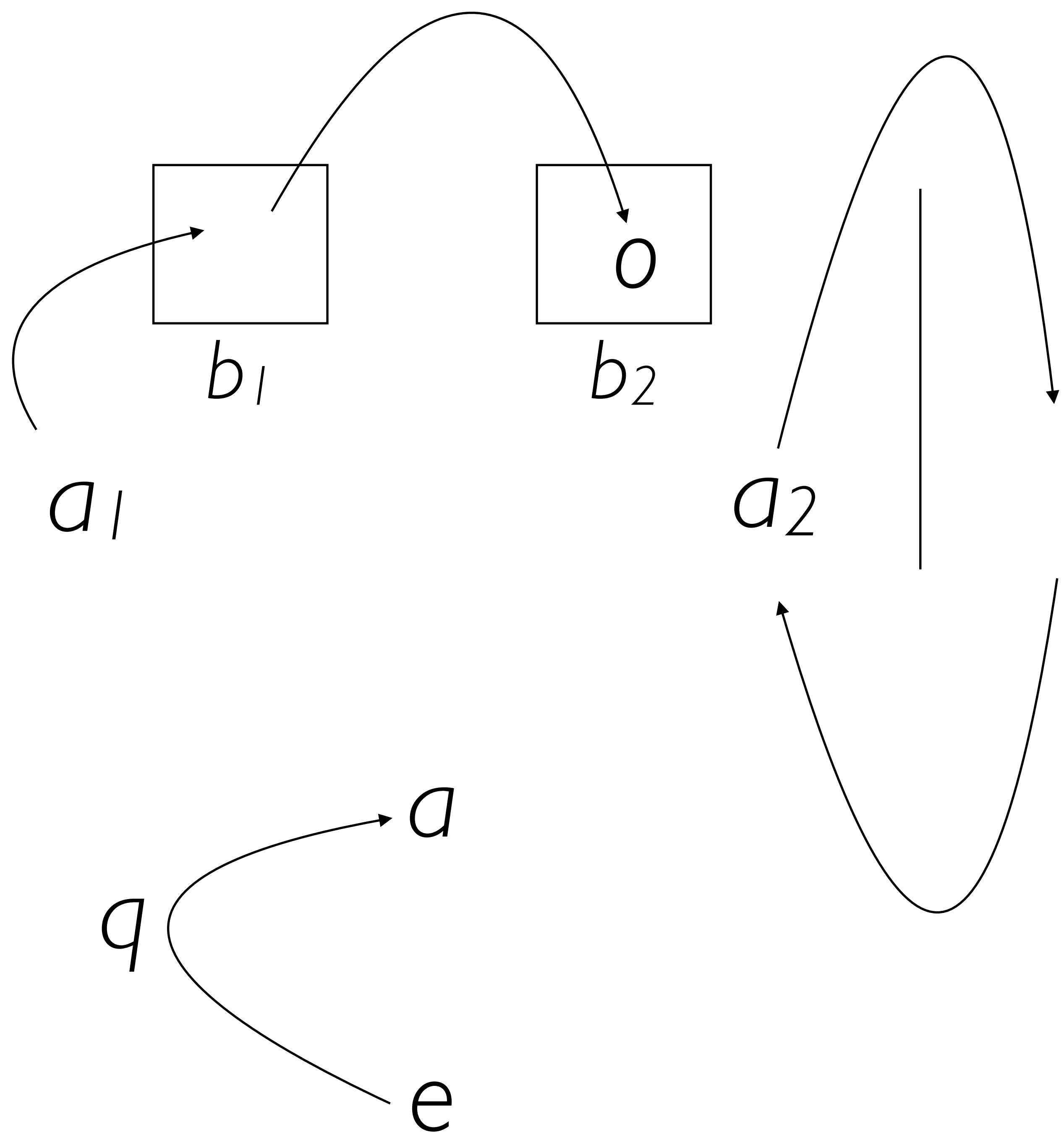


Framework for $FBT^0{}_1$

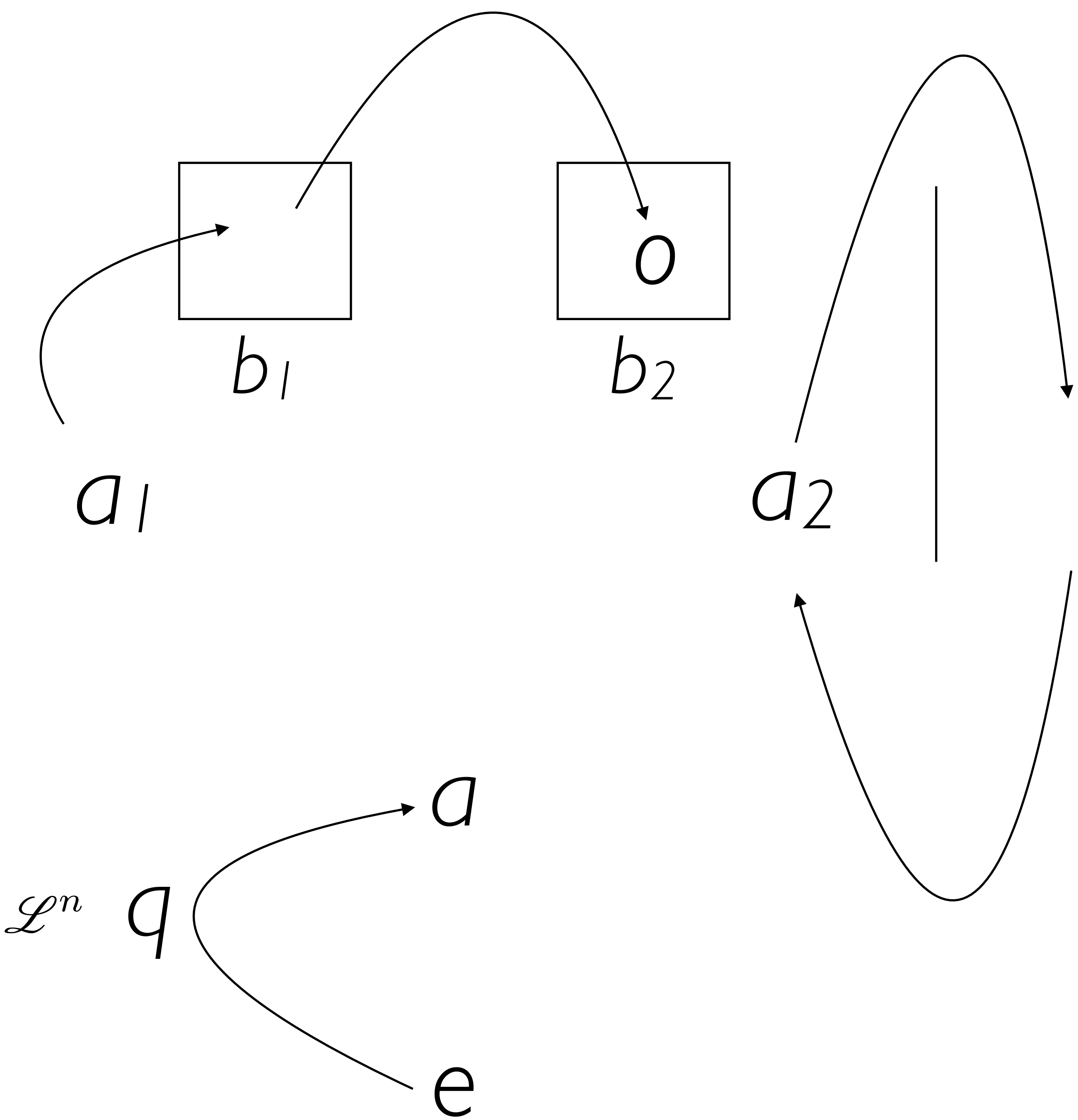
(five timepoints)



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

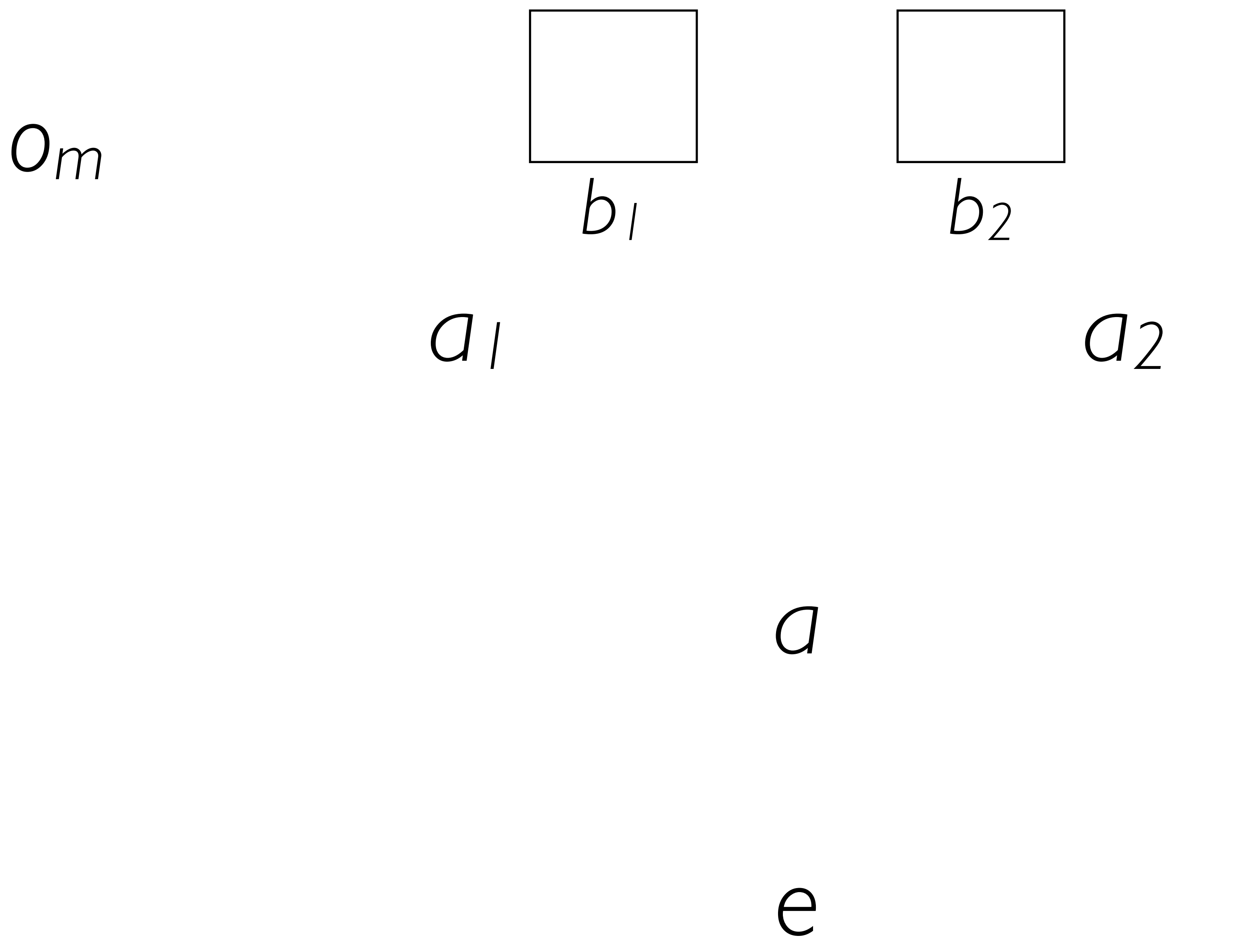


Framework for $FBT^0{}_1$
(five timepoints)

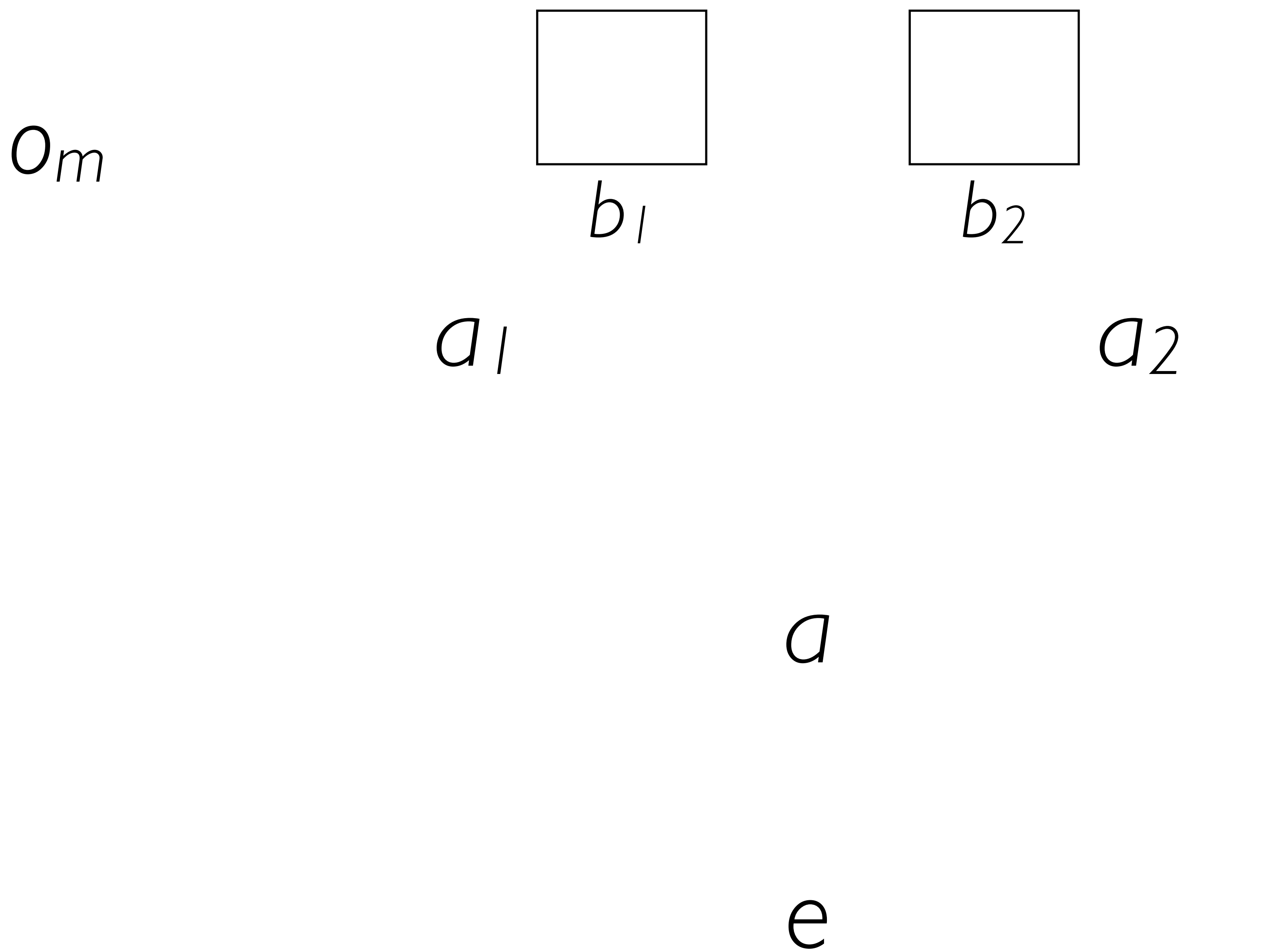


q a formula in modal $\mathscr{L}^n$

Framework for FBT^{1}_{1}

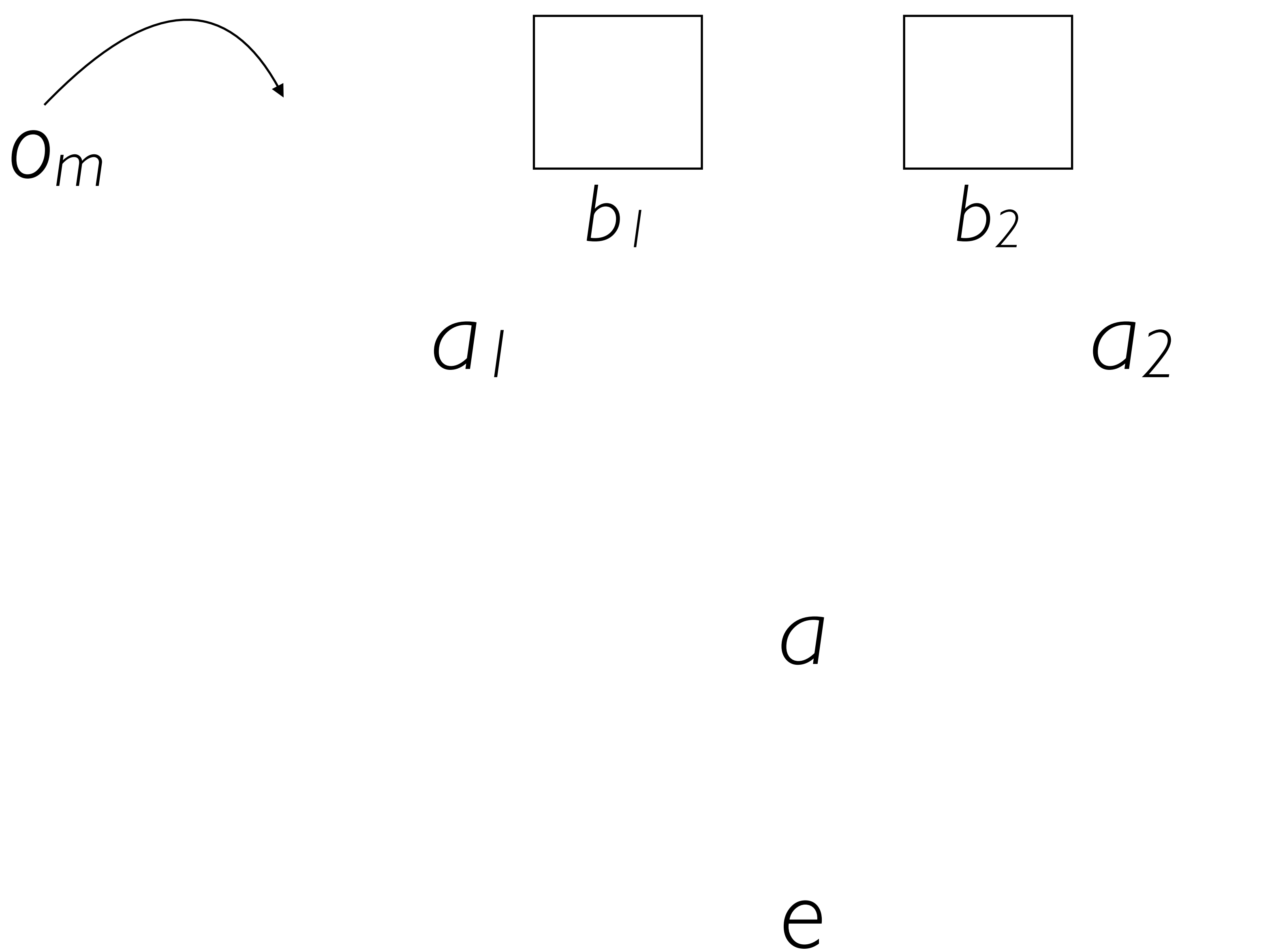


Framework for FBT^1_1
(six timepoints)

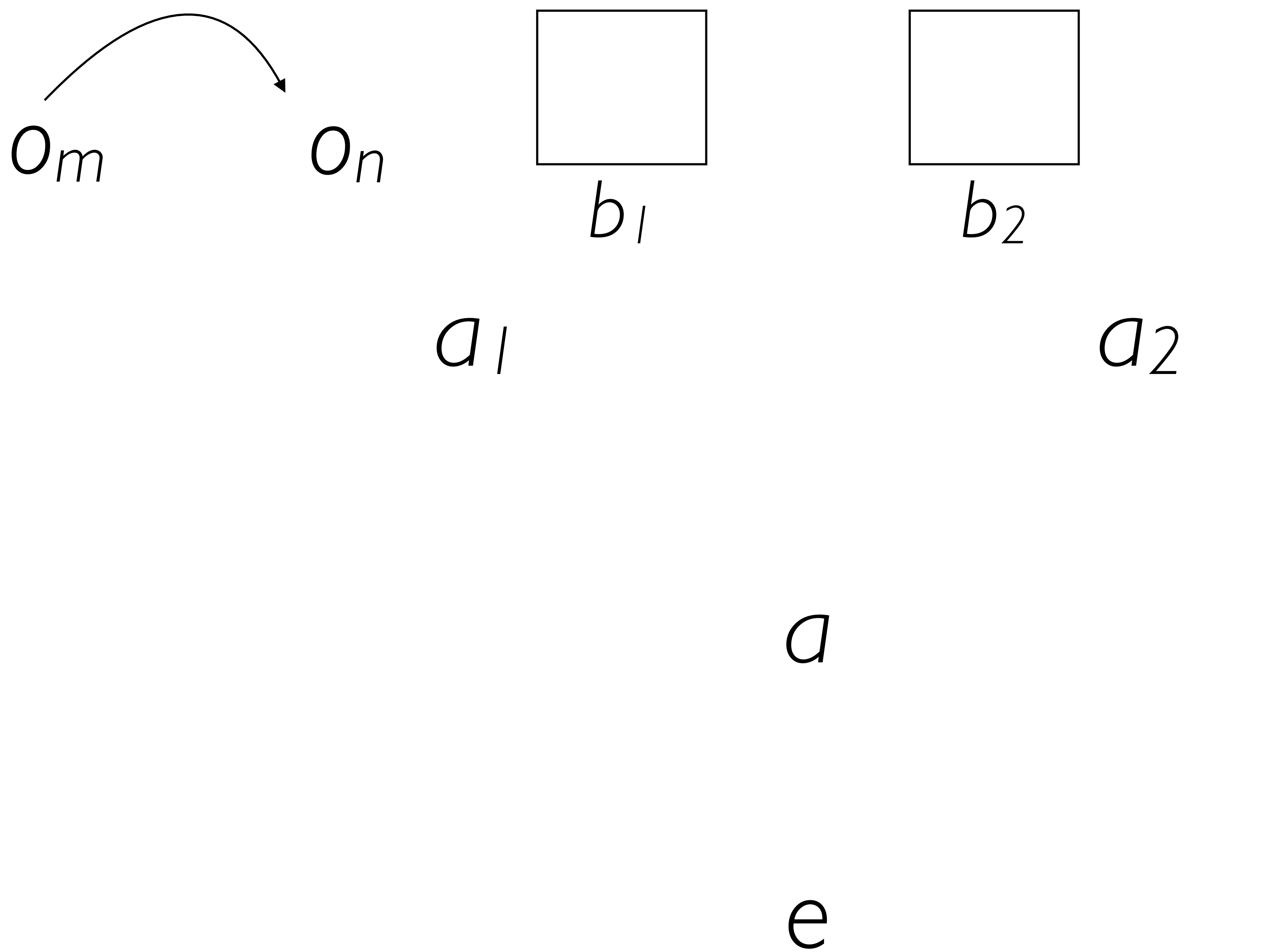


Framework for FBT^1_1

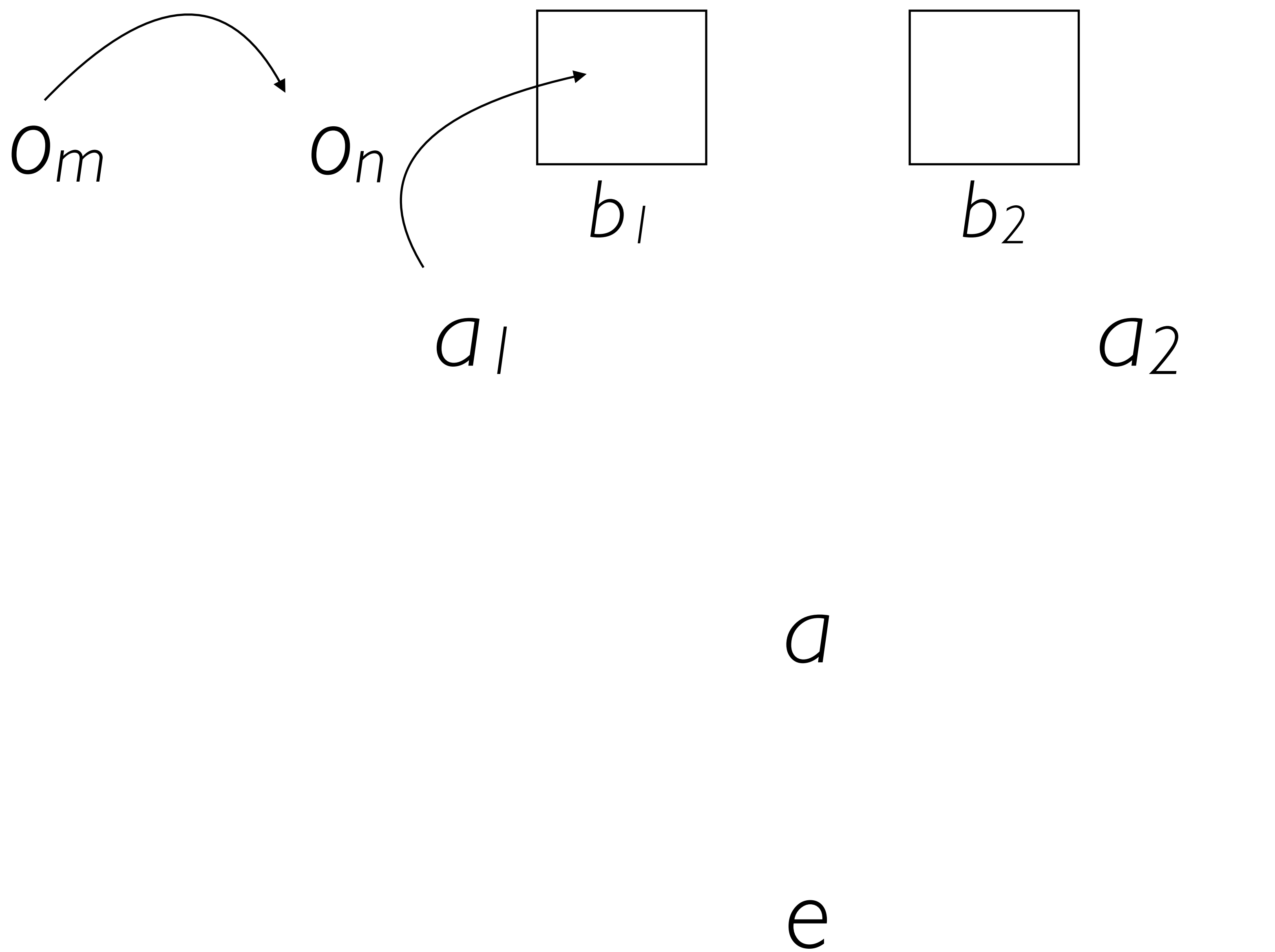
(six timepoints)



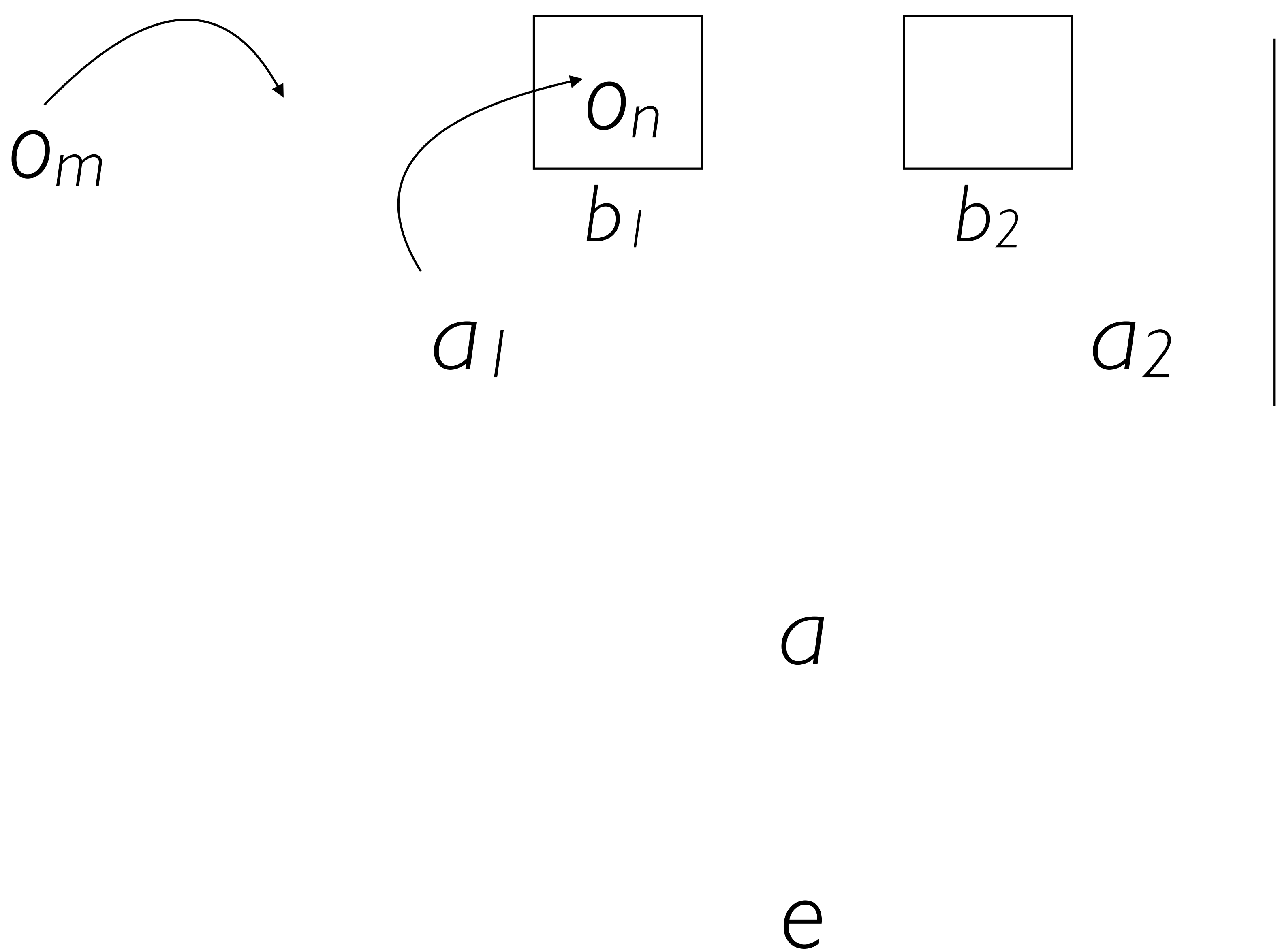
Framework for FBT^{1}_{1}
(six timepoints)



Framework for FBT^1_1
(six timepoints)

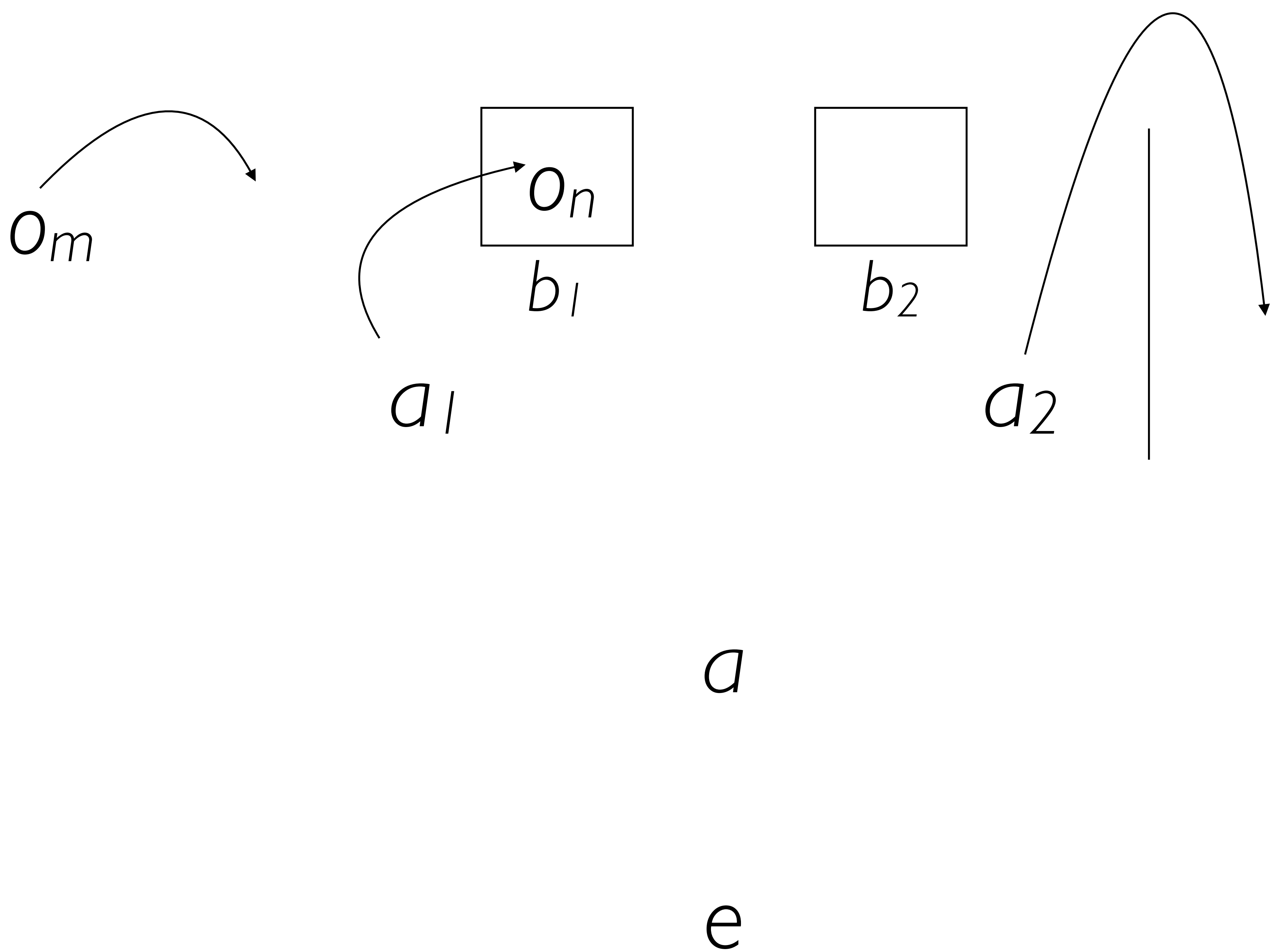


Framework for $FBT^1{}_1$
(six timepoints)



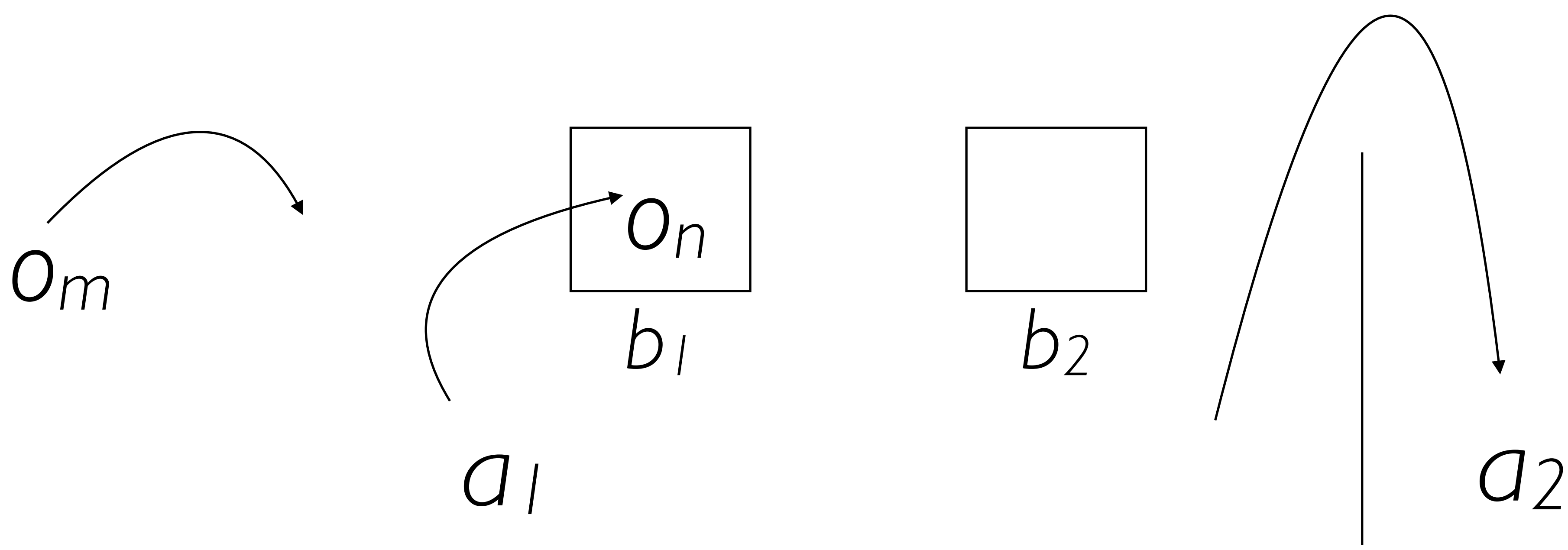
Framework for FBT^1_1

(six timepoints)



Framework for FBT¹₁

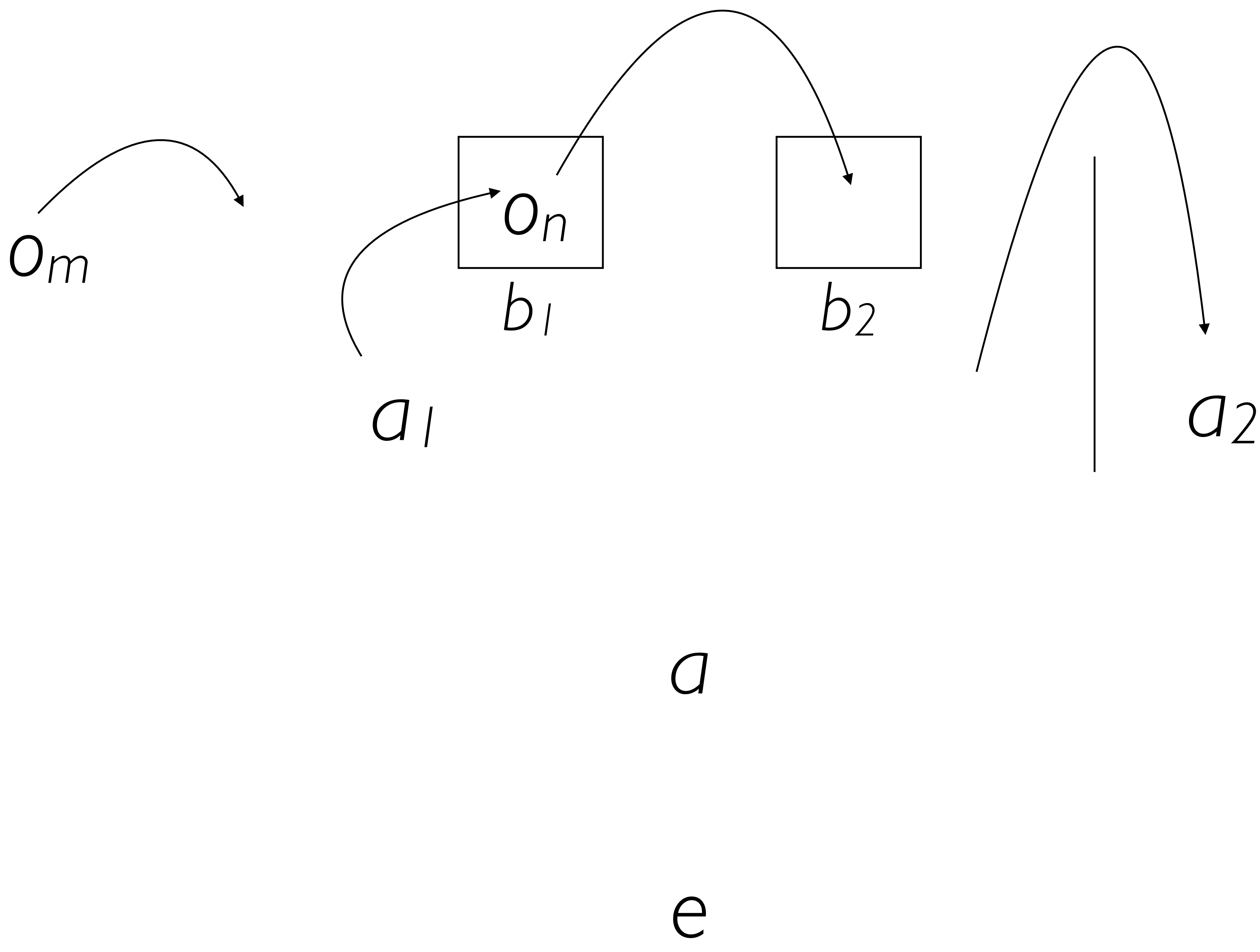
(six timepoints)



a

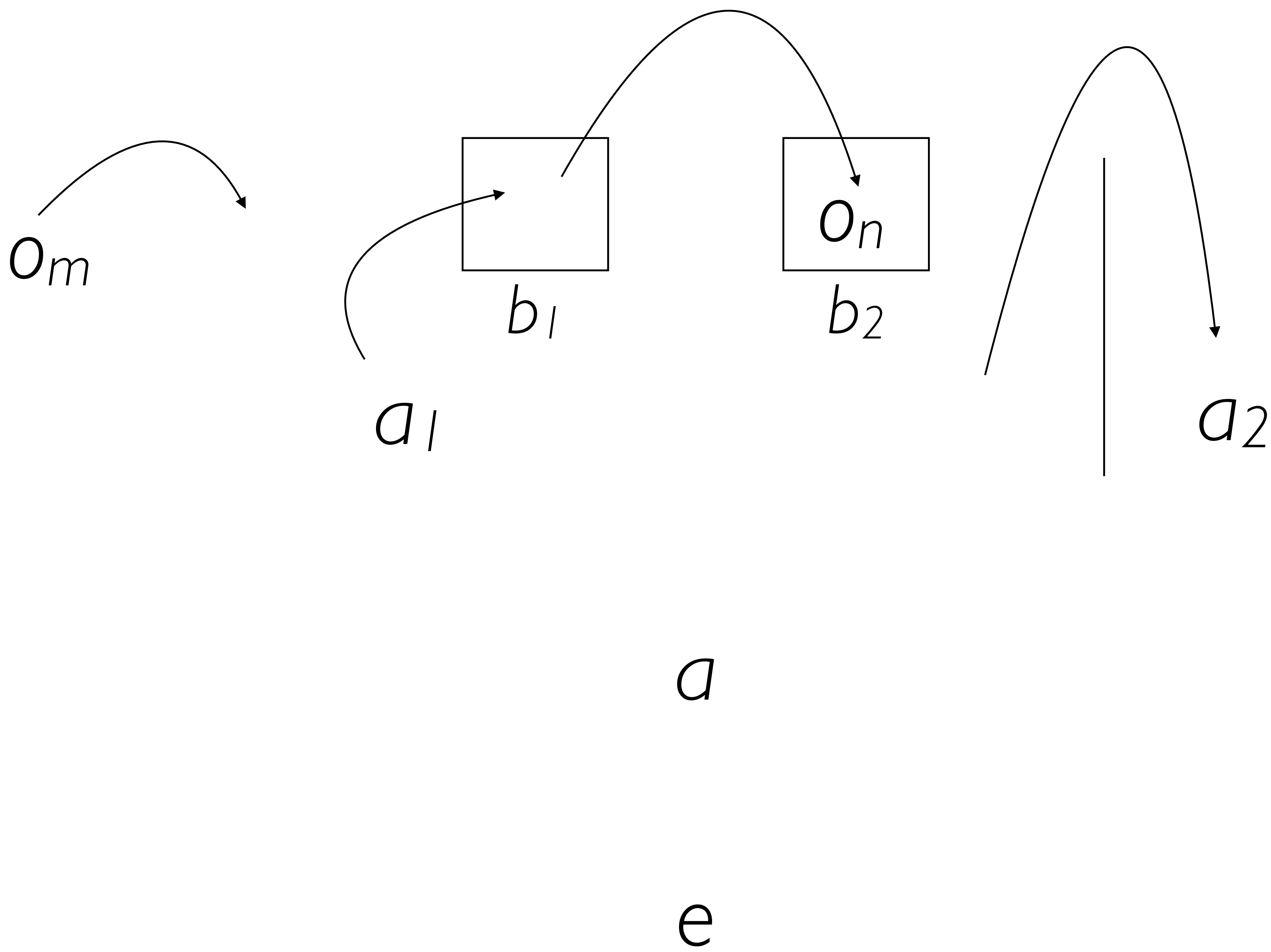
e

Framework for FBT^1_1
(six timepoints)



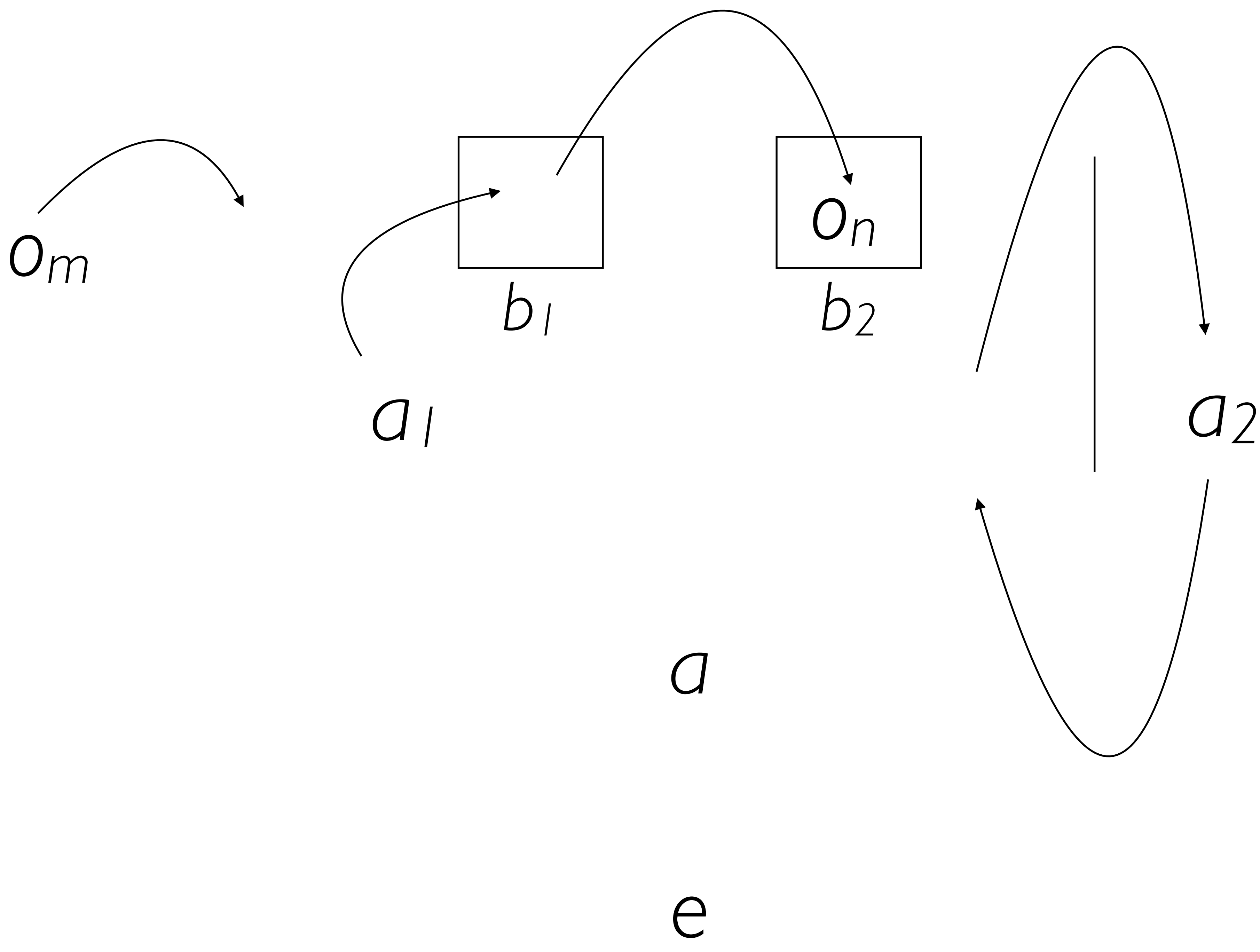
Framework for FBT¹₁

(six timepoints)

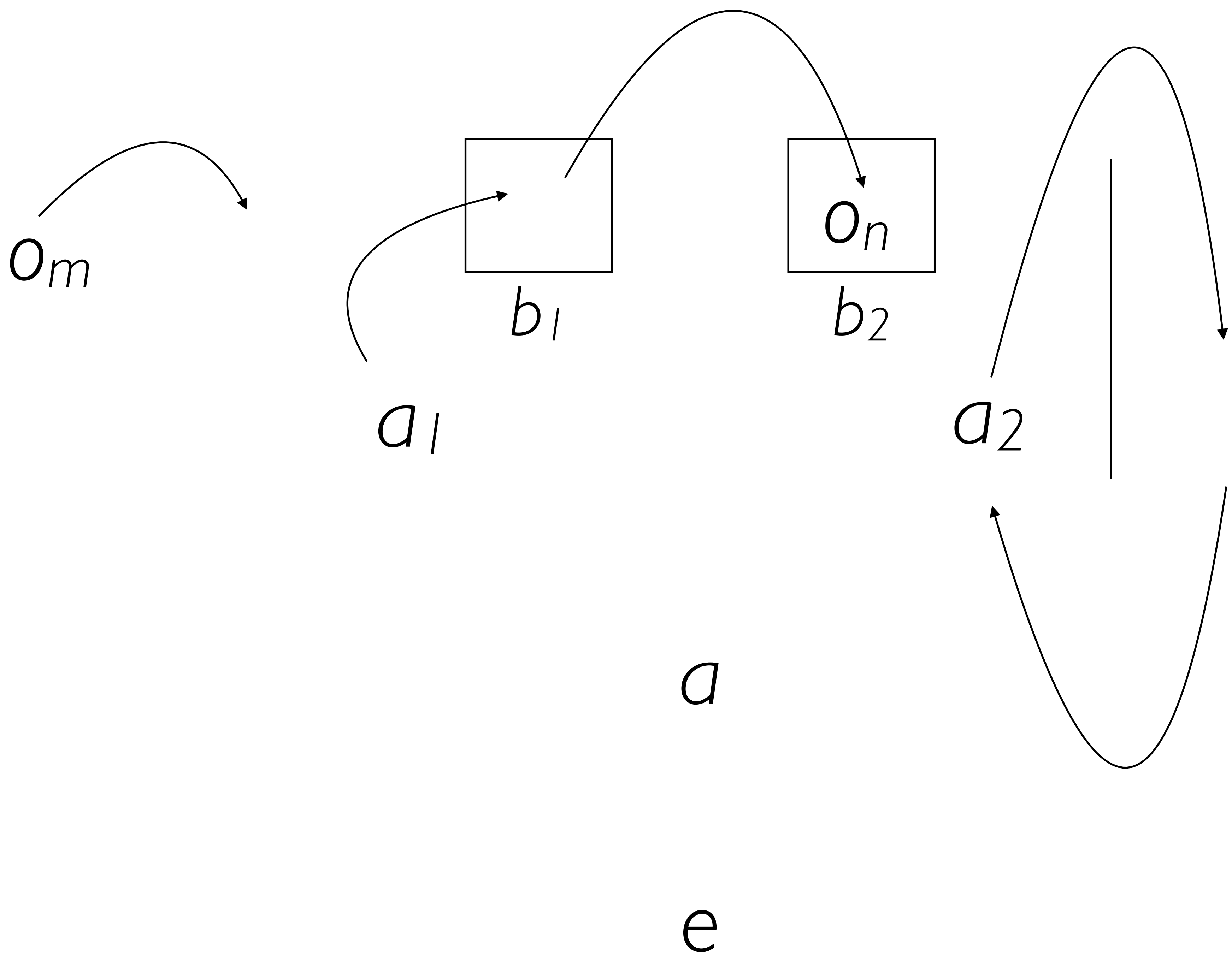


Framework for $FBT^1{}_1$

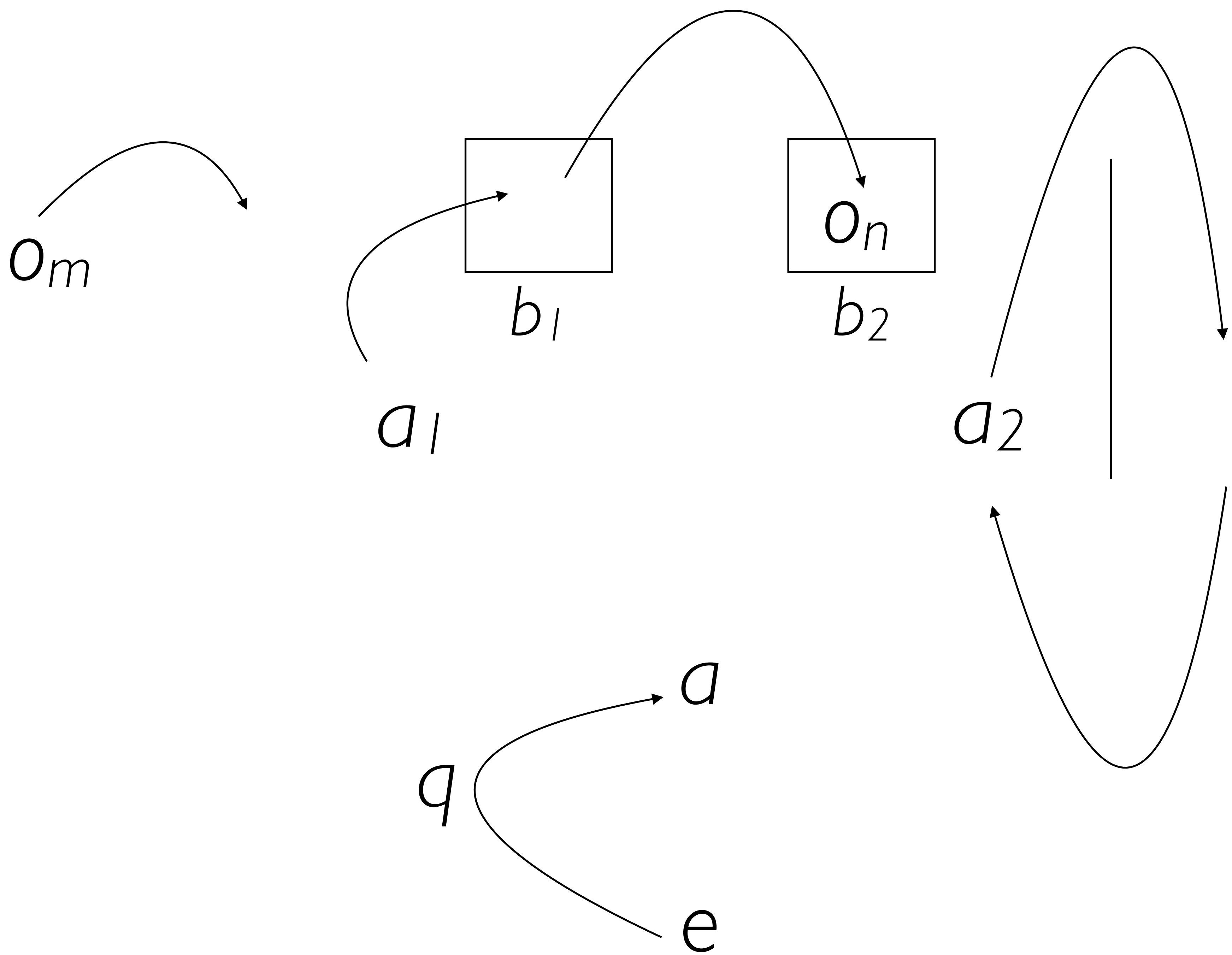
(six timepoints)



Framework for FBT'1 (six timepoints)

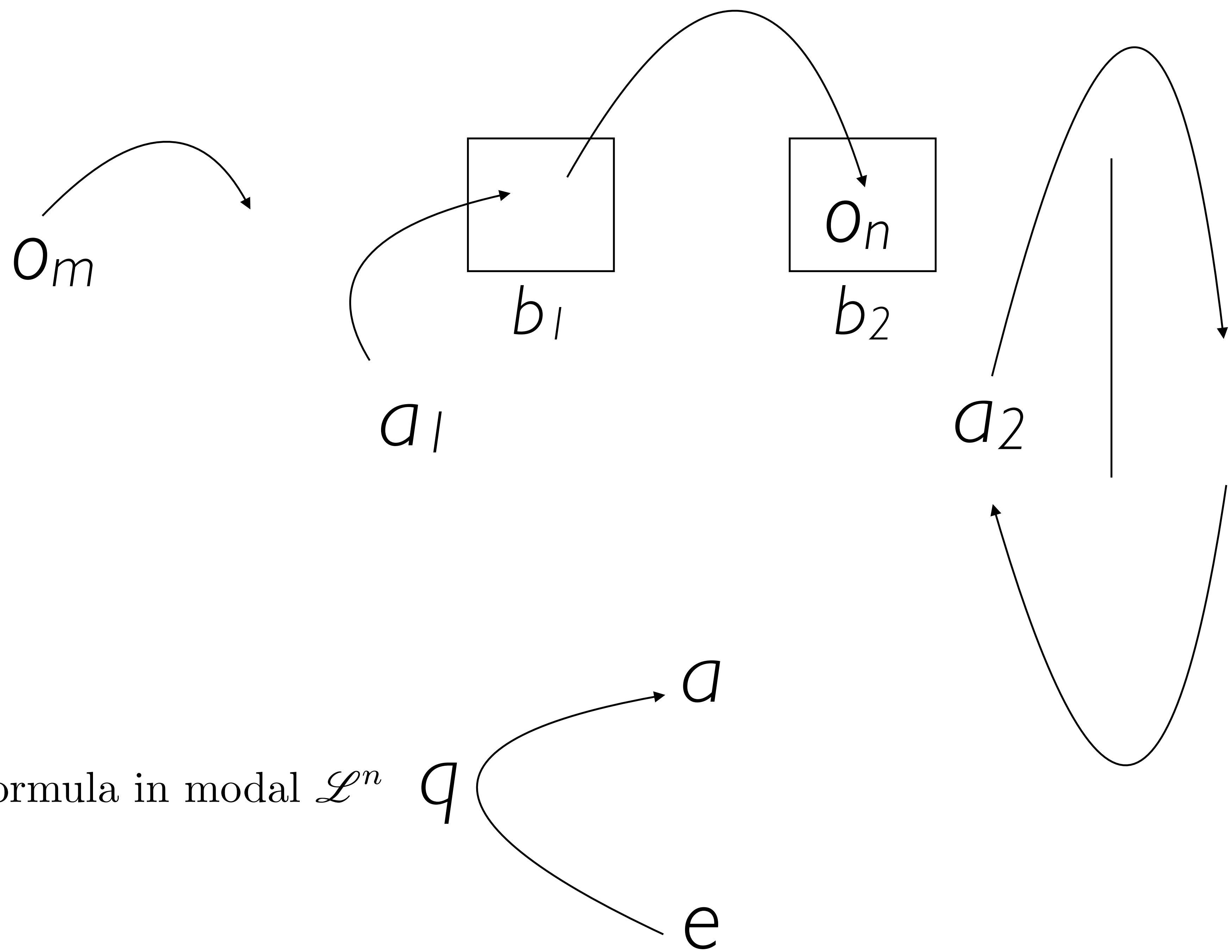


Framework for FBT'₁



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.

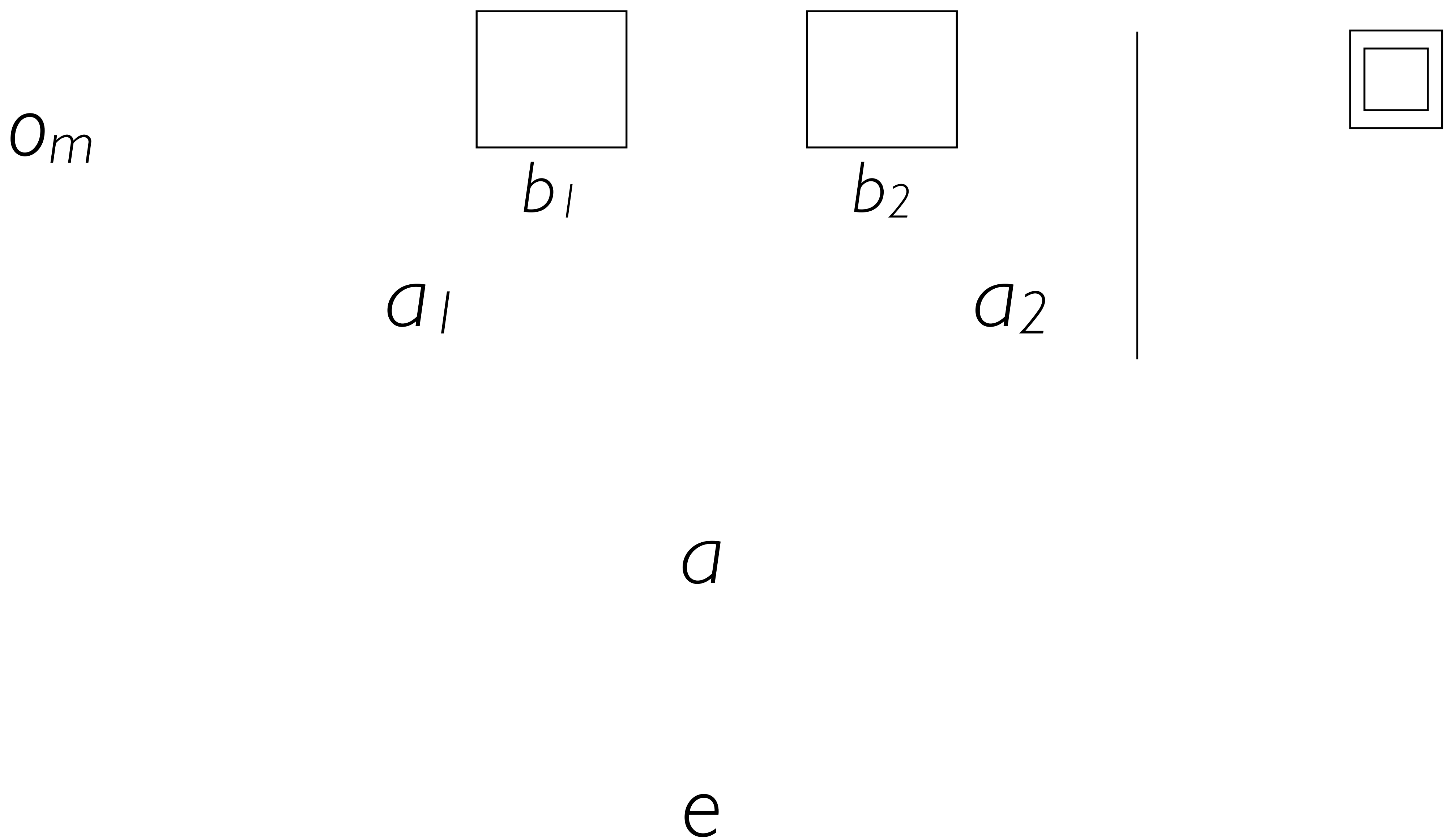
1 Introduction

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.

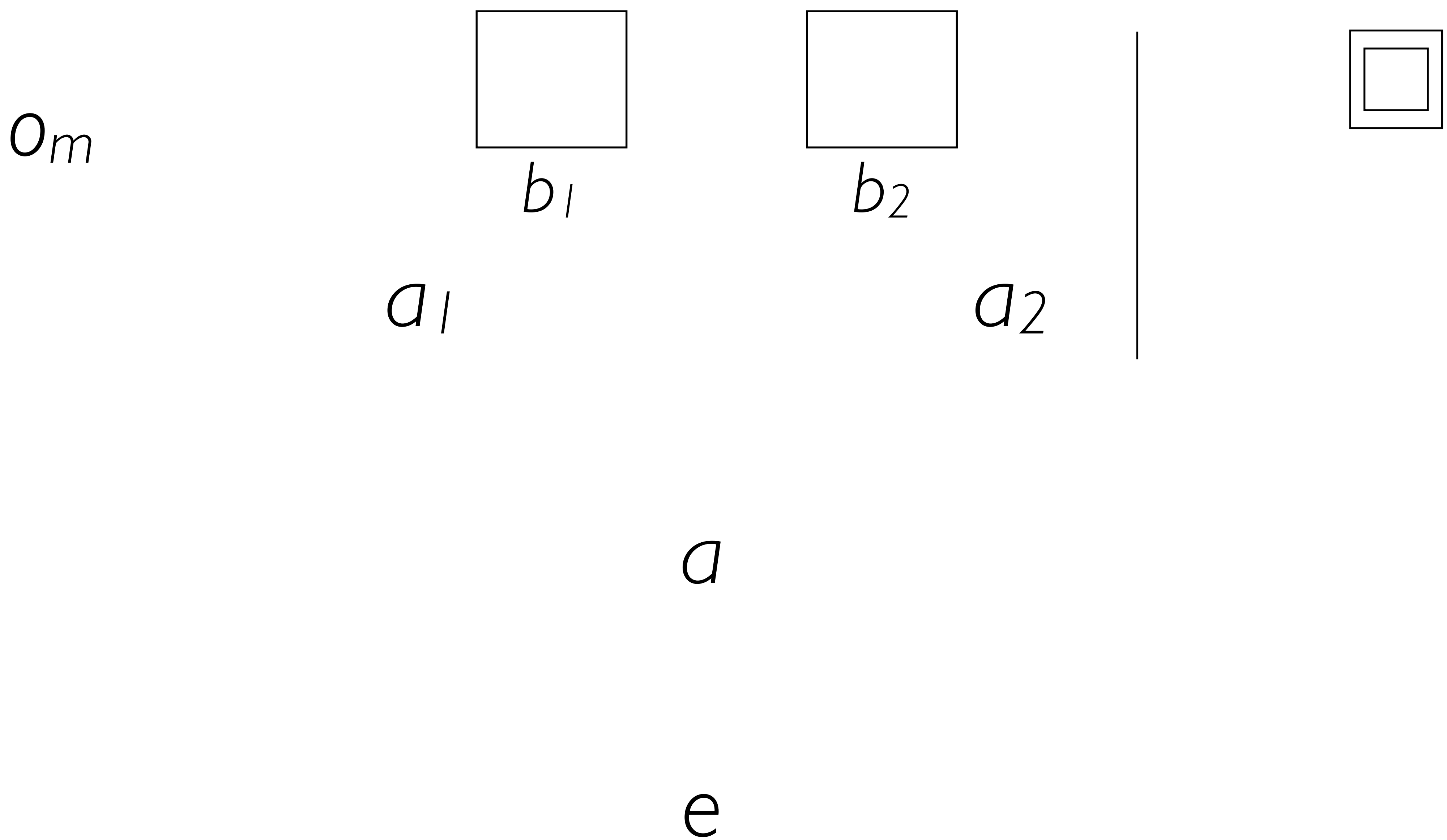
Artificial agents without a mastery of CSP would be severely handicapped in their interactions with humans. This could present problems not only for artificial agents trying to interpret human behavior, but also for artificial agents trying to interpret the behavior of one another. When a system exhibits 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$^{1}_{2}$

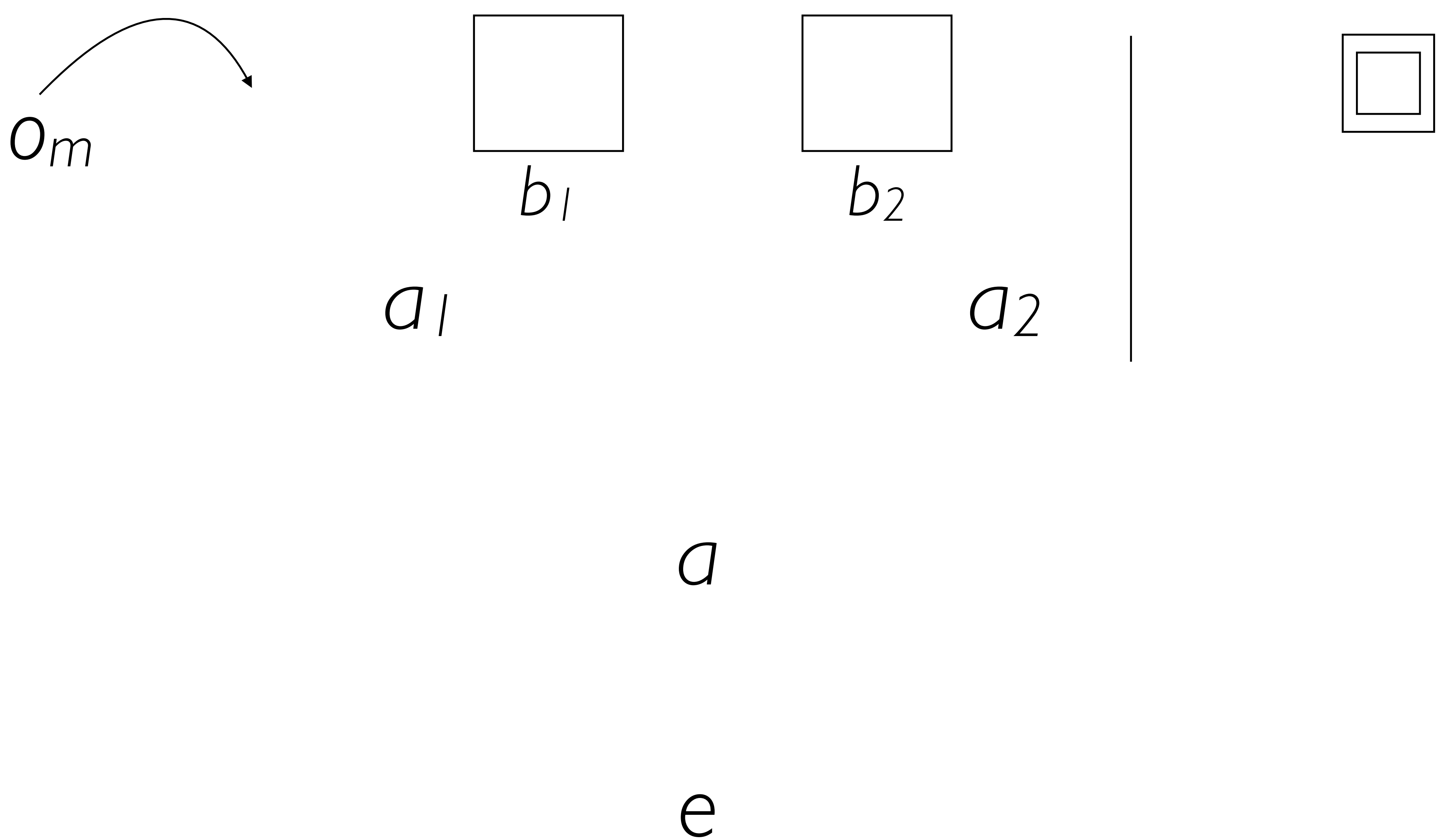


Framework for FBT^1_2
(seven timepoints)



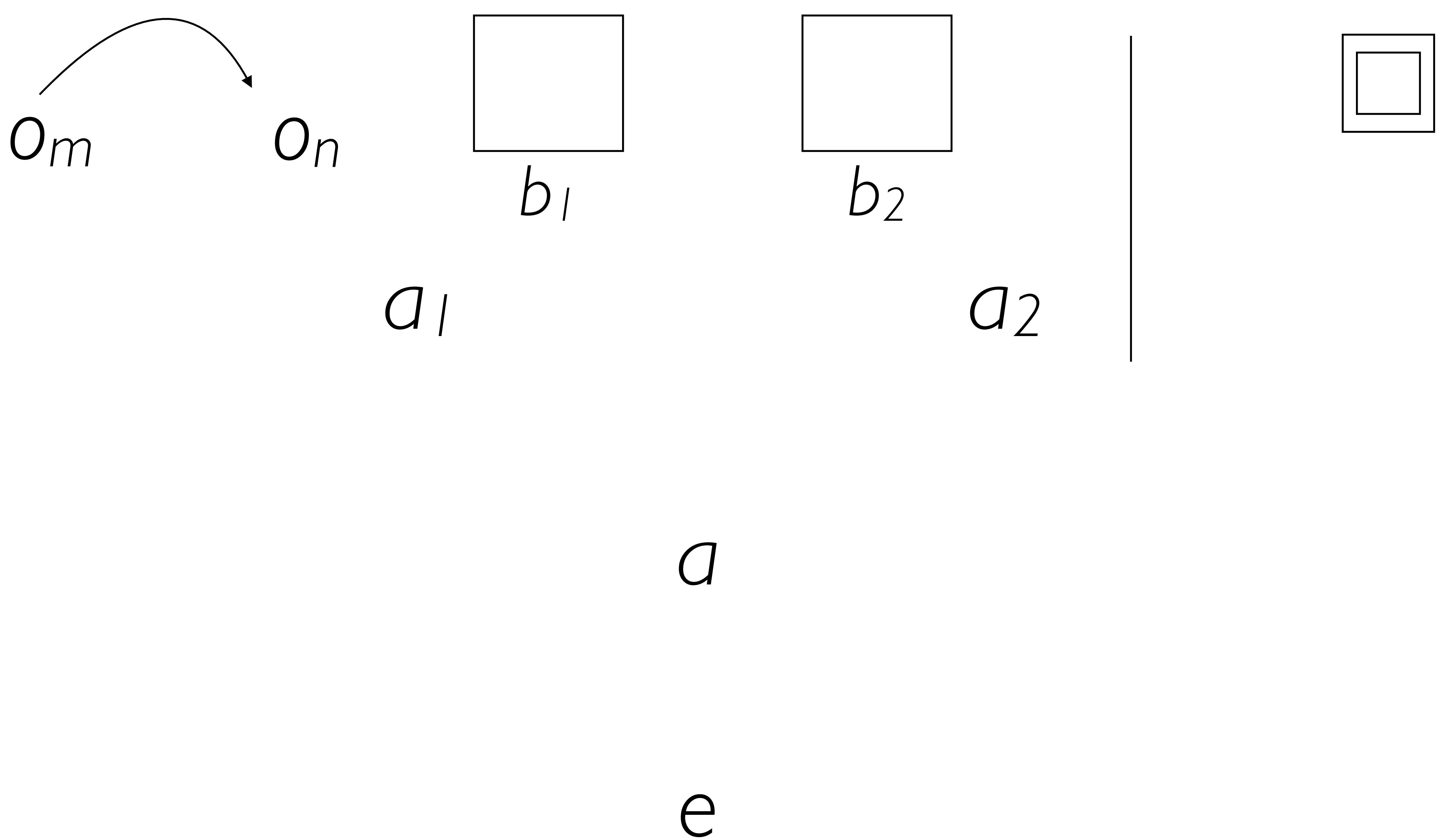
Framework for FBT^1_2

(seven timepoints)



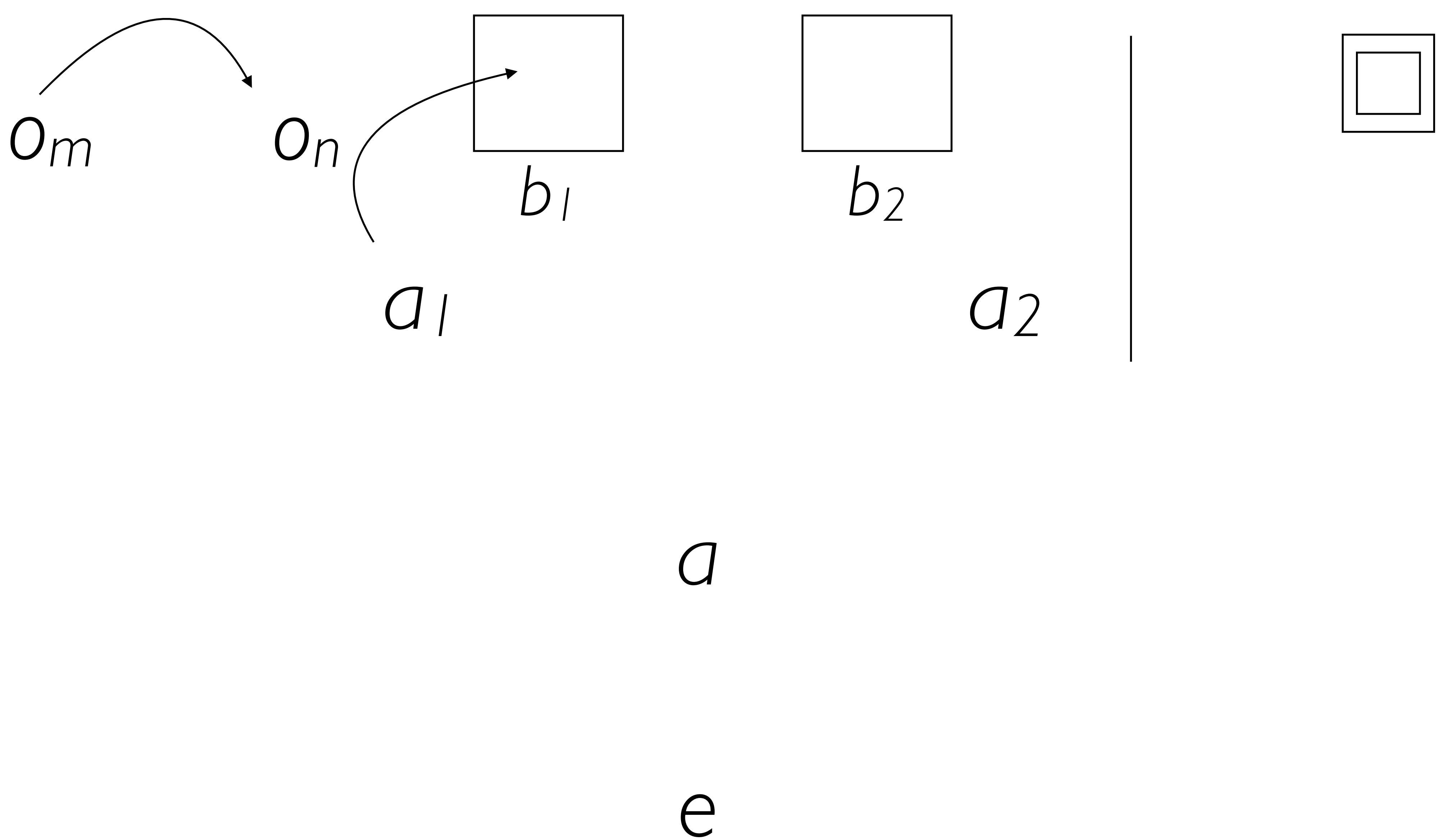
Framework for $\text{FBT}^{1}{}_{2}$

(seven timepoints)



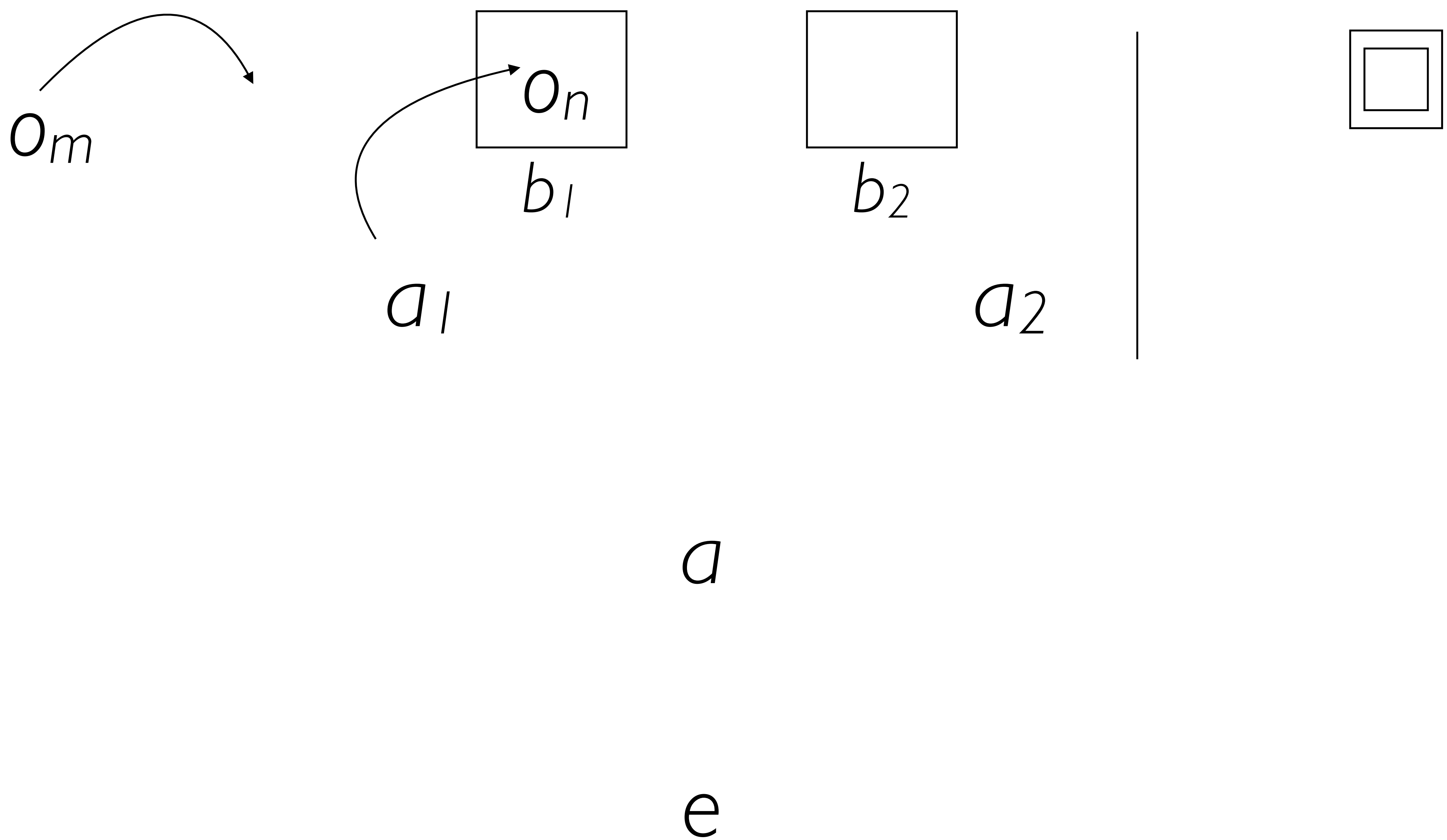
Framework for FBT$^1{}_2$

(seven timepoints)

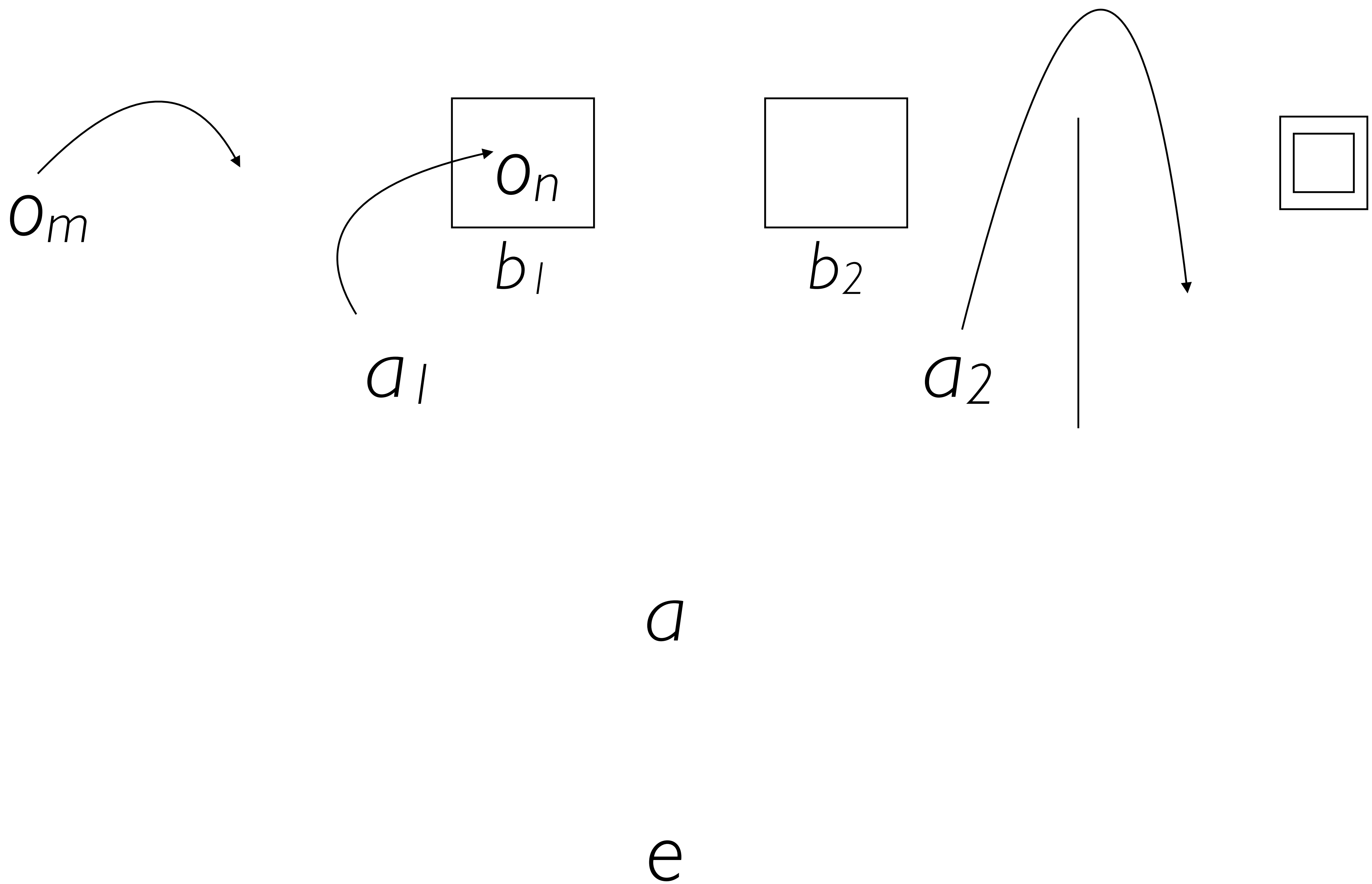


Framework for FBT^1_2

(seven timepoints)

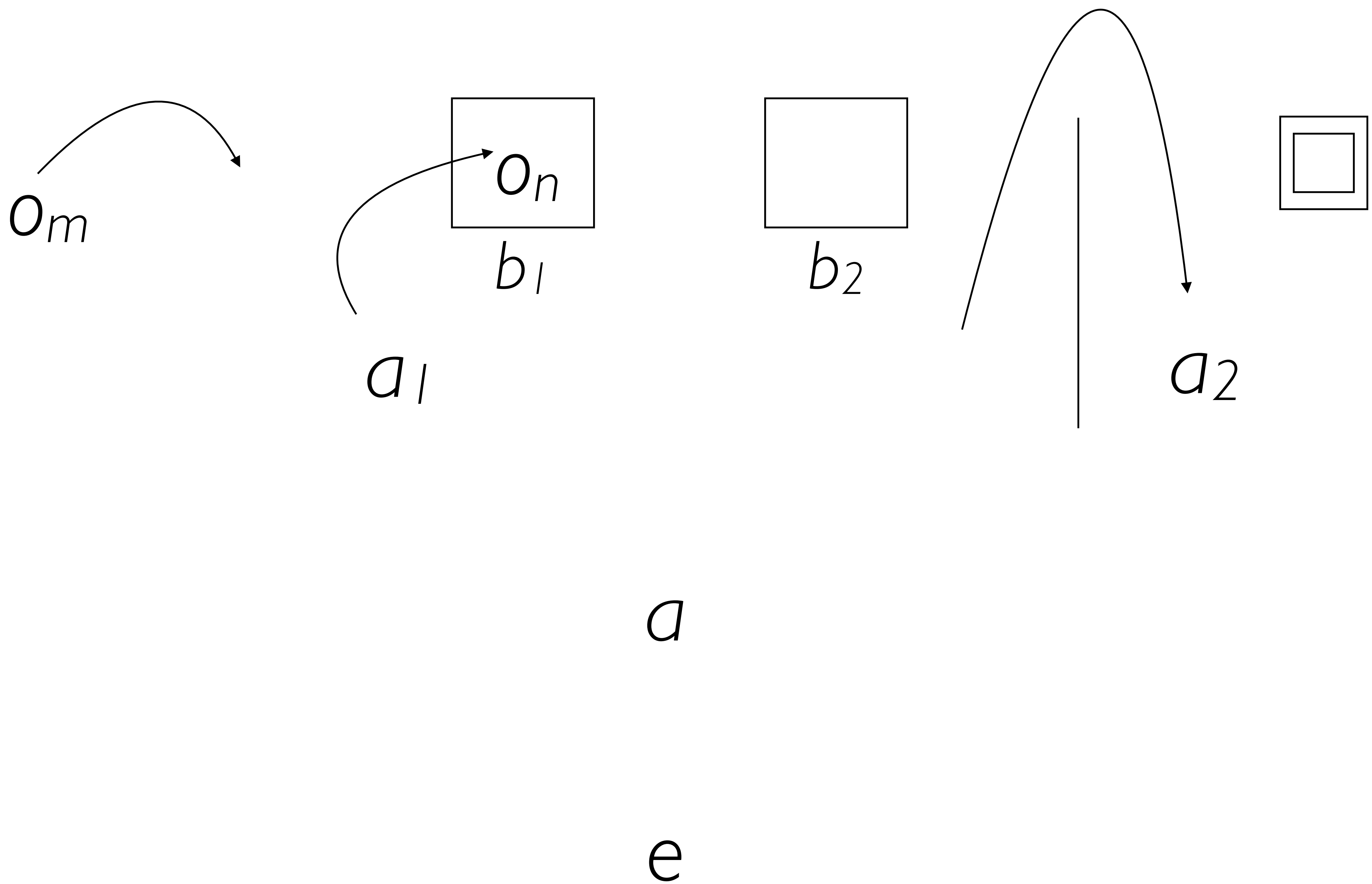


Framework for FBT^1_2
(seven timepoints)



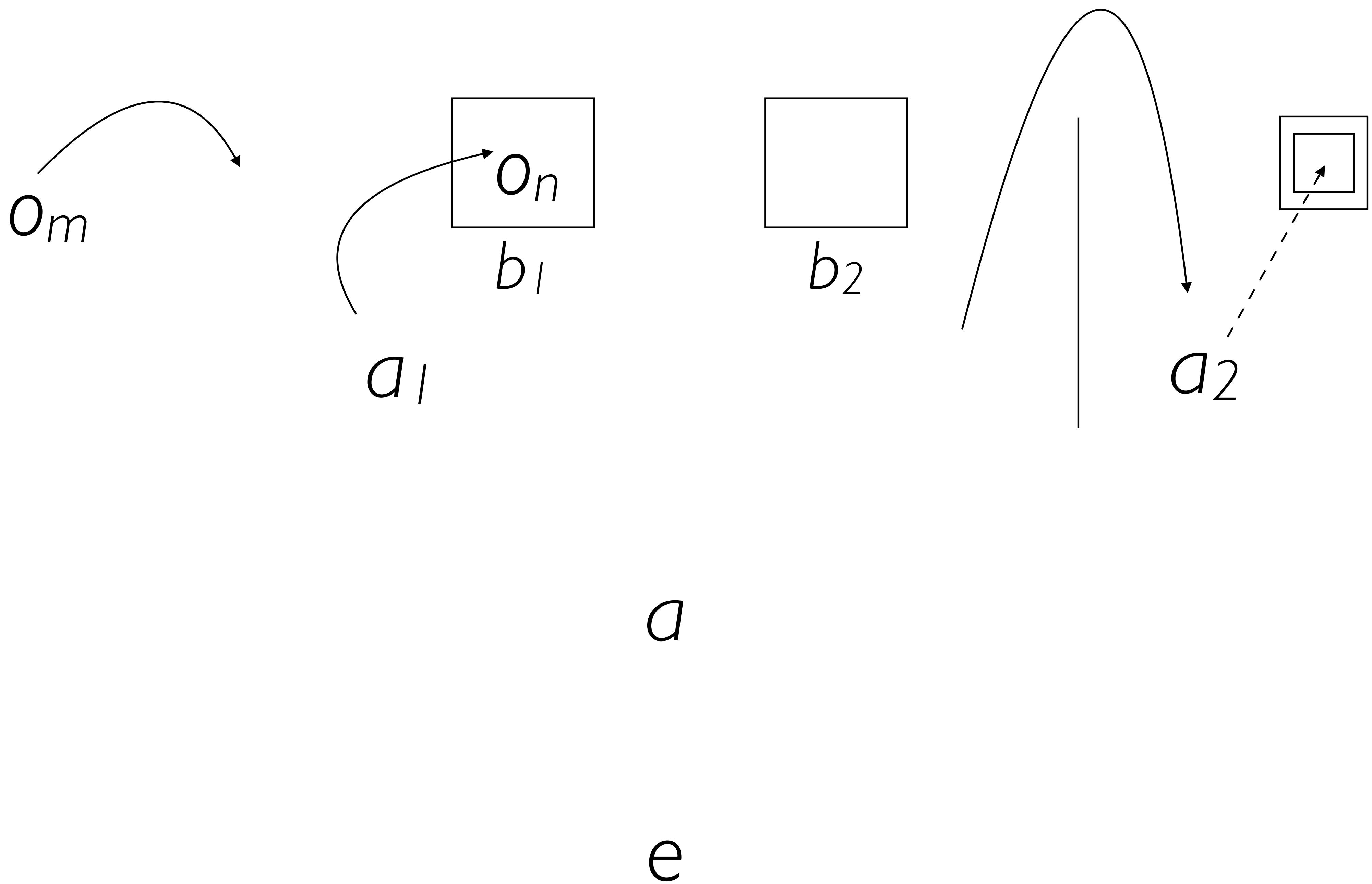
Framework for $FBT^1{}_2$

(seven timepoints)

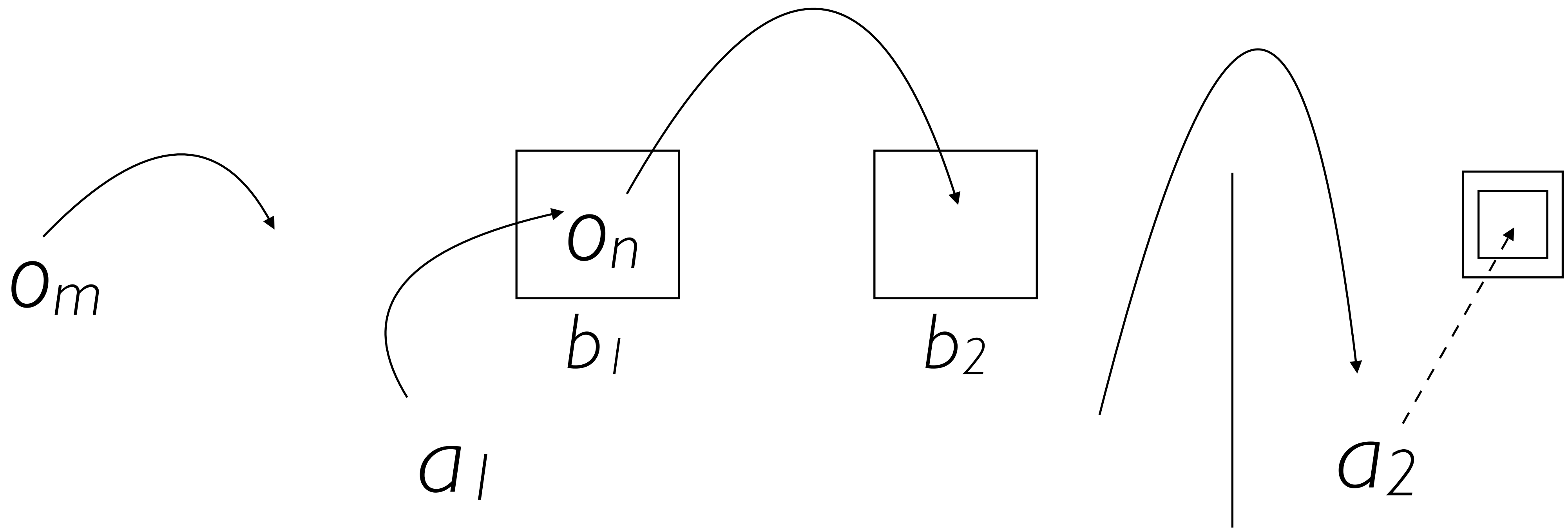


Framework for FBT$^1{}_2$

(seven timepoints)



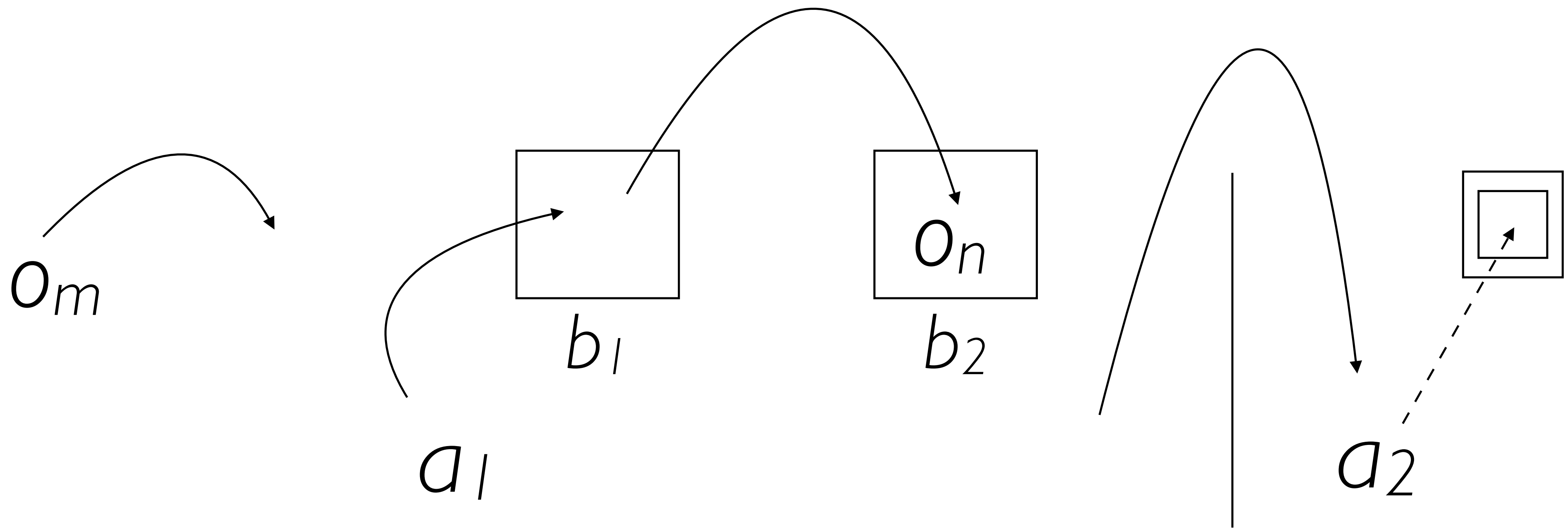
Framework for FBT^1_2
(seven timepoints)



a

e

Framework for FBT^1_2
(seven timepoints)

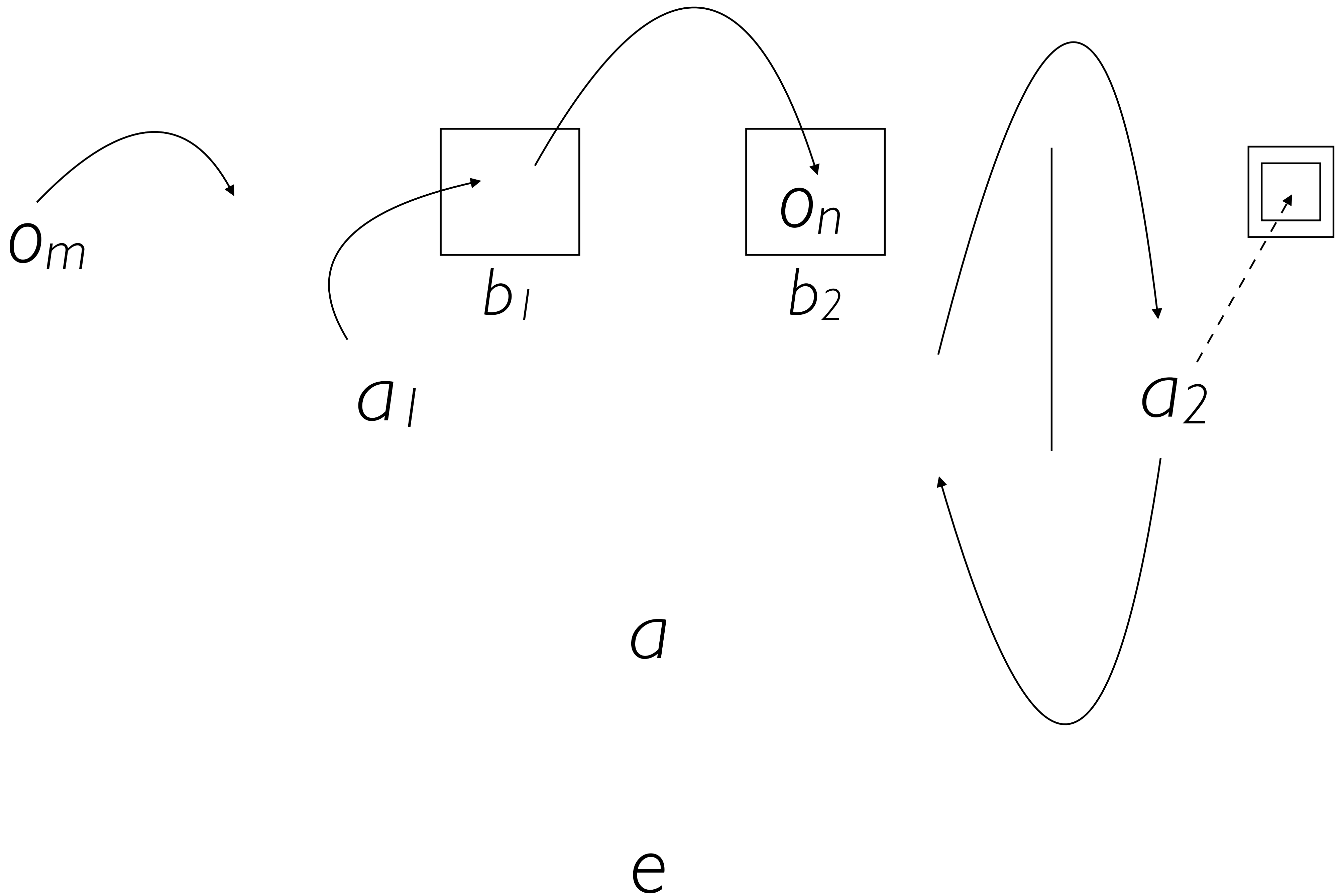


a

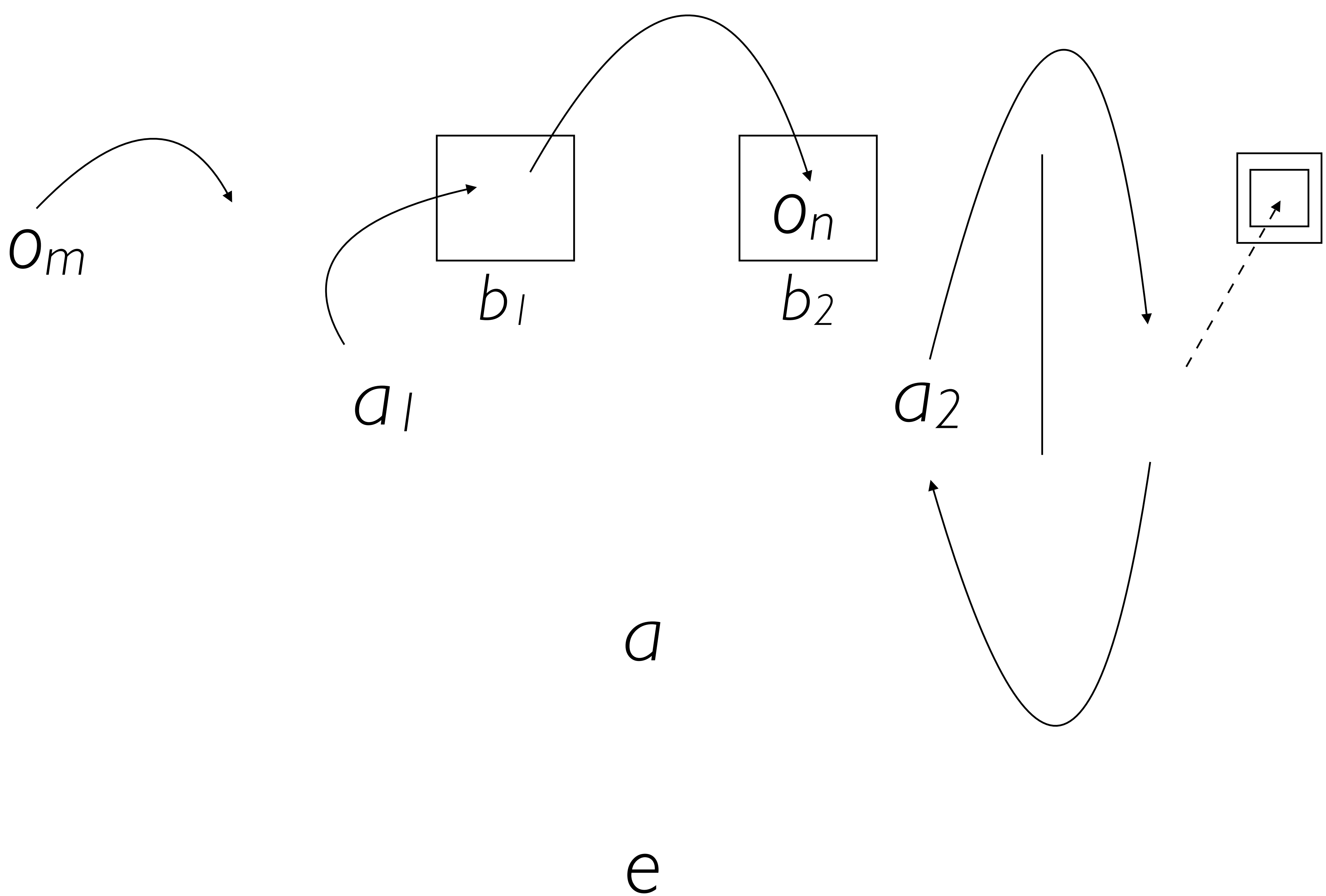
e

Framework for FBT$^1{}_2$

(seven timepoints)

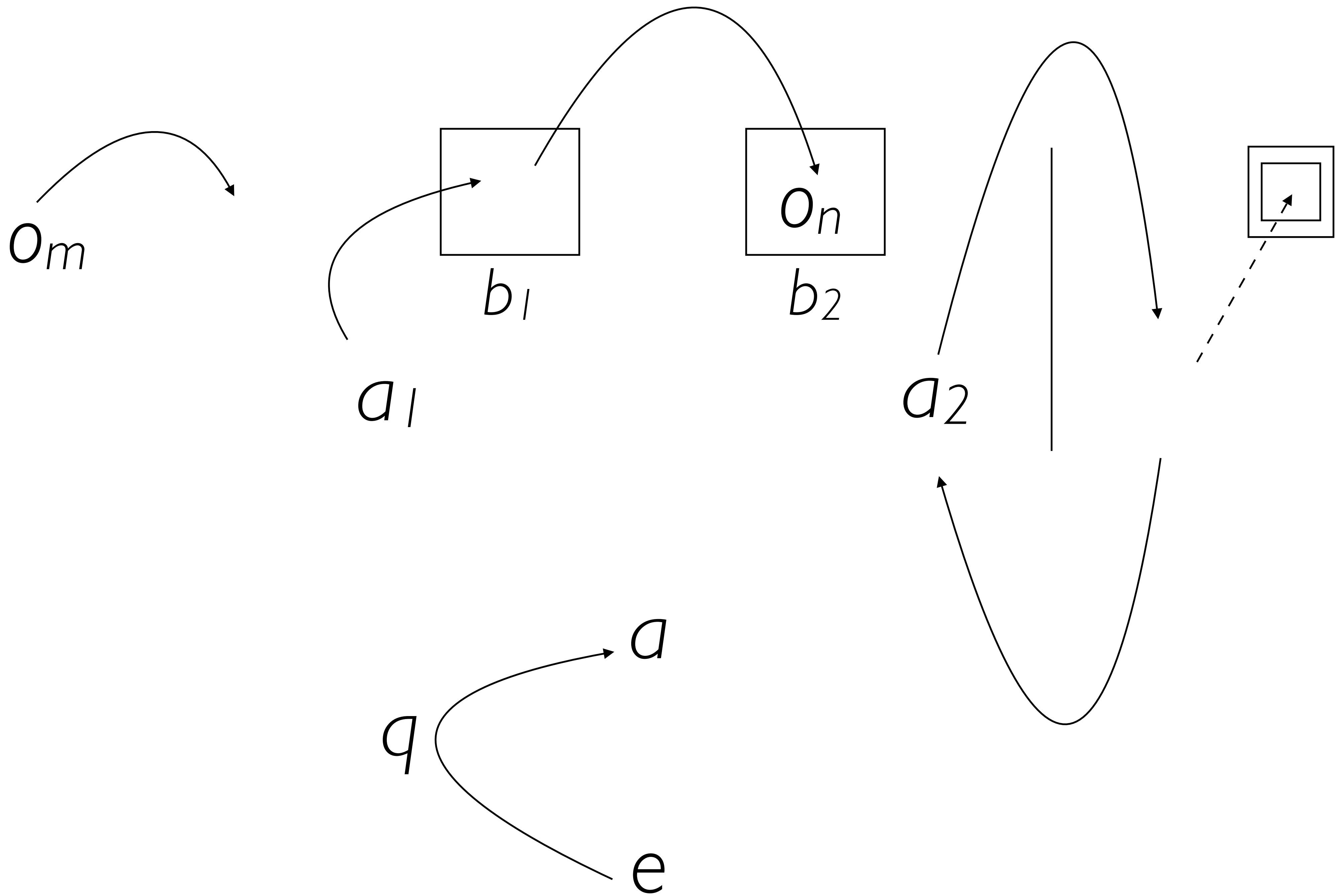


Framework for $FBT^1{}_2$
(seven timepoints)



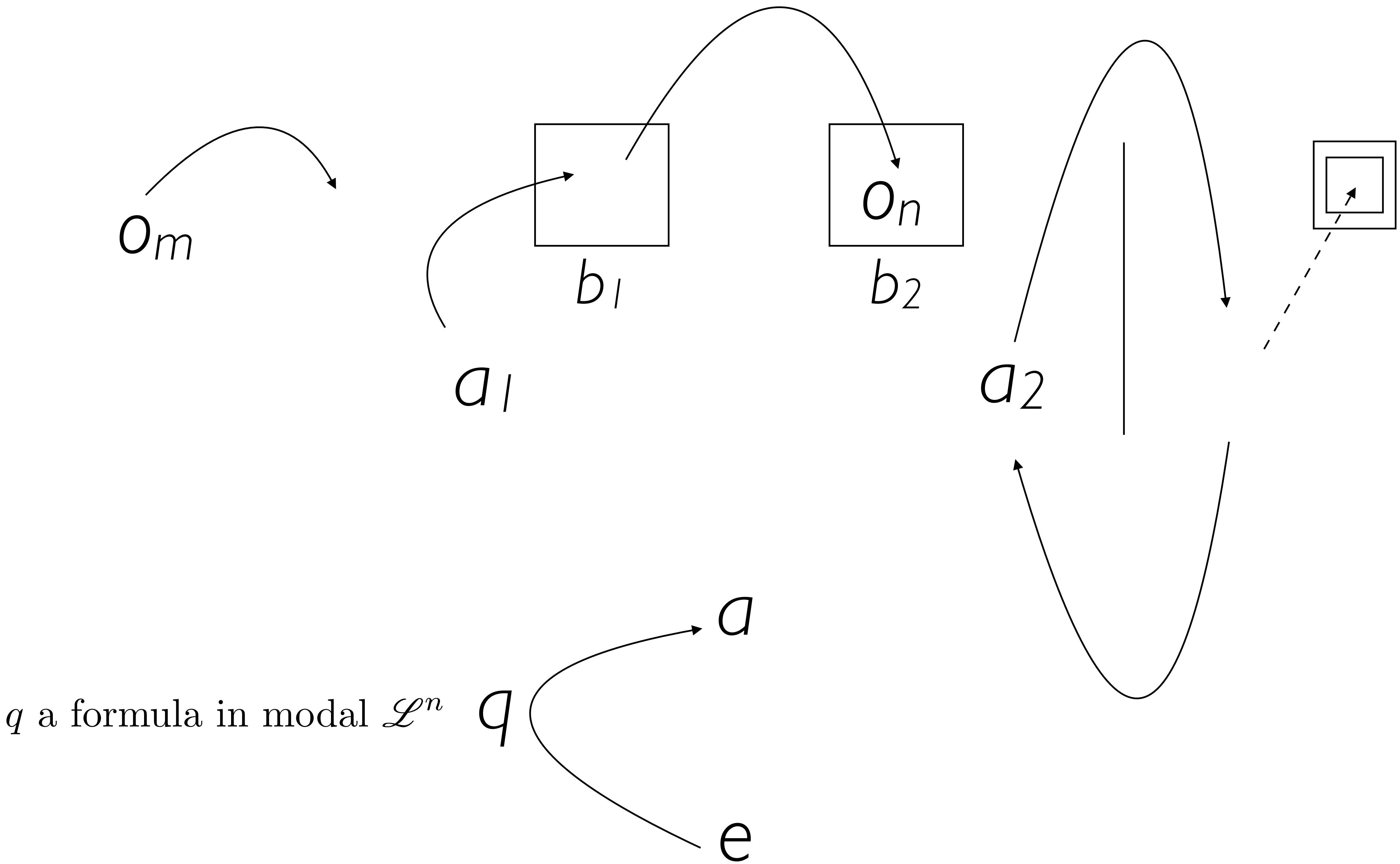
Framework for $FBT^1{}_2$

(seven timepoints)

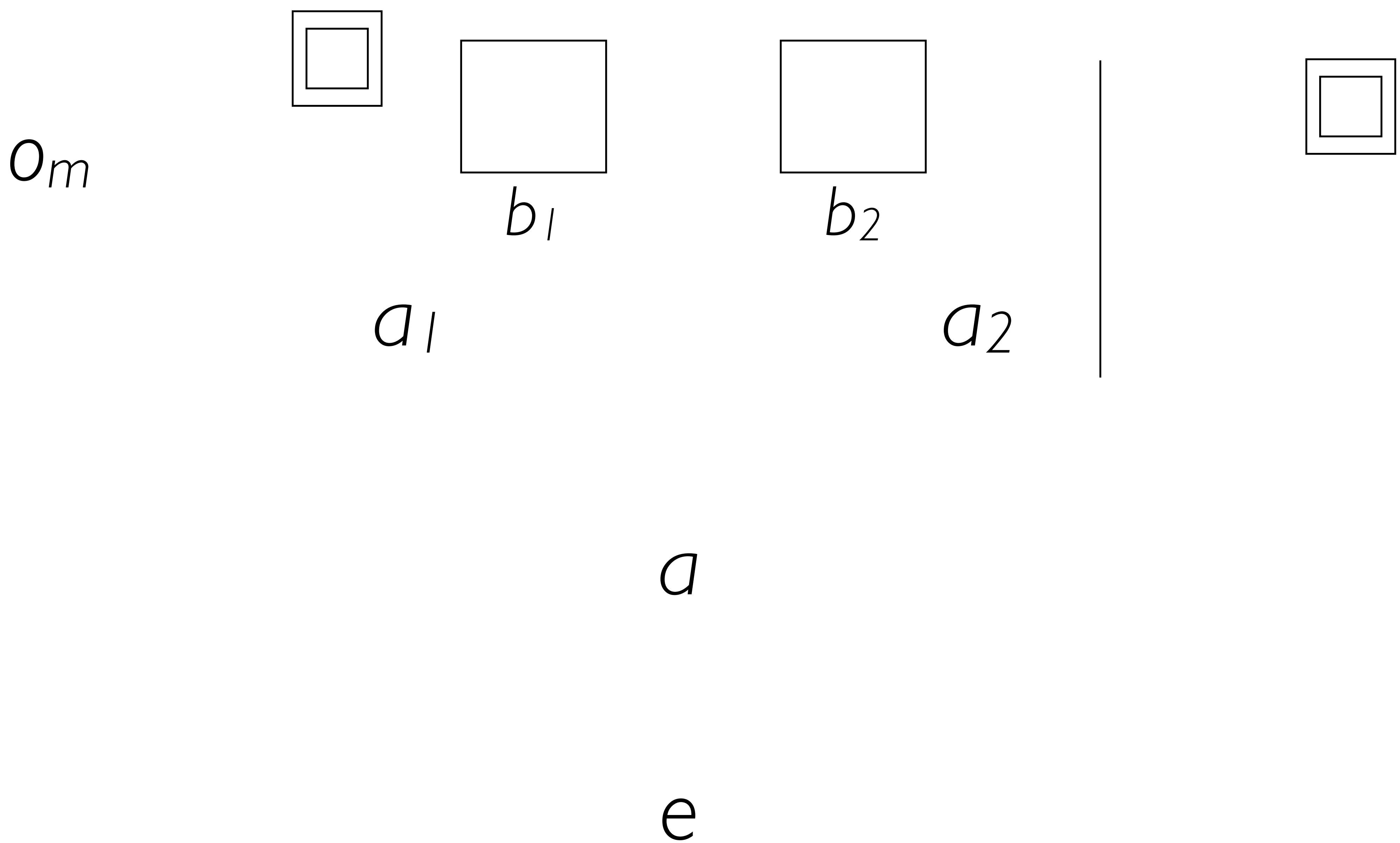


Framework for $FBT^1{}_2$

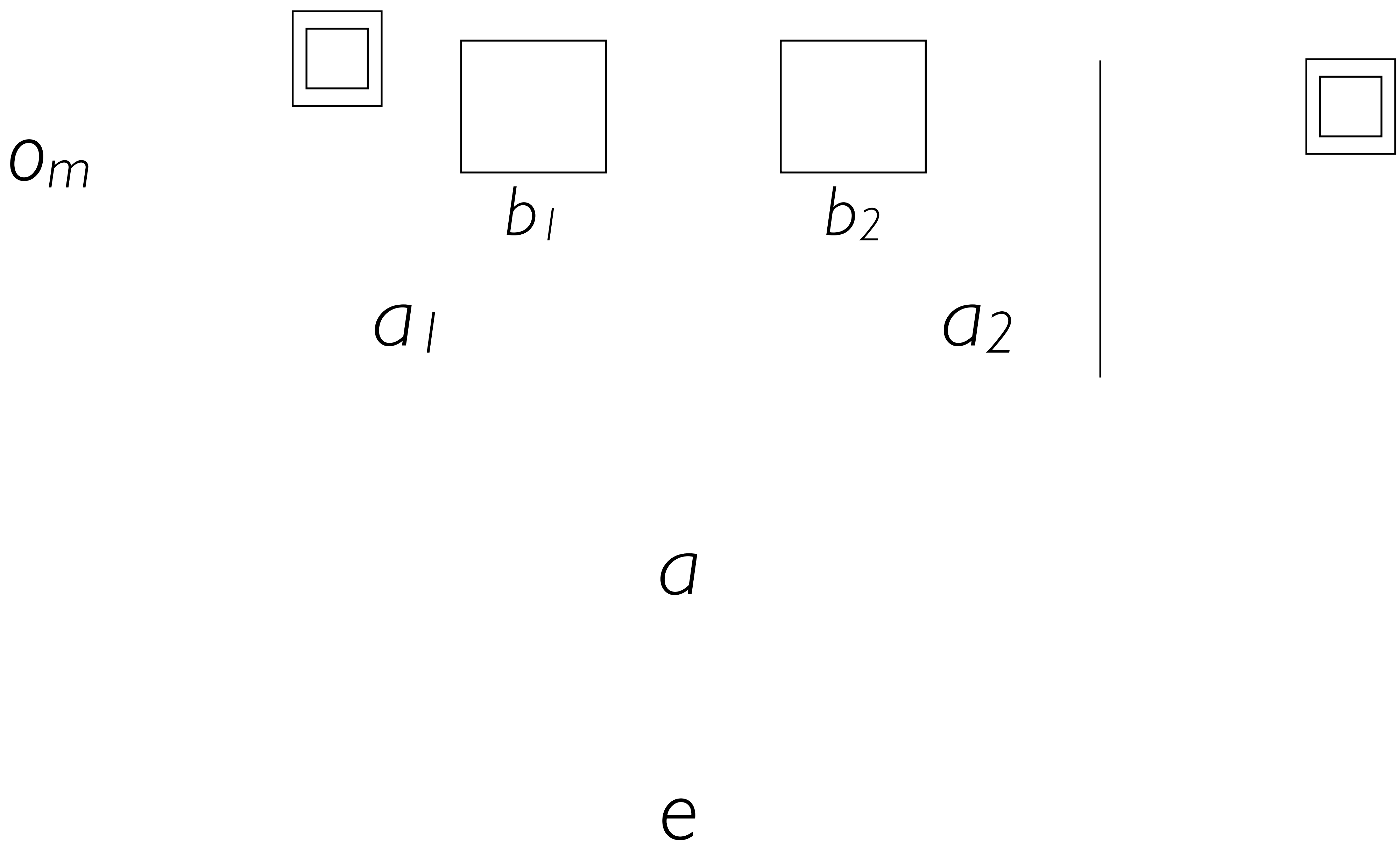
(seven timepoints)



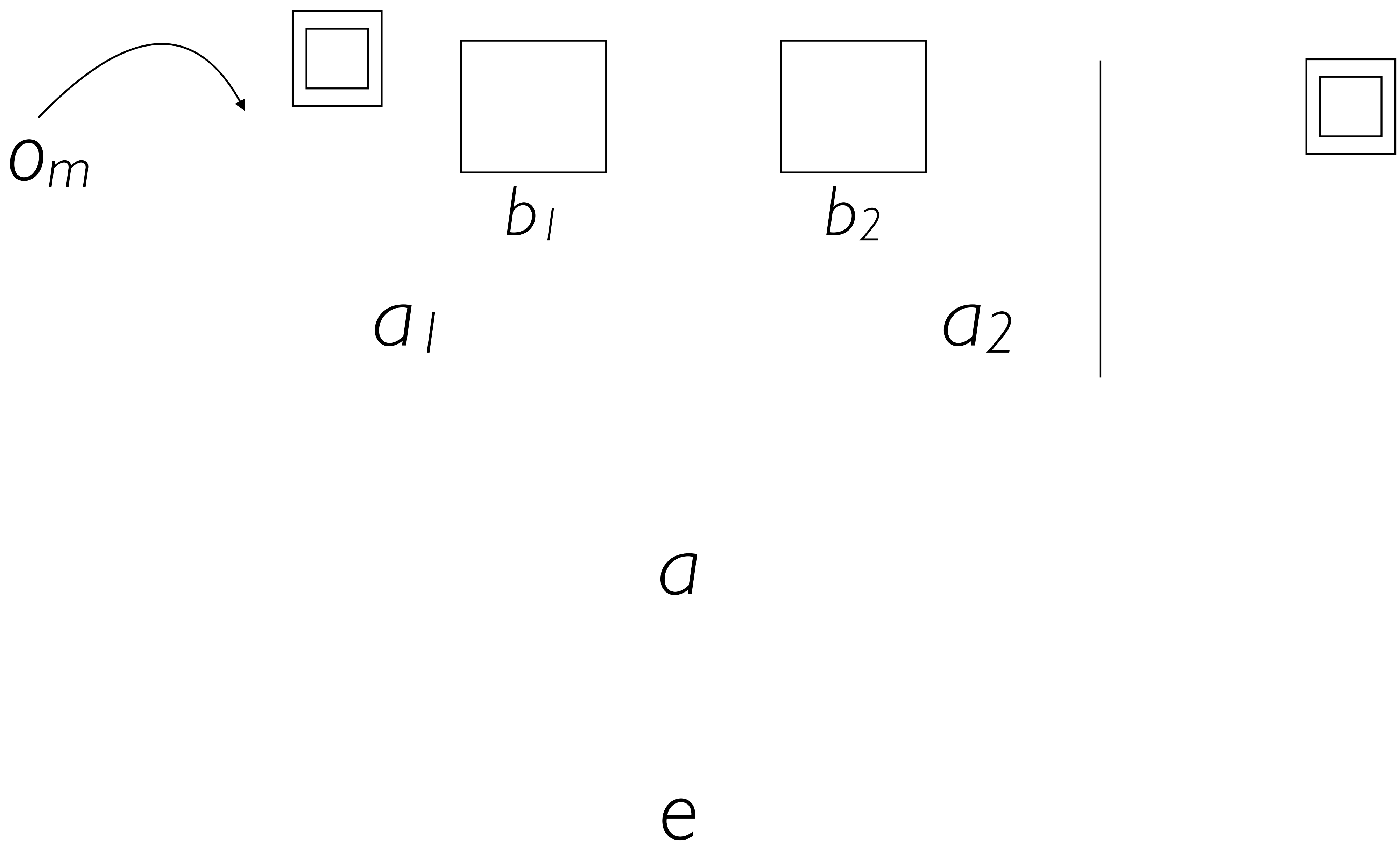
Framework for FBT^1_3



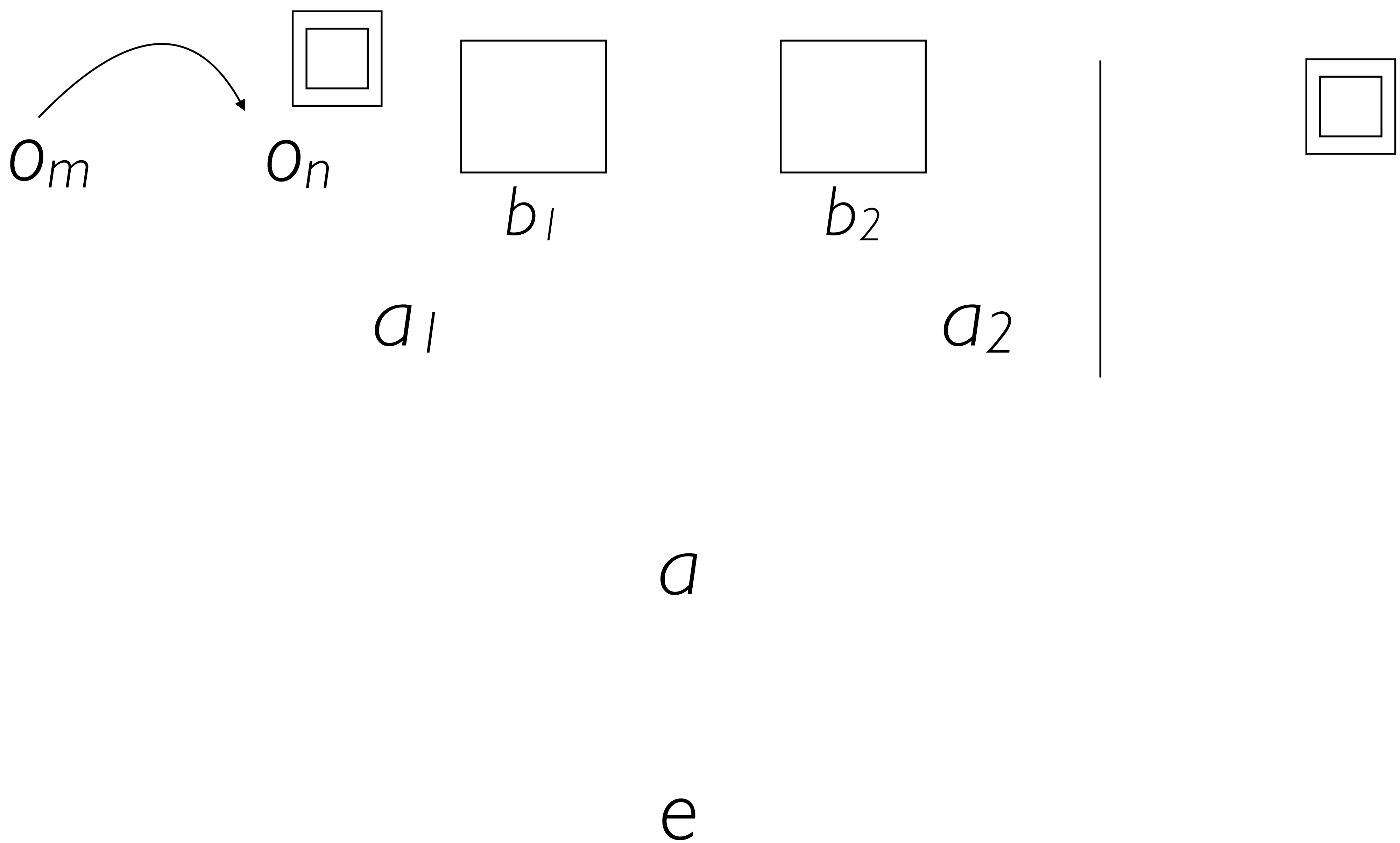
Framework for FBT^1_3
(eight timepoints)



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

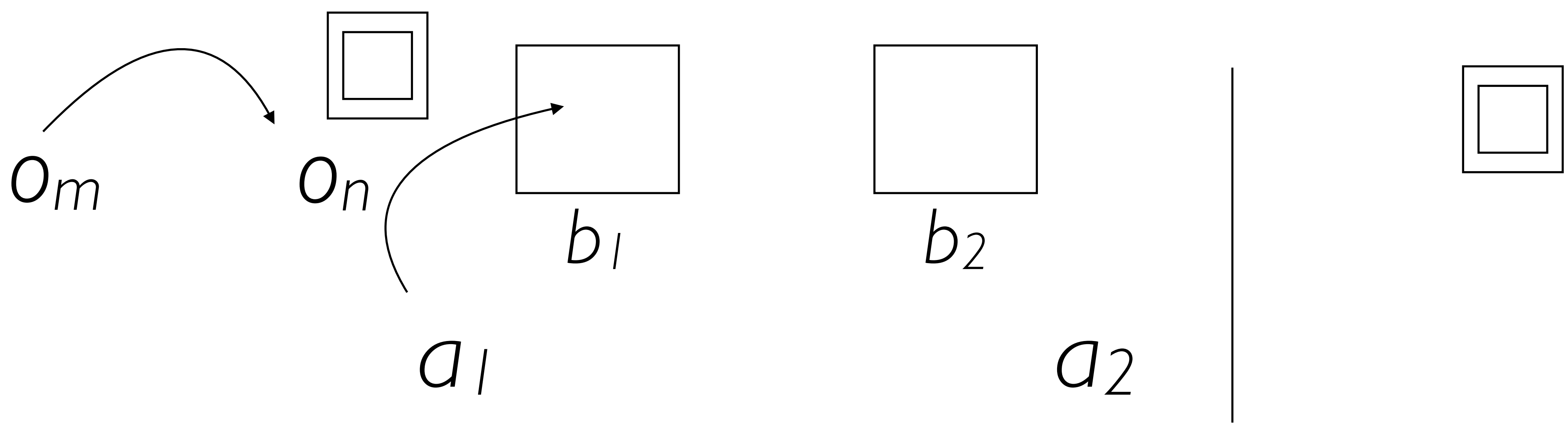


Framework for FBT^1_3
(eight timepoints)



Framework for FBT^1_3

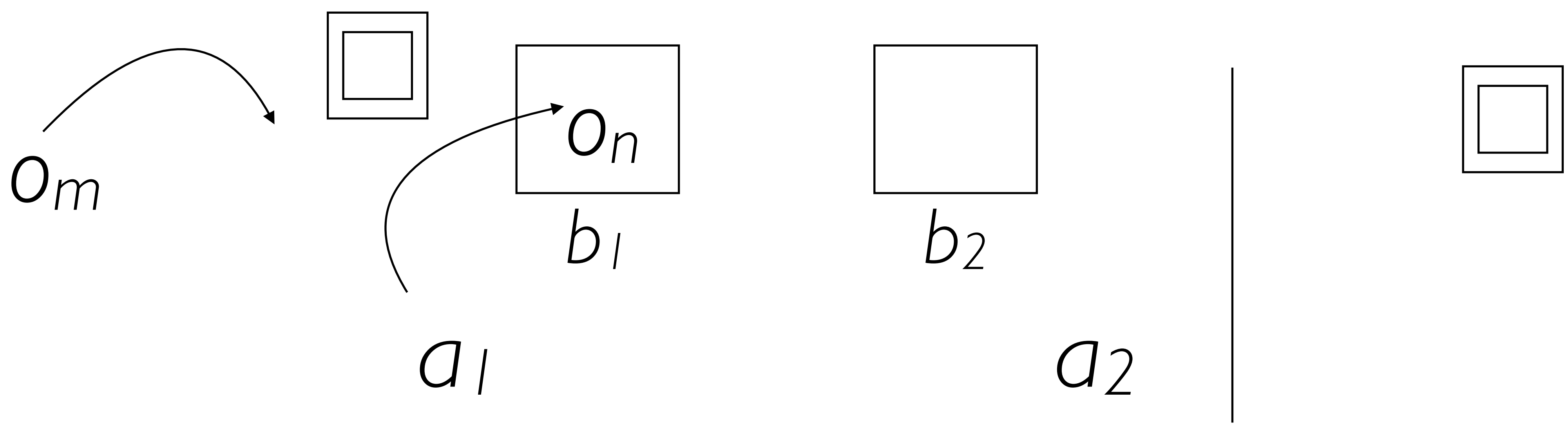
(eight timepoints)



a

e

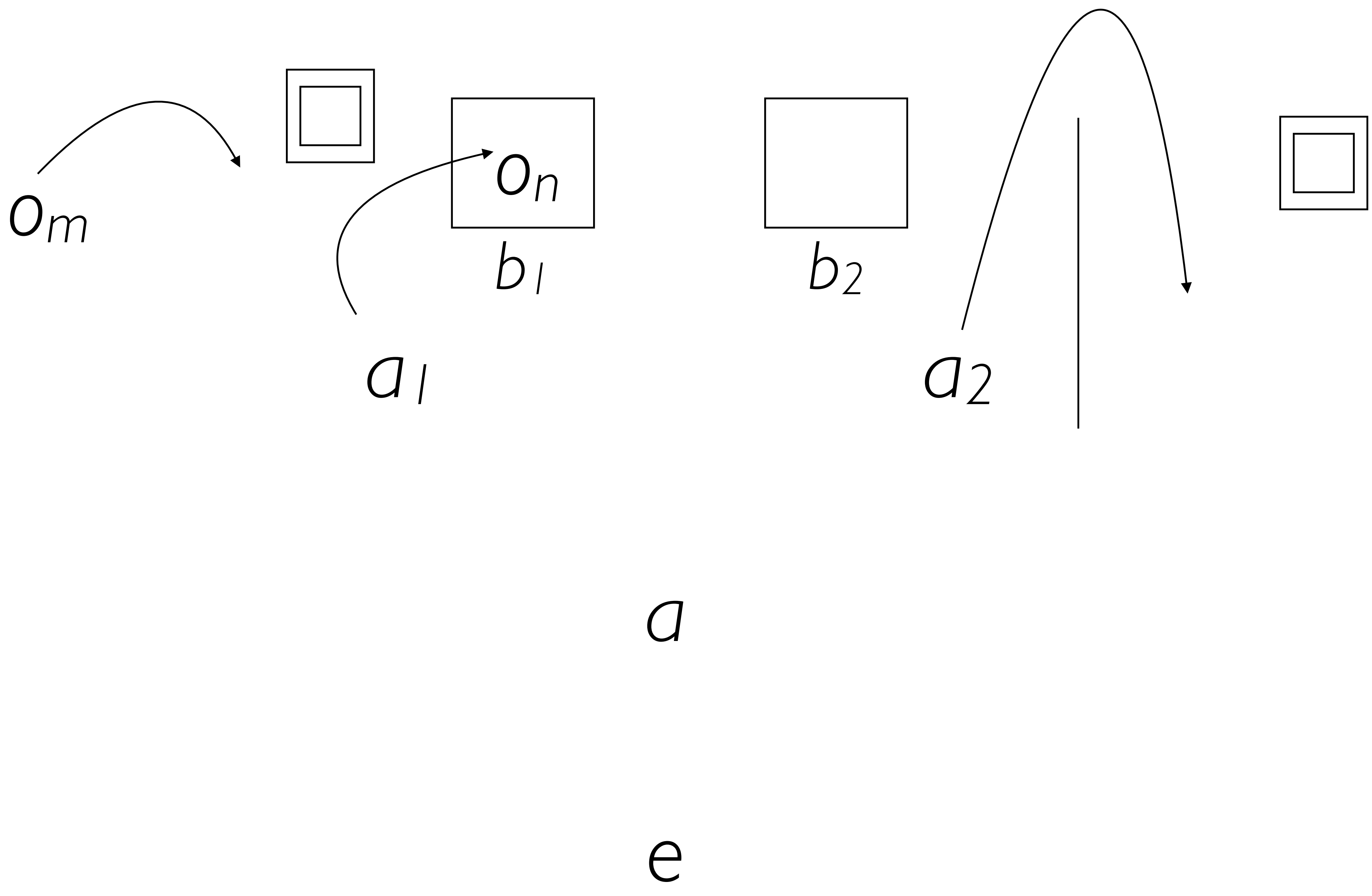
Framework for FBT$^1_3$
(eight timepoints)



a

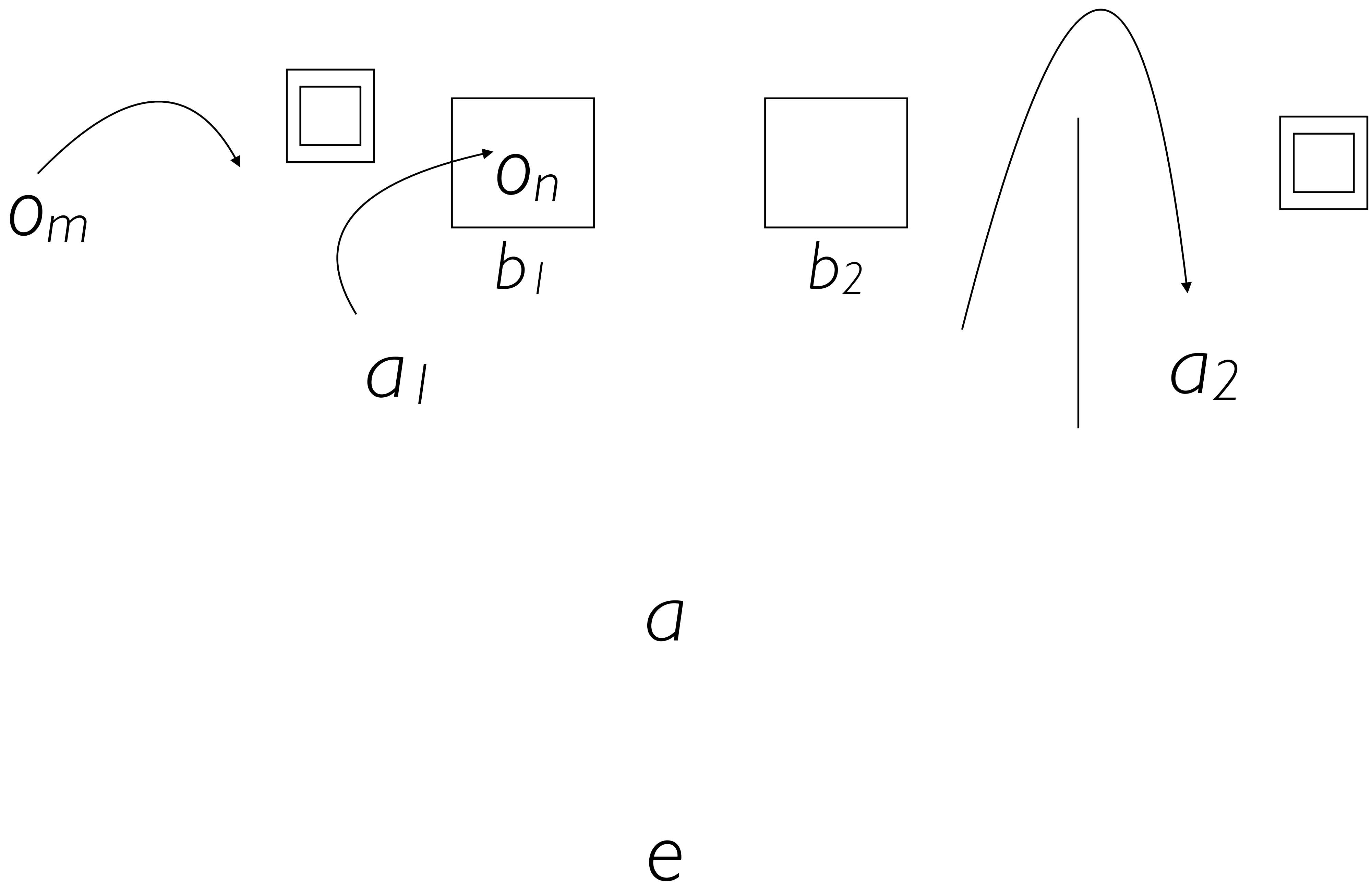
e

Framework for FBT^1_3
(eight timepoints)



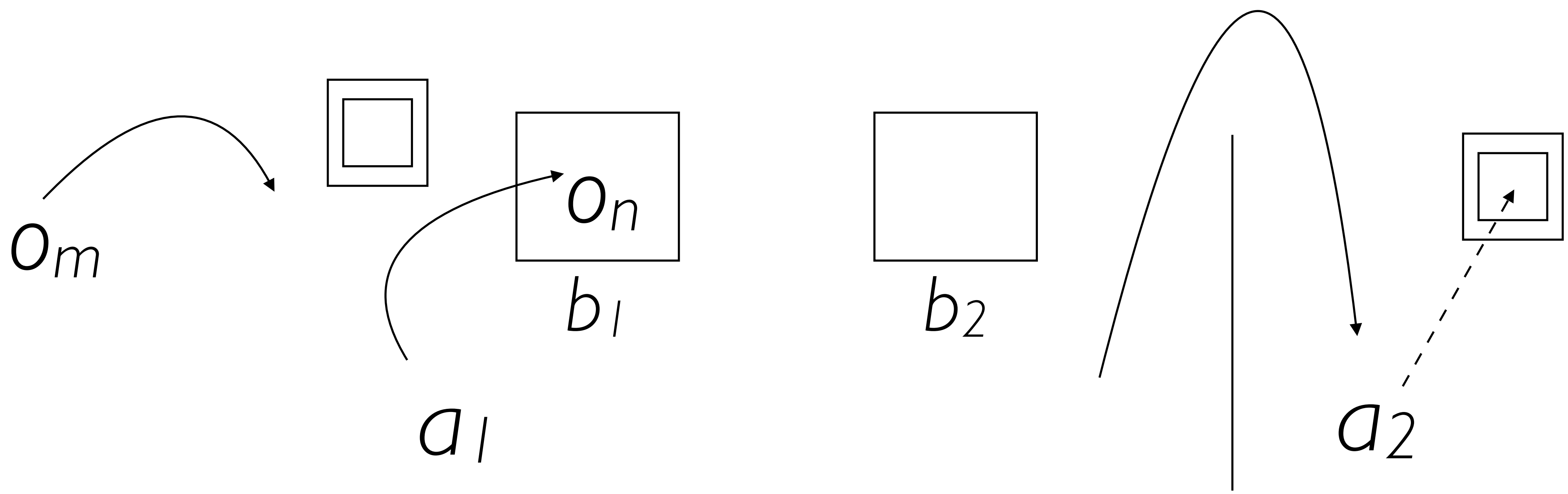
Framework for FBT^1_3

(eight timepoints)



Framework for FBT^{1}_{3}

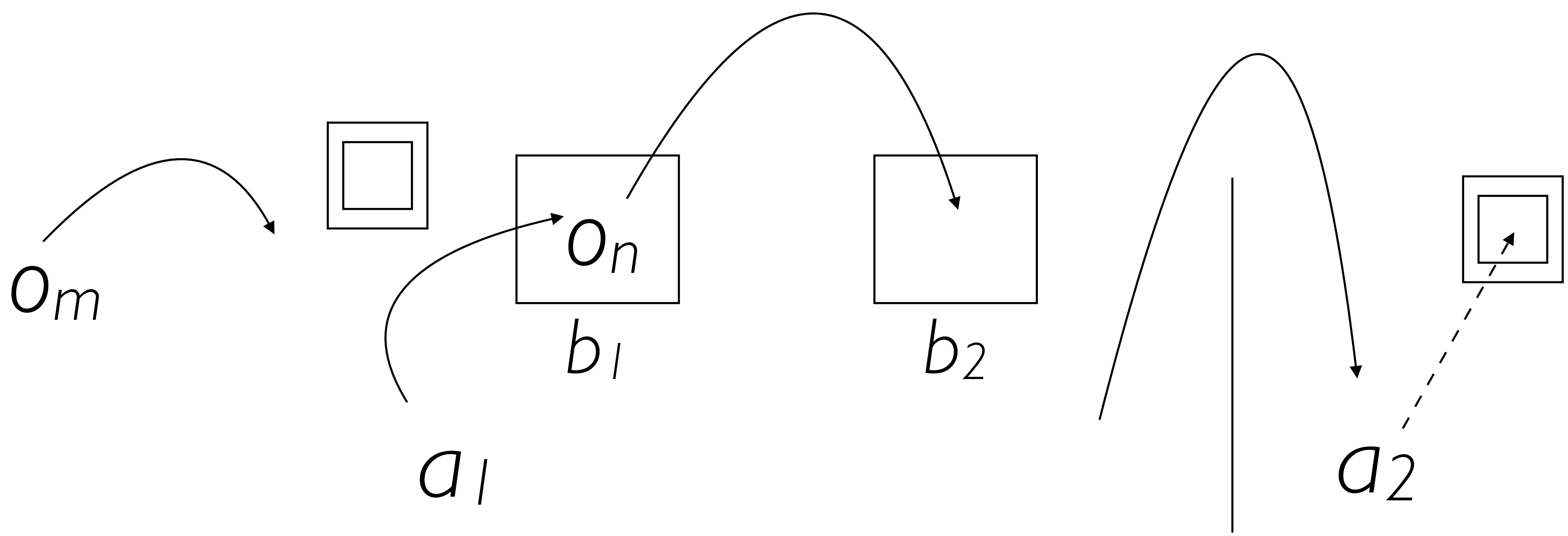
(eight timepoints)



a

e

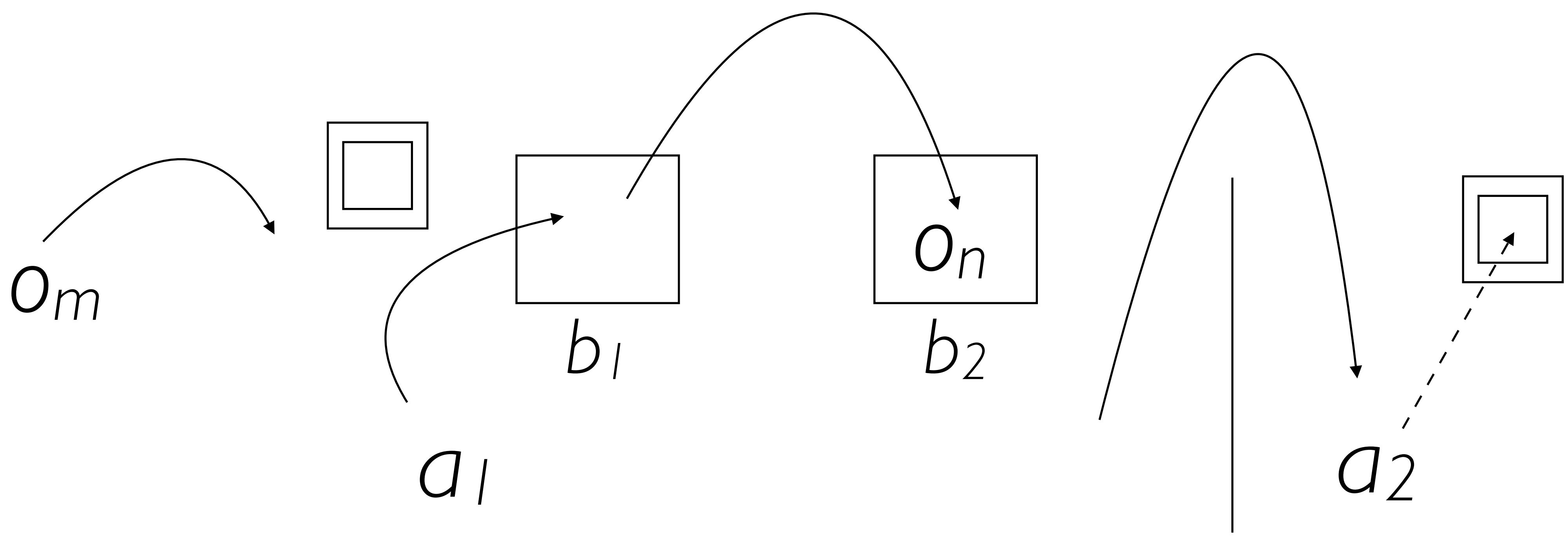
Framework for FBT^1_3
(eight timepoints)



a

e

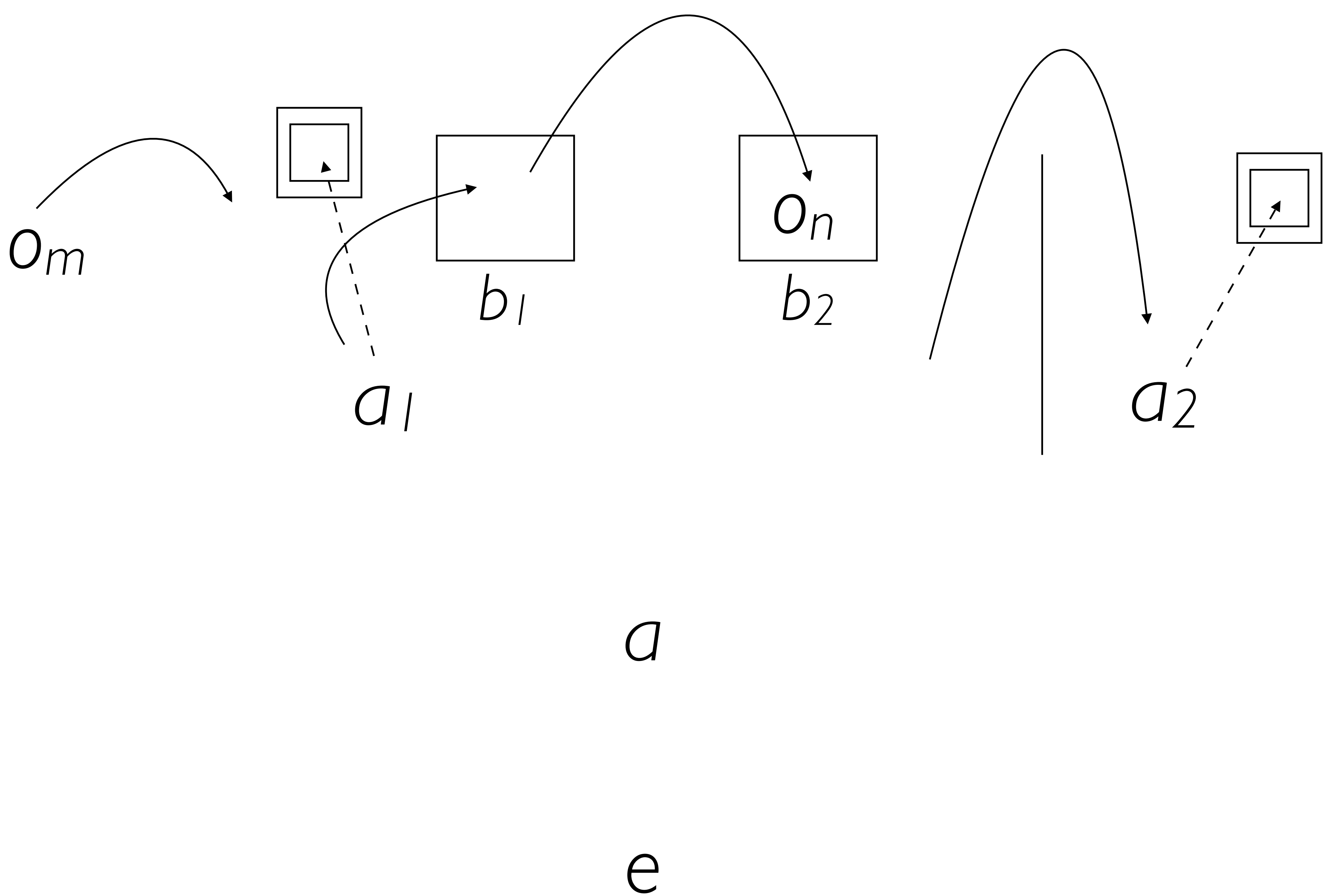
Framework for $FBT^1{}_3$
(eight timepoints)



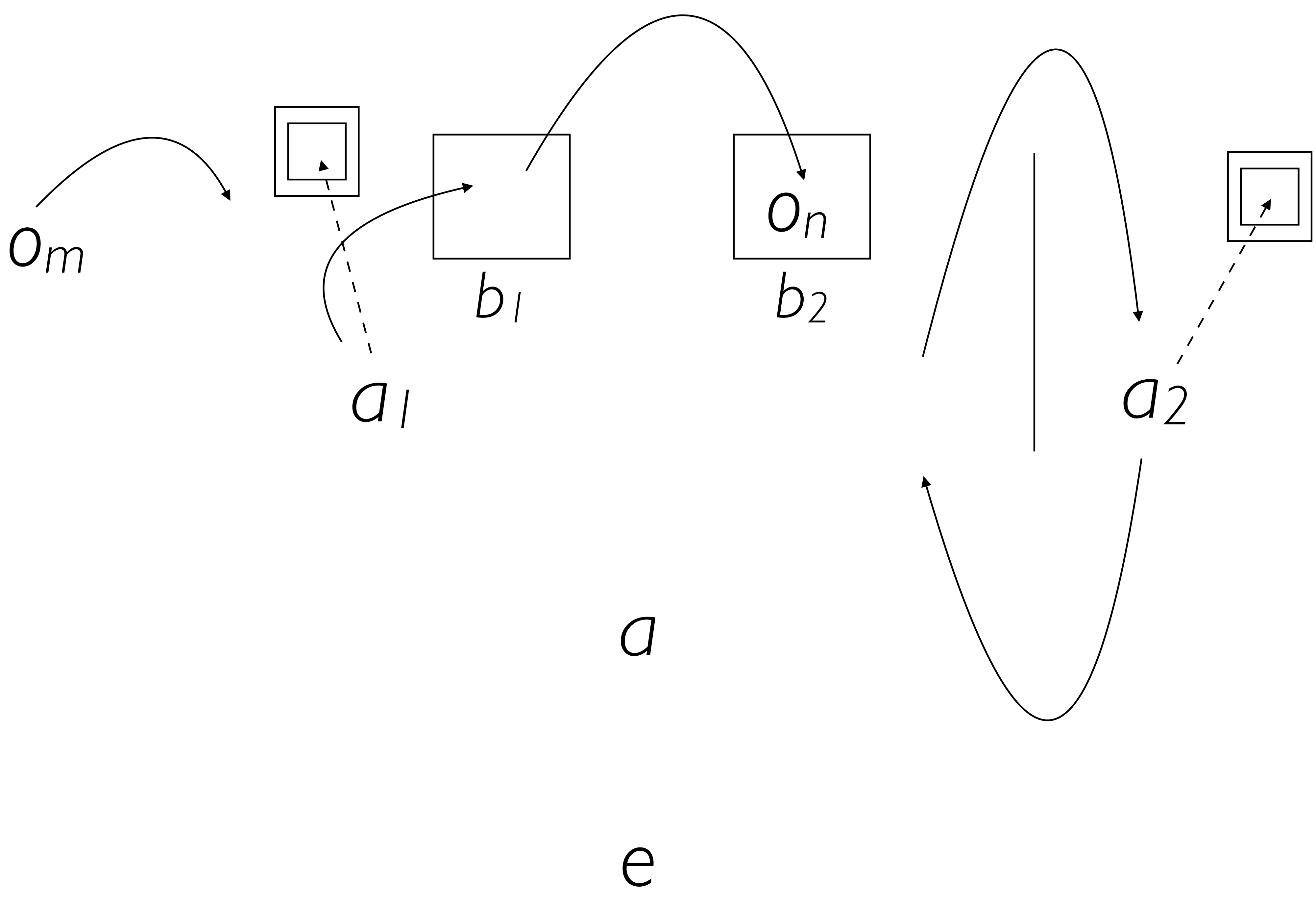
a

e

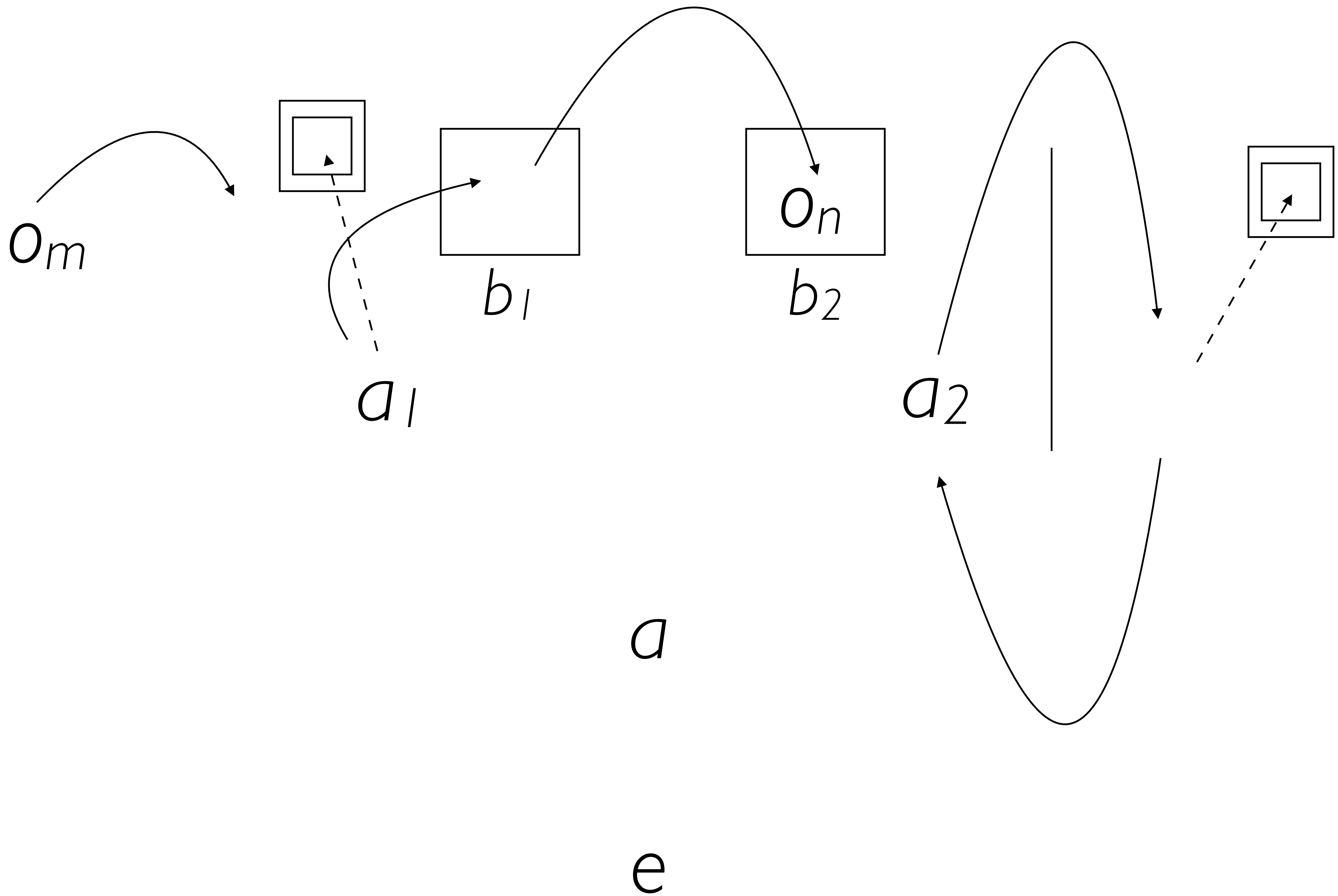
Framework for $FBT^1{}_3$
(eight timepoints)



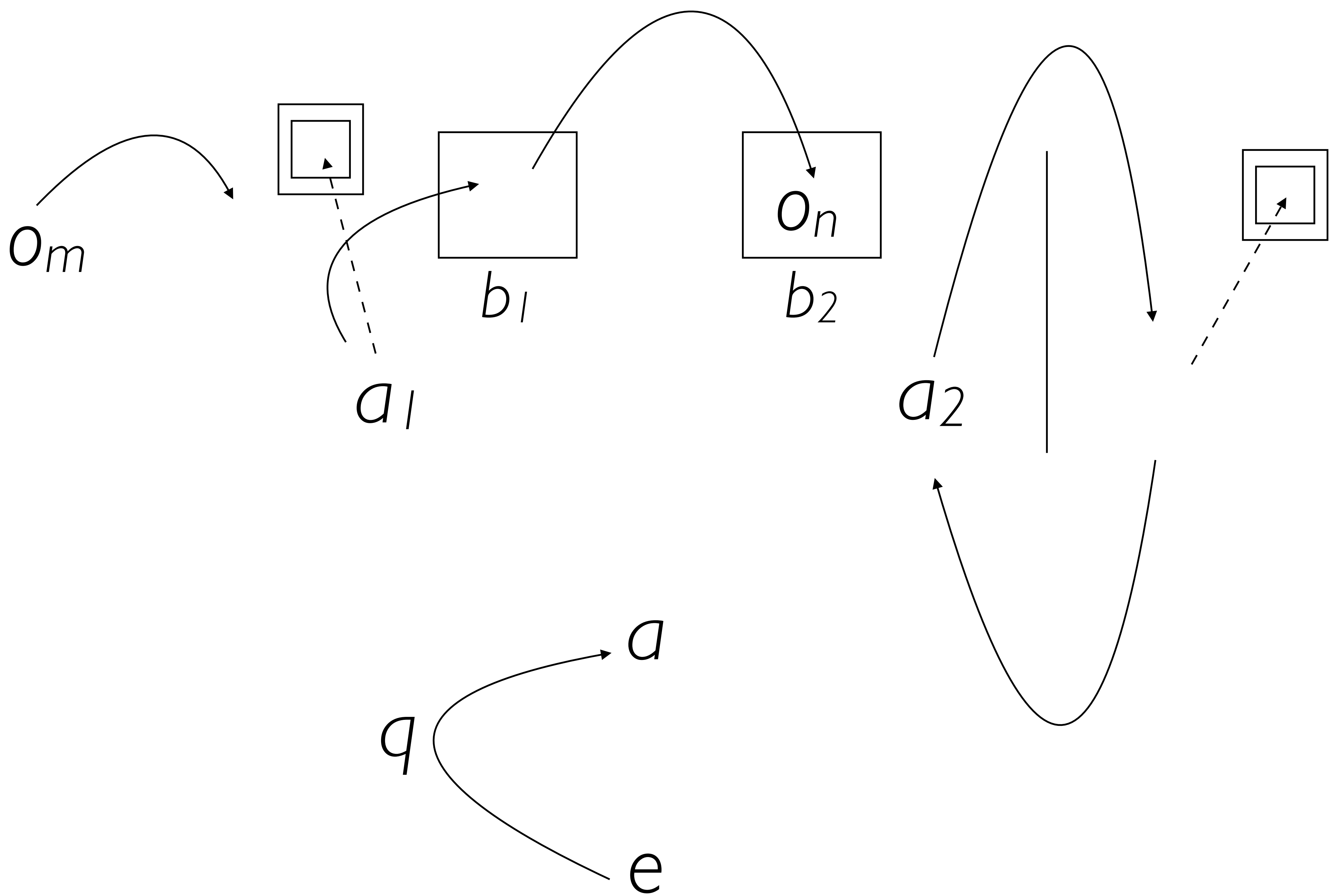
Framework for FBT^1_3
(eight timepoints)



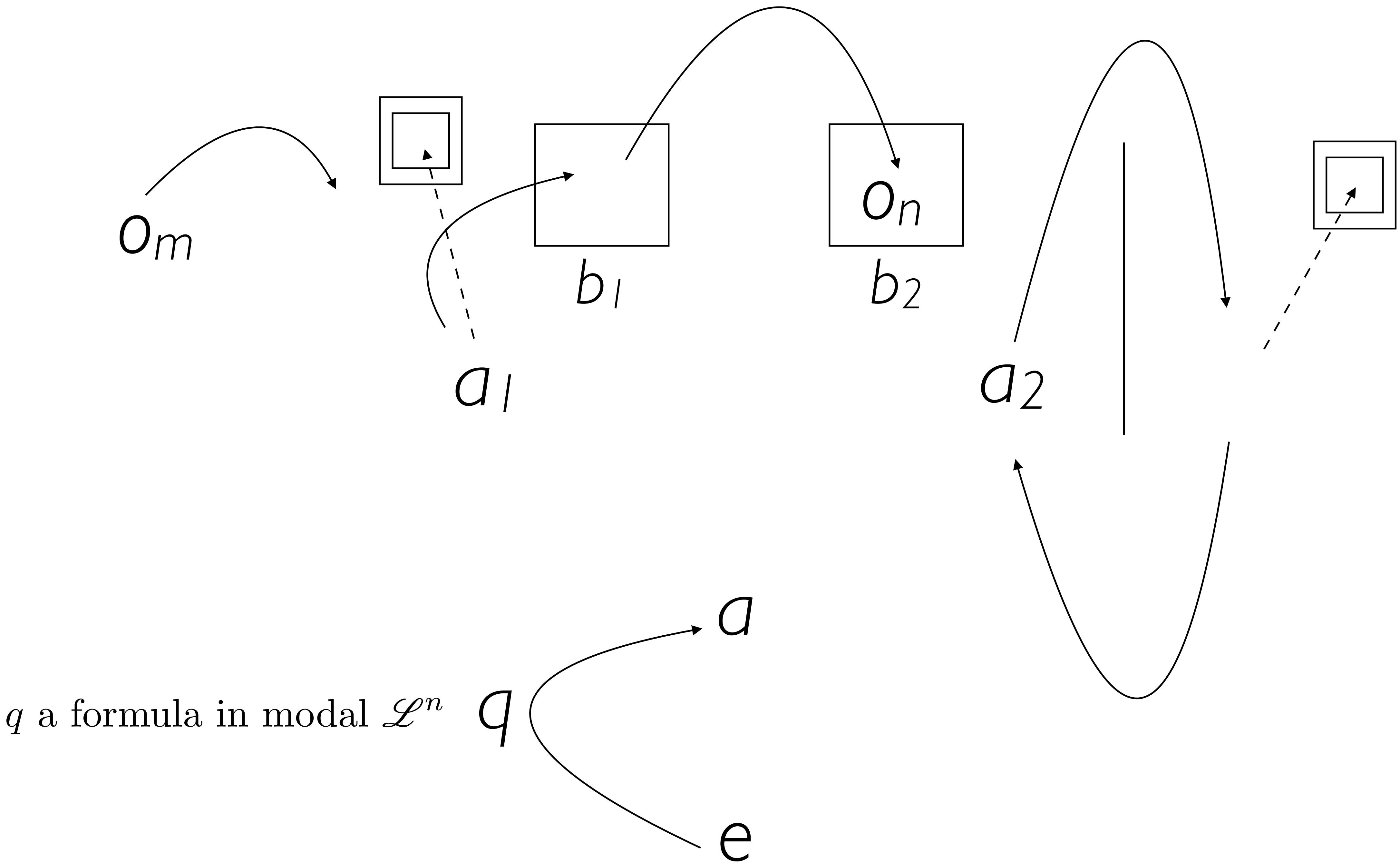
Framework for FBT^1_3
(eight timepoints)



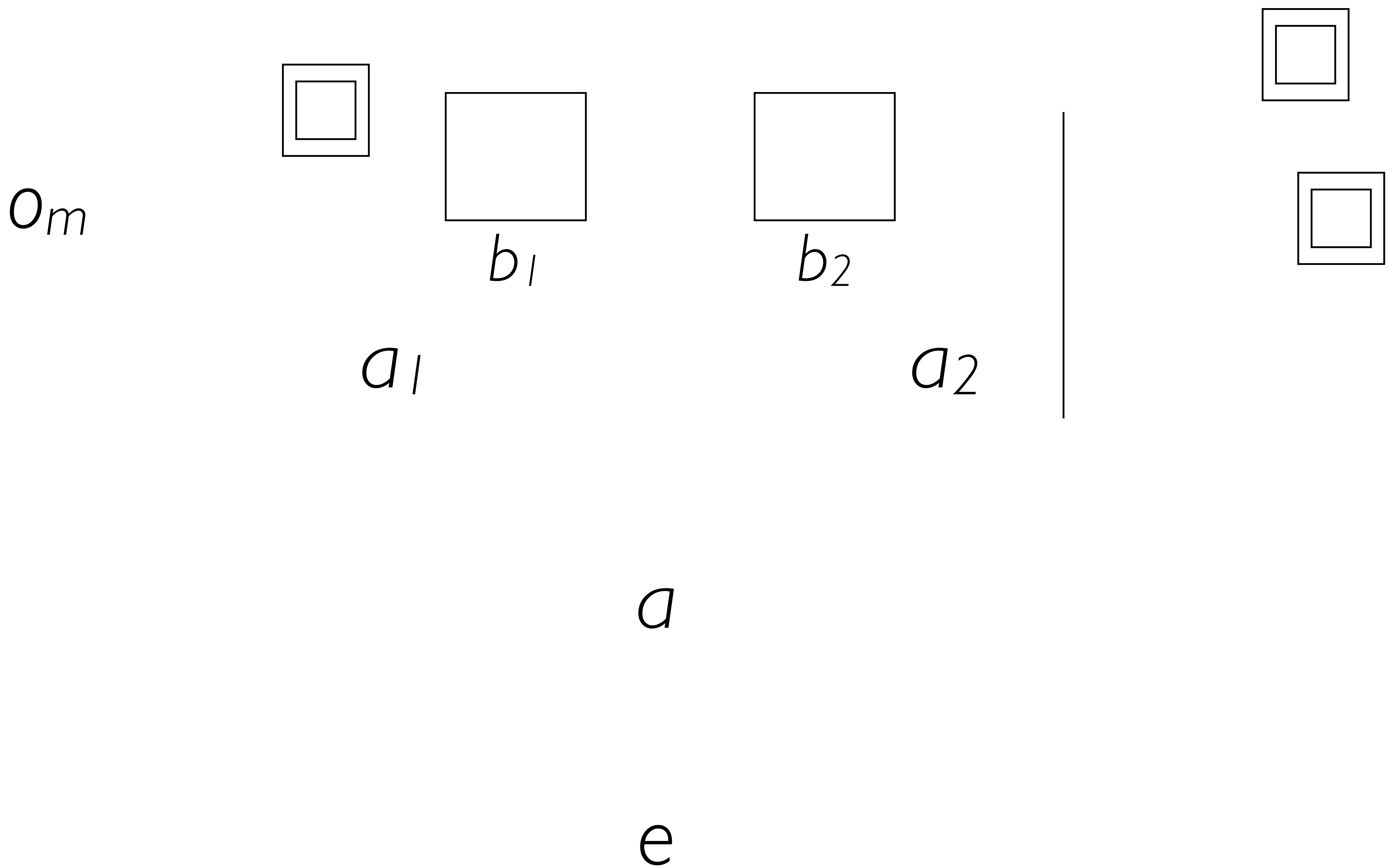
Framework for $FBT^1{}_3$
(eight timepoints)



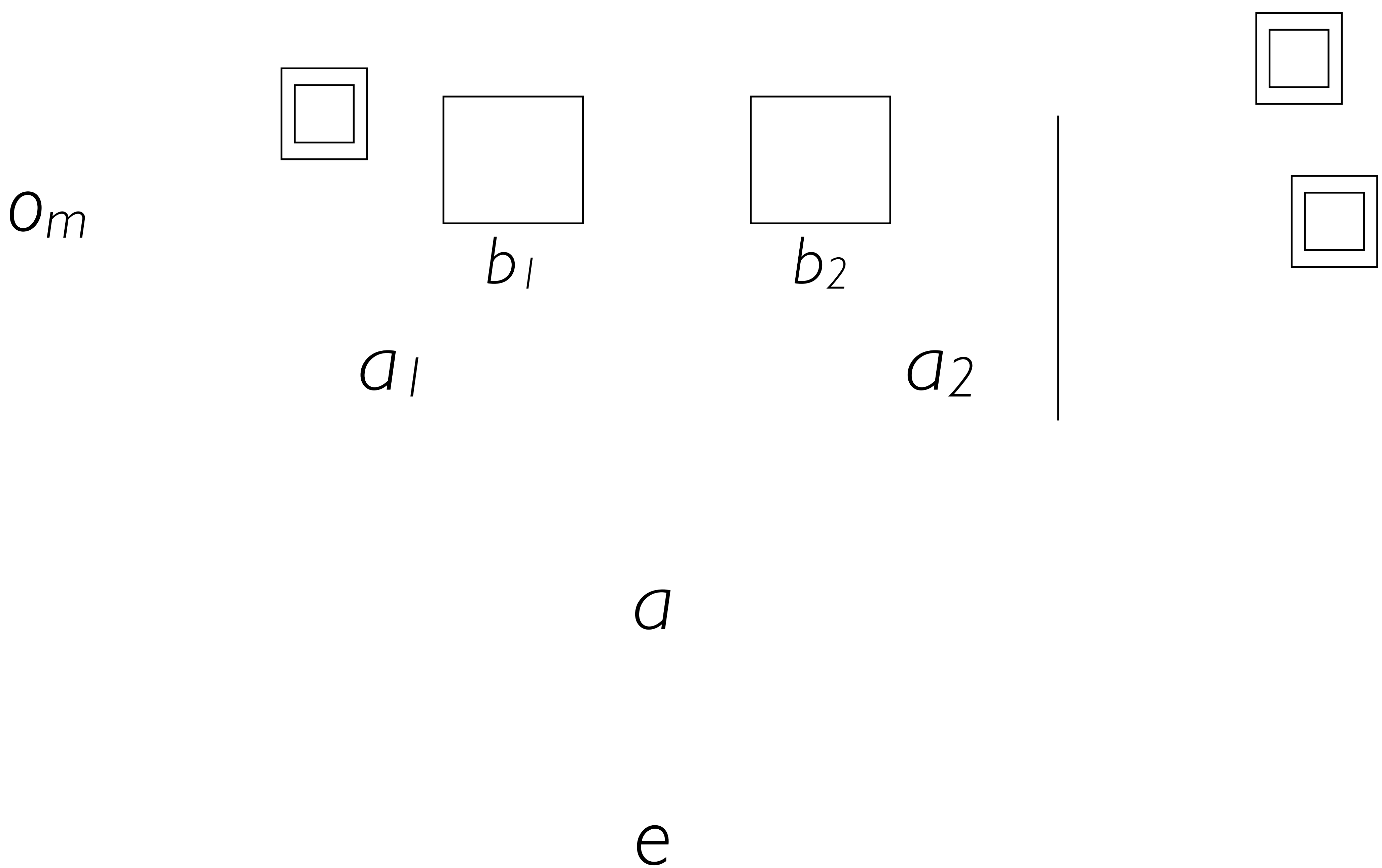
Framework for $FBT^1{}_3$
(eight timepoints)



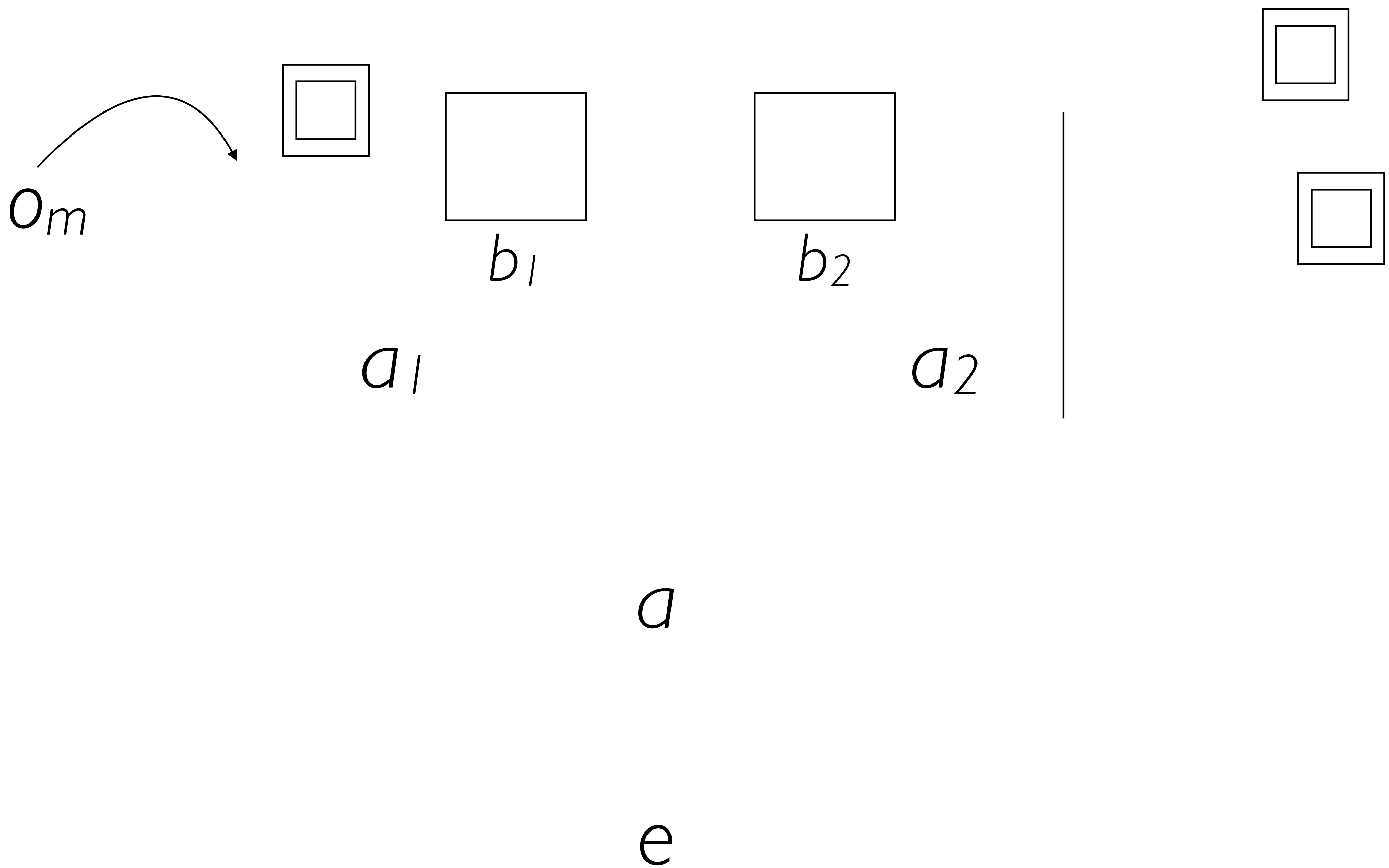
Framework for FBT^1_4



Framework for FBT^1_4
(nine timepoints)

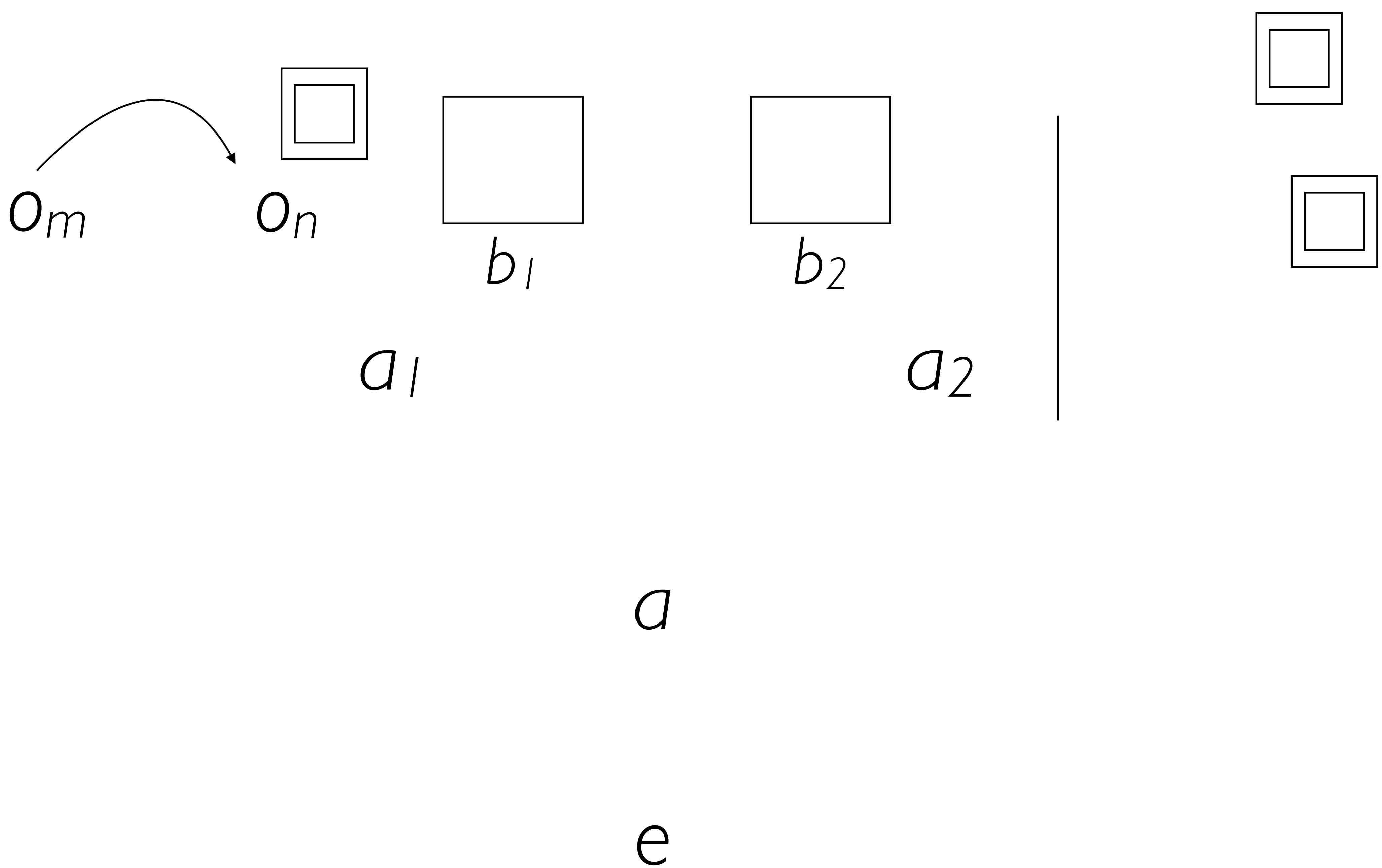


Framework for FBT$^1_4$
(nine timepoints)



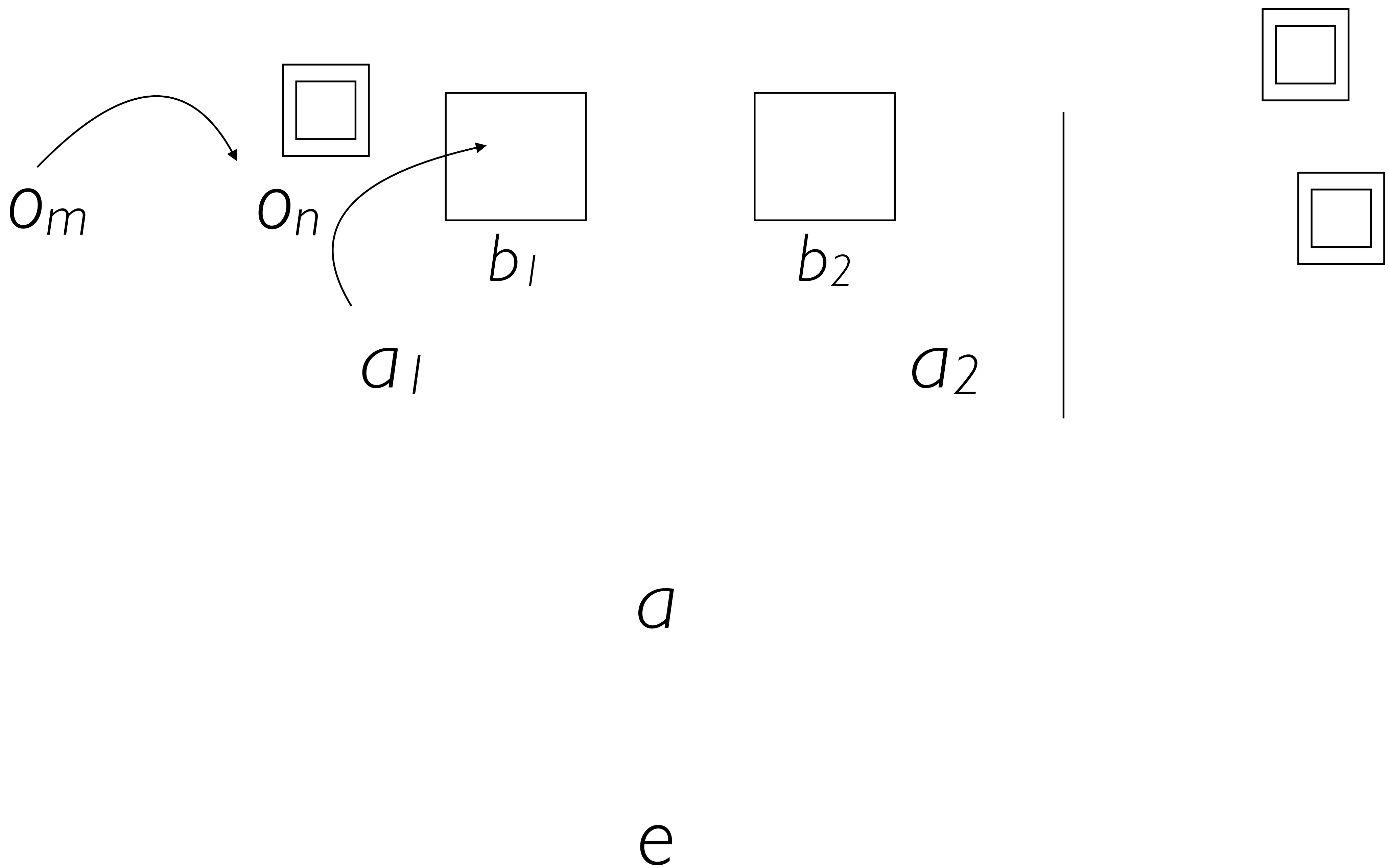
Framework for FBT^{1}_{4}

(nine timepoints)

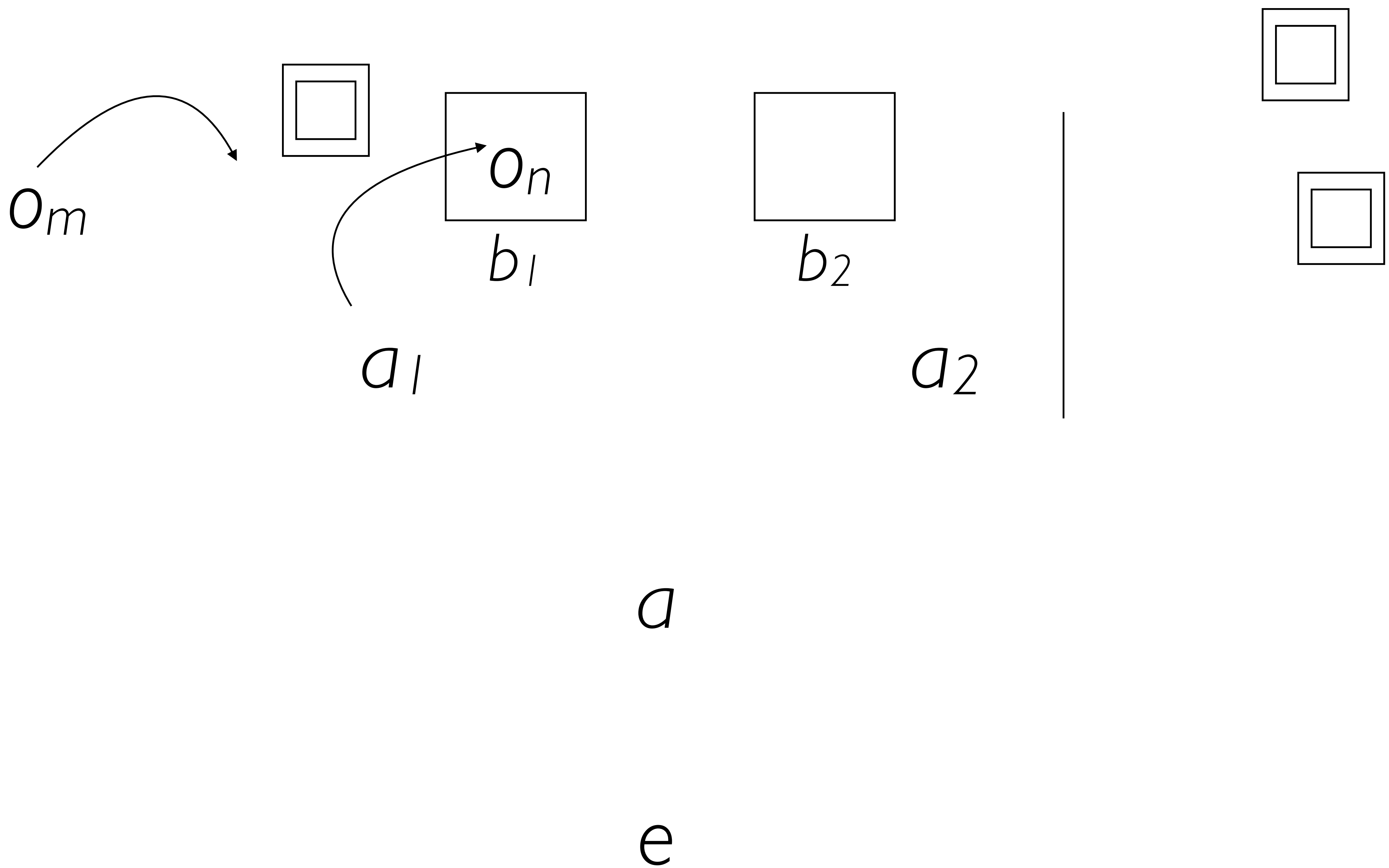


Framework for FBT^1_4

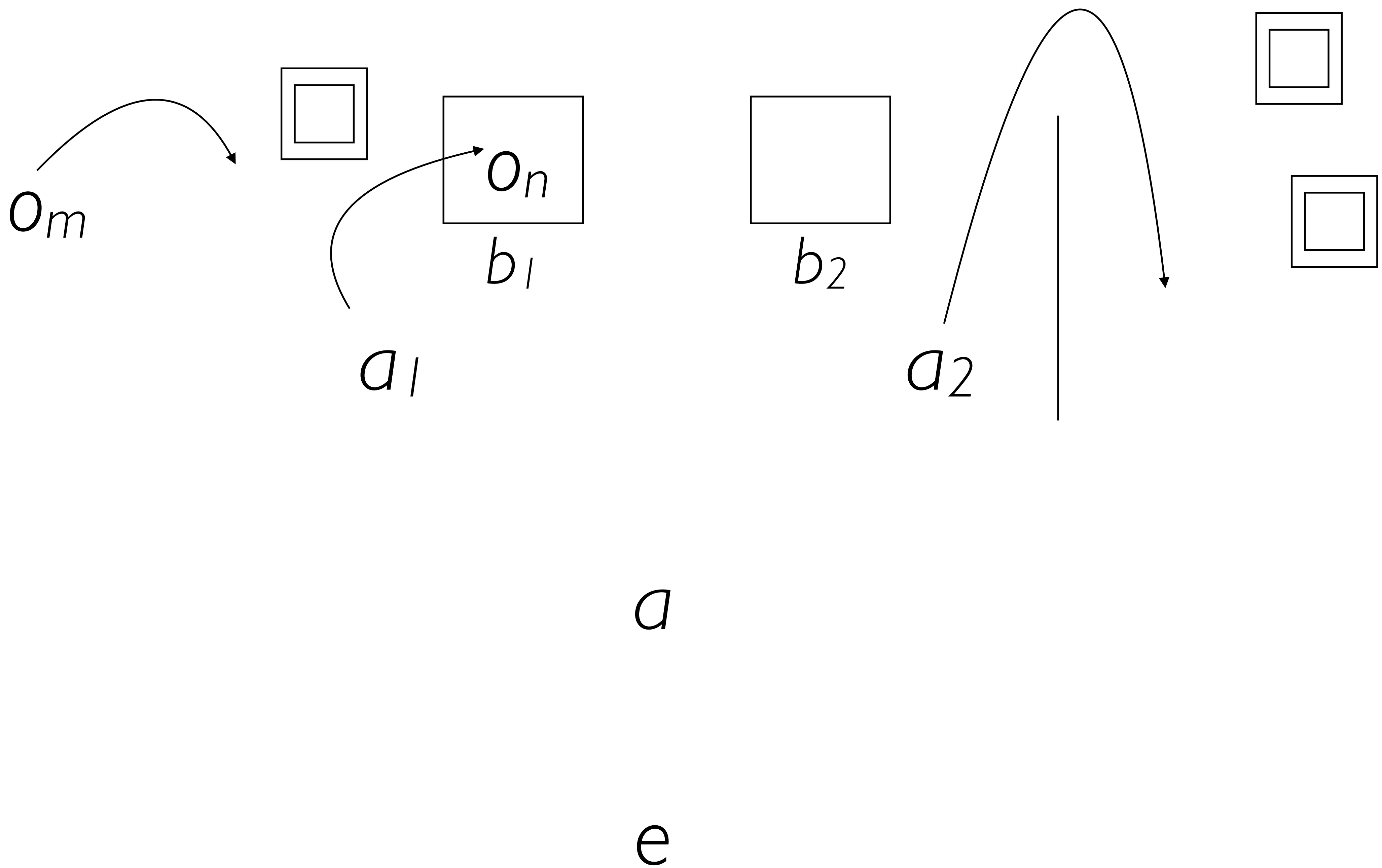
(nine timepoints)



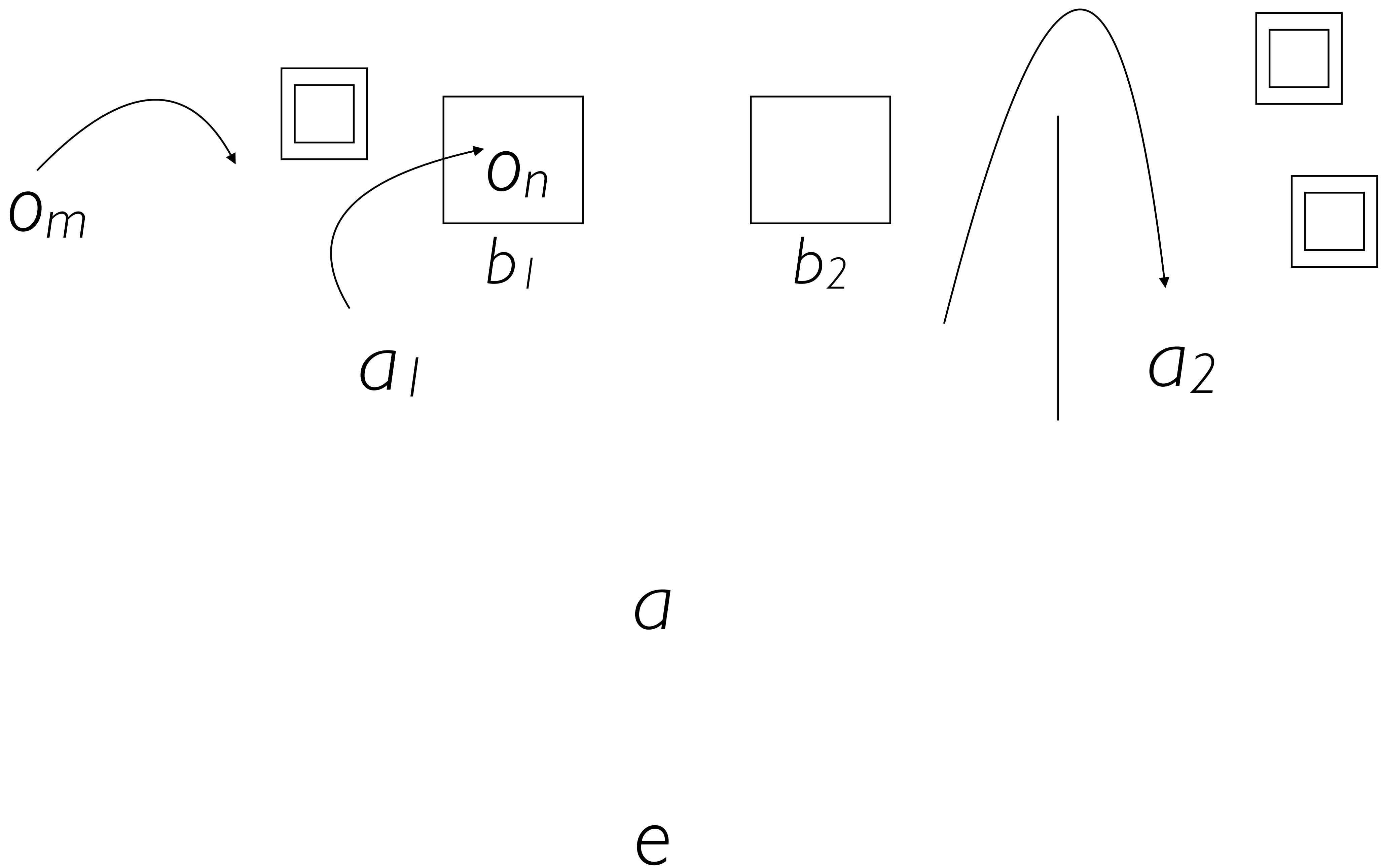
Framework for FBT^{1}_{4}
(nine timepoints)



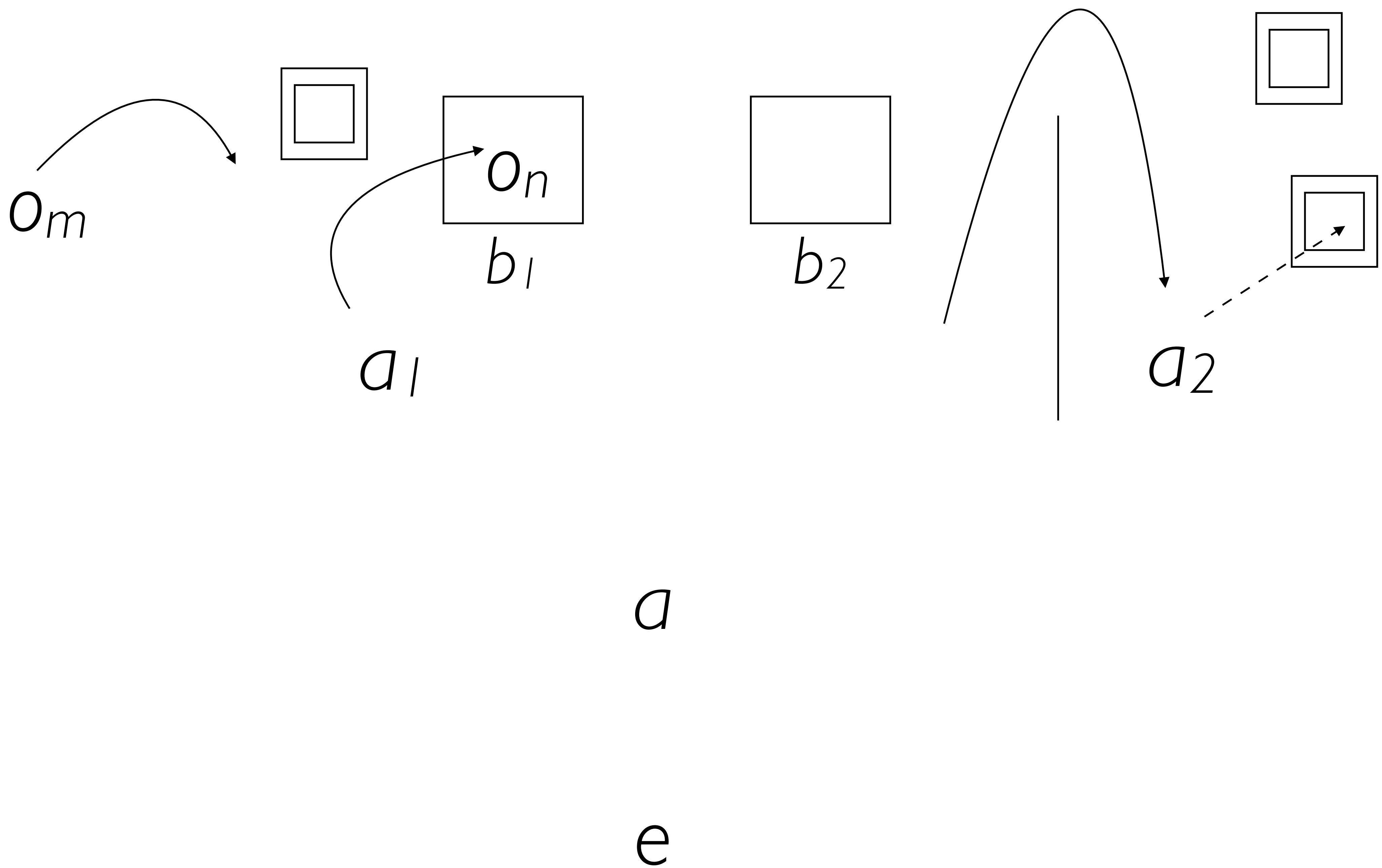
Framework for FBT^1_4
(nine timepoints)



Framework for $FBT^1{}_4$
(nine timepoints)

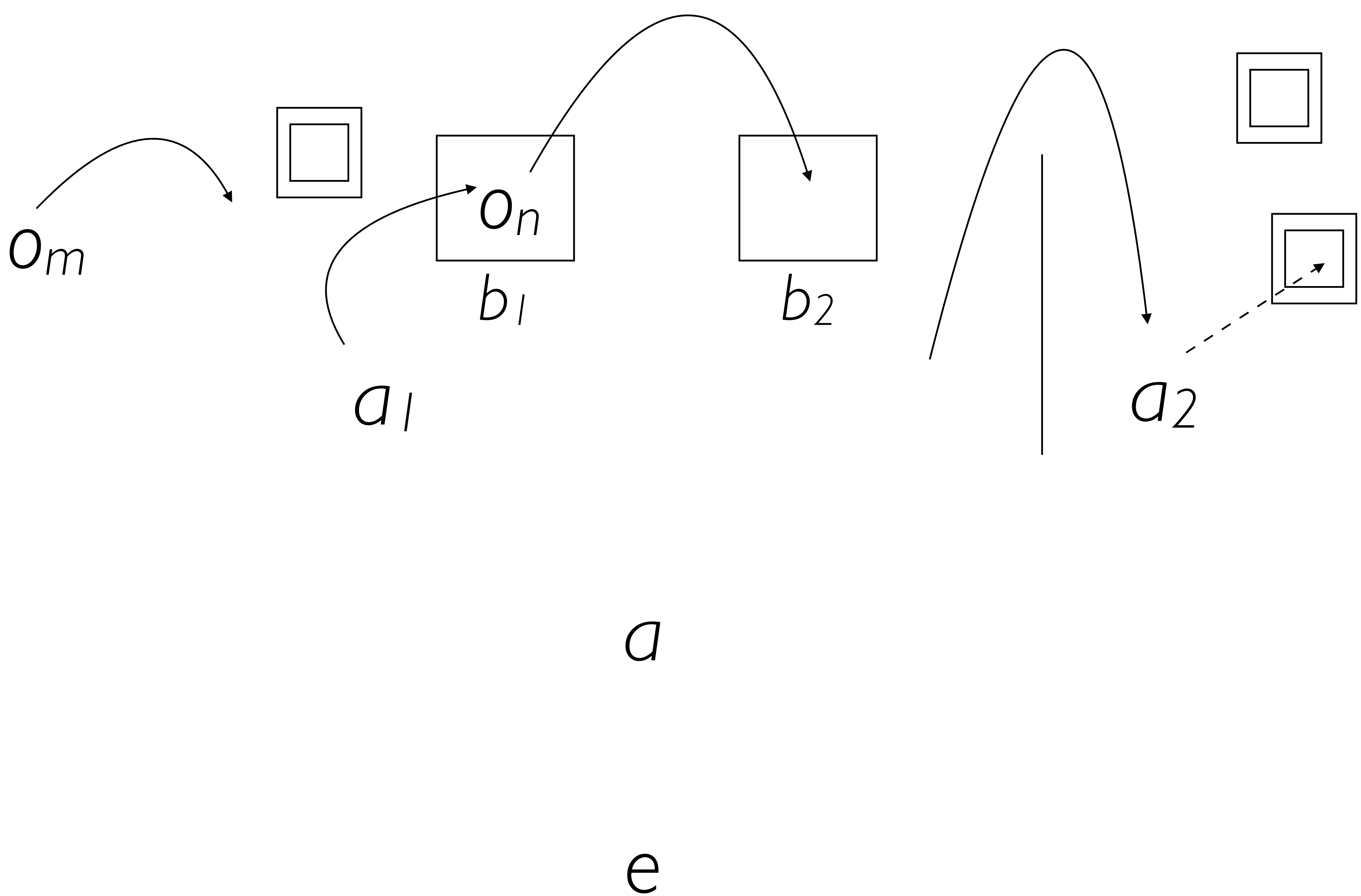


Framework for FBT$^{1}_{4}$
(nine timepoints)

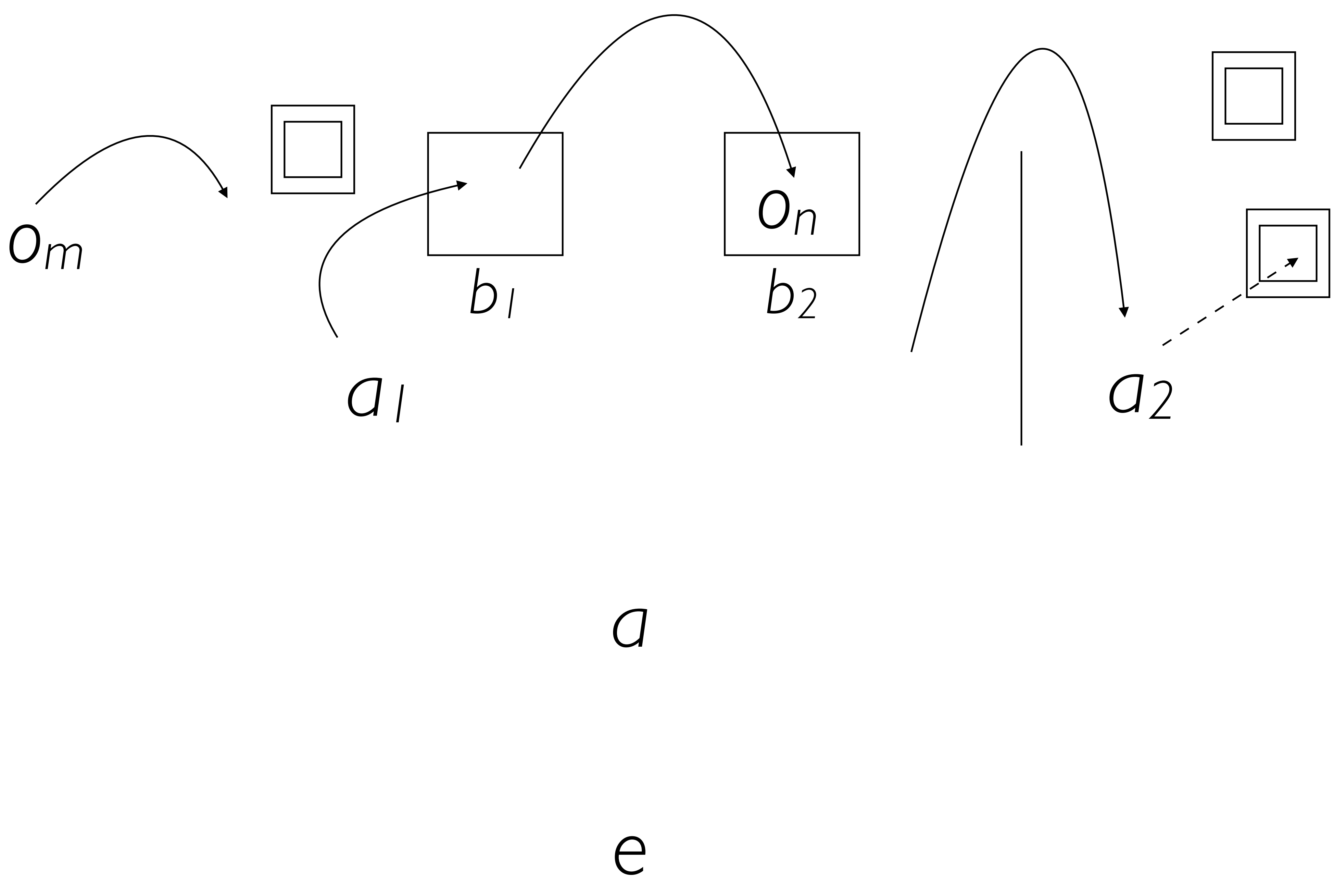


Framework for FBT^{1}_{4}

(nine timepoints)

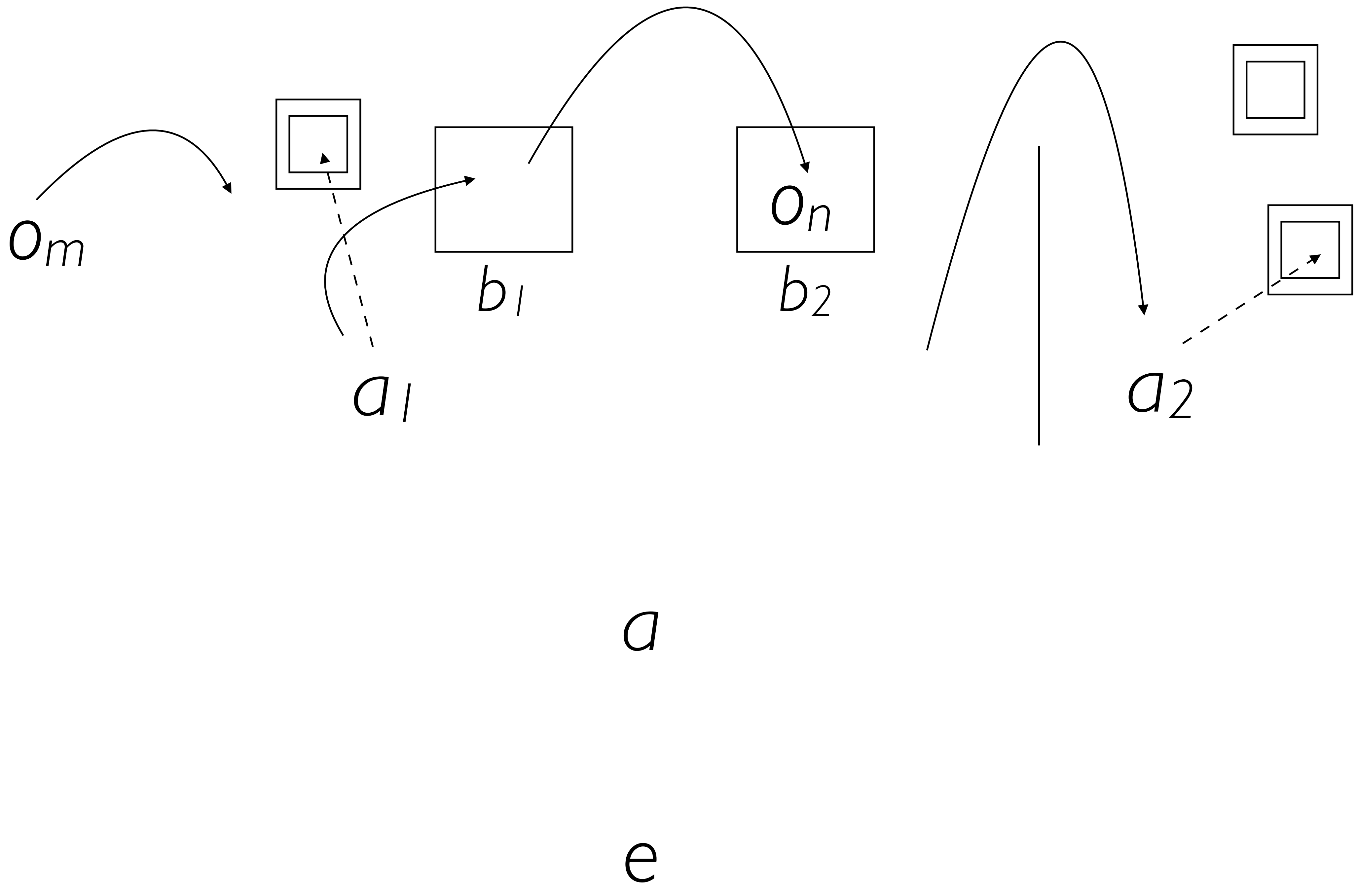


Framework for FBT^1_4
(nine timepoints)

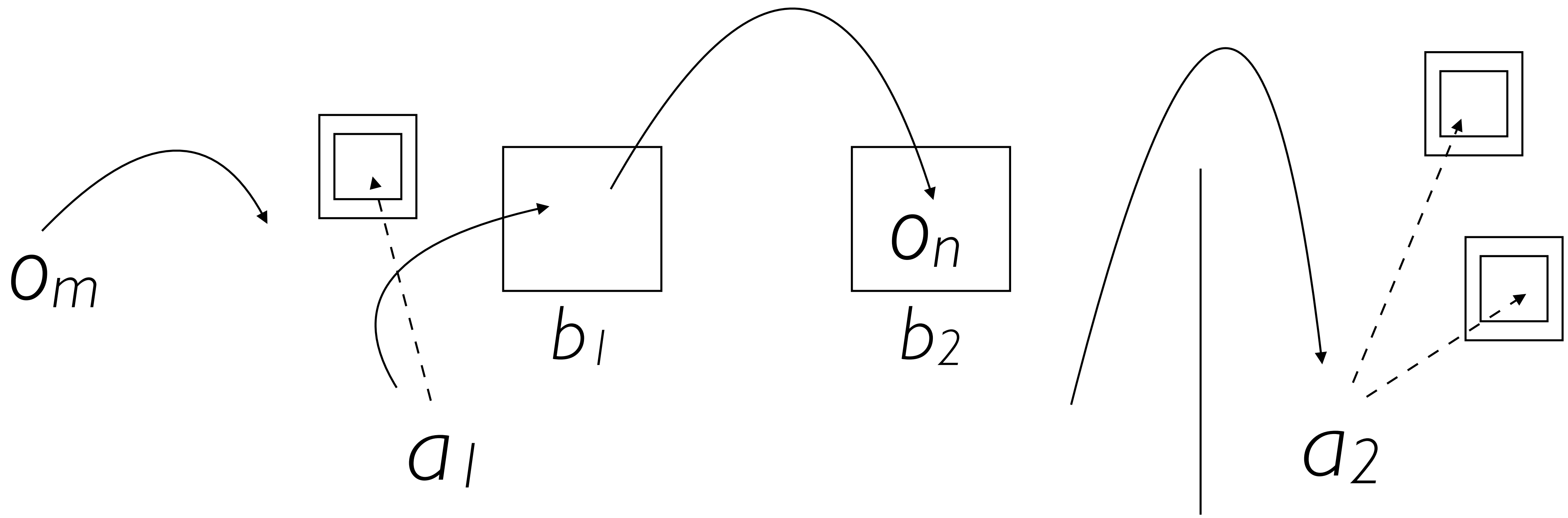


Framework for FBT$^1_4$

(nine timepoints)



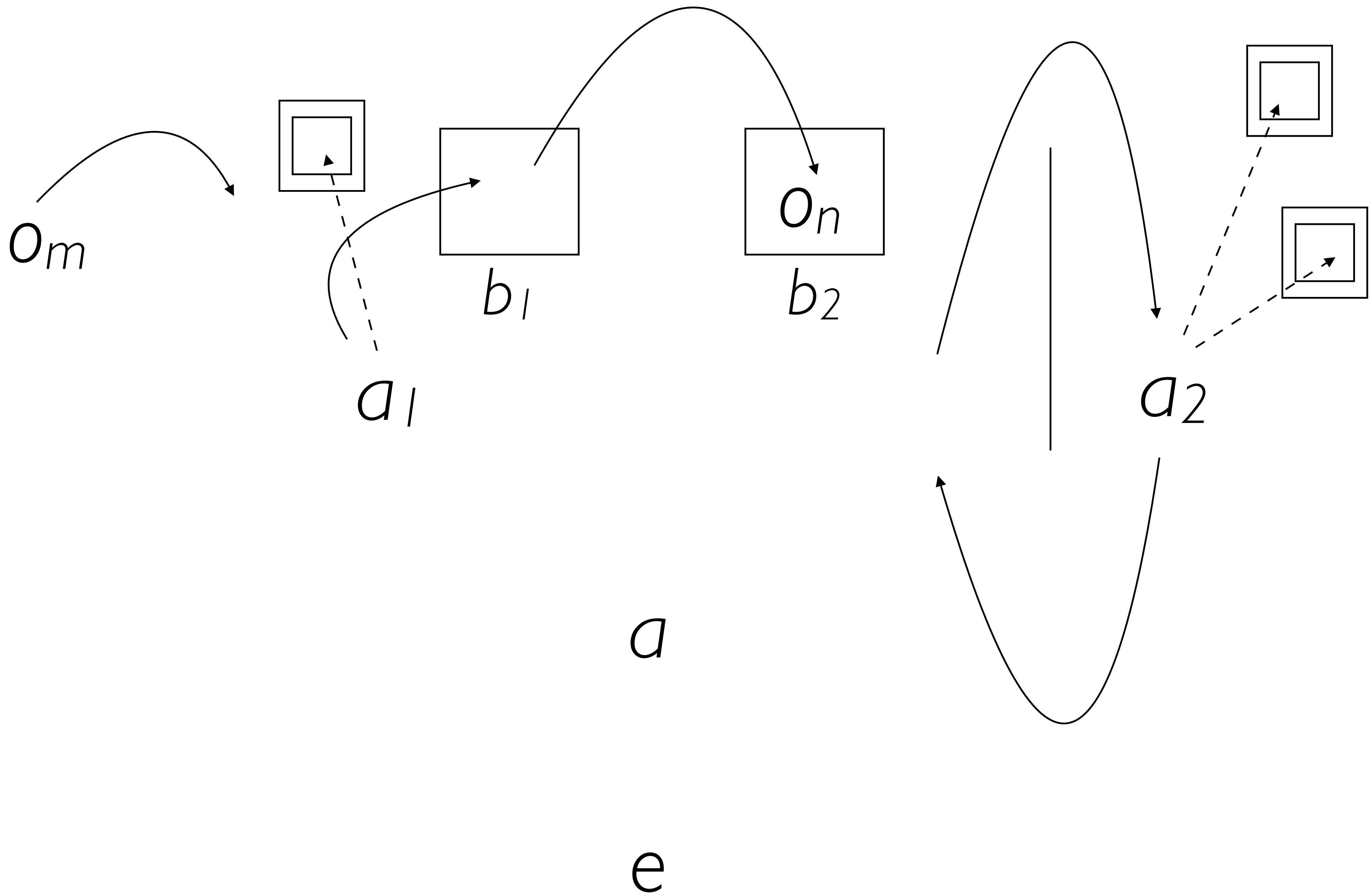
Framework for $FBT^1{}_4$
(nine timepoints)



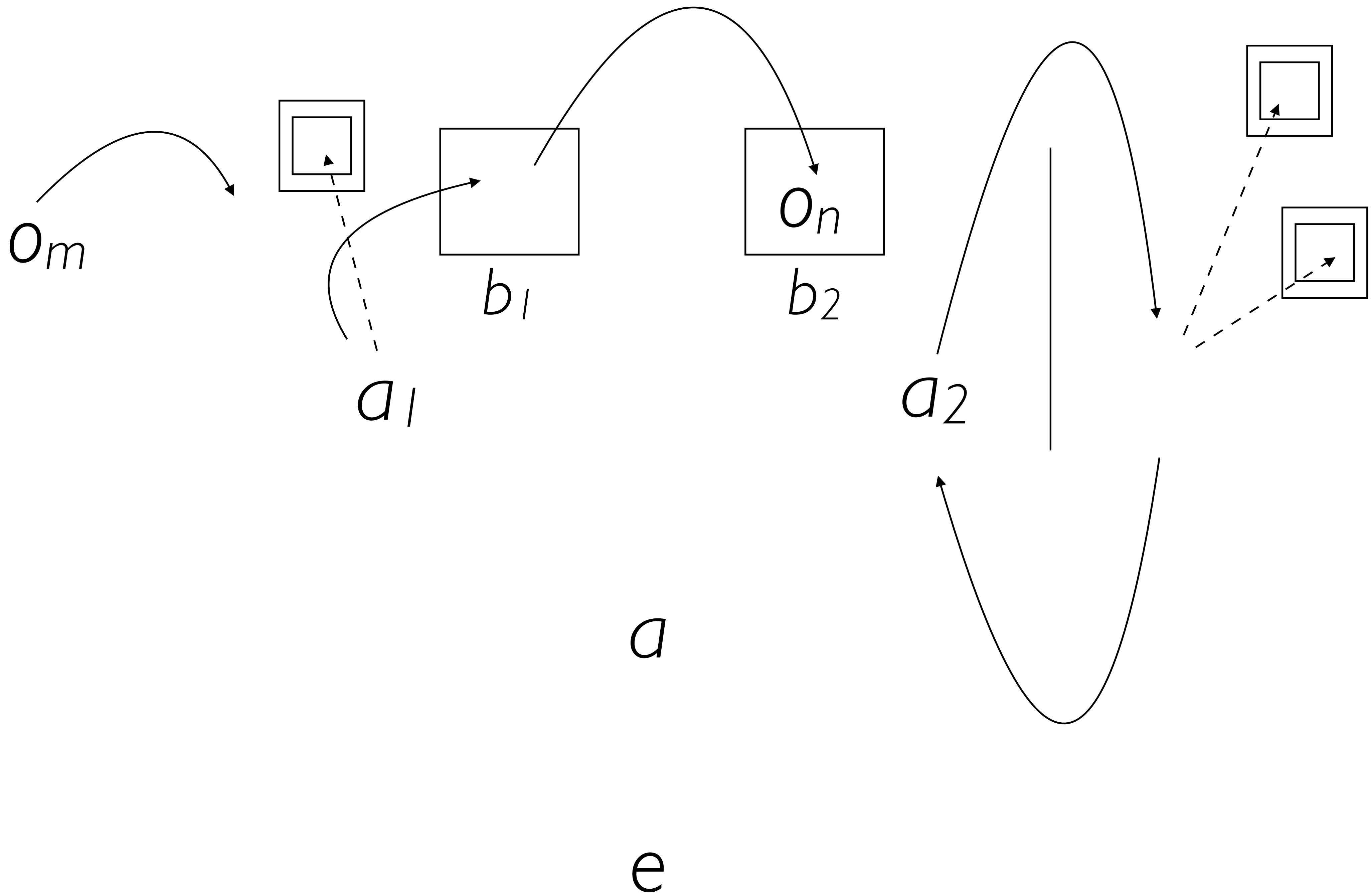
a

e

Framework for $FBT^1{}_4$
(nine timepoints)

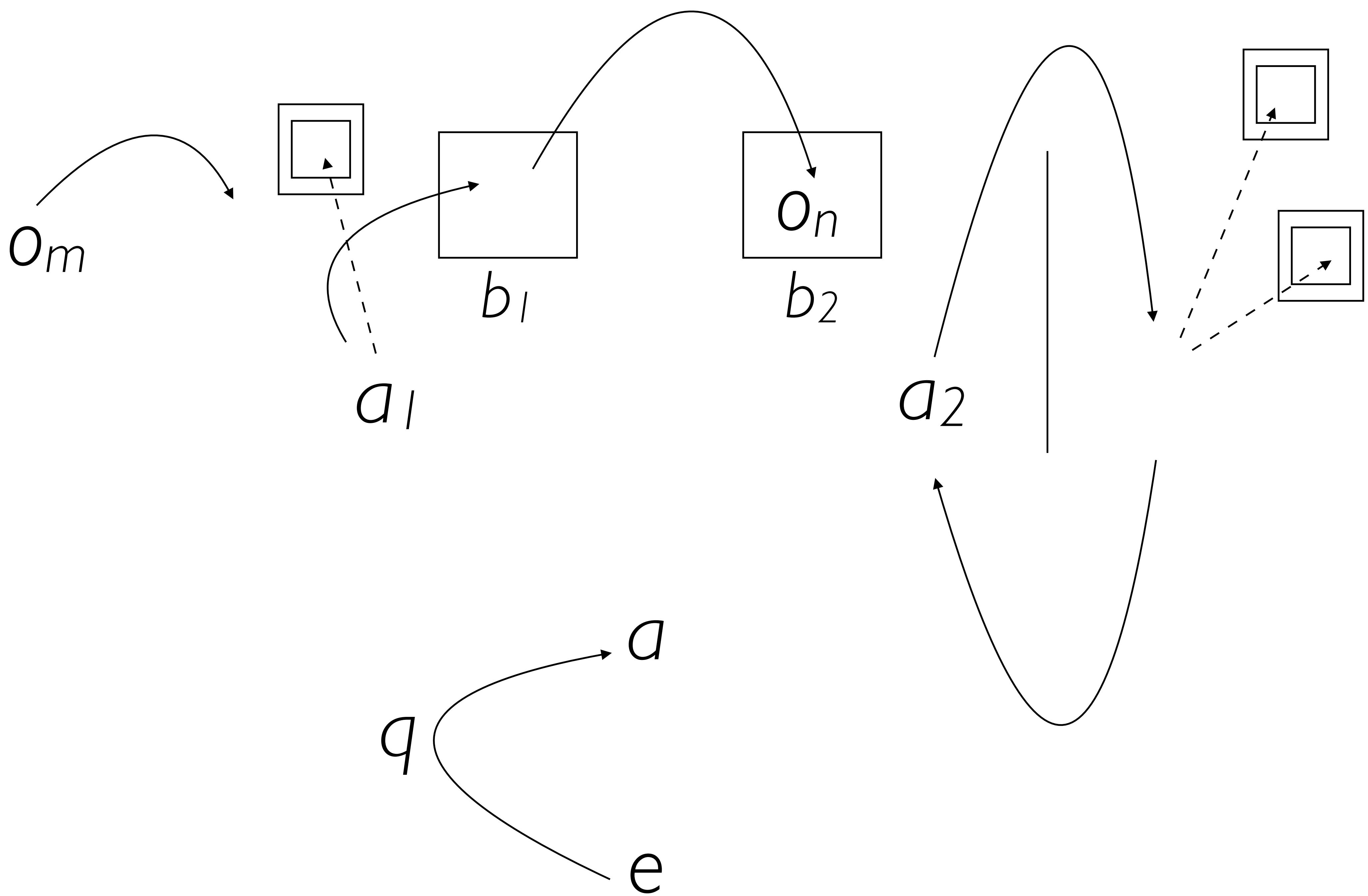


Framework for $FBT^1{}_4$
(nine timepoints)



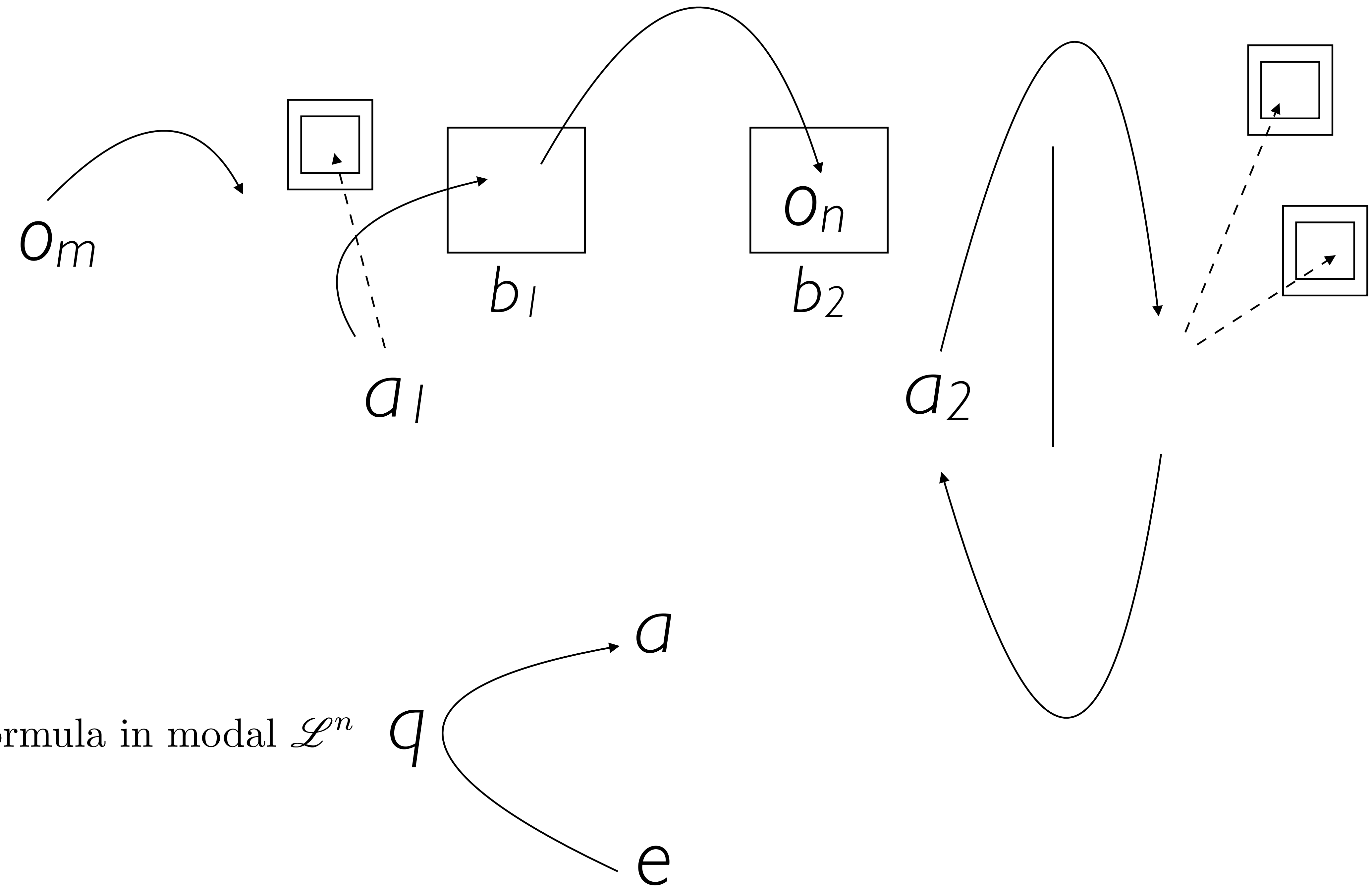
Framework for FBT^1_4

(nine timepoints)



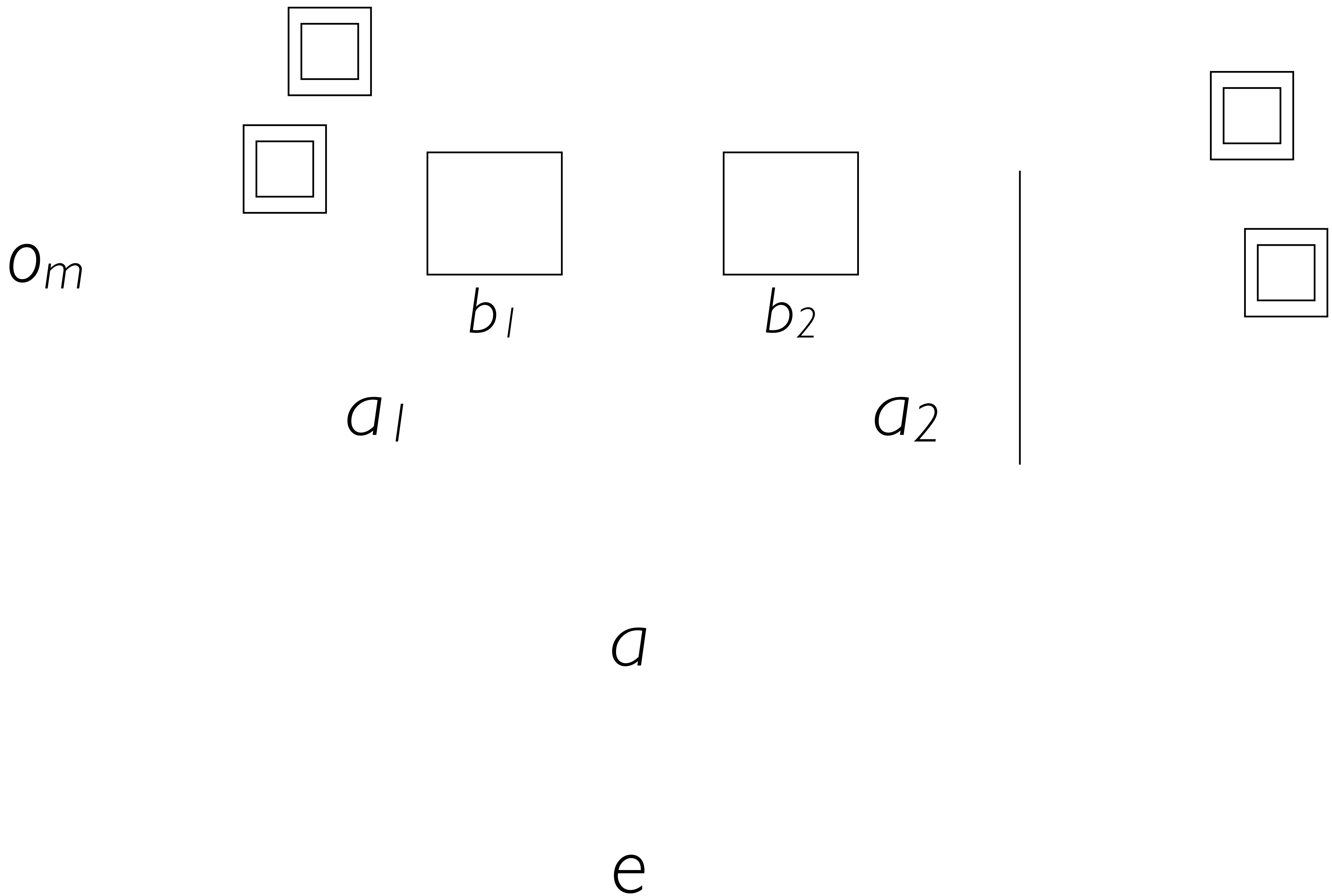
Framework for FBT$^{1}_{4}$

(nine timepoints)

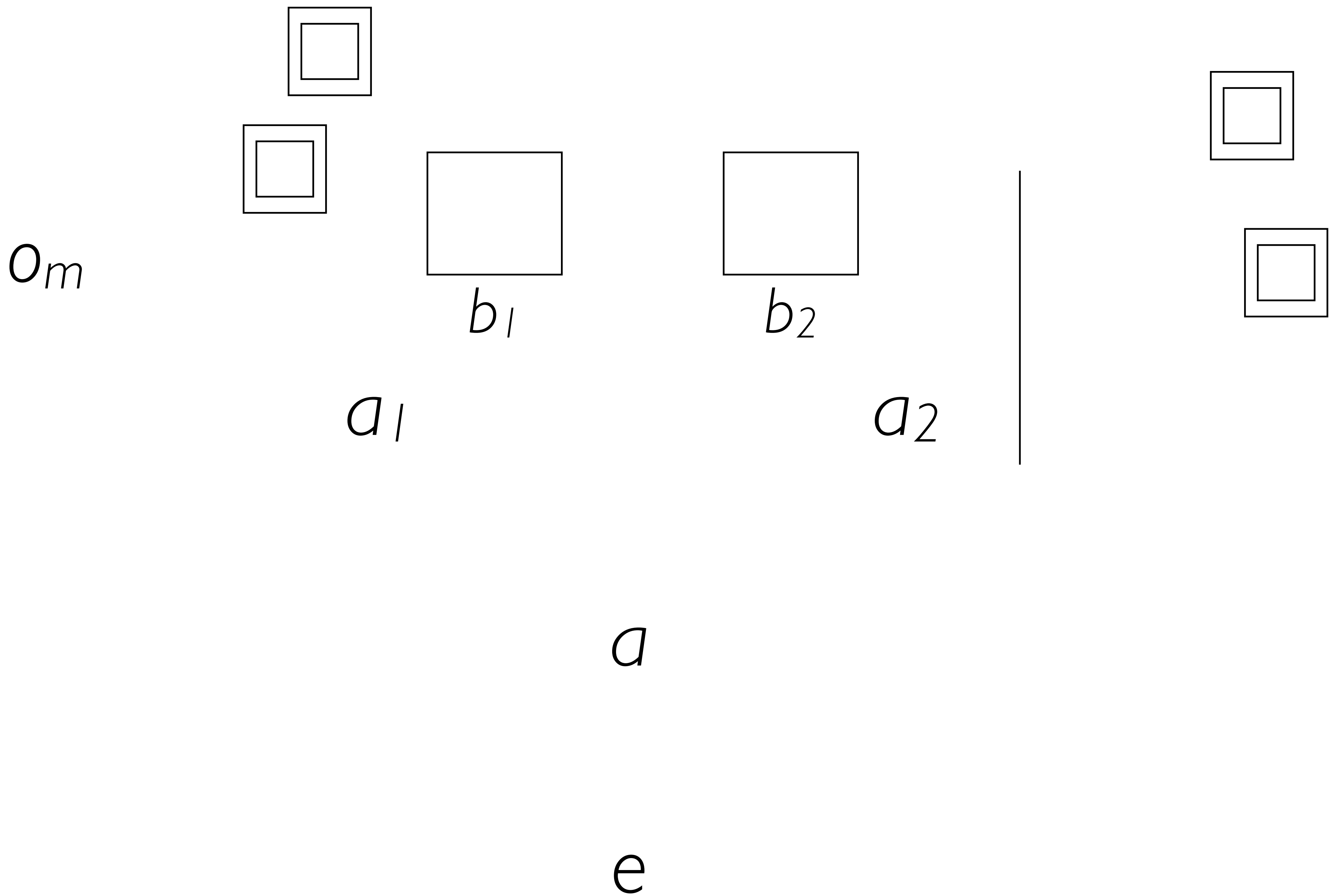


q a formula in modal $\mathscr{L}^n$

Framework for FBT^{1}_{5}

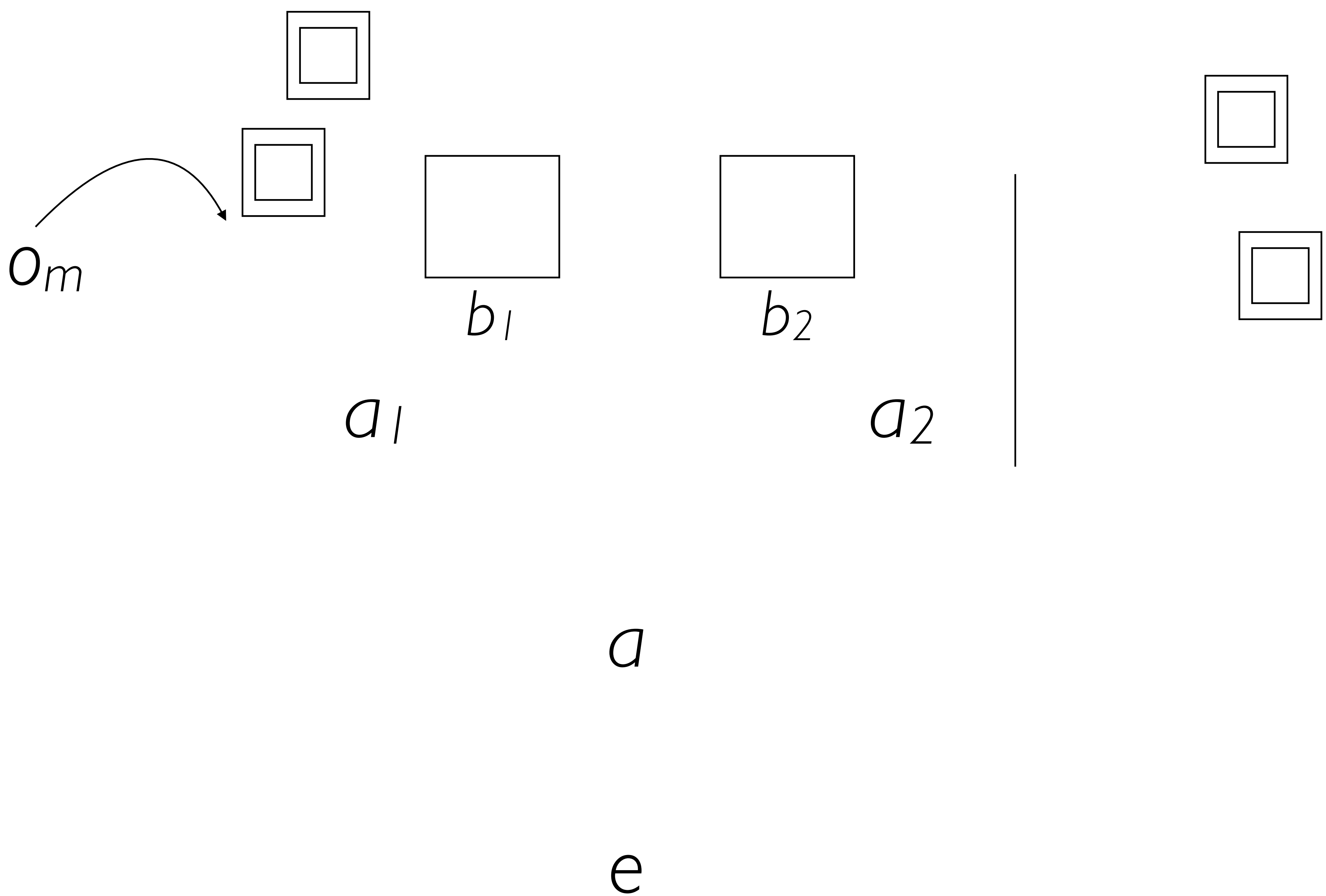


Framework for FBT$^{1}_{5}$
(ten timepoints)



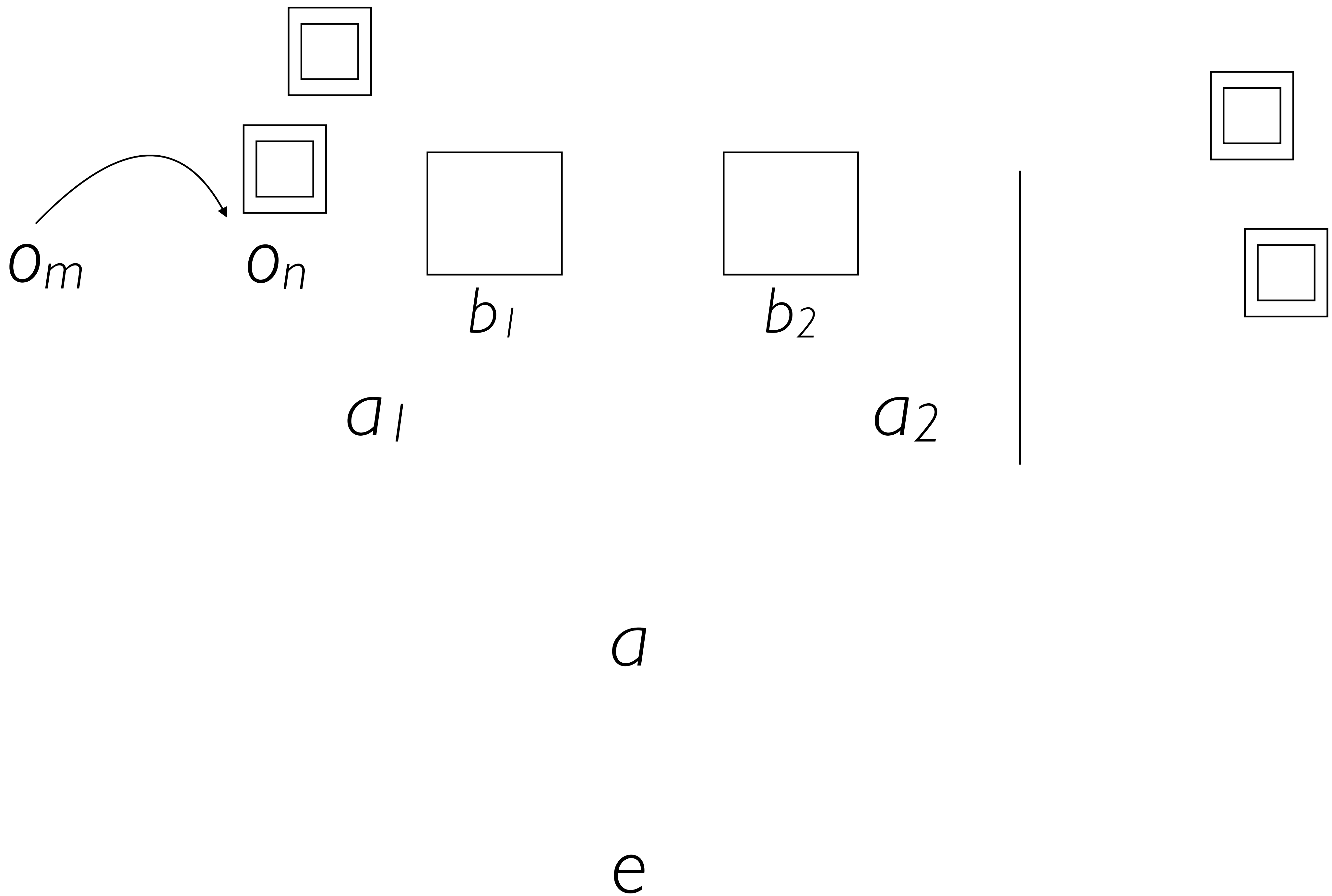
Framework for FBT^1_5

(ten timepoints)



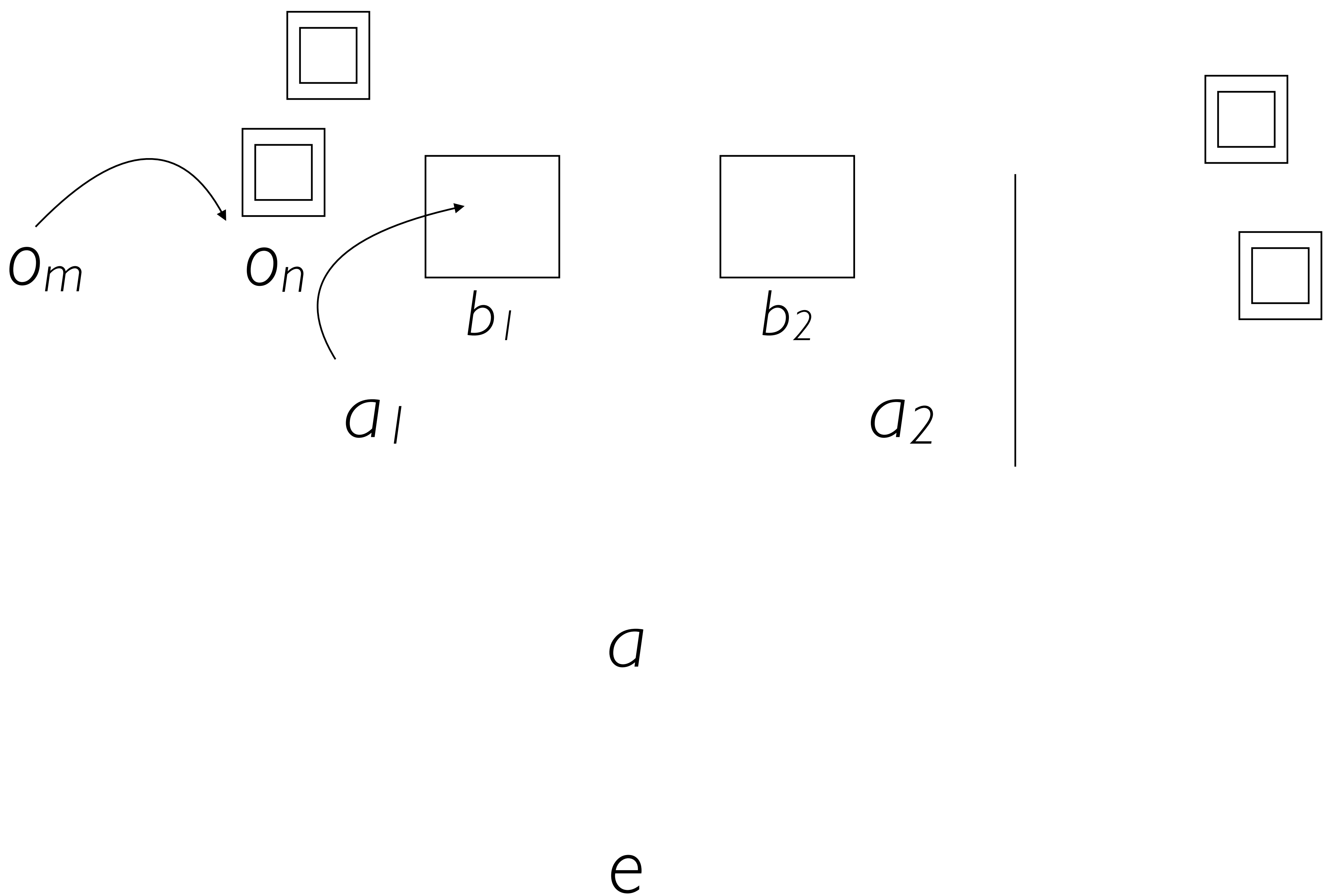
Framework for FBT^1_5

(ten timepoints)

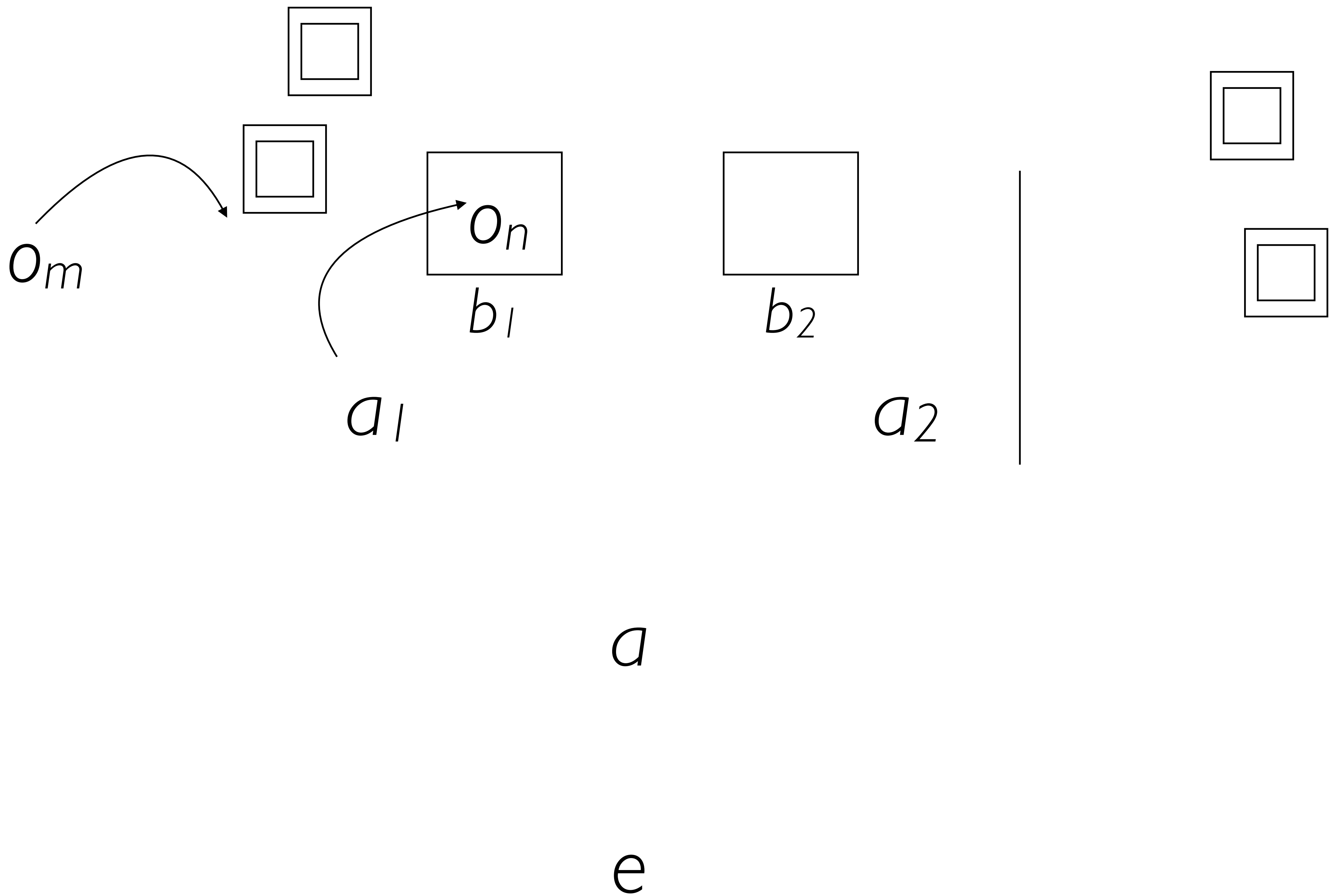


Framework for FBT^{1}_{5}

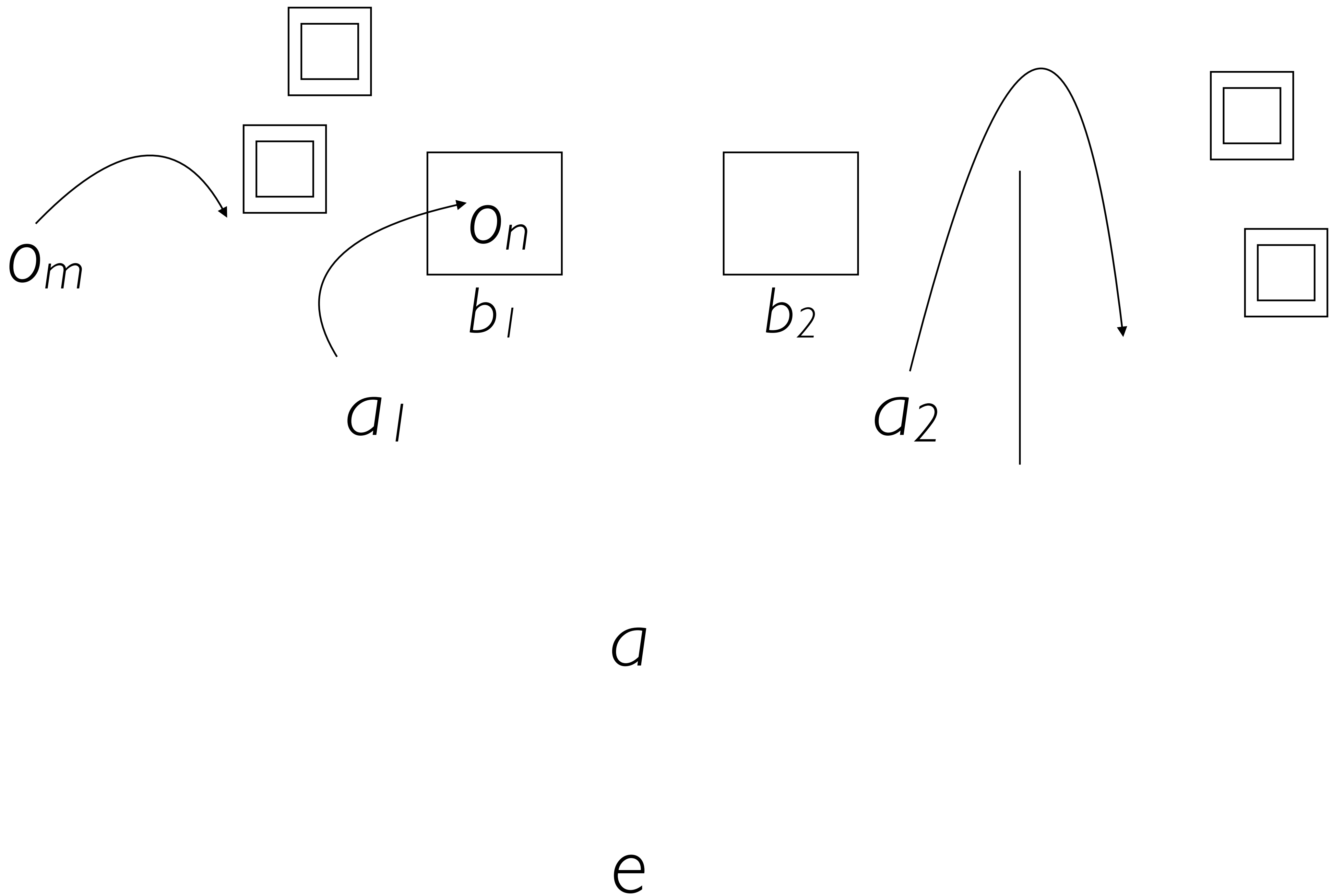
(ten timepoints)



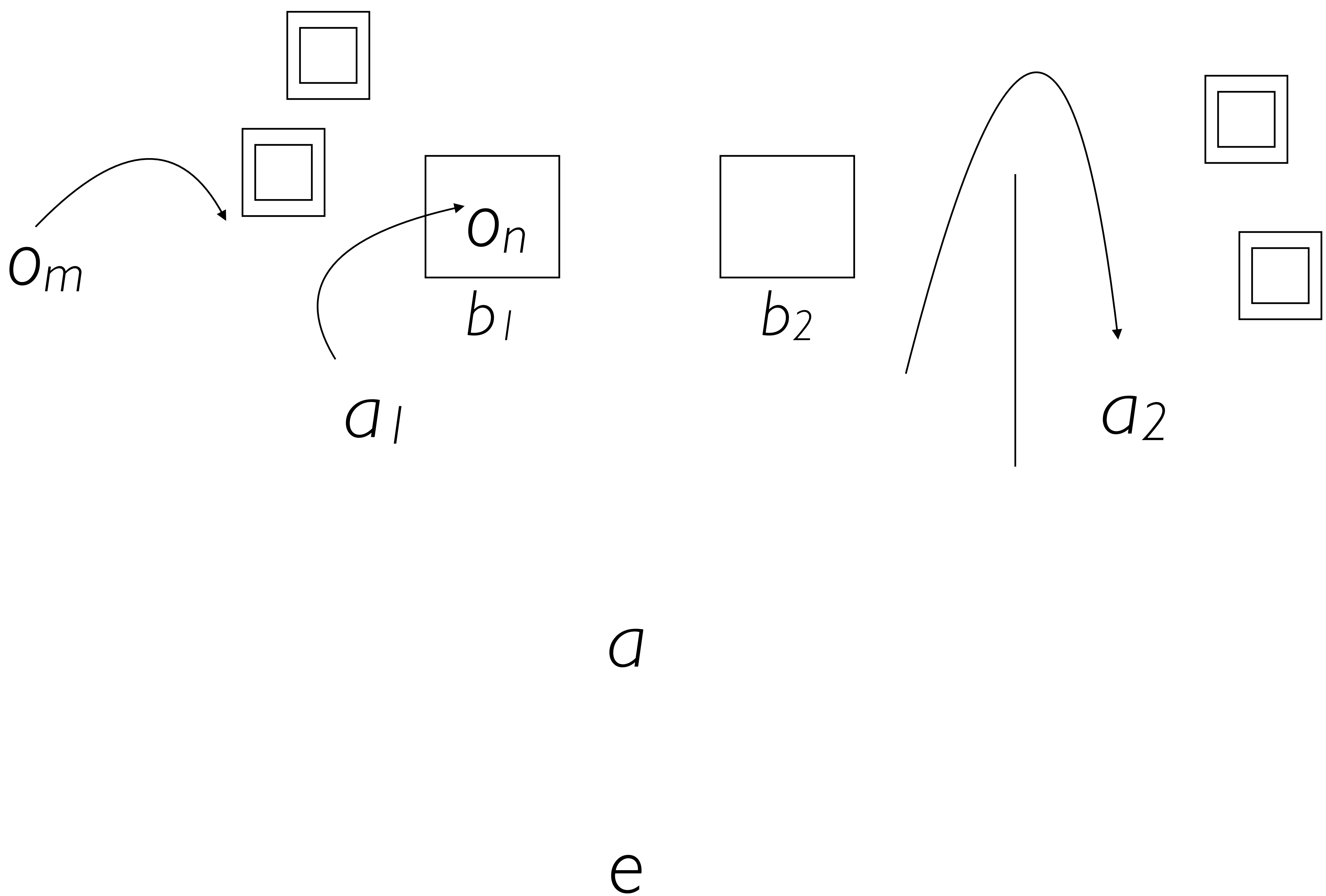
Framework for FBT^1_5
(ten timepoints)



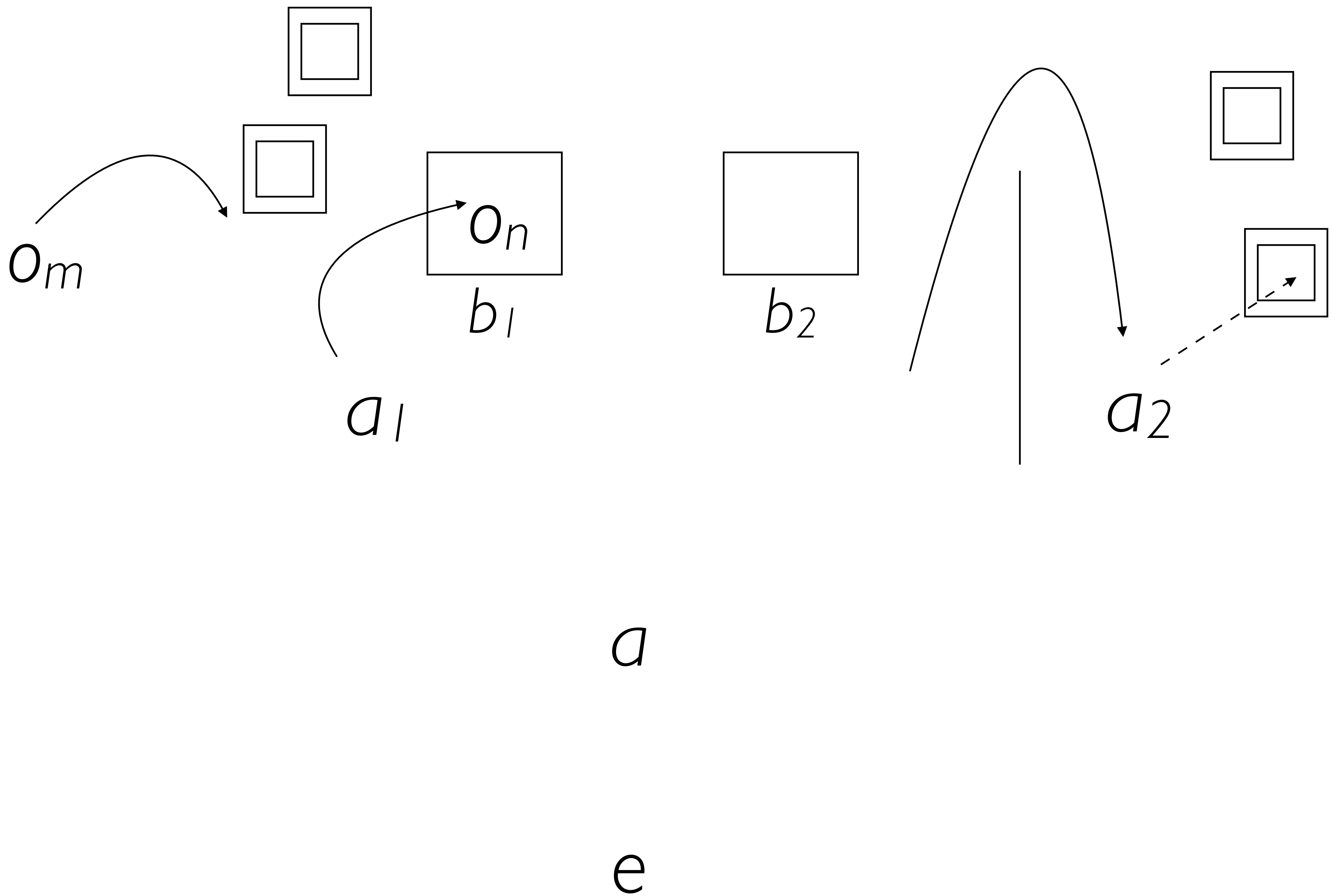
Framework for FBT$^1_5$
(ten timepoints)



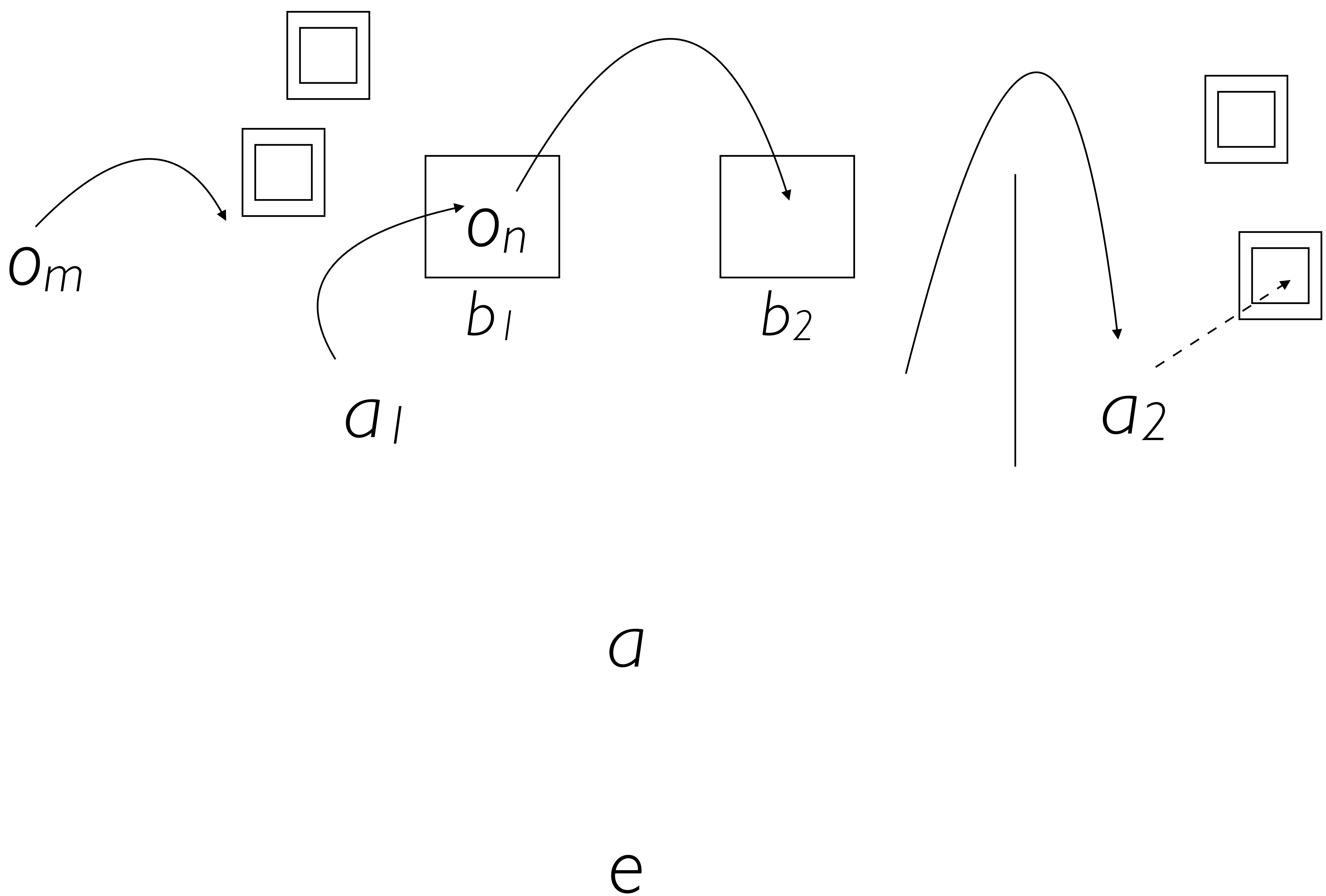
Framework for $FBT^1{}_5$
(ten timepoints)



Framework for FBT$^1{}_5$
(ten timepoints)

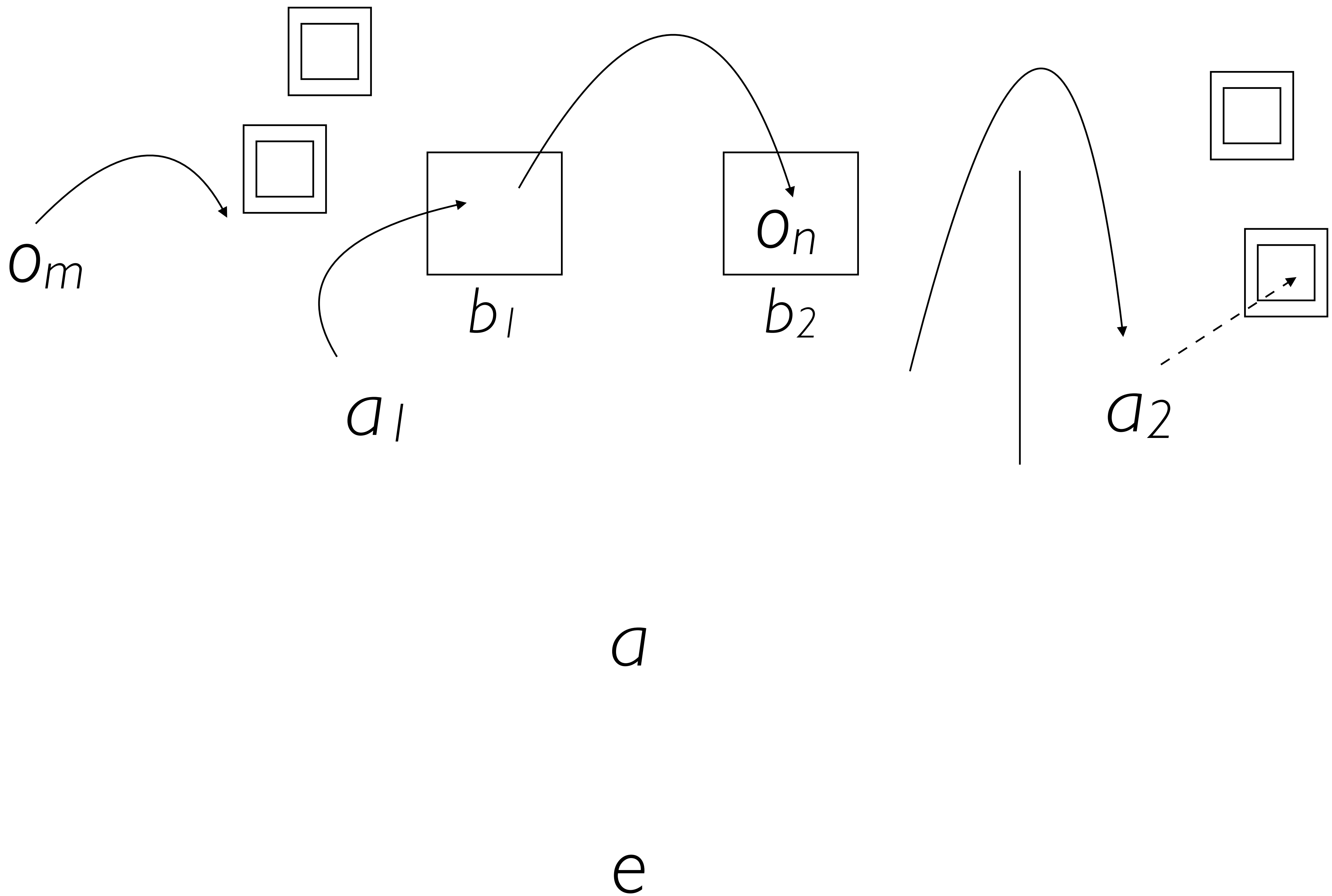


Framework for FBT^{1}_{5}
(ten timepoints)



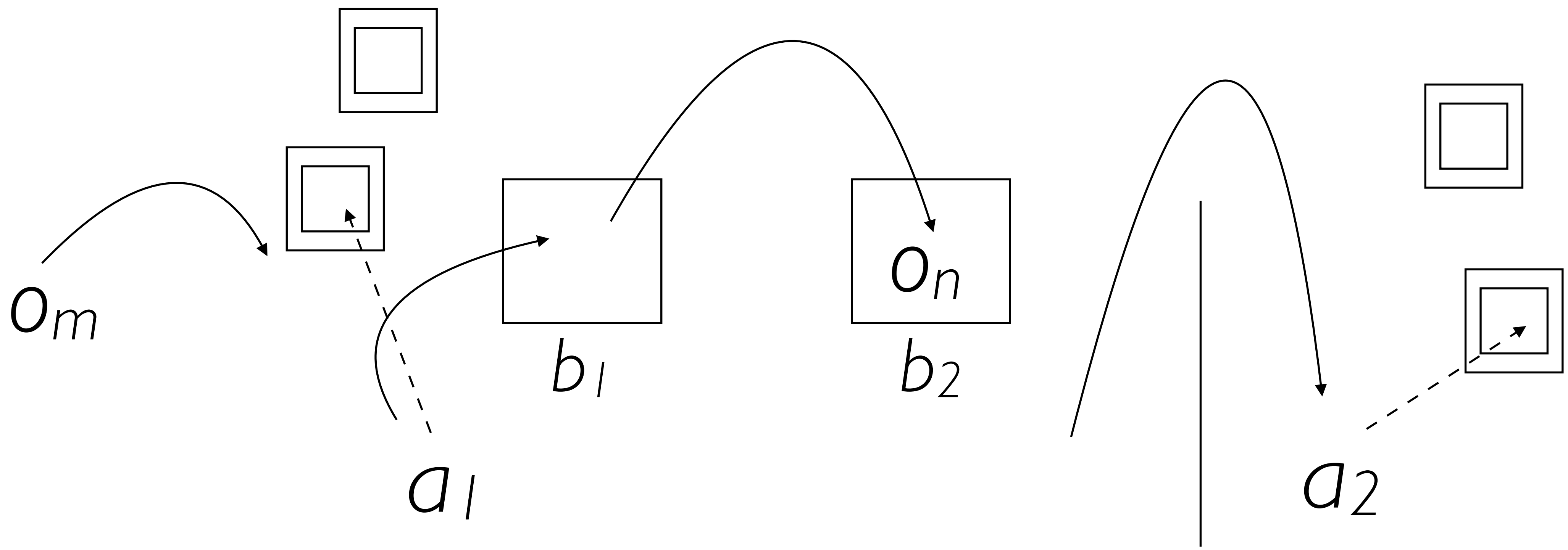
Framework for FBT^1_5

(ten timepoints)



Framework for FBT$^1{}_5$

(ten timepoints)

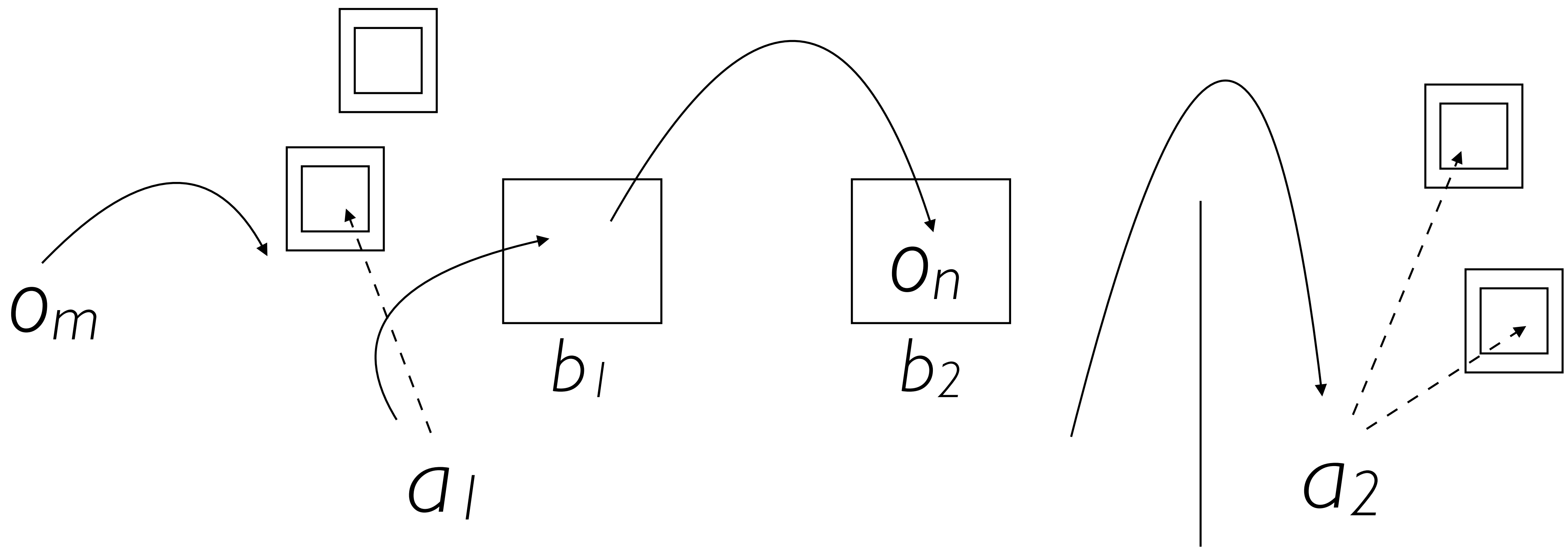


a

e

Framework for $FBT^1{}_5$

(ten timepoints)

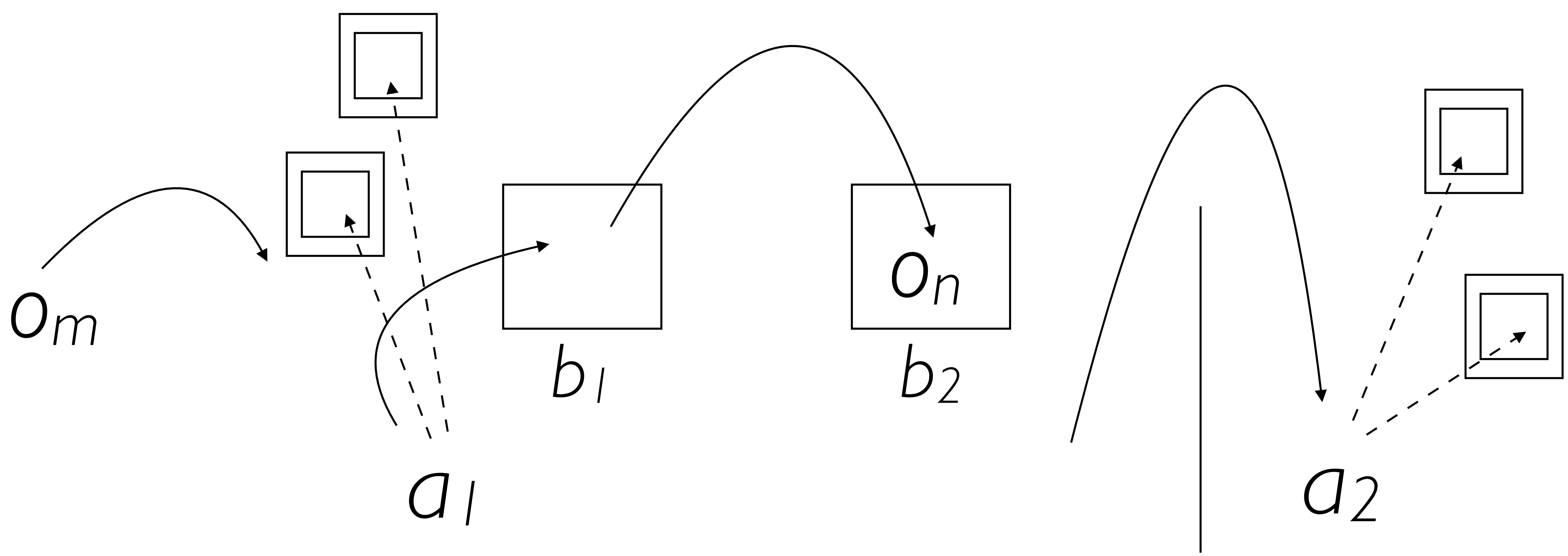


a

e

Framework for FBT^1_5

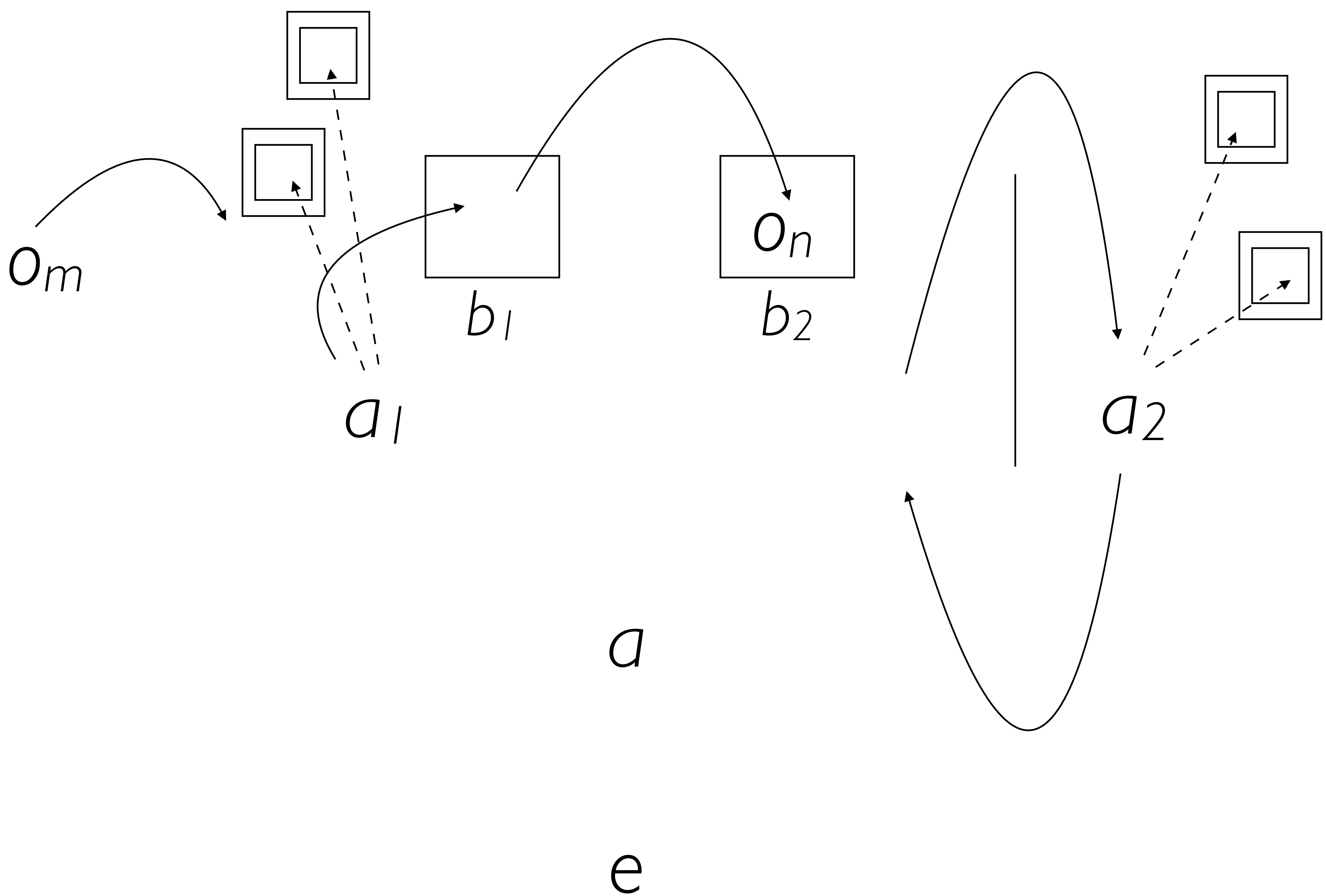
(ten timepoints)



a

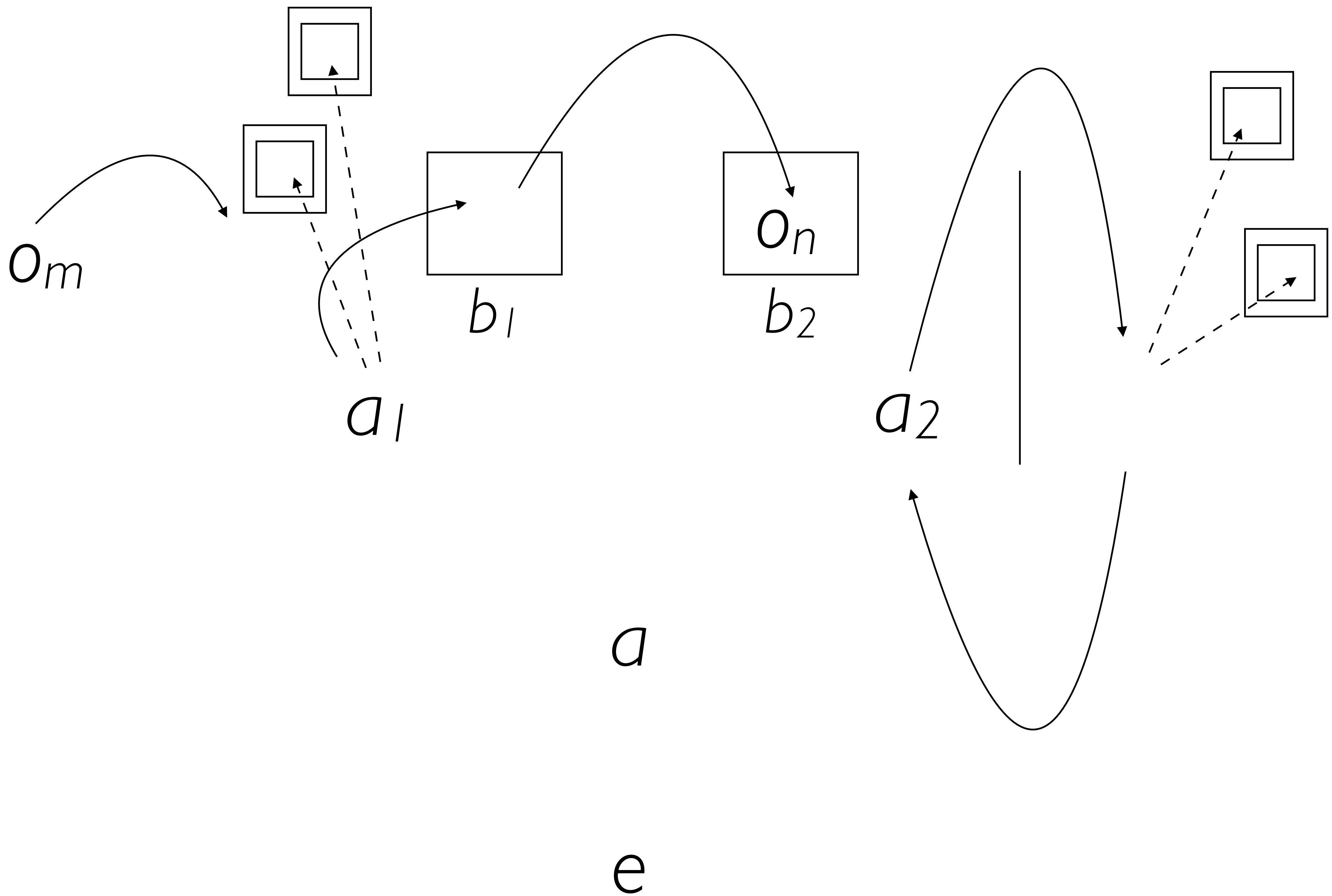
e

Framework for $FBT^1{}_5$
(ten timepoints)

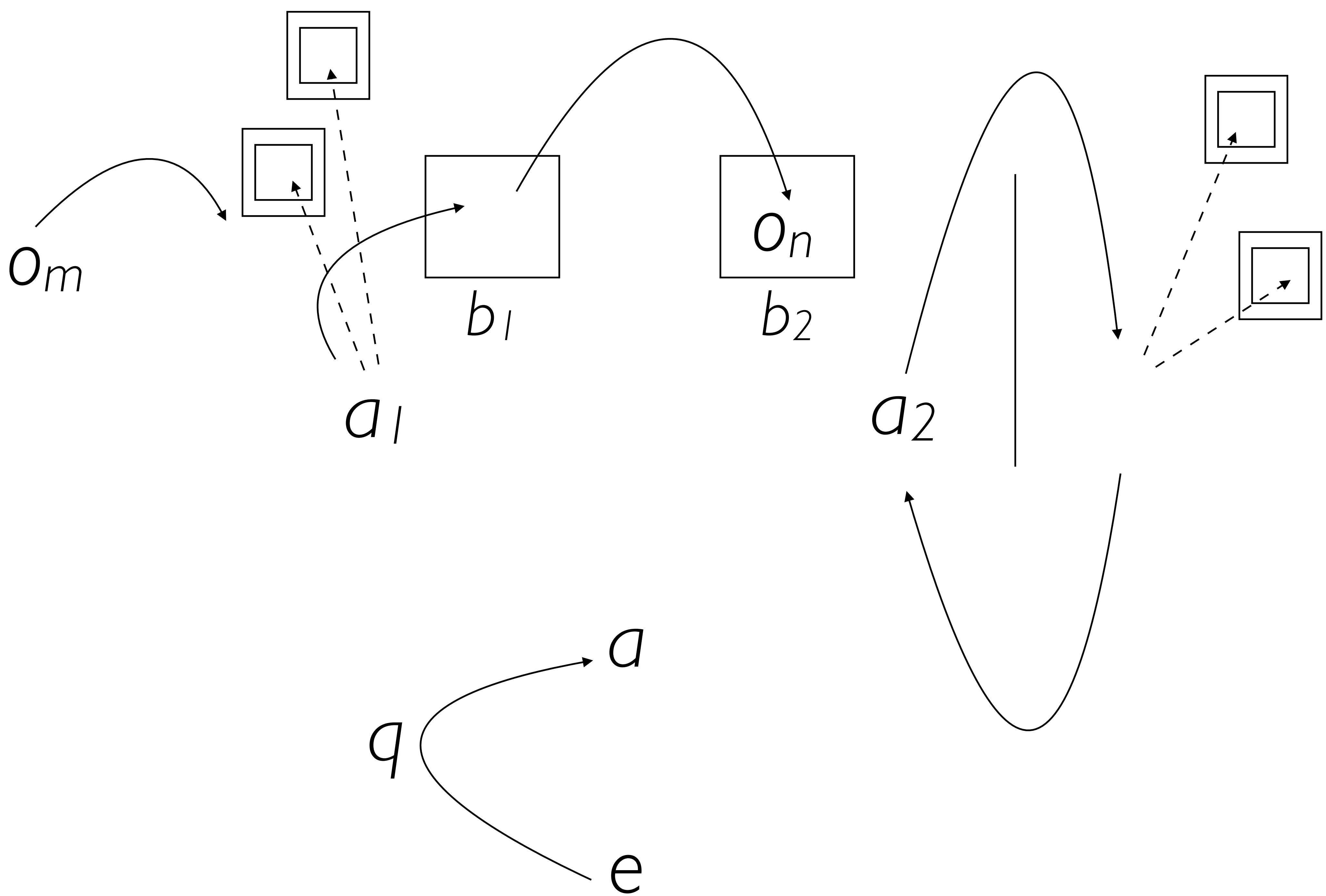


Framework for FBT^1_5

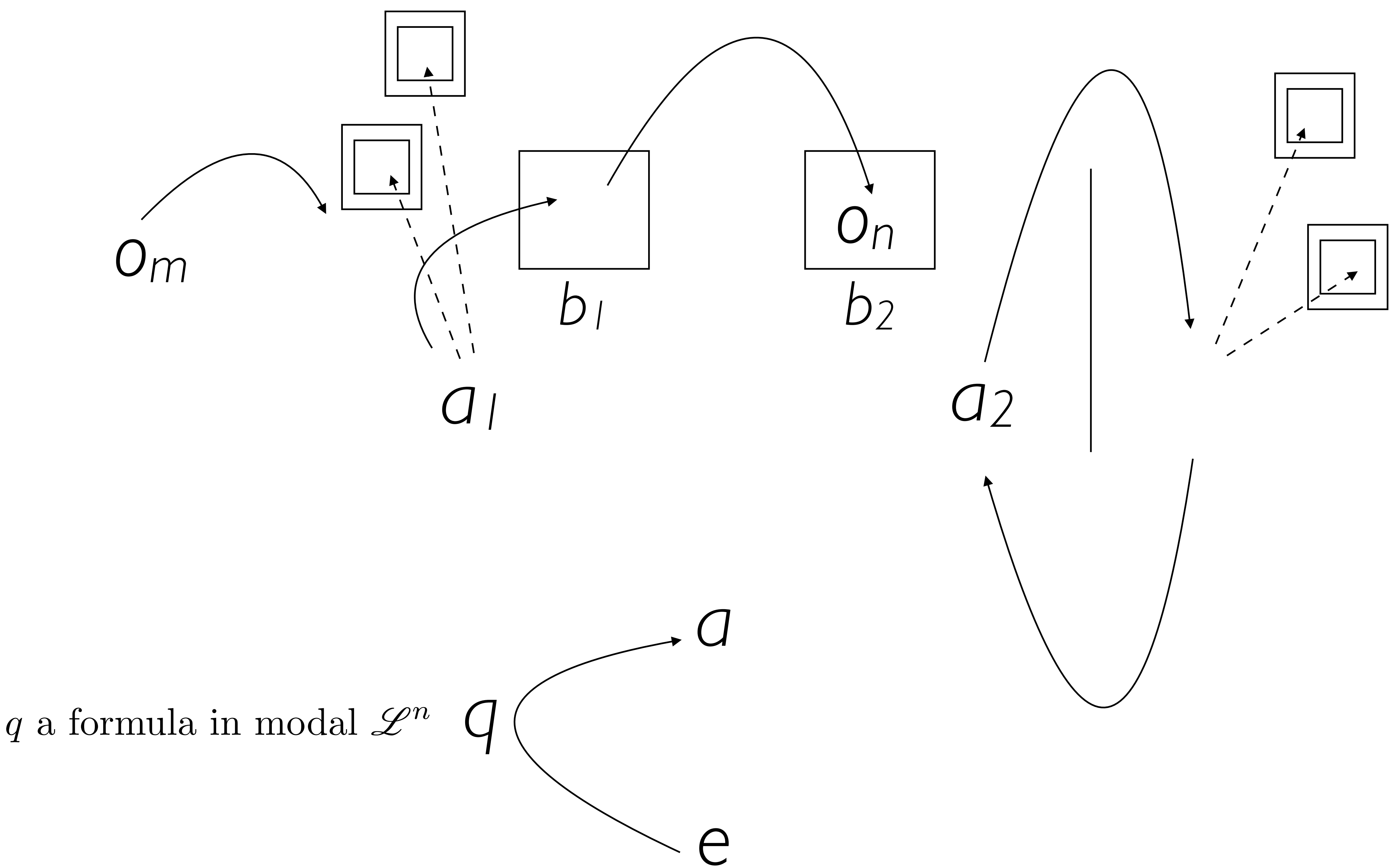
(ten timepoints)



Framework for $FBT^1{}_5$
(ten timepoints)



Framework for $FBT^1{}_5$
(ten timepoints)



Humans Can Succeed

Neurobiologically normal, nurtured, educated, and sufficiently motivated humans can correctly answer any relevant query *q* for the infinite progression, and prove that their answer is correct. For the obvious subclass of queries (the form of which appear in the box below), they can prove and exploit the following lemma.

Lemma: Suppose $\mathrm{FBT}_k, k \in \mathbb{Z}^+$, holds; (i.e. that level k of FBT holds). Then, if k is even, $\mathbf{B}_2\mathbf{B}_1 \ldots \mathbf{B}_2\ \iota$, where there are $k+1$ iterated $\mathbf{B}_i$ operators; otherwise $\mathbf{B}_1\mathbf{B}_2 \ldots \mathbf{B}_1\mathbf{B}_2\ \iota$, where there again there are $k+1$ iterated $\mathbf{B}_i$ operators.

Passing to Probing Mastery of the Specific Subclass

Experimenter to *a*: "At level *k*, from which box will a_2 attempt to retrieve the objects o_n? Prove it!"

Theoretical Machine Success on Infinite FBT!

Theorem: $\forall q \in \mathscr{C}\mathscr{C}, \mathfrak{M}$ can correctly answer and justify q. I.e., $\mathfrak{M}$ can pass FBT_ω.

Ok, so this logic machine exists in the *mathematical* universe; but does there exist an *implemented* machine with this power?

Theoretical Machine Success on Infinite FBT!

Theorem: $\forall q \in \mathscr{C}\mathscr{C}, \mathfrak{M}$ can correctly answer and justify q. I.e., $\mathfrak{M}$ can pass FBT_ω.

Ok, so this logic machine exists in the *mathematical* universe; but does there exist an *implemented* machine with this power?

Simulation Courtesy of …

ShadowProver!

Level 1

```
: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))))))}
```

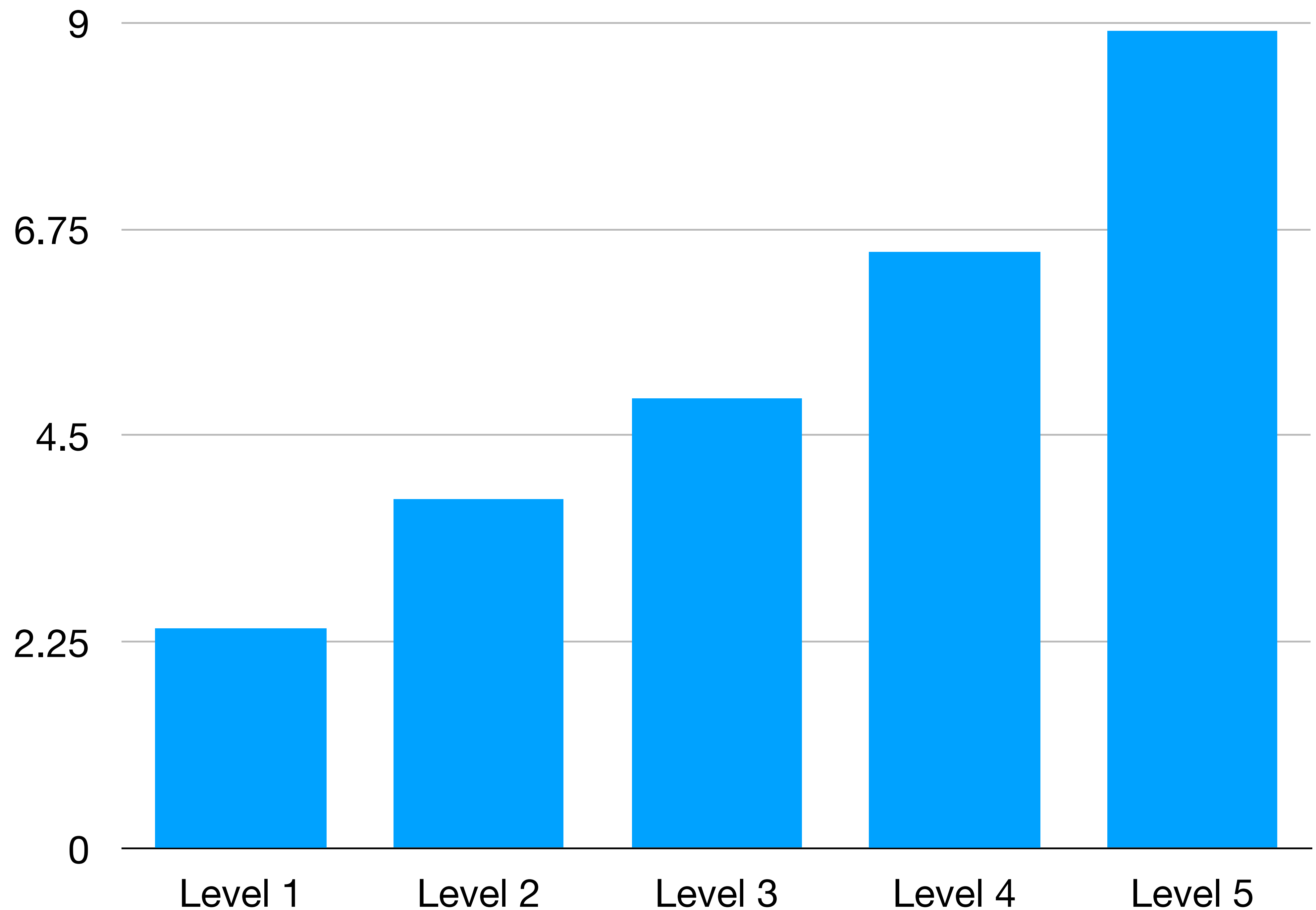
Level 5

```
{: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
...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 ...
...l 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))))))))}
```

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

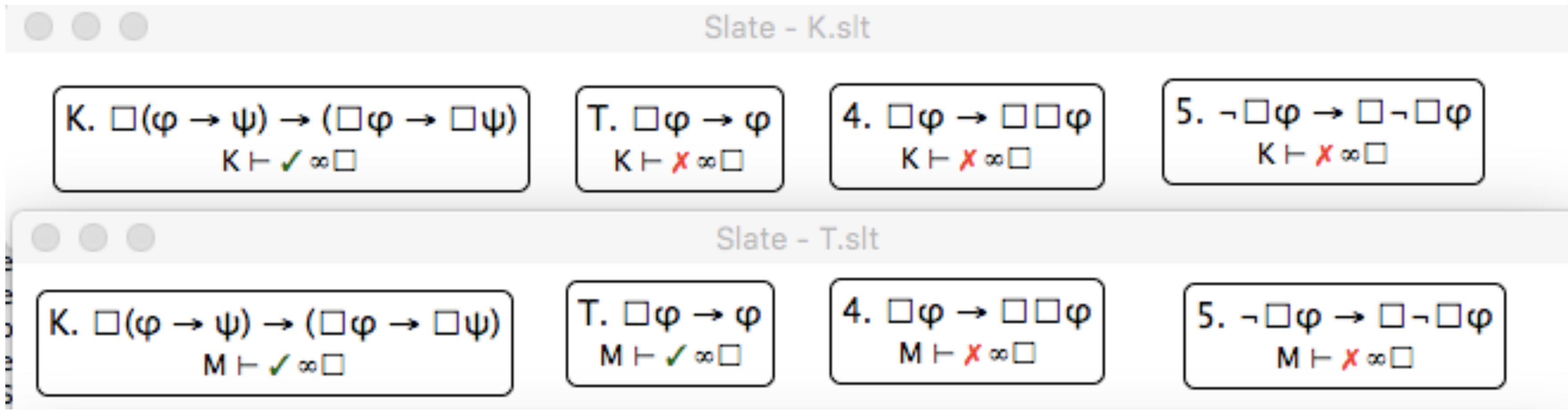
```
/Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java ...
objc[16653]: Class JavaLaunchHelper is implemented in both /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java (0x102a2d4c0) and /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/jre/lib/libinstrument.dylib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

Simulation of Level 5 in Real Time

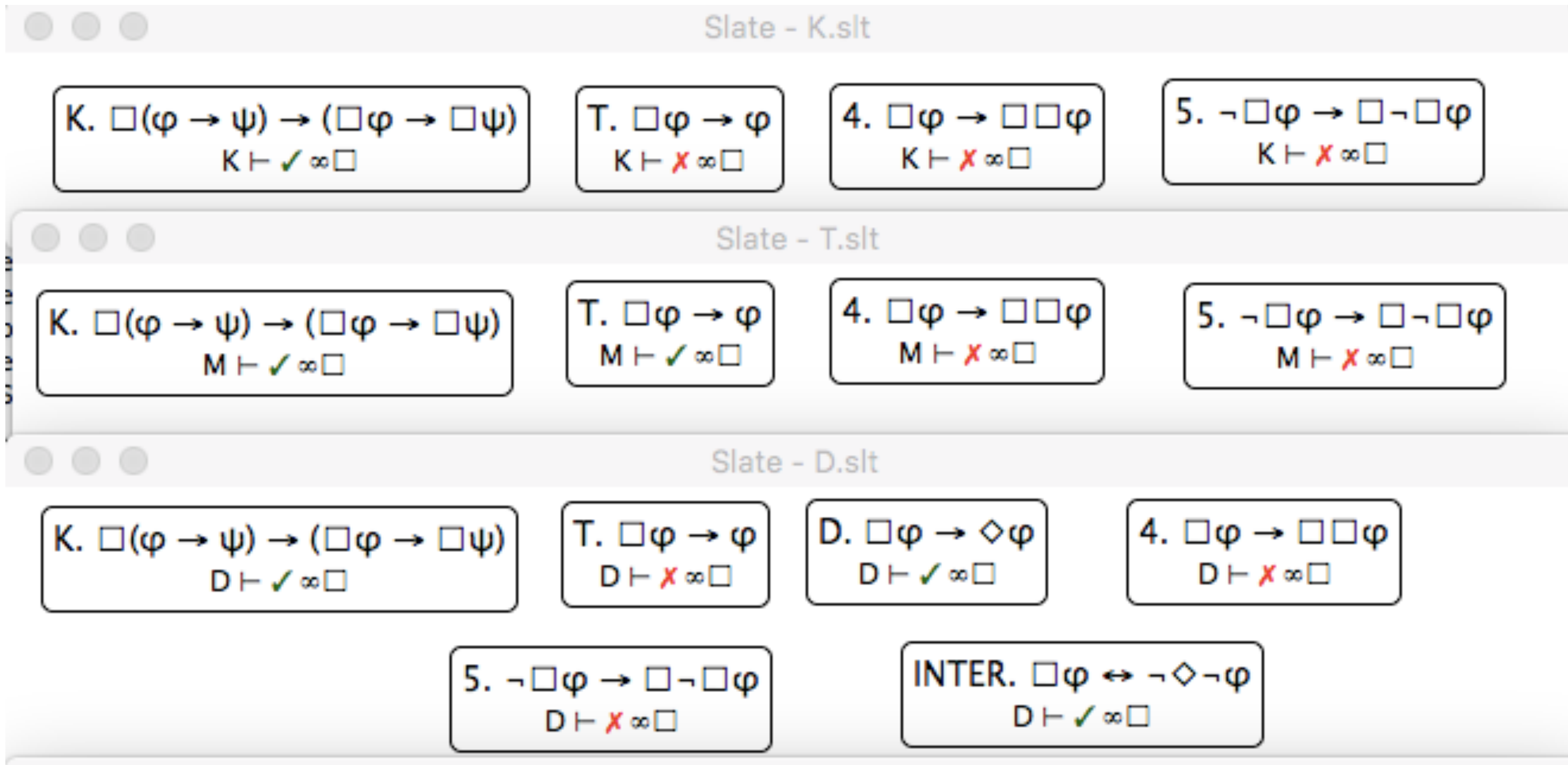
```
/Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java ...
objc[16653]: Class JavaLaunchHelper is implemented in both /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/bin/java (0x102a2d4c0) and /Library/Java/JavaVirtualMachines/jdk1.8.0_131.jdk/Contents/Home/jre/lib/libinstrument.dylib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

Encapsulation

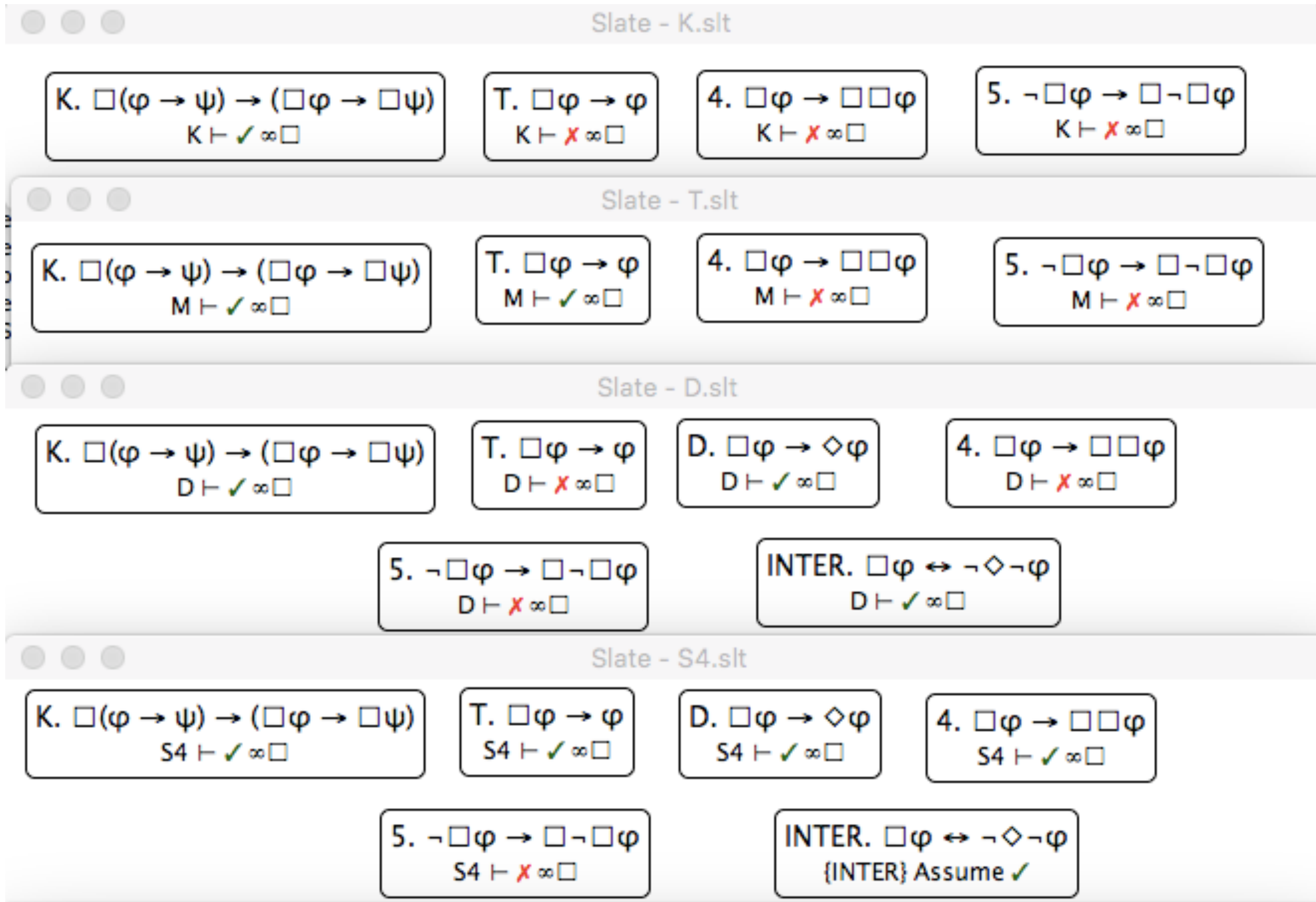
Encapsulation



Encapsulation



Encapsulation



Encapsulation

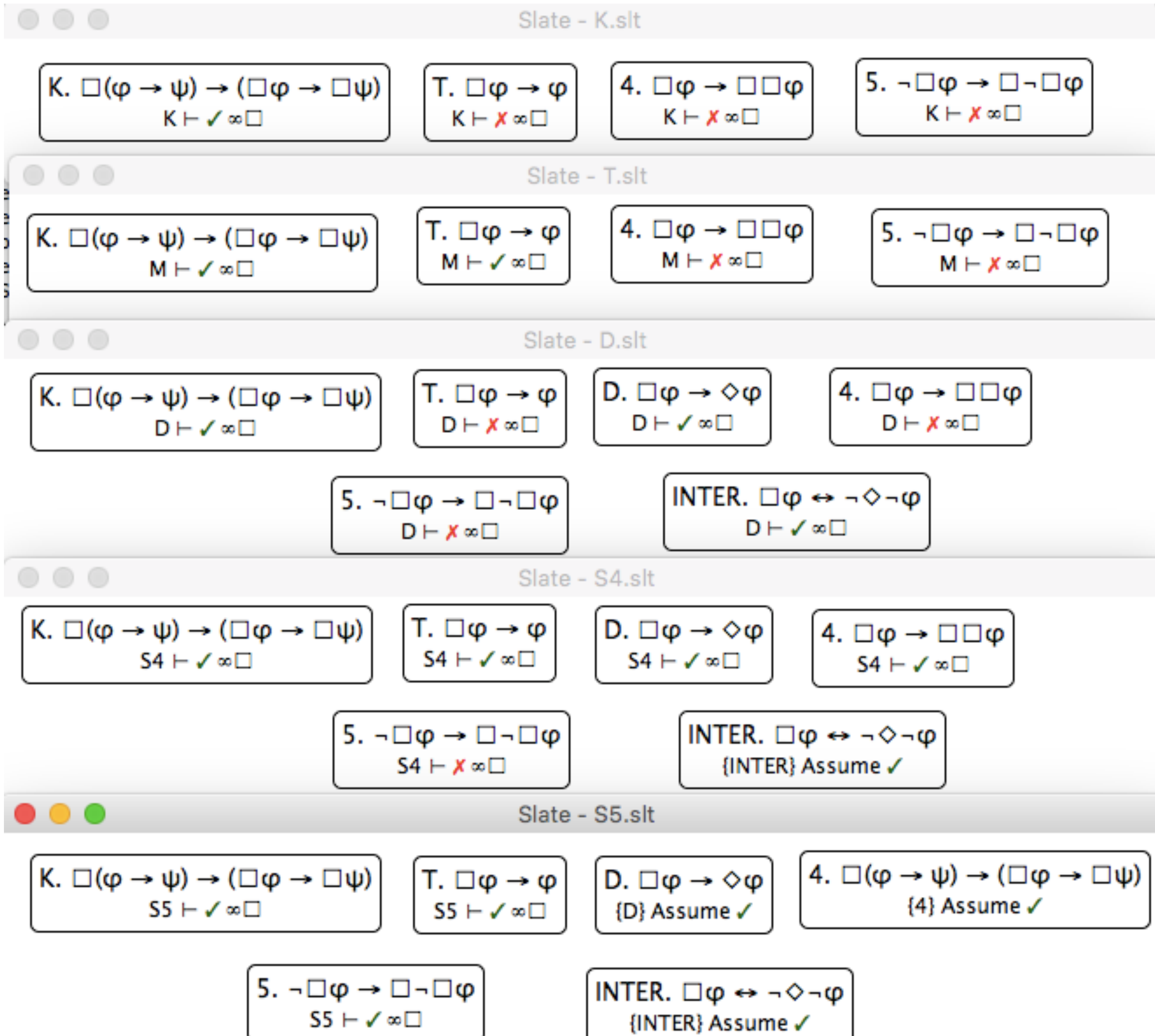
K

T

D

4 = S4

5 = S5



Encapsulation

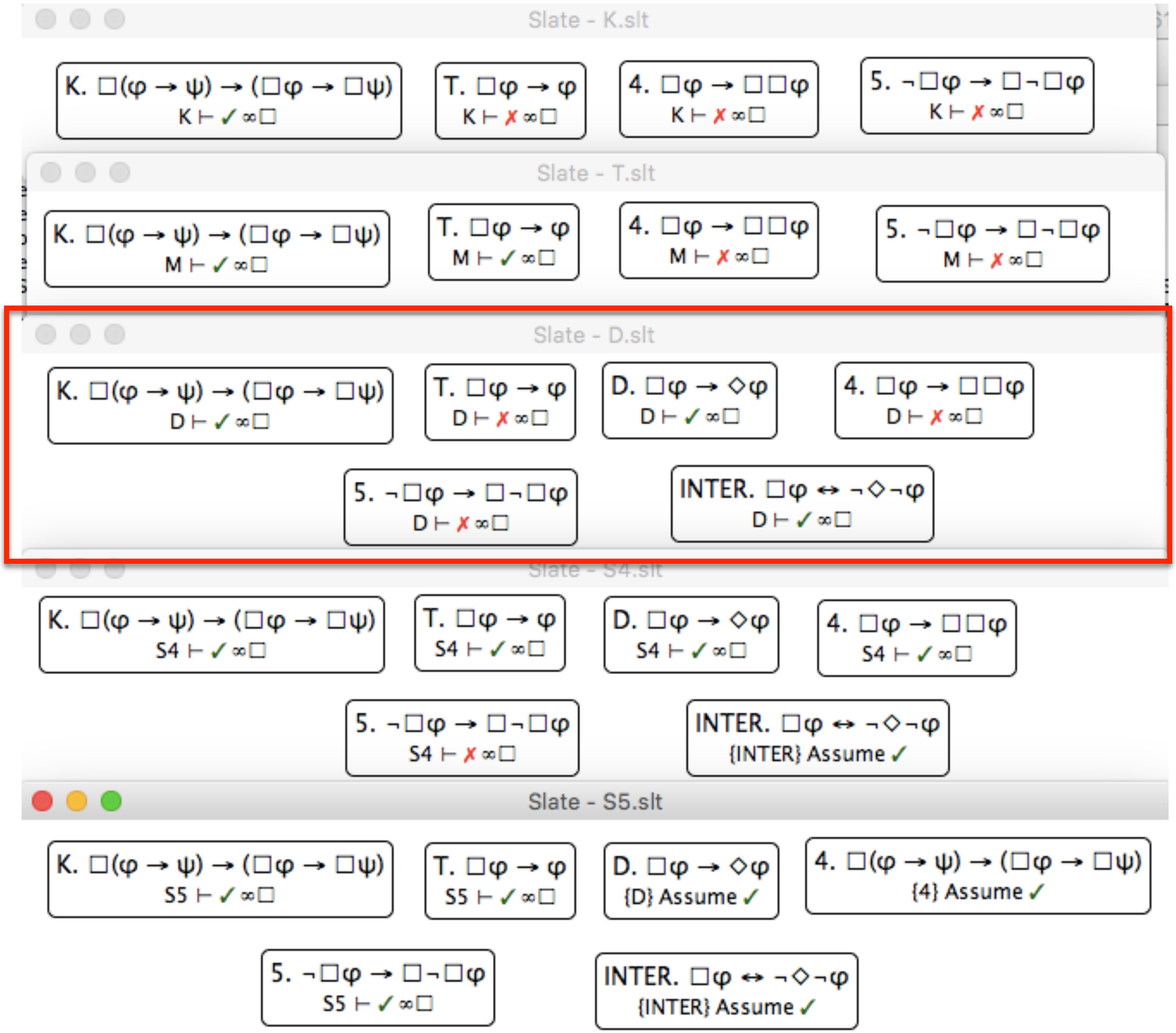
K

T

D

4 = S4

5 = S5



HyperSlate®

Encapsulation

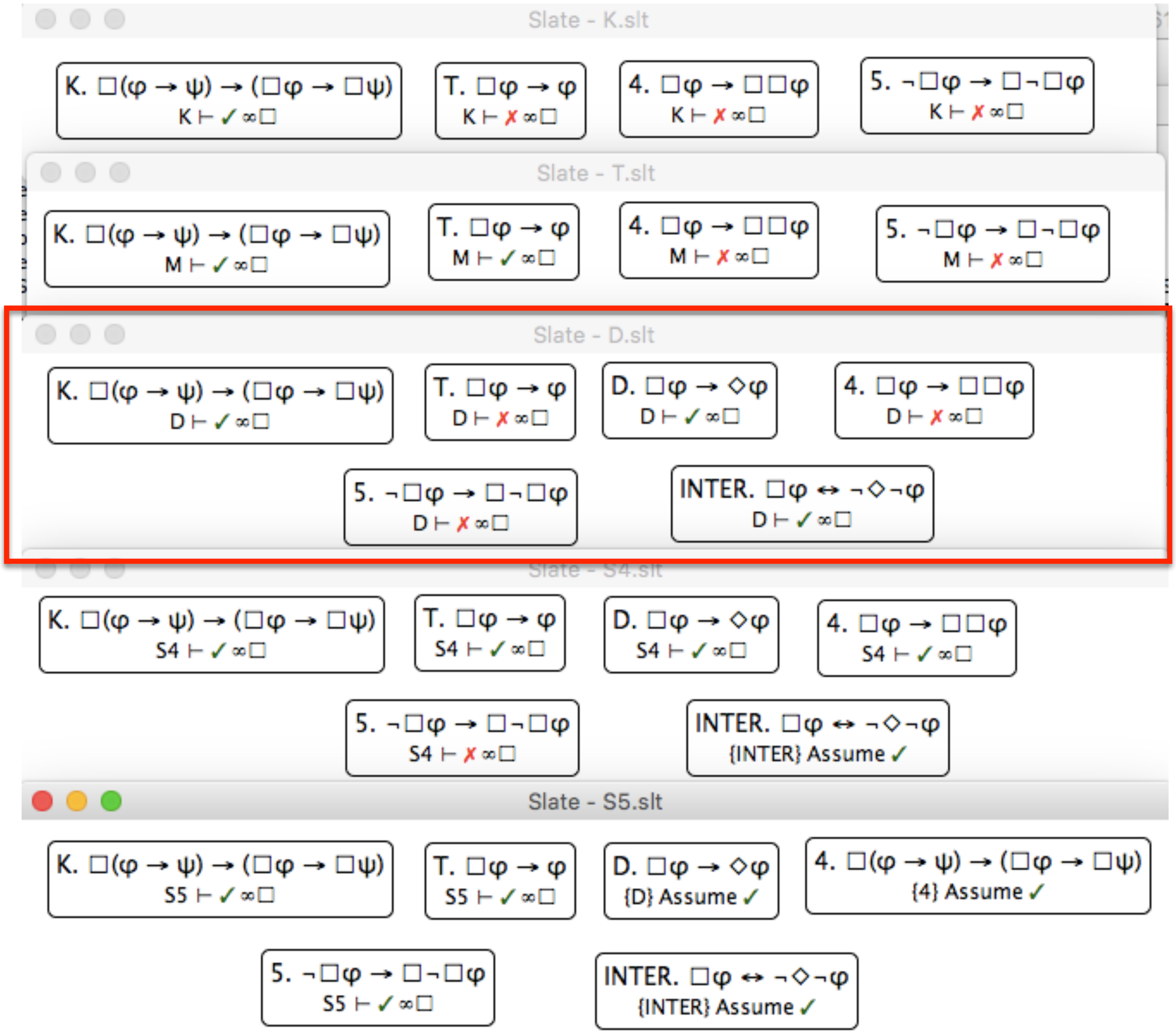
K

T

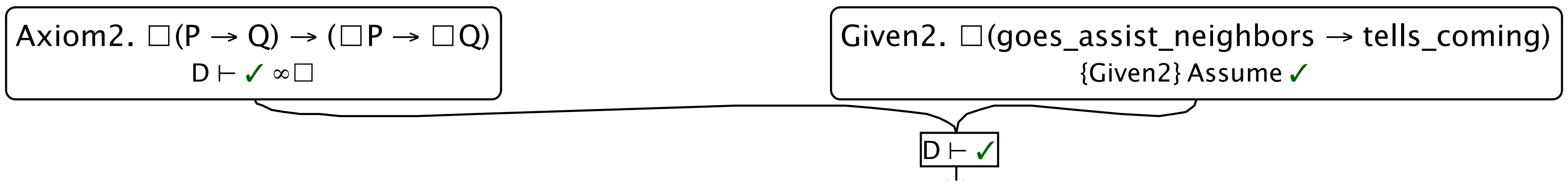
D

4 = S4

5 = S5



Chisholm's Paradox

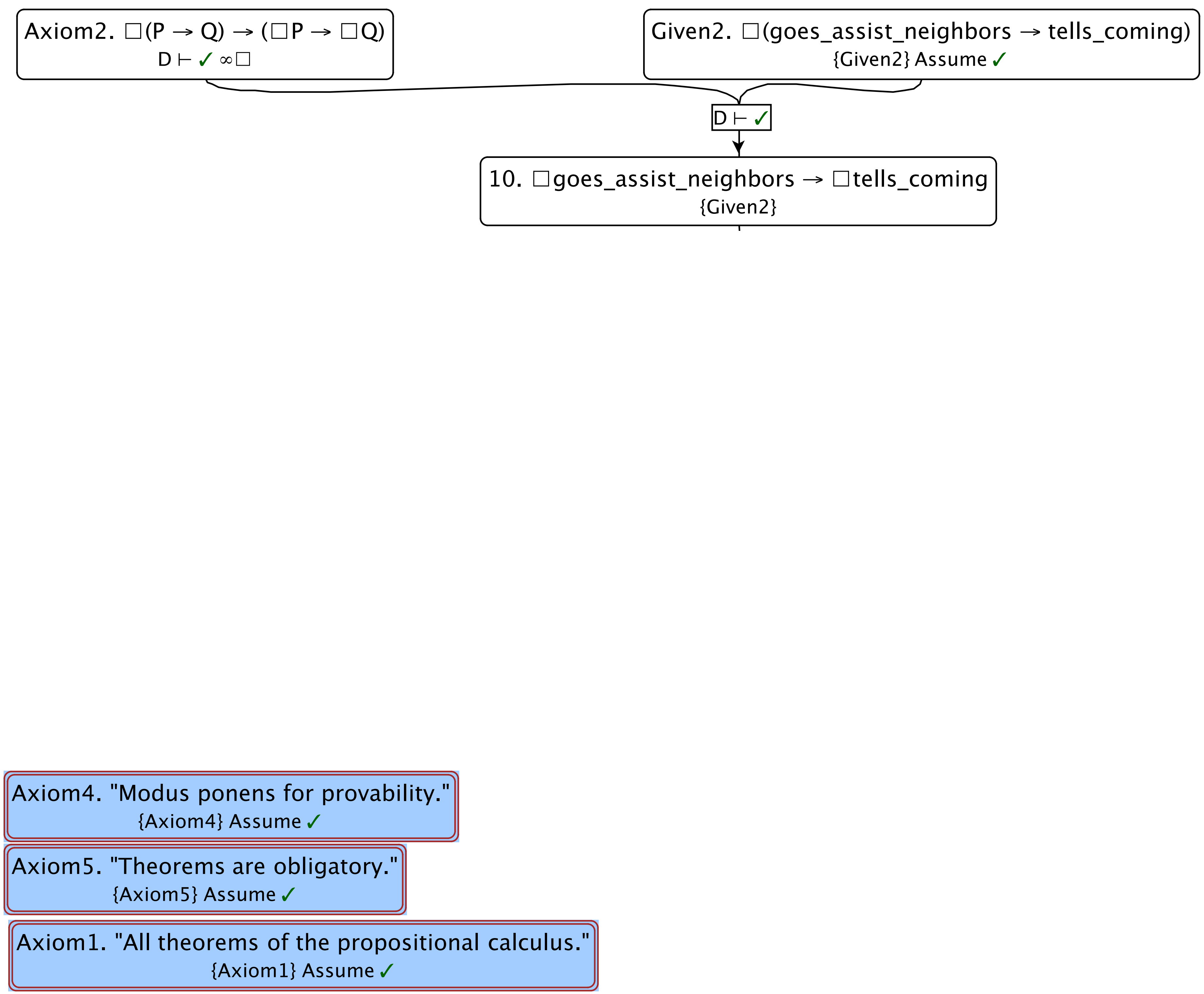


Axiom4. "Modus ponens for provability."
{Axiom4} Assume ✓

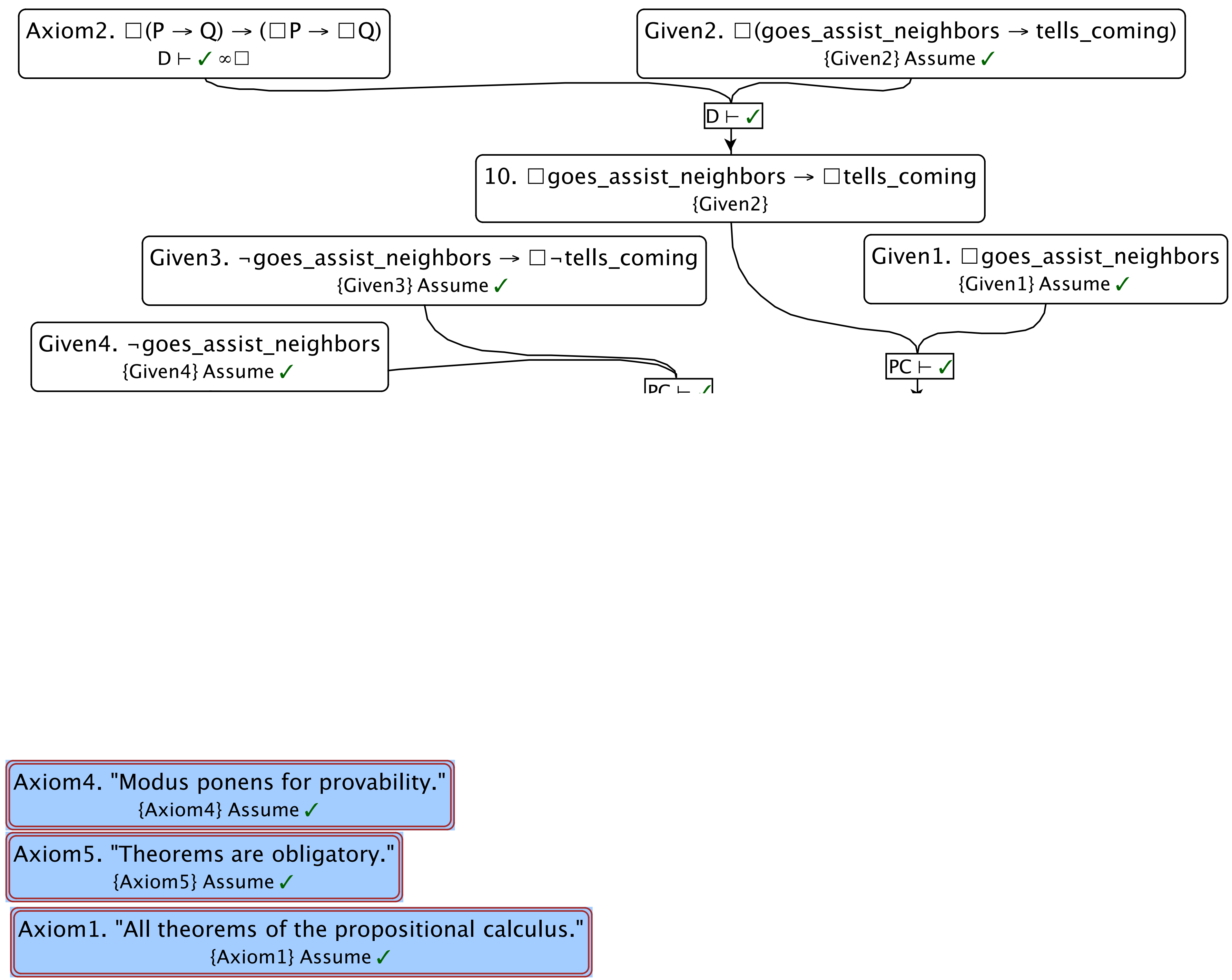
Axiom5. "Theorems are obligatory."
{Axiom5} Assume ✓

Axiom1. "All theorems of the propositional calculus."
{Axiom1} Assume ✓

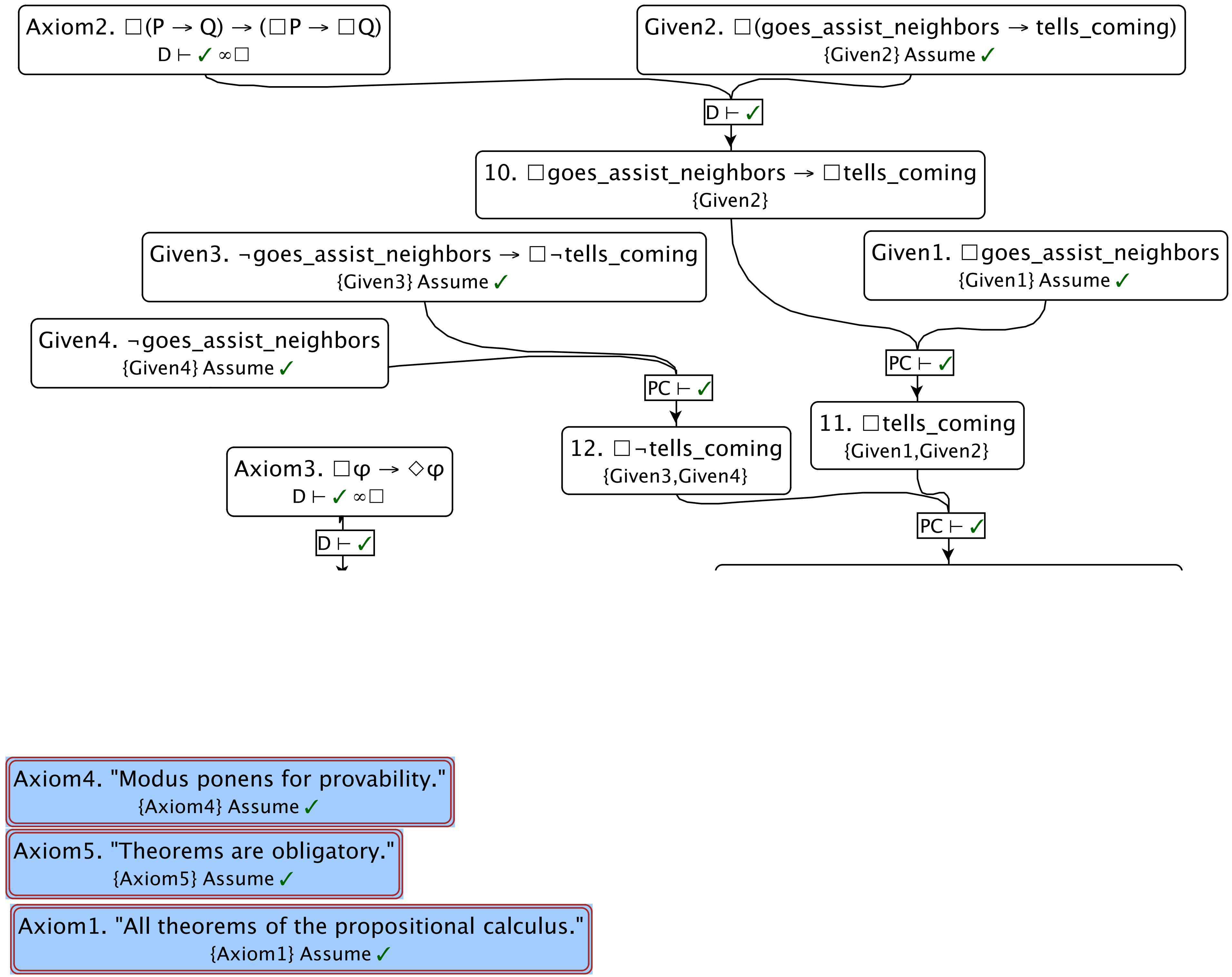
Chisholm’s Paradox



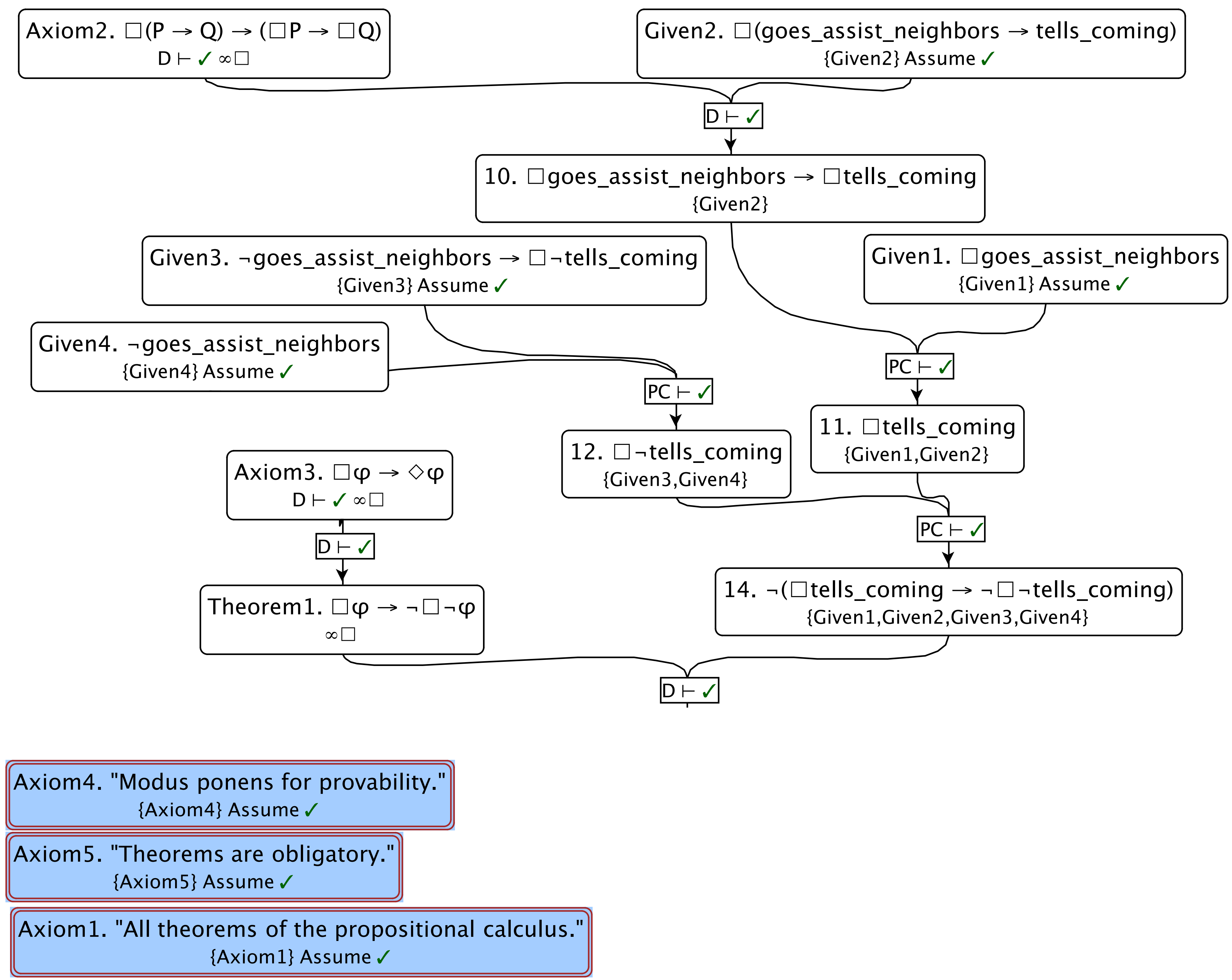
Chisholm's Paradox



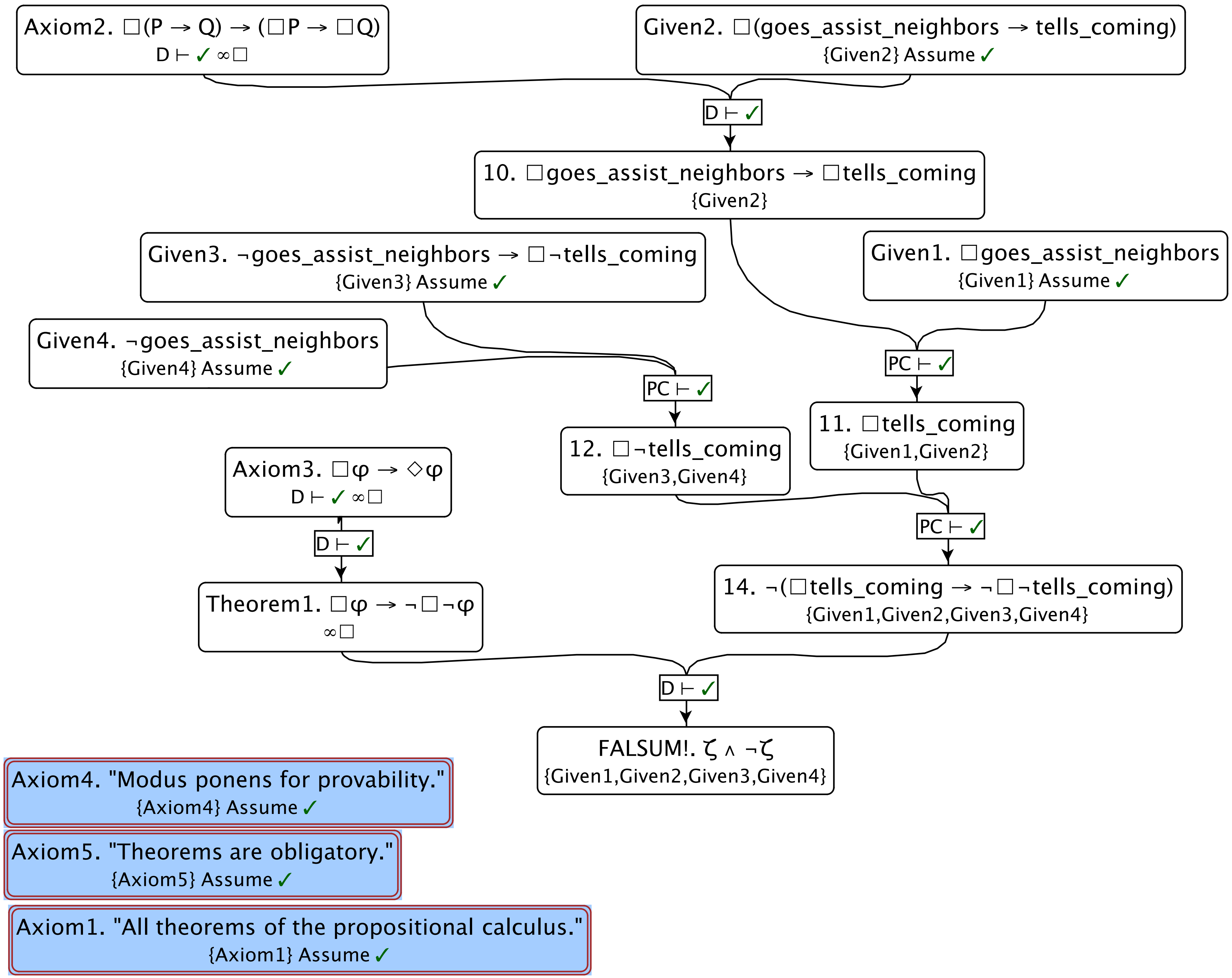
Chisholm's Paradox



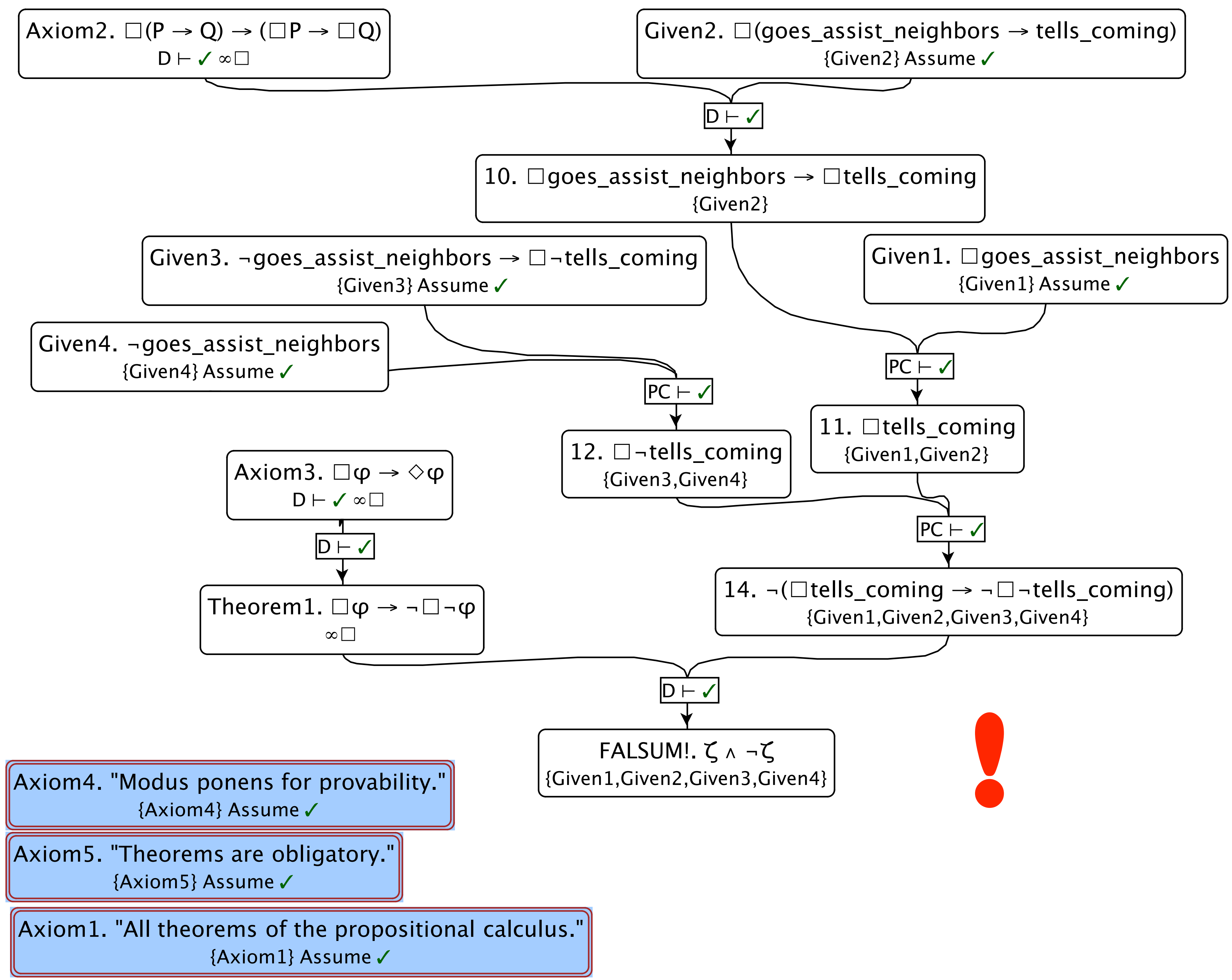
Chisholm's Paradox



Chisholm's Paradox



Chisholm's Paradox



HyperSlate®

Review: Encapsulation

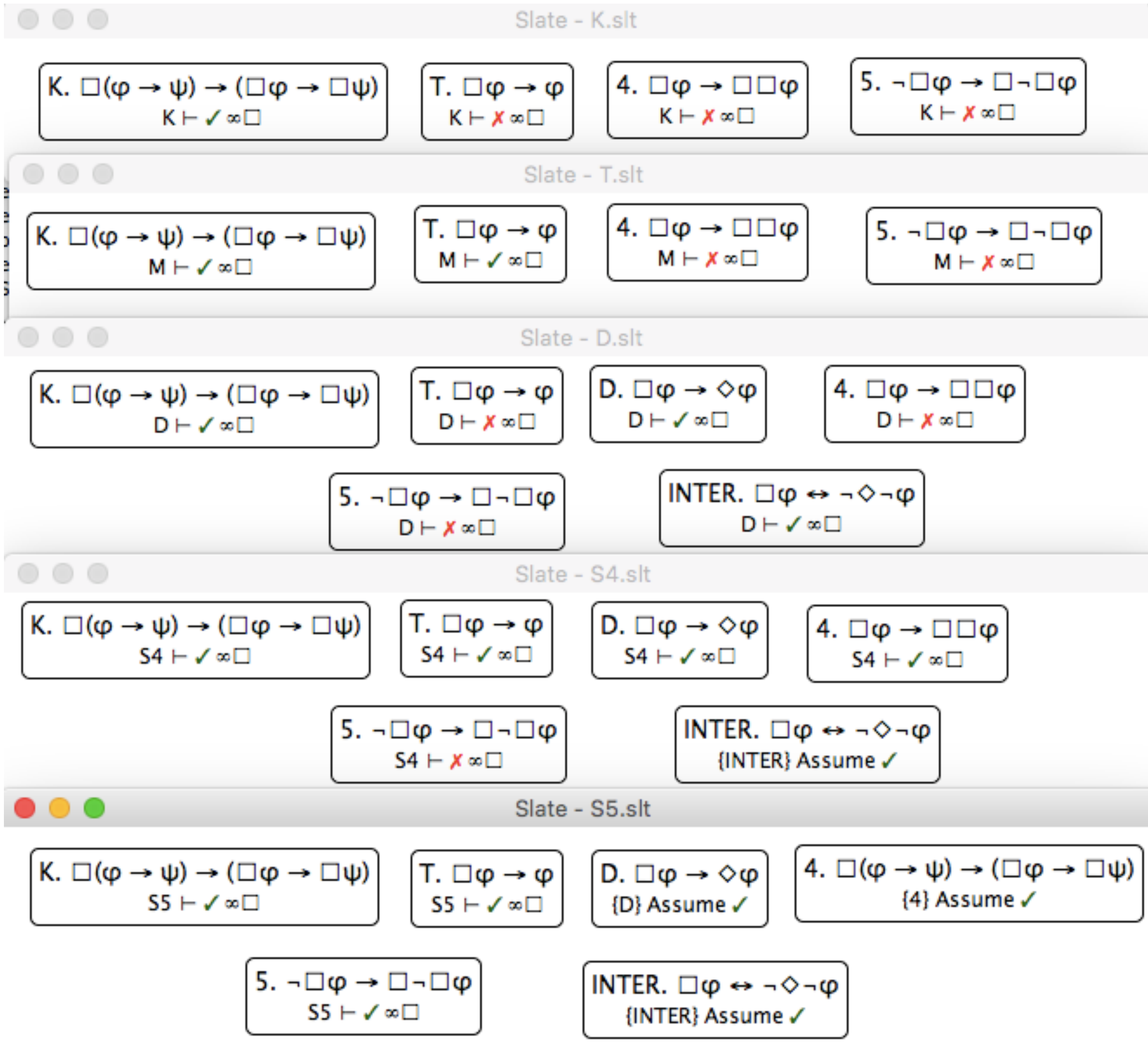
K

T

D

4 = S4

5 = S5



HyperSlate®

Review: Encapsulation

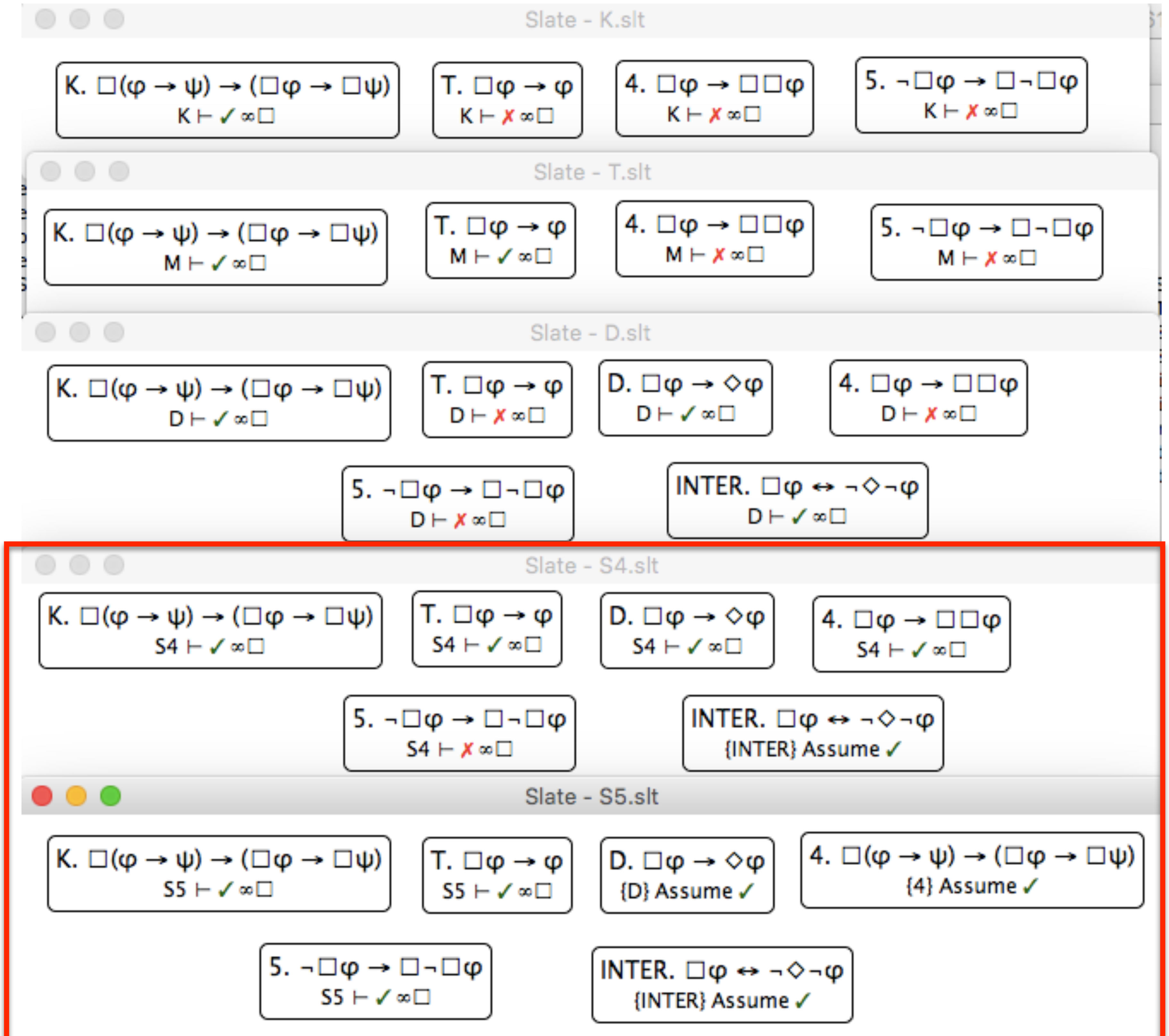
K

T

D

4 = S4

5 = S5



HyperSlate®

Review: Encapsulation

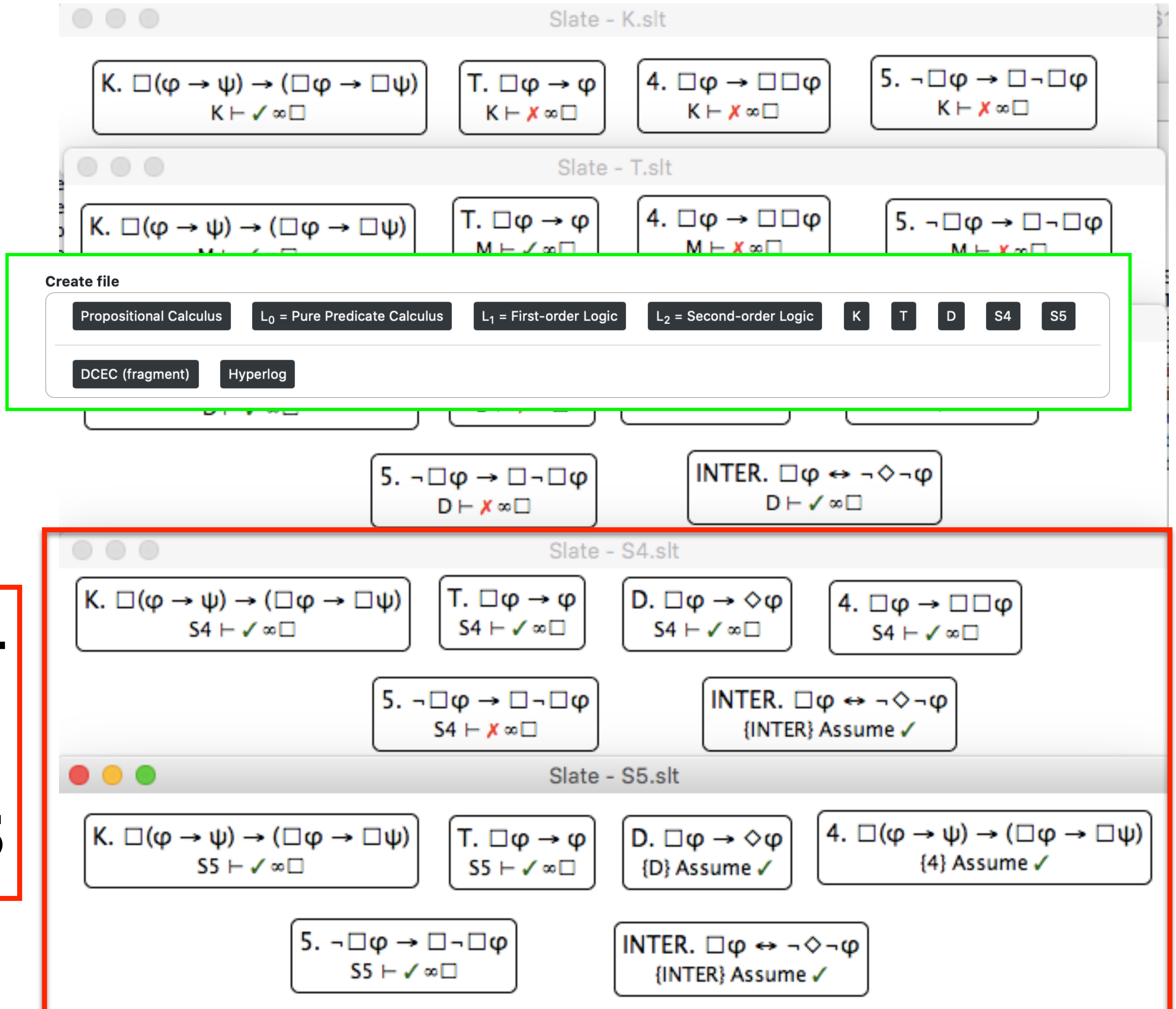
K

T

D

4 = S4

5 = S5



HyperSlate®

Review: Encapsulation

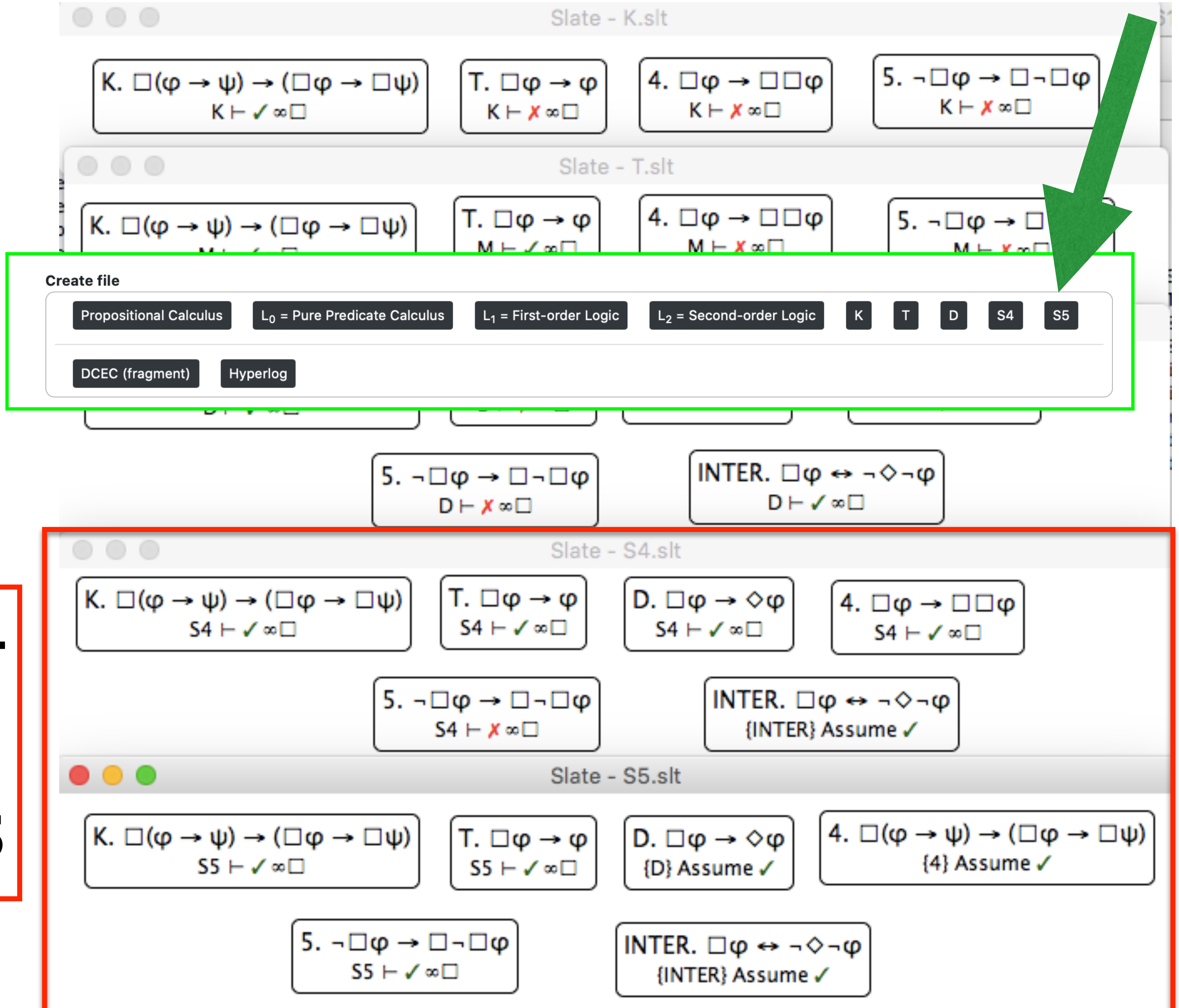
K

T

D

4 = S4

5 = S5



Det er en logikk for hvert problem!