### On to Intensional Logics

### **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

IFLAII 4/1/2021



# On those extensional logics, quickly ...

### FOL



## $\sqrt{FOL}$ Epistemic + FOL $_{\mathbf{B}_{d}\mathbf{B}_{v}\mathbf{B}_{d}Vv}$

# $\begin{array}{c} \checkmark \text{FOL} \\ \text{Epistemic + FOL} \text{ (for coverage of "killer" robots, later)} \\ \mathbf{B}_{d}\mathbf{B}_{v}\mathbf{B}_{d}Vv \end{array}$

# $\begin{array}{c} \checkmark \text{FOL} \\ \checkmark \text{Epistemic + FOL} \text{ (for coverage of "killer" robots, later)} \\ \mathbf{B}_{d}\mathbf{B}_{v}\mathbf{B}_{d}Vv \end{array}$

## $\sqrt{\text{FOL}}$ $\sqrt{\text{Epistemic}_{\mathbf{B}_d\mathbf{B}_v\mathbf{B}_dVv}} + \text{FOL}_{\text{(for coverage of "killer" robots, later)}}$

TOL

 $\exists X[X(j) \land \neg X(m) \land S(X)]$ 

### √ FOL

### √ Epistemic + FOL (for coverage of "killer" robots, later)

 $\mathbf{B}_d \mathbf{B}_v \mathbf{B}_d V v$ 

TOL

 $\exists X[X(j) \land \neg X(m) \land S(X)]$ 



### √ FOL

### $\sqrt{\text{Epistemic} + \text{FOL}}$ (for coverage of "killer" robots, later)

TOL

 $\exists X[X(j) \land \neg X(m) \land S(X)]$ 



The Contemporary Craft of Creating Characters Meets Today's Cognitive Architectures: A Case Study in Expressivity\*

Selmer Bringsjord • John Licato • Alexander Bringsjord version of 0121161500NY

### Contents

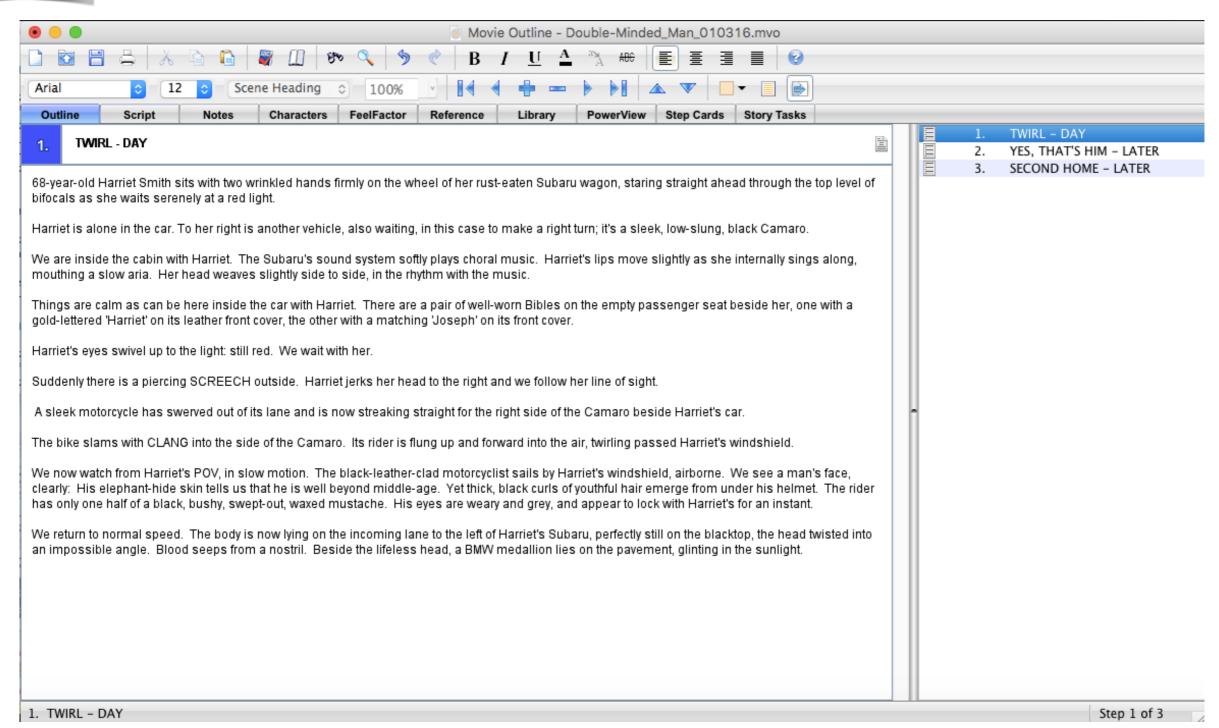
Abstract	2			
1 Introduction				
2 What is Expressivity?				
3 Double-Minded Man, and Expressivity Challenges Therefrom 3.1 Gist of Double-Minded Man 3.2 Basic Demands of Movie Outline 3 3.3 Harriet, Joseph, and a Pair of Expressivity Challenges	6			
4 The "No Way" Category and the Expressivity Challenges				
5 The "Maybe" Catgory ACT-R, and the Expressivity Challenges				
6 CLARION and the Expressivity Challenges 6.1 Overview of CLARION 6.1.1 NACS — the Non-Action-Centered Subsystem 6.2 The Expressivity of CLARION 6.2.1 Representing Structured Knowledge in General 6.2.2 FOL-level Expressivity in CLARION 6.2.3 Challenges (1) and (2) Met	9 11 11 12			
7 Concerns & Objections; Replies 7.1 Concern 1: "What is the purpose?" 7.2 Concern 2: "What exactly are these 'challenges'?" 7.3 Concern 3: "Can CLARION handles Double-Minded Man?" 7.4 Concern 4: "What about the character-centric Al of Mateas et al.?"				
8 Conclusion				
References				

<sup>\*</sup>We are indebted to two anonymous referees, and editor Jeremy Turner, for insightful suggestions and objections.

Any remaining deficiencies are due to our own failings.

```
The Centemporary Crist of Creating Characters
Motes Today & Cognitive Architecture:
A Cose Today is Expressively
Start Propage of Jain Leave. Section Propage
section of STATE ACCUSATION CONTRACTOR C
```

## The Contemporary Code of Counting Characters Motor Today & Cognitive Architectures: A Case Study in Engineering of the Counting Characters and Counti



```
The Centemporary Crist of Creating Characters
Motes Today & Cognitive Architecture:
A Cose Body is Expressively
Start Property of Jain Expressively
Start Designation
Start Start Designation
St
```

```
The Centemporary Cody of Considing Classactors
Meets Today's Cognitive Architectures:
A Case Study in Experiency (1974)
Solare Braggied Labit Latin - Stansactor Braggied
woman of IEEE Latin - Stansactor Braggied

Stansactor | 2

Latin - Latin - Stansactor Braggied

Latin - Latin - Latin - Stansactor Braggied

Latin - L
```

Double-Minded Man by

S Bringsjord & A Bringsjord

DRAFT #5 © June 30 2016

Selmer.Bringsjord@gmail.com

```
The Centemporary Crist of Creating Characters
Motes Today & Cognitive Architecture:
A Cose Body is Expressively
Start Property of Jain Expressively
Start Designation
Start Start Designation
St
```

## The Contemporary Cody of Conding Connectors Market Today Cognitive Manifestiments A Case Study in Expensivity\* Scient Program of the Connector Scient Program

### Double-Minded Man

### 1. TWIRL - DAY

68-year-old Harriet Smith sits with two wrinkled hands firmly on the wheel of her rust-eaten Subaru wagon, staring straight ahead through the top level of bifocals as she waits serenely at a red light.

Harriet is alone in the car. To her right is another vehicle, also waiting, in this case to make a right turn; it's a sleek, low-slung, black Camaro.

We are inside the cabin with Harriet. The Subaru's sound system softly plays choral music. Harriet's lips move slightly as she internally sings along, mouthing a slow aria. Her head weaves slightly side to side, in the rhythm with the music.

Things are calm as can be here inside the car with Harriet. There are a pair of well-worn Bibles on the empty passenger seat beside her, one with a gold-lettered 'Harriet' on its leather front cover, the other with a matching 'Joseph' on its front cover.

Harriet's eyes swivel up to the light; still red. We wait with her.

Suddenly there is a piercing SCREECH outside. Harriet jerks her head to the right and we follow her line of sight.

A sleek motorcycle has swerved out of its lane and is now streaking straight for the right side of the Camaro beside Harriet's car.

The bike slams with CLANG into the side of the Camaro. Its rider is flung up and forward into the air, twirling passed Harriet's windshield.

We now watch from Harriet's POV, in slow motion. The black-leather-clad motorcyclist sails by Harriet's windshield, airborne. We see a man's face, clearly: His elephant-hide skin tells us that he is well beyond middle-age. Yet thick, black curls of youthful hair emerge from under his helmet. The rider has only one half of a black, bushy, swept-out, waxed mustache. His eyes are weary and grey, and appear to lock with Harriet's for an instant.

We return to normal speed. The body is now lying on the incoming lane to the left of Harriet's Subaru, perfectly still on the blacktop, the head twisted into an impossible angle. Blood seeps from a nostril. Beside the lifeless head, a BMW medallion lies on the pavement, glinting in the sunlight.

## The Contemporary Craft of Creating Characters Meets Today's Cognitive Architectures A Case Today's Expressive of Comments South Today South Comments South Today South Claims Associate Tringing of woman of PETITION OF TODAY SOUTH COMMENTS SOUTH COMMENTS SOUTH CO

### Double-Minded Man

### 1. TWIRL - DAY

68-year-old Harriet Smith sits with two wrinkled hands firmly on the wheel of her rust-eaten Subaru wagor staring straight ahead through the top level of bifocals as she waits serenely at a red light.

Harriet is alone in the car. To her right is another vehicle, also waiting, in this case to make a right turn; it's a sleek, low-slung, black Camaro.

We are inside the cabin with Harriet. The Subaru's sound system softly plays choral music. Harriet's lips move slightly as she internally sings along, mouthing a slow aria. Her head weaves slightly side to side, in the rhythm with the music.

Things are calm as can be here inside the car with Harriet. There are a pair of well-worn Bibles on the empty passenger seat beside her, one with a gold-lettered 'Harriet' on its leather front cover, the other with a matching 'Joseph' on its front cover.

Harriet's eyes swivel up to the light; still red. We wait with her

Suddenly there is a piercing SCREECH outside. Harriet jerks her head to the right and we follow her line of sight.

A sleek motorcycle has swerved out of its lane and is now streaking straight for the right side of the Camaro beside Harriet's car.

The bike slams with CLANG into the side of the Camaro. Its rider is flung up and forward into the air, twirling passed Harriet's windshield.

We now watch from Harriet's POV, in slow motion. The black-leather-clad motorcyclist sails by Harriet's windshield, airborne. We see a man's face, clearly: His elephant-hide skin tells us that he is well beyond middle-age. Yet thick, black curls of youthful hair emerge from under his helmet. The rider has only one half of a black, bushy, swept-out, waxed mustache. His eyes are weary and grey, and appear to lock with Harriet's for an instant.

We return to normal speed. The body is now lying on the incoming lane to the left of Harriet's Subaruperfectly still on the blacktop, the head twisted into an impossible angle. Blood seeps from a nostril. Beside the lifeless head, a BMW medallion lies on the pavement, glinting in the sunlight.

### 1. TWIRL - DAY

68-year-old Harriet Smith sits with two wrinkled hands firmly on the wheel of her rust-eaten Subaru wagon staring straight ahead through the top level of bifocals as she waits serenely at a red light.

Harriet is alone in the car. To her right is another vehicle, also waiting, in this case to make a right turn; it's a sleek, low-slung, black Camaro.

We are inside the cabin with Harriet. The Subaru's sound system softly plays choral music. Harriet's lips

### $\exists X[X(joseph) \land \neg X(m(harriet, joseph)) \land Sleazy(X)]$

empty passenger seat beside her, one with a gold-lettered 'Harriet' on its leather front cover, the other with a matching 'Joseph' on its front cover.

Harriet's eves swivel up to the light; still red. We wait with her

Suddenly there is a piercing SCREECH outside. Harriet jerks her head to the right and we follow her line of sight.

A sleek motorcycle has swerved out of its lane and is now streaking straight for the right side of the Camaro beside Harriet's car.

The bike slams with CLANG into the side of the Camaro. Its rider is flung up and forward into the air twirling passed Harriet's windshield.

We now watch from Hamet's POV, in slow motion. The black-leather-clad motorcyclist sails by Hamet's windshield, airborne. We see a man's face, clearly: His elephant-hide skin tells us that he is well beyond middle-age. Yet thick, black curls of youthful hair emerge from under his helmet. The rider has only one half of a black, bushy, swept-out, waxed mustache. His eyes are weary and grey, and appear to lock with Hamet's for an instant.

We return to normal speed. The body is now lying on the incoming lane to the left of Harriet's Subaruperfectly still on the blacktop, the head twisted into an impossible angle. Blood seeps from a nostril. Beside the lifeless head, a BMW medallion lies on the pavement, glinting in the sunlight.



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

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

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

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

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

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

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

 $\mbox{FOL} \quad \exists x [Llama(x) \land Llama(b) \land Likes(x,b) \land Llama(fatherOf(x))]$ 

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

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

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

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

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

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

Things x and y, along with the father of x, share a certain property; and, x  $R^2$ s 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**  $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$ 

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

Things x and y, along with the father of x, share a certain property; and, x  $R^2$ s 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**  $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$ 

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

Things x and y, along with the father of x, share a certain property; and, x  $R^2$ s 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**  $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$ 

**TOL** 
$$\exists x, y \; \exists R, R^2[R(x) \land R(y) \land R^2(x, y) \land Positive(R^2) \land R(fatherOf(x))]$$
  
Things  $x$  and  $y$ , along with the father of  $x$ , share a certain property; and,  $x \; R^2$ s  $y$ , 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 
$$\exists x[Llama(x) \land Llama(b) \land Likes(x,b) \land Llama(fatherOf(x))]$$
  
 $\mathscr{L}_1$  There's some thing which is a llama and likes  $b$  (which is also a llama), and whose father is a llama too.

**ZOL** 
$$Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$$
 $\mathscr{L}_0$   $a$  is a llama, as is  $b$ ,  $a$  likes  $b$ , and the father of  $a$  is a llama as well.

•

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

Things x and y, along with the father of x, share a certain property; and,  $x R^2$ s 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).

 $\mathcal{L}_2$ 

 $\mathscr{L}_1$ 

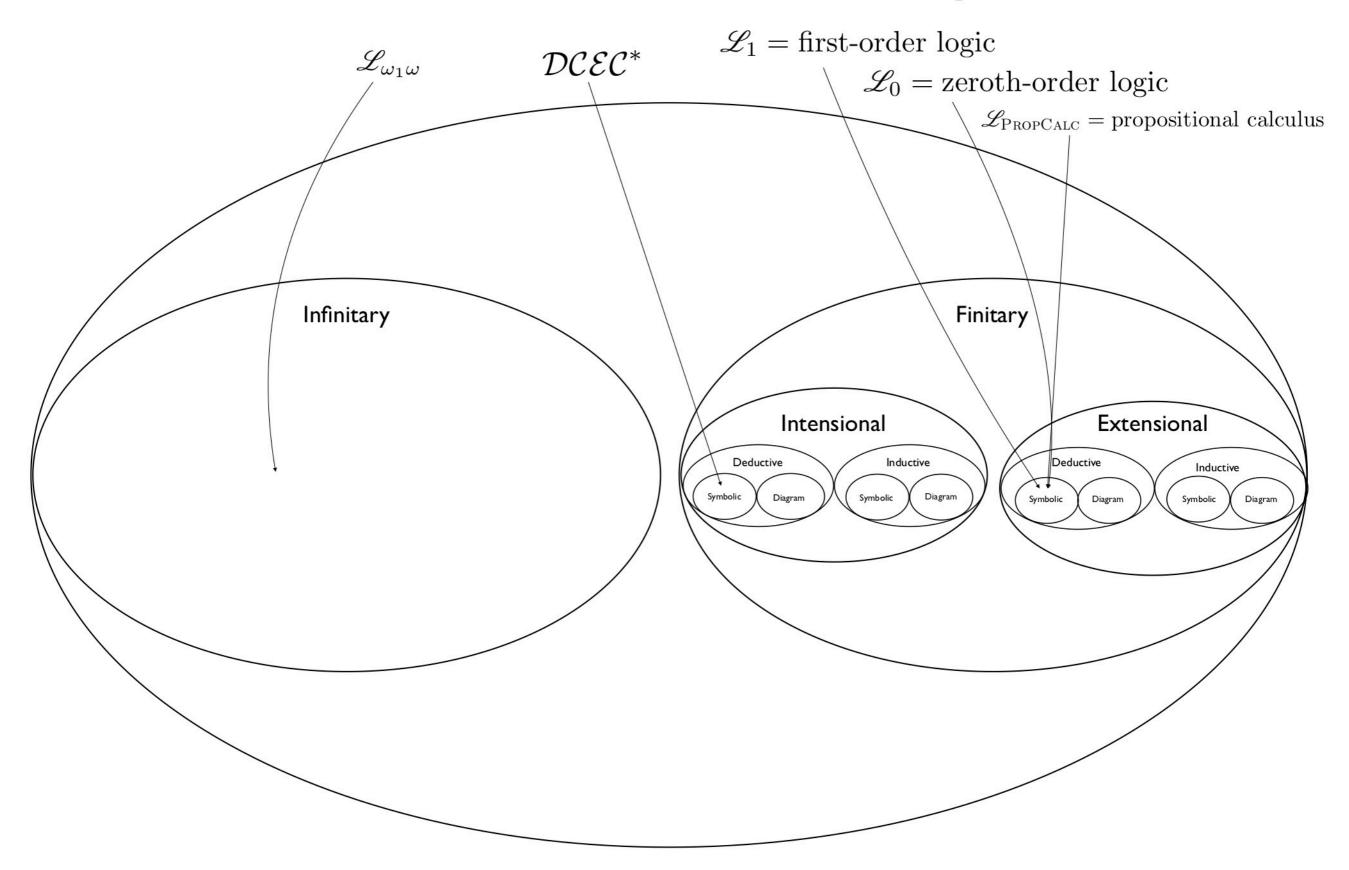
 $\mathscr{L}_0$ 

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

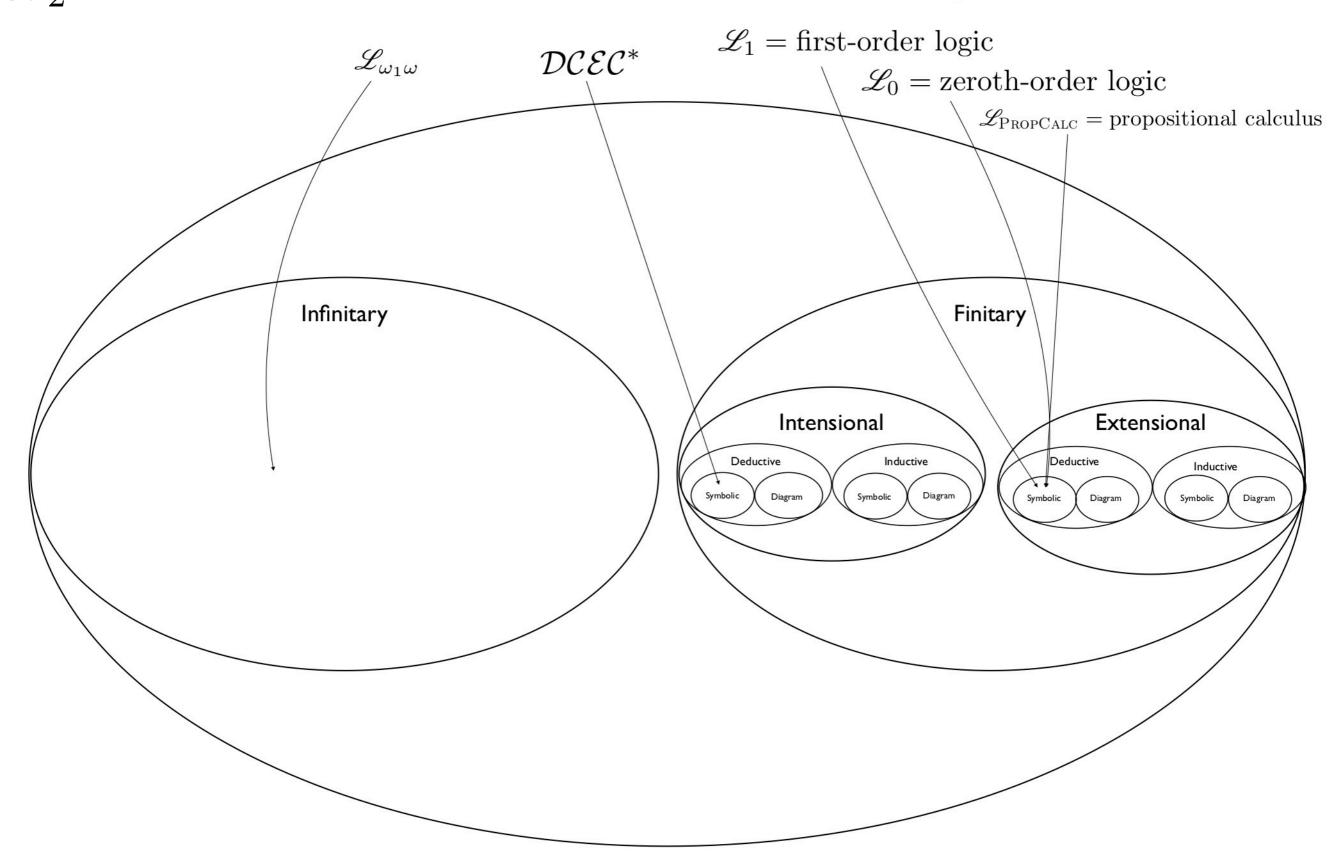
There's some thing which is a llama and likes b (which is also a llama), and whose father is a llama too.

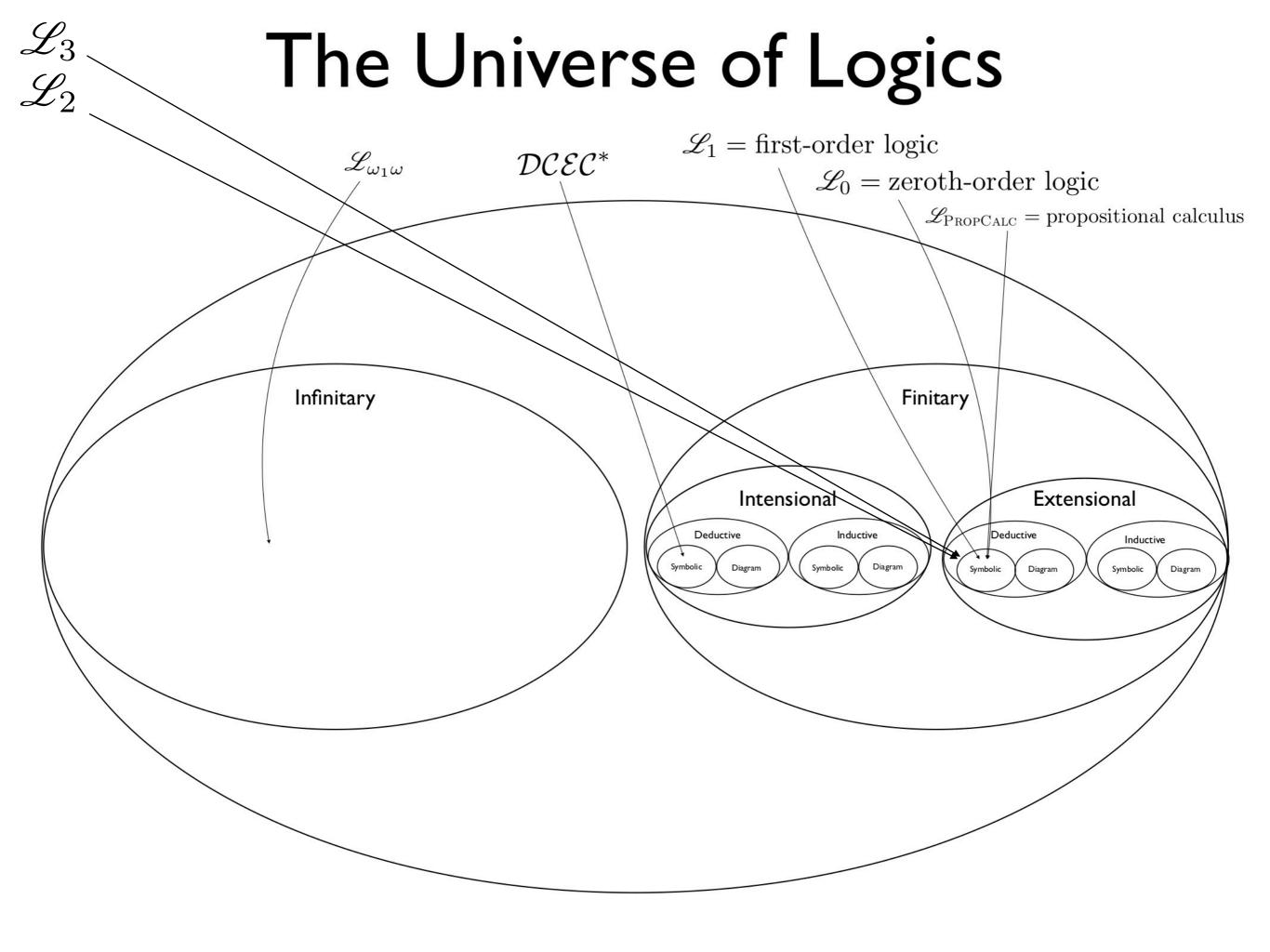
**ZOL**  $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$ 

## The Universe of Logics



## The Universe of Logics





## Climbing the k-order Ladder

•

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

Things x and y, along with the father of x, share a certain property; and,  $x R^2$ s 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).

 $\mathcal{L}_2$ 

 $\mathscr{L}_1$ 

 $\mathscr{L}_0$ 

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

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

•

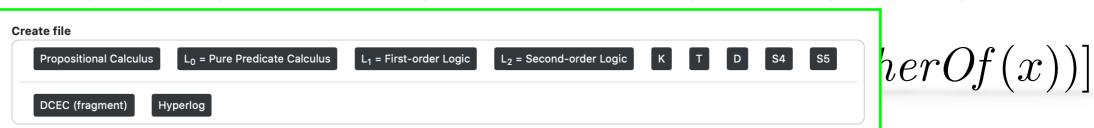
TOL

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

 $\mathscr{L}_3$ 

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

SOL L



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

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

 $\mathscr{L}_0$ 

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

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

 $\mathcal{L}_3$ 

Things x and y, along with the father of x, share a certain property; and,  $x R^2s y$ , where x = x + y + y = 0 a positive property.

SOL L Propositional Calculus L<sub>0</sub> = Pure Predicate Calculus L<sub>1</sub> = First-order Logic L<sub>2</sub> = Second-order Logic K T D S4 S5

DCEC (fragment) Hyperlog

herOf(x))]

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

**FOL** 

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

 $\mathscr{L}_1$ 

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

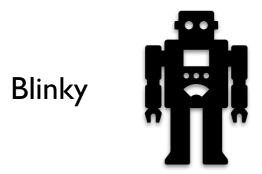
ZOL

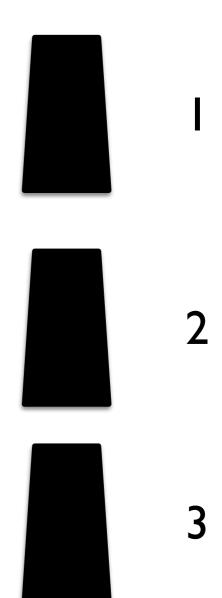
 $Llama(a) \wedge Llama(b) \wedge Likes(a,b) \wedge Llama(fatherOf(a))$ 

 $\mathscr{L}_0$ 

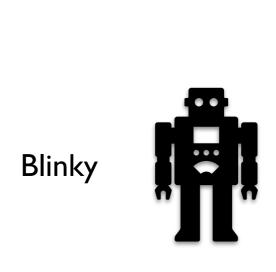
a is a llama, as is b, a likes b, and the father of a is a llama as well.

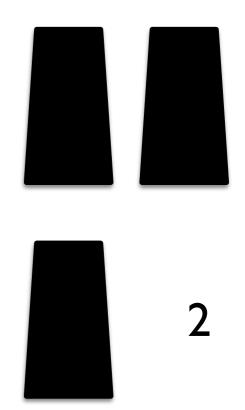
# Blinky as portal to intensional logics ...

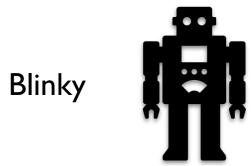


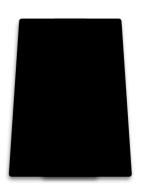


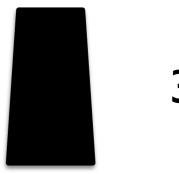


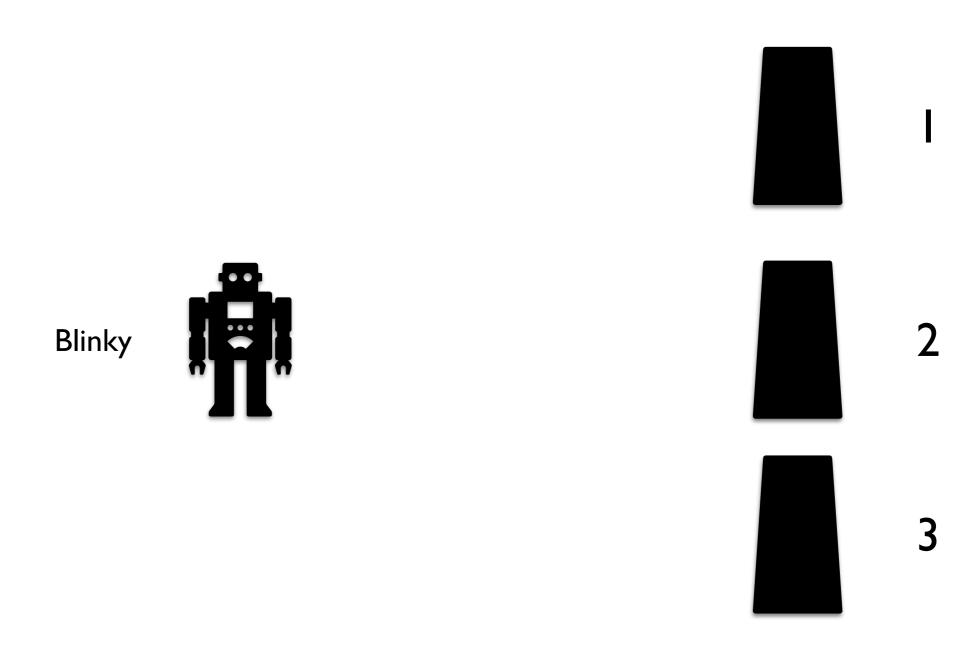




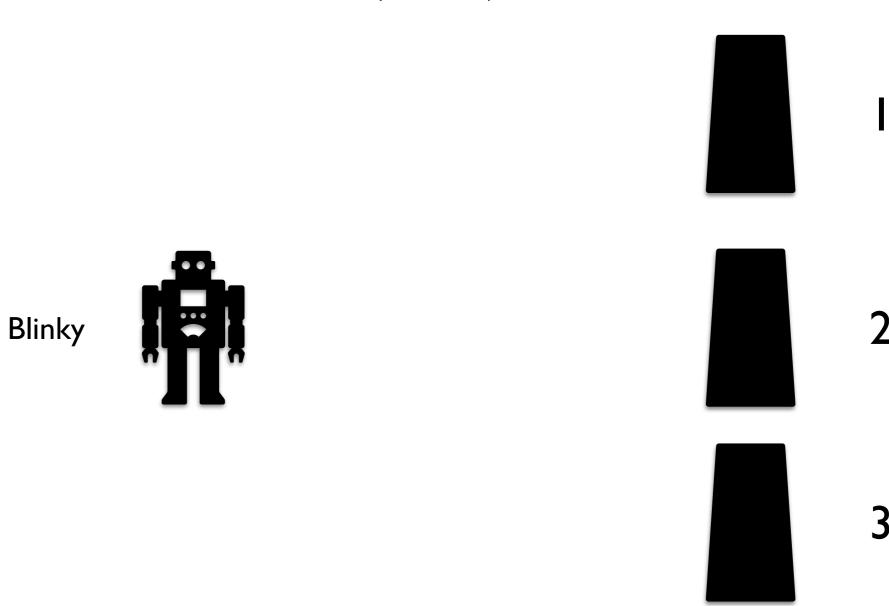


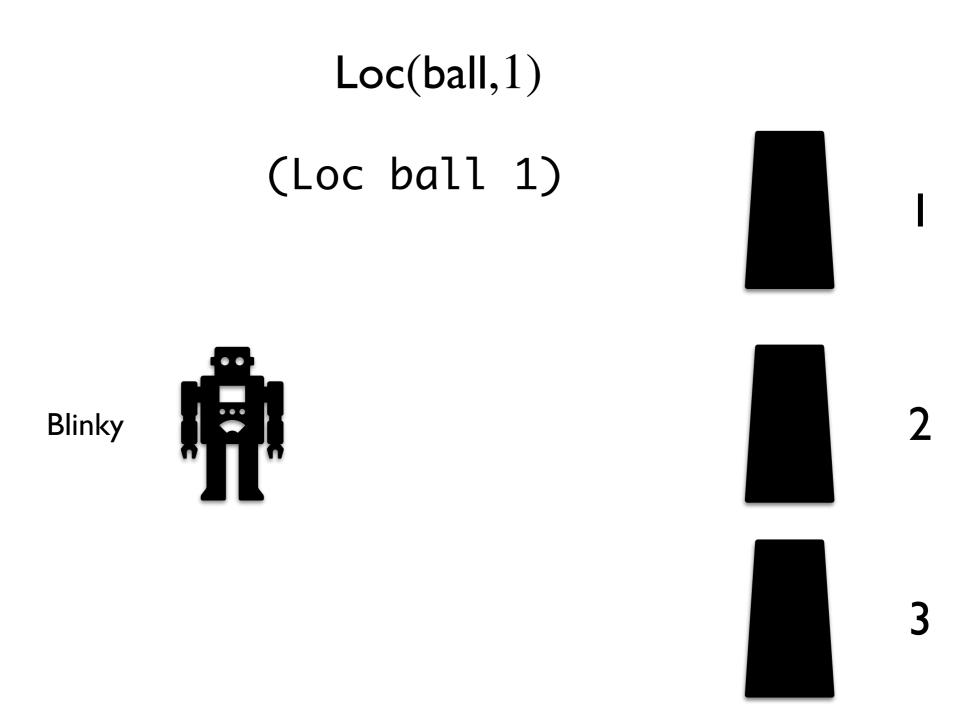


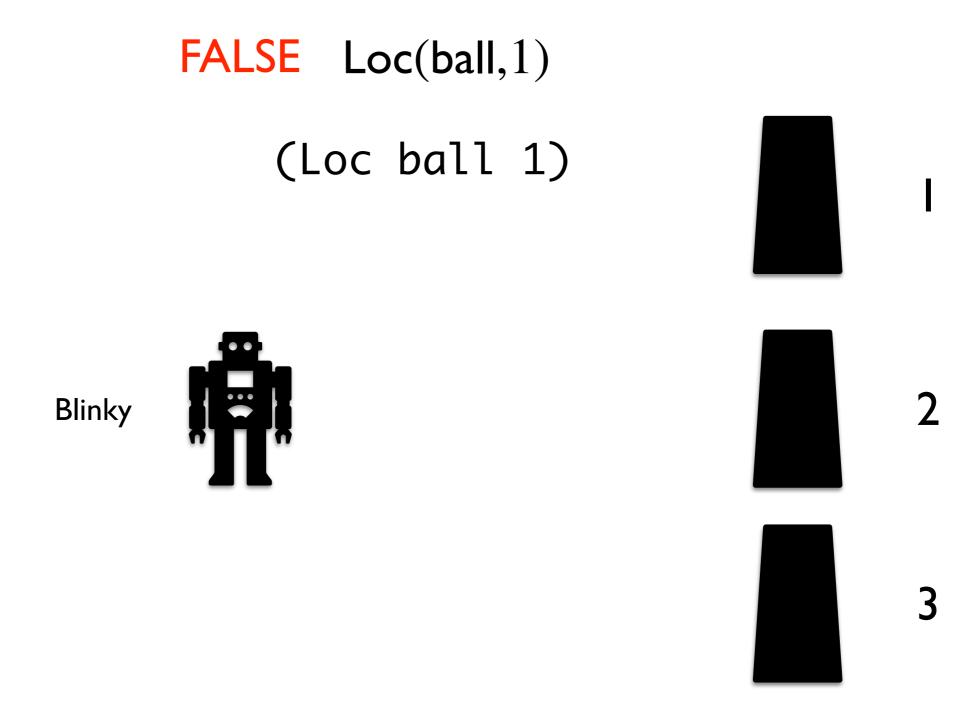




Loc(ball,1)







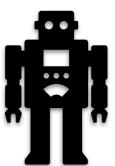
## FALSE Loc(ball,1) (Loc ball 1) 2 Blinky

#### **FALSE**

(Loc ball 1)



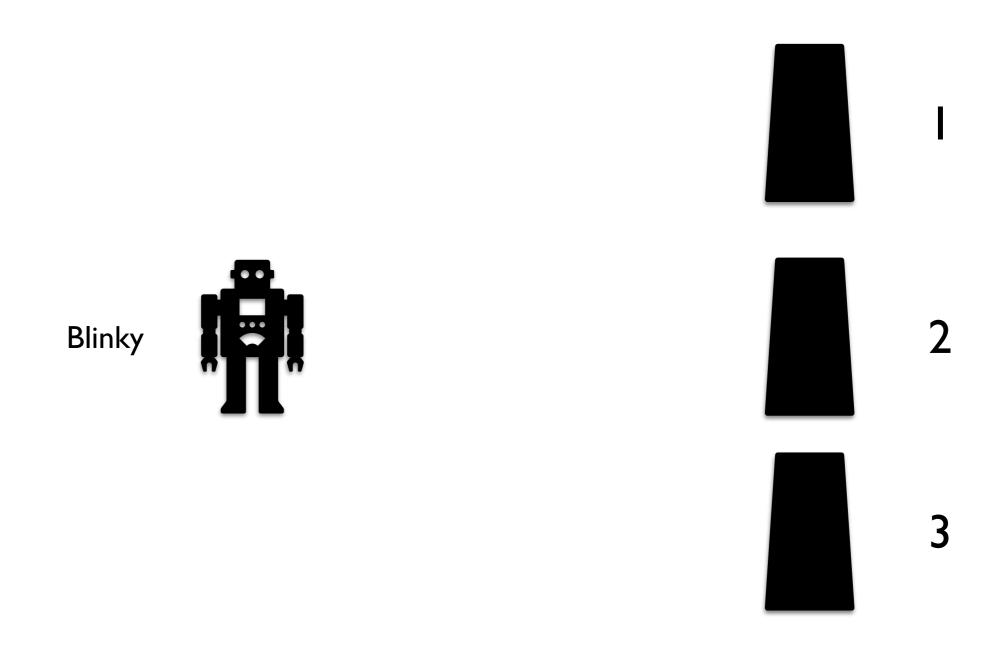




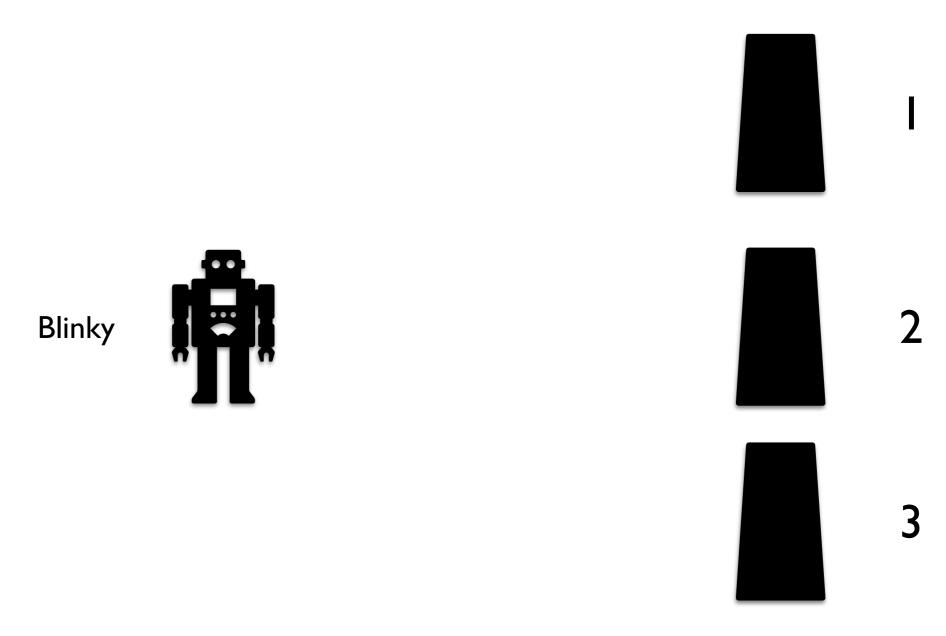
2



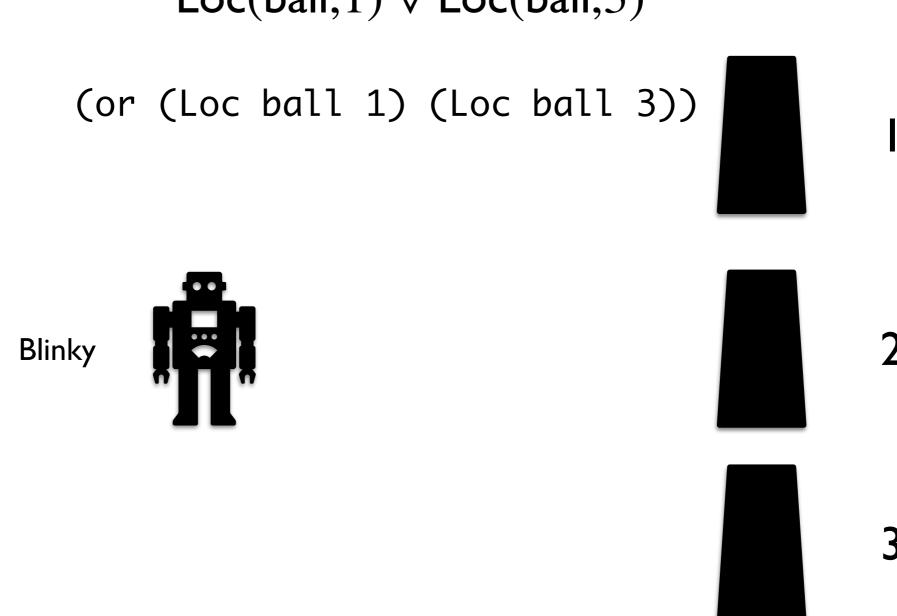
(Loc ball 1) Blinky



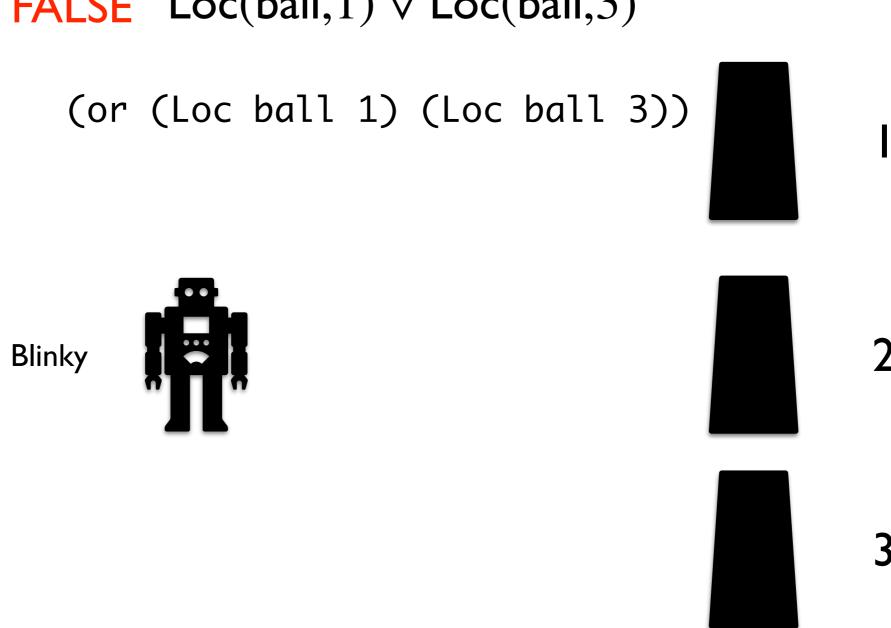
$$Loc(ball,1) \lor Loc(ball,3)$$





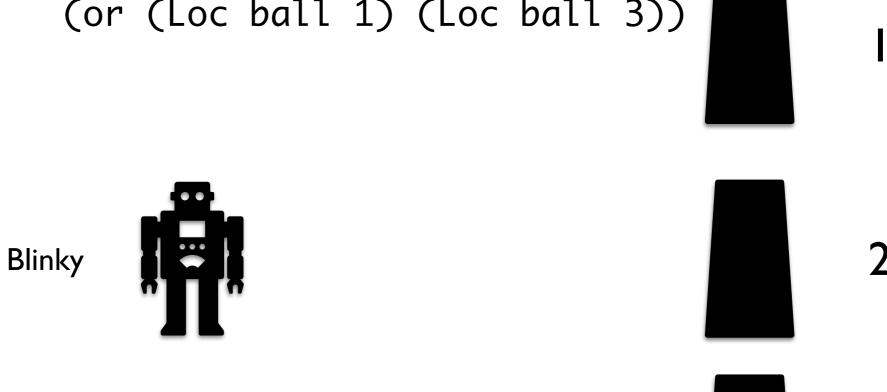




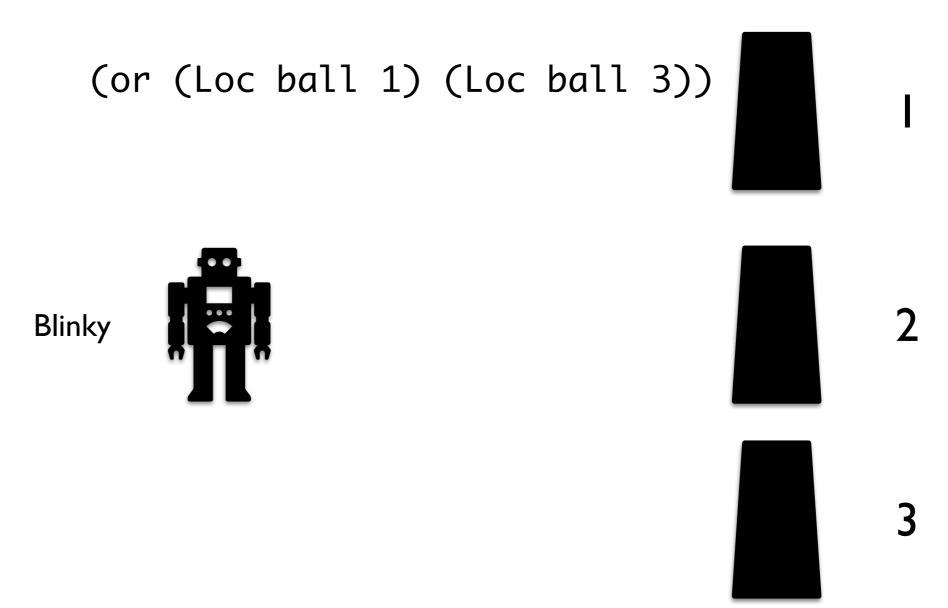


#### FALSE Loc(ball,1) \times Loc(ball,3)

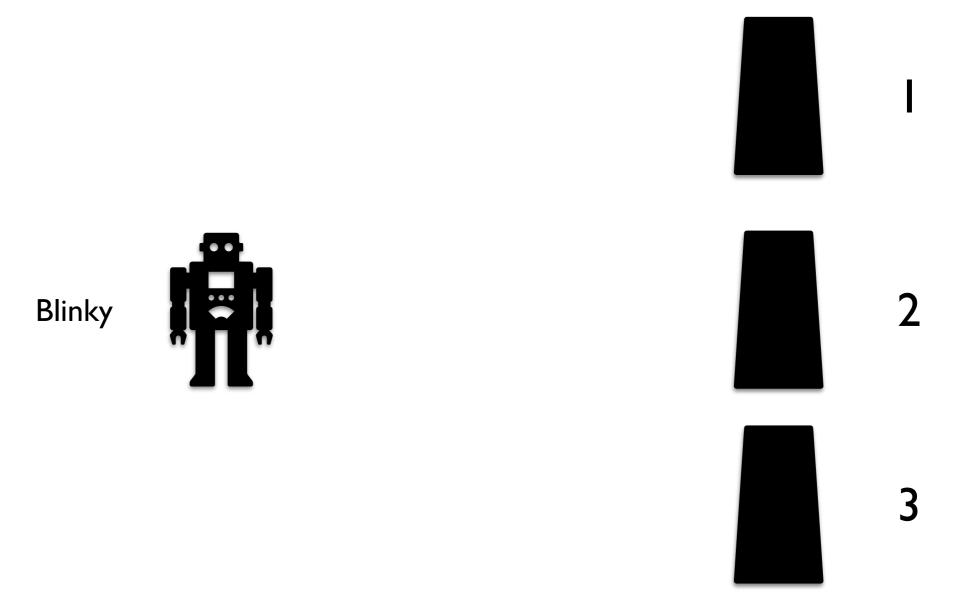
(or (Loc ball 1) (Loc ball 3))



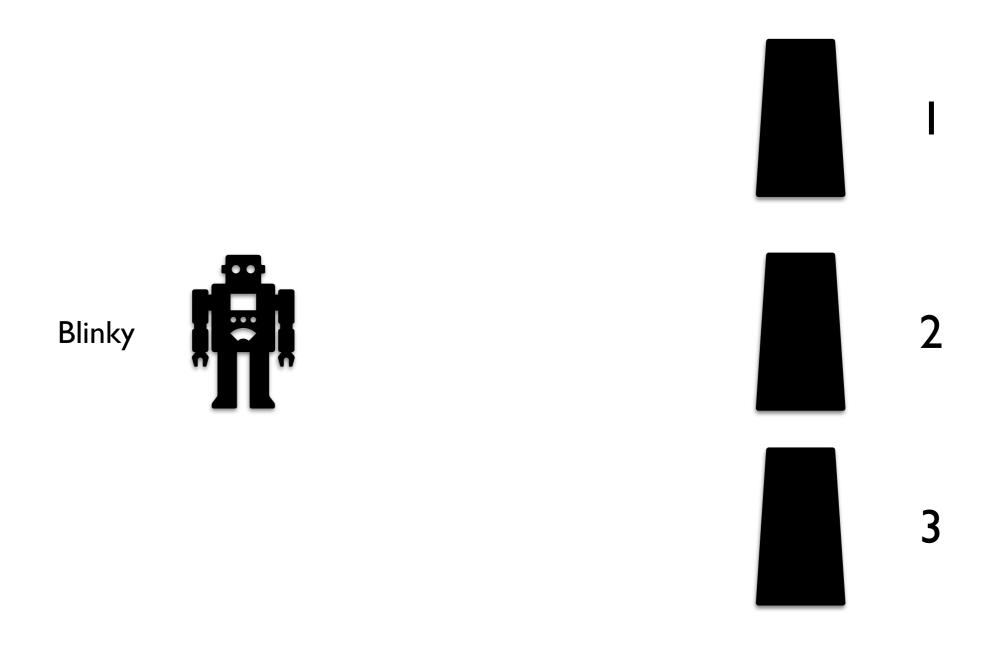
#### **FALSE**



#### **FALSE**

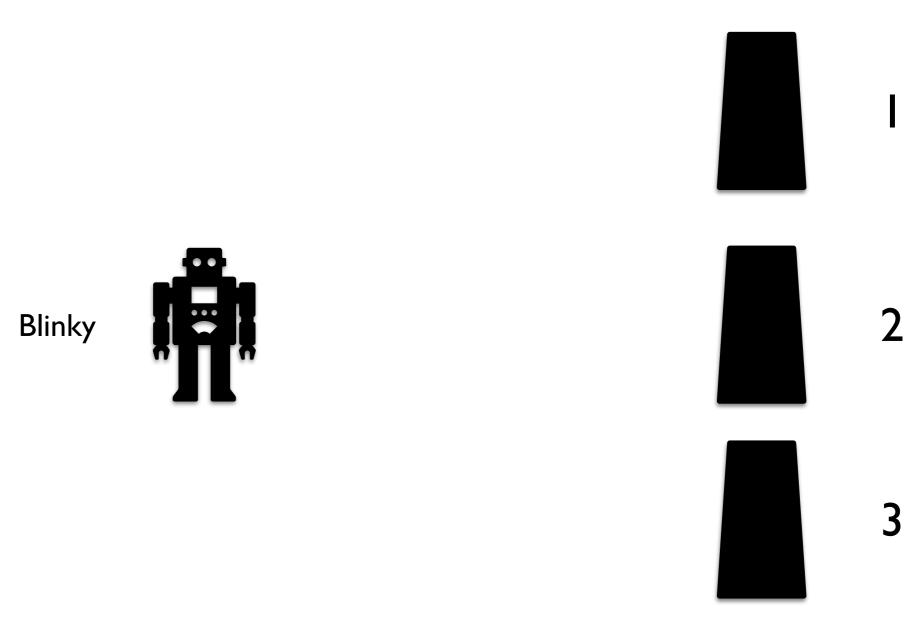


Blinky believes that the ball is in the cup at location #1.



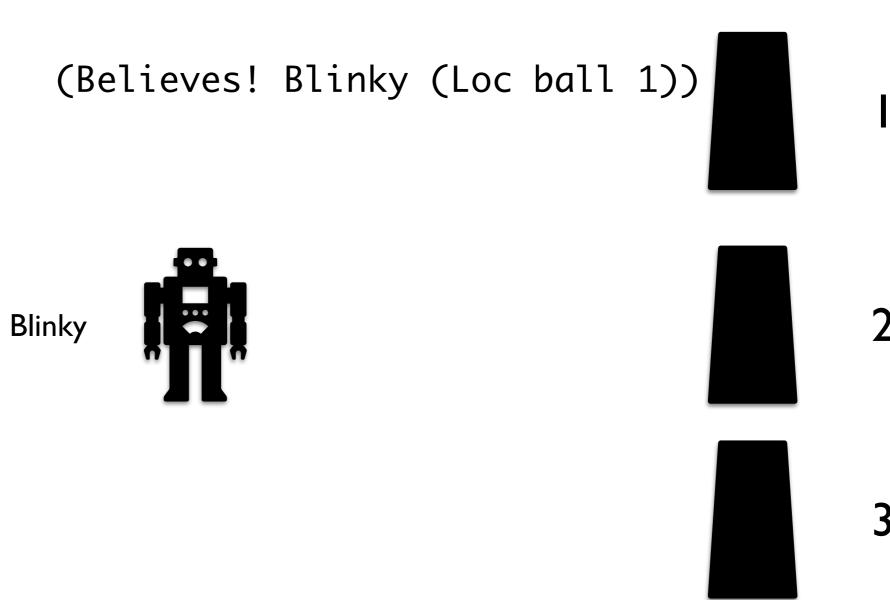
Blinky believes that the ball is in the cup at location #1.

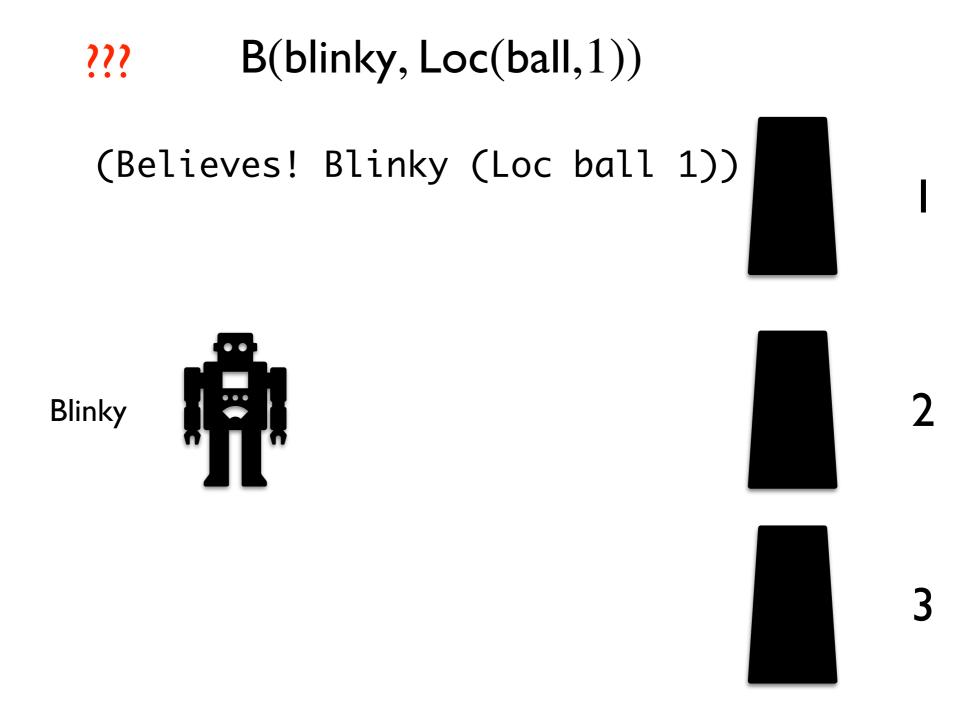
#### B(blinky, Loc(ball, 1))



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

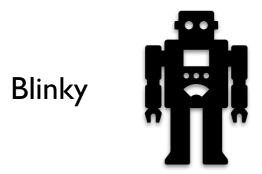
#### B(blinky, Loc(ball, 1))

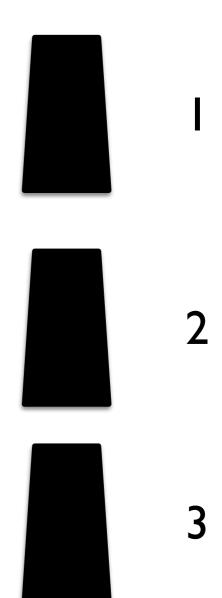




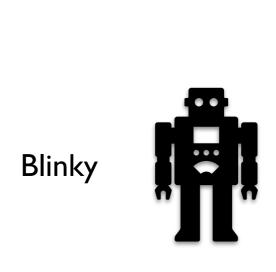


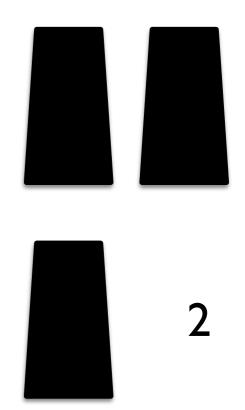
In extensional logics, what is denoted is conflated with meaning (the latter being naïvely compositional), but intensional attitudes like *believes*, *knows*, *hopes*, *fears*, etc cannot be represented and reasoned over smoothly (e.g. without fear of inconsistency rising up).



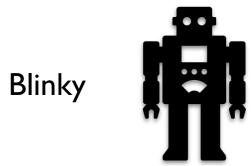


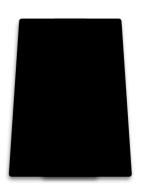


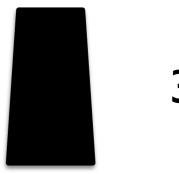


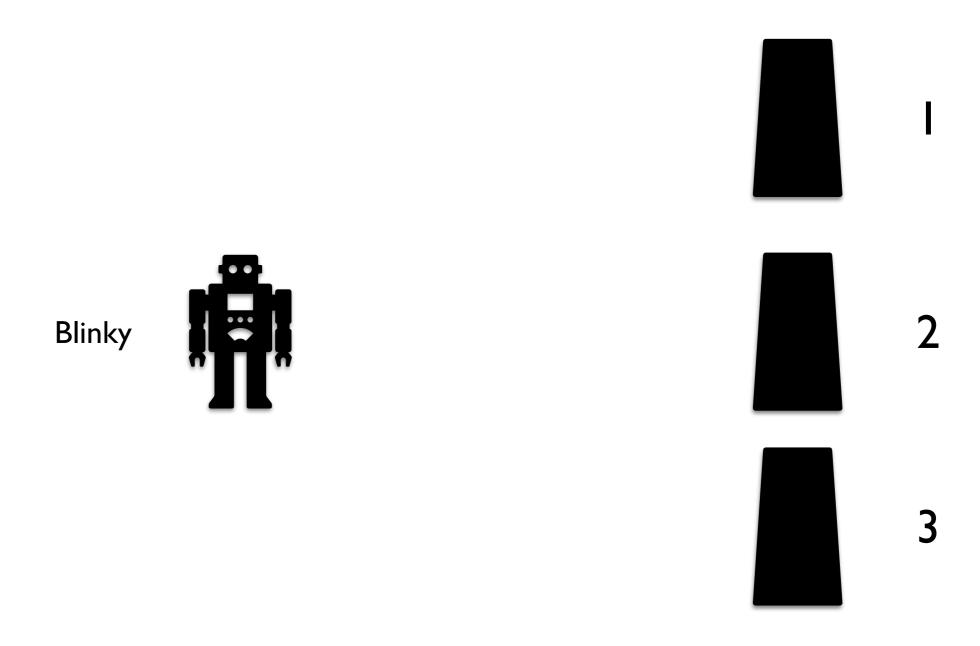


3

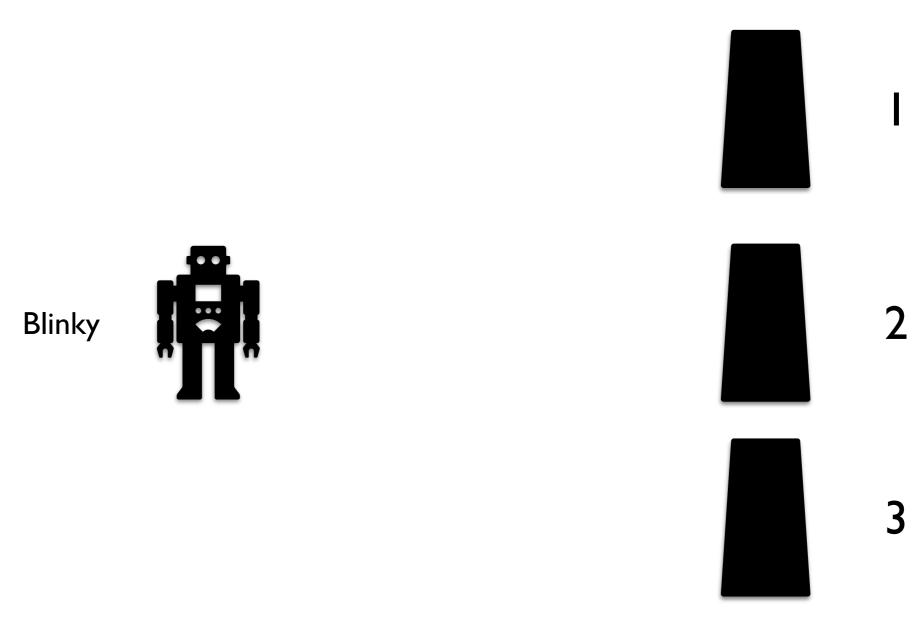






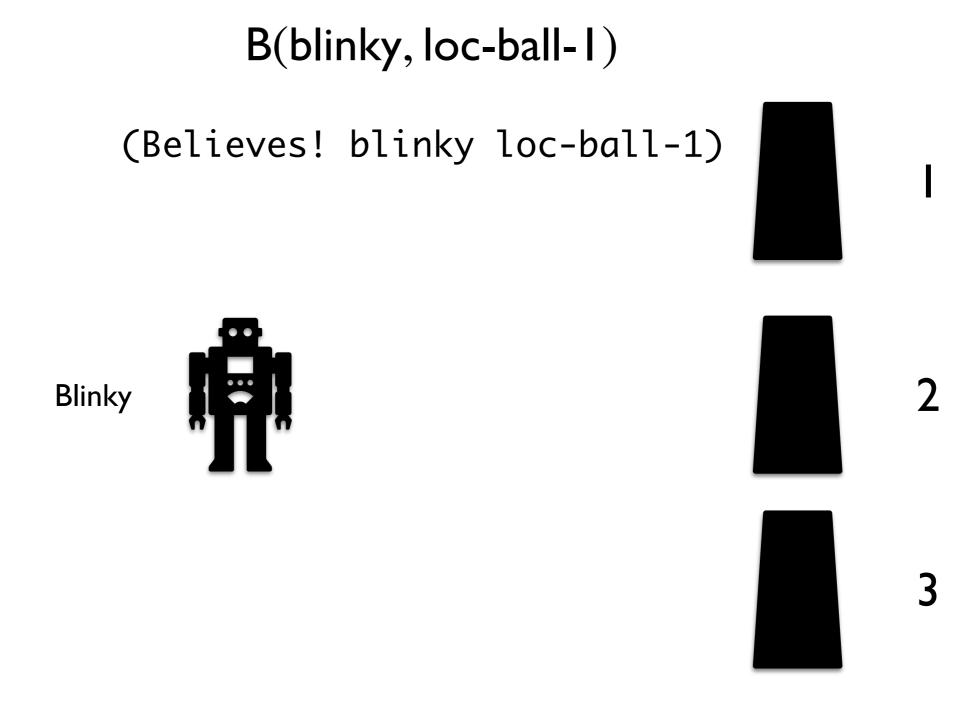


#### B(blinky, loc-ball-1)

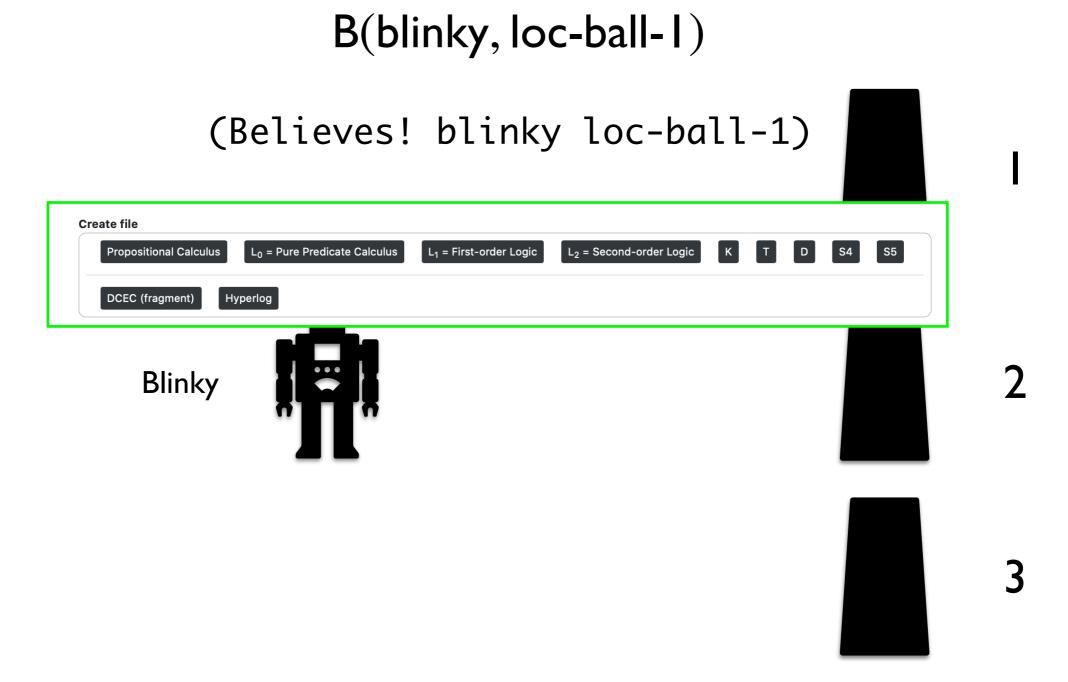




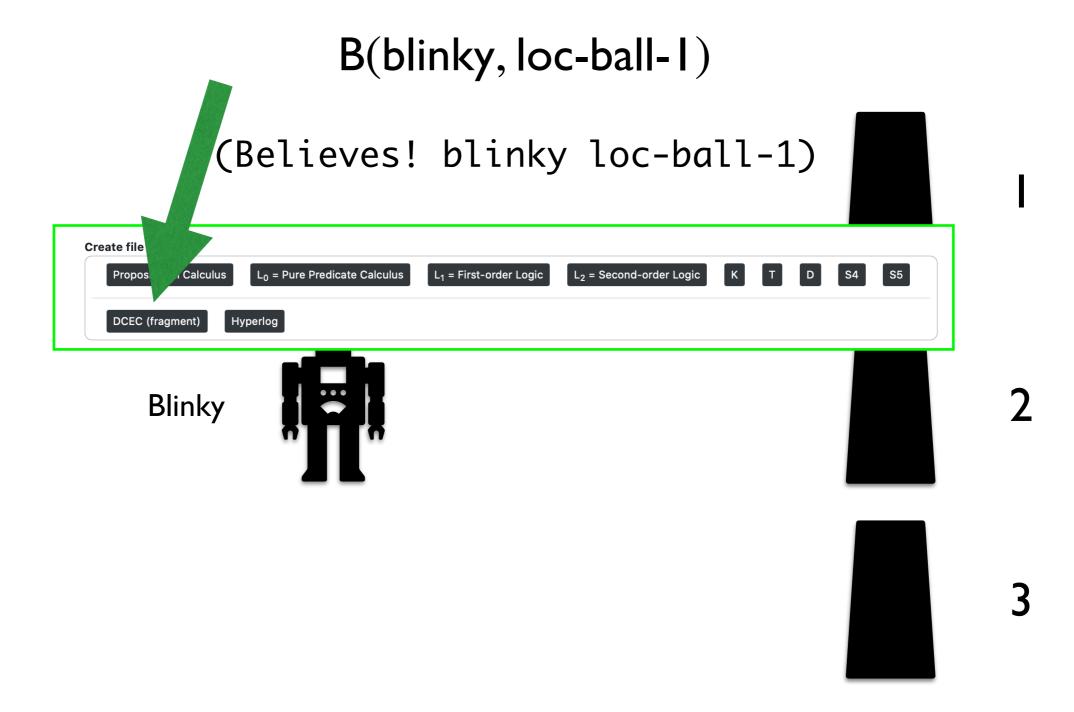
(Believes! blinky loc-ball-1)



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



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



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

# False Belief Task Demands Intensional Logic ...

#### False Belief Task Demands Intensional Logic ...







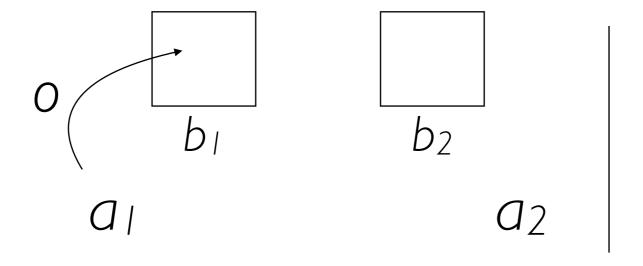
## False Belief Task Demands Intensional Logic ...

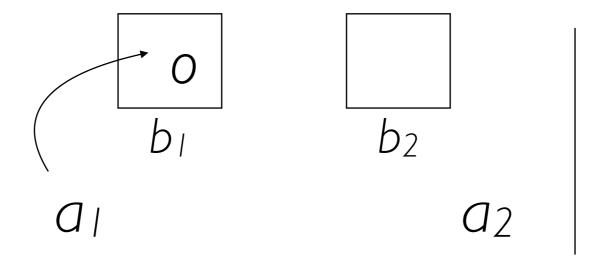


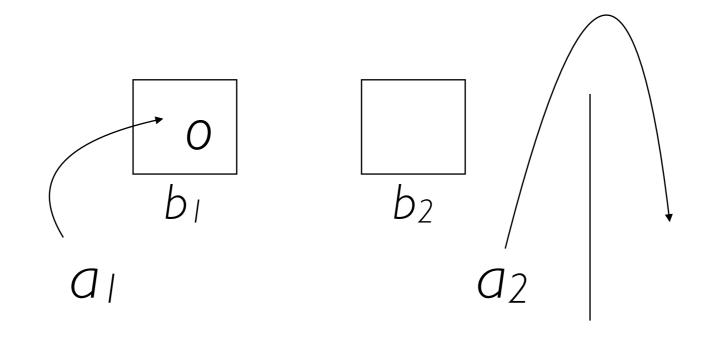
#### Framework for FBT<sup>0</sup><sub>1</sub>

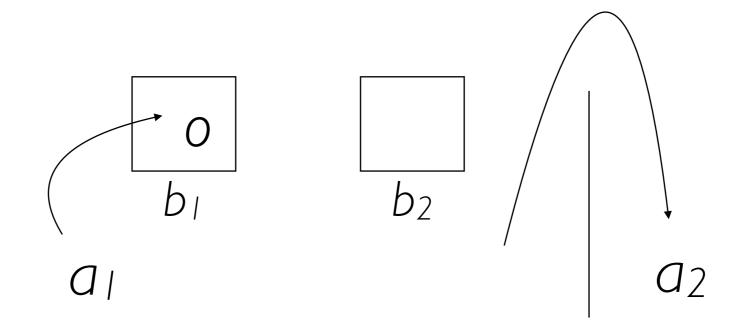
a

е

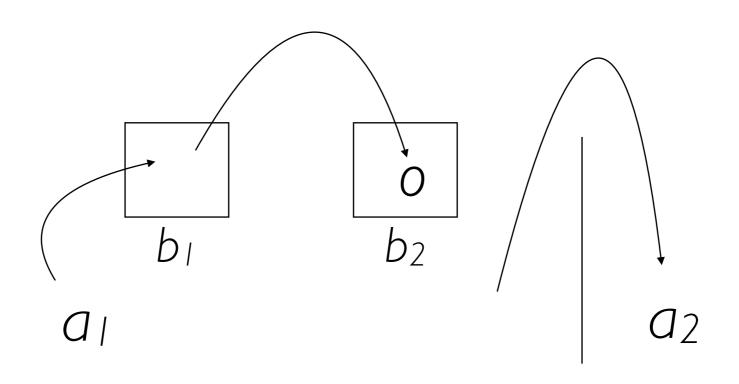






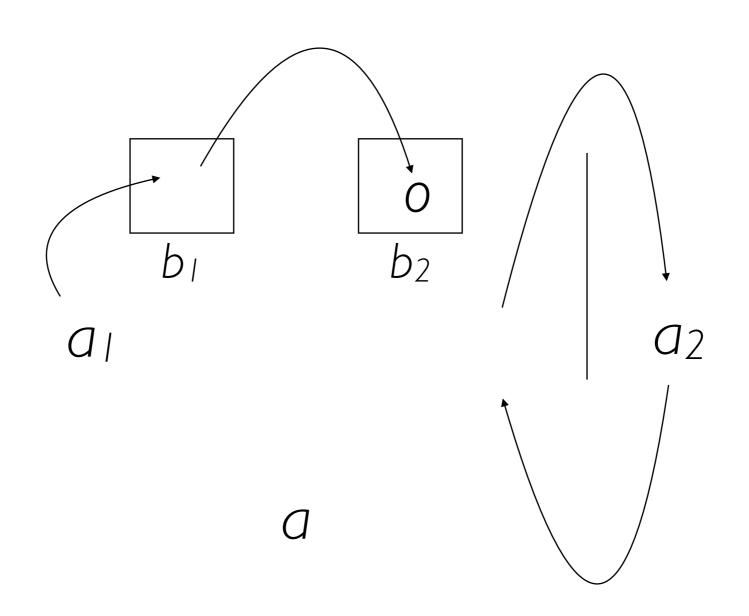


 $a_1$   $b_2$   $a_2$ 



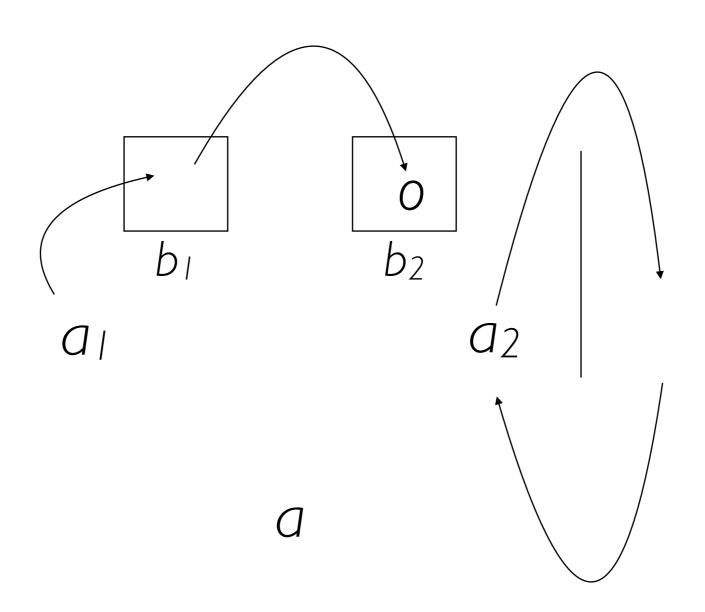
#### Framework for FBT<sup>0</sup><sub>1</sub>

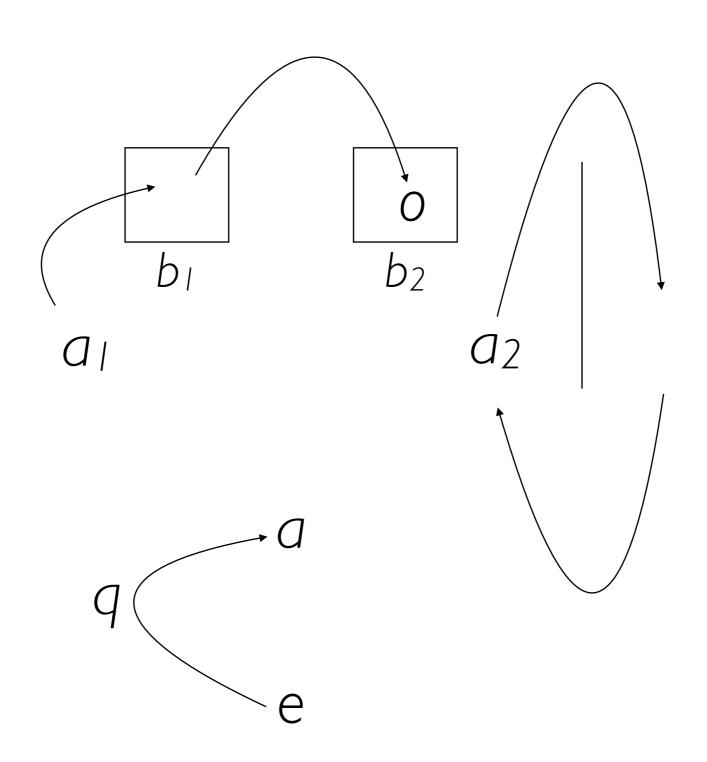
(five timepoints)



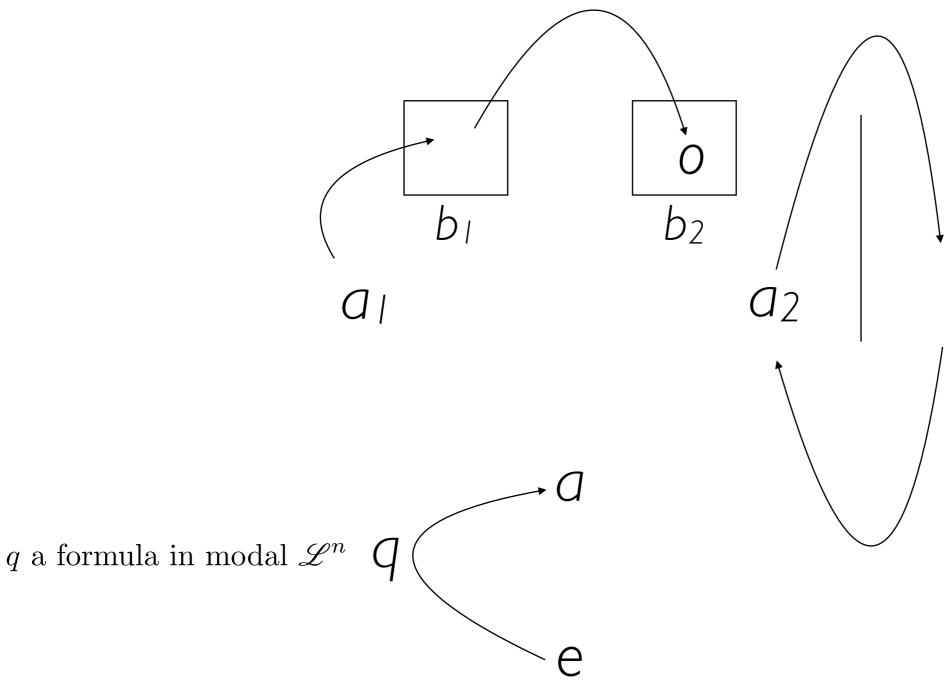
#### Framework for FBT<sup>0</sup><sub>1</sub>

(five timepoints)





(five timepoints)



#### Framework for FBT<sup>1</sup>

$$O_{m}$$

$$\begin{bmatrix} b_{1} & b_{2} \\ a_{1} & a_{2} \end{bmatrix}$$

 $\mathcal{Q}$ 

e

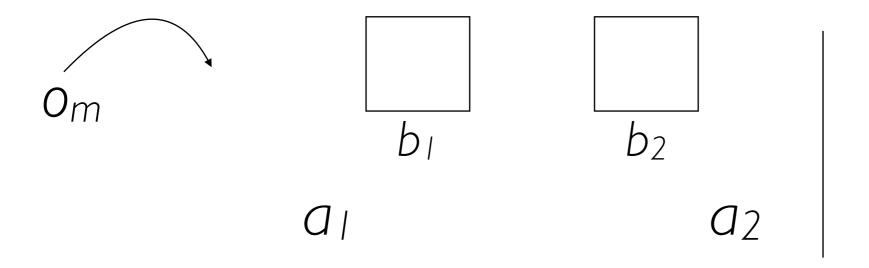
### Framework for FBT<sub>1</sub> (six timepoints)

a

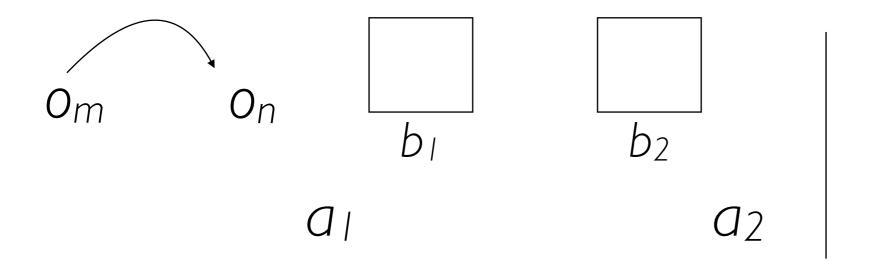
e

#### Framework for FBT<sub>1</sub>

(six timepoints)

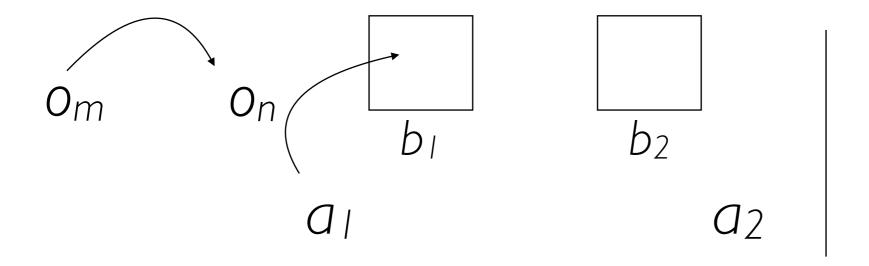


(six timepoints)

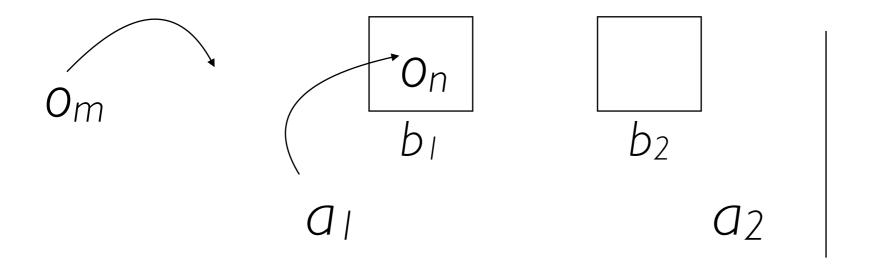


 $\mathcal{Q}$ 

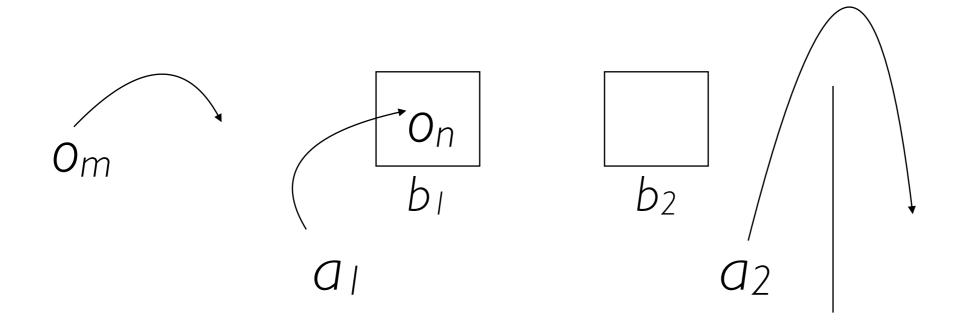
### Framework for FBT1

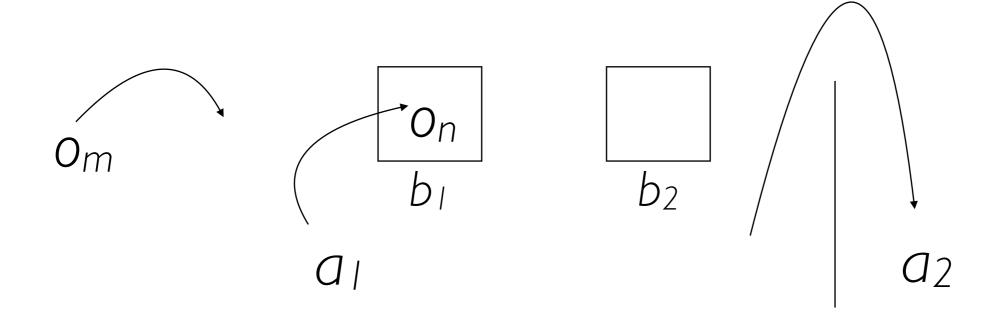


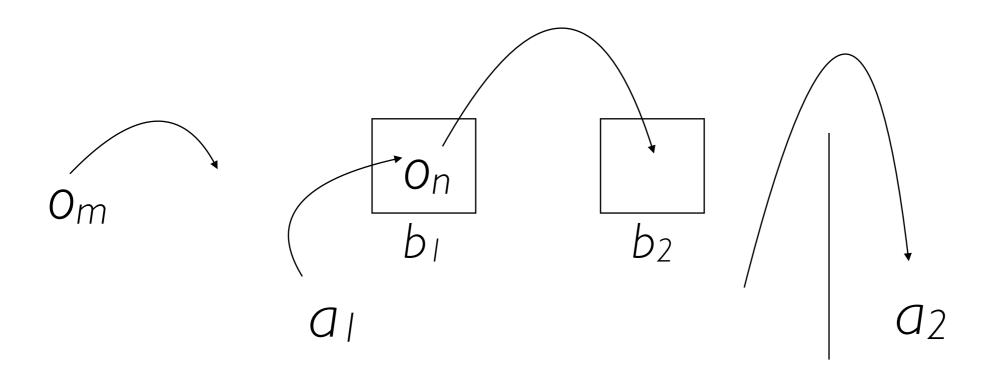
(six timepoints)



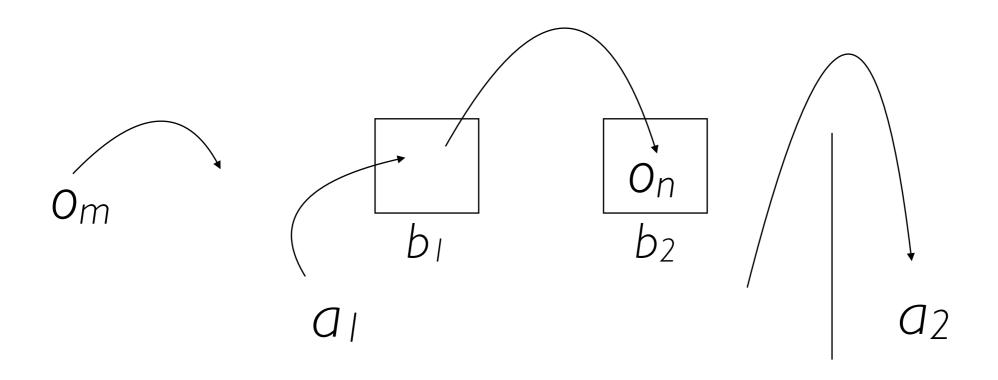
(six timepoints)

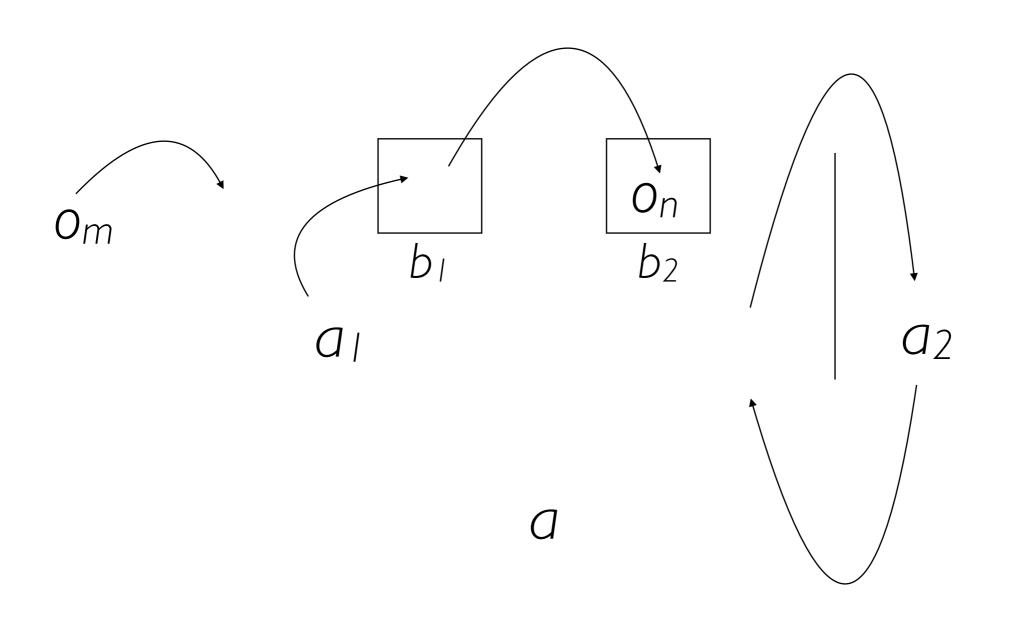


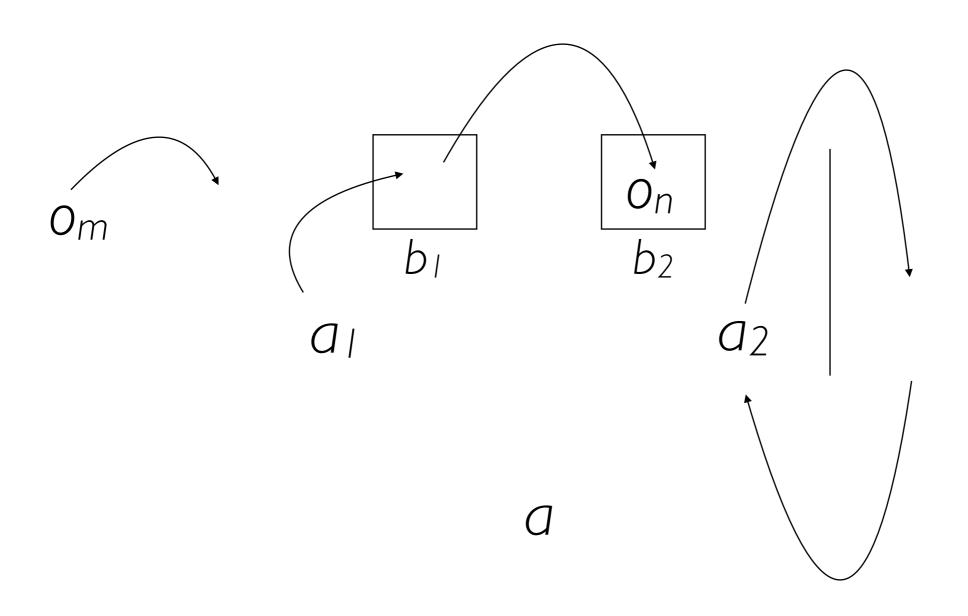


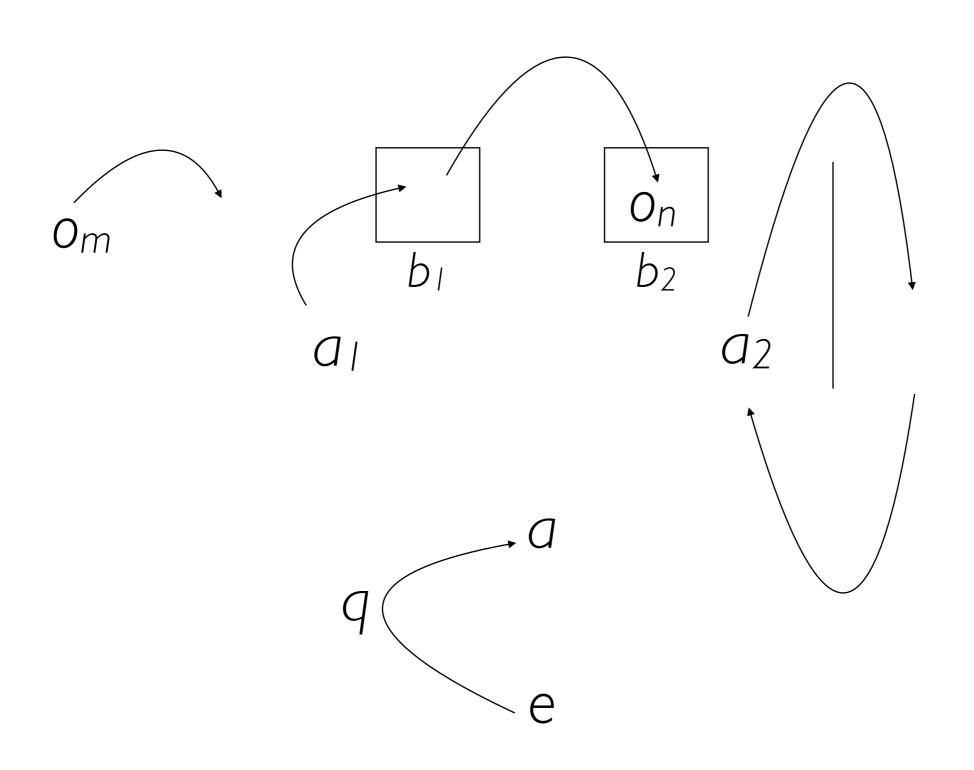


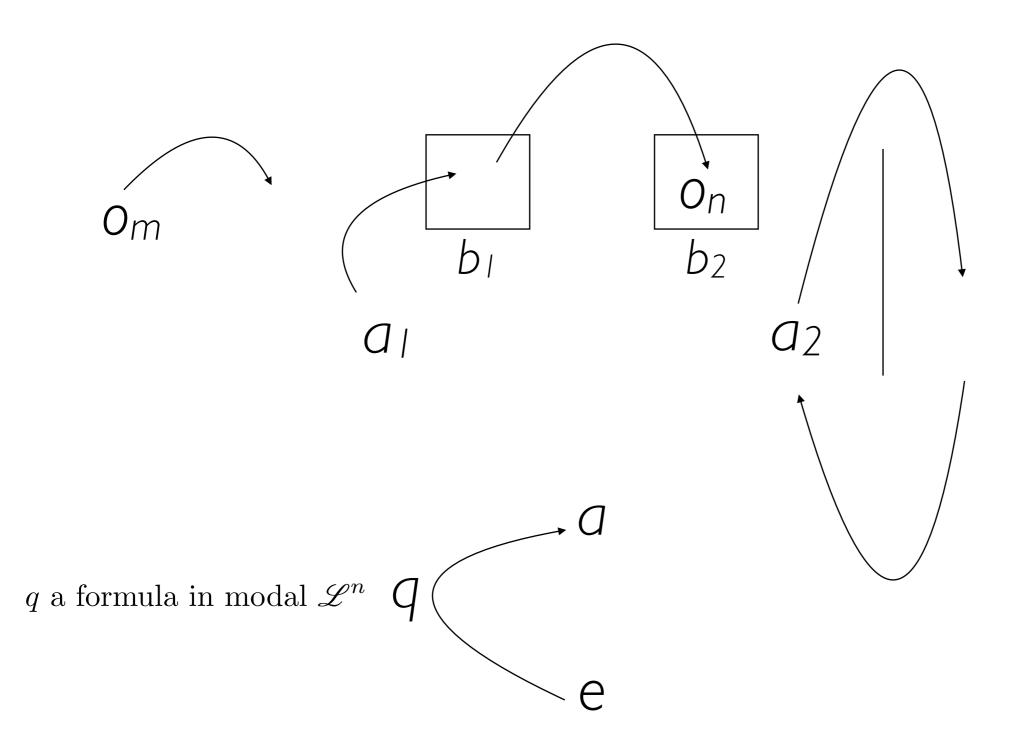
(six timepoints)











# 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, a a will carry out a a whose that a a a a a knows that a a a fears that a a a a discovers that a 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

<sup>&</sup>lt;sup>1</sup> 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<sup>1</sup><sub>2</sub>

 $\mathcal{Q}$ 

e

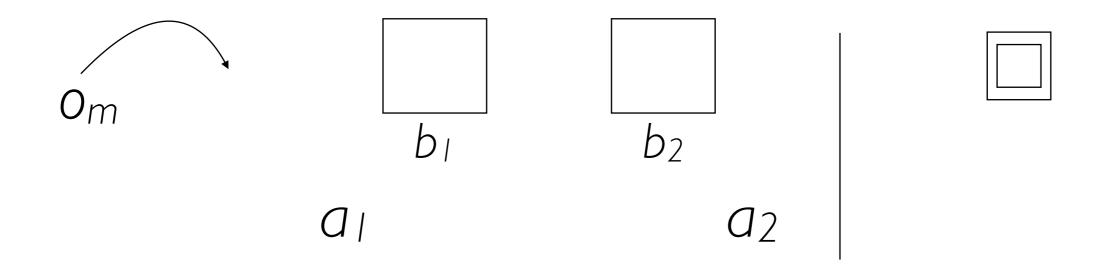
# Framework for FBT<sup>1</sup><sub>2</sub> (seven timepoints)

 $O_{m}$   $\begin{bmatrix} b_{1} & b_{2} \\ a_{1} & a_{2} \end{bmatrix}$ 

a

9

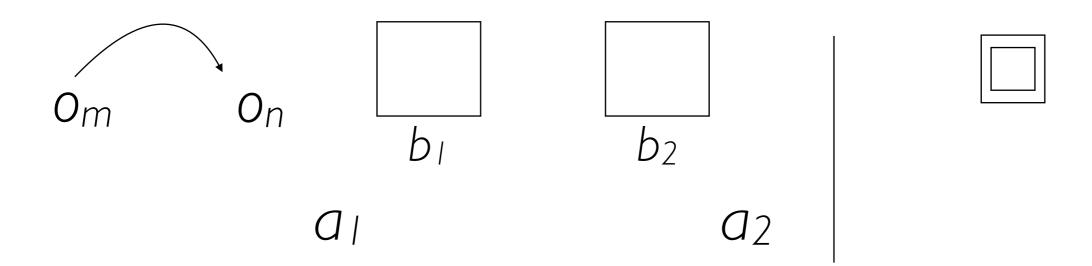
(seven timepoints)



 $\mathcal{Q}$ 

e

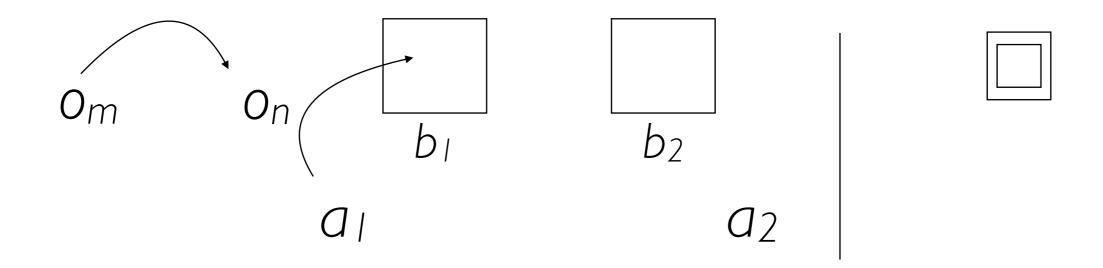
# Framework for FBT<sub>2</sub> (seven timepoints)



a

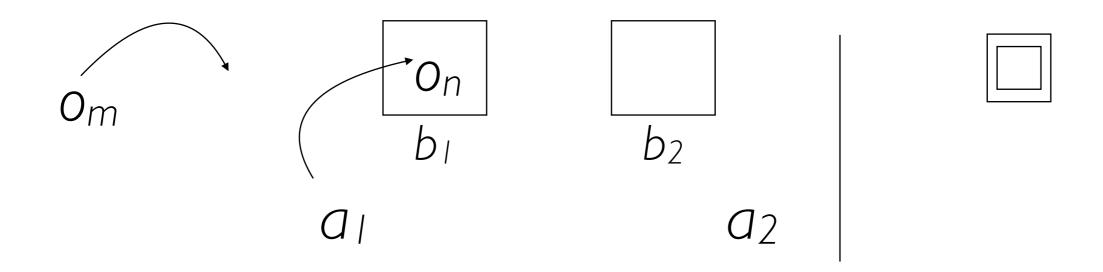
e

(seven timepoints)

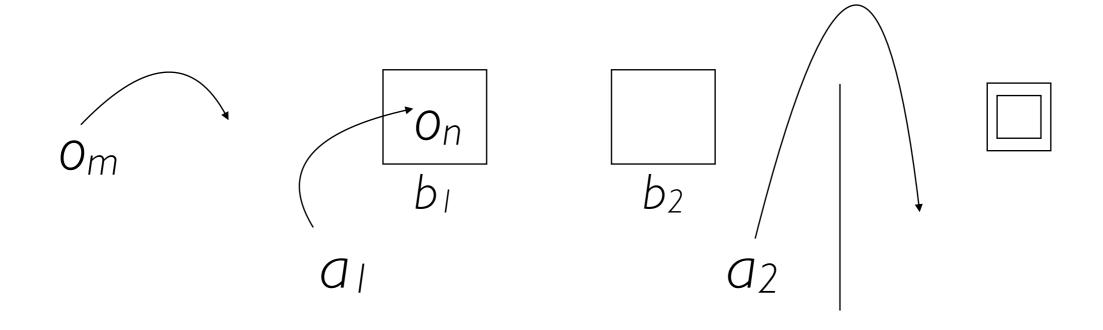


 $\mathcal{Q}$ 

(seven timepoints)

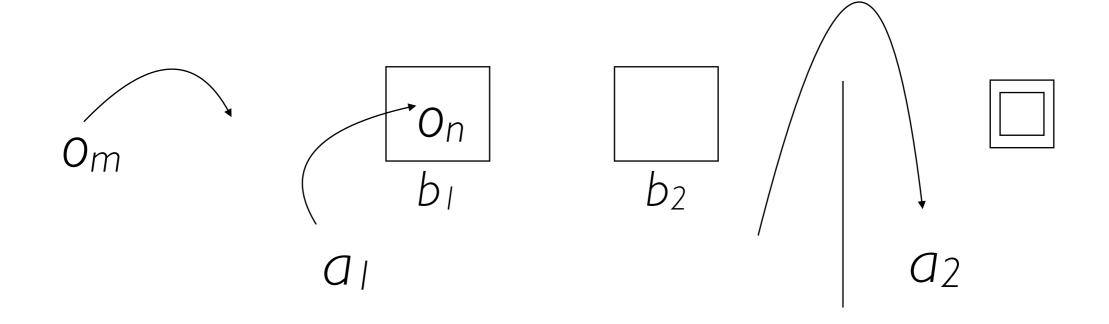


(seven timepoints)



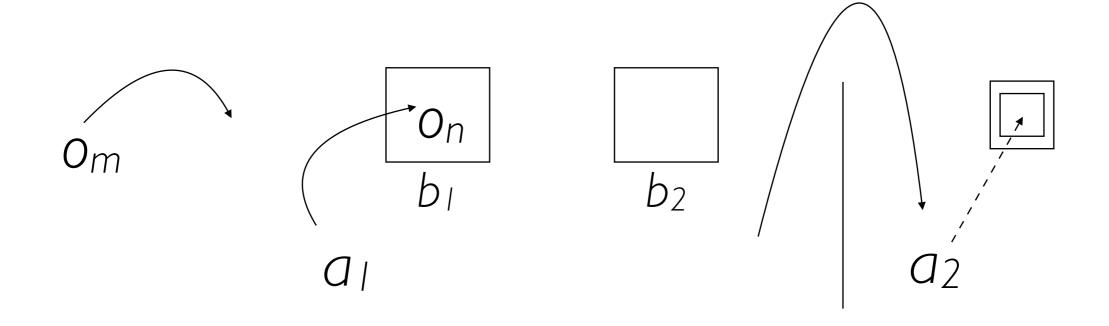
 $\mathcal{Q}$ 

(seven timepoints)

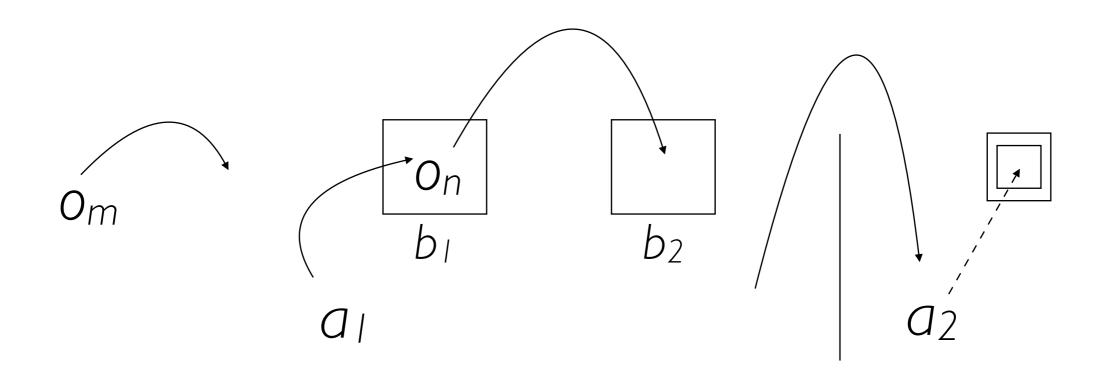


 $\mathcal{Q}$ 

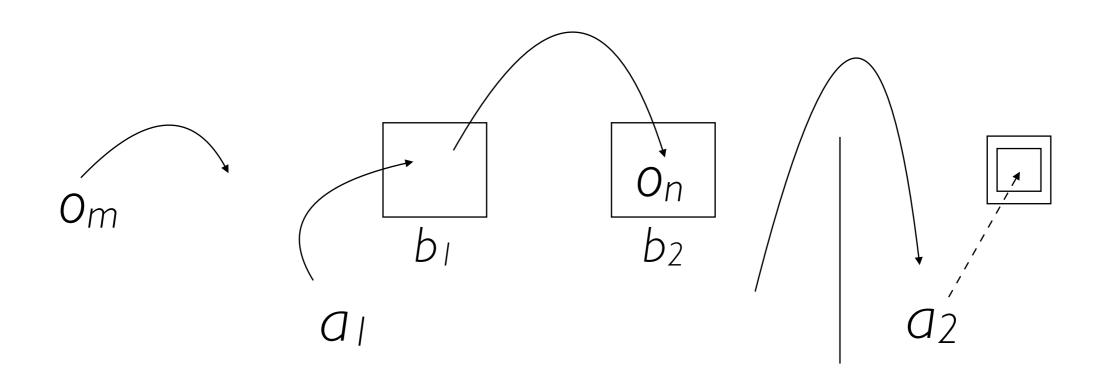
(seven timepoints)



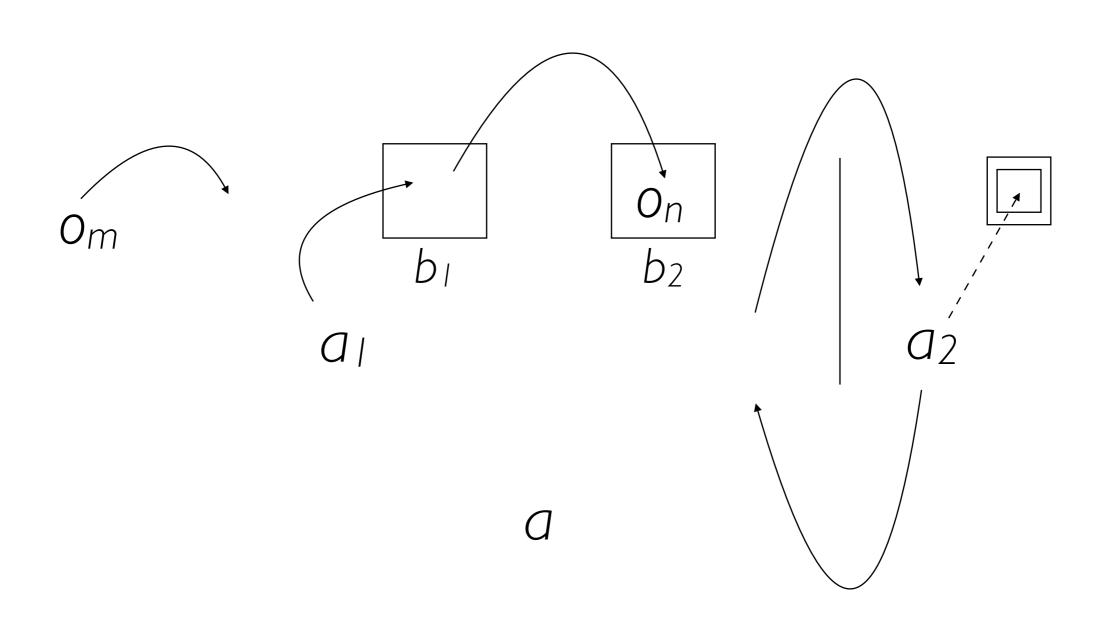
(seven timepoints)

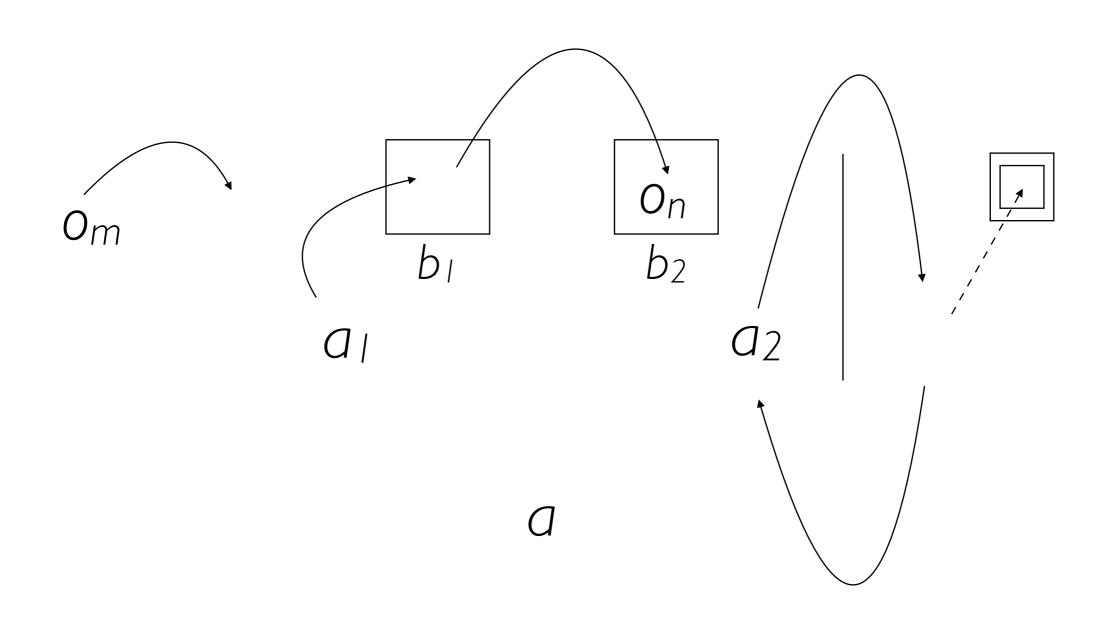


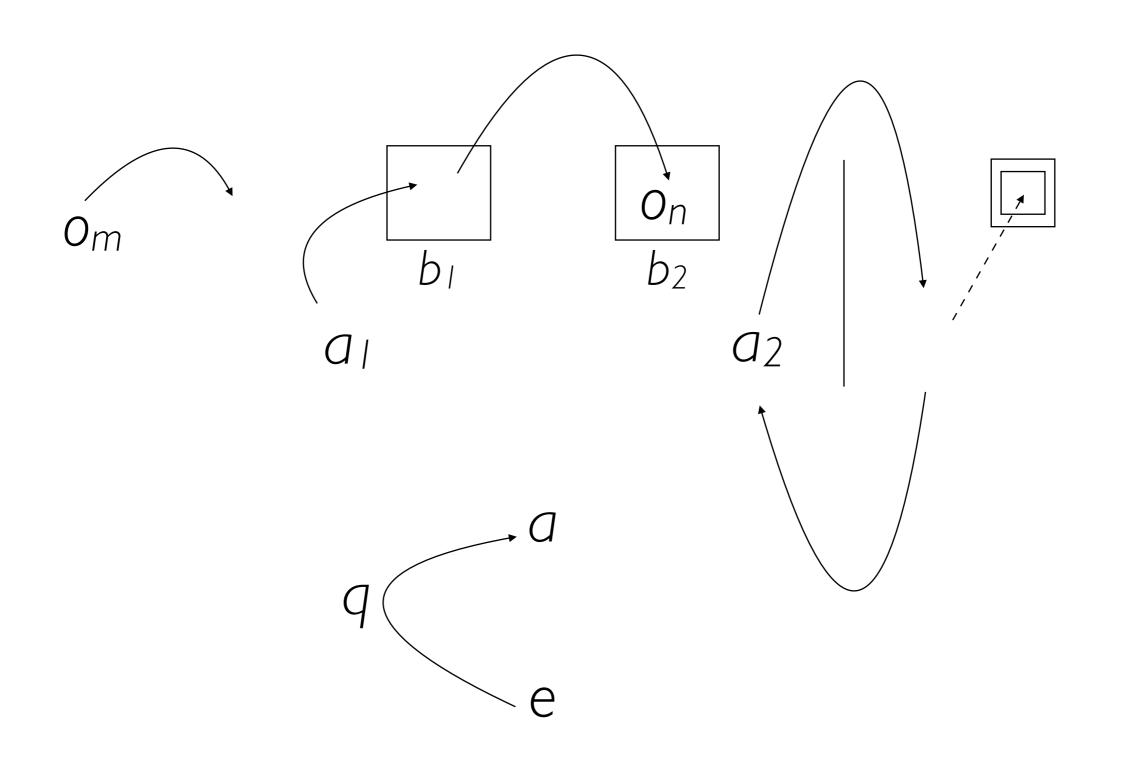
(seven timepoints)

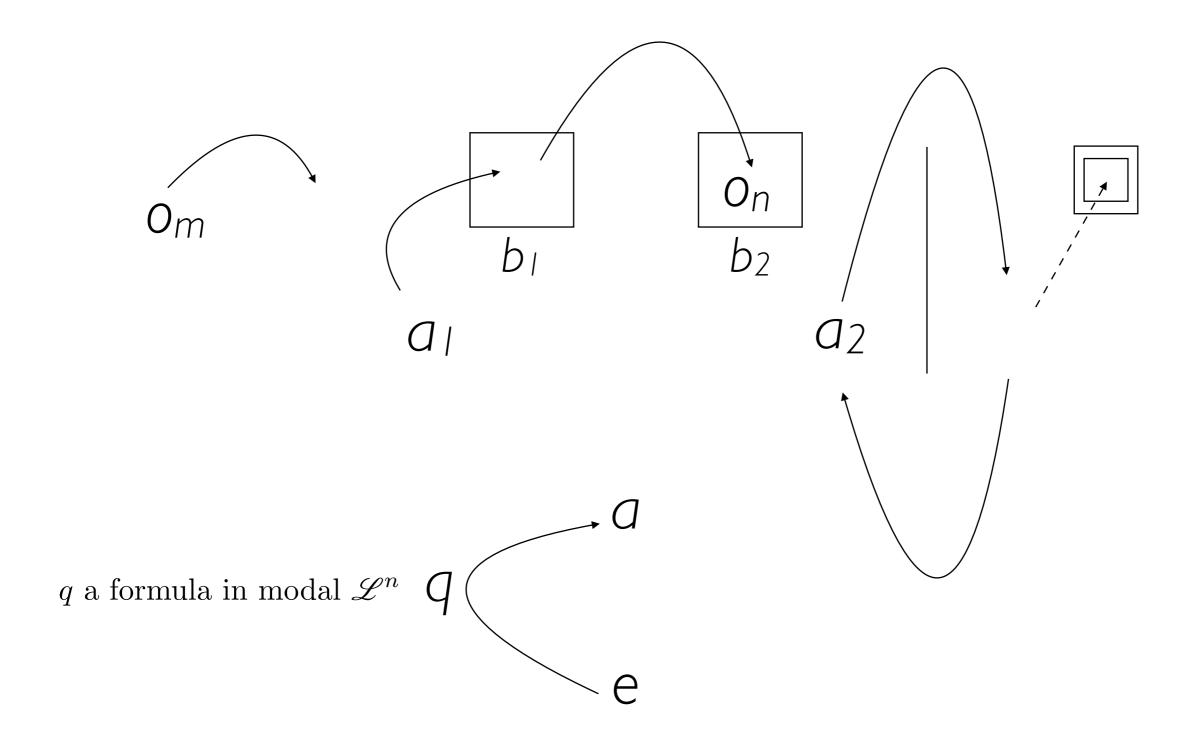


 $\mathcal{Q}$ 









$$Om$$

$$b_1$$

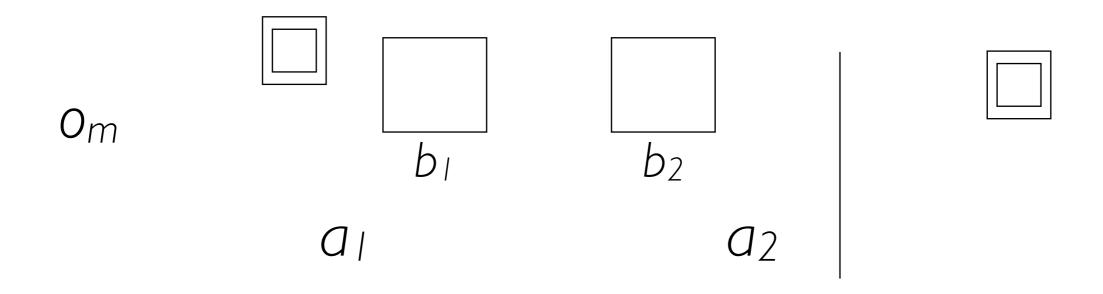
$$a_1$$

$$a_2$$

 $\mathcal{Q}$ 

9

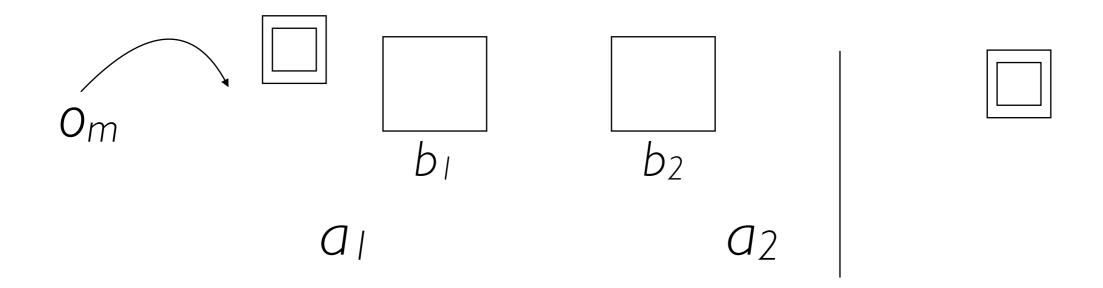
# Framework for FBT<sup>1</sup><sub>3</sub> (eight timepoints)



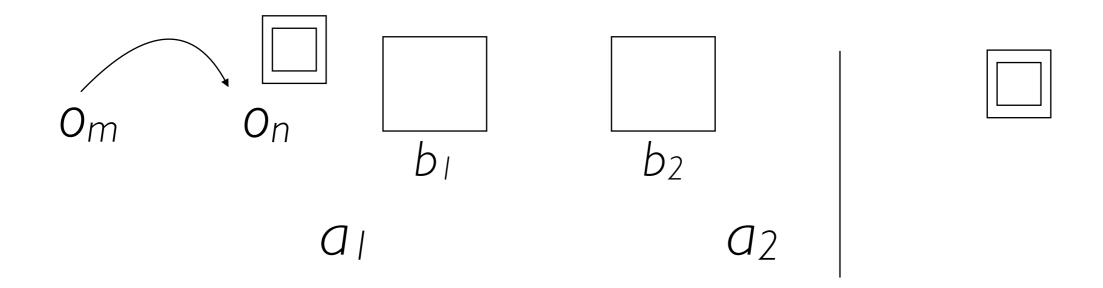
 $\mathcal{Q}$ 

e

# Framework for FBT<sup>1</sup>3 (eight timepoints)



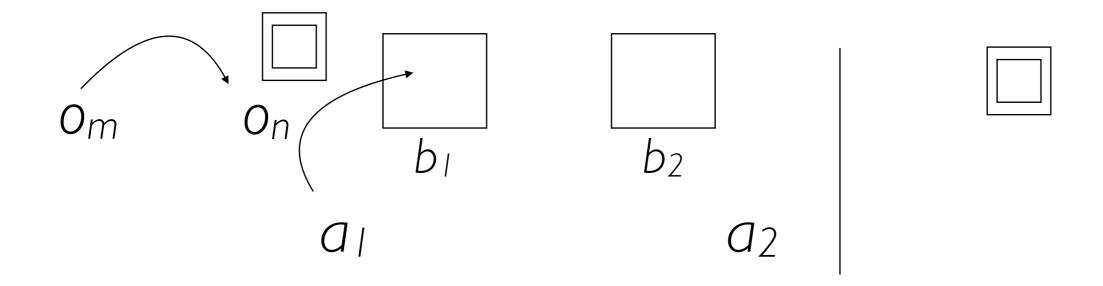
# Framework for FBT<sup>1</sup>3 (eight timepoints)



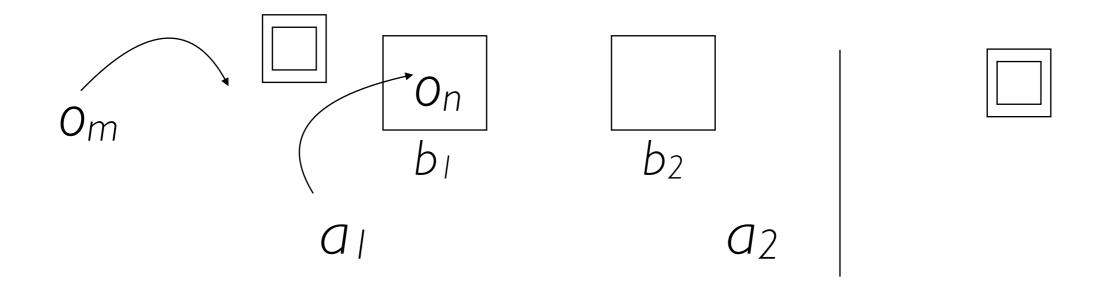
a

e

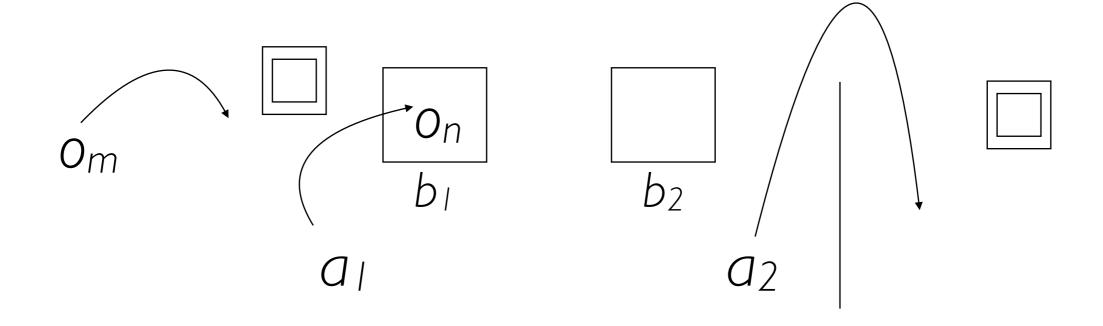
(eight timepoints)



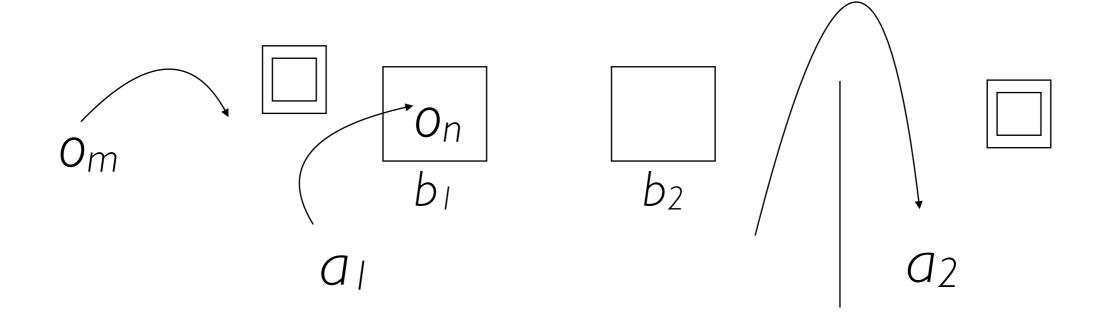
(eight timepoints)



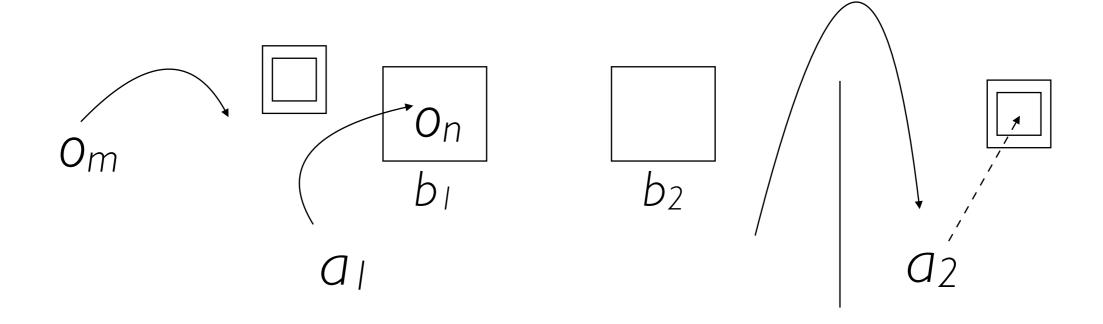
(eight timepoints)



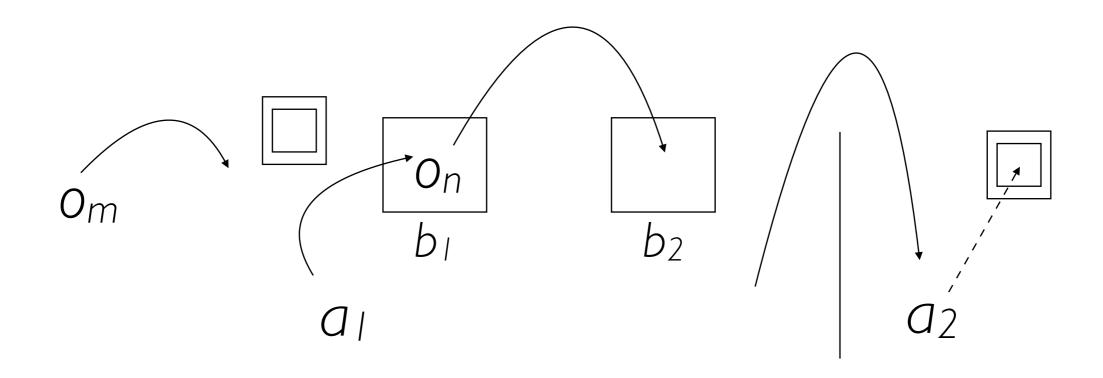
(eight timepoints)



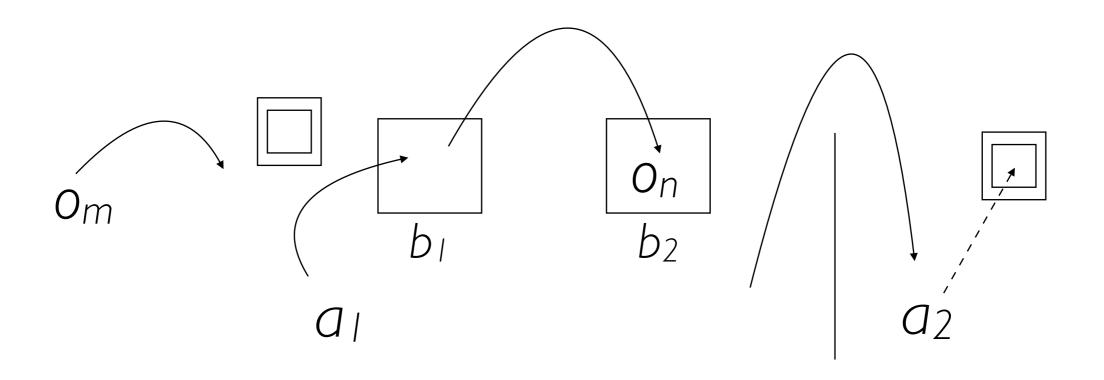
(eight timepoints)



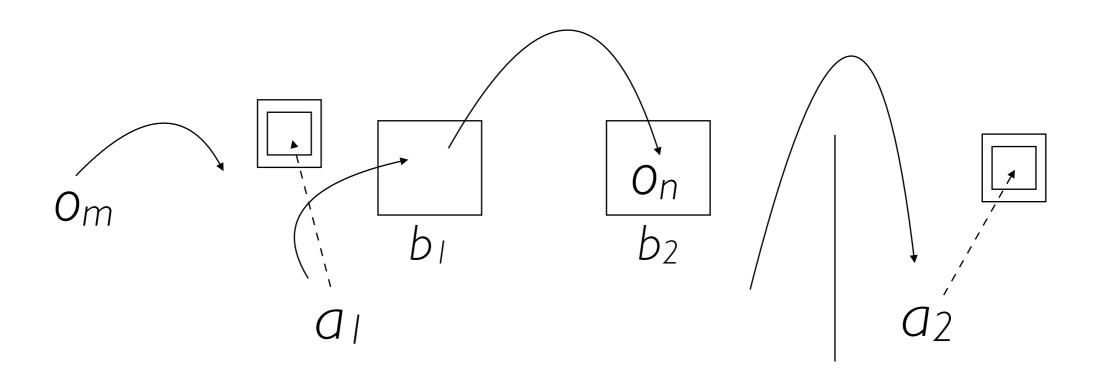
(eight timepoints)



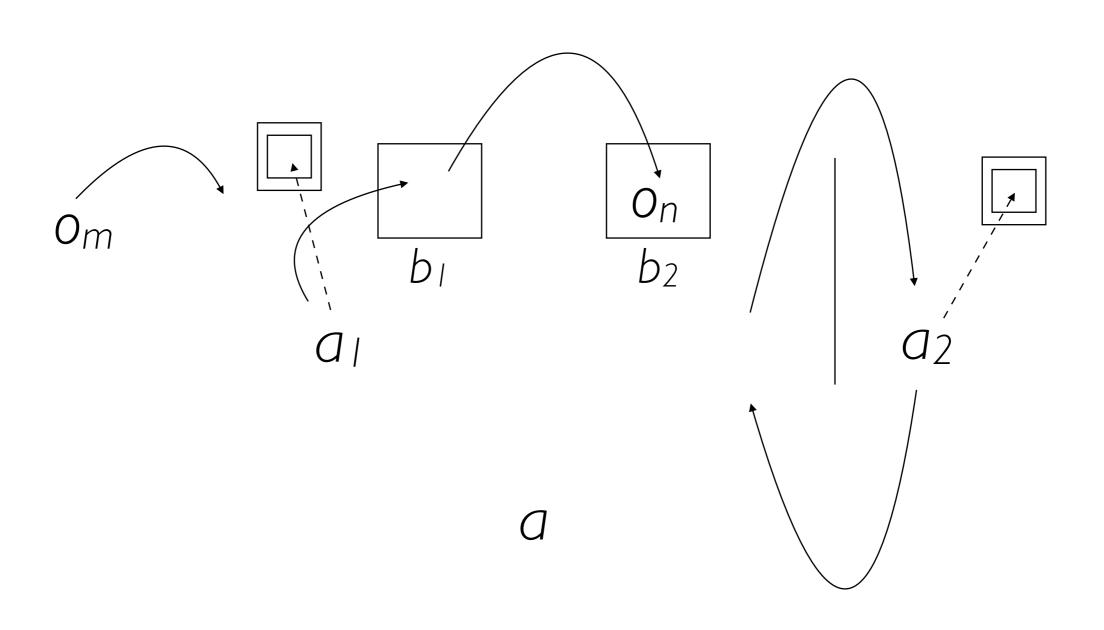
(eight timepoints)

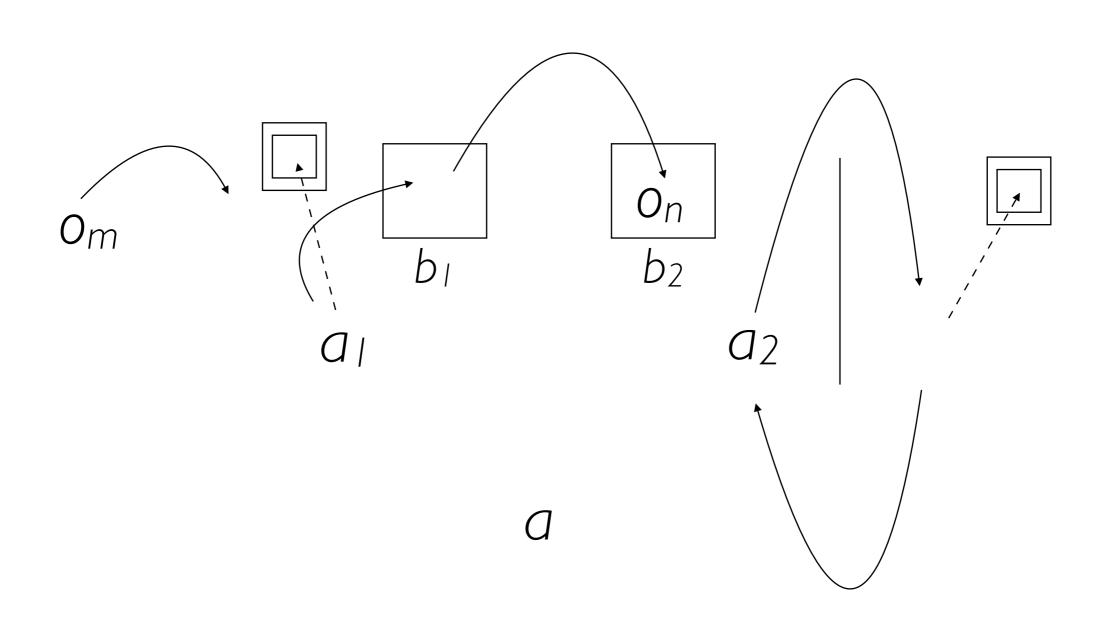


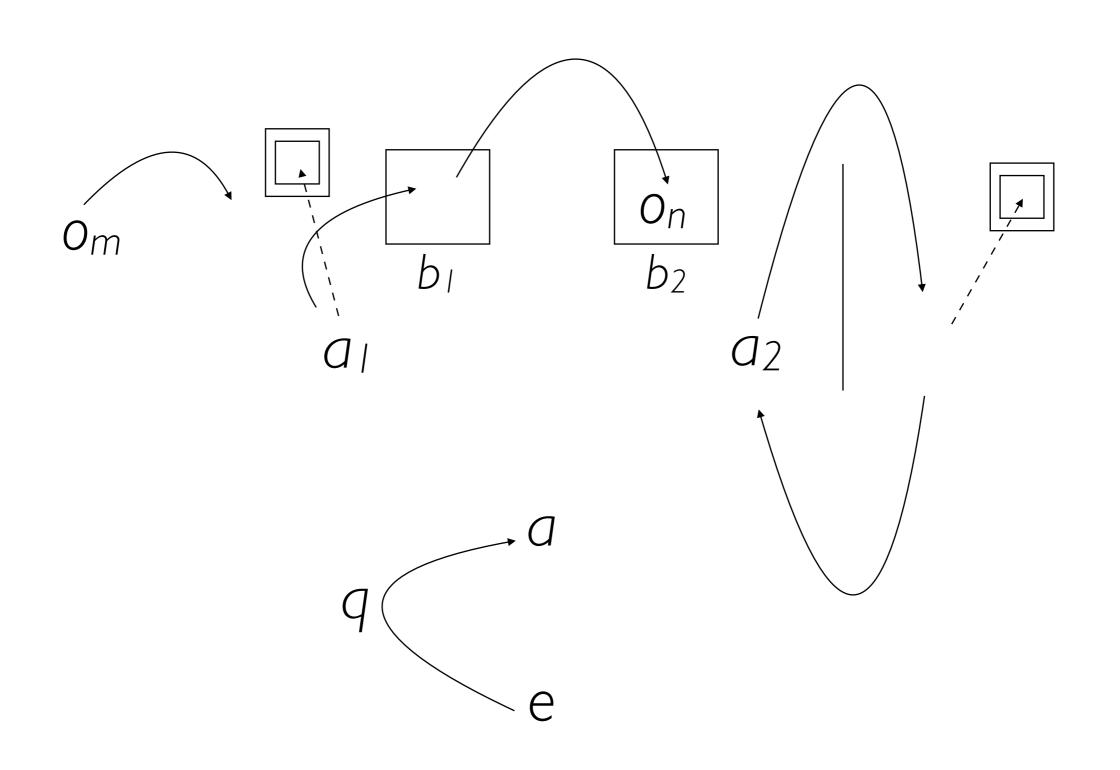
(eight timepoints)

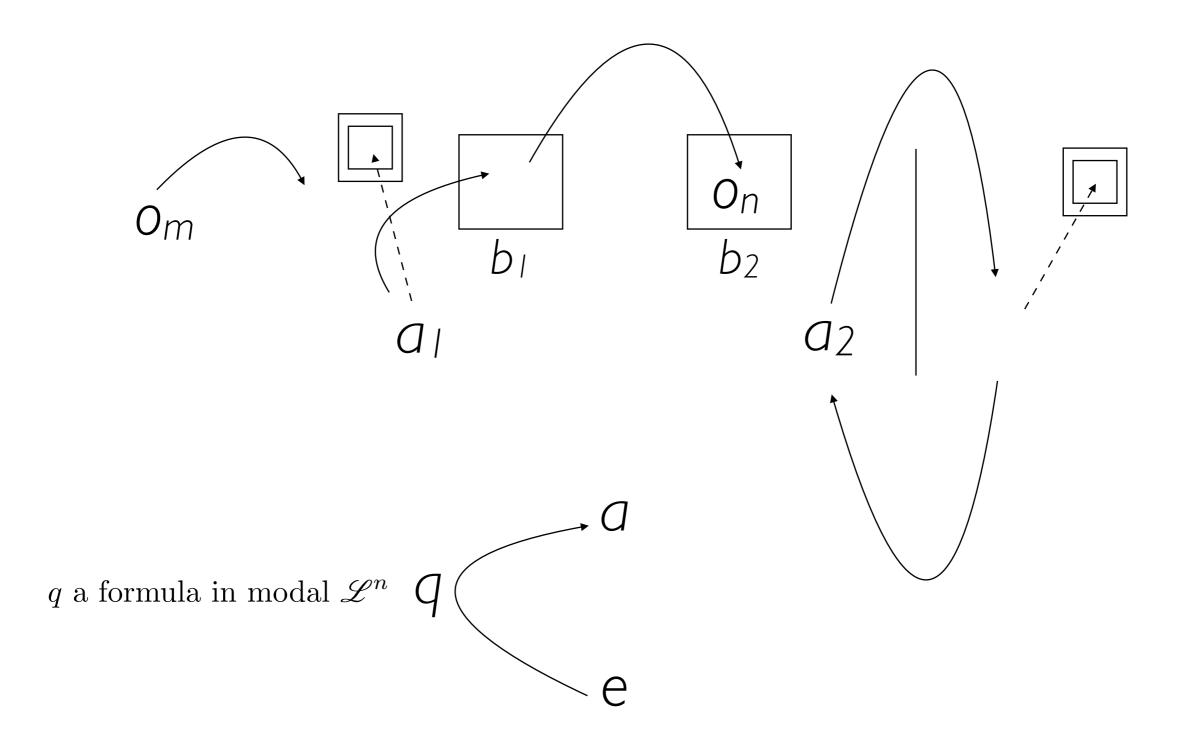


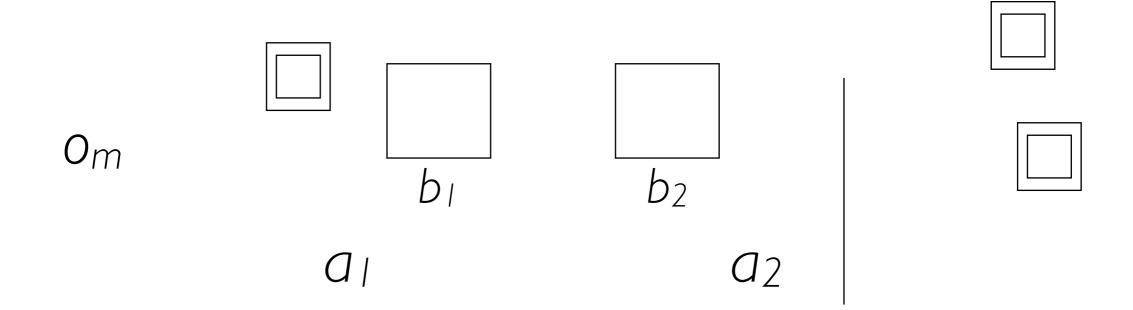
 $\mathcal{Q}$ 





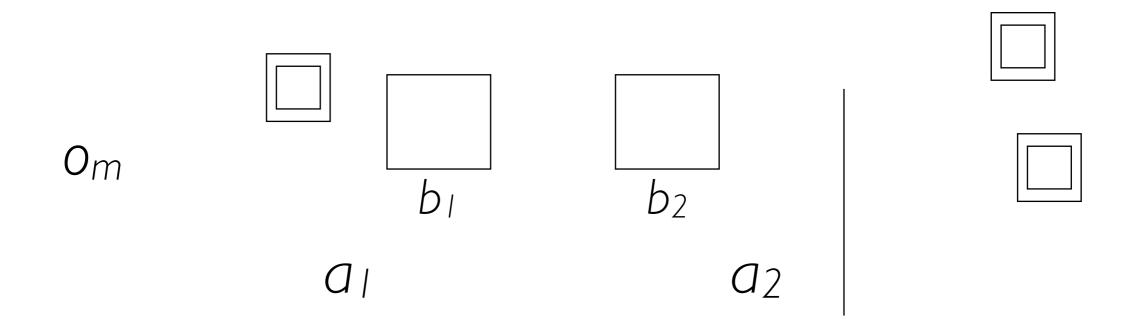






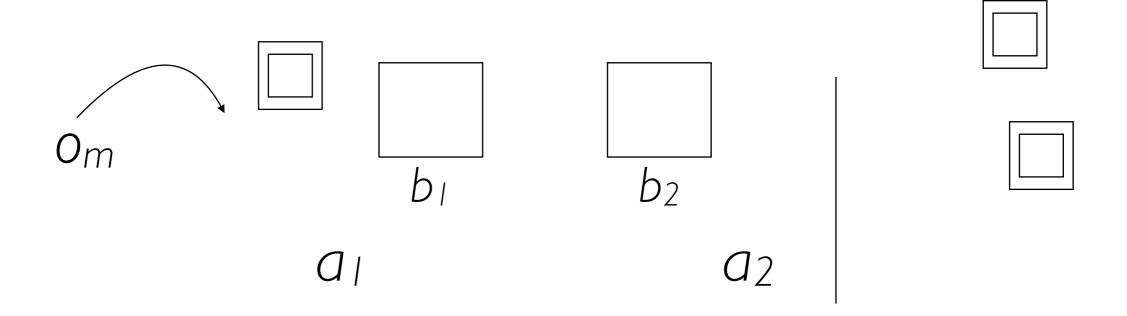
a

# Framework for FBT 4 (nine timepoints)



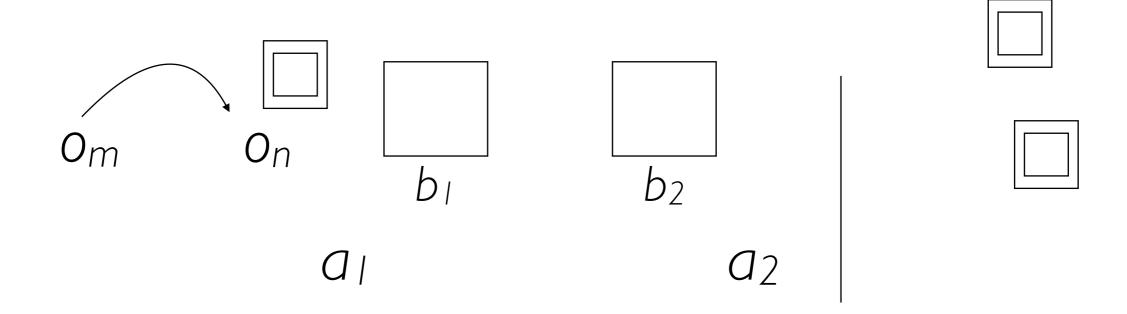
 $\mathcal{Q}$ 

# Framework for FBT 4 (nine timepoints)



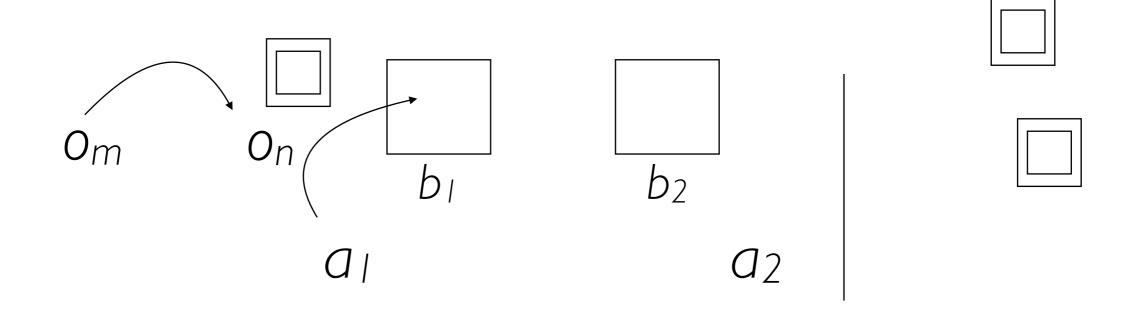
a

# Framework for FBT 4 (nine timepoints)

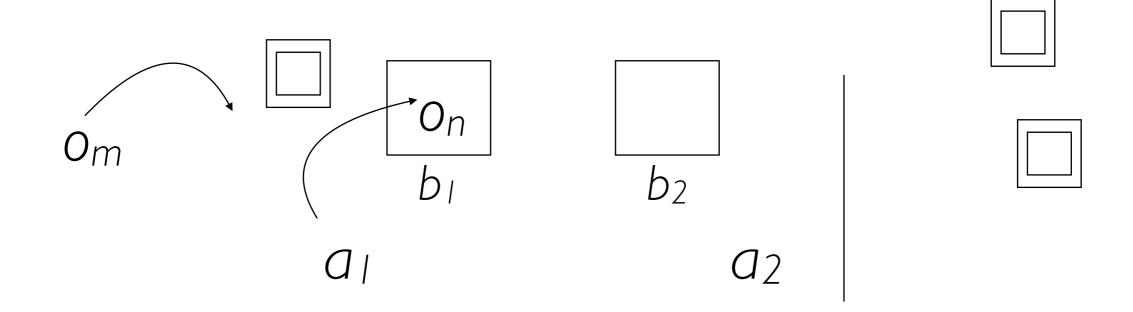


U

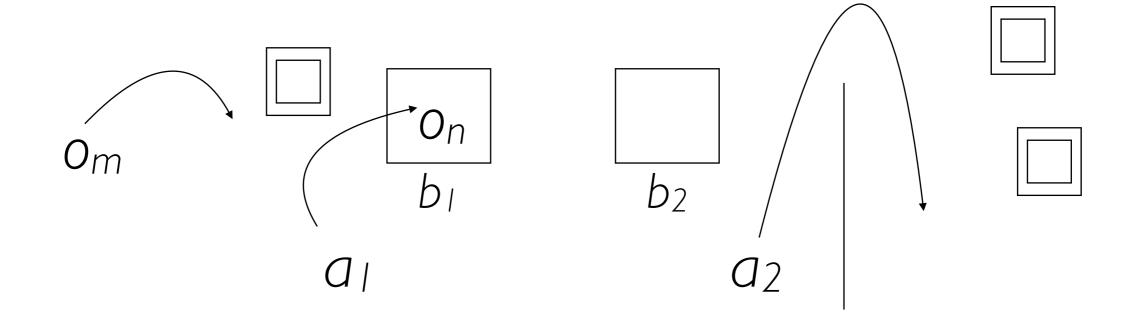
(nine timepoints)



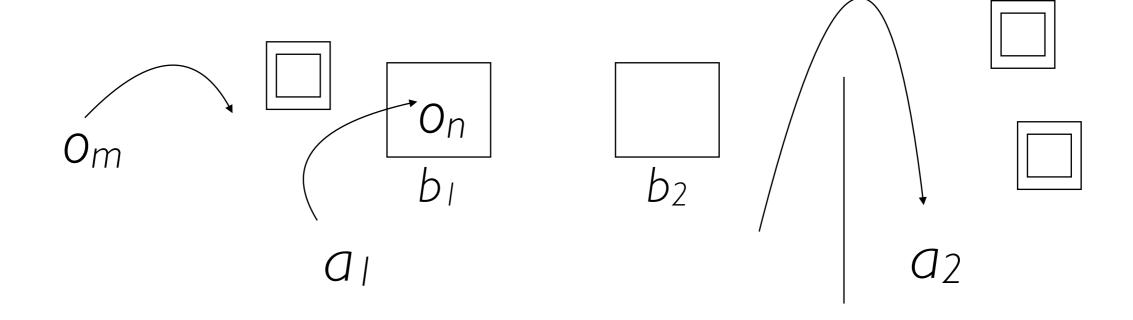
(nine timepoints)



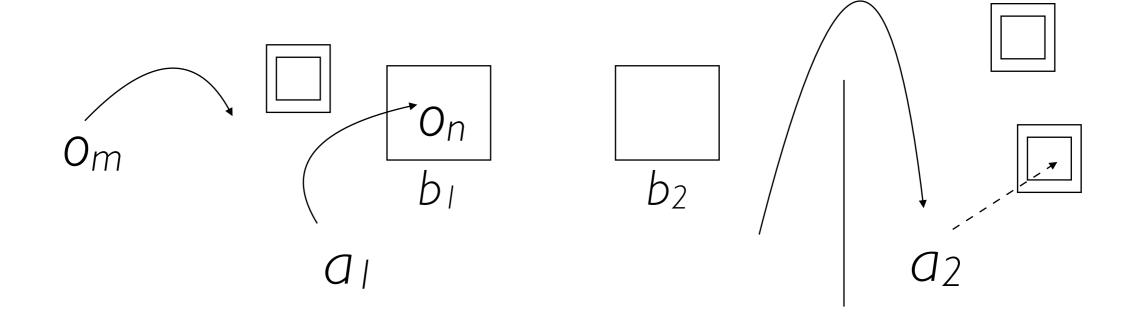
(nine timepoints)



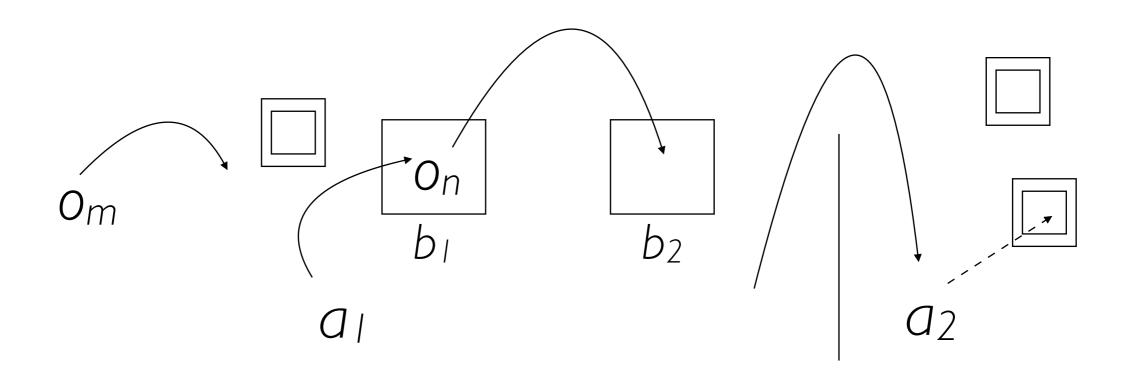
(nine timepoints)



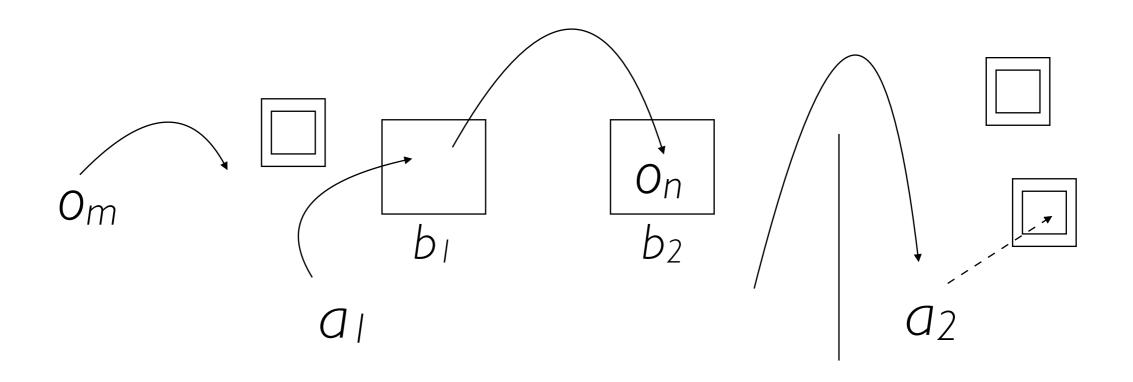
(nine timepoints)



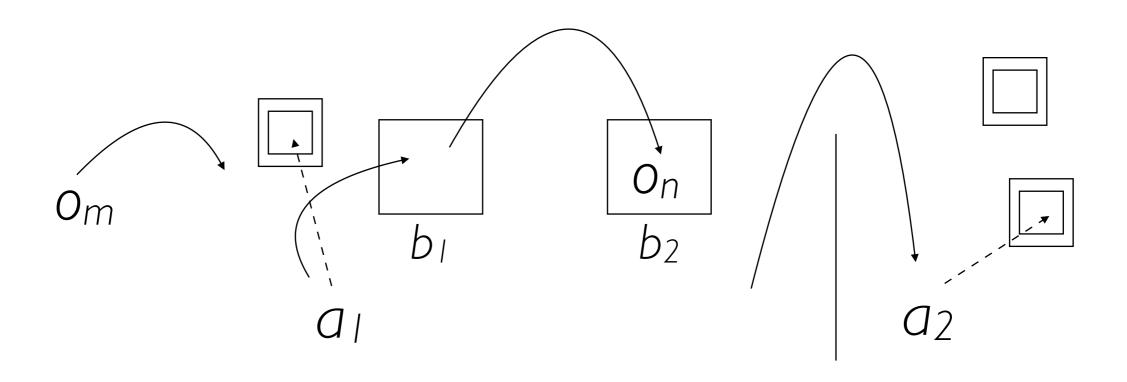
(nine timepoints)



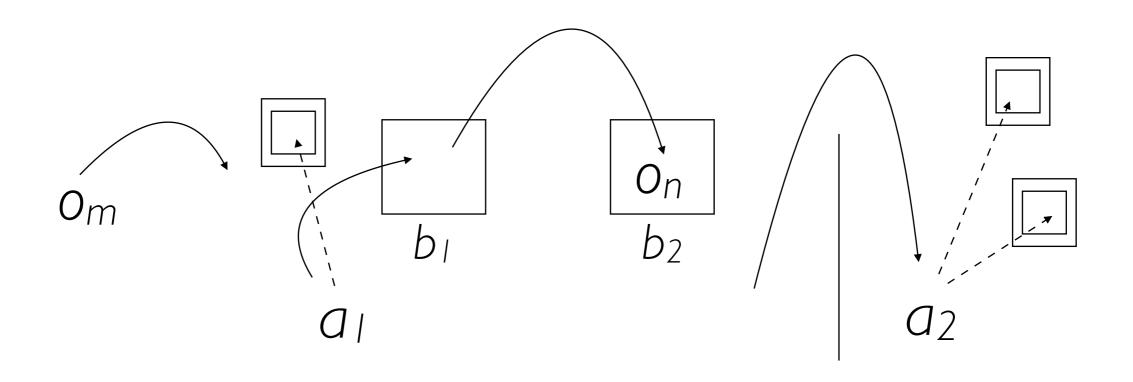
(nine timepoints)

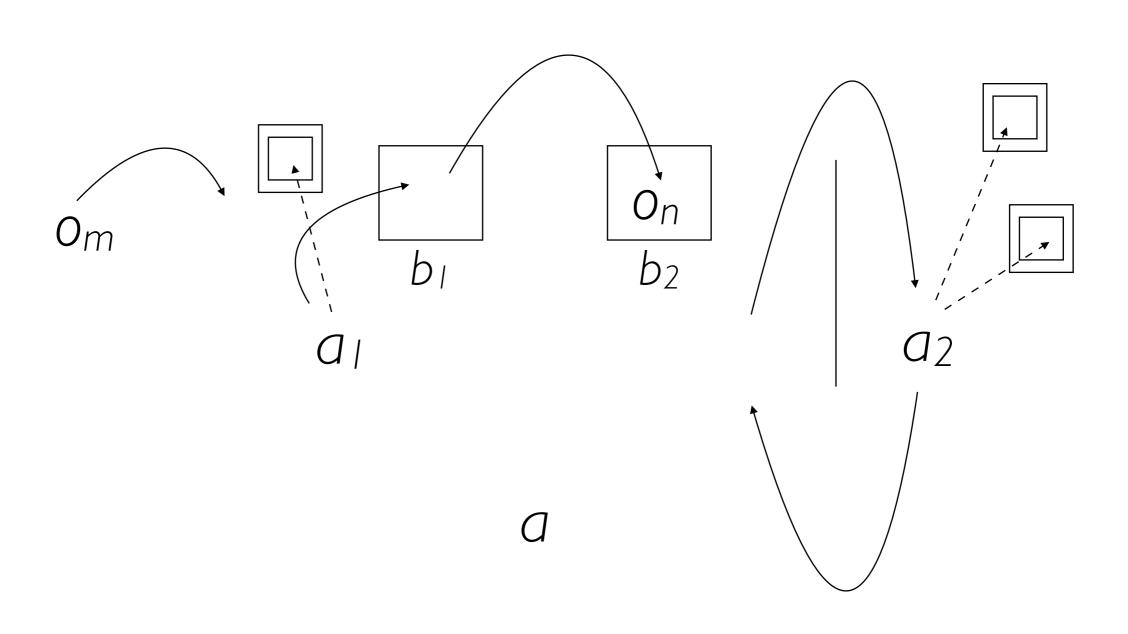


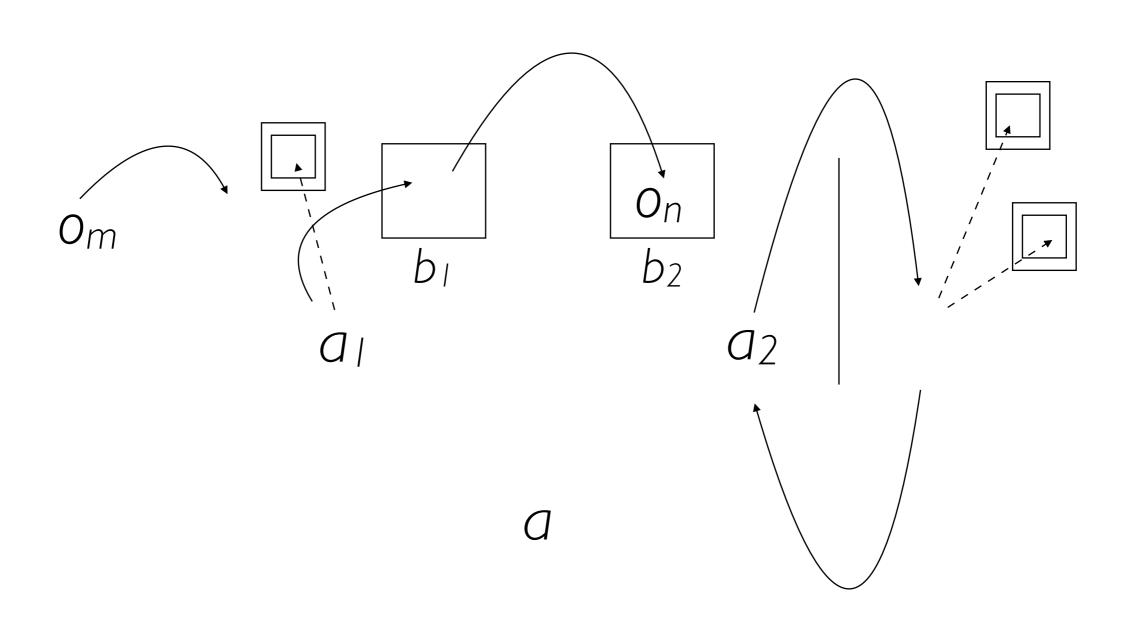
(nine timepoints)

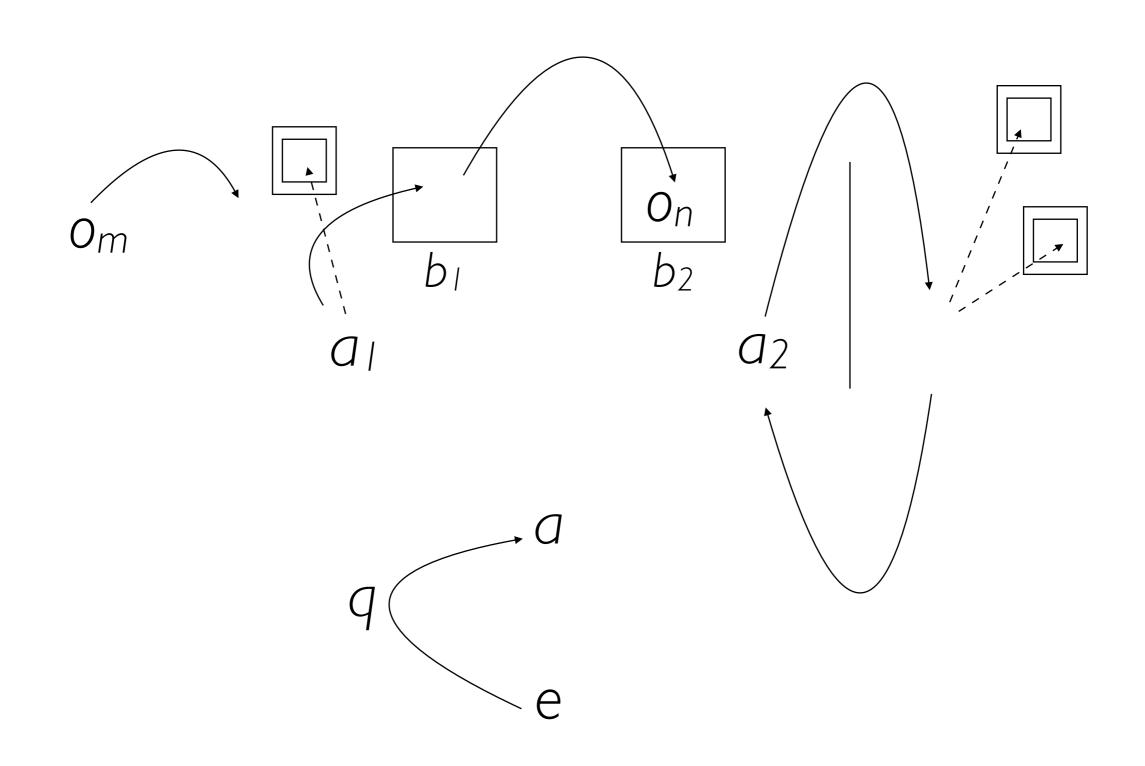


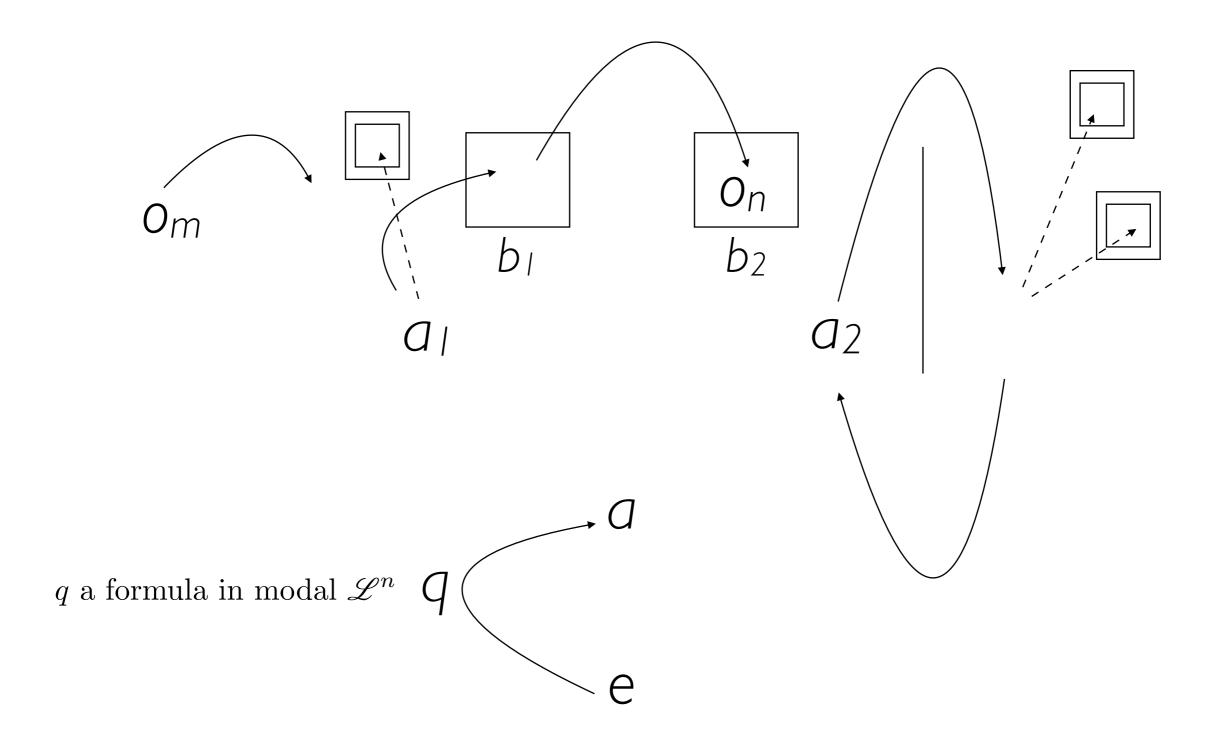
(nine timepoints)

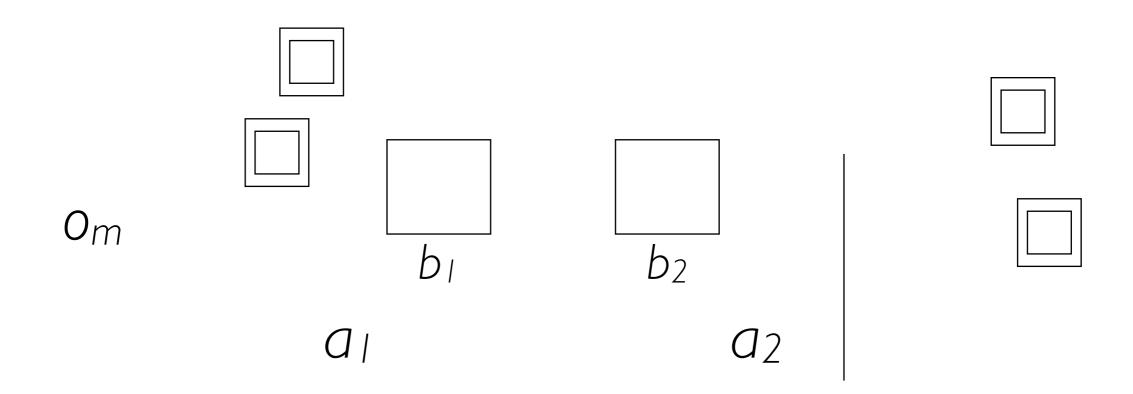






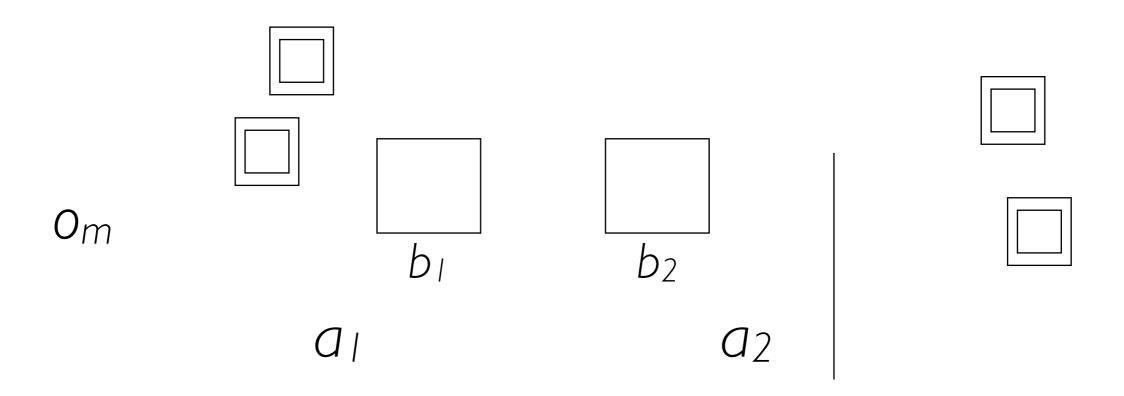






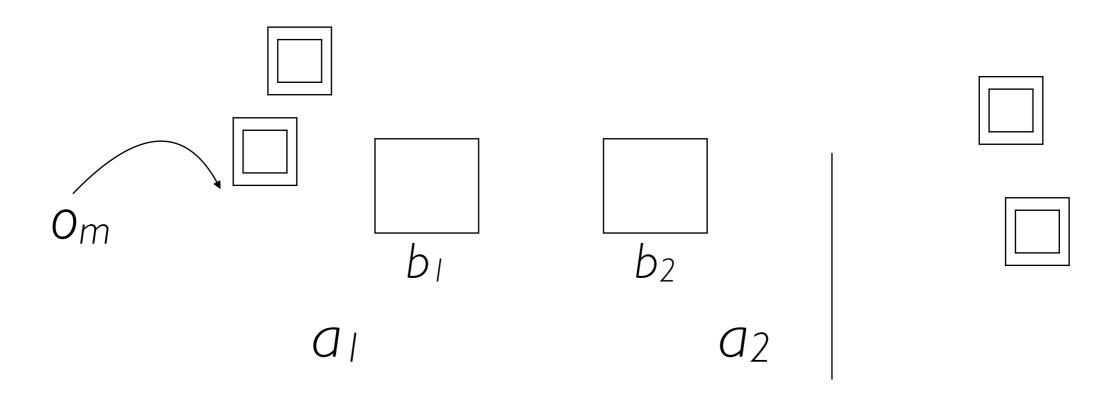
a

(ten timepoints)

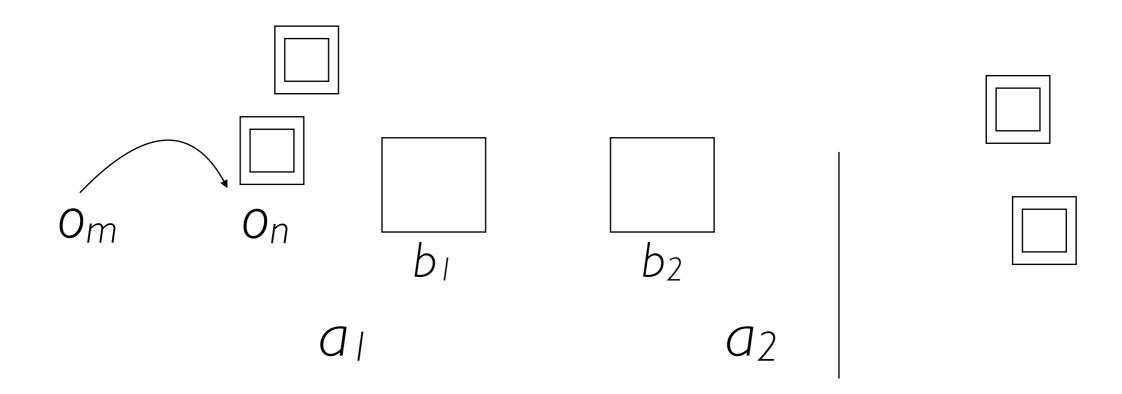


a

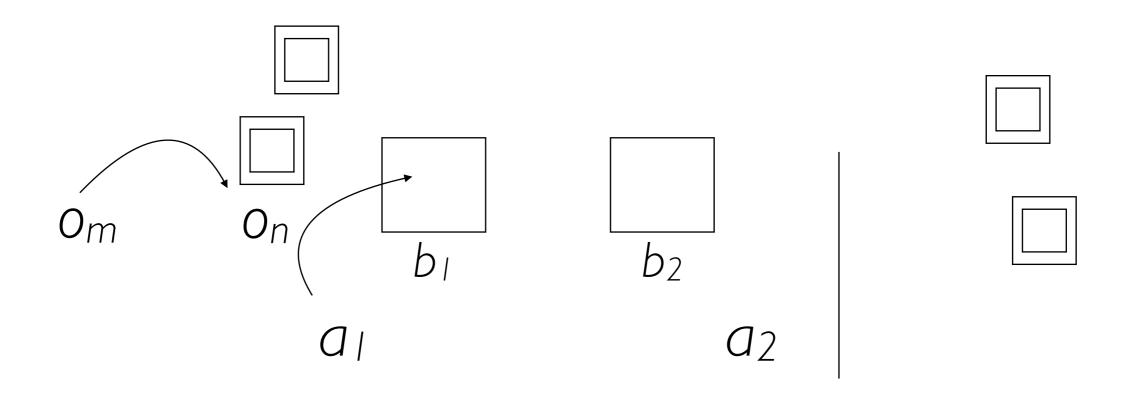
(ten timepoints)



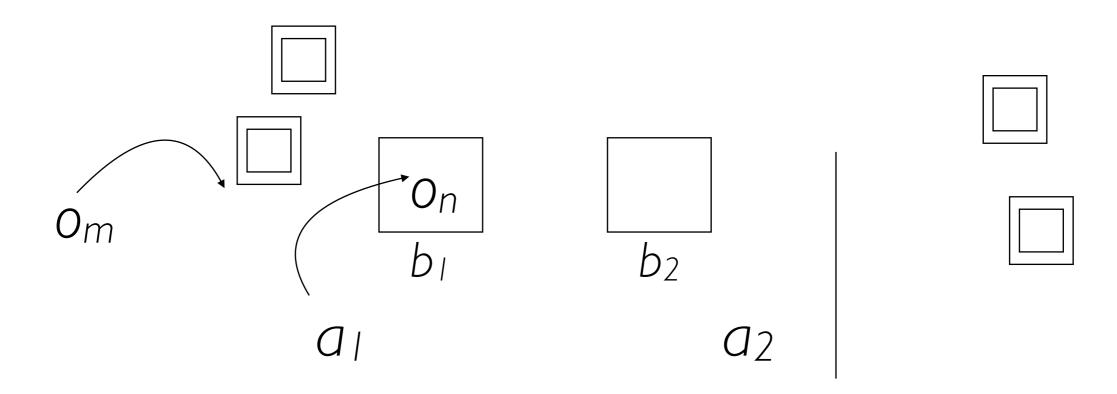
(ten timepoints)



(ten timepoints)

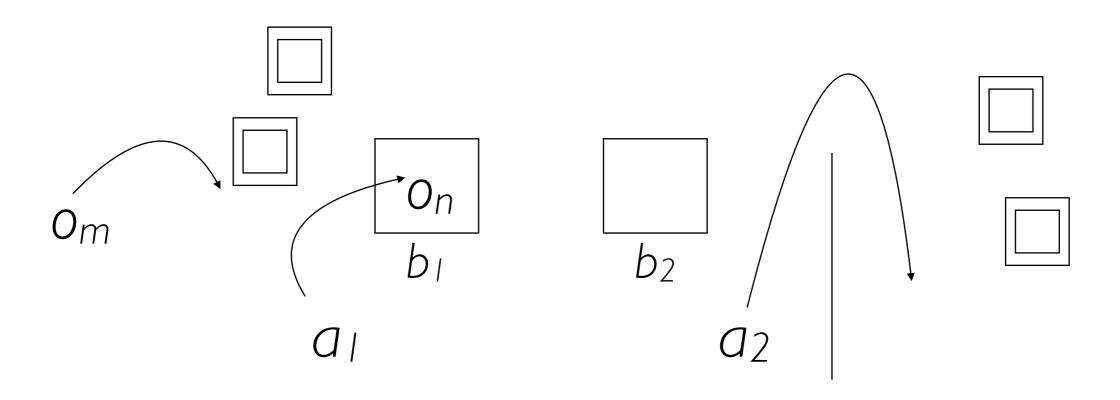


(ten timepoints)

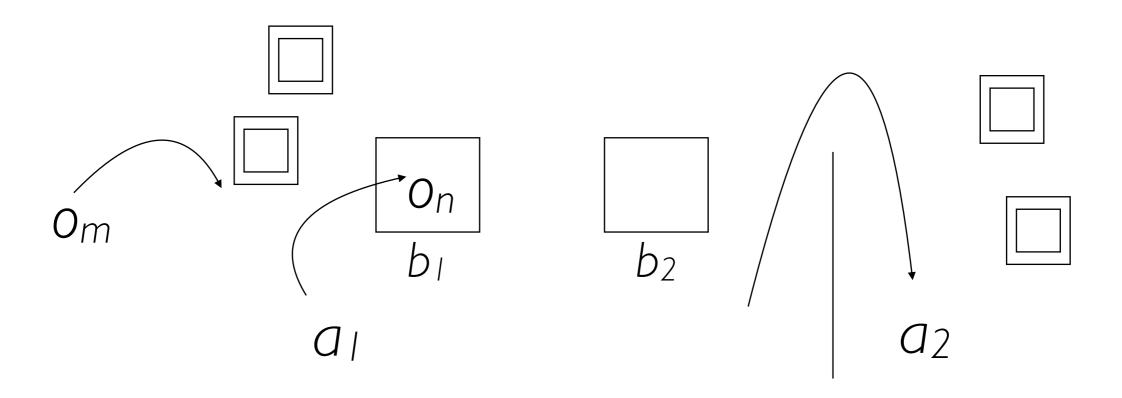


 $\mathcal{Q}$ 

(ten timepoints)

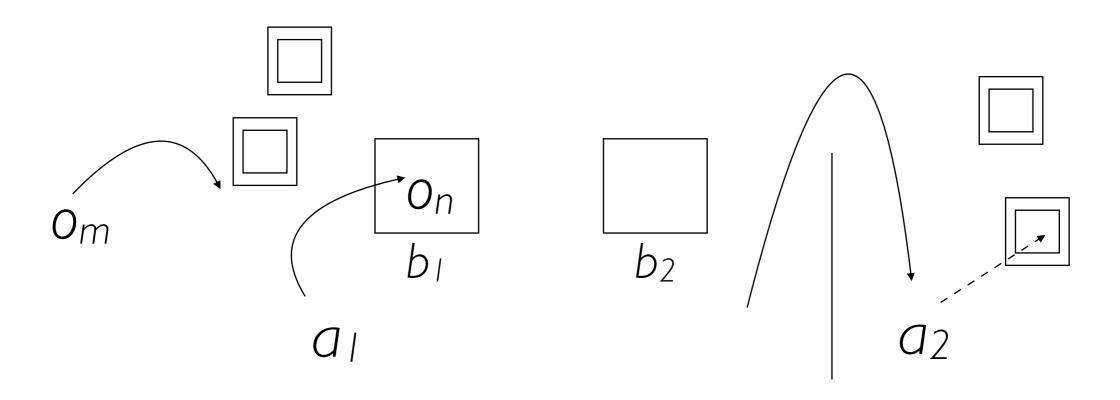


(ten timepoints)

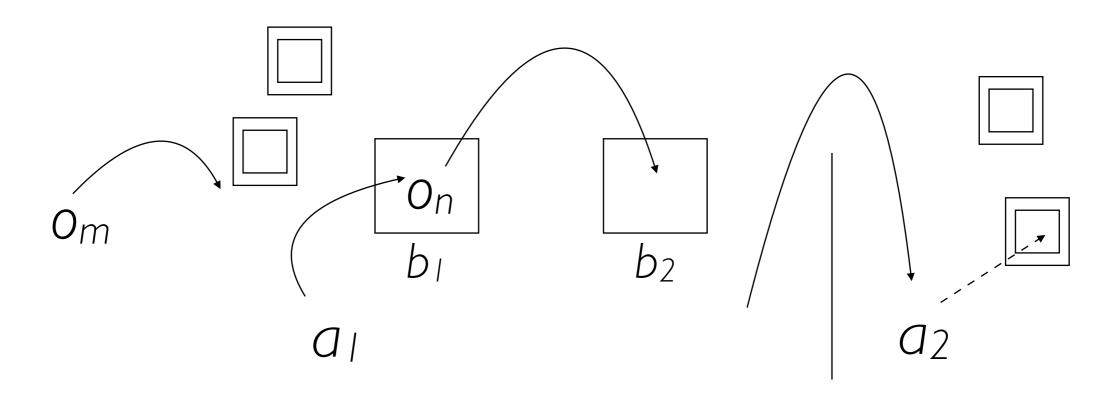


 $\mathcal{Q}$ 

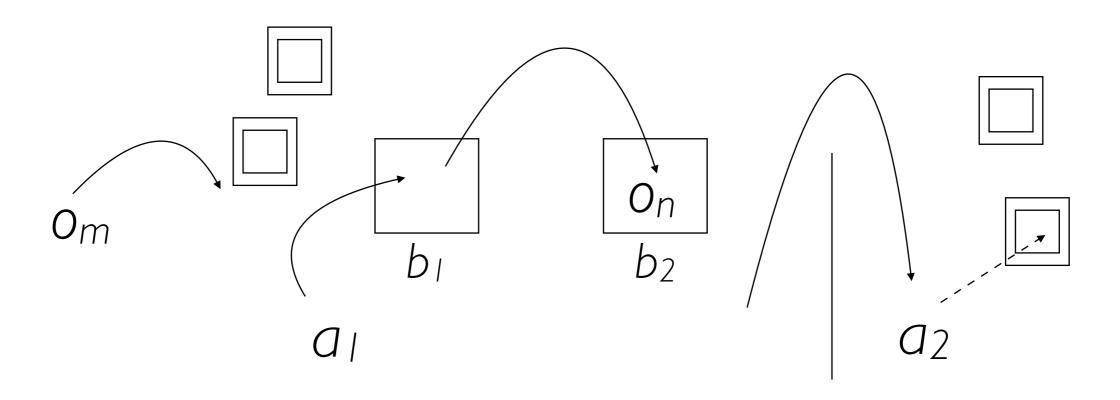
(ten timepoints)



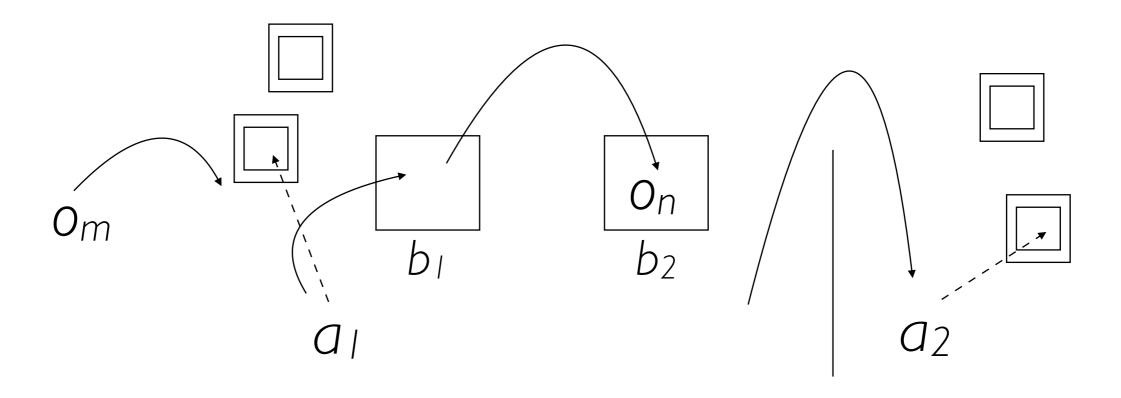
(ten timepoints)



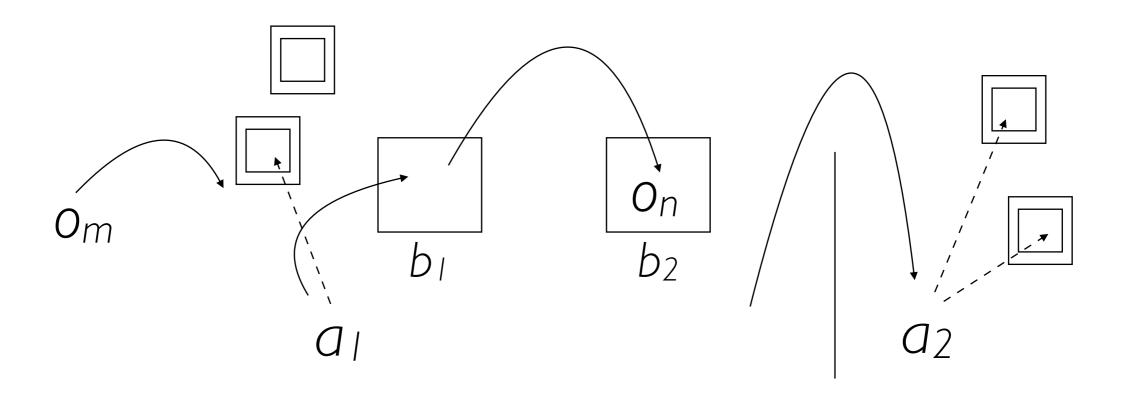
(ten timepoints)



(ten timepoints)

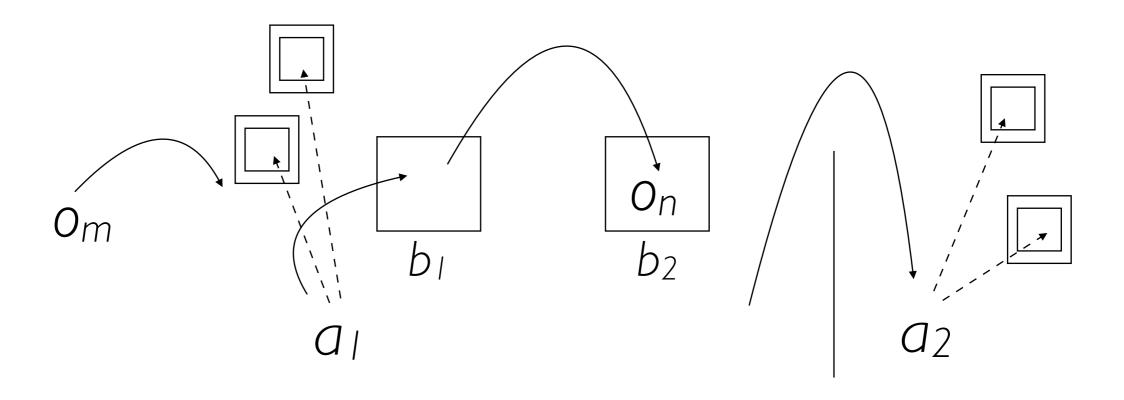


(ten timepoints)

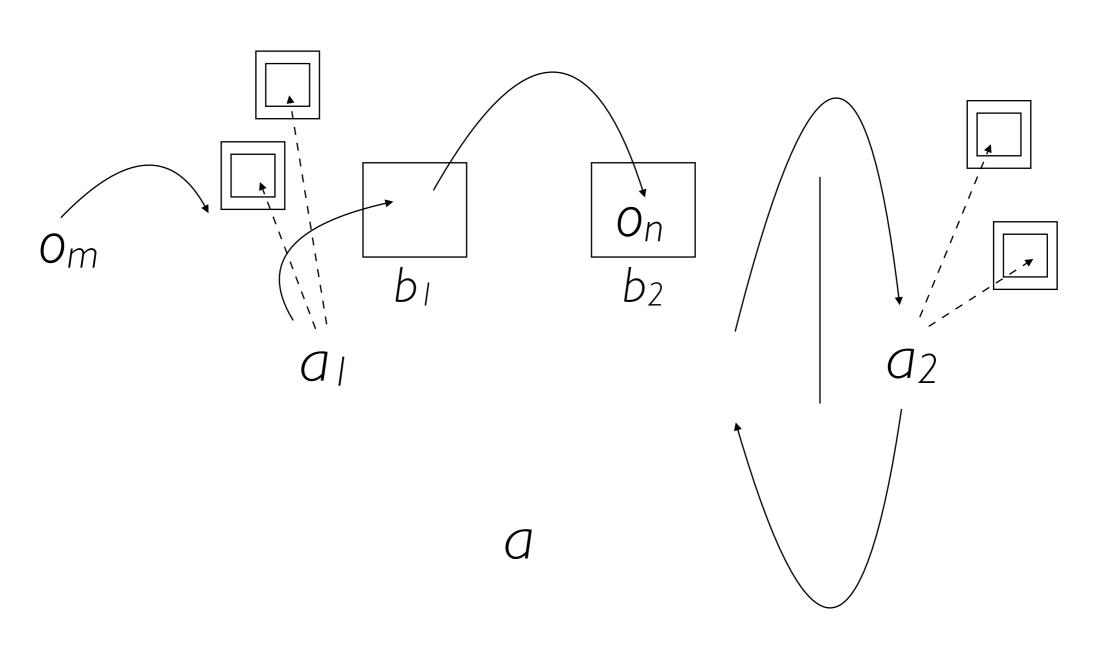


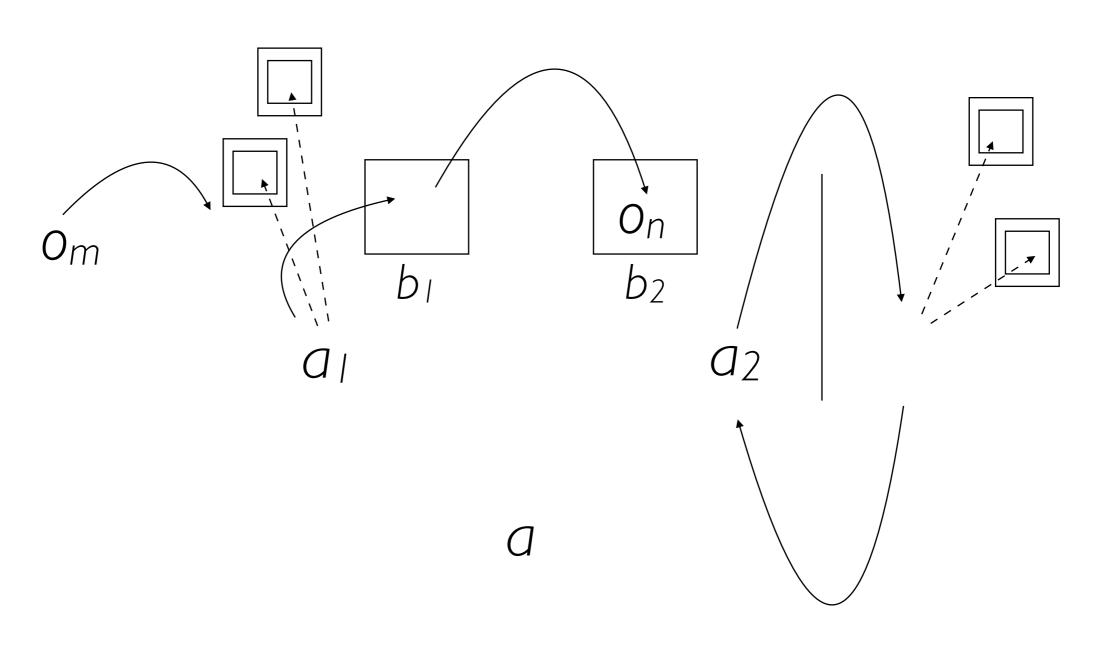
a

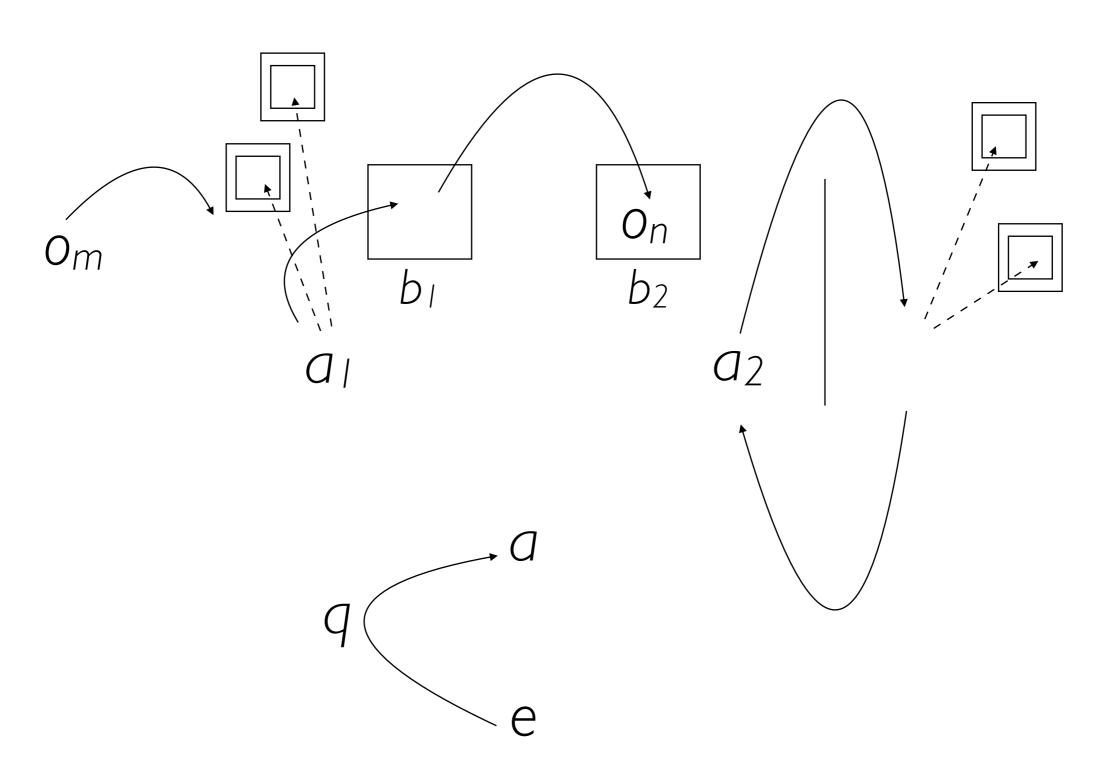
(ten timepoints)

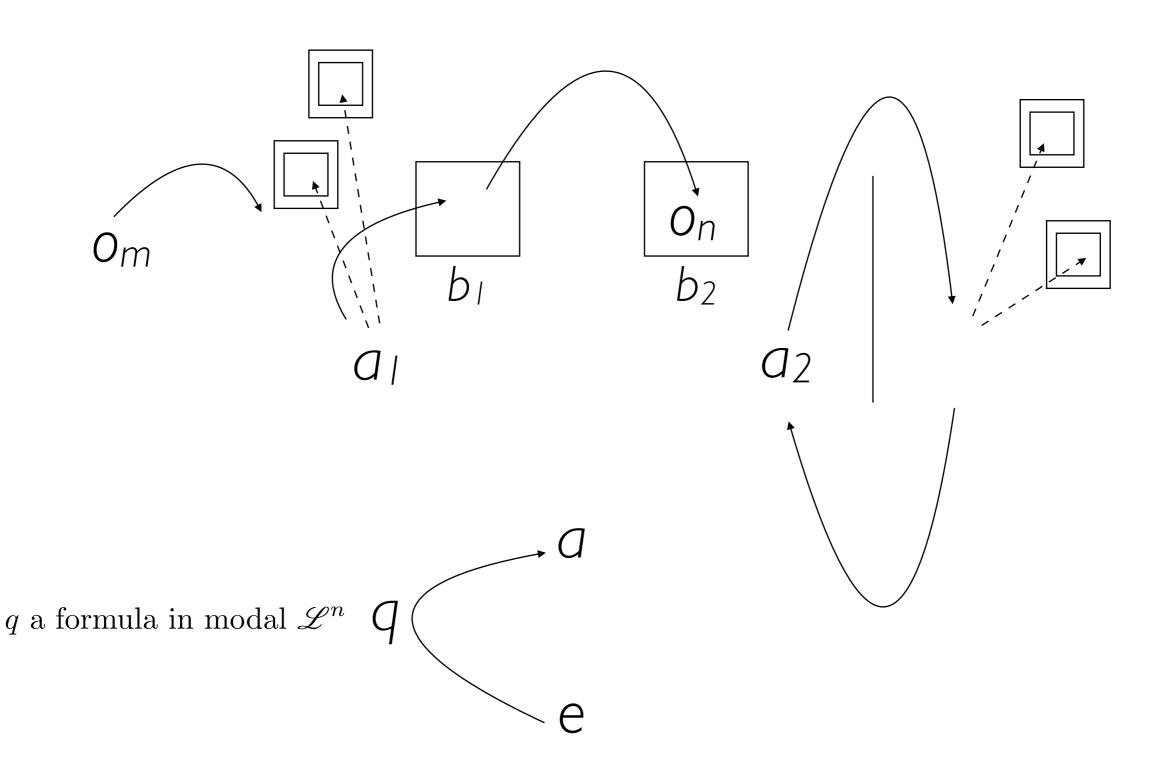


a









## 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  $\operatorname{FBT}_k, k \in \mathbb{Z}^+$ , holds; (i.e. that level k of  $\operatorname{FBT}$  holds). Then, if k is even,  $\mathbf{B}_2\mathbf{B}_1\dots\mathbf{B}_2\ \iota$ , where there are k+1 iterated  $\mathbf{B}_i$  operators; otherwise  $\mathbf{B}_1\mathbf{B}_2\dots\mathbf{B}_1\mathbf{B}_2\ \iota$ , 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**:  $\forall q \in \mathscr{CC}, \mathfrak{M}$  can correctly answer and justify q. I.e.,  $\mathfrak{M}$  can pass  $\mathrm{FBT}_{\omega}$ .

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**:  $\forall q \in \mathscr{CC}, \mathfrak{M}$  can correctly answer and justify q. I.e.,  $\mathfrak{M}$  can pass  $\mathrm{FBT}_{\omega}$ .

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 I

```
"Level 1: False Belief Task "
:name
:description "Agent al puts an object o into bl 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: al 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)))
             (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))}
:goal
```

```
"Level 2: False Belief Task "
{:name
 :description "Agent al 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.
              "Monday July 22, 2019"
 :date
 :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)))}
              (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3))))}
⊣ :goal
```

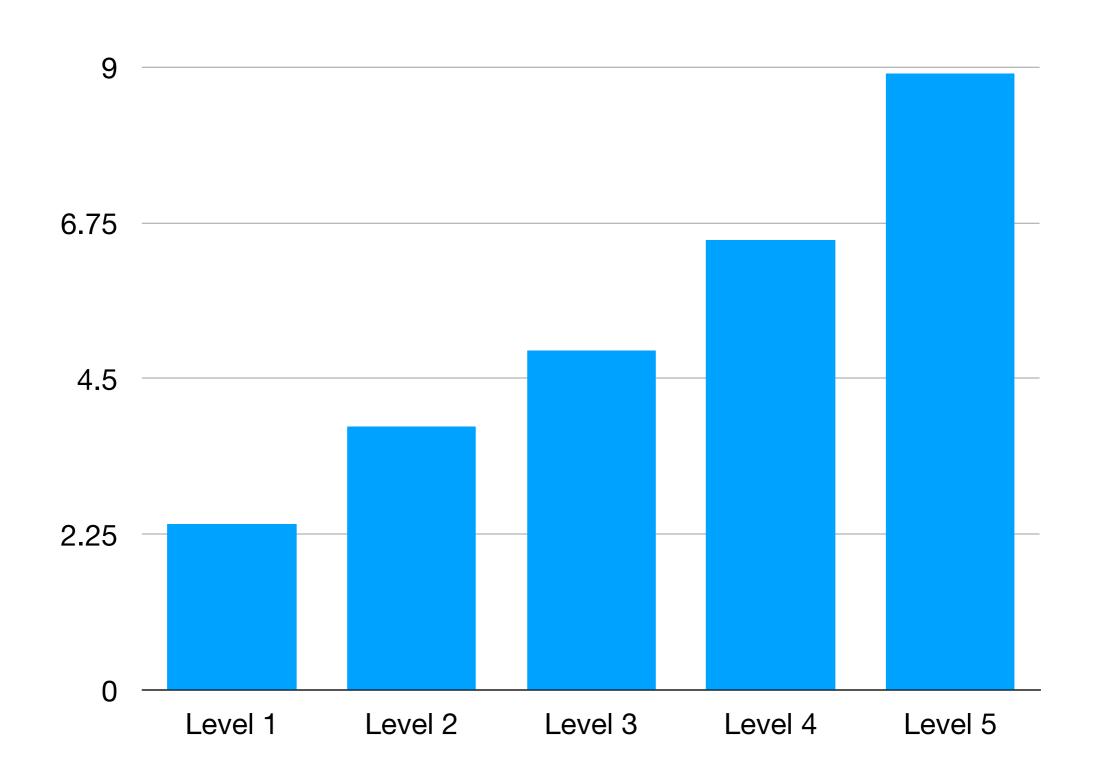
```
{:name
              "Level 3: False Belief Task "
 :description "Agent al puts an object o into b1 in plain view of a2.
               Agent a2 then leaves, and in the absence of a2, a1 moves o
               from b1 into b2; this movement isn't perceived by a2. Agent
               a2 now returns, 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)))}
              (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))))}
:goal
```

```
"Level 4: False Belief Task "
 :description "Agent al 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)))}
              (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))))}
 :goal
```

```
{:name
             "Level 5: False Belief Task "
:description "Agent al puts an object o into bl 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 bl."
              Level5 Belief: a1 believes a2 believes a1 believes a2 believes a1 believes a2 believes o is in b1.
             "Monday July 22, 2019"
:date
:assumptions {
               :P1 (Perceives! a1 t1 (Perceives! a2 t1 (Perceives! a1 t1 (Perceives! a2 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
                                    (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)))}
             (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (Believes! a1 t3 (Believes! a2 t3 (holds (In o b1) t3)))))))
:goal
```

(Common! to (In o bi) ti) To (Bellows) of the same of th ·9081 C (Common! to (and (e to to)) (e to to))))

#### Time (in seconds) to Prove



#### Simulation of Level 5 in Real Time

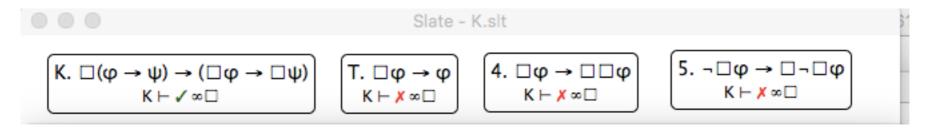
/Library/Java/JavaVirtualMachines/jdk1.8.0\_131.jdk/Contents/Home/bin/java ...

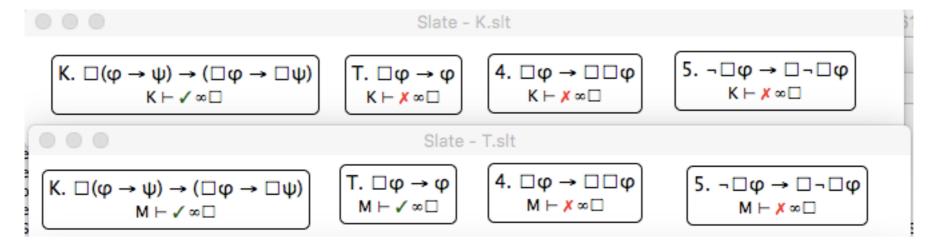
objc[16653]: Class\_JavaLaunchHelper is implemented in both /Library/Java/JavaVirtualMachines/jdk1.8.0\_131.jdk/Contents/Home/jre/lib/libinstrument.dylib (0x102ab94e0)

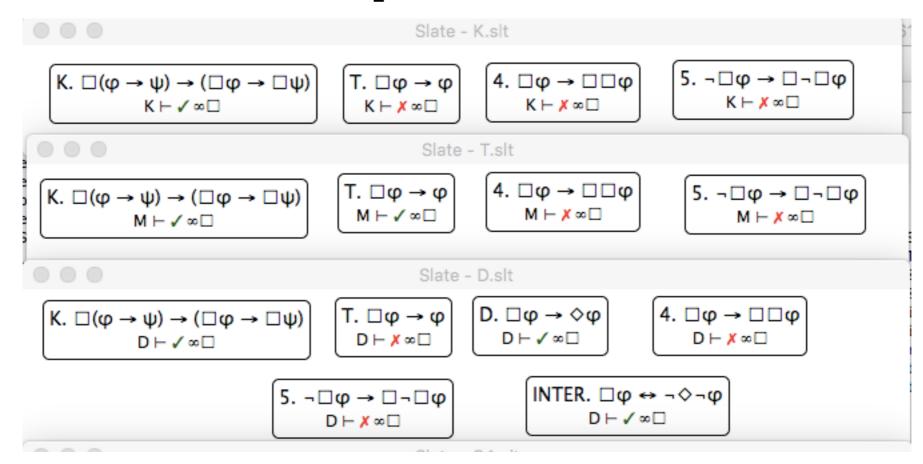
#### Simulation of Level 5 in Real Time

/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/jre/lib/libinstrument.dylib (0x102ab94e0)







$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

$$K \vdash \chi \otimes \Box$$

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

$$K \vdash \chi \otimes \Box$$

$$Slate - T.slt$$

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

$$M \vdash \chi \otimes \Box$$

$$Slate - D.slt$$

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

$$D \vdash \chi \otimes \Box$$

$$D \vdash \chi \otimes \Box$$

$$Slate - D.slt$$

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

$$D \vdash \chi \otimes \Box$$

$$D \vdash \chi \otimes \Box$$

$$Slate - S4.slt$$

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

$$D \vdash \chi \otimes \Box$$

$$Slate - S4.slt$$

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

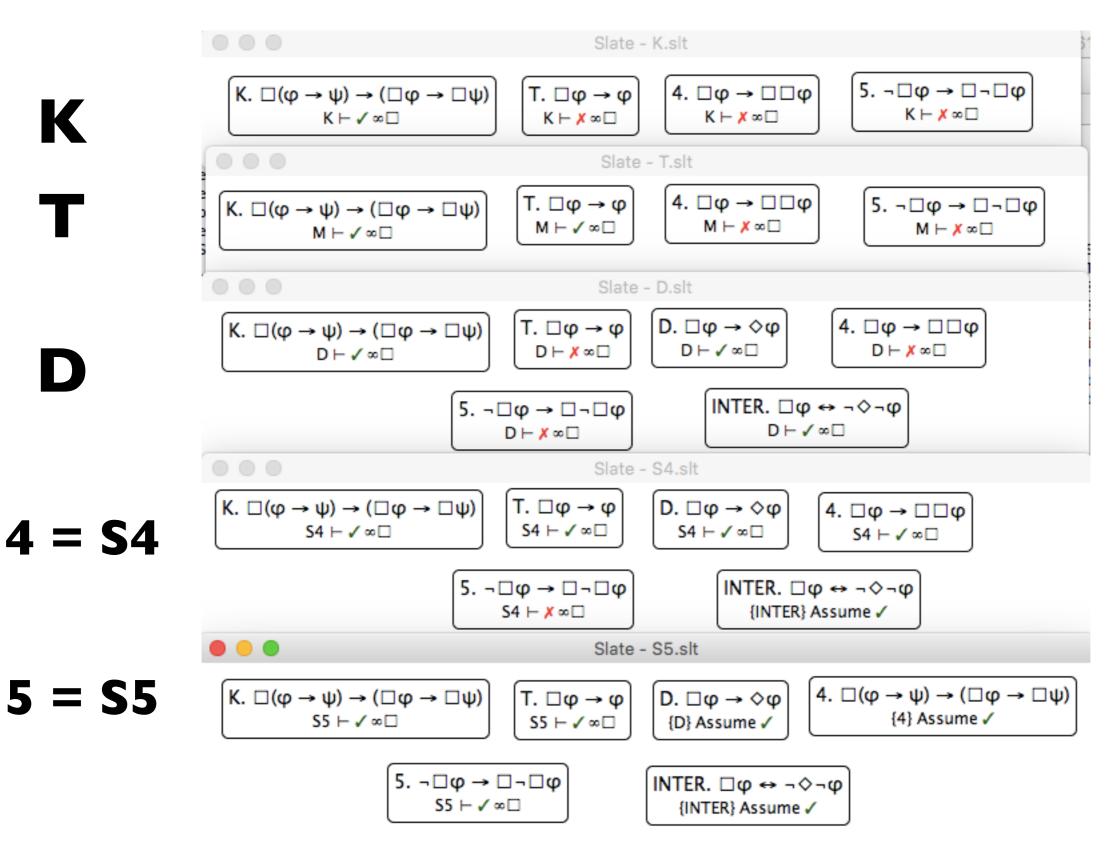
$$Slate - S4.slt$$

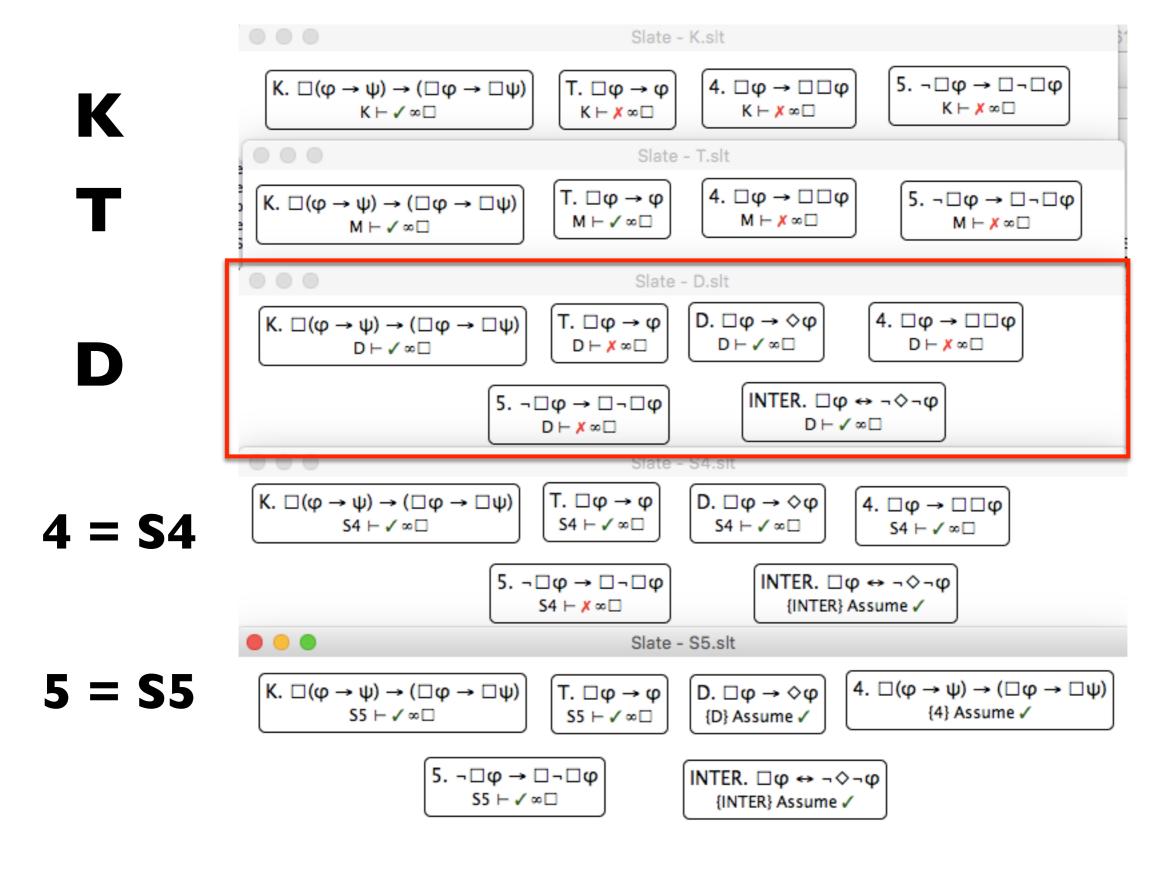
$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

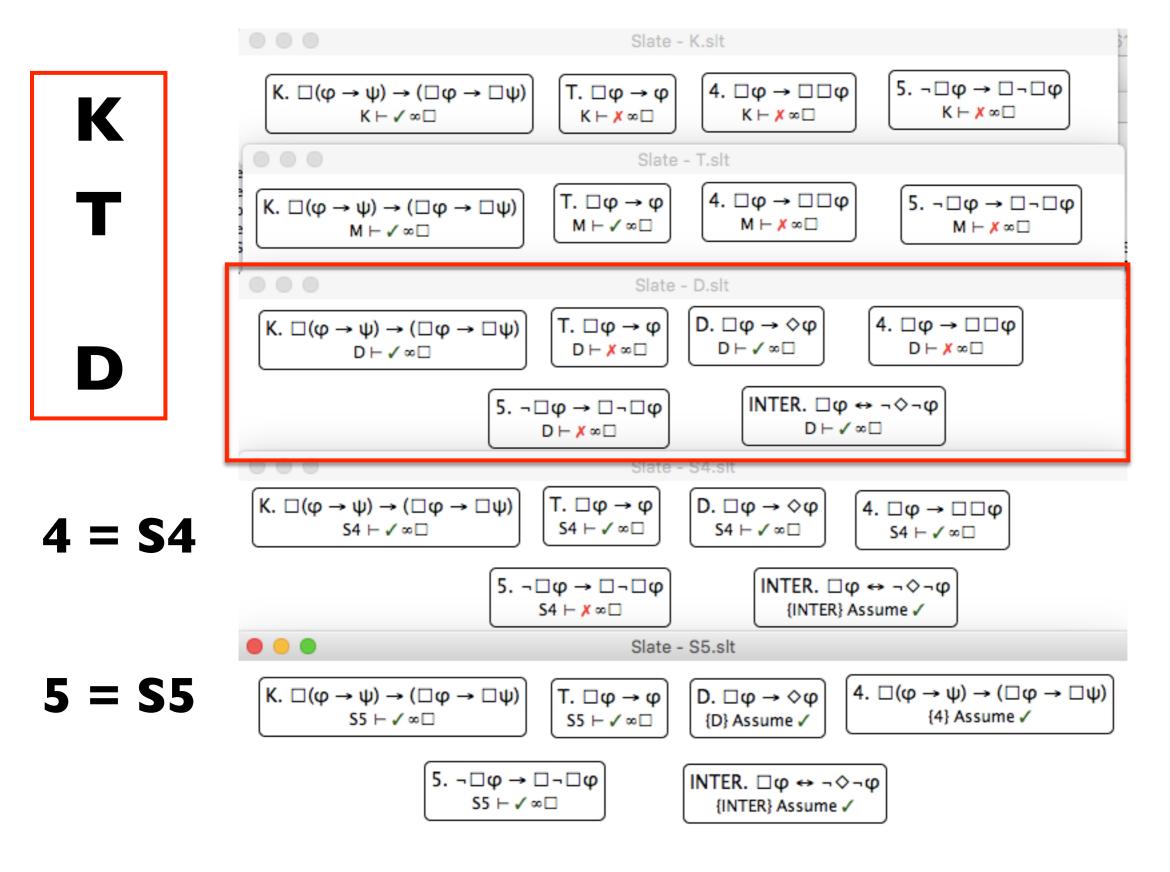
$$Slate - S4.slt$$

$$S1.slt - S4.s$$

K

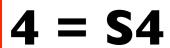




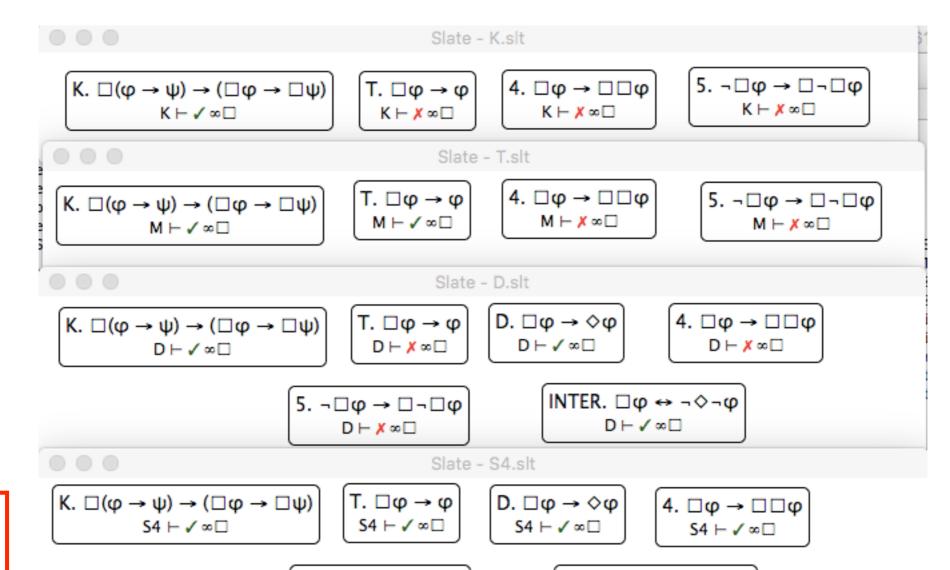


K

D



$$5 = S5$$



Slate - S5.slt

$$K. \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$$

S5  $\vdash \checkmark \infty \square$ 

$$\begin{bmatrix} \mathsf{T.} \ \Box \varphi \to \varphi \\ \mathsf{SS} \vdash \checkmark \infty \Box \end{bmatrix}$$

5.  $\neg \Box \phi \rightarrow \Box \neg \Box \phi$ 

 $S4 \vdash x \infty \square$ 

D. 
$$\Box \phi \rightarrow \Diamond \phi$$
 {D} Assume  $\checkmark$ 

4. 
$$\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$$
  
{4} Assume  $\checkmark$ 

INTER. □φ ↔ ¬⋄¬φ
{INTER} Assume ✓

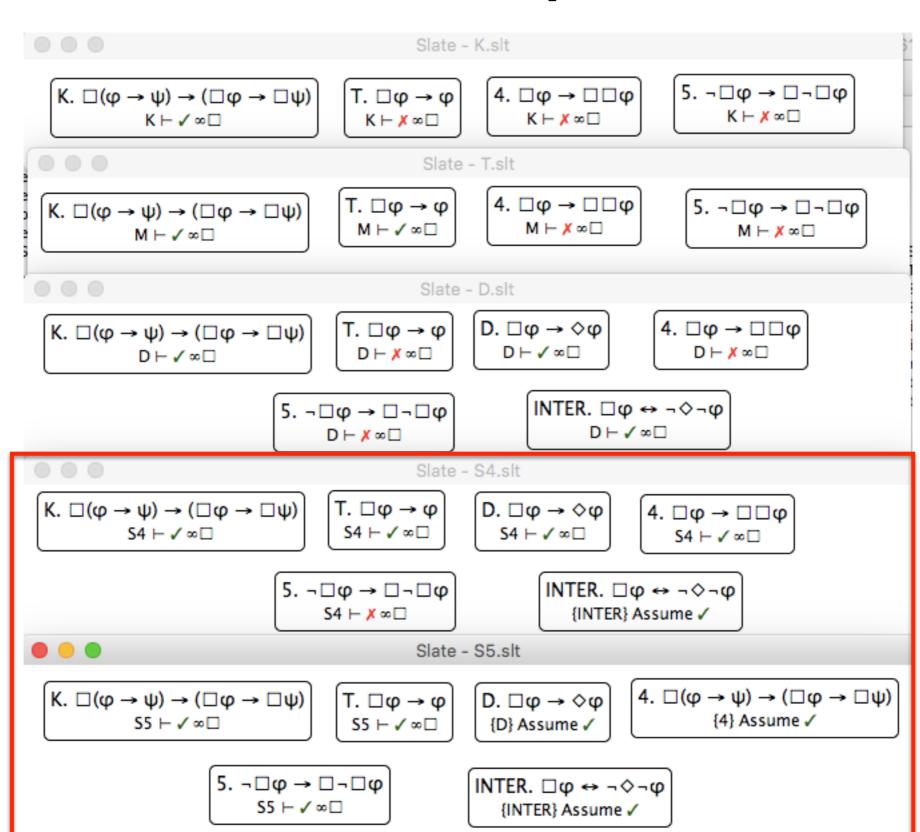
K

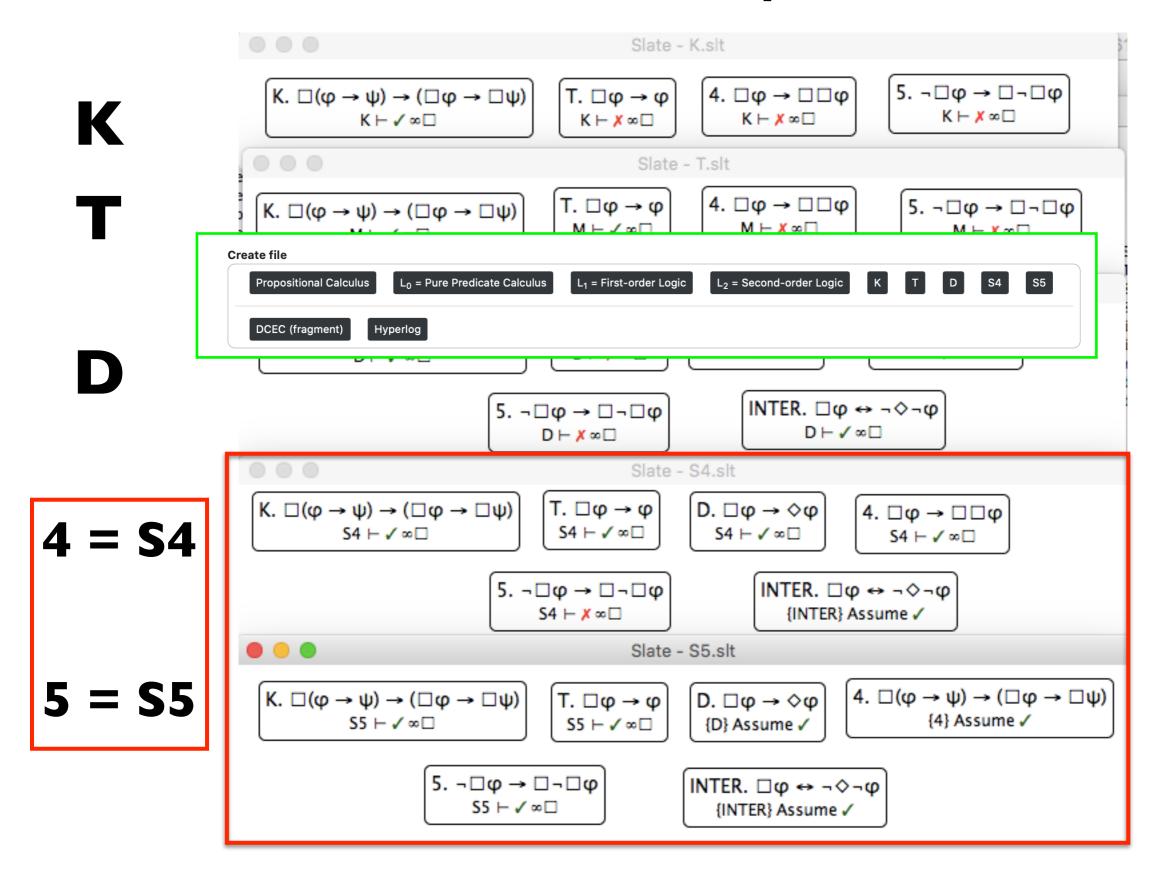
T

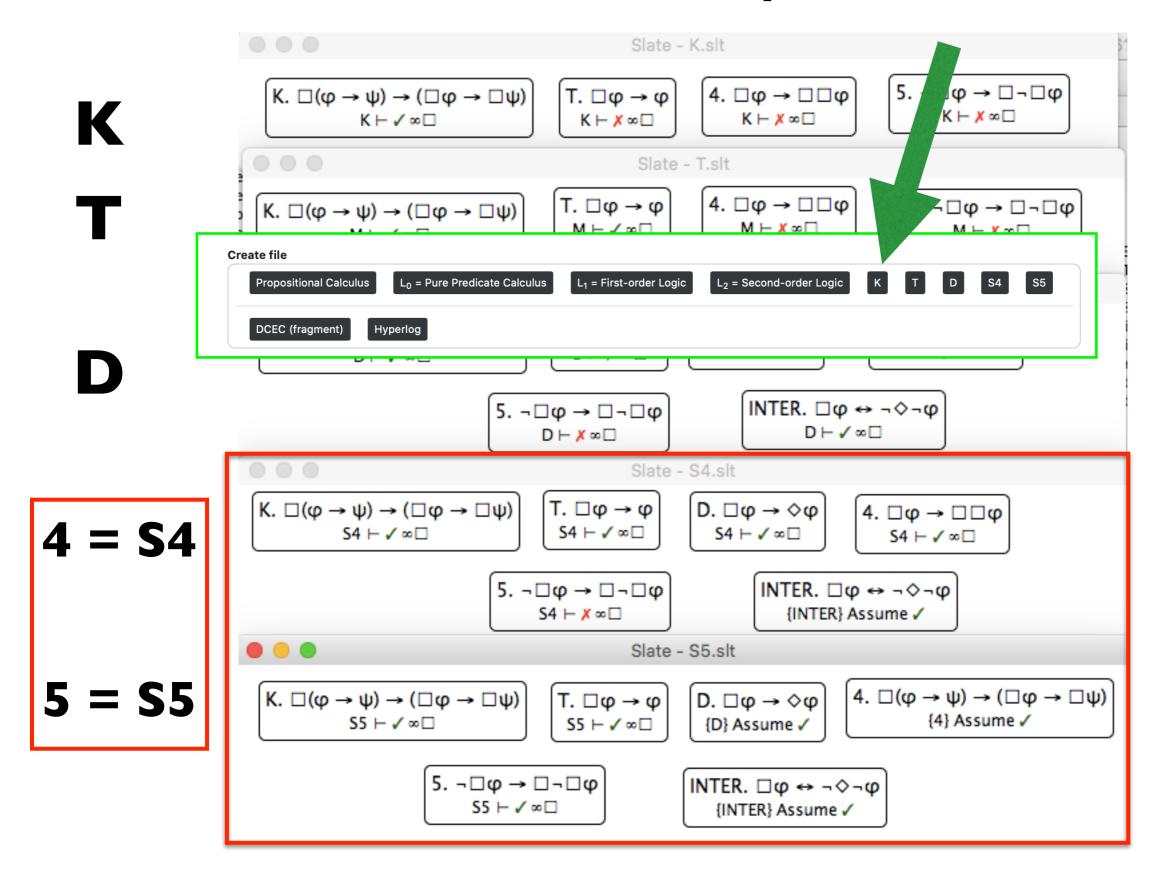
D

$$4 = $4$$

$$5 = S5$$







# Det er en logikk for hvert problem!