A latent variable model of configural conditioning

Aaron C. Courville
Robotics Institute, CMU

Work with: Nathaniel D. Daw, UCL (Gatsby),
David S. Touretzky, CMU and
Geoff Gordon, CMU

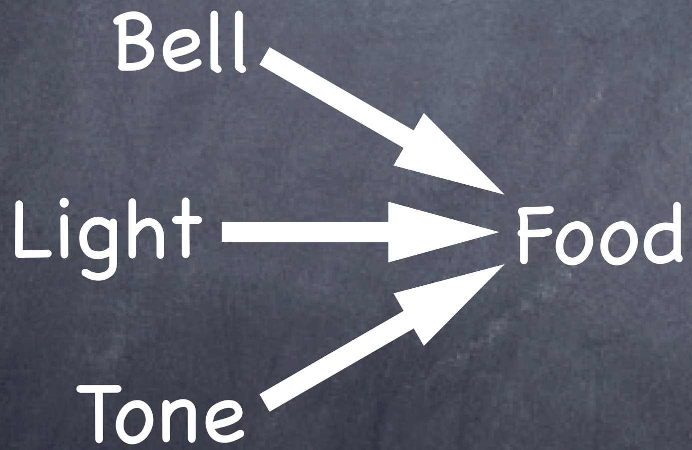
Similarity & Discrimination in Animal Learning

- Similarity: How do animals respond to novel patterns of stimuli?
- Discrimination: How do animals learn to discriminate between overlapping patterns of stimuli?
- We recognize these issues as tradeoff between generalization and data-fitting

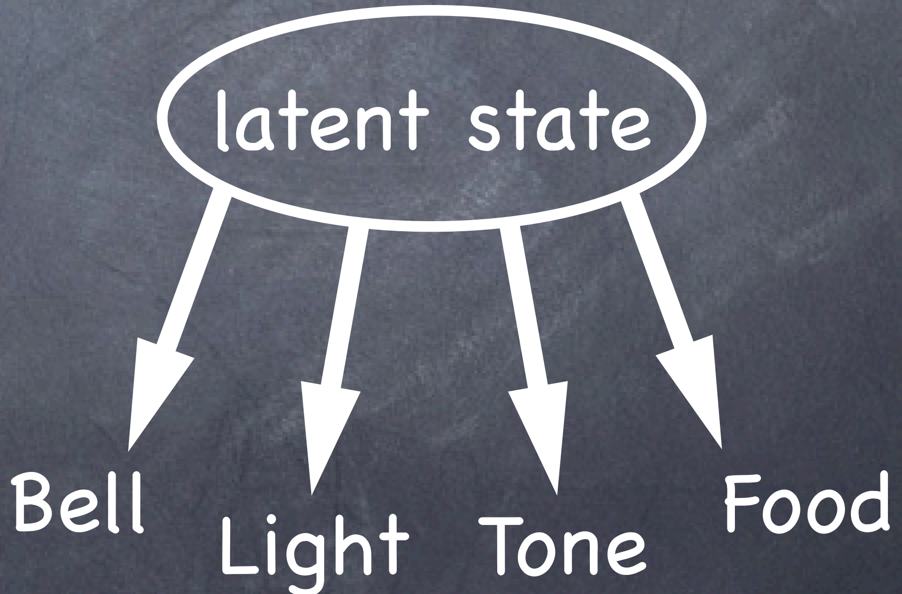

TONGUE OUT
SALIVA BEGINS
TO FLOW
DIGESTIVE JUICES
BEGIN TO WORK

Perspectives on modeling conditioning

Discriminative



Generative



Models of Conditioning: Rescorla-Wagner (1972)

- Predicts reinforcement intensity as a linear function of stimuli, X = [A (light), B (bell), ...].

$$V = \sum_i w_i X_i$$

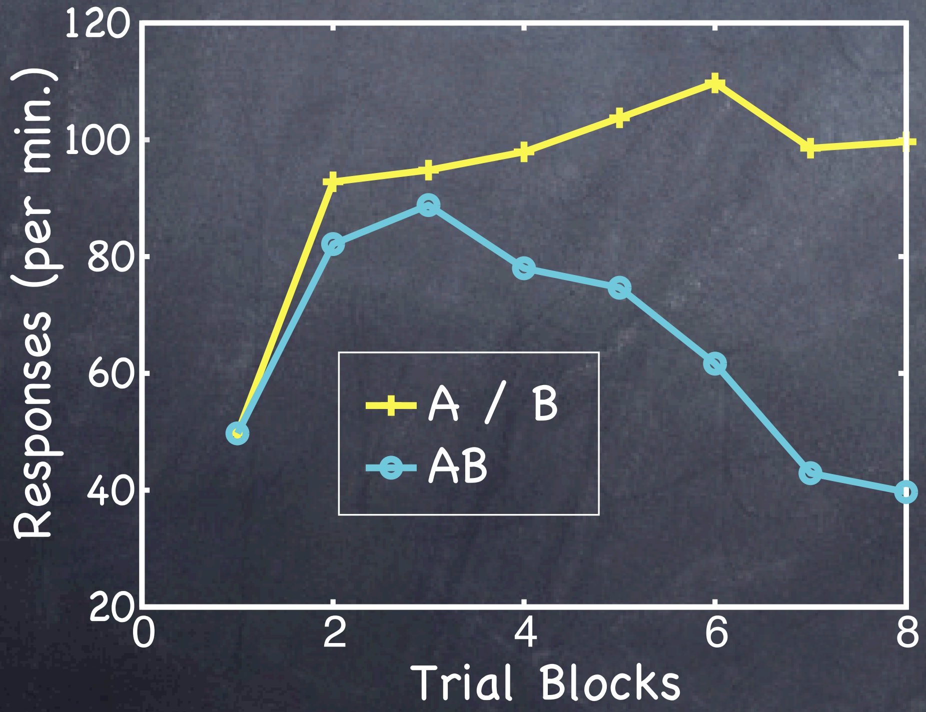
- Learning rule is gradient descent on prediction error

$$\Delta w_i = \alpha_i \beta (r - V) X_i$$

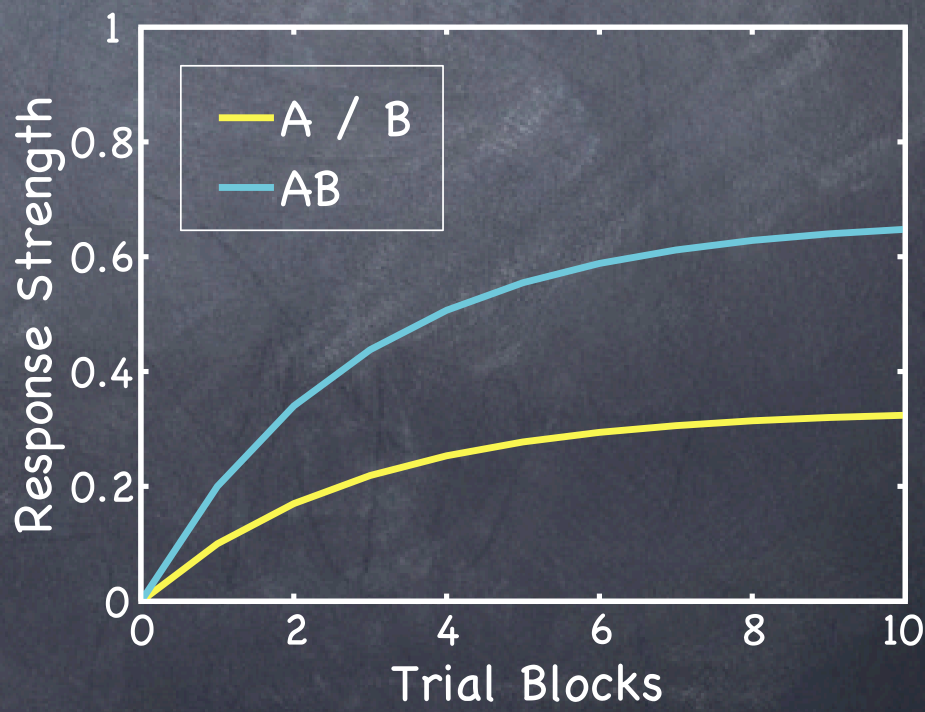
Stimulus configurations

- Configural conditioning: discrimination and generalization between patterns of stimuli.

Training: (XOR) A+ B+ AB-



Rescorla-Wagner:



Modeling Configurations

- Two dominant perspectives:
  1. Added elements RW, [W&R, 1972]
  2. Configural model [Pearce, 1994]

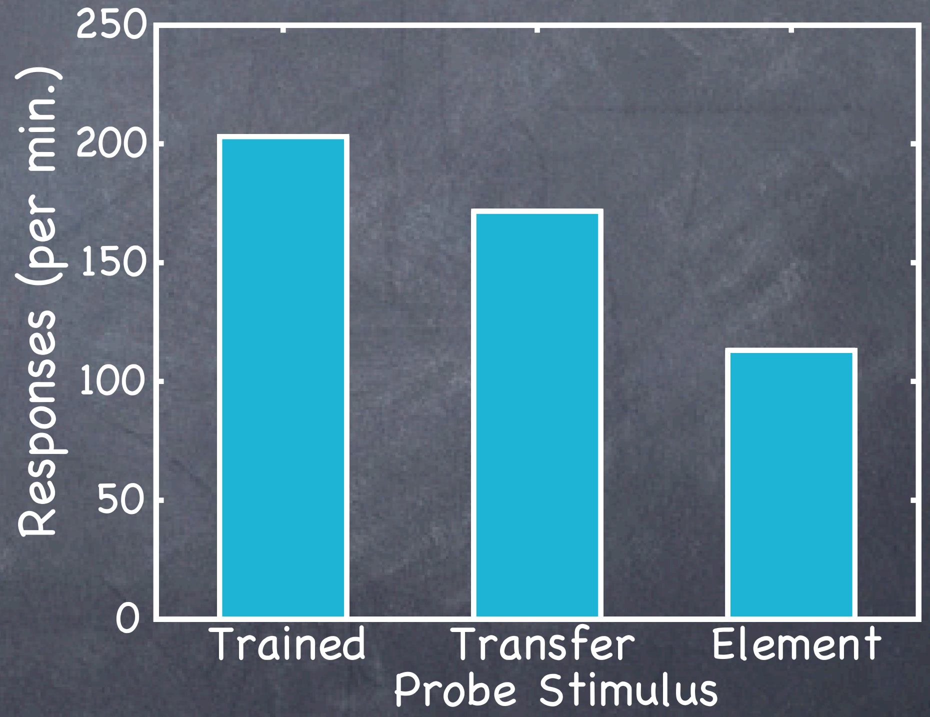
- Augment stimulus representation with configural unit. eg. XOR: X=[A,B,AB].

- Which units are active? Observe AB
  - [W&R 1972]: All units present. X=[A=1,B=1,AB=1]
  - [Pearce 1994]: Graded activation by generalization rule. X=[A=.5,B=.5,AB=1]

Expt. 1 Paired Compounds

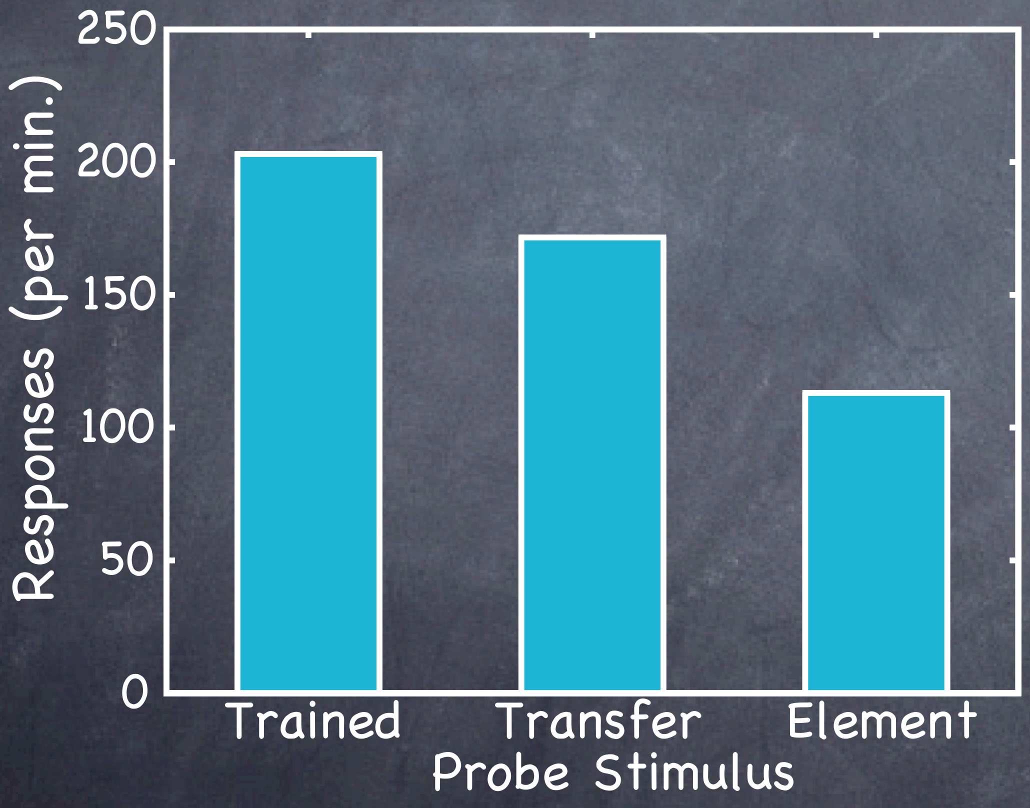
[Rescorla, 2003]

Training Trials:	Test Trials:	
AB+ CD+	Trained:	AB, CD
	Transfer:	AD, BC
	Elements:	A, B, C, D

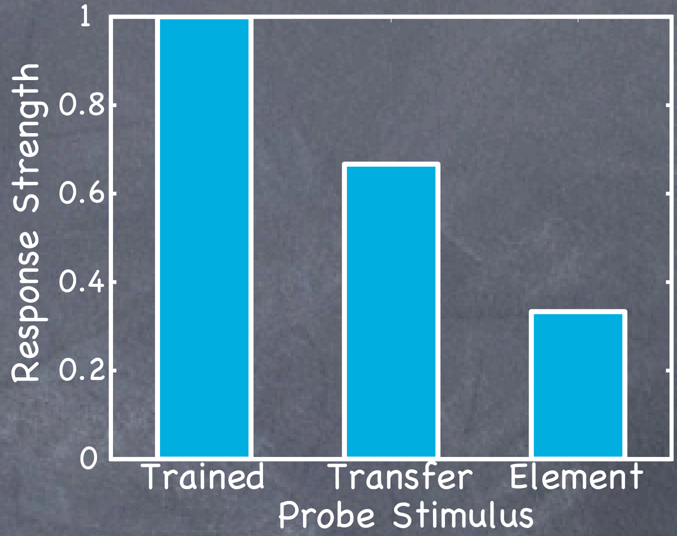


Modeling Paired Compounds

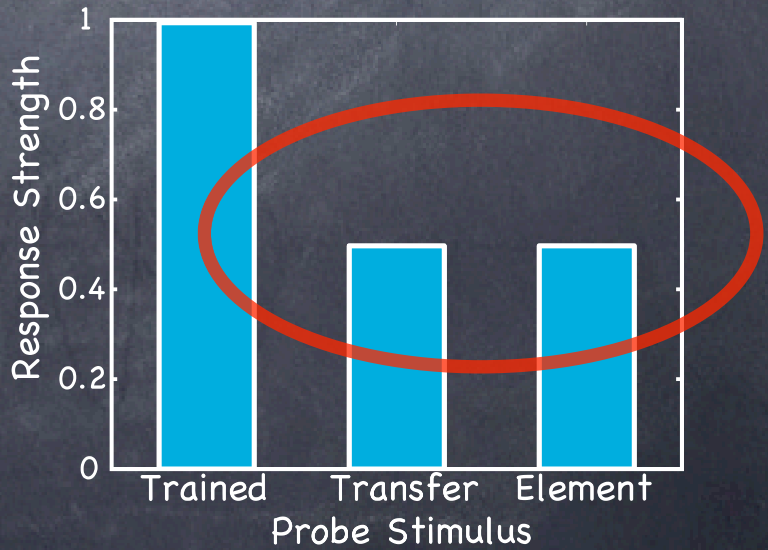
Training: AB+ CD+



Rescorla-Wagner:



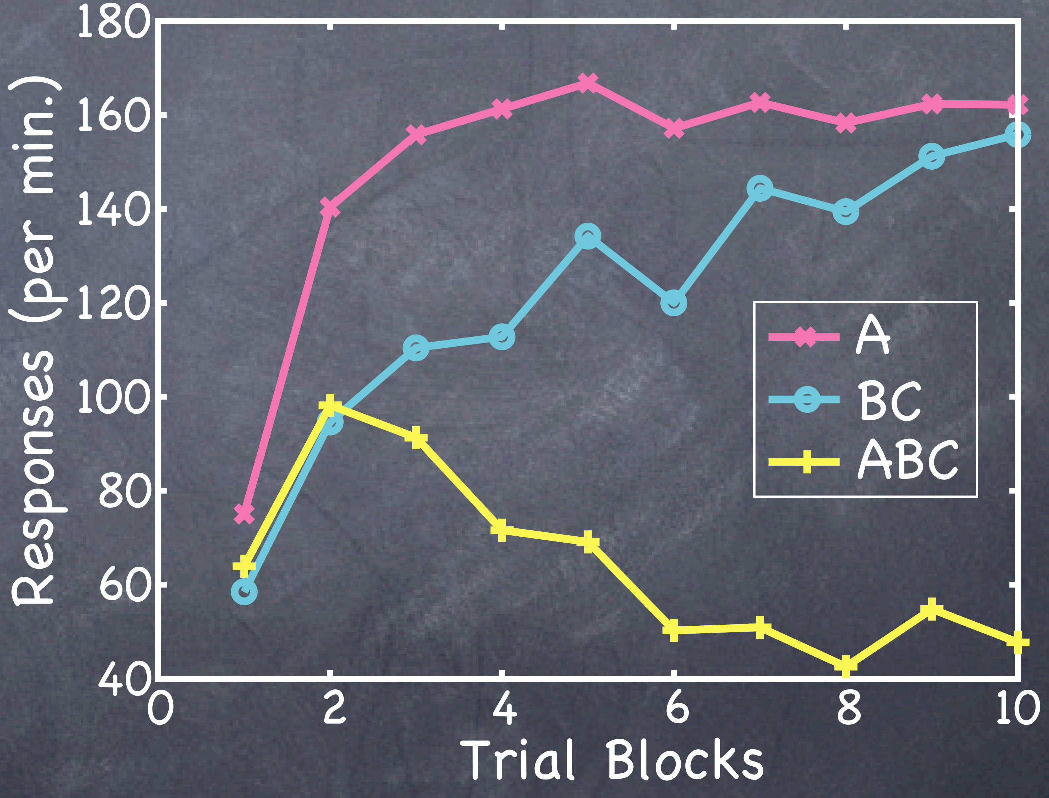
Pearce:



Expt. 2 Asymmetric XOR

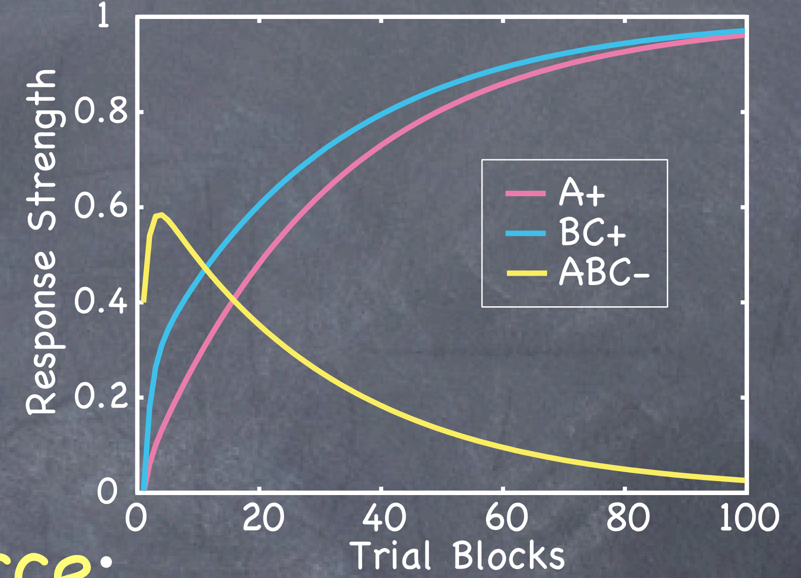
[Redhead & Pearce, 1995]

Training Trials:	Test Trials:
A+	A
BC+	BC
ABC-	ABC

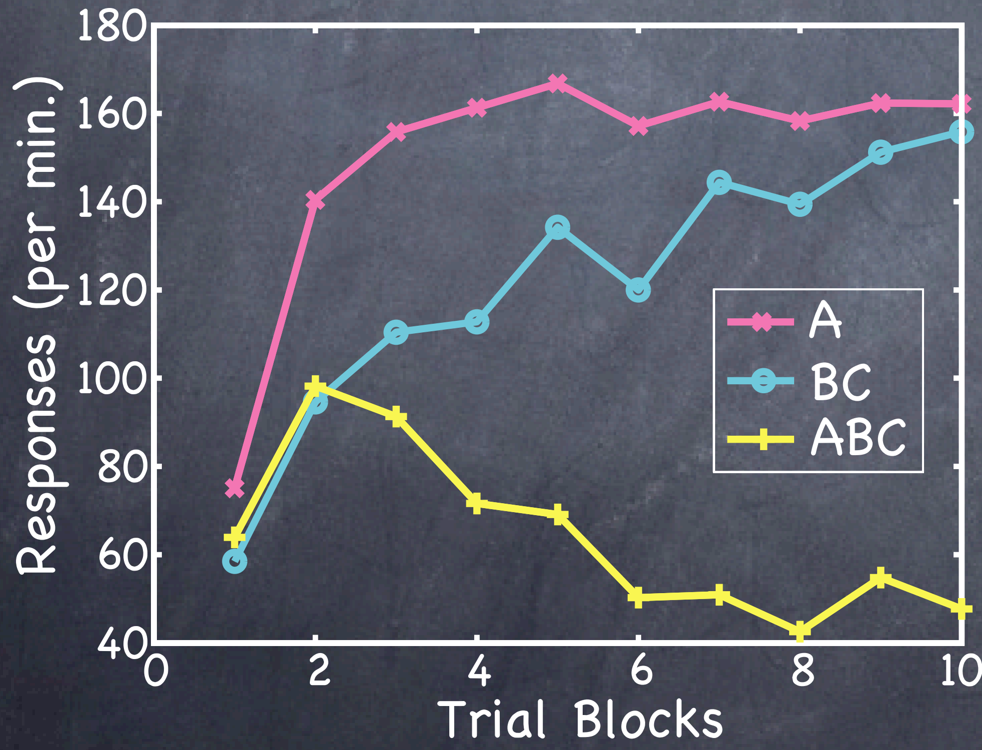


Modeling Asymmetric XOR

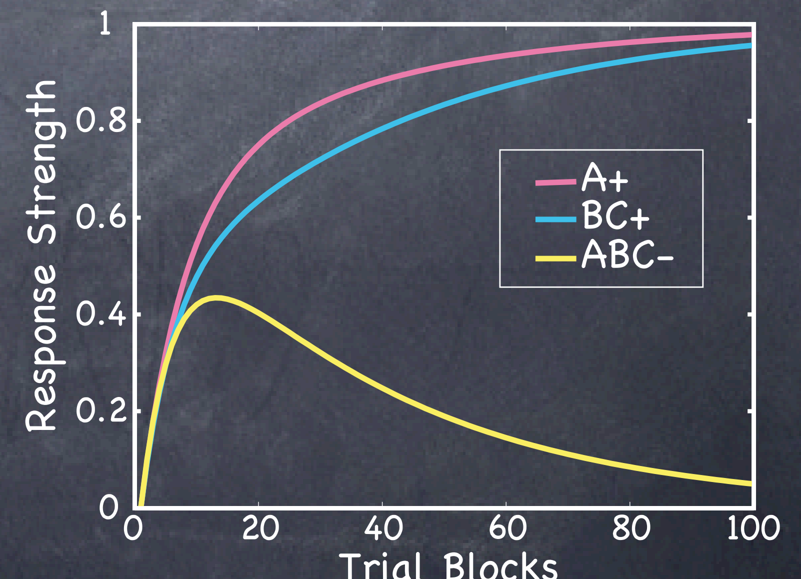
Rescorla-Wagner:



Training: A+ BC+ ABC-



Pearce:



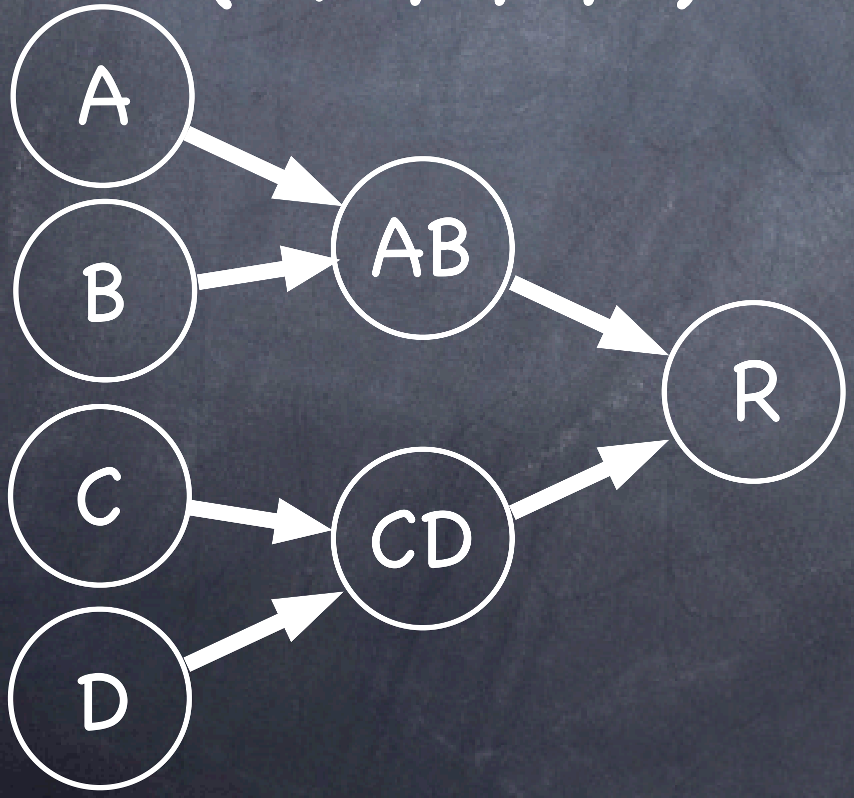
Issues in Modeling Configural Conditioning

- How do we choose between the two models?
- Similarity: How to measure similarity between patterns of stimuli?
- Discrimination: How do we choose a representation that is flexible enough?
- A formal Bayesian approach can guide us

Perspectives on modeling conditioning

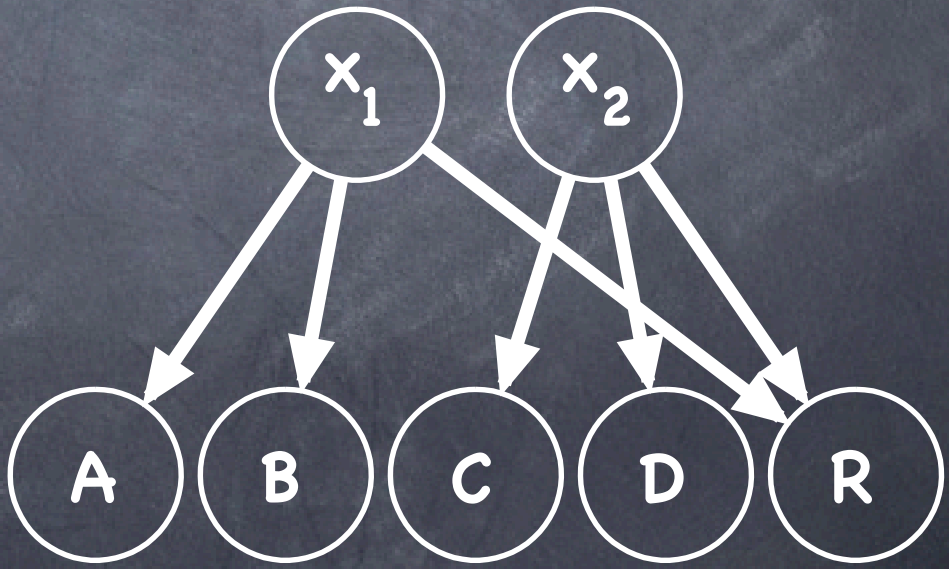
Discriminative

P(R|A,B,C,D)



Generative

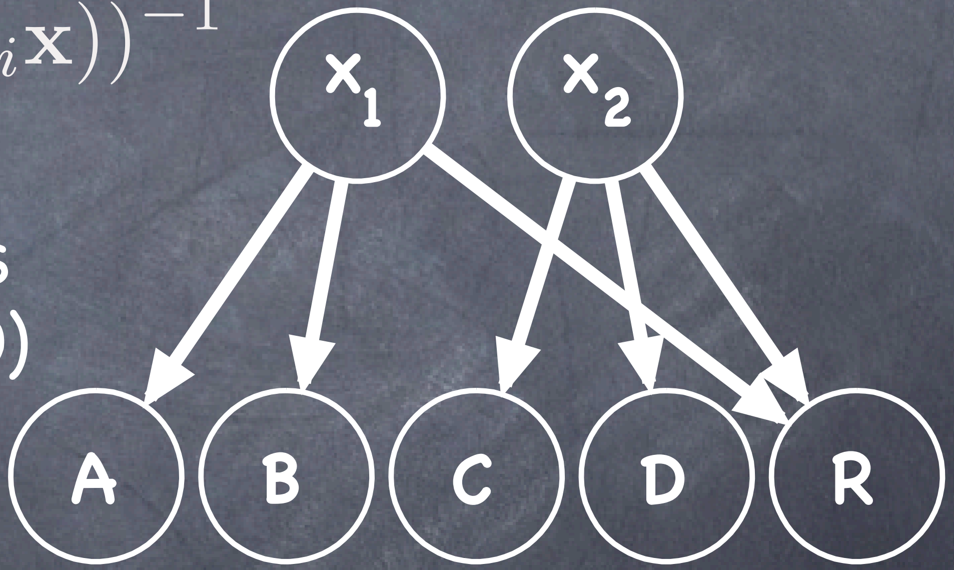
P(R,A,B,C,D)



A latent variable model

- Generative model: sigmoid belief network.

$$P(S_i \mid \mathbf{x}) = (1 + exp(-\mathbf{w}_i \mathbf{x}))^{-1}$$



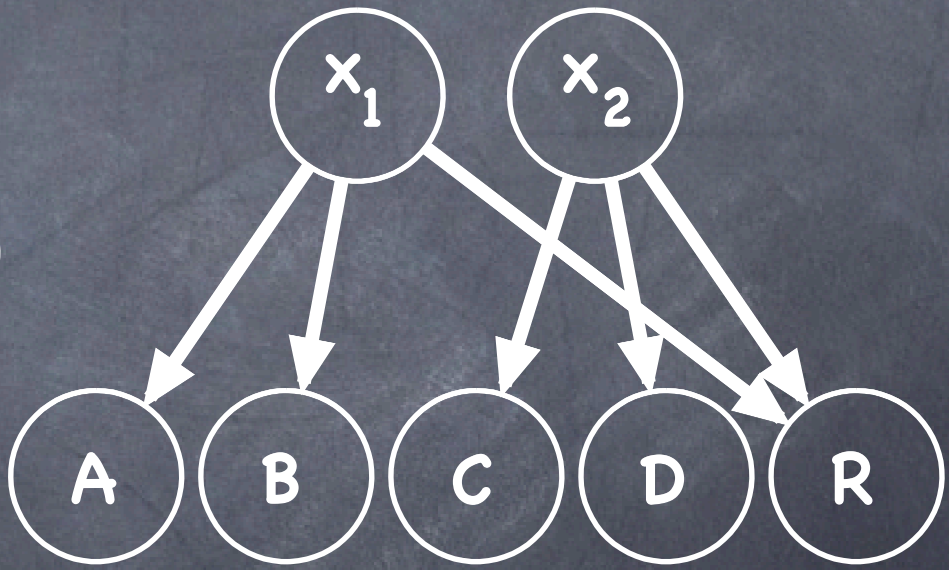
- Stimuli and Latent variables are binary (on = 1, off = 0)

-

- Latent variables ⇨ configural unit
  - correlate stimuli

Model Inference

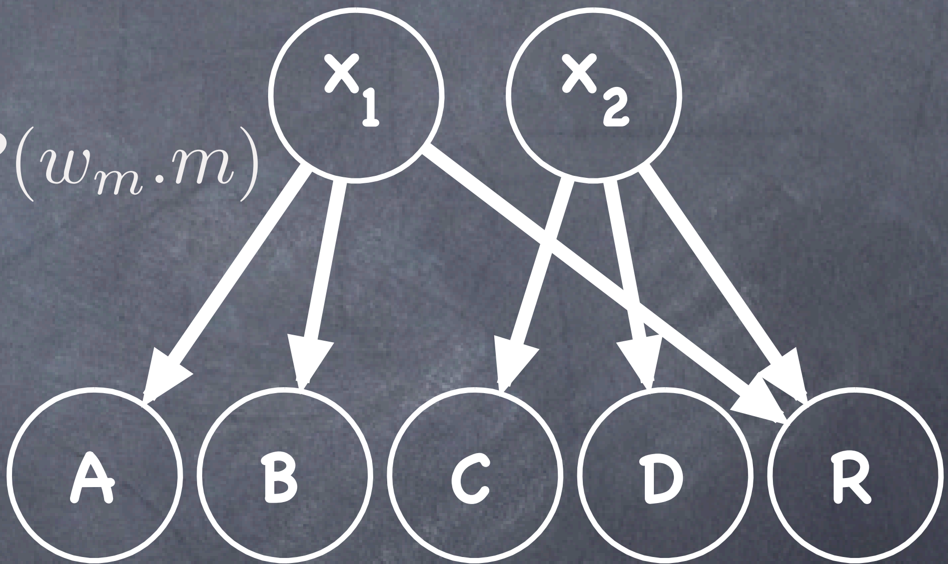
- Learning: P(w,m | D)
- Prediction: P(R | Stim,D)



Learning in the L.V. model

- Learning = Bayesian inference over weights & model structure conditional on training data

$$P(w_m, m \mid \mathcal{D}) \propto P(\mathcal{D} \mid w_m, m) P(w_m . m)$$

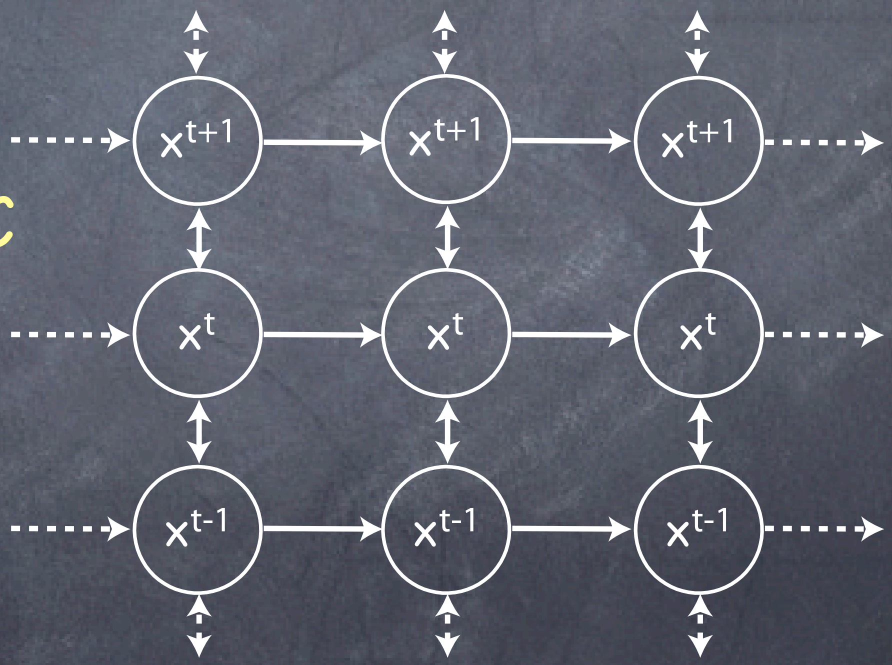


- Latent variable is unknown and unwanted so we compute the marginal likelihood:

$$P(\mathcal{D} \mid w_m, m) = \prod_t \sum_x \prod_i P(S_{t,i} \mid x, w_m, m) P(x \mid w_m, m)$$

Approximate inference

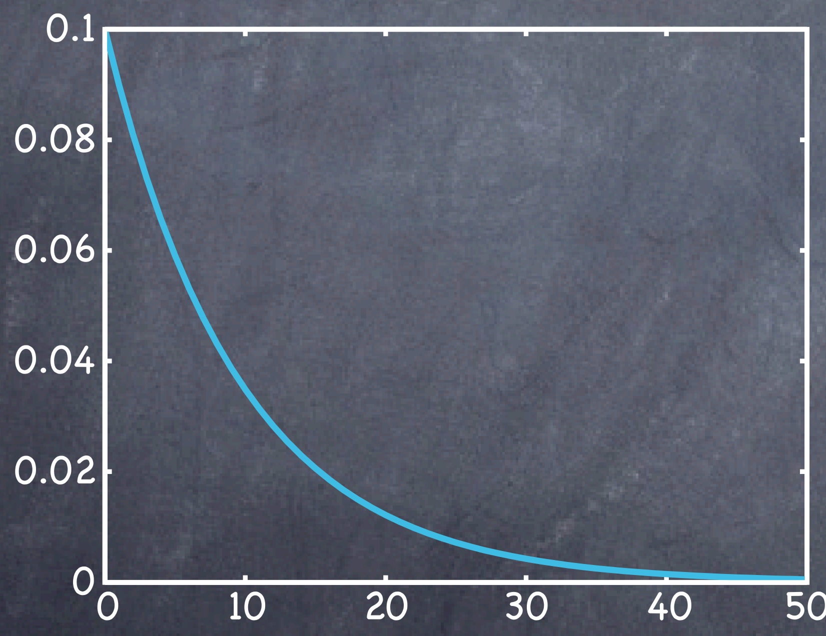
- Inference is analytically intractable: use reversible-jump MCMC
- reversible-jump mixes slowly: Exchange MCMC method to help



L.V. model priors

Prior over number of latent variables:

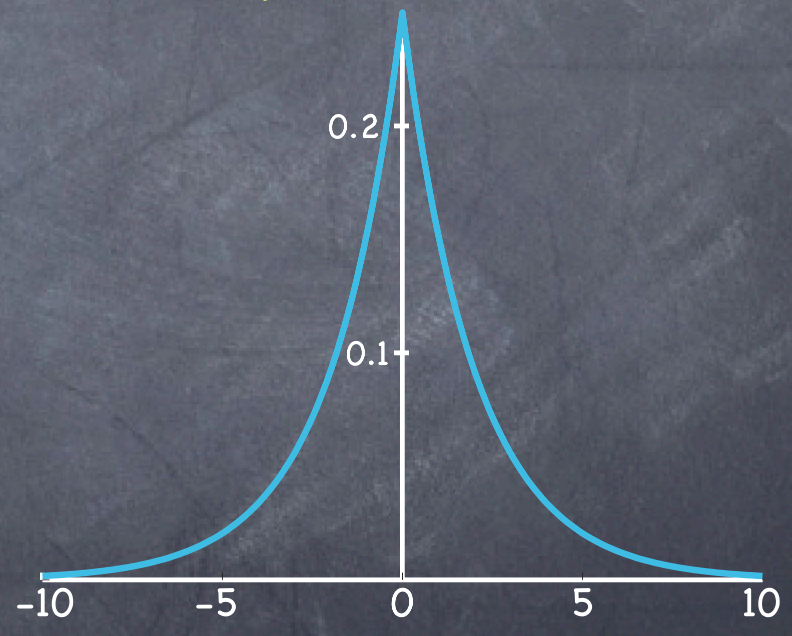
Geometric(0.1)



num. latent variables

Prior over weight magnitudes:

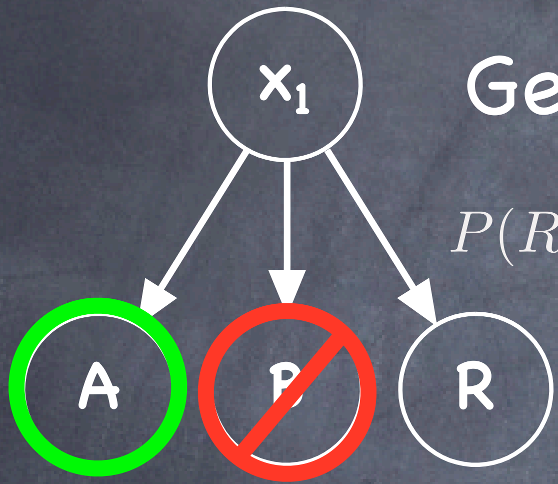
Laplace(2.0)



weight

Additional assumption: Stimuli are a priori rare

Prediction



Generalization => inference over latents

$$P(R \mid A, \bar{B}, m, \mathbf{w}_m) = \sum_{x_1} P(R \mid x_1, m, \mathbf{w}_m) P(\mathbf{x} \mid A, \bar{B}, m, \mathbf{w}_m)$$

$$P(x_1 \mid A, \bar{B}, m, \mathbf{w}_m) \propto P(A \mid x_1, m, \mathbf{w}_m) P(\bar{B} \mid x_1, m, \mathbf{w}_m) P(x_1 \mid m, \mathbf{w}_m)$$

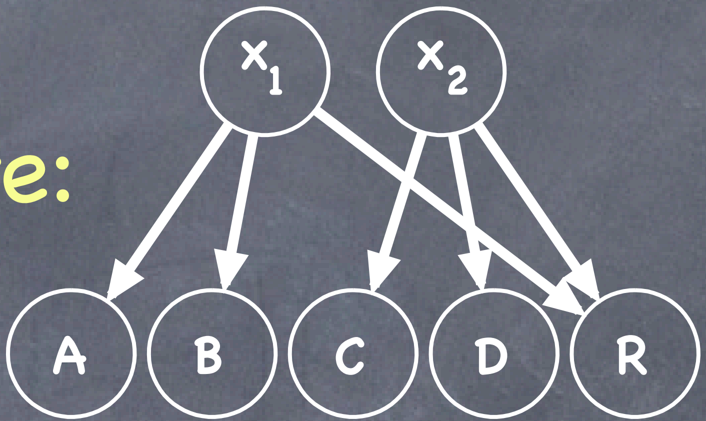
Posterior reinforcement prediction: marginalize over choice of weights and model structure.

$$P(R \mid Stim, m, \mathcal{D}) = \int P(R \mid Stim, m, \mathbf{w}_m, \mathcal{D}) P(\mathbf{w}_m \mid m, \mathcal{D}) \, d\mathbf{w}_w$$

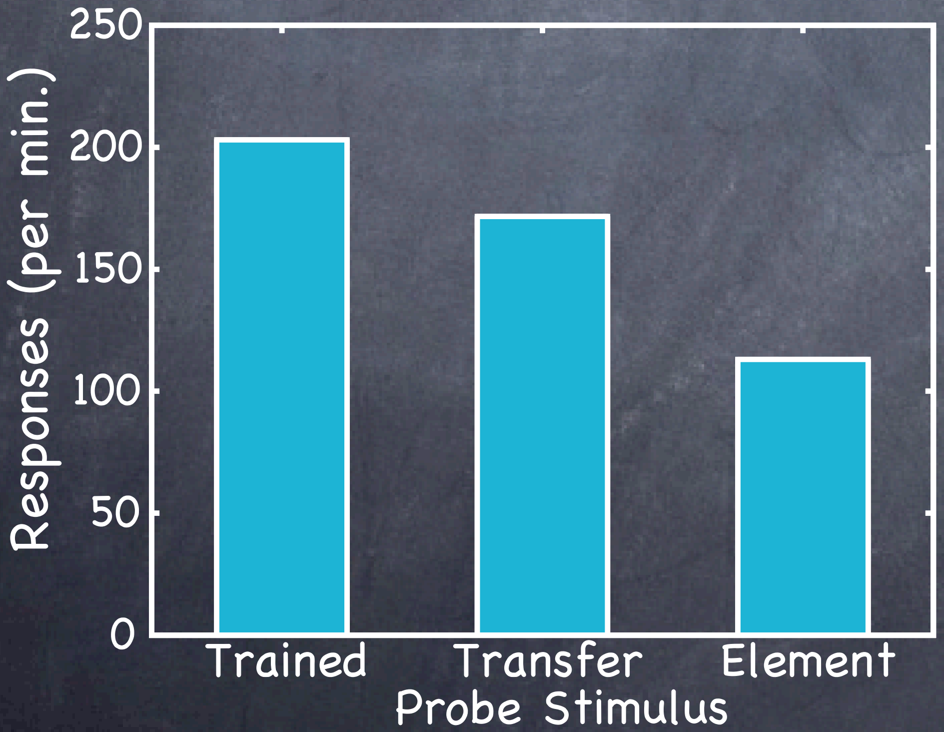
$$P(R \mid Stim, \mathcal{D}) = \sum_m P(R \mid Stim, m, \mathcal{D}) P(m \mid \mathcal{D})$$

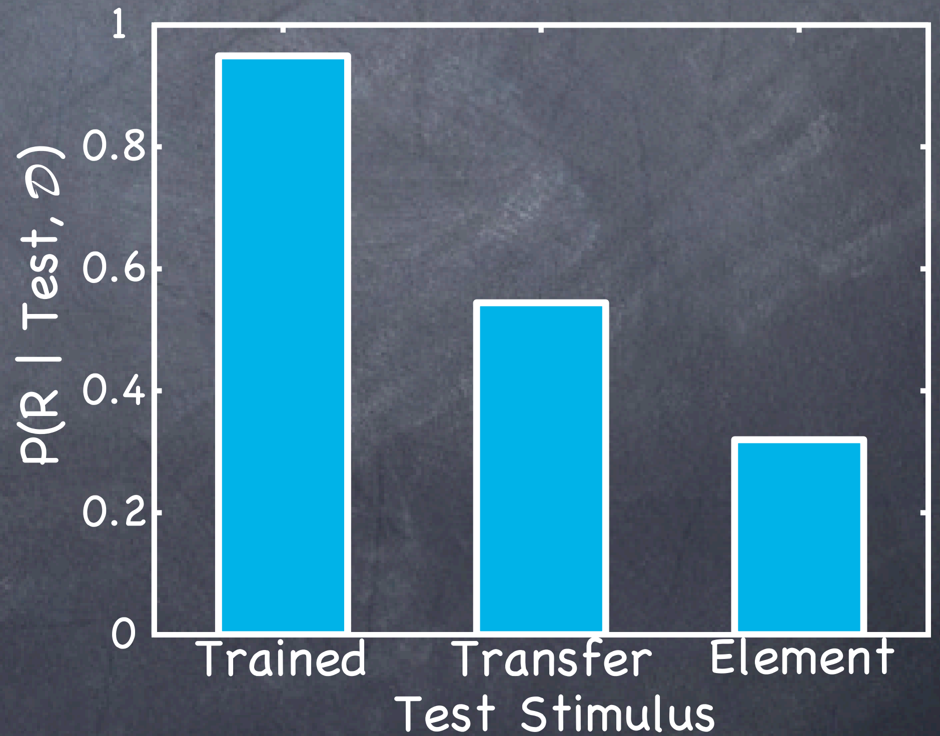
L.V. Model of Paired Compounds

MAP Model Structure:



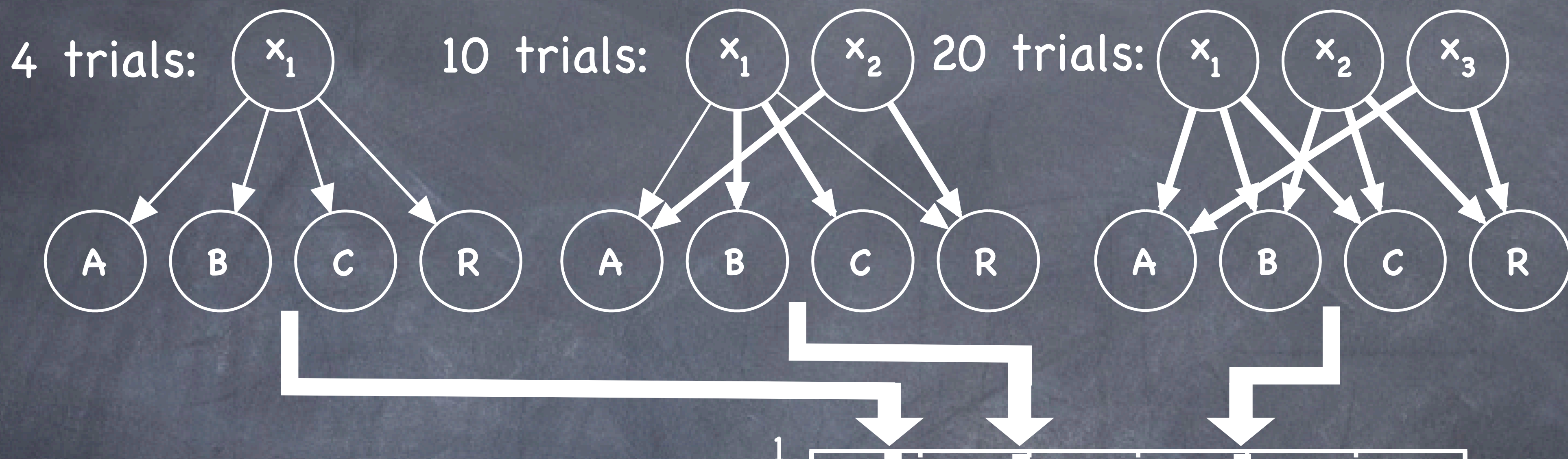
Training: AB+ CD+



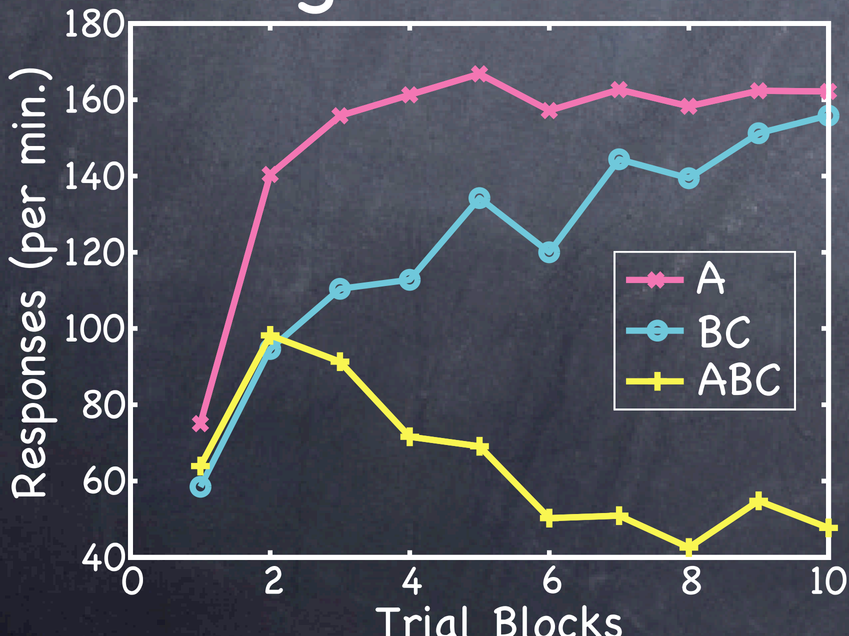


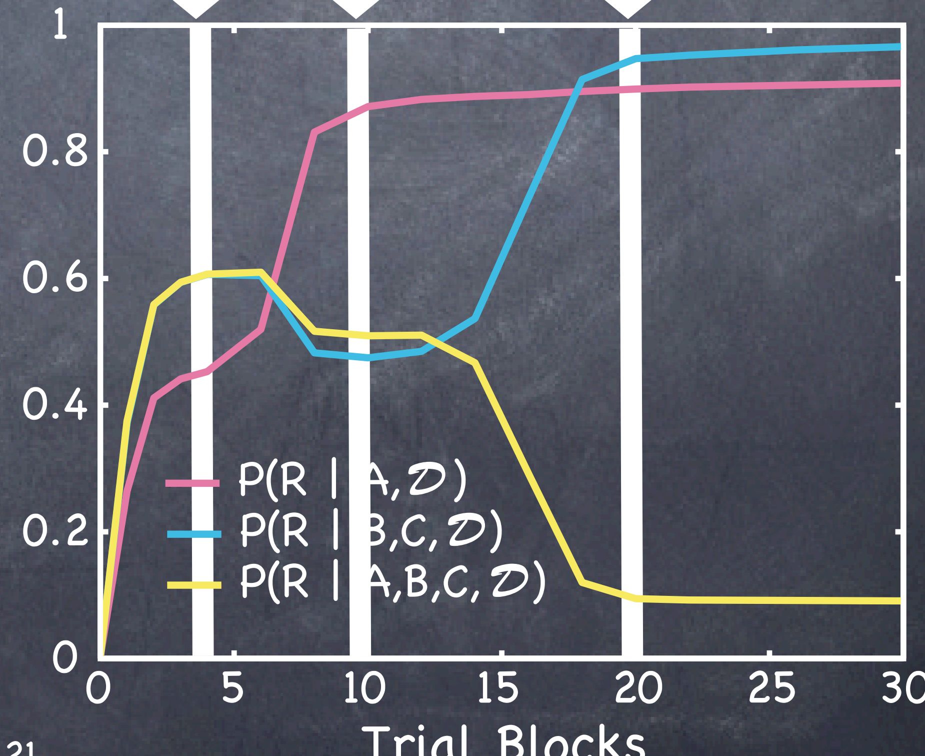
L.V. Model of Asymmetric XOR

MAP Model Structures:



Training: A+ BC+ ABC-





What's a configuration?

- Can account for experiments that are traditionally deemed "Configural Conditioning"

- Previous models cast configuration as result of stimuli being trained together.

- We view it as the result of model complexity pressures to group stimuli.

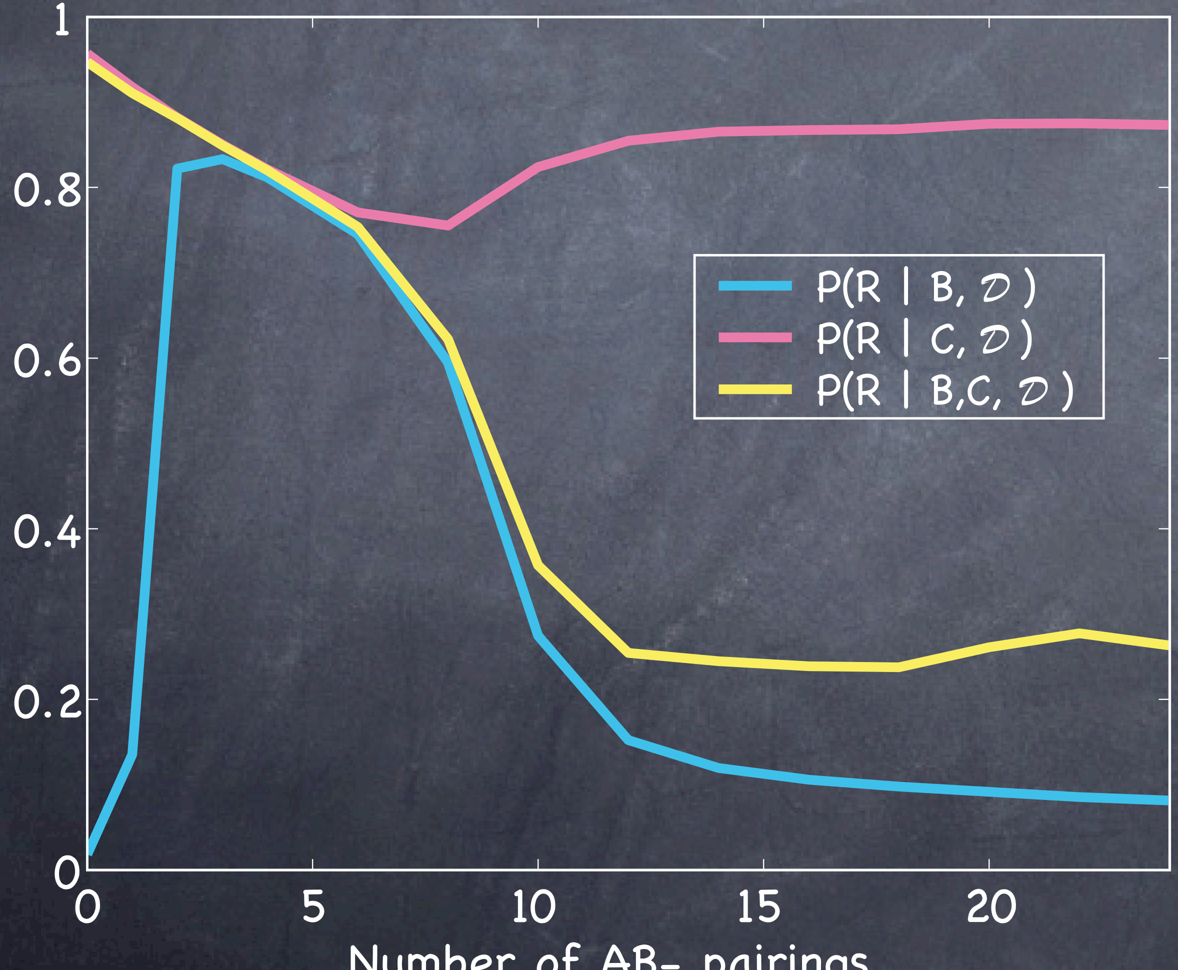
Expt: Second-order conditioning versus Conditioned Inhibition

[Yin et al, 1994]

Group	A+	AB-	C+	Test ➡ Result	Test ➡ Result
No B	96	0	8	B ➡ _	BC ➡ Resp.
Few B	96	4	8	B ➡ Resp.	BC ➡ Resp.
Many B	96	48	8	B ➡ _	BC ➡ _

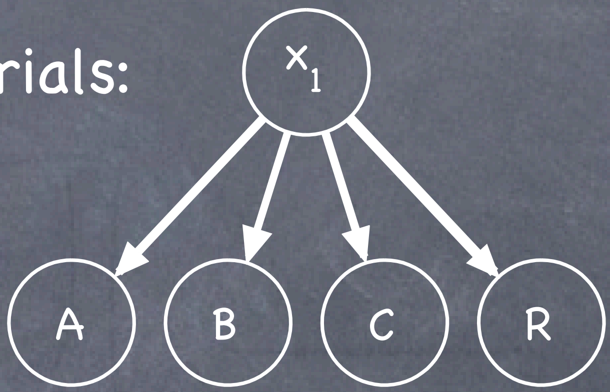
Bayesian Model of Second-Order Conditioning / Conditioned Inhibition

Training: A+ AB- C+

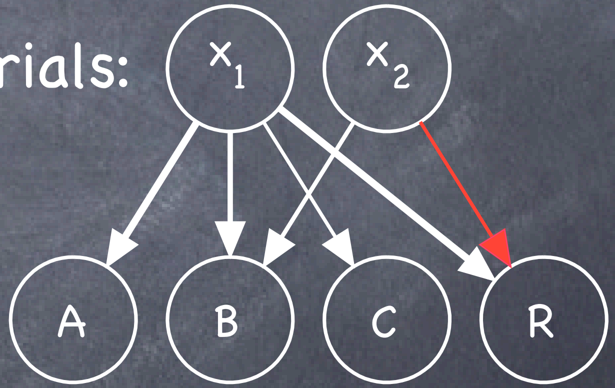


MAP Model Structure:

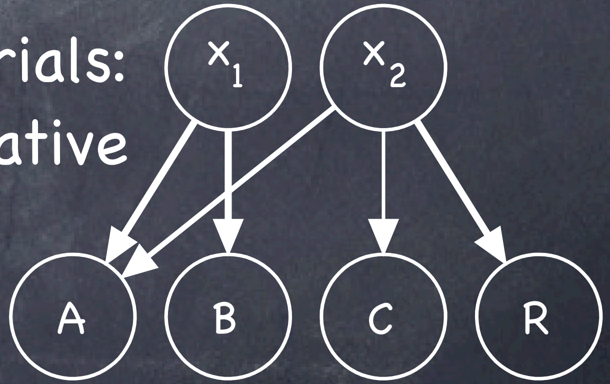
4 trials:



18 trials:



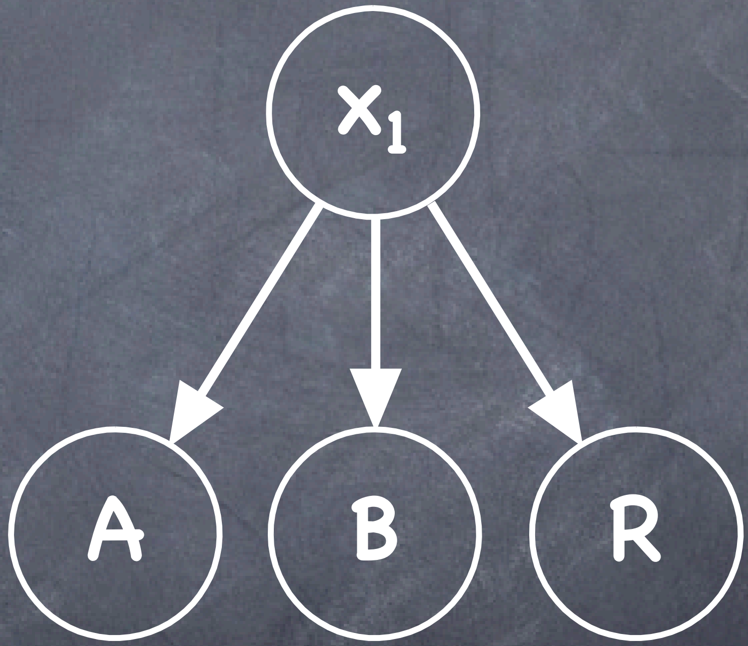
18 trials: Alternative



Dealing with Reinforcement

Are reinforcers really just like other stimulus?

Train: A+ B+

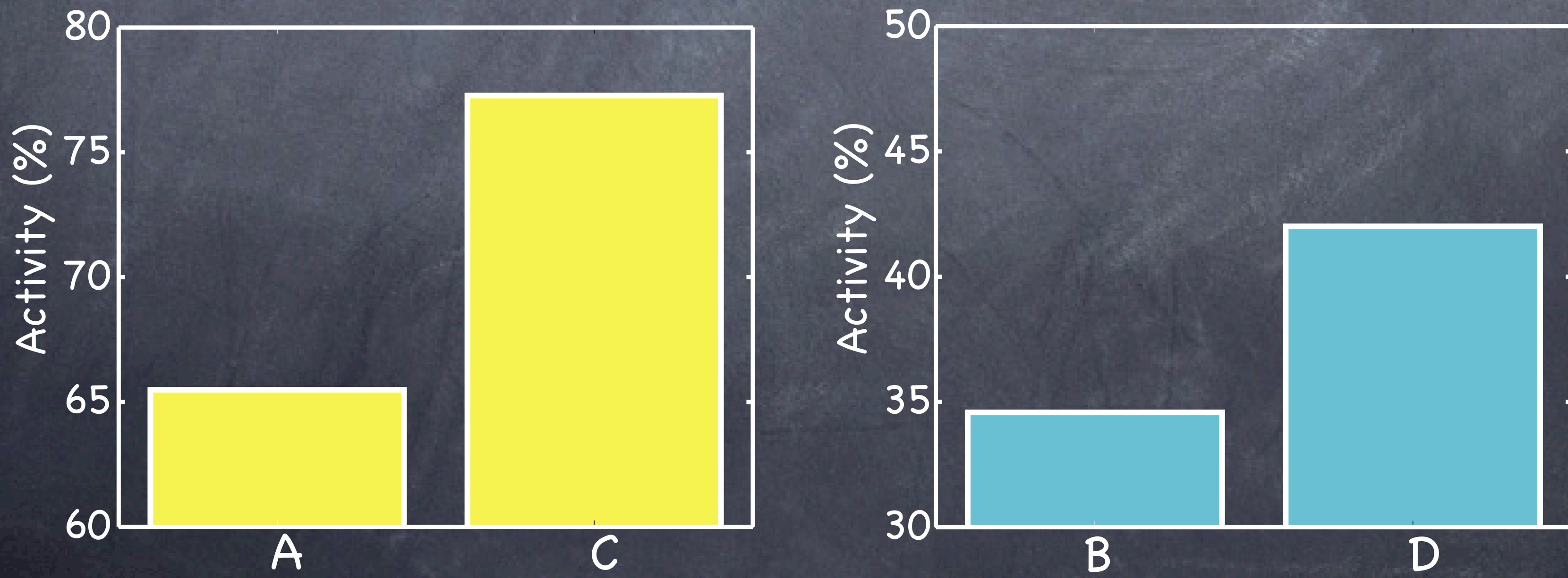


Do animals do this?

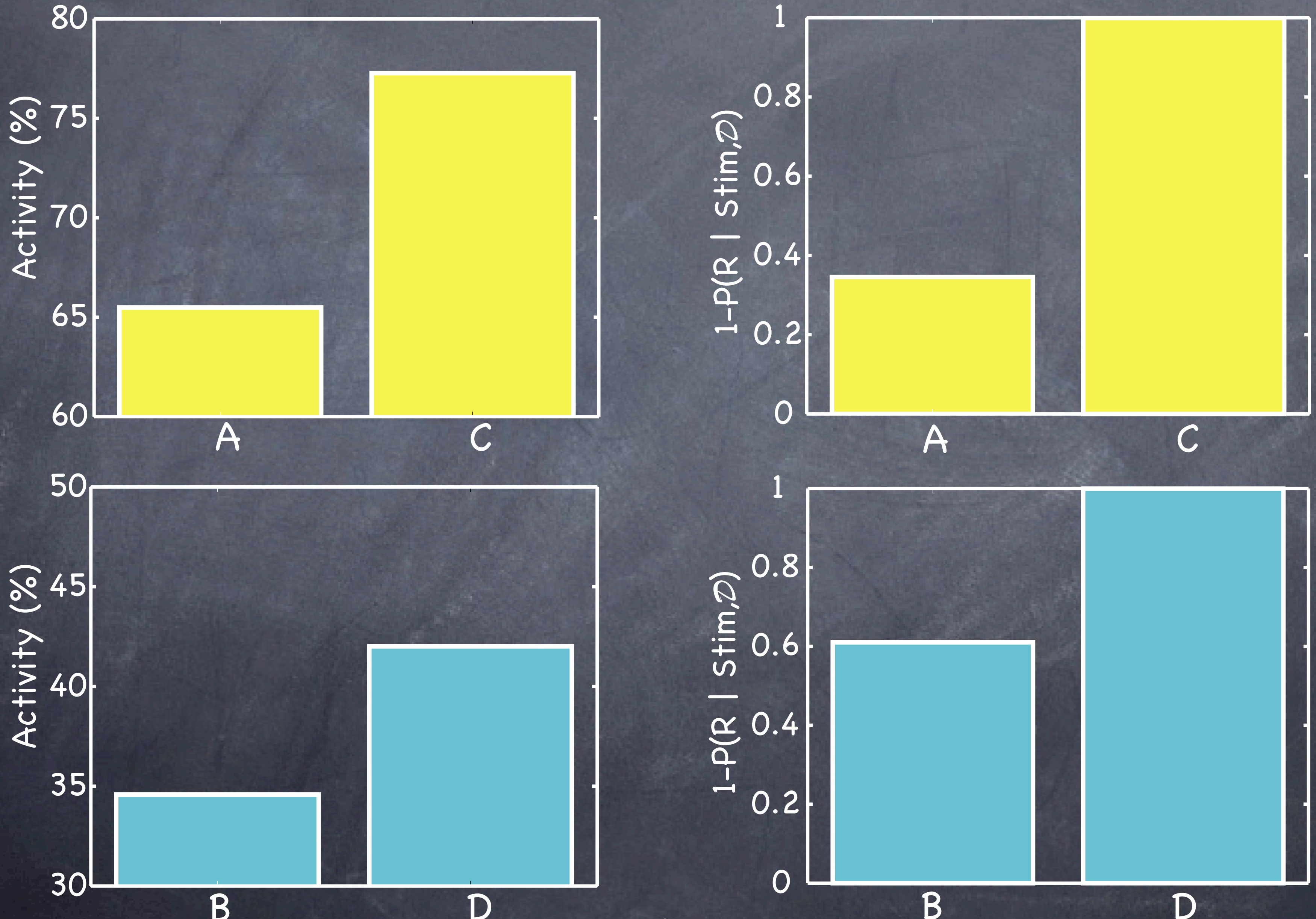
Acquired Relational Equivalence

[Honey & Watt, 1999]

Biconditional training		Revaluation	Test
AY-food	AZ-no food	A-shock	A vs C
BY-food	BZ-no food		
CY-no food	CZ-food	C-no shock	B vs D
DY-no food	DZ-food		

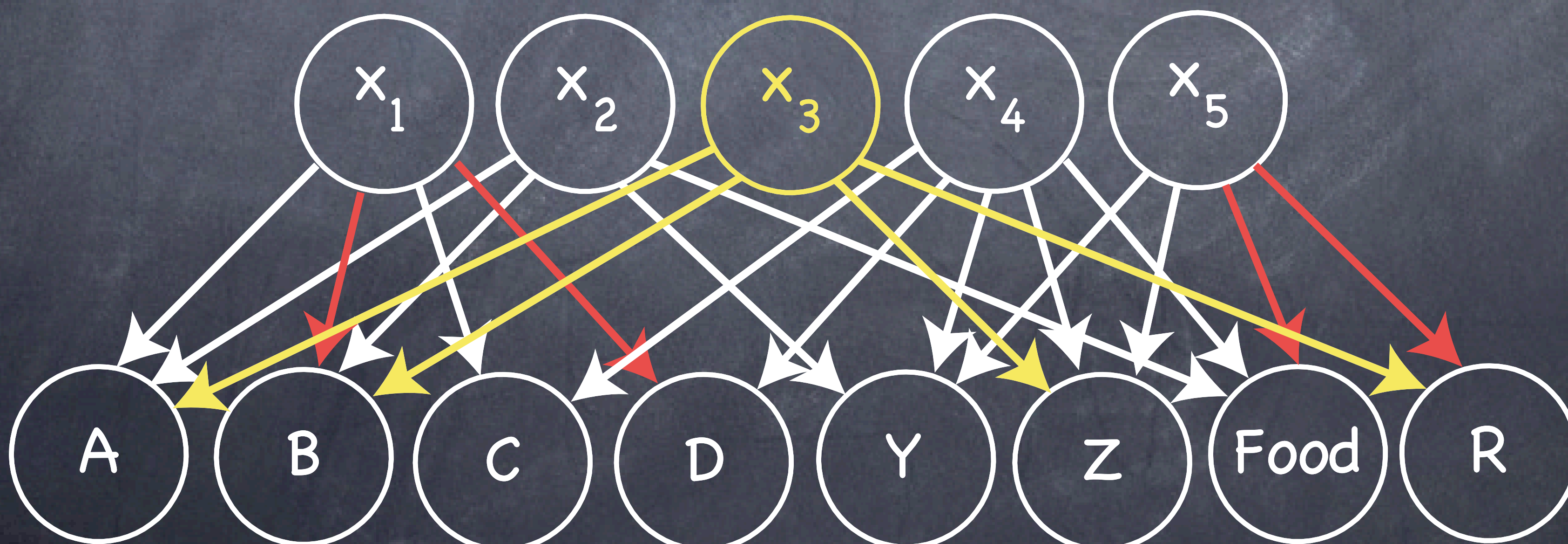


Modeling Acq. Rel. Equiv.



Modeling Acq. Rel. Equiv.

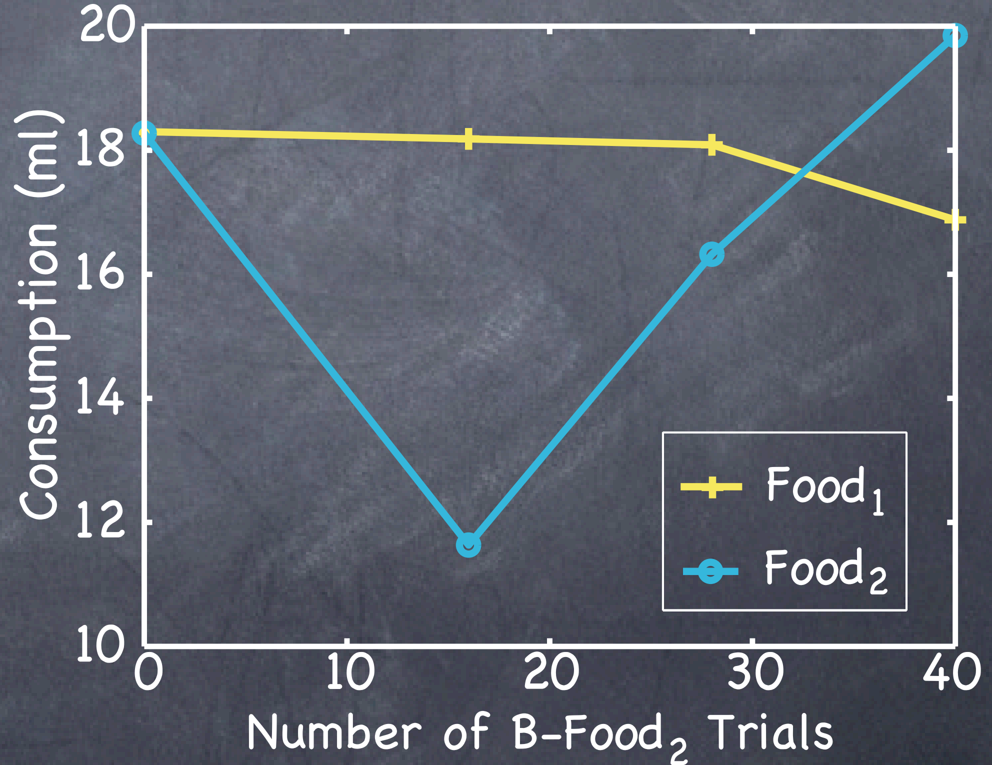
Biconditional training		Revaluation	Test
AY-food	AZ-no food	A-shock	A vs C
BY-food	BZ-no food		
CY-no food	CZ-food	C-no shock	B vs D
DY-no food	DZ-food		



Expt: Food Devaluation

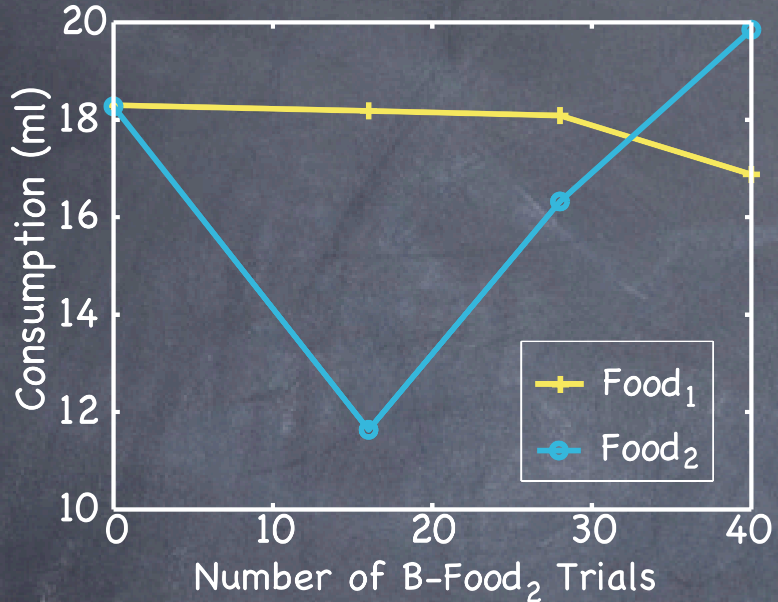
[Holland, 1998]

<table>
<tr><th colspan="3">Training Trials</th><th rowspan="3">Test</th></tr>
<tr><th colspan="2">Phase 1</th><th>Phase 2</th></tr>
<tr><th>A-F1</th><th>B-F2</th><th>B-R</th></tr>
<tr><td rowspan="4">16</td><td>0</td><td rowspan="4">6</td><td rowspan="4">F1
F2</td></tr>
<tr><td>16</td></tr>
<tr><td>28</td></tr>
<tr><td>40</td></tr>
</table>

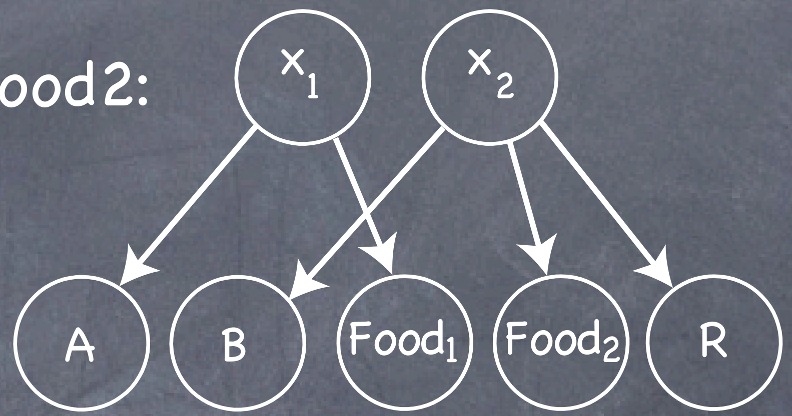


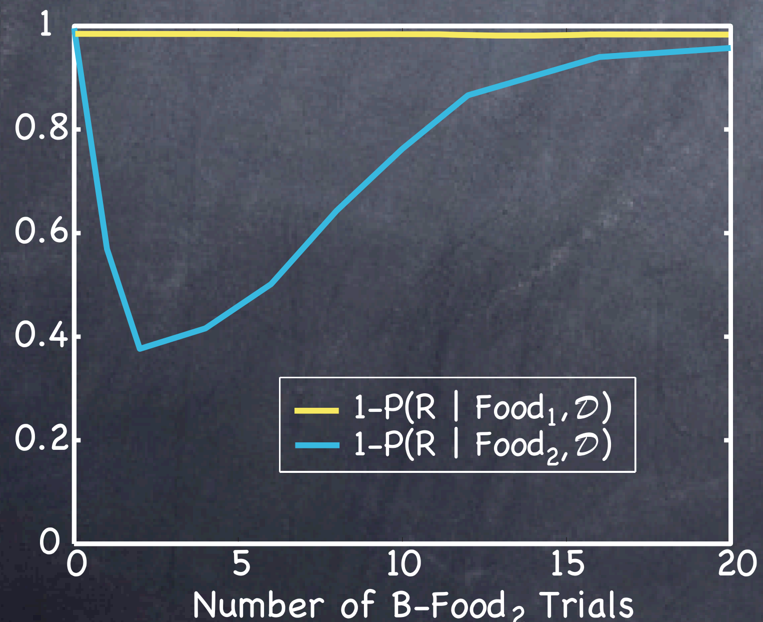
L.V. model of Devaluation

A-Food1 / B-Food2, B-R

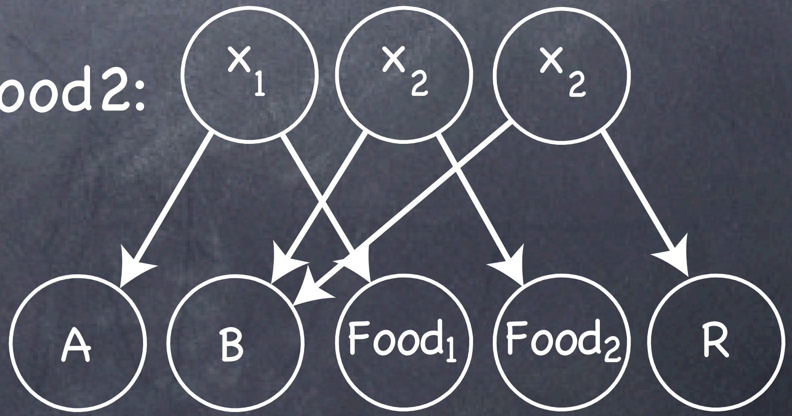


Few B-Food2:





Many B-Food2:

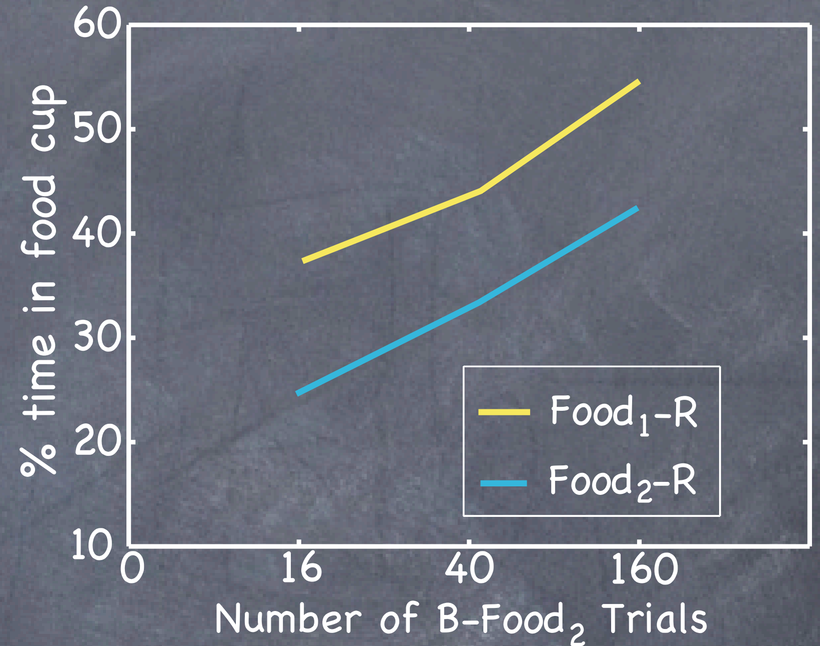


Not the whole story...

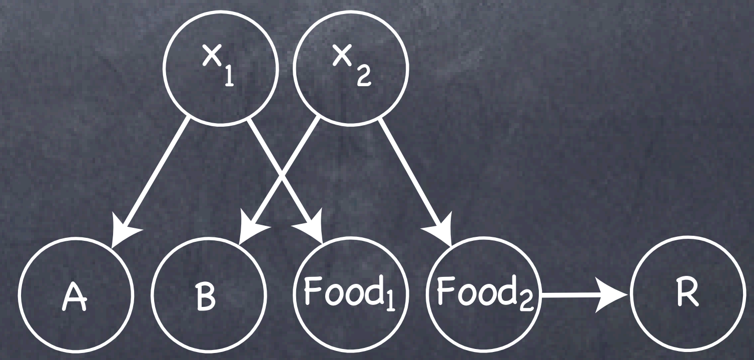
(Variant on Devaluation)

[Holland, 1998]

Training Trials			Test
Phase 1		Phase 2	
A-F1	B-F2	F(1,2)-R	
16	16	2	B
	40		
	160		



Model Structure?



Future Directions

- Explore the priors: Experimentally manipulatable.
- Remove independent trial assumption.

Modeling Change

- Should reflect our understanding of how the world is believed to change.
- Example: Causal model parameter drift.
- The marginal distribution of the diffusion process should reflect your prior

Conclusions

- Similarity and Discrimination are recognized as the tradeoff between compexity and data fidelity arising in Bayesians inference.
- A latent variable is a natural (causal) setting for the study of classical conditioning.
- Account for configural conditionng data and more.