How to grow a mind: statistics, structure and abstraction

Josh Tenenbaum
MIT Department of Brain and Cognitive Sciences
CSAIL

Funding: ONR, ARO, AFOSR, IARPA, Google, Qualcomm

Acknowledgments

Tom Griffiths
Charles Kemp

Noah Goodman
Vikash Mansinghka

Dan Roy
Cameron Freer

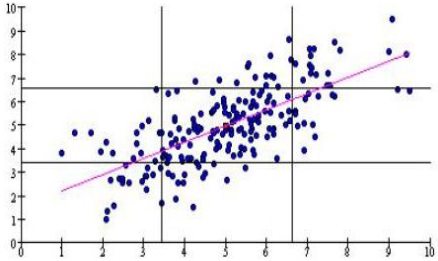
Peter Battaglia
Chris Baker
Tomer Ullman
Steve Piantadosi

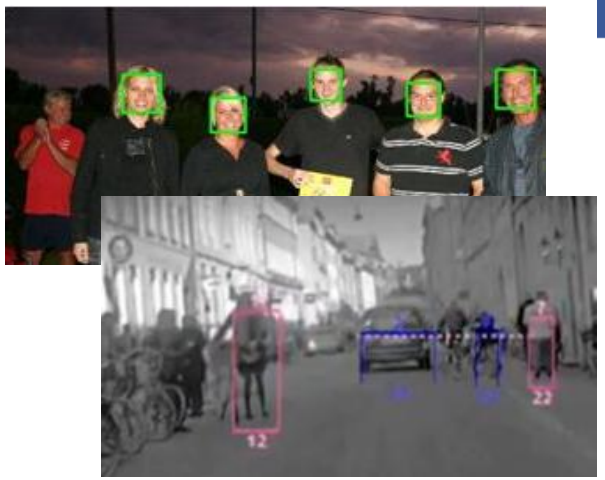
The goal

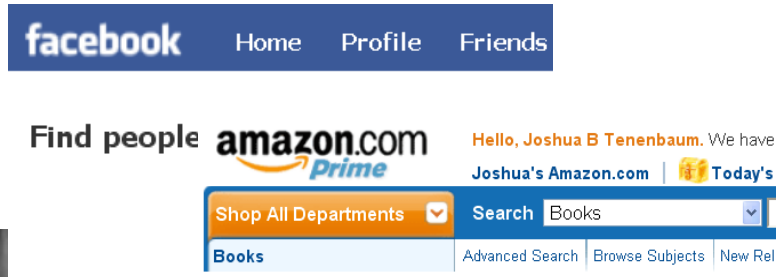
“Reverse-engineering the mind”

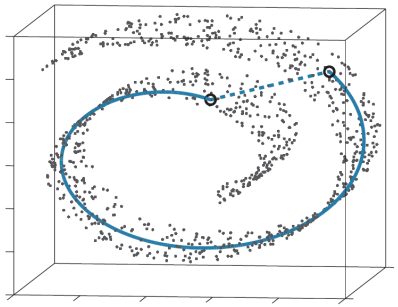
- Understand human intelligence in our best engineering terms, and use that knowledge to engineer more human-like intelligence in artificial systems.

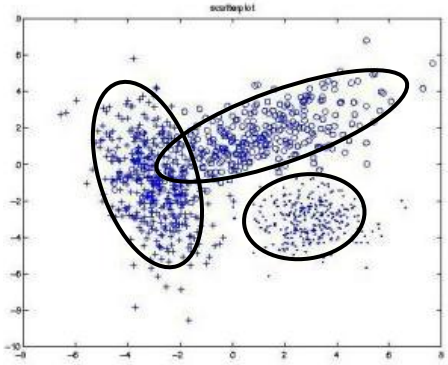
A success story: AI Technologies "statistics on a grand scale"

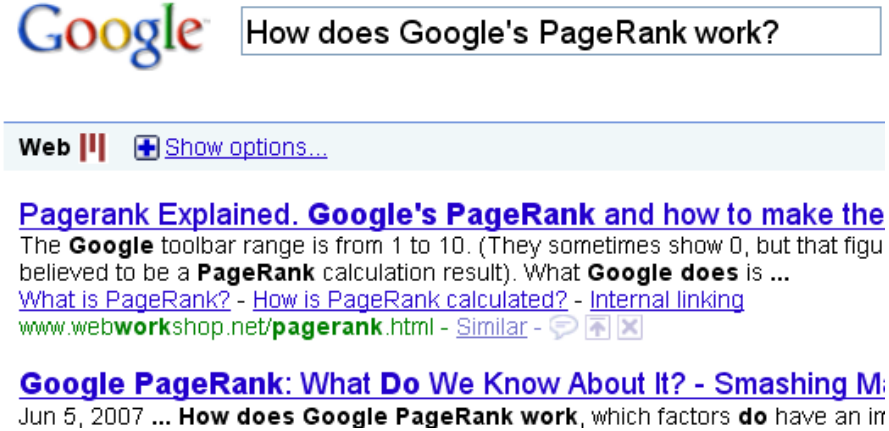




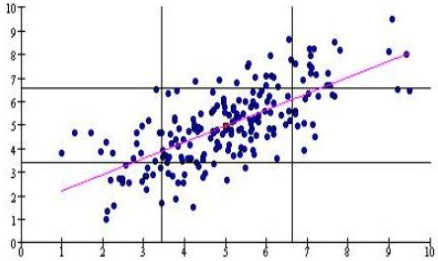


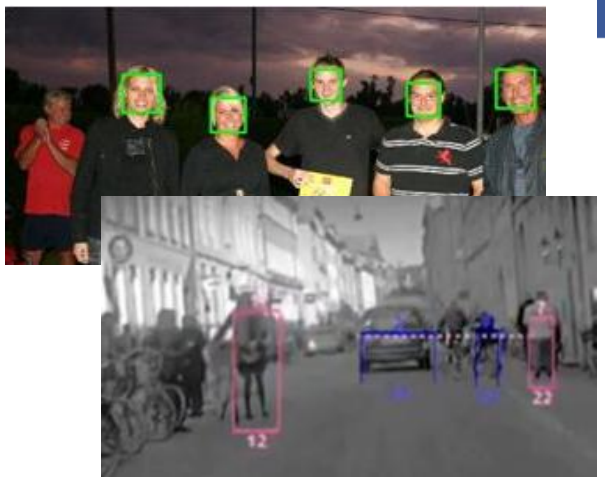


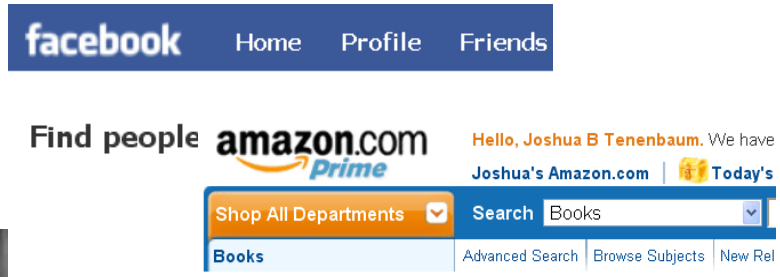


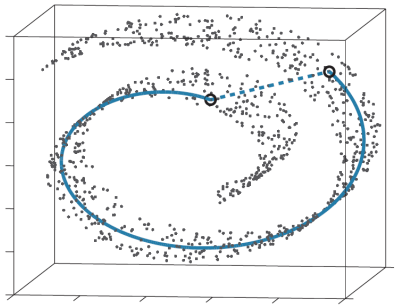


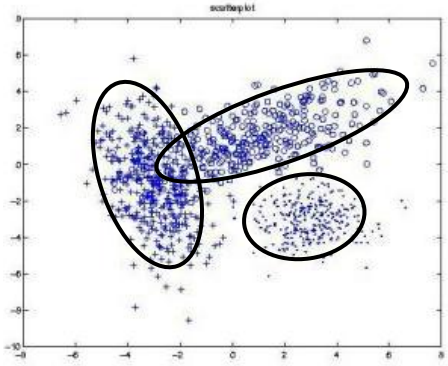
A success story: AI Technologies "statistics on a grand scale"

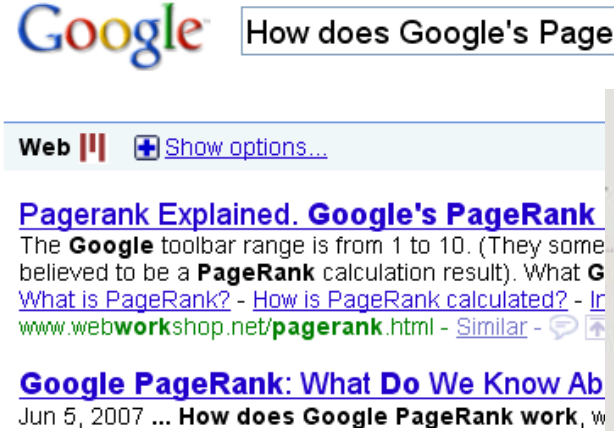












The big question

How does the mind get so much out of so little?

The big question

How does the mind get so much out of so little?

Our minds build rich models of the world and make strong generalizations from input data that is sparse, noisy, and ambiguous – in many ways far too limited to support the inferences we make.

How do we do it?

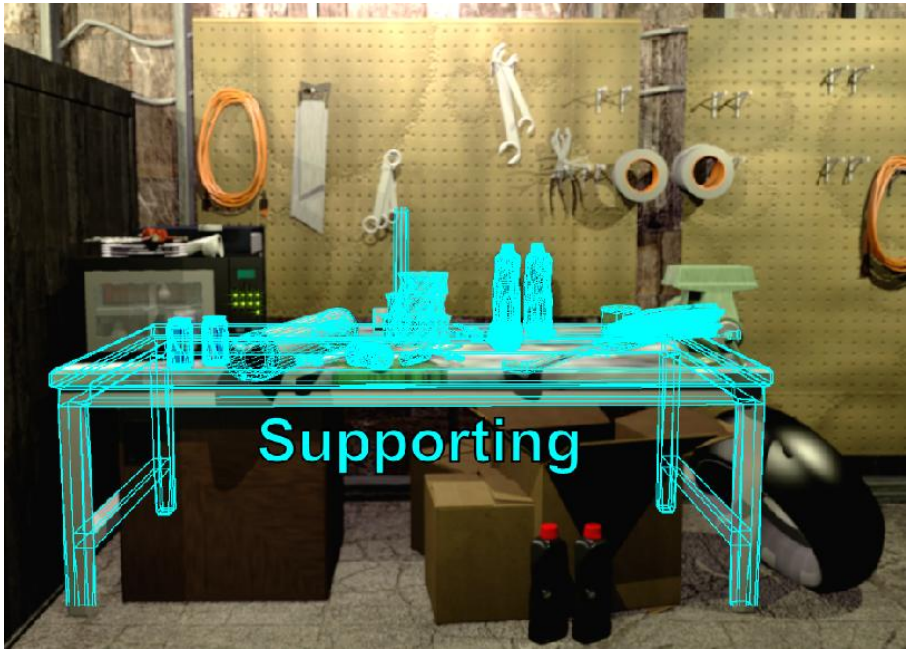
Learning from very few examples

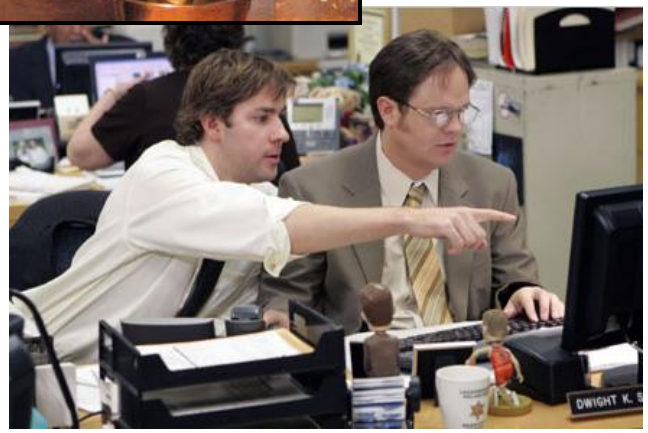
The big question

How does the mind get so much out of so little?

- Learning about kinds of objects and their properties.
- Learning the meanings of words, phrases, and sentences.
- Inferring causal relations.
- Intuitive theories of physics, psychology, and other common-sense domains.

Common-sense understanding

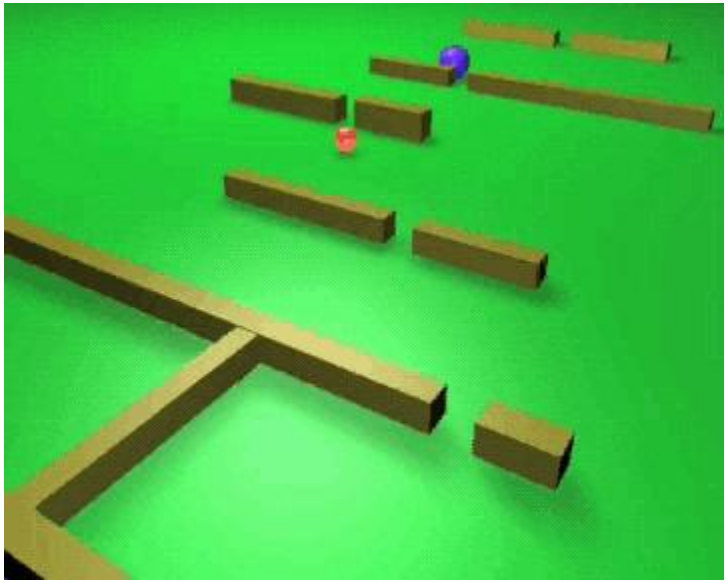




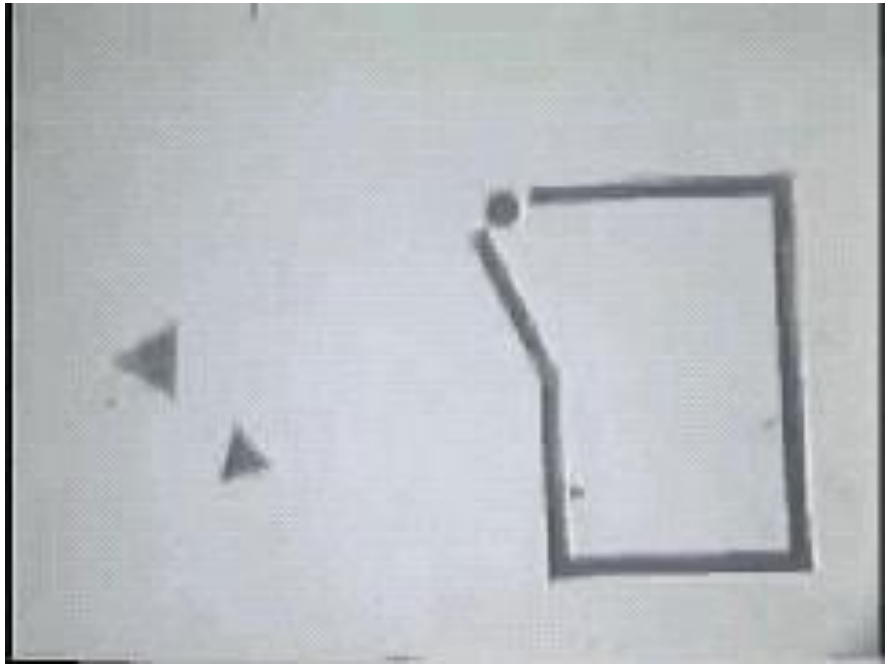
Common-sense understanding

Common-sense understanding

(Southgate and Csibra)

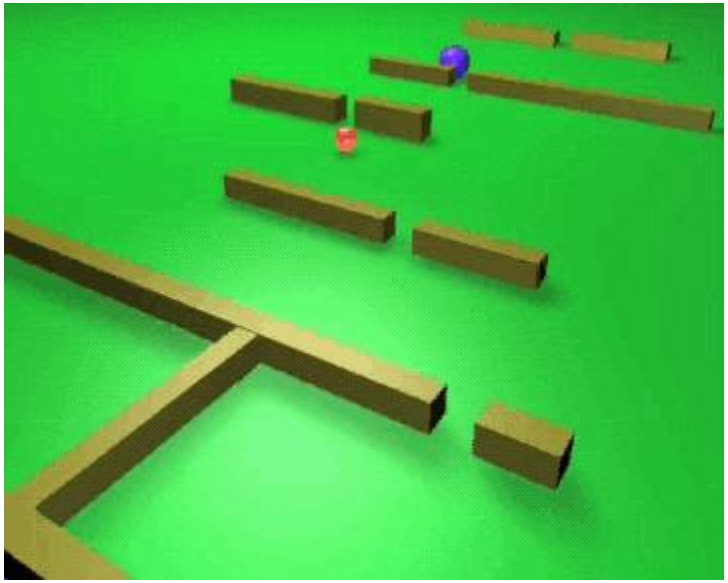


(Heider and Simmel)

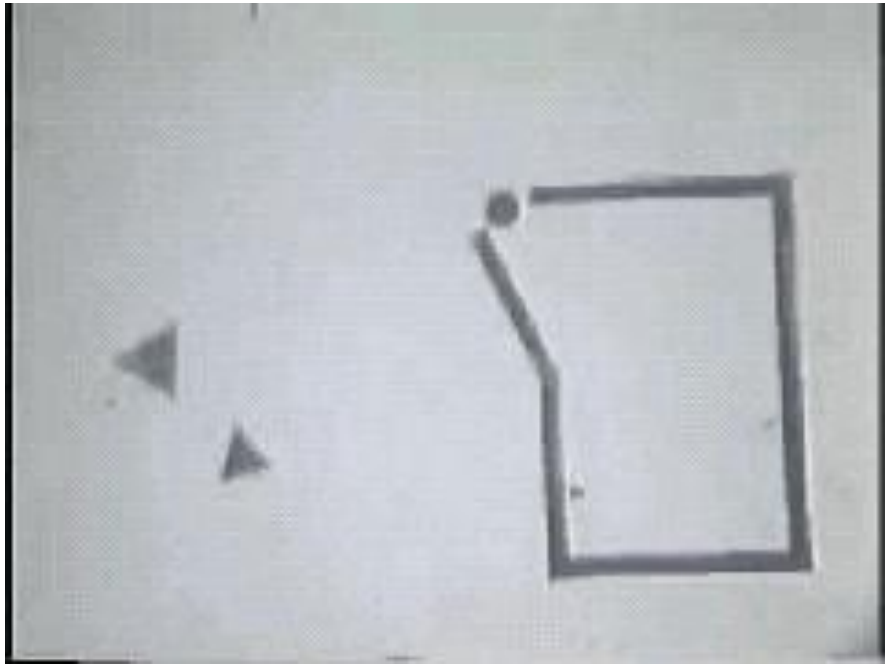


Common-sense understanding

(Southgate and Csibra)



(Heider and Simmel)



Intuitive physics: objects, forces and masses
Intuitive psychology: beliefs and desires
Intuitive sociology: us and them
Intuitive morality: good and bad

… all from 9 or 10 numbers over time.

The approach: statistics meets knowledge

1. How does abstract knowledge guide learning and inference from sparse data?

2. What are the forms and contents of that knowledge?

3. How is that knowledge itself acquired?

The approach: statistics meets knowledge

1. How does abstract knowledge guide learning and inference from sparse data?

   (Approximate) Bayesian inference in probabilistic models.

$$P(h \mid d) = \frac{P(d \mid h)P(h)}{\sum_{h_i \in H} P(d \mid h_i)P(h_i)}$$

2. What are the forms and contents of that knowledge?

   Probabilities defined over a range of structured representations: graphs, grammars, predicate logic, schemas… *programs*.

3. How is that knowledge itself acquired?

   Hierarchical Bayesian models, with inference at multiple levels of abstraction ("learning to learn").

   Learning as (hierarchical Bayesian) program induction.

Is probability even an appropriate basis for modeling cognition?

- Kahneman and Tversky “Heuristics and biases” research program (1970’s-1980’s; 2002 Nobel Prize in Economics).
  - “People aren’t Bayesian.”
- Slovic, Fischhoff, and Lichtenstein (1976): “It appears that people lack the correct programs for many important judgmental tasks.... it may be argued that we have not had the opportunity to evolve an intellect capable of dealing conceptually with uncertainty.”
- Stephen Jay Gould (1992): “Our minds are not built (for whatever reason) to work by the rules of probability.”

Basic cognitive capacities as intuitive Bayesian statistics

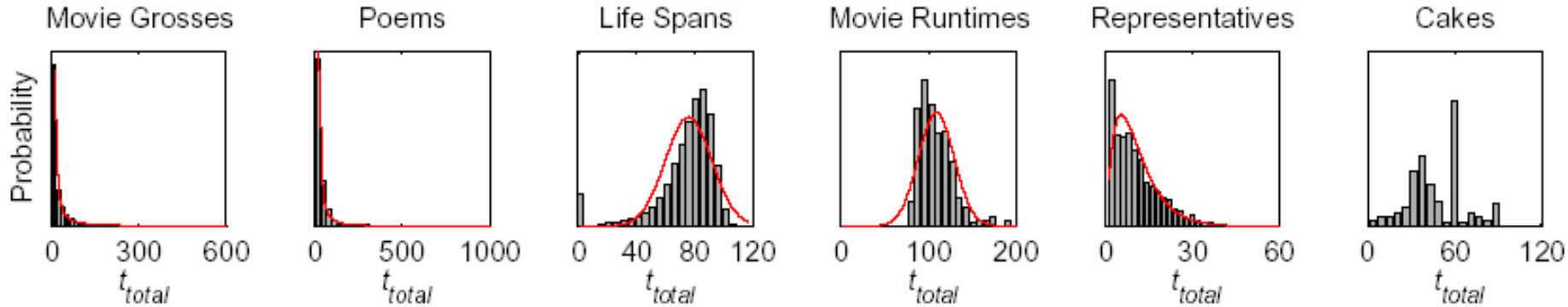
- Similarity (Tenenbaum & Griffiths, *BBS* 2001; Kemp & Tenenbaum, *Cog Sci* 2005)
- Representativeness and evidential support (Tenenbaum & Griffiths, *Cog Sci* 2001)
- Causal judgment (Steyvers et al., 2003; Griffiths & Tenenbaum, *Cog. Psych.* 2005)
- Coincidences and causal discovery (Griffiths & Tenenbaum, *Cog Sci* 2001; *Cognition* 2007; *Psych. Review*, in press)
- Diagnostic inference (Krynski & Tenenbaum, *JEP: General* 2007)
- Predicting the future (Griffiths & Tenenbaum, *Psych. Science* 2006)

Everyday prediction problems

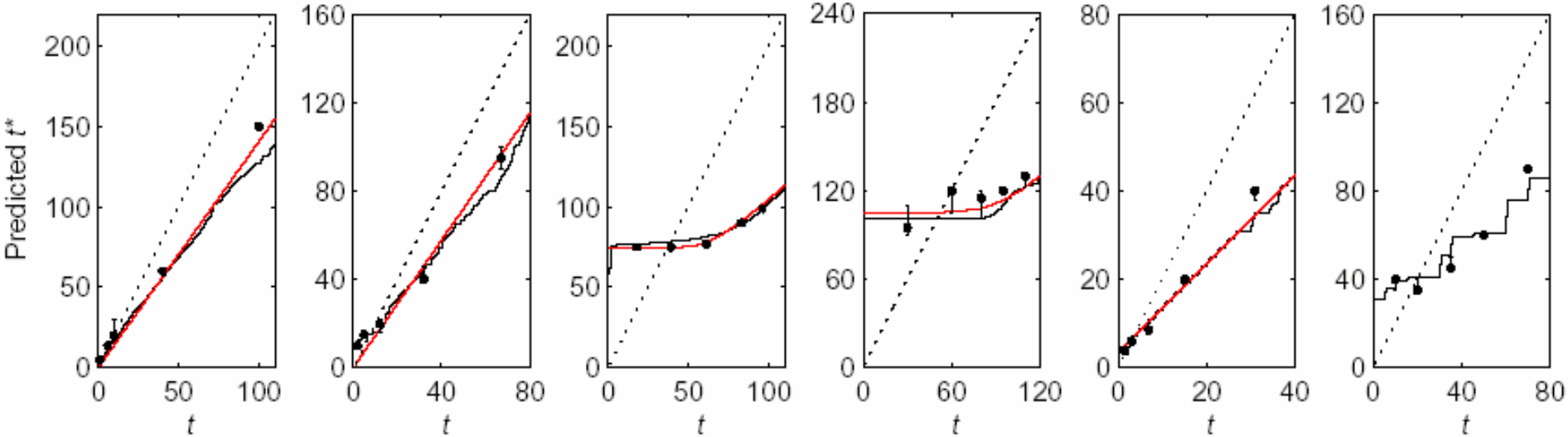
(Griffiths & Tenenbaum, Psych Science 2006)

- You read about a movie that has made $60 million to date. How much money will it make in total?
- You see that something has been baking in the oven for 34 minutes. How long until it's ready?
- You meet someone who is 78 years old. How long will they live?
- Your friend quotes to you from line 17 of his favorite poem. How long is the poem?
- You meet a US congressman who has served for 11 years. How long will he serve in total?
- You encounter a phenomenon or event with an unknown extent or duration, t_{total}, at a random time or value of $t < t_{total}$. What is the total extent or duration t_{total}?

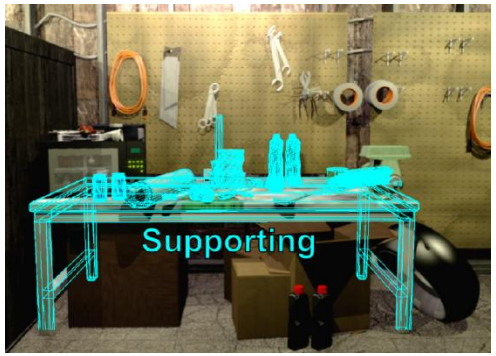
Priors $P(t_{total})$ based on empirically measured durations or magnitudes for many real-world events in each class:

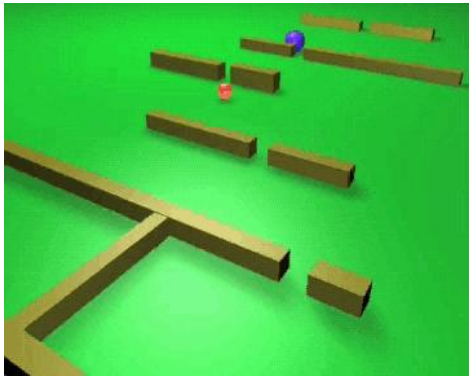


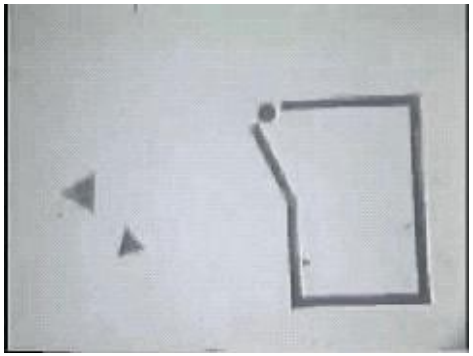
Median human judgments of the total duration or magnitude t_{total} of events in each class, given one random observation at a duration or magnitude t, versus Bayesian predictions (median of $P(t_{total}|t)$).



Scaling up to the hard problems







What is the right prior?
What is the right hypothesis space?
How do people acquire that background knowledge?

Probabilistic graphical models (e.g., Bayes nets)

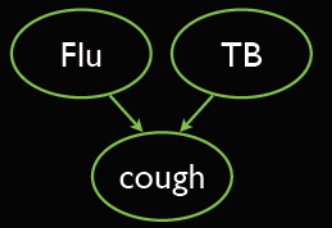
Provide a general language for representing a complex set of probabilistic dependencies in a domain.

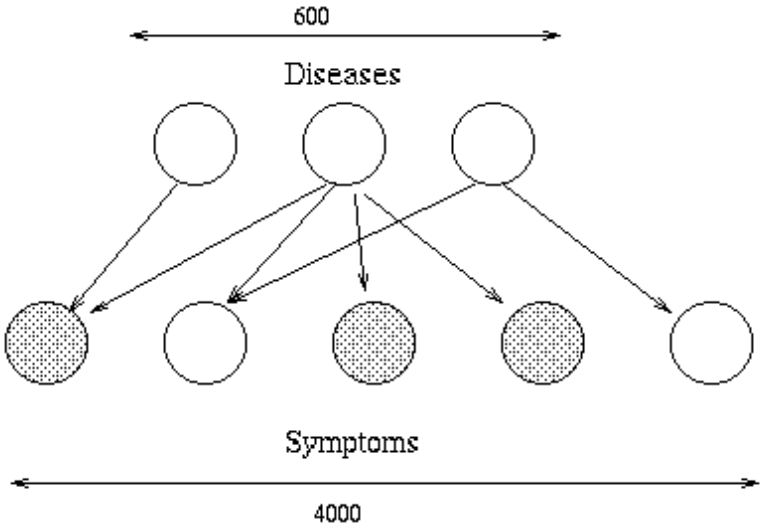
- Graph structure encodes only the essential (irreducible) dependencies, often corresponding to our intuitions about direct causes.

Support general algorithms for explanation, prediction, learning, action planning via probabilistic inference.

P(flu)=0.2
P(TB)=0.01

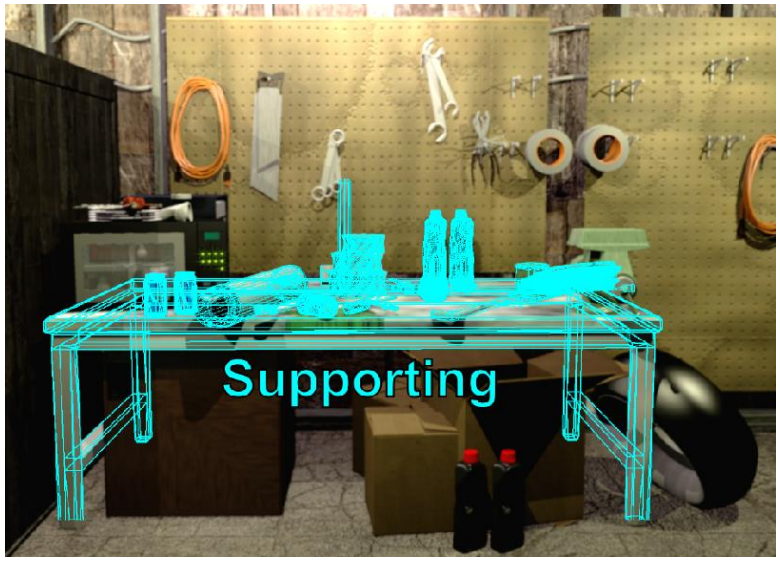
P(cough|flu,TB)	TB	no TB
flu	0.8	0.8
no flu	0.8	0.1

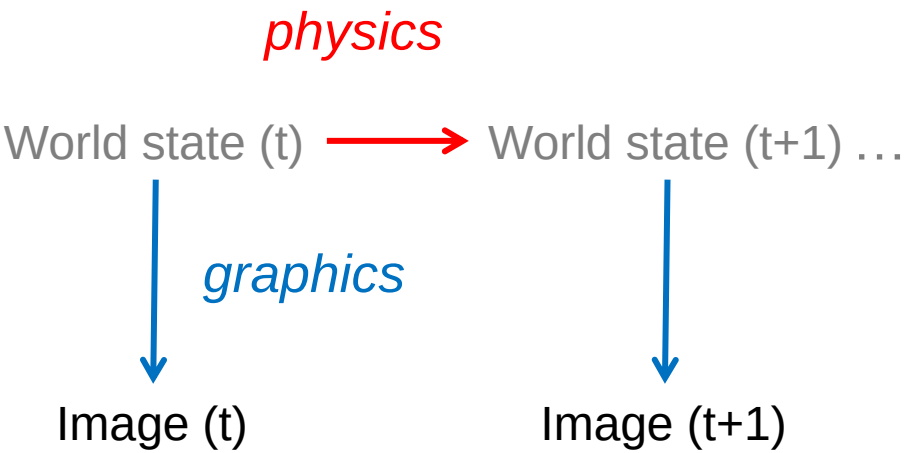




Towards probabilistic programs

A causal graph is too impoverished to capture the causal processes underlying intuitive physical and psychological reasoning. Need a more powerful probabilistic language…

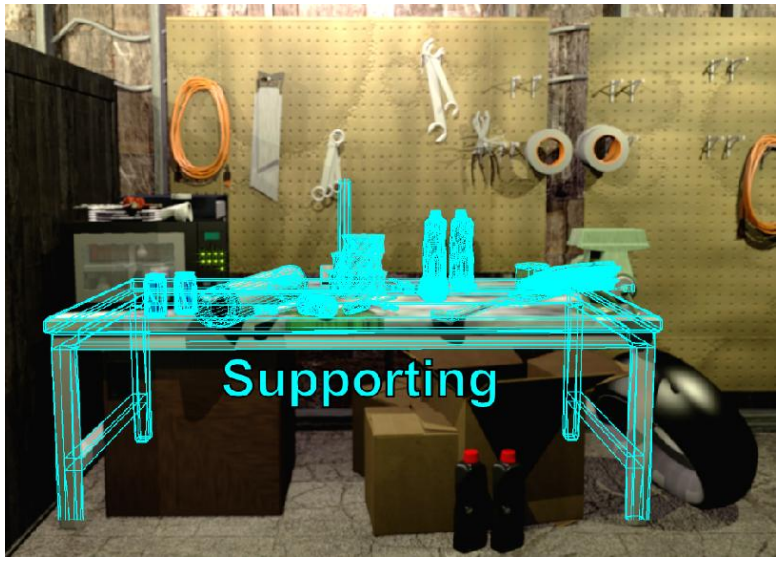


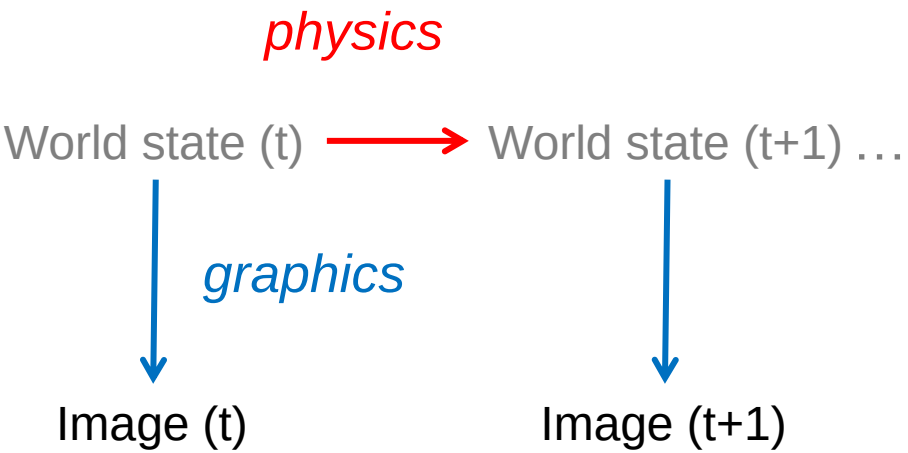


Capture *physics* and *graphics* as *programs*: probabilistic programs that can be run forwards for prediction and planning, and backwards for inference and explanation and learning, via conditional simulation.

Towards probabilistic programs

A causal graph is too impoverished to capture the causal processes underlying intuitive physical and psychological reasoning. Need a more powerful probabilistic language…

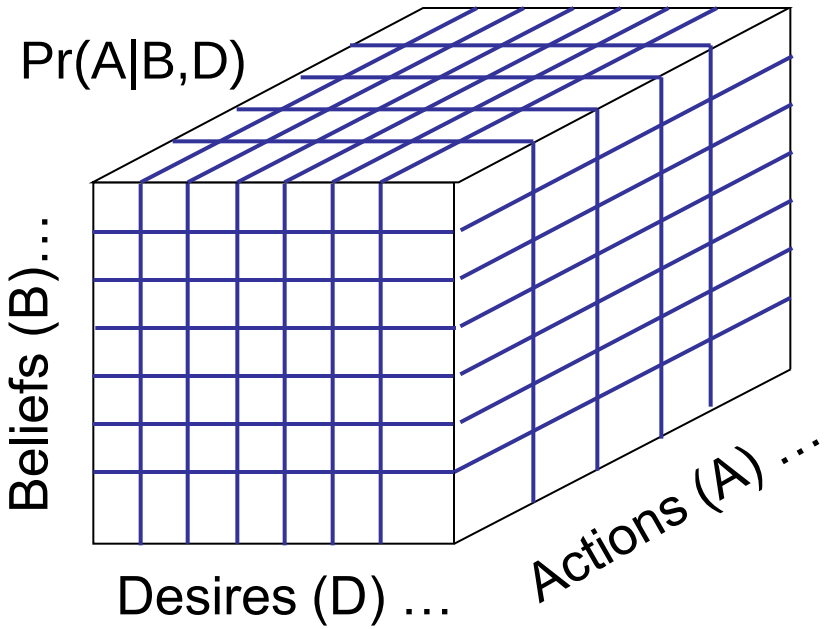


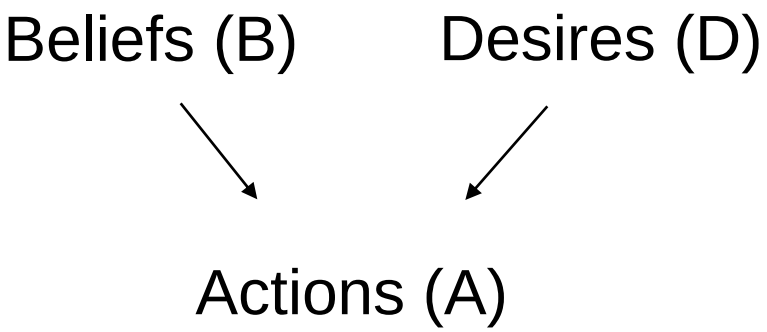


Probabilistic programs: representations for causal processes that are generative, relational, recursive, composable, and computationally universal.

Towards probabilistic programs

A causal graph is too impoverished to capture the causal processes underlying intuitive physical and psychological reasoning. Need a more powerful probabilistic language…

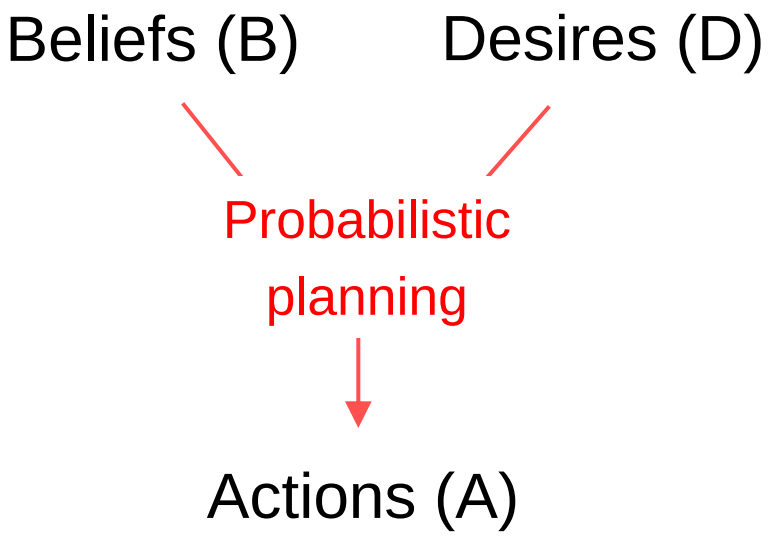




Towards probabilistic programs

A causal graph is too impoverished to capture the causal processes underlying intuitive physical and psychological reasoning. Need a more powerful probabilistic language…

"Intentional agents will tend to choose sequences of actions that they expect (in light of their beliefs) to lead to their desires being achieved, as efficiently and effectively as possible."



Probabilistic programming languages

e.g., Church (Goodman, Mansinghka, et al. 2008): a LISP-like functional language for describing rich generative models as systems of stochastic functions (like Bayes nets, but with arbitrary symbolic nodes and arrows), and performing generic inference on these models (e.g., MCMC).

Causal networks

```
(define cause (mem (lambda (a b)  (flip 0.5))))
(define spontaneous (mem (lambda (a t) (flip 0.01))))
(define do (mem (lambda (a t) (uniform-draw (pair '() (values a))))))
(defi
;;a n
(defi
(defi
  (or
```

Physical objects

```
(define objects (repeat (poisson 1.0) gensym))
(define depth (mem (lambda (object time) (depth object (- time 1)))))
(define location (mem (lambda (ob
                        (+ (drift
(define (drift) (uniform-draw (li
(define extent (mem (lambda (obje
(define (object-seen location tim
              (argmin depth
                      (map (lamb
(define (view location time) (obj
```

Rational agents

```
(define (choose-action state)
  (lex-query
   '((action (action-prior)))
   'action
   '(flip (normalize-reward
            (sample-reward action state)))))

(define (sample-reward action state)
  (let ((next-state (state-transition state action)))
      (+ (reward next-state)
         (if (terminal? next-state)
             0
             (sample-reward
              (choose-action next-state)
              next-state)))))
```

Probabilistic programming languages

e.g., Church (Goodman, Mansinghka, et al. 2008): a LISP-like functional language for describing rich generative models as systems of stochastic functions (like Bayes nets, but with arbitrary symbolic nodes and arrows), and performing generic inference on these models (e.g., MCMC).

```
(define cause (mem (lambda (a b)  (flip 0.5))))
(define spontaneous (mem (lambda (a t) (flip 0.01))))
(define do (mem (lambda (a t) (uniform-draw (pair '() (values a))))))
```

Causal networks

```
(define objects (repeat (poisson 1.0) gensym))

(define depth (mem (lambda (object time) (depth object (- time 1)))))

(define location (mem (lambda (ob
                                (+ (drift
(define (drift) (uniform-draw (li
(define extent (mem (lambda (obje

(define (object-seen location tim
                (argmin depth
                        (map (lamb

(define (view location time) (obj
```

Physical objects

```
(define (choose-action state)
  (lex-query
   '((action (action-prior)))
   'action
   '(flip (normalize-reward
            (sample-reward action state)))))

(define (sample-reward action state)
  (let ((next-state (state-transition state action)))
      (+ (reward next-state)
         (if (terminal? next-state)
             0
             (sample-reward
              (choose-action next-state)
              next-state)))))
```

Rational agents

Precursors: The mid-20th century view of minds as machines

Precursors: The mid-20th century view of minds as machines

Kenneth Craik

(1914-1945)

Precursors: The mid-20th century view of minds as machines

Kenneth Craik

(1914-1945)

The Nature of Explanation (1943):

One of the most fundamental properties of thought is its power of predicting events…. It enables us, for instance, to design bridges with a sufficient factor of safety instead of building them haphazard and waiting to see whether they collapse… If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it. Most of the greatest advances of modern technology have been instruments which extended the scope of our sense-organs, our brains or our limbs. Such are telescopes and microscopes, wireless, calculating machines, typewriters, motor cars, ships and aeroplanes. Is it not possible, therefore, that our brains themselves utilize comparable mechanisms to achieve the same ends and that these mechanisms can parallel phenomena in the external world as a calculating machine can parallel the development of strains in a bridge?

Precursors: The mid-20th century view of minds as machines

Kenneth Craik

(1914-1945)

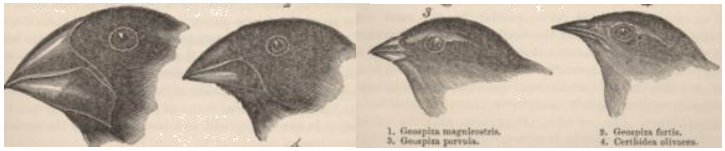
The Nature of Explanation (1943):

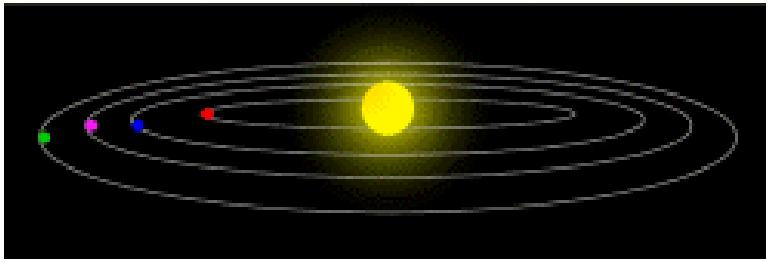
One of the most fundamental properties of thought is its power of predicting events…. It enables us, for instance, to design bridges with a sufficient factor of safety instead of building them haphazard and waiting to see whether they collapse… If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, aling with the present and ch fuller, safer, and more hich face it. Most of the have been instruments rgans, our brains or our pes, wireless, calculating and aeroplanes. Is it not elves utilize comparable nd that these mechanisms orld as a calculating strains in a bridge?

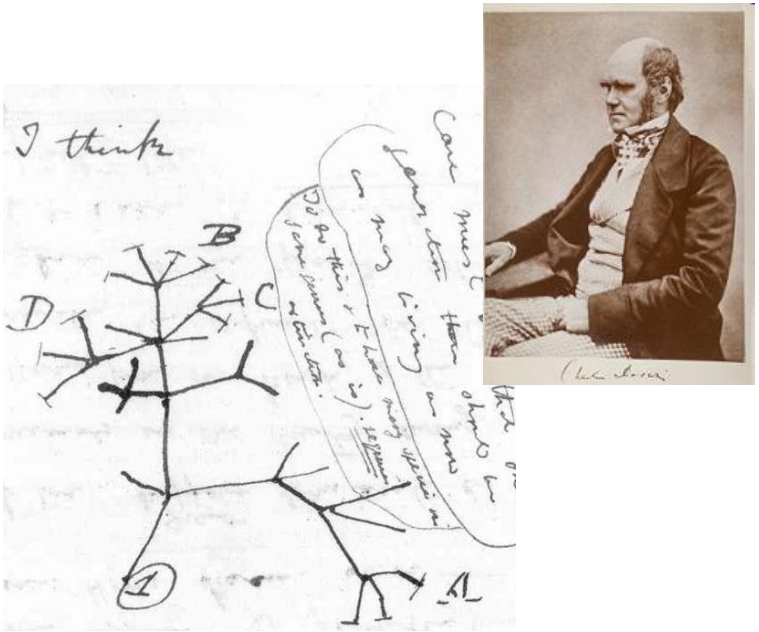
Some later 20th century AI precursors – "mental models", "analog reasoning", "qualitative reasoning":

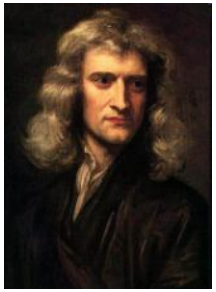
Forbus, Gentner, de Kleer & Brown, Funt, Gardin & Meltzer, Johnson-laird, Kuipers, Langley, Winston, Sloman, …

Precursors: Theory construction as probabilistic program induction

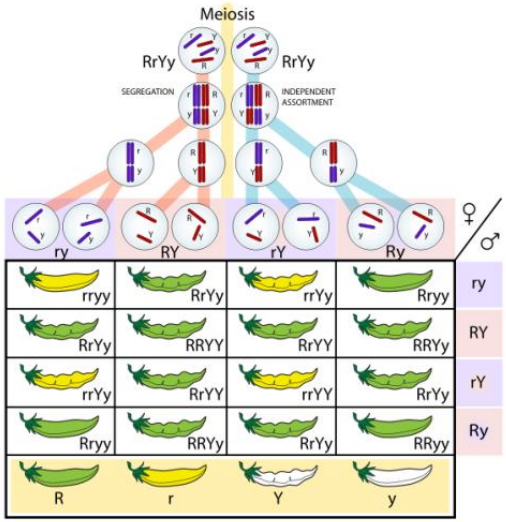








$$F = \frac{GMm}{r^2}$$



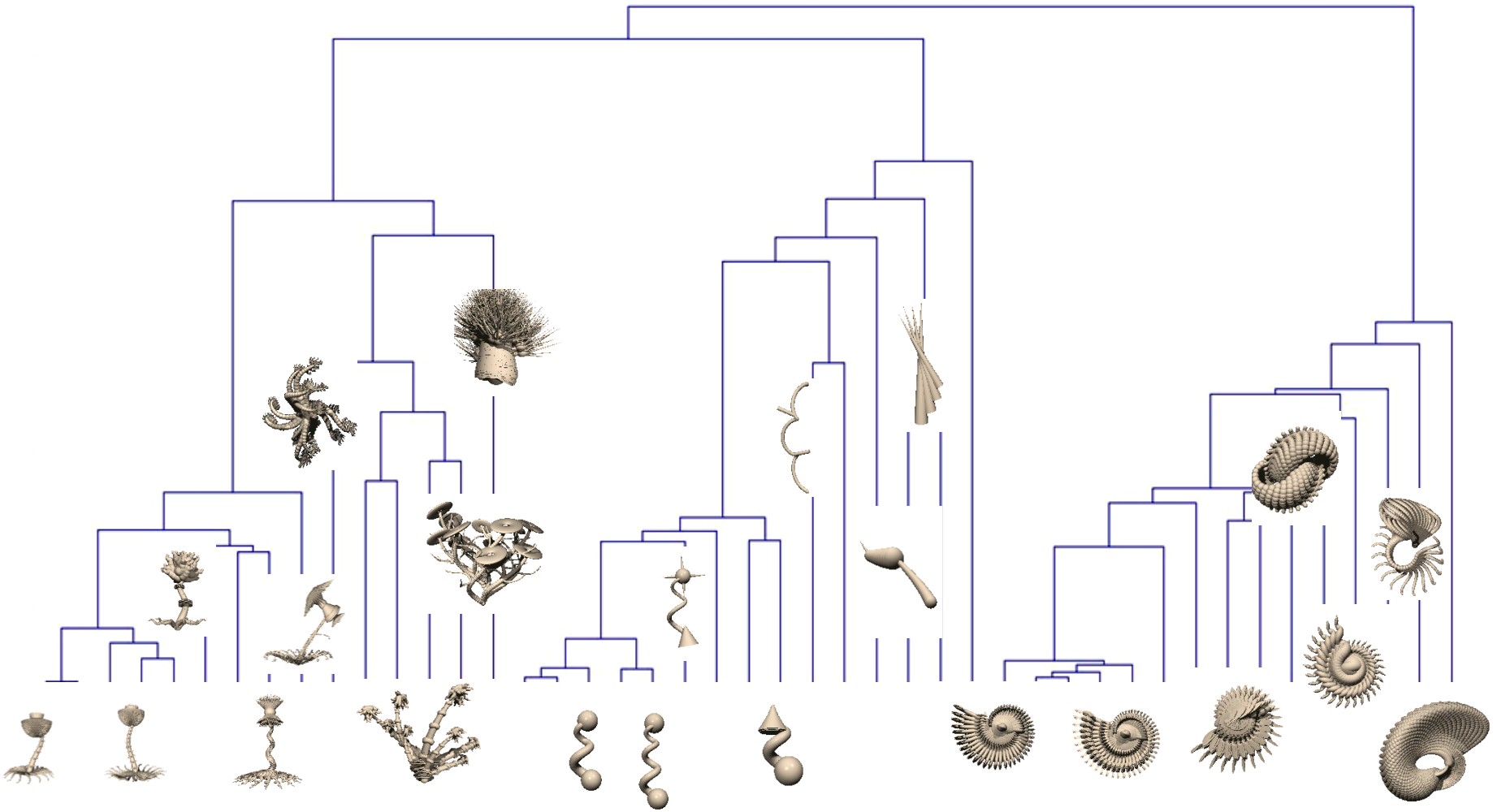
C.f. Ross King on formalizing theories in science

Learning from very few examples

(Xu & Tenenbaum, *Psych Review*, 2007)

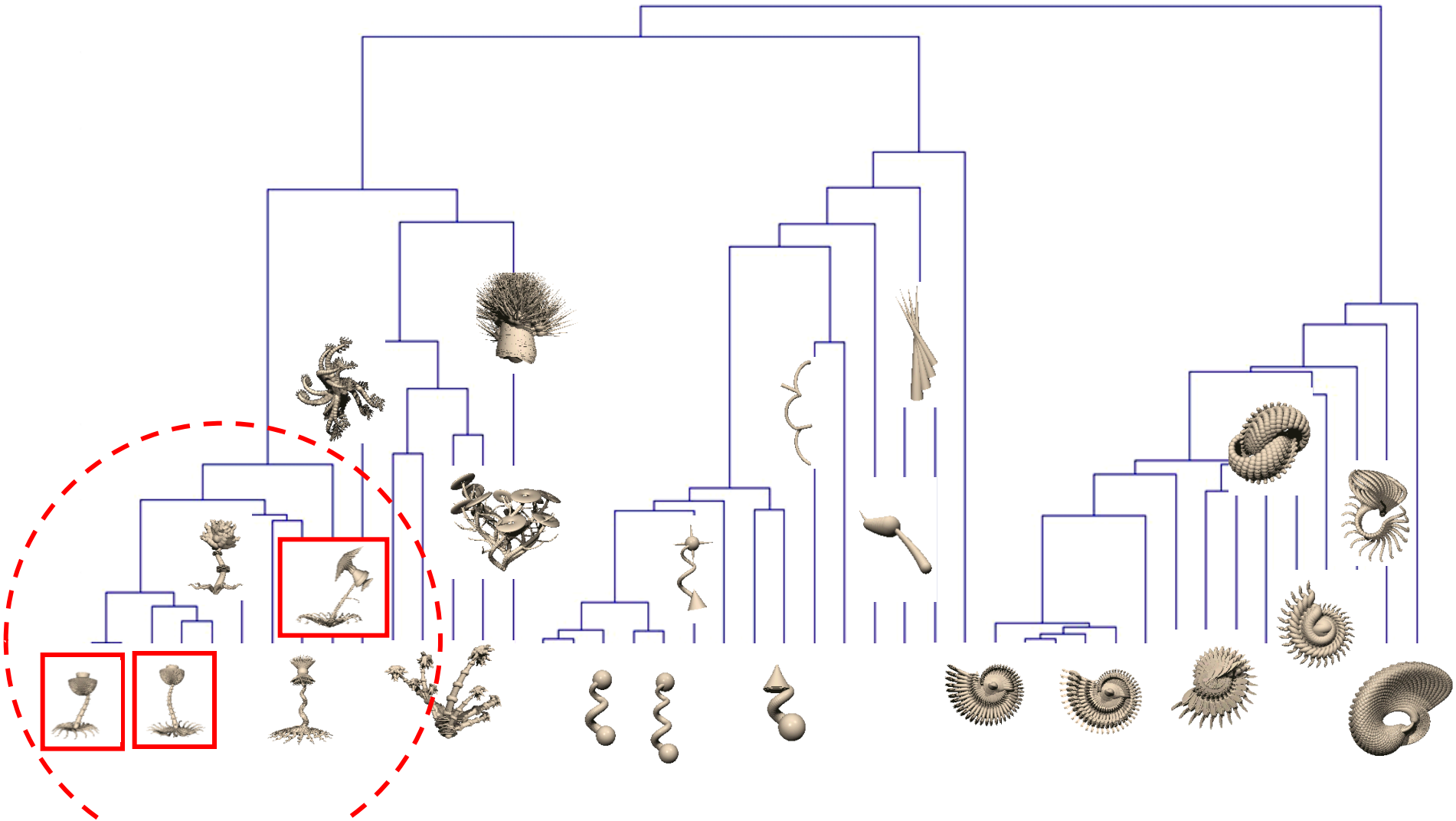
Learning from very few examples

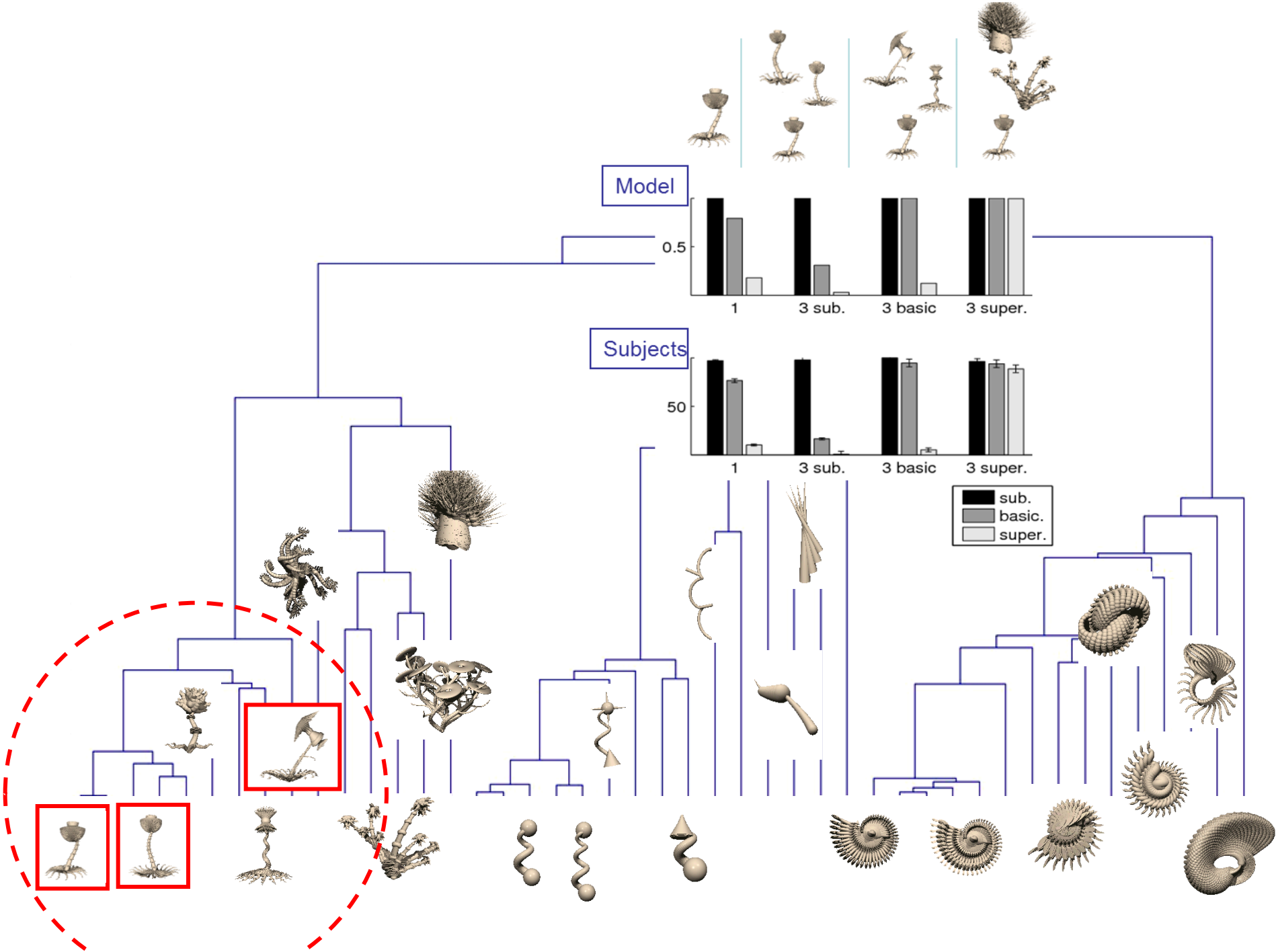
(Xu & Tenenbaum, *Psych Review*, 2007)



Learning from very few examples

(Xu & Tenenbaum, *Psych Review*, 2007)



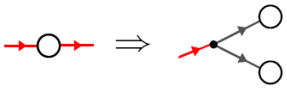

Model
0.5
1
3 sub.
3 basic
3 super.
Subjects
50
1
3 sub.
3 basic
3 super.
sub.
basic.
super.

Discovering the structural form of a domain

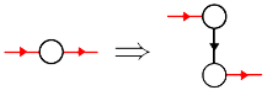
(Kemp & Tenenbaum, *PNAS* 2008; *Psych Review*, 2009)

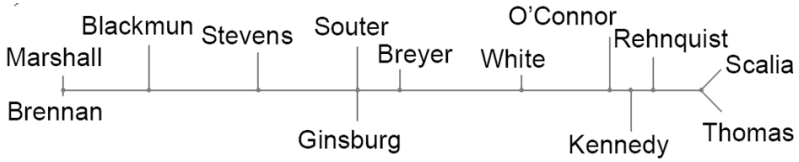
Abstract principles

tree:

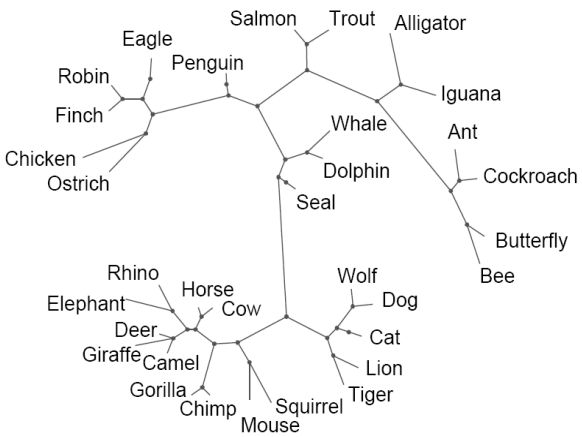


chain:

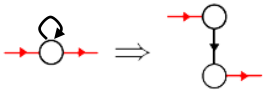


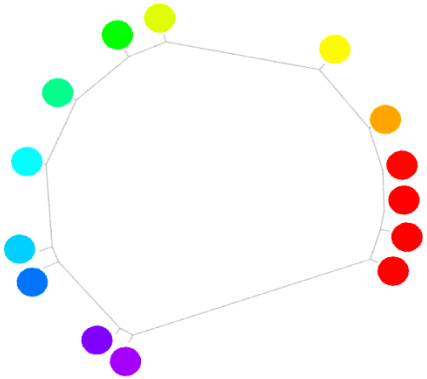


Model

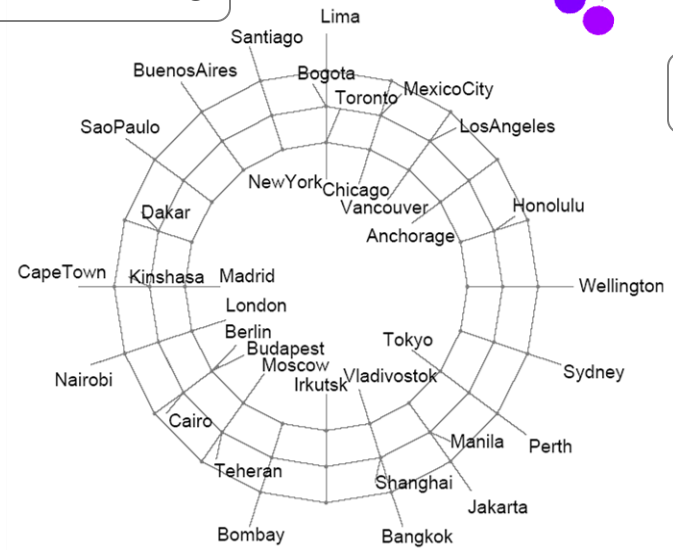


ring:

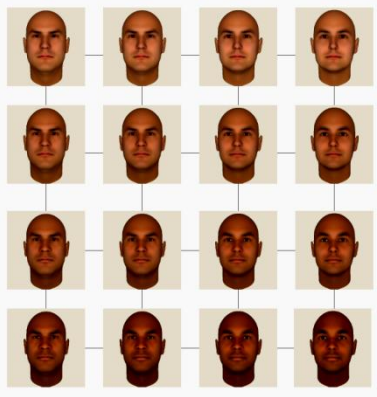




chain x ring



chain x chain



Data

features

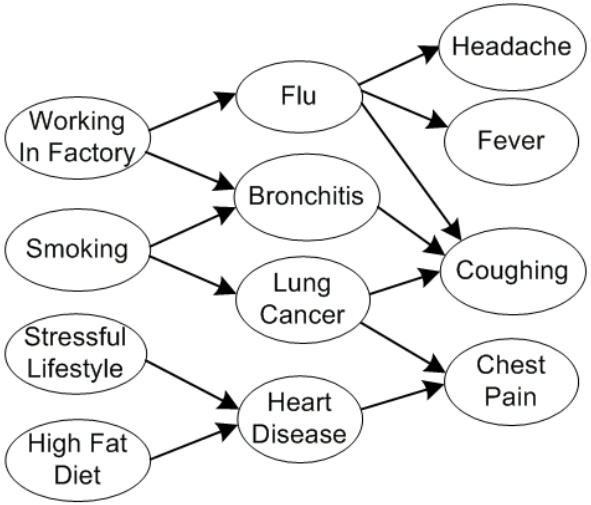
animals

Causal learning and reasoning

Causal model

↓

Event data



Patient 1: Stressful lifestyle
Chest Pain

Patient 2: Smoking
Coughing

Patient 3: Working in factory
Chest Pain

. . .

(Mansinghka, Kemp, Tenenbaum, Griffiths, UAI 2006)

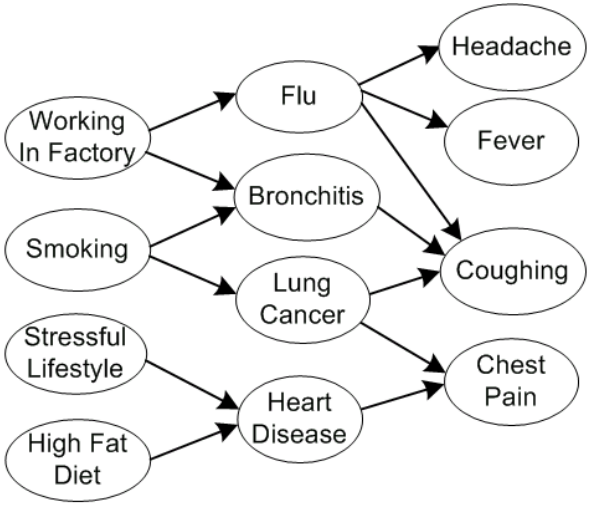
Causal learning and reasoning

Causal schema

Behaviors → Diseases → Symptoms

Behaviors	Diseases	Symptoms
high-fat diet working in factory …	heart disease lung cancer …	coughing chest pain …

Causal model



Event data

Patient 1: Stressful lifestyle
Chest Pain

Patient 2: Smoking
Coughing

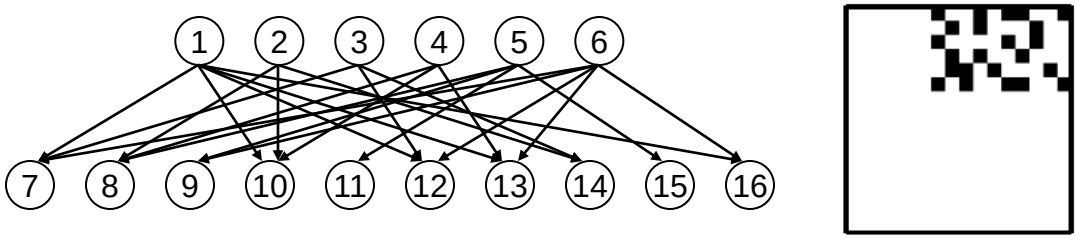
Patient 3: Working in factory
Chest Pain

. . .

Cut down hypothesis space from size 521,939,651,343,829, 405,020,504,063 to 131,072

(Mansinghka, Kemp, Tenenbaum, Griffiths, UAI 2006)

Ground-truth causal network



Causal model

↓

Event data

20 | 80 | 1000 | # samples

recovered model

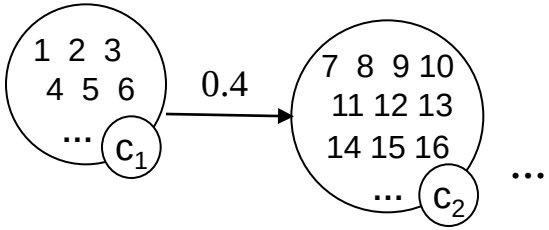
Causal schema

↓

Causal model

↓

Event data

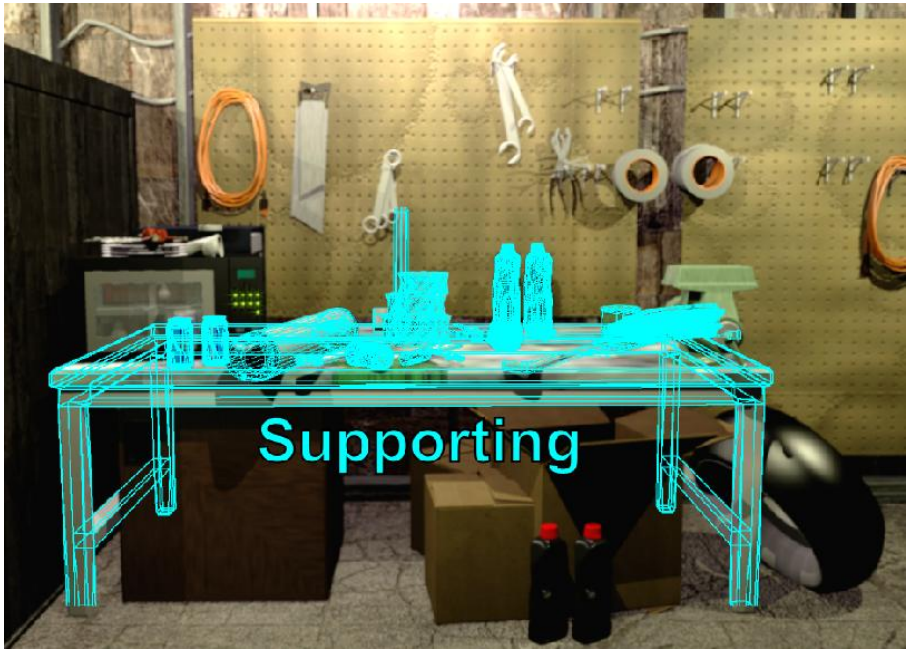


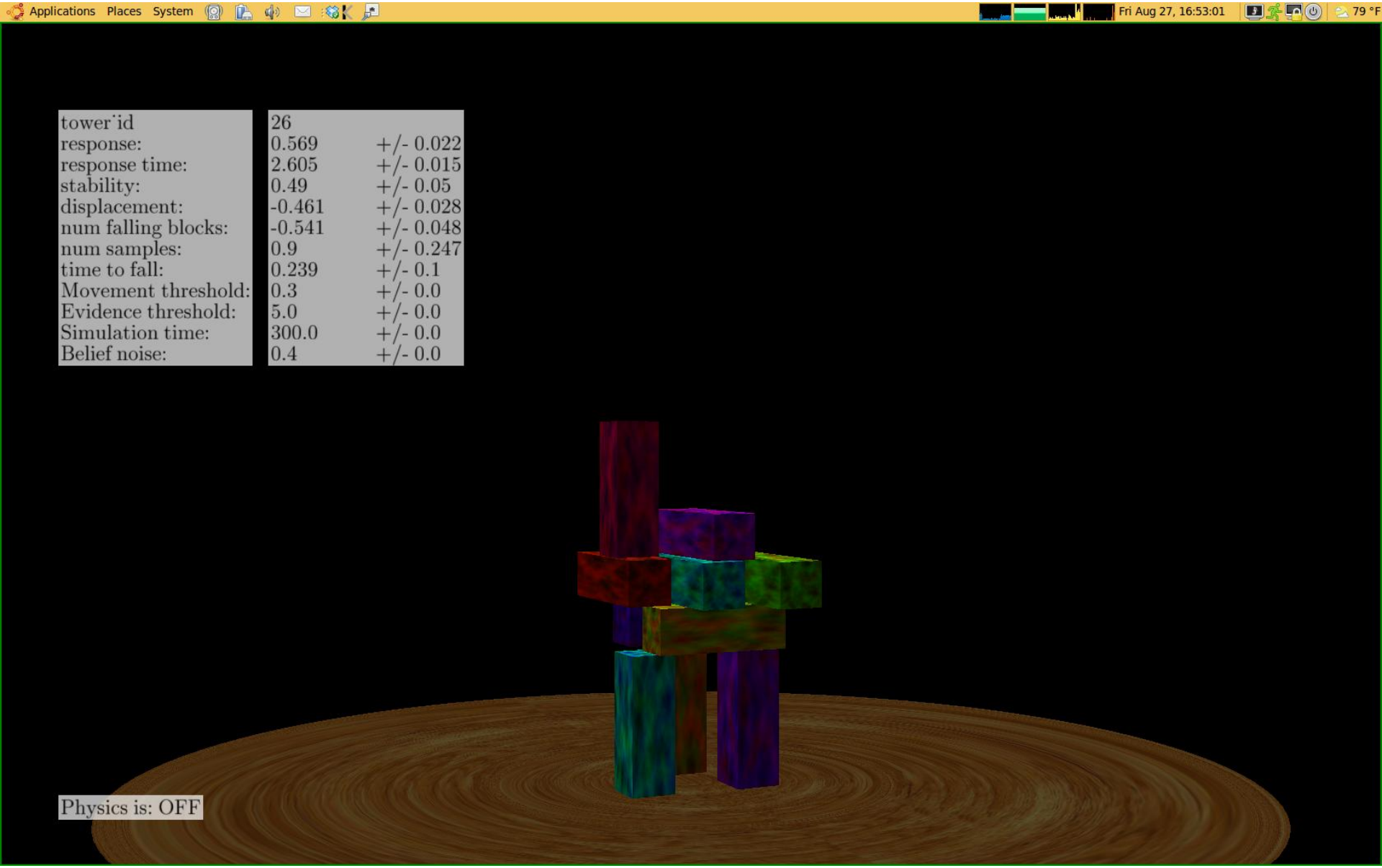
"blessing of abstraction"

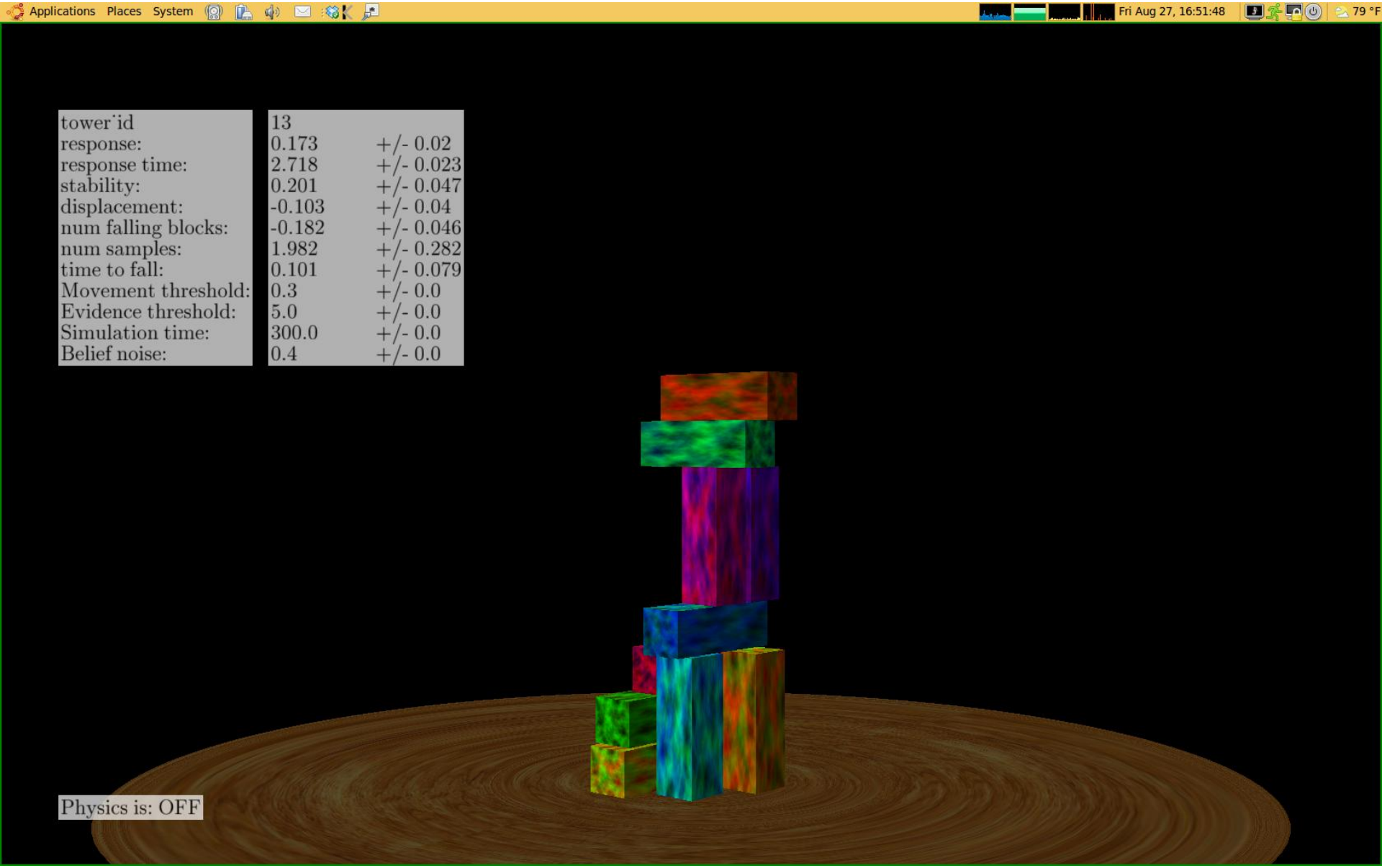
recovered model

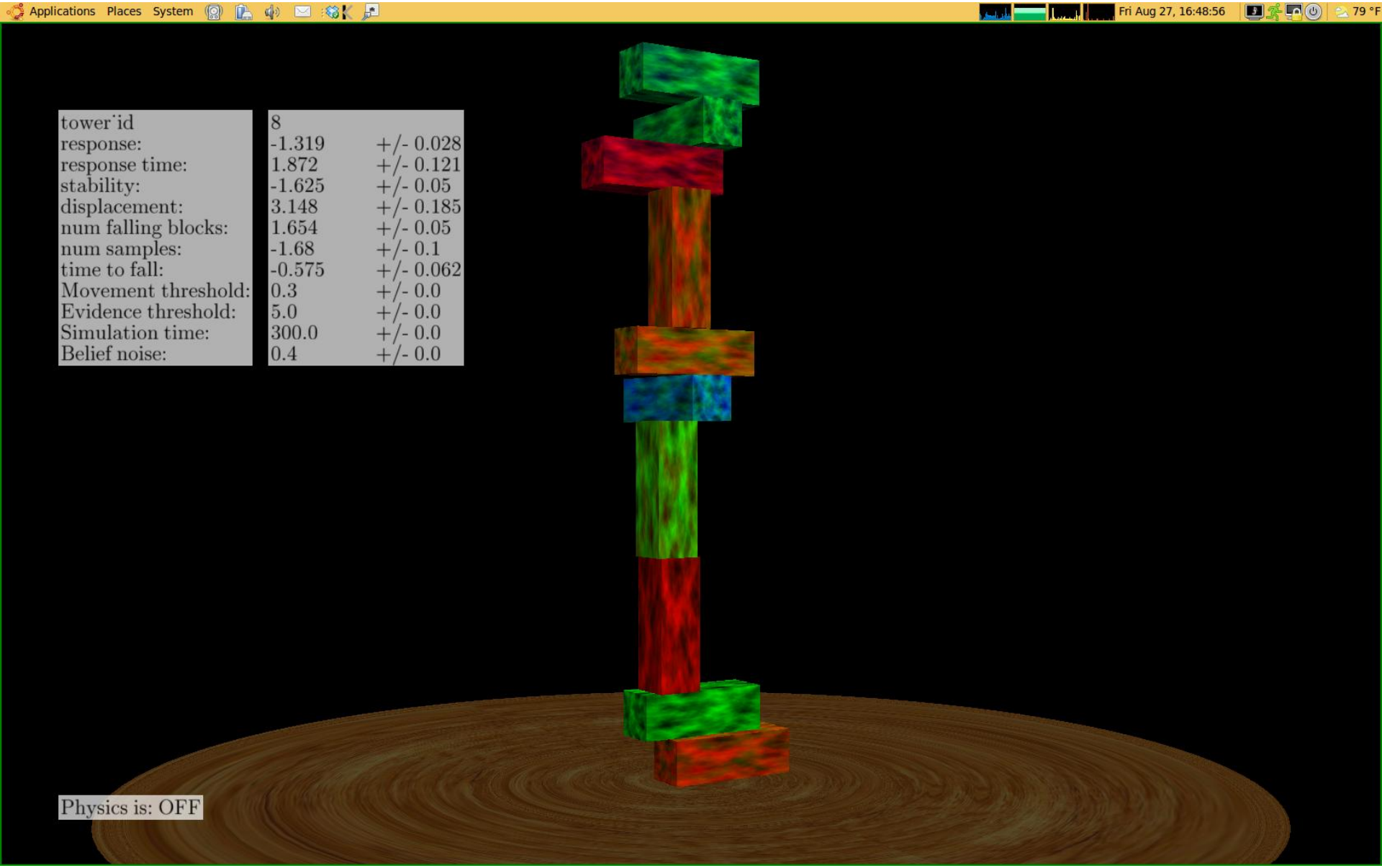
(Mansinghka, Kemp, Tenenbaum, Griffiths, UAI 2006)

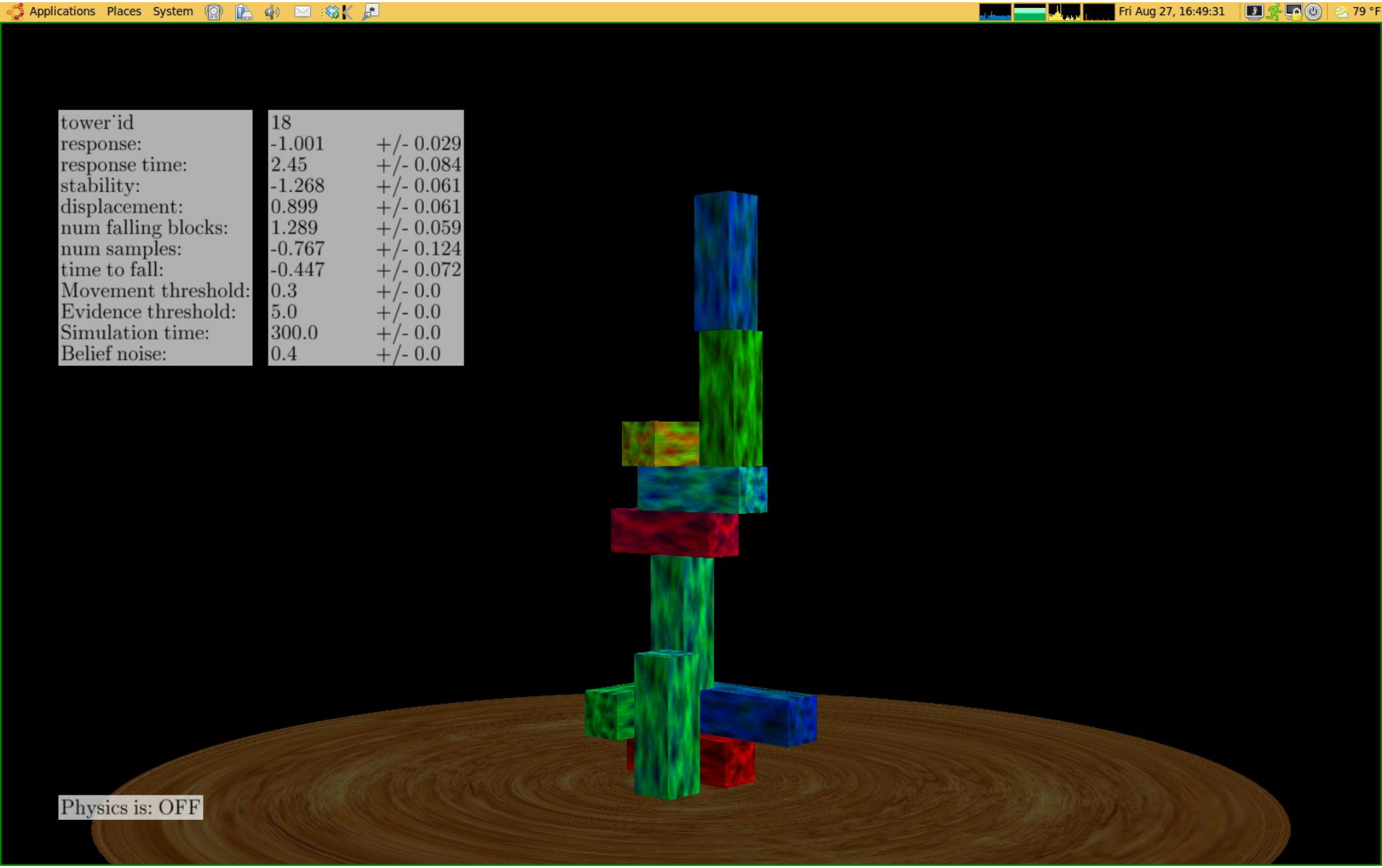
Intuitive physics: stability

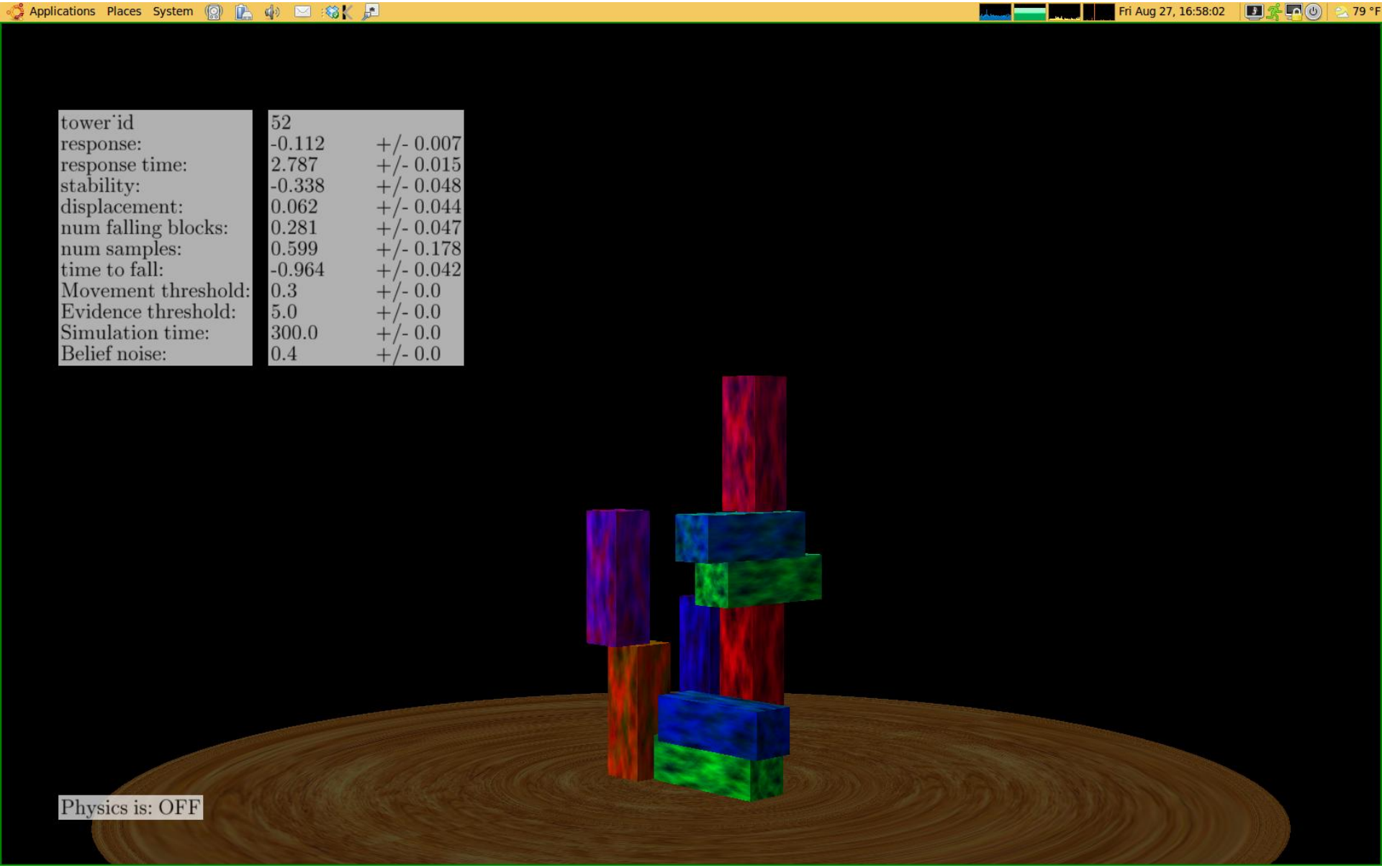


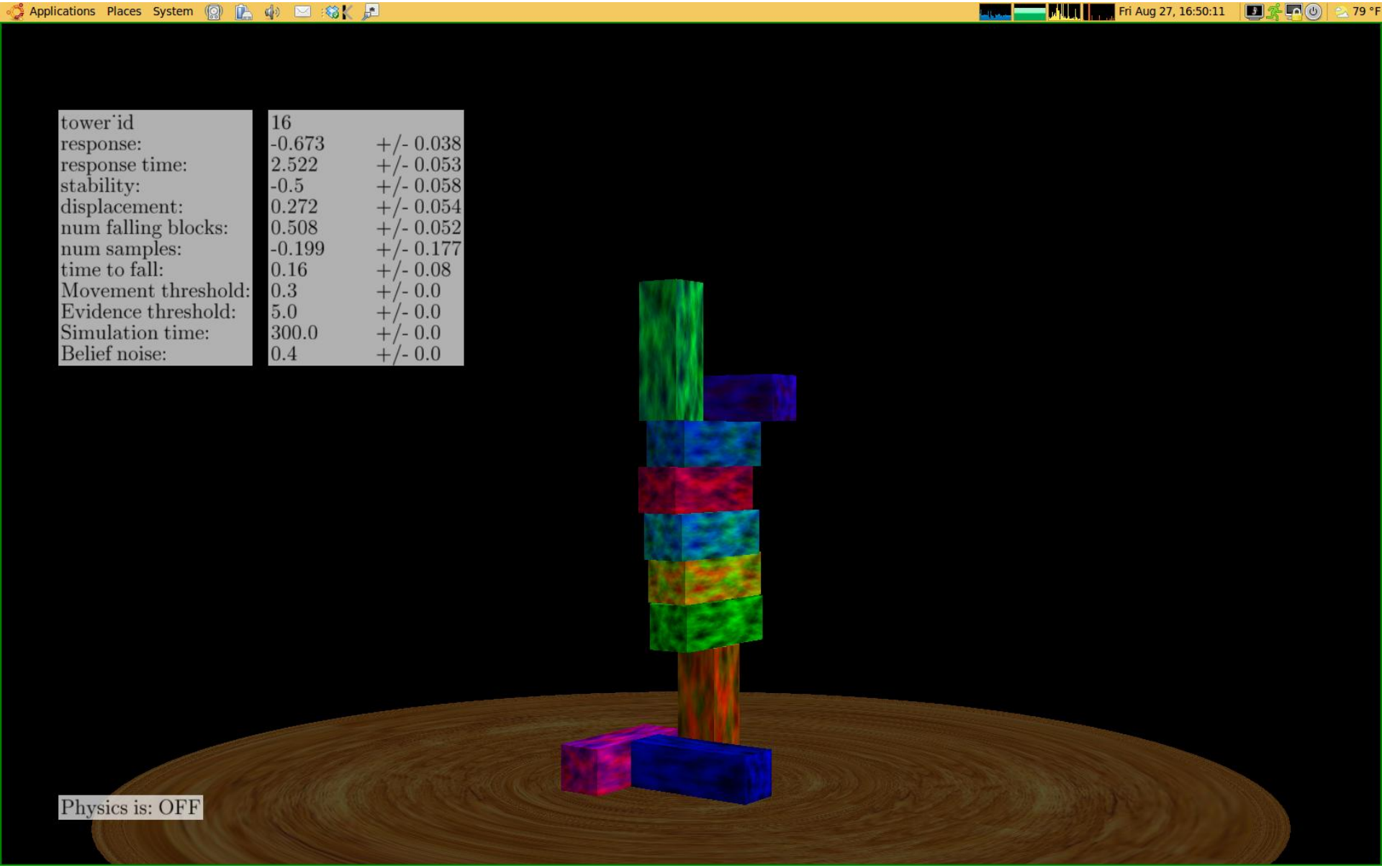

Applications Places System
Fri Aug 27, 16:53:01
79 °F
tower id 26
response: 0.569 +/- 0.022
response time: 2.605 +/- 0.015
stability: 0.49 +/- 0.05
displacement: -0.461 +/- 0.028
num falling blocks: -0.541 +/- 0.048
num samples: 0.9 +/- 0.247
time to fall: 0.239 +/- 0.1
Movement threshold: 0.3 +/- 0.0
Evidence threshold: 5.0 +/- 0.0
Simulation time: 300.0 +/- 0.0
Belief noise: 0.4 +/- 0.0
Physics is: OFF


Applications Places System
Fri Aug 27, 16:51:48
79 °F
tower id 13
response: 0.173 +/- 0.02
response time: 2.718 +/- 0.023
stability: 0.201 +/- 0.047
displacement: -0.103 +/- 0.04
num falling blocks: -0.182 +/- 0.046
num samples: 1.982 +/- 0.282
time to fall: 0.101 +/- 0.079
Movement threshold: 0.3 +/- 0.0
Evidence threshold: 5.0 +/- 0.0
Simulation time: 300.0 +/- 0.0
Belief noise: 0.4 +/- 0.0
Physics is: OFF


Applications Places System
Fri Aug 27, 16:48:56
79 °F
tower id 8
response: -1.319 +/- 0.028
response time: 1.872 +/- 0.121
stability: -1.625 +/- 0.05
displacement: 3.148 +/- 0.185
num falling blocks: 1.654 +/- 0.05
num samples: -1.68 +/- 0.1
time to fall: -0.575 +/- 0.062
Movement threshold: 0.3 +/- 0.0
Evidence threshold: 5.0 +/- 0.0
Simulation time: 300.0 +/- 0.0
Belief noise: 0.4 +/- 0.0
Physics is: OFF


Applications Places System
Fri Aug 27, 16:49:31
79 °F
tower id 18
response: -1.001 +/- 0.029
response time: 2.45 +/- 0.084
stability: -1.268 +/- 0.061
displacement: 0.899 +/- 0.061
num falling blocks: 1.289 +/- 0.059
num samples: -0.767 +/- 0.124
time to fall: -0.447 +/- 0.072
Movement threshold: 0.3 +/- 0.0
Evidence threshold: 5.0 +/- 0.0
Simulation time: 300.0 +/- 0.0
Belief noise: 0.4 +/- 0.0
Physics is: OFF


Applications Places System
Fri Aug 27, 16:58:02
79 °F
tower id 52
response: -0.112 +/- 0.007
response time: 2.787 +/- 0.015
stability: -0.338 +/- 0.048
displacement: 0.062 +/- 0.044
num falling blocks: 0.281 +/- 0.047
num samples: 0.599 +/- 0.178
time to fall: -0.964 +/- 0.042
Movement threshold: 0.3 +/- 0.0
Evidence threshold: 5.0 +/- 0.0
Simulation time: 300.0 +/- 0.0
Belief noise: 0.4 +/- 0.0
Physics is: OFF


Applications Places System
Fri Aug 27, 16:50:11
79 °F
tower id 16
response: -0.673 +/- 0.038
response time: 2.522 +/- 0.053
stability: -0.5 +/- 0.058
displacement: 0.272 +/- 0.054
num falling blocks: 0.508 +/- 0.052
num samples: -0.199 +/- 0.177
time to fall: 0.16 +/- 0.08
Movement threshold: 0.3 +/- 0.0
Evidence threshold: 5.0 +/- 0.0
Simulation time: 300.0 +/- 0.0
Belief noise: 0.4 +/- 0.0
Physics is: OFF

Modeling stability judgments

(Battaglia, Hamrick et al.)

Prob.
approx.
Newton

… World state (t-1) → World state (t) → World state (t+1) …

Prob. approx. rendering

Image (t-1)

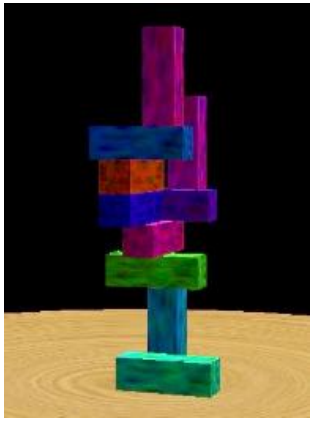


Image (t)

Image (t+1)

Modeling stability judgments

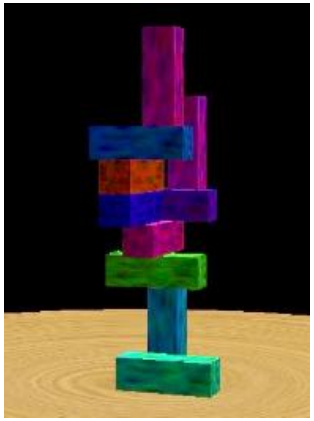
(Battaglia, Hamrick et al.)

Prob.
approx.
Newton

… World state (t-1) → World state (t) → World state (t+1) …

Prob. approx. rendering

Image (t-1) Image (t) Image (t+1)



Modeling stability judgments

(Battaglia, Hamrick et al.)

Prob.
approx.
Newton

… World state (t-1) → World state (t) → World state (t+1) …

Prob. approx. rendering

Image (t-1)

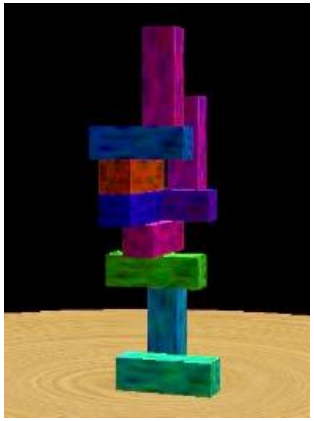


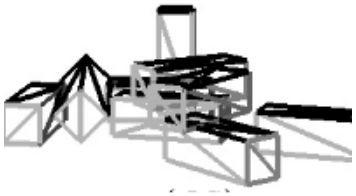
Image (t)

Image (t+1)

Modeling stability judgments

(Battaglia, Hamrick et al.)

Prob.
approx.
Newton



… World state (t-1) → World state (t) → World state (t+1) …

Prob. approx. rendering

Image (t-1)

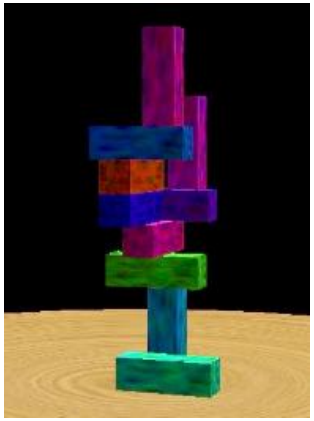


Image (t)

Image (t+1)

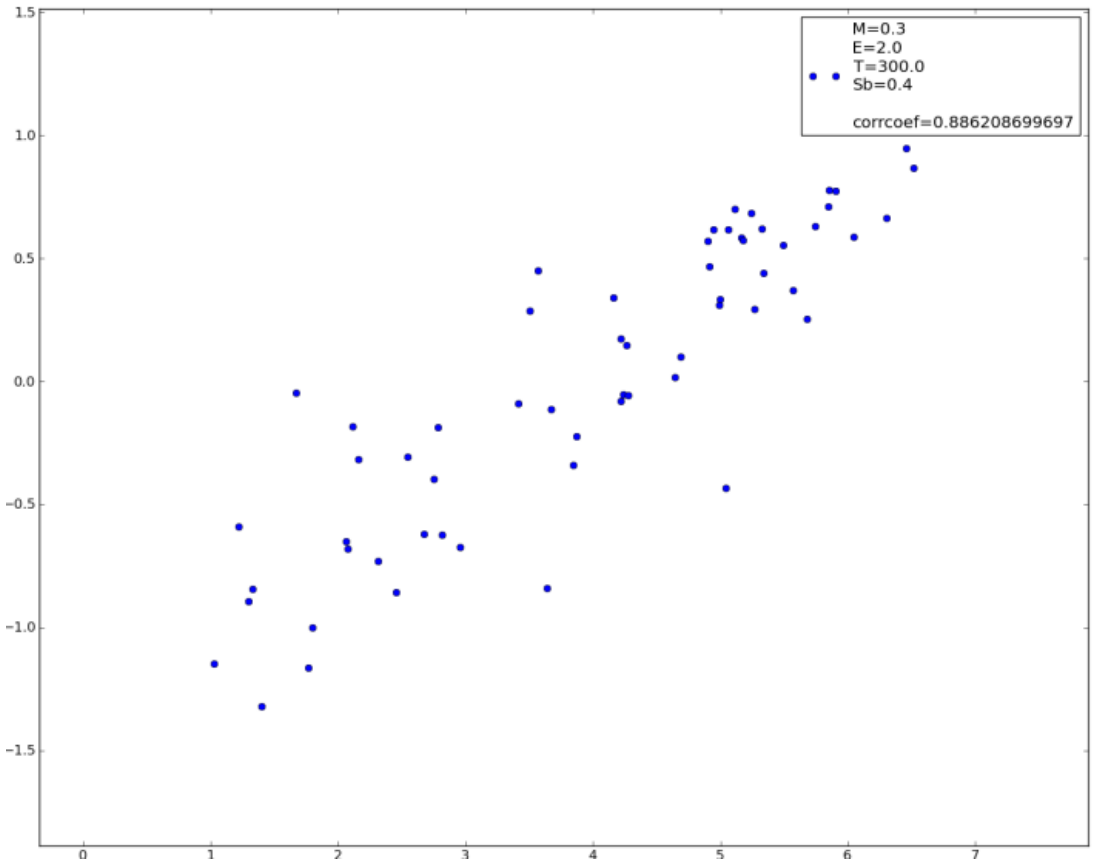
σ = state uncertainty

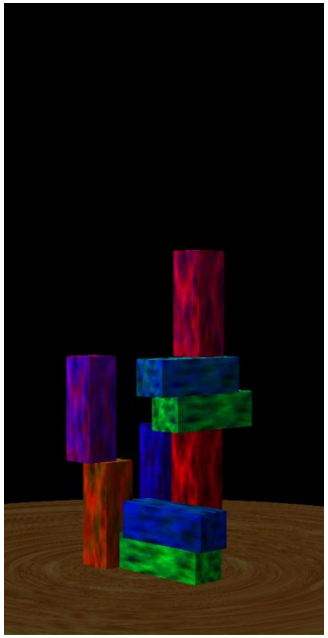
τ = latent force magnitude

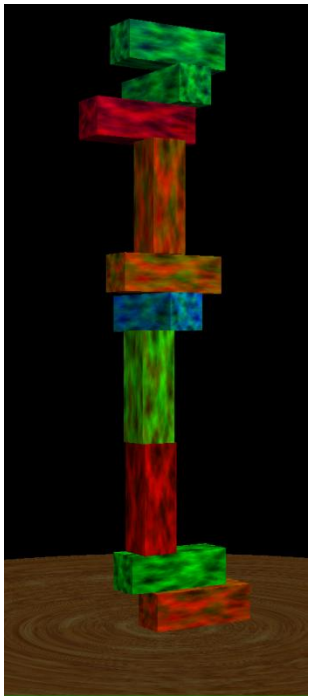
Modeling stability judgments

(Battaglia, Hamrick et al.)

Mean human stability judgment







Model prediction

(expected proportion of tower that will fall)

$\sigma = 0.05$

$\tau = 0$

Physical simulation, or visual/geometric heuristic?

The flexibility of common sense
(c.f. “infinite use of finite means”, Turing test)

Will the tower fall? What proportion of the tower will fall?

Which way will the tower fall? How far will the blocks fall?

If you bump the table, will more red blocks or yellow blocks fall over? ... What if you bump this side rather than that side? … What if the blocks have different shapes and sizes? ... What if there are rails here, but not there? …

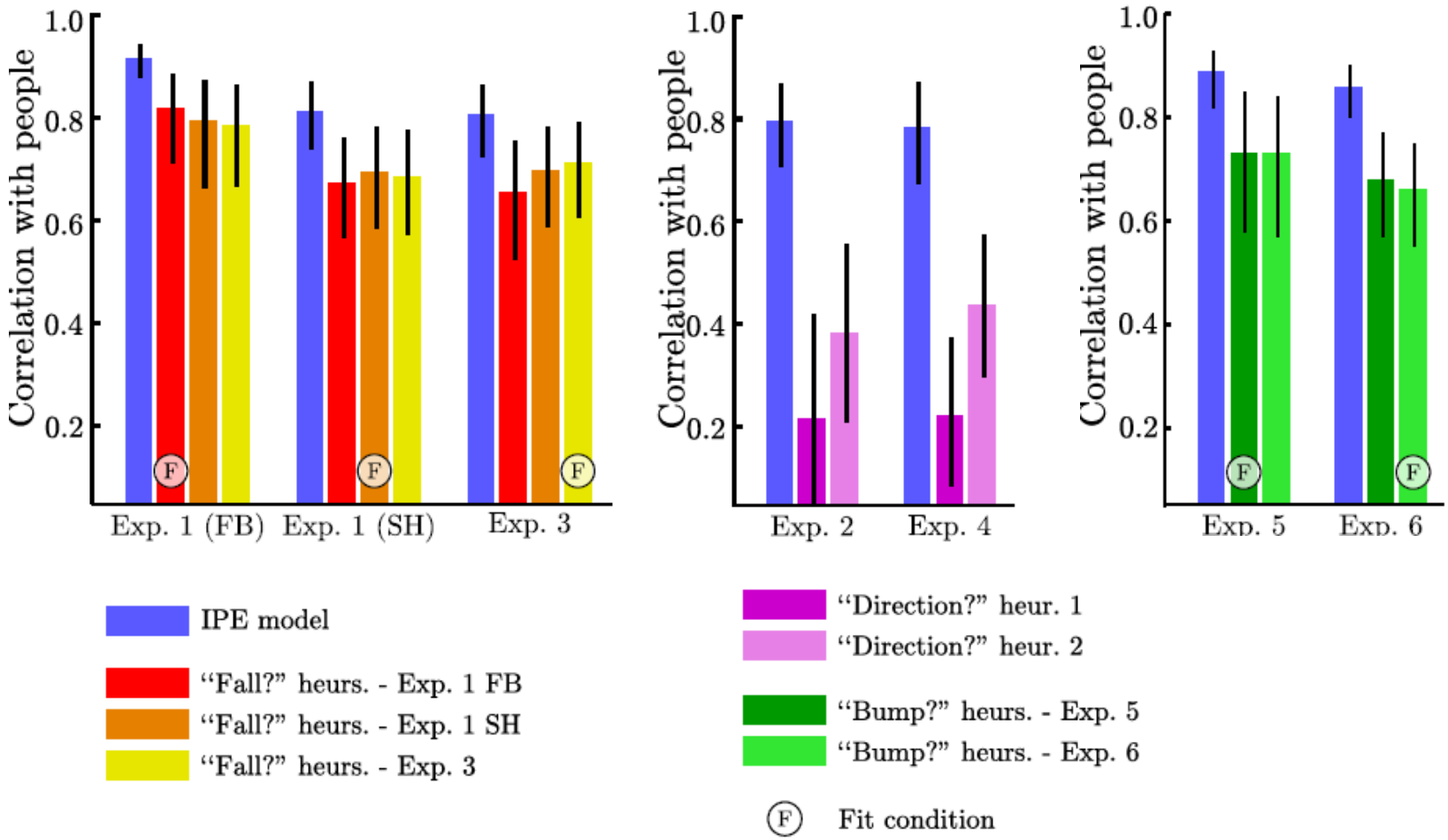
Did adding this block cause the tower to fall?

If this block had (not) been present, would the tower (still) have fallen over?

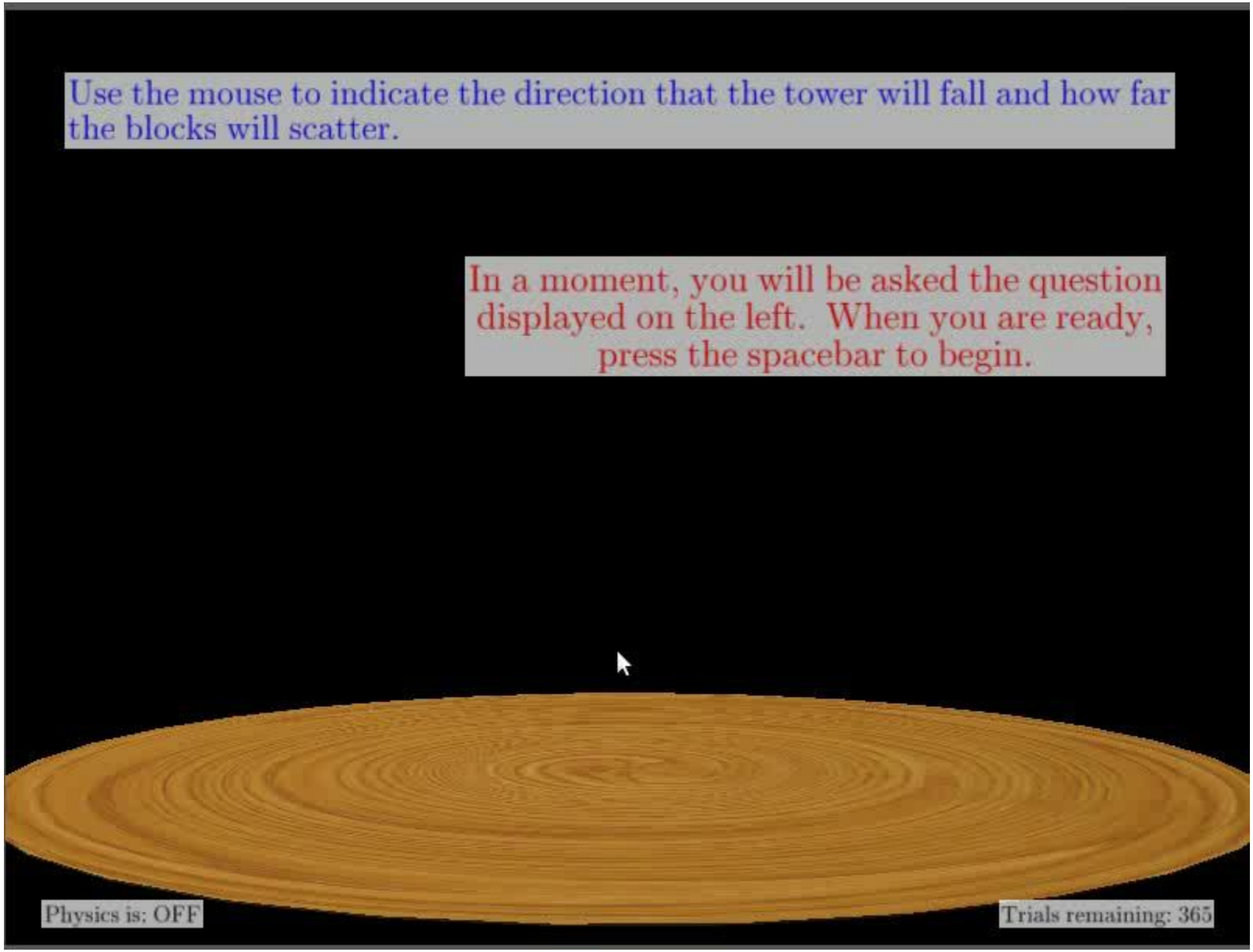
If red blocks are ten times heavier than yellow blocks, how will that change things?

Which of these blocks is heavier or lighter than the others?

Physical simulation, or visual/geometric heuristic?

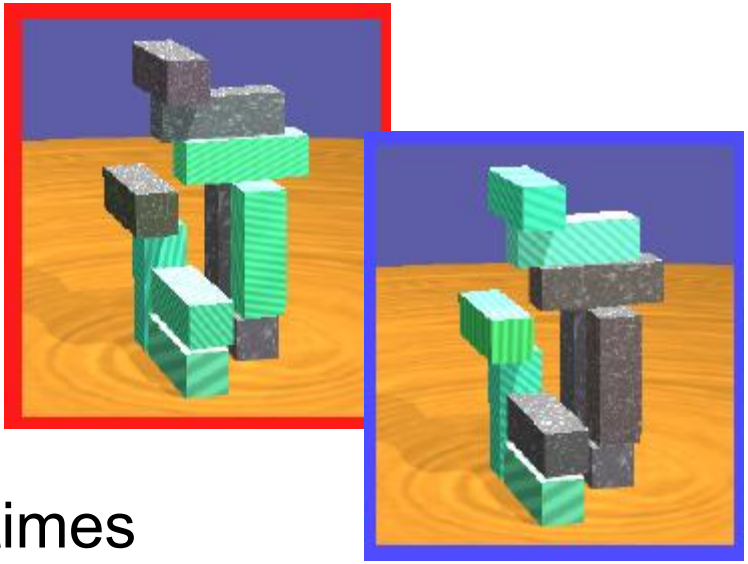


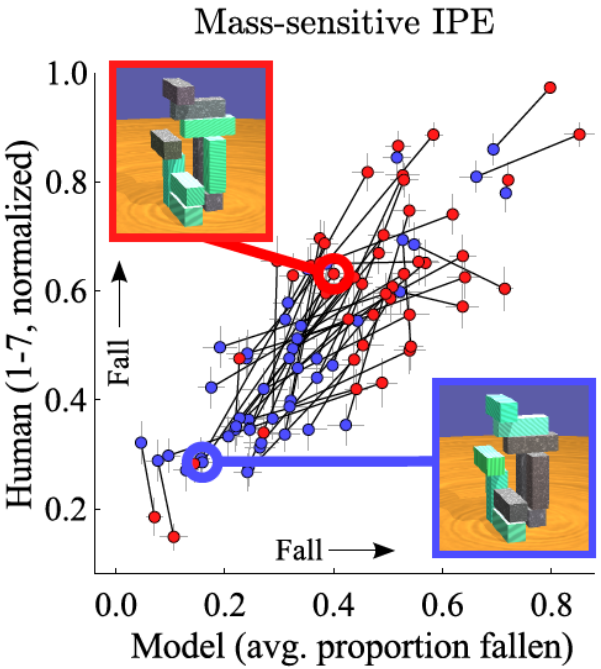
Direction and distance of fall



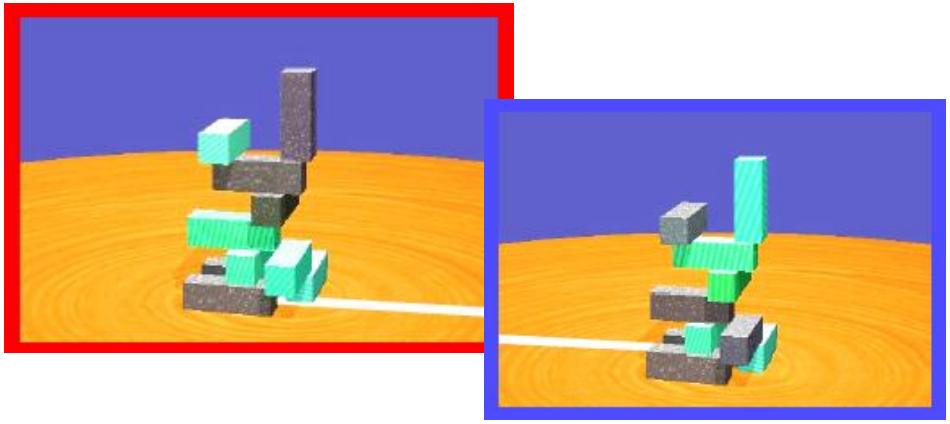
Mass-sensitive predictions

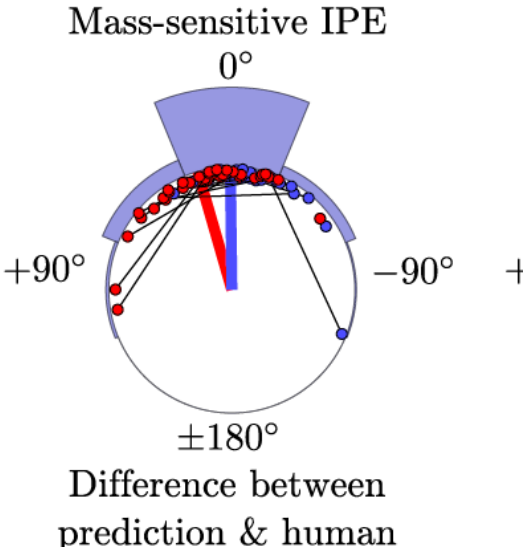
This …





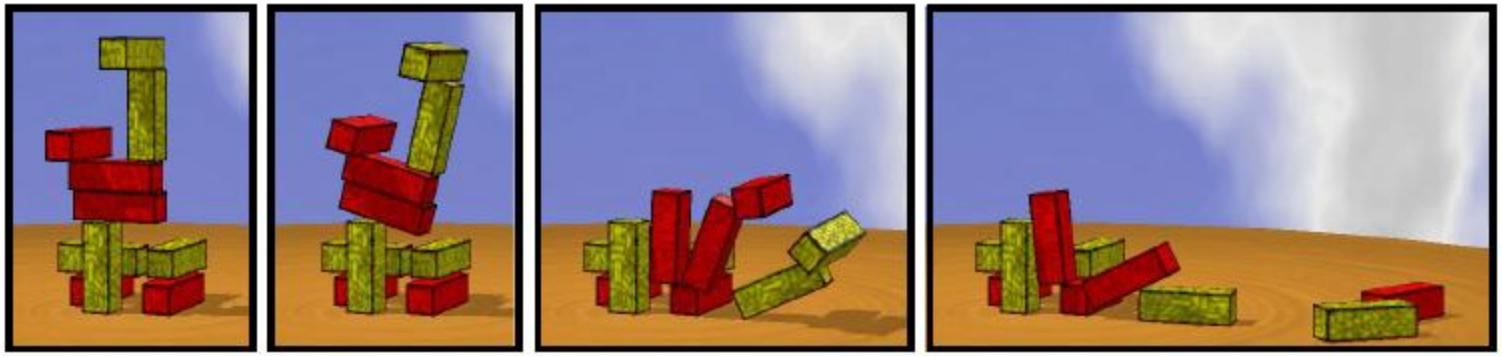
… is 10 times heavier than this:



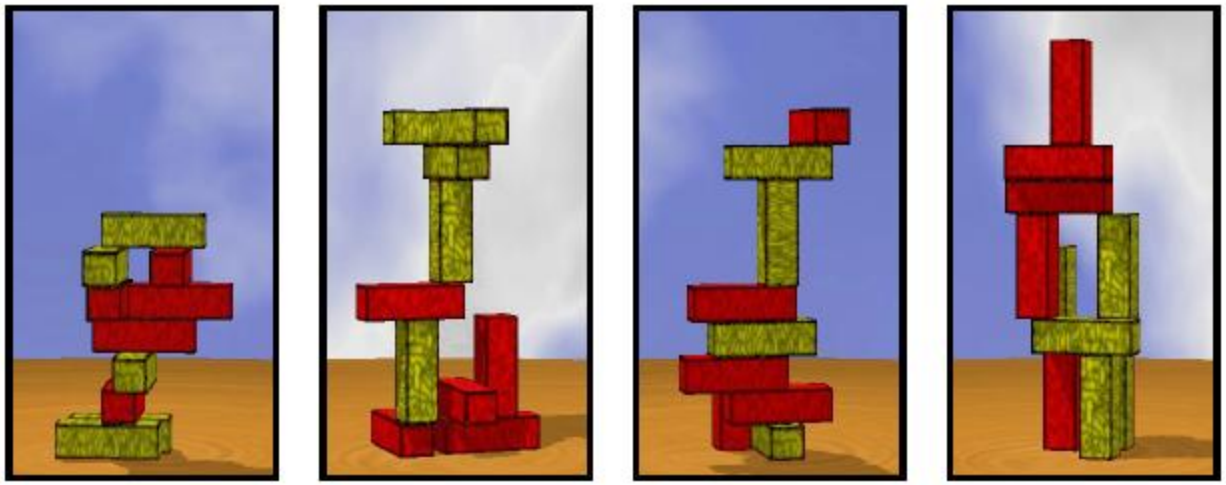


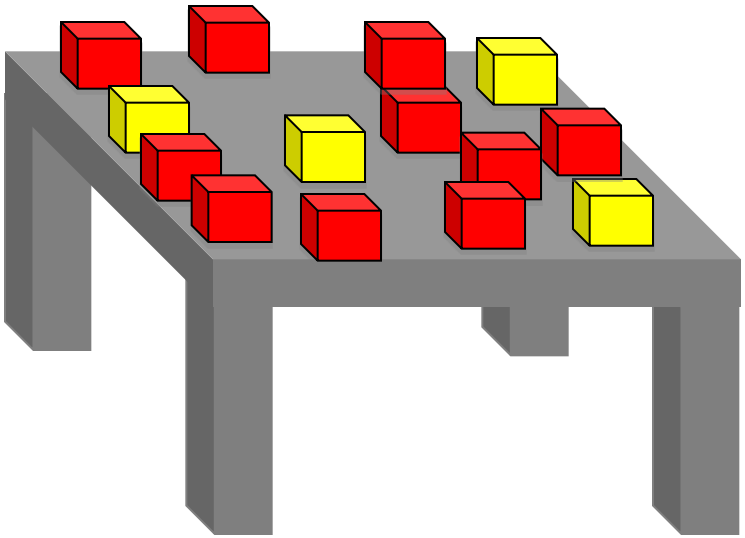
Learning dynamical parameters

Given that this tower fell, which color is heavier?

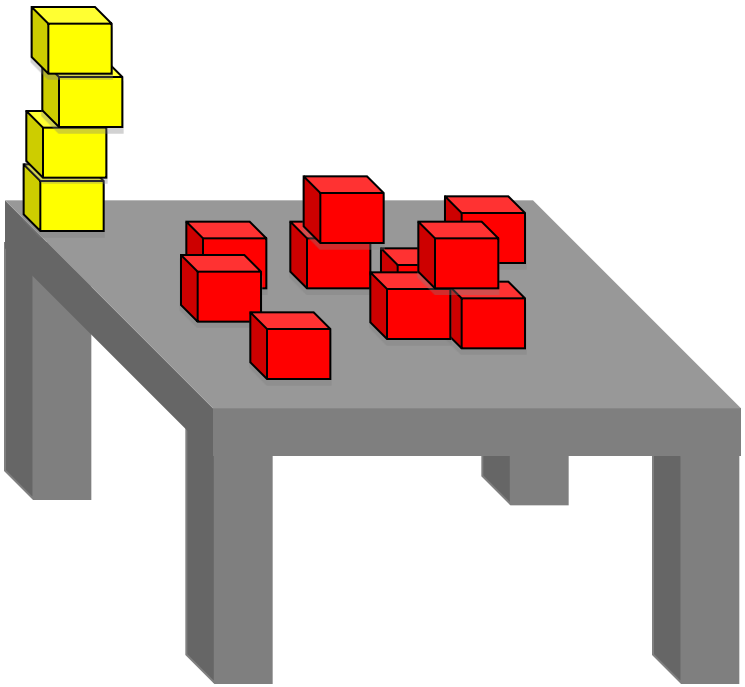


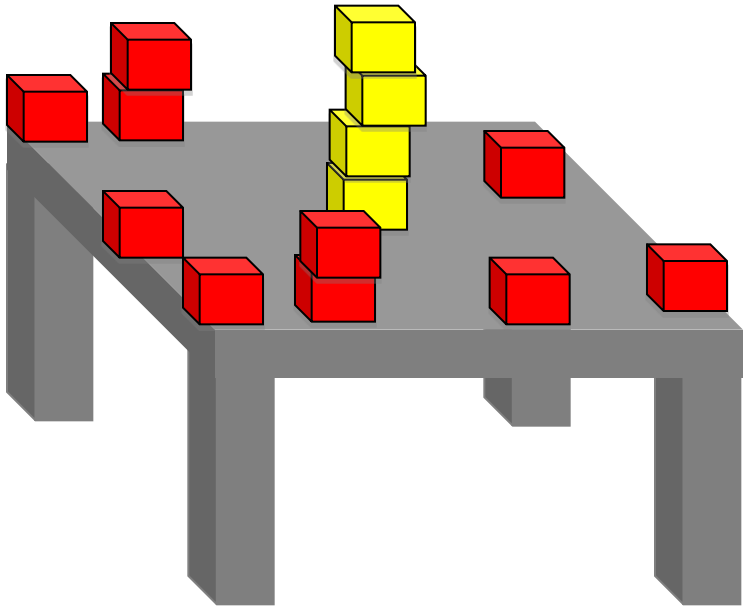
These are stable towers. Which color is heavier?

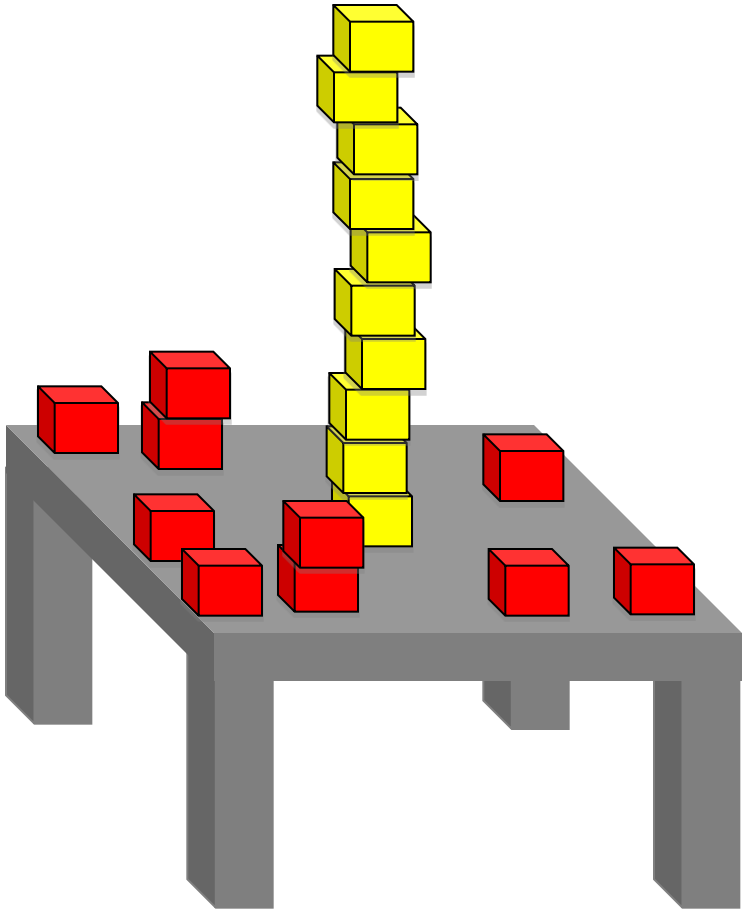


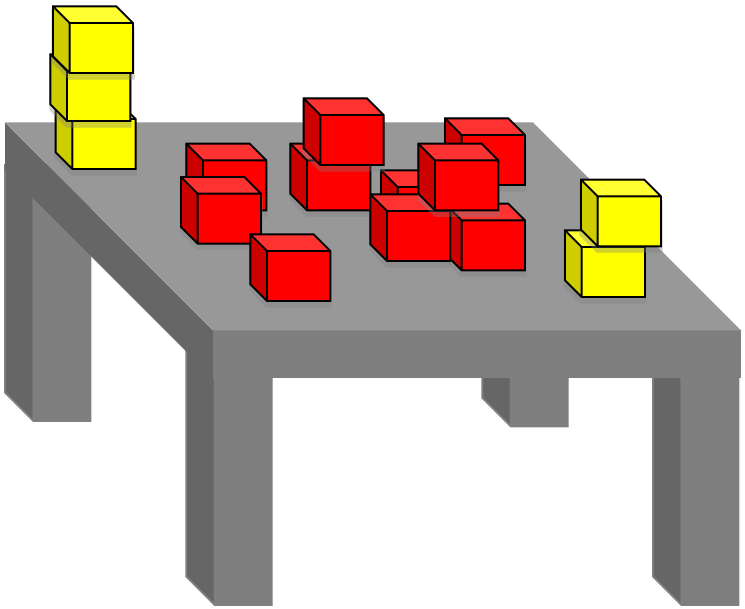


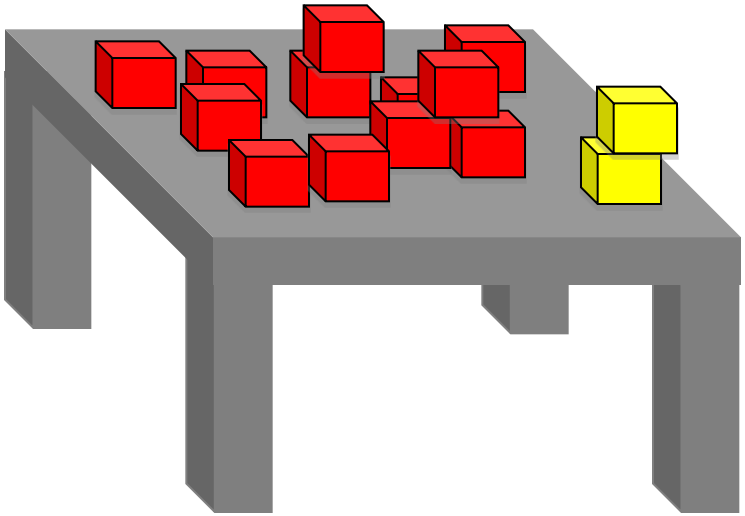
If the table is bumped hard enough to knock some of the blocks onto the floor, is it more likely to be red blocks or yellow blocks?

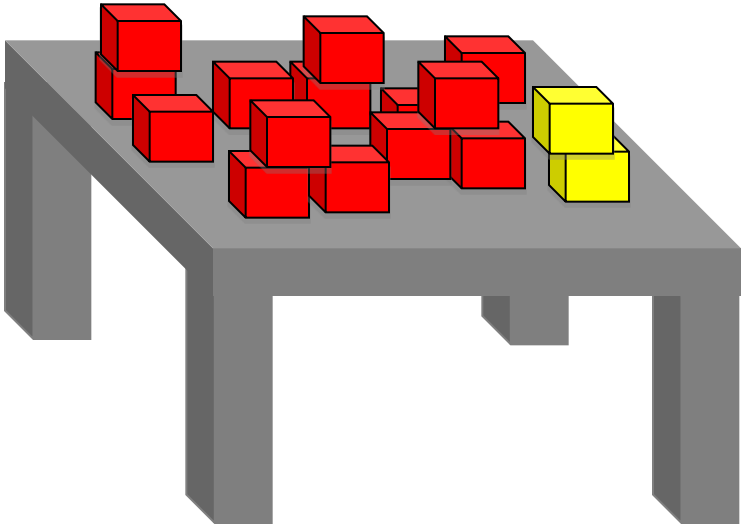


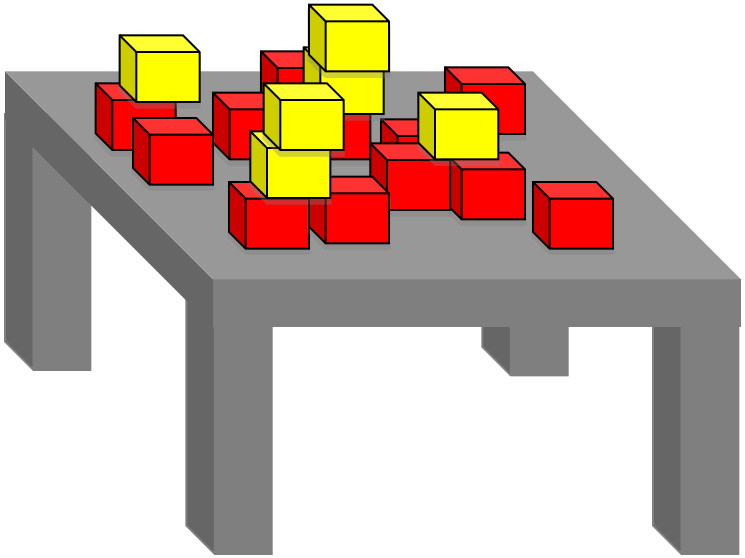




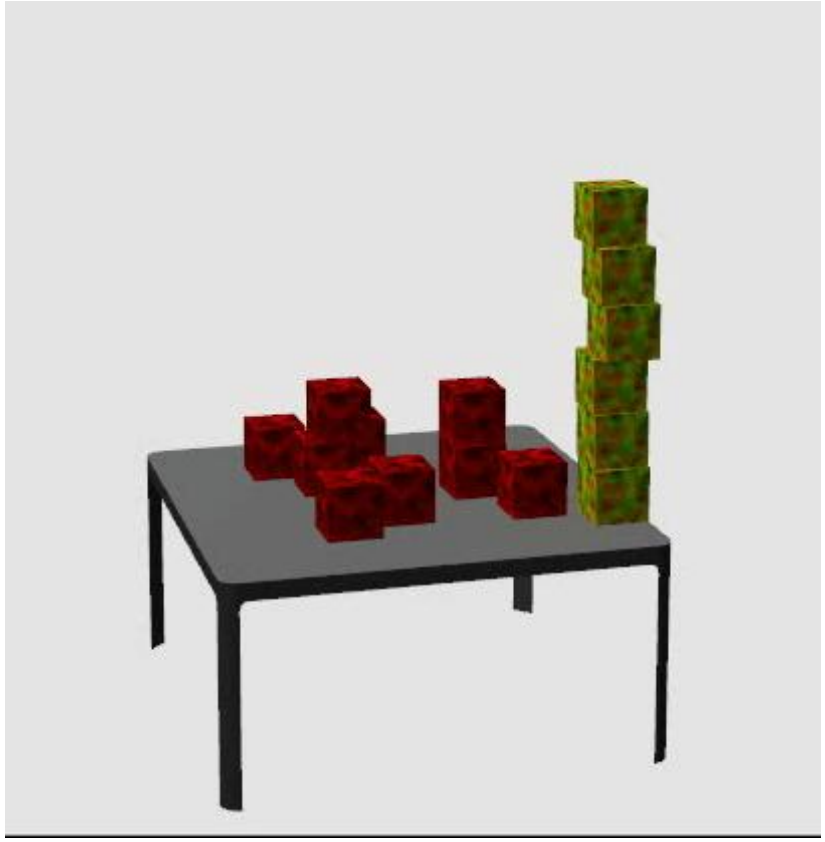




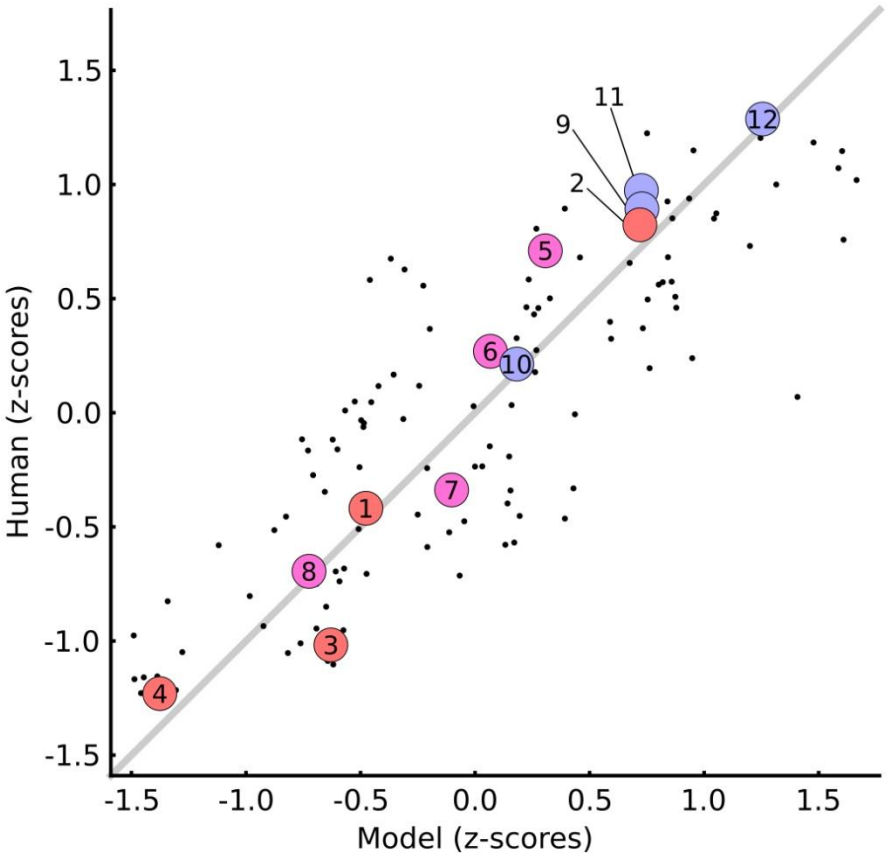


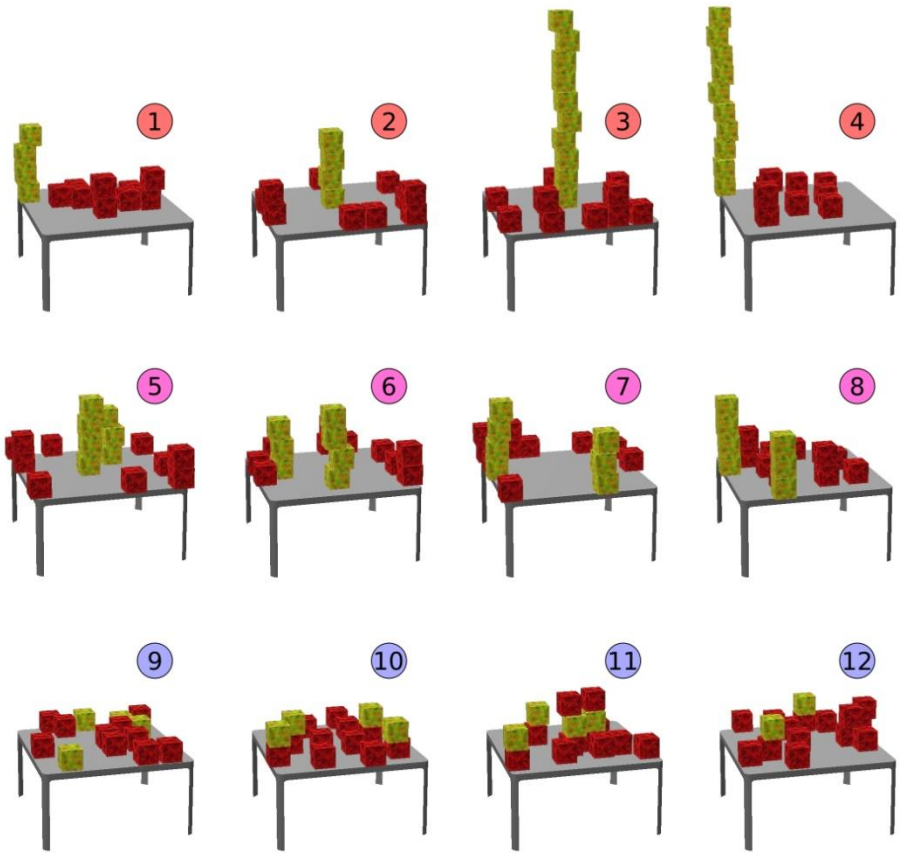


If you bump the table...



If you bump the table…





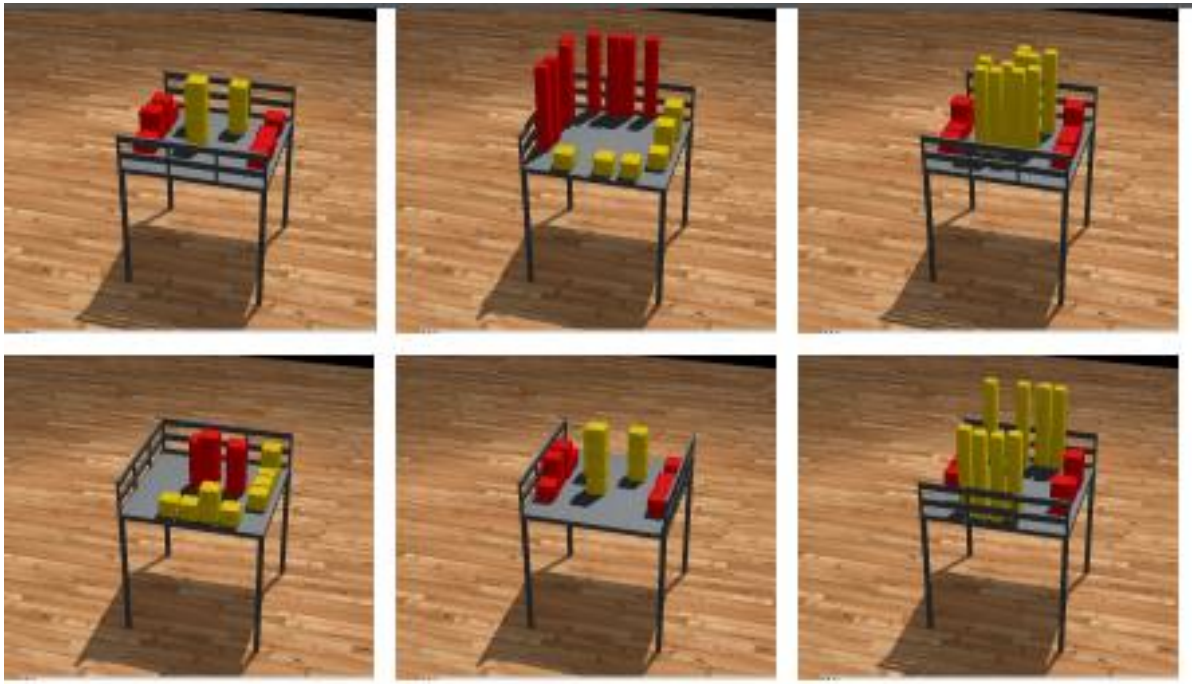
100% yellow

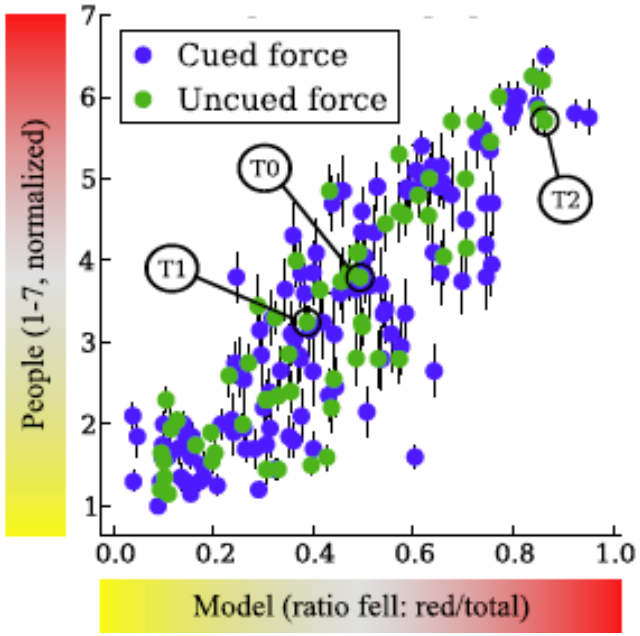
100% red

Model simulates table "bumps" integrating over a range of force magnitudes and directions. (R = 0.84)

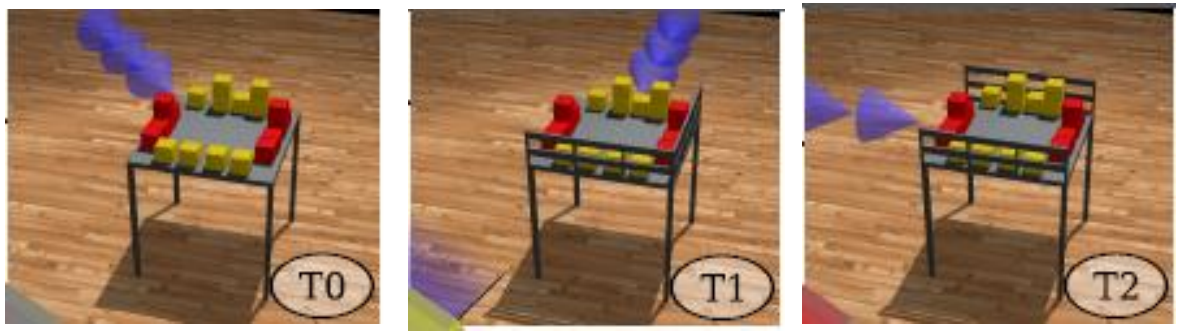
Varying objects, constraints, forces

Uncued forces





Cued forces

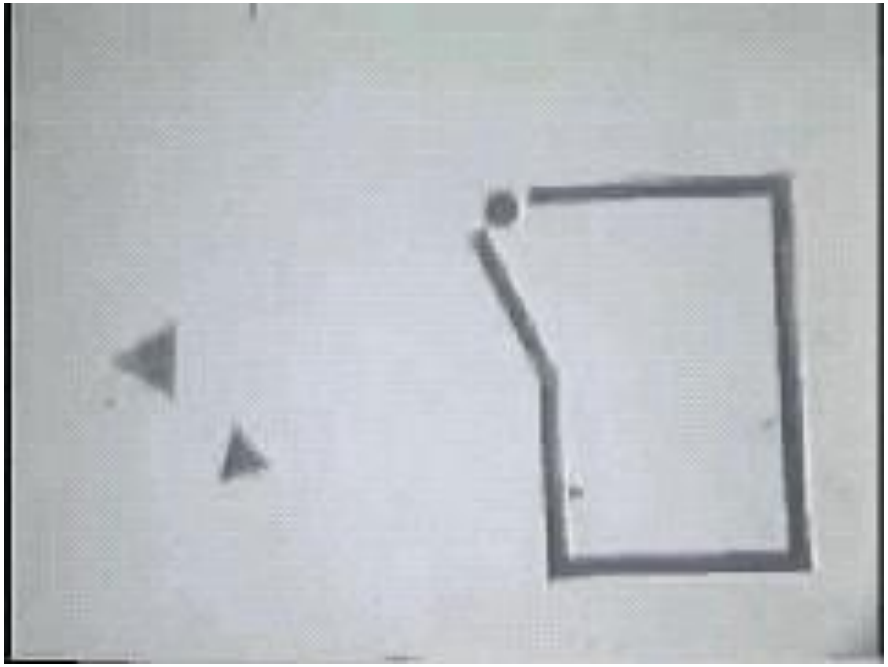


Intuitive psychology

Beliefs (B) Desires (D)

Probabilistic planning

Actions (A)



Heider and Simmel, 1944

Intuitive psychology

Beliefs (B) Desires (D)

Probabilistic planning

Actions (A)

Actions i

States j

In state j, choose action i* =

$$\arg\max_i \sum_{j'} p^i_{j,j'} \, u_{j'}$$

"Inverse economics"

"Inverse optimal control"

"Inverse reinforcement learning"

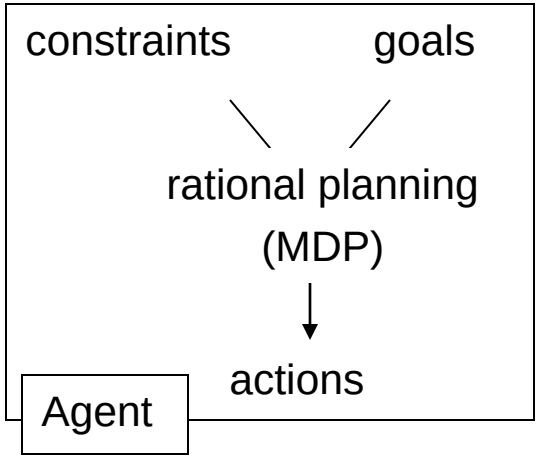
"Inverse Bayesian decision theory"

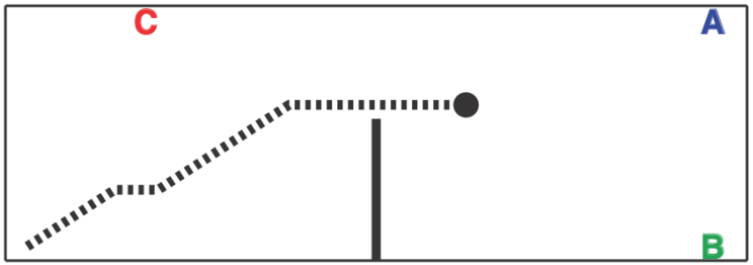
(Lucas & Griffiths; Jern & Kemp; Tauber & Steyvers; Rafferty & Griffiths; Goodman & Baker; Goodman & Stuhlmuller; Bergen, Evans & Tenenbaum …

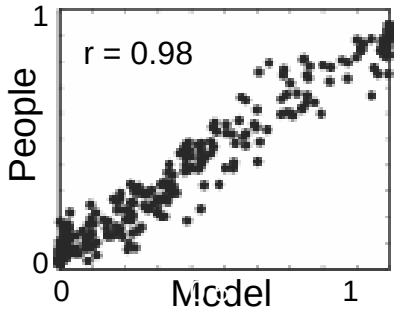
Ng & Russell; Todorov; Rao; Fox; Ziebart, Dey & Bagnell…)

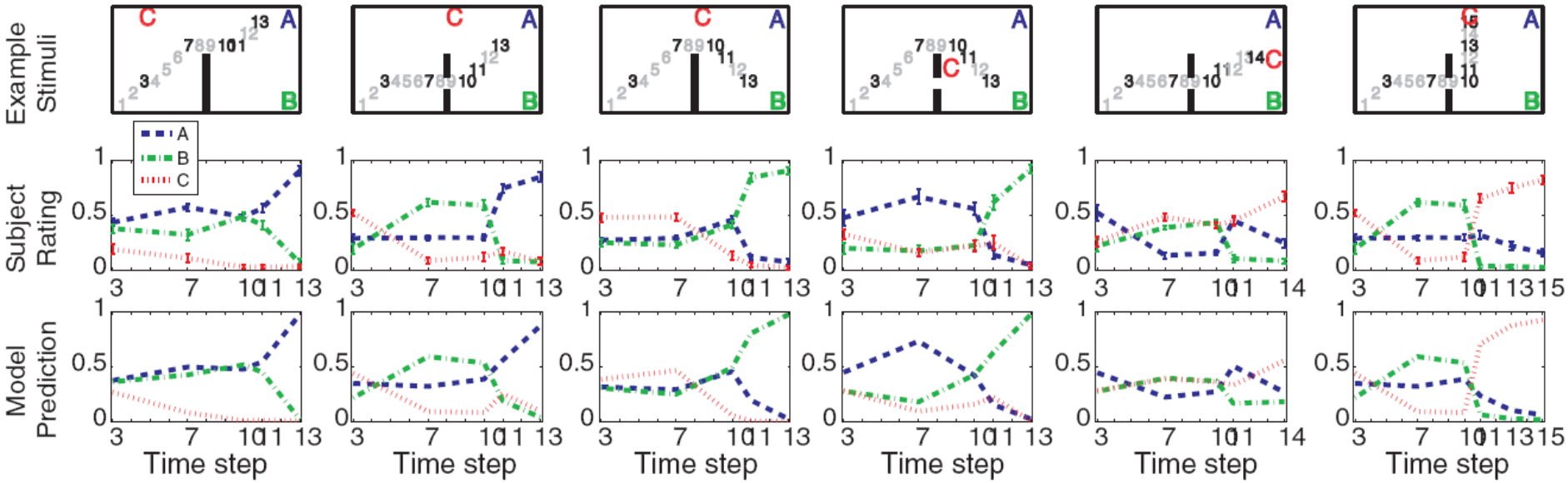
Goal inference as inverse probabilistic planning

(Baker, Tenenbaum & Saxe, *Cognition* 2009)



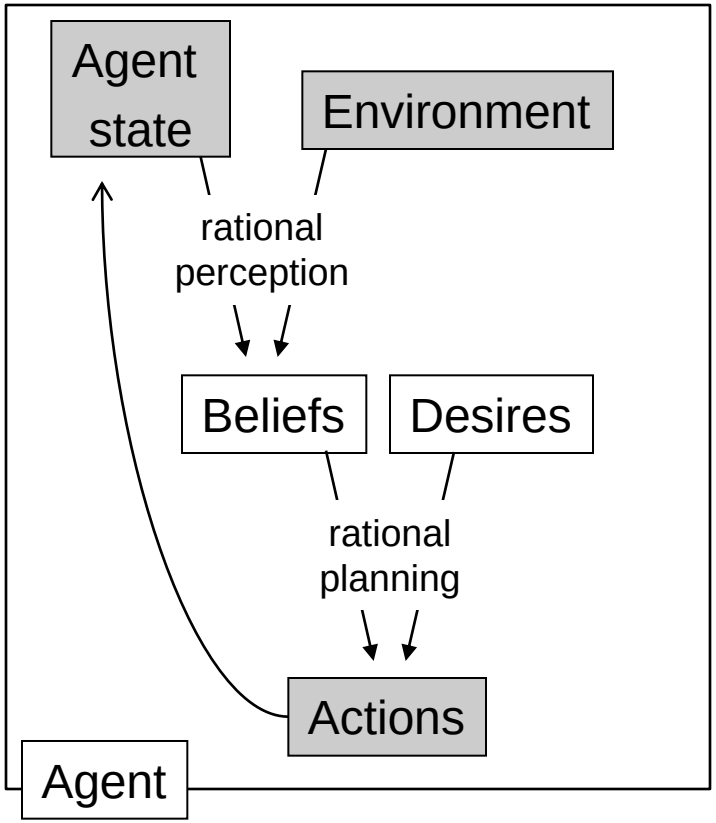




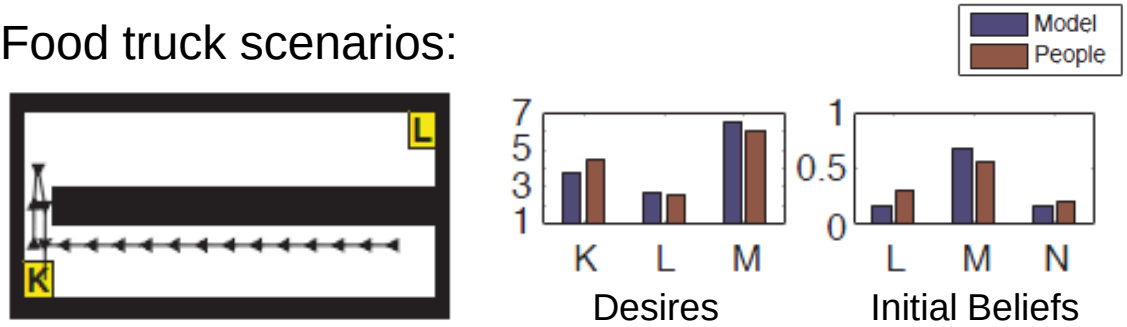


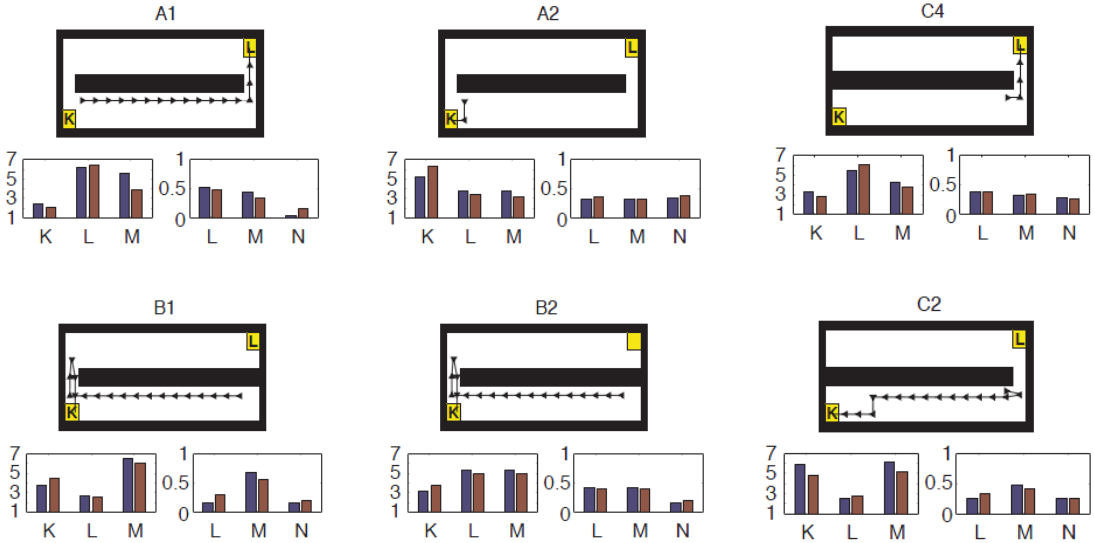
Theory of mind:
Joint inferences about beliefs and preferences

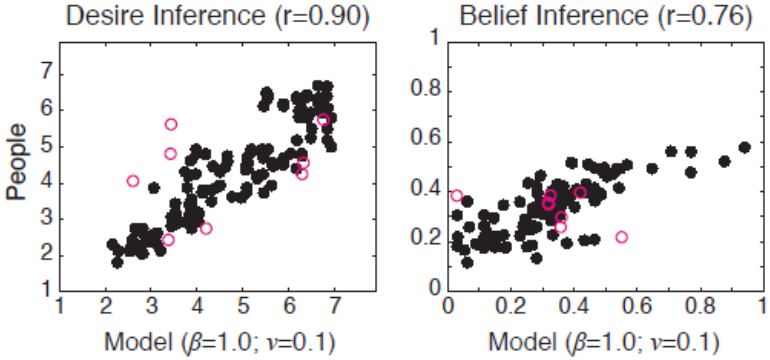
(Baker, Saxe & Tenenbaum, Cog Sci 2011)



Food truck scenarios:

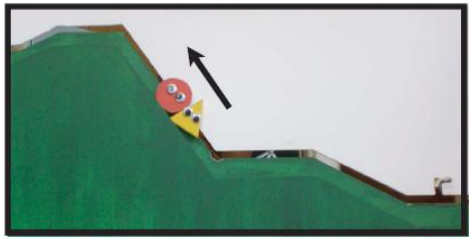


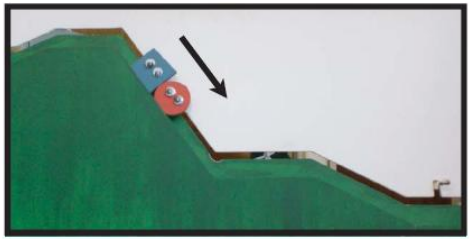




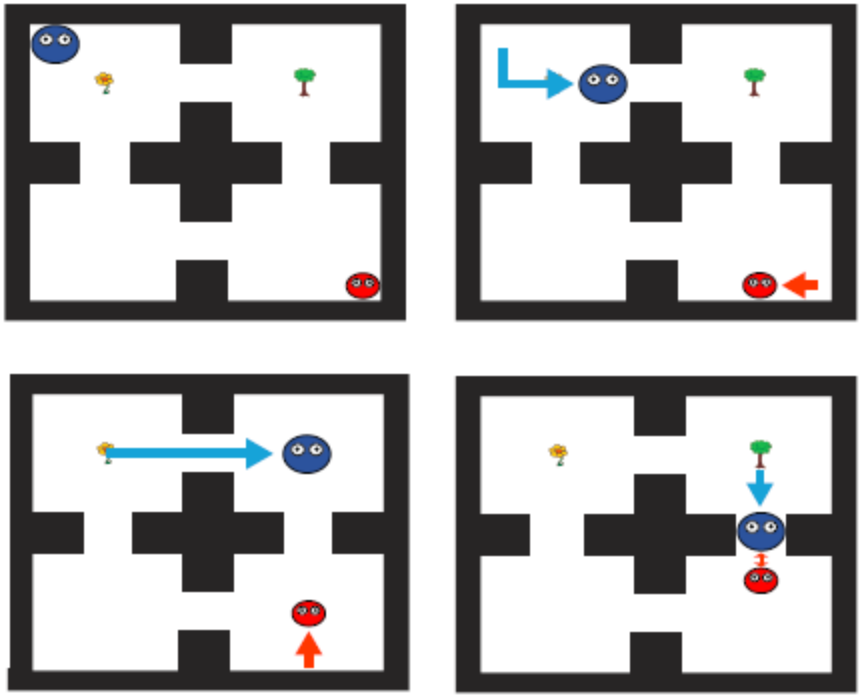
Social goals, moral evaluation

(Hamlin, Kuhlmeier, Wynn & Bloom)

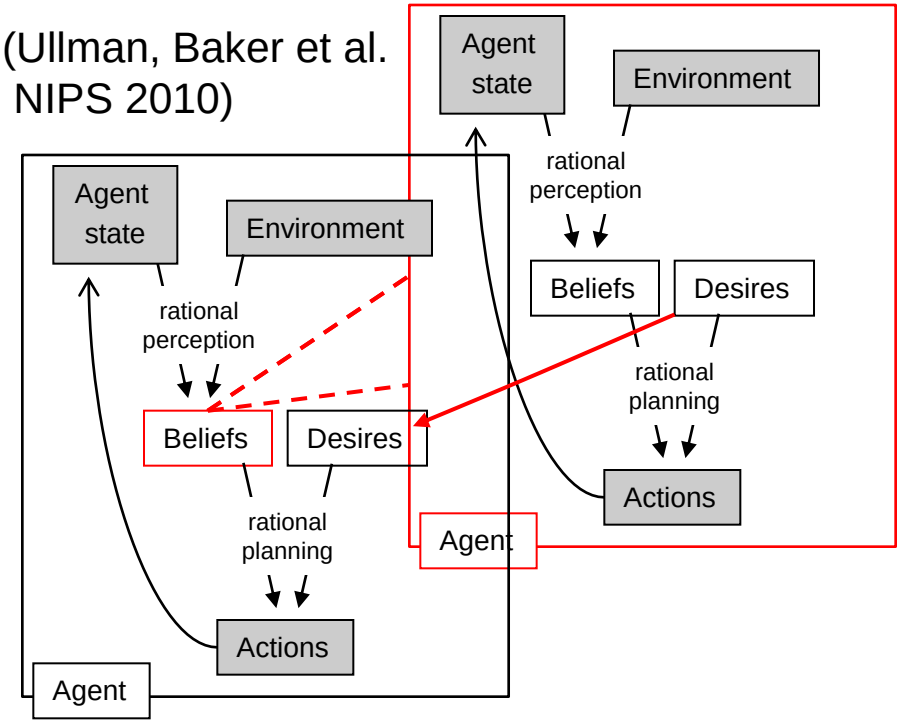


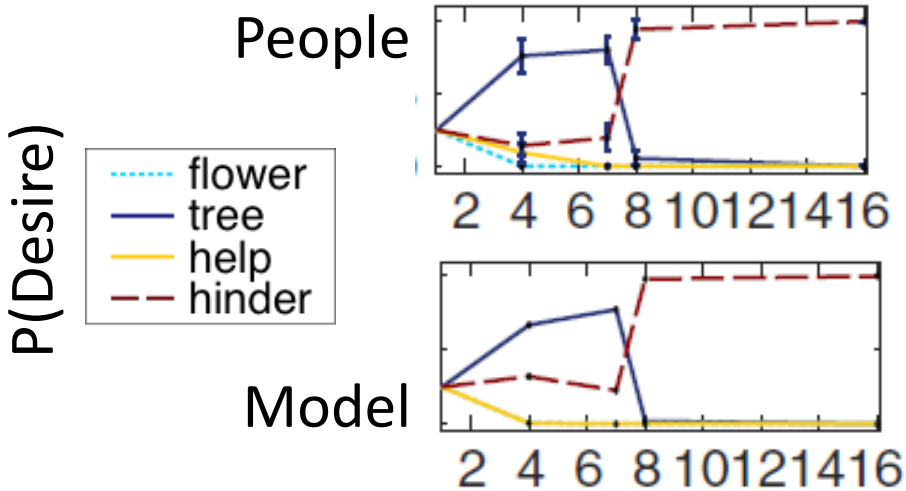


Helping **Hindering**

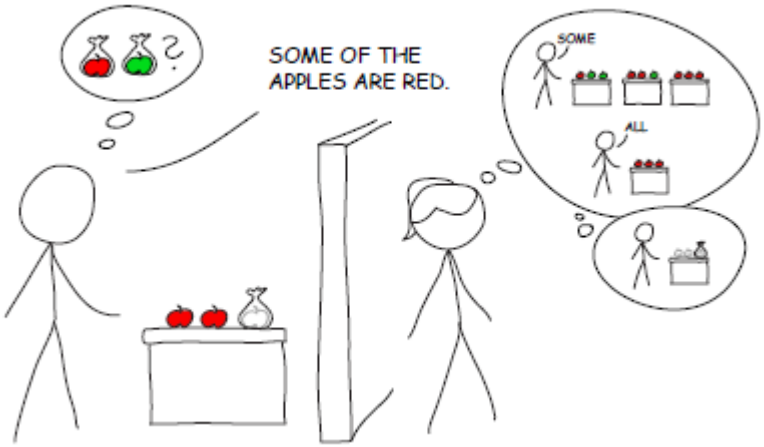


(Ullman, Baker et al. NIPS 2010)





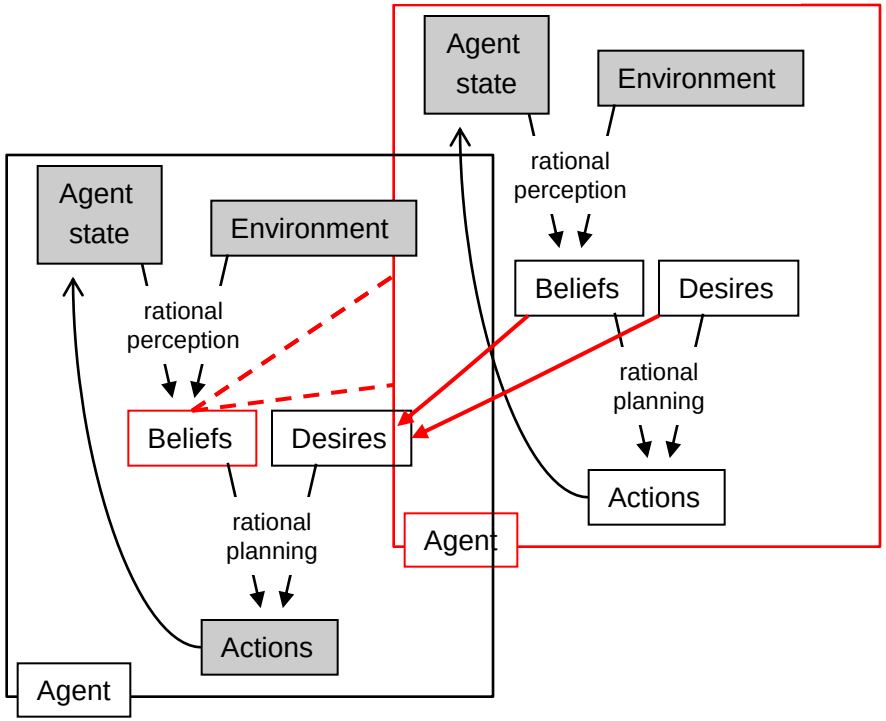
Pragmatic inference in language



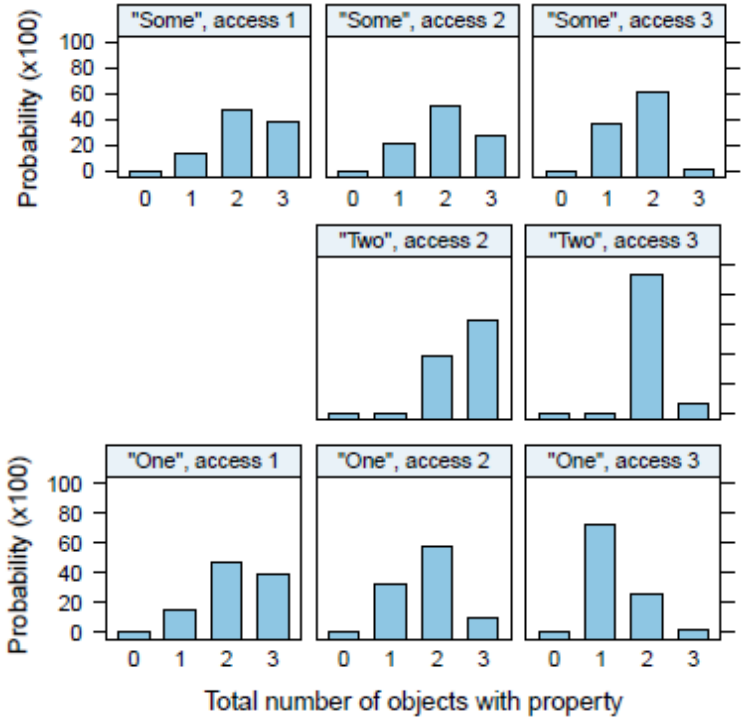
(Goodman & Stuhlmuller, CogSci 2012)

Letters to Laura's company almost always have checks inside. Today Laura received 3 letters.

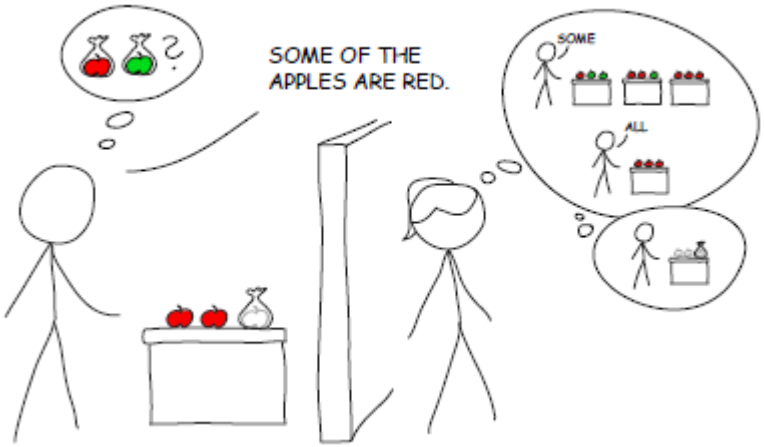
Laura tells you on the phone: "I have looked at 2 of the 3 letters. Two of the letters have checks inside." How many of the 3 letters do you think have checks inside?



Model:



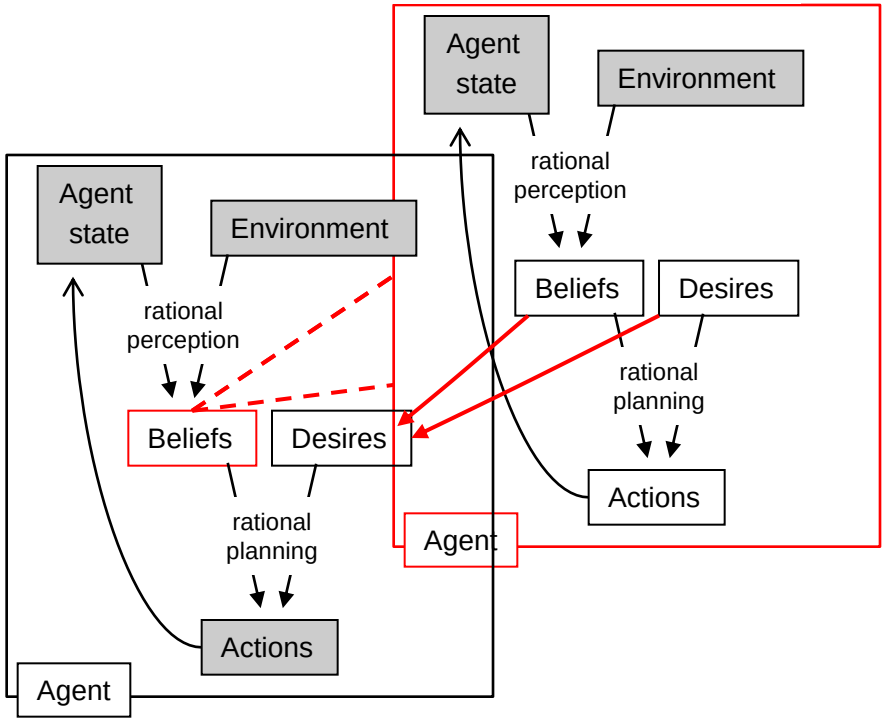
Pragmatic inference in language



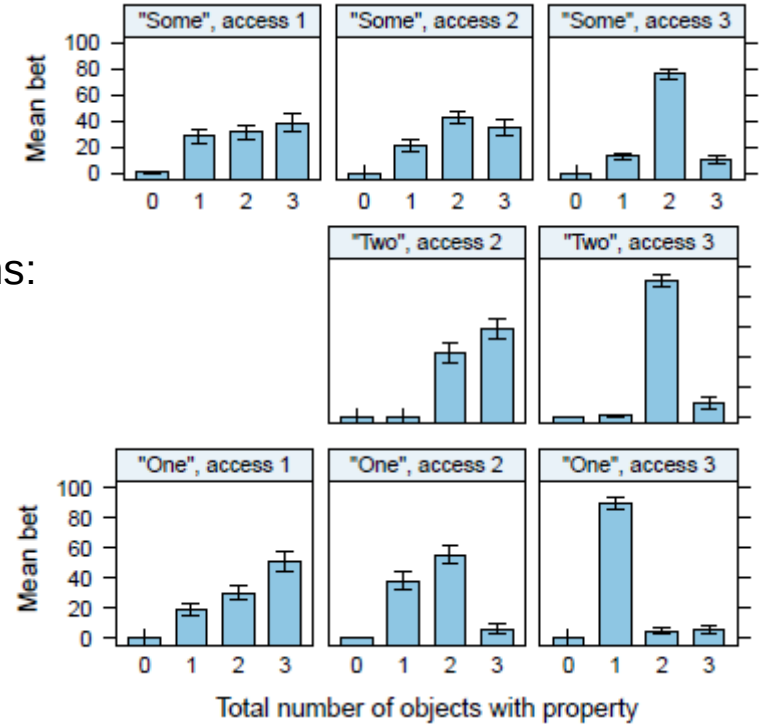
(Goodman & Stuhlmuller, CogSci 2012)

Letters to Laura's company almost always have checks inside. Today Laura received 3 letters.

Laura tells you on the phone: "I have looked at 2 of the 3 letters. Two of the letters have checks inside." How many of the 3 letters do you think have checks inside?



Humans:



Intuitive psychology as inverse decision-making/planning

- Simple goal inference: Baker, Tenenbaum & Saxe, 2006, 2007, 2009
- Social goals: Baker, Goodman & Tenenbaum, 2008; Ullman, Baker, Macindoe, Evans, Goodman, Tenenbaum, 2010
- Learning about preferences: Lucas, Griffiths, Xu & Fawcett, 2008; Bergen, Evans & Tenenbaum 2010
- Theory of mind (joint inference of beliefs and desires): Baker, Saxe & Tenenbaum, 2011; Tauber & Steyvers, 2011
- Learning what is where from observing other agents' exploration: Jara-Ettinger, Baker, Tenenbaum, 2012.
- Goal inference for gaze following, imitation: Rao et al., Friesen & Rao, 2011.
- Using intuitive psychology for causal learning: Goodman, Baker and Tenenbaum, 2009; Goodman & Baker, in prep.
- Mental state reasoning with influence diagrams: Jern & Kemp, 2011.
- Theory of mind in natural language pragmatics: Goodman & Stuhlmuller, 2012, Frank & Goodman, 2012
- Recursive multi-agent models for story understanding and game playing: Bergen, Stuhlmuller, Goodman & Tenenbaum, in prep., Hedden & Zhang, 2002; Doshi, Goodie et al. 2009, 2010; Yoshida, Dolan & Friston, 2008, 2010.

Frontiers

(Mansinghka, Roy, Freer, Goodman, Stuhlmuller…)

- Effective algorithms for universal approximate inference in probabilistic programming languages.
  - Sampling-based, dynamic programming, variational... Connections to programming languages (e.g. static analysis)?
- Theory of universal inference in probabilistic programs.
  - Computability, complexity, efficiency under resource bounds… From theory to algorithms?
- Learning as program induction(!). Warm up problems…
  - Graph grammars for structural form.
  - Learning schemas for causal networks.
  - Learning simple logical theories.
  - Learning functional aspects of language: determiners, quantifiers, prepositions, adverbs.
  - Learning programs to describe dynamics of objects and substances.

Learning natural number

(Piantadosi, Goodman, Tenenbaum, *Cognition*, in press)

Count-list-knower:	recite counting routine
One-knower (2;10):	"one"
Two-knower (3;0):	"one," "two"
Three-knower (3;5):	"one," "two," "three"
CP-knower (> 3;7):	"one," "two," "three," "four," "five," "six," "seven," ...

"Can you count the balloons?

"Can you give me three balloons?

(Spelke 2003; Wynn 1990, 1992; Sarnecka & Lee, 2009; Lee & Sarnecka, 2010b, 2010a)

Ingredients of number knowledge

Set-theoretic functions	union – union of two sets intersection – intersection of two sets select – choose an element of a set set-difference – set-difference
Logic	and – logical conjunction or – logical disjunction not – logical negation if – logical conditional
Small cardinalities (Wynn 1992)	singleton? – check if a set has one element doubleton? – check if a set has two elements tripleton? – check if a set has three elements
Counting routine (Fuson 1988)	next – next word in the counting routine prev – previous word in the counting routine equal-word? – word equality
Recursion (Hauser, Chomsky, Fitch 2002)	L – recursion

Three-knower

```
λ S . (if (singleton? S)
          "one"
          (if (doubleton? S)
              "two"
              (if (tripleton? S)
              "three"
              undef))
```

2-not-1-knower

```
λ S . (if (doubleton? S)
          "two"
          undef)
```

Singular-Plural

```
λ S . (if (singleton? S)
          "one"
          "two")
```

CP-knower

```
λ S . (if (singleton? S)
          "one"
          (next (L (set-difference S
                                   (select S)))))
```

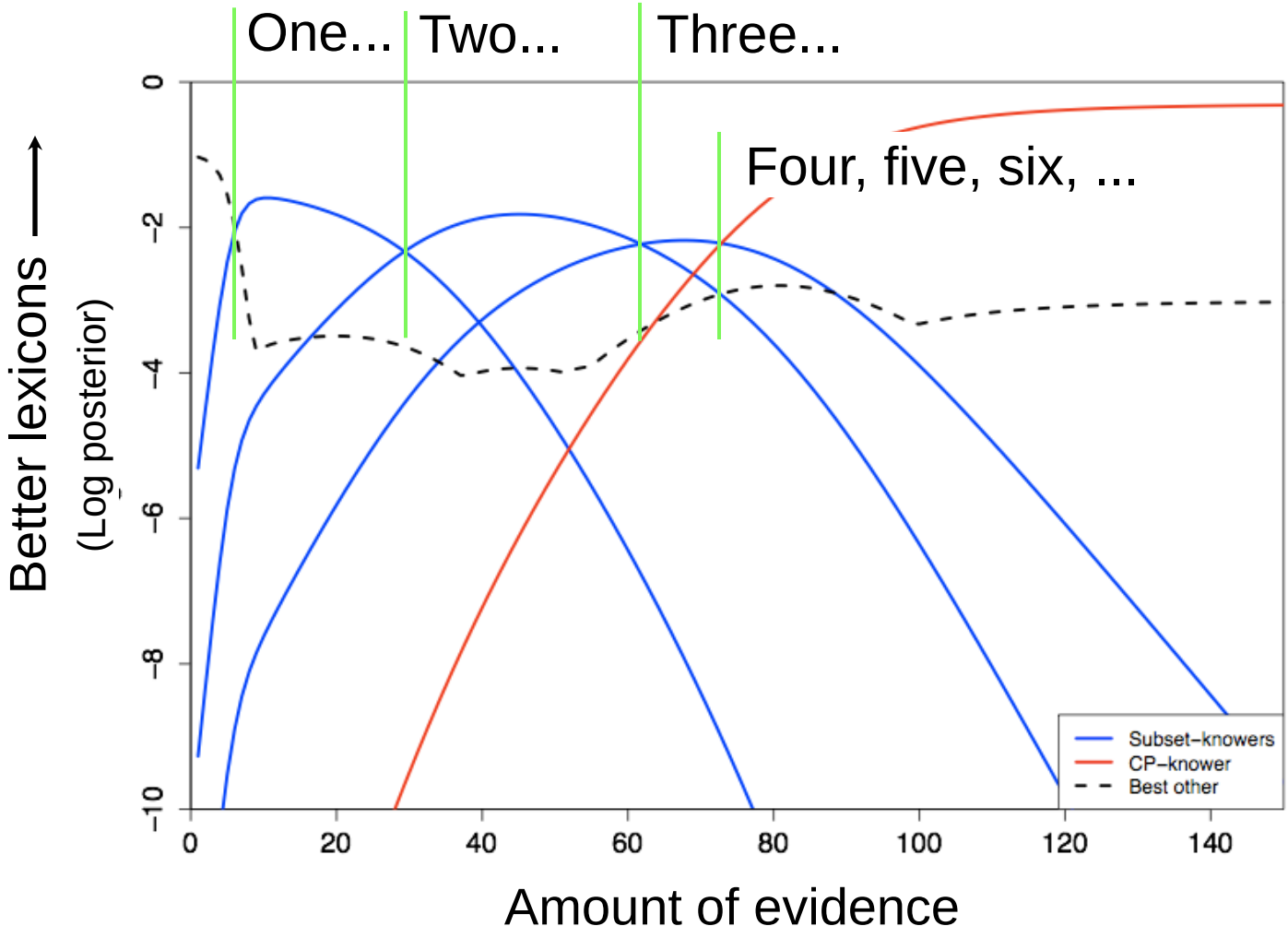
2N-knower

```
λ S . (if (singleton? S)
          "one"
          (next (next (L (set-difference S (select S))))))
```

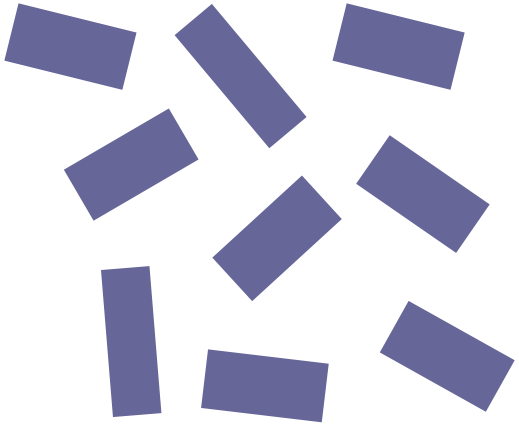
Mod-5

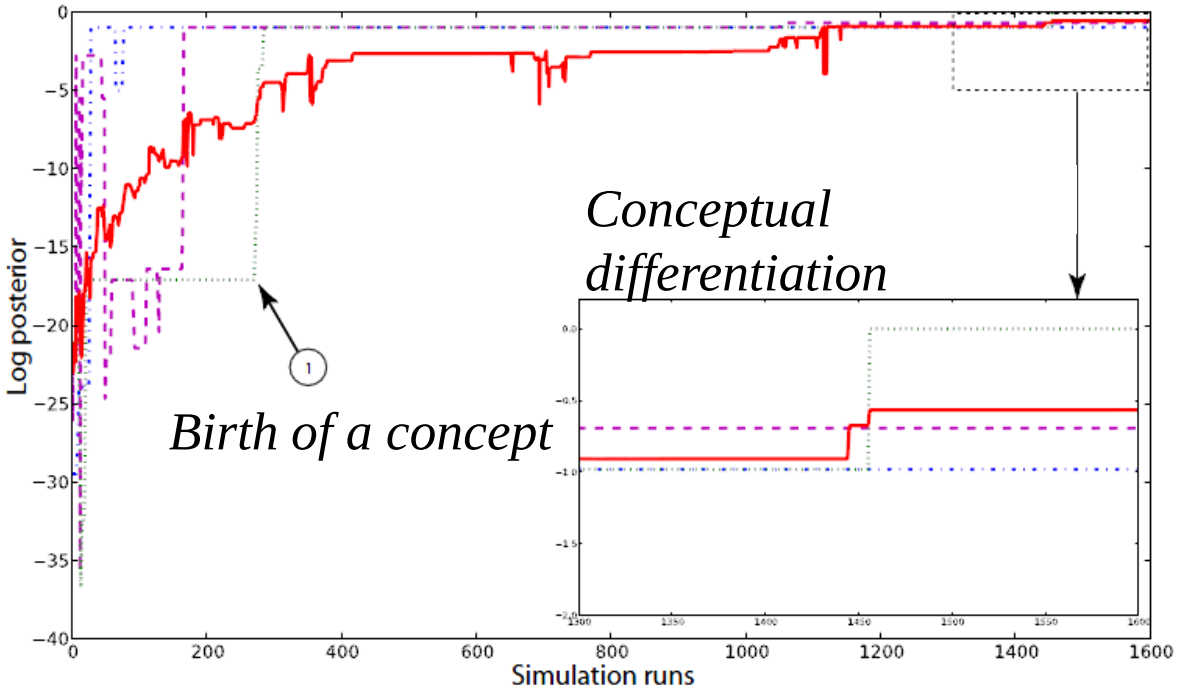
```
λ S . (if (or (singleton? S)
              (equal-word? (L (set-difference S)
                              (select S))
                           "five"))
          "one"
          (next (L (set-difference S
                                   (select S)))))
```

Learning results



Theory learning as stochastic search
(Ullman, Goodman, Tenenbaum, 2010, in press)





A simple theory of magnetism:

Rule 1: Interacts(X,Y) ⟵ p(X) ∧ p(Y)

A less simple but better alternative:

Rule 1: Interacts(X,Y) ⟵ p(X) ∧ p(Y)

Rule 2: interacts(X,Y) ⟵ p(X) ∧ q(Y)

Rule 3: interacts(X,Y) ⟵ interacts(Y,X)

Theories generated from a probabilistic grammar.

Metropolis-Hastings search w/ grammar-based proposals.

Inducing physics programs

(Ullman, Stuhlmuller, Goodman, Tenenbaum, in prep)

Conclusions

How does the mind get so much from so little, in learning and reasoning about objects, categories, causes, scenes, events?

A toolkit for learning and reasoning where statistics meets abstract knowledge:

- *Bayesian inference* in probabilistic generative models.
- Probabilistic models defined over a range of *structured representations*: graphs, grammars, schemas, predicate logic…
- *Hierarchical models*, with inference at multiple levels of abstraction.
- *Probabilistic programs*: computationally universal representations for causal processes that are relational, recursive, composable.

Towards a computational account of human common sense.

- The "common-sense core": intuitive theories of the physical world, intentional agents, and their interaction that emerge early in infancy.
- How these theories are used to perceive, reason, predict, plan, learn and communicate, and might themselves be learned…

Some thoughts for AI researchers interested in human-like AI:

- Focus on the common-sense core.
- Abstract knowledge is essential but need not be wired in. Think of how knowledge can grow flexibly and quickly in response to experience.
- Probabilistic programs are the way to go. Reach out to graphics, physical simulation, algorithms, programming languages, theory, circuits.