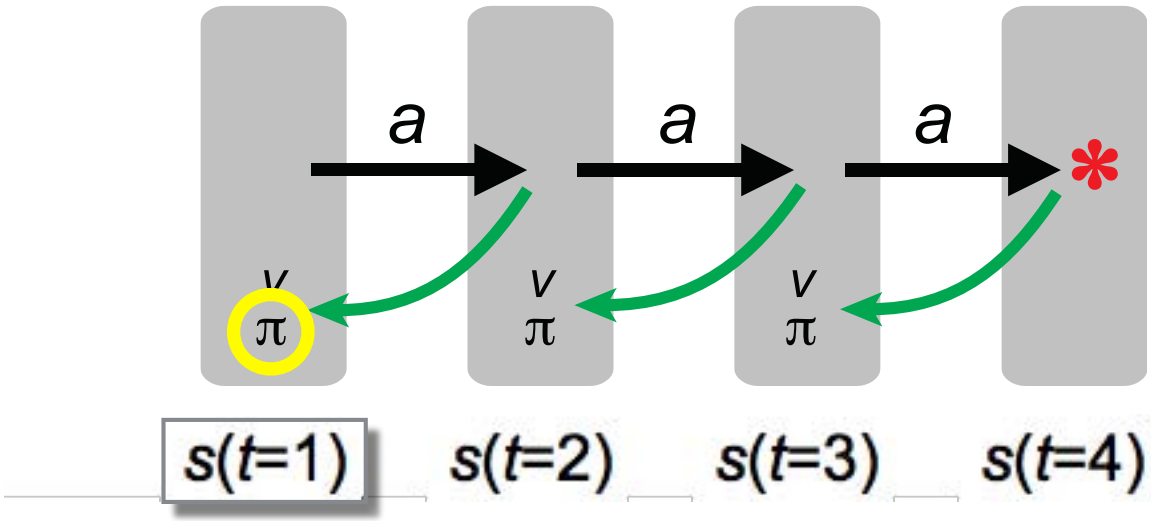
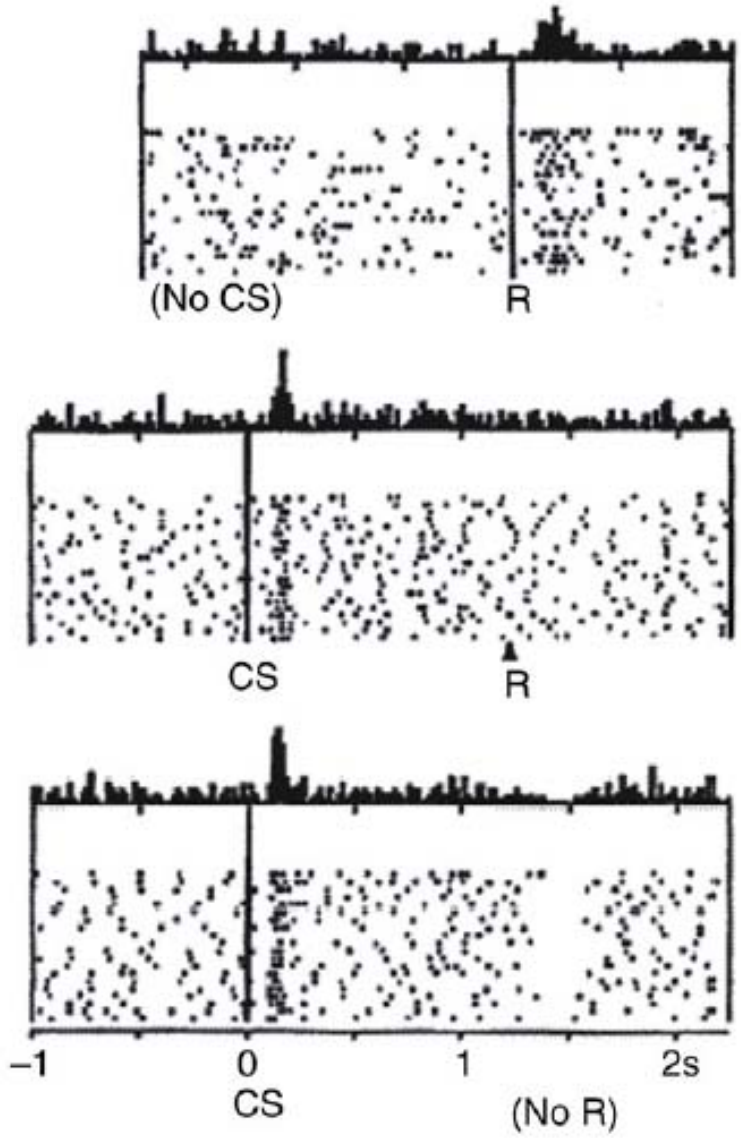
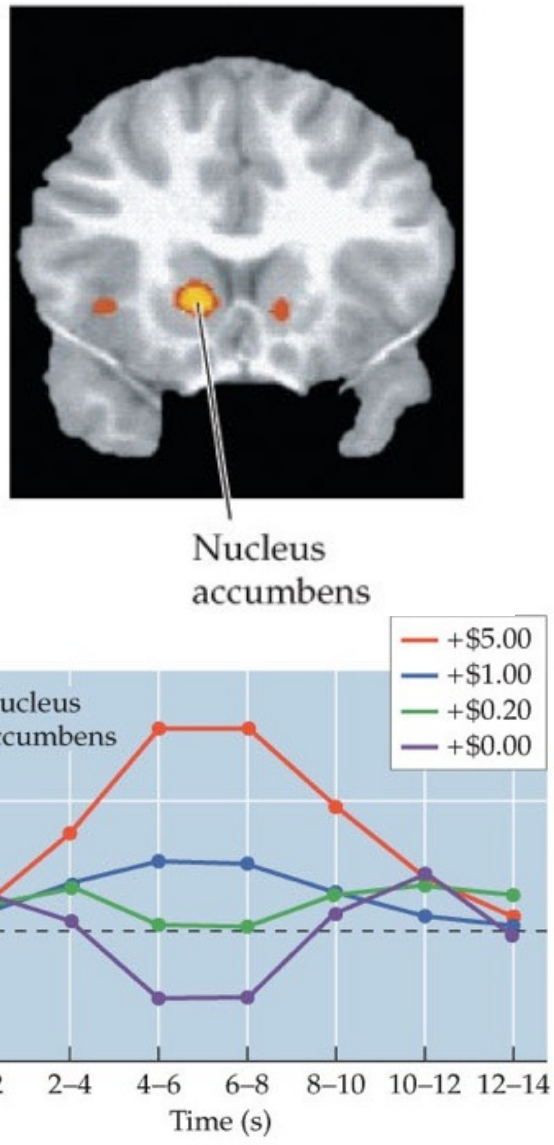
Hierarchical Reinforcement Learning and Human Behavior

Matthew Botvinick
Princeton Neuroscience Institute
and Department of Psychology
Princeton University

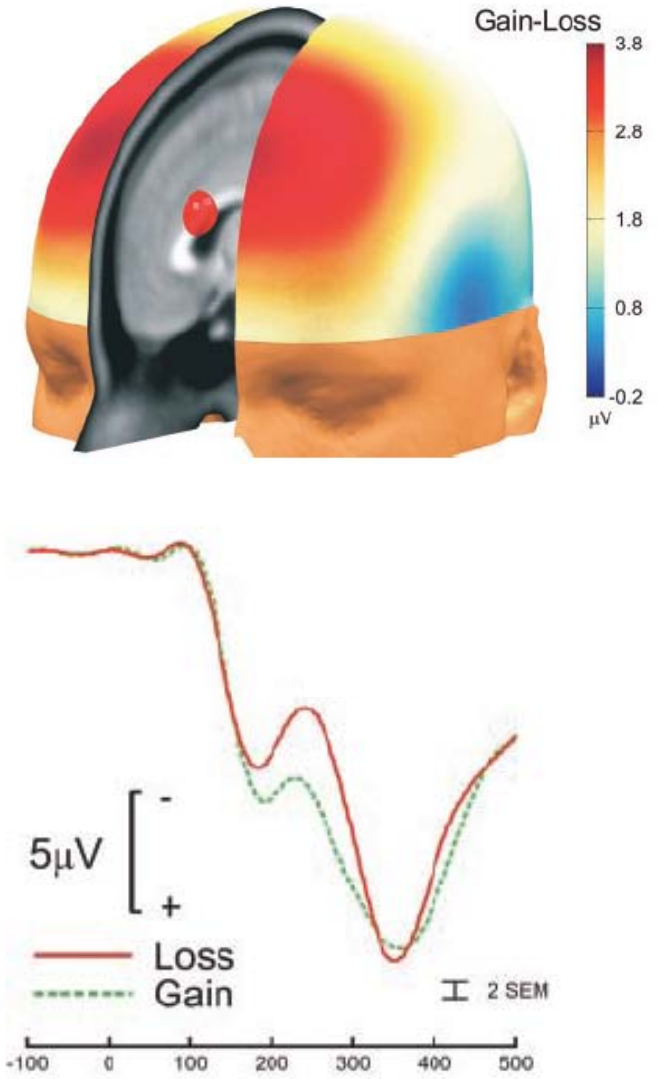

a
a
a
v
π
v
π
v
π
s(t=1)
s(t=2)
s(t=3)
s(t=4)



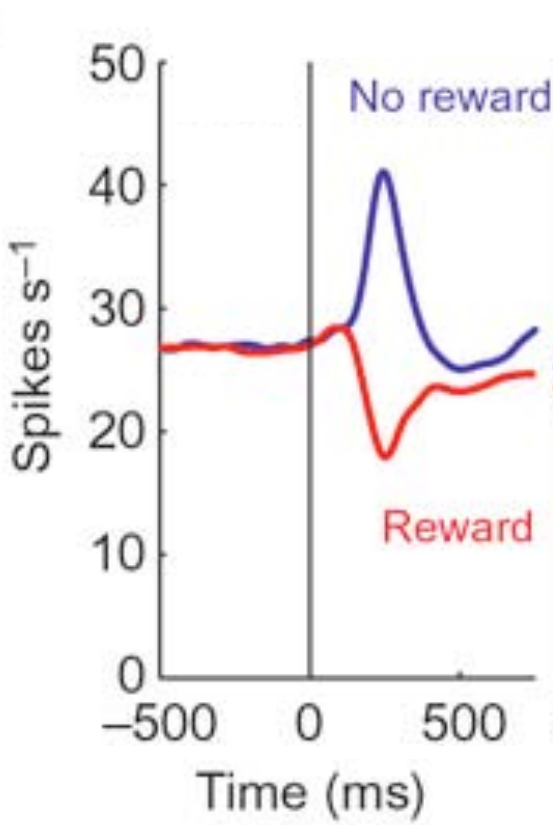
Schultz et al., *Science*, 1997



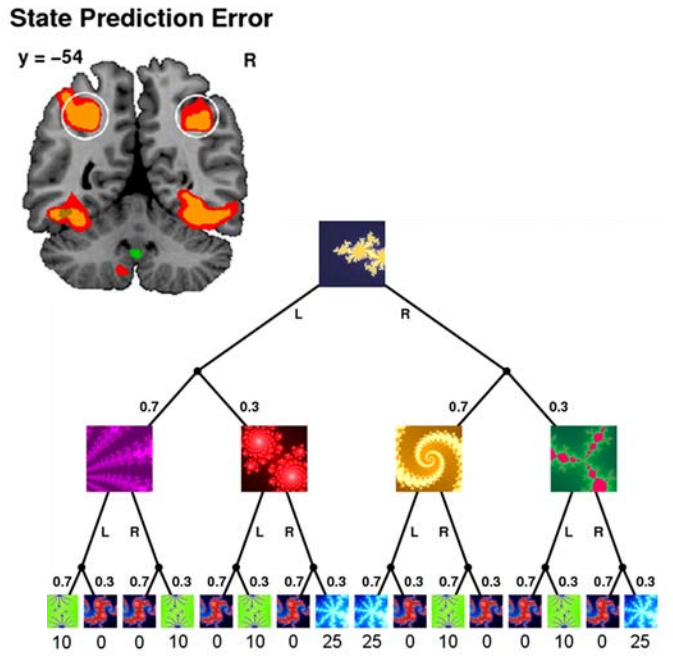
Knutson et al., *NeuroReport*, 2001



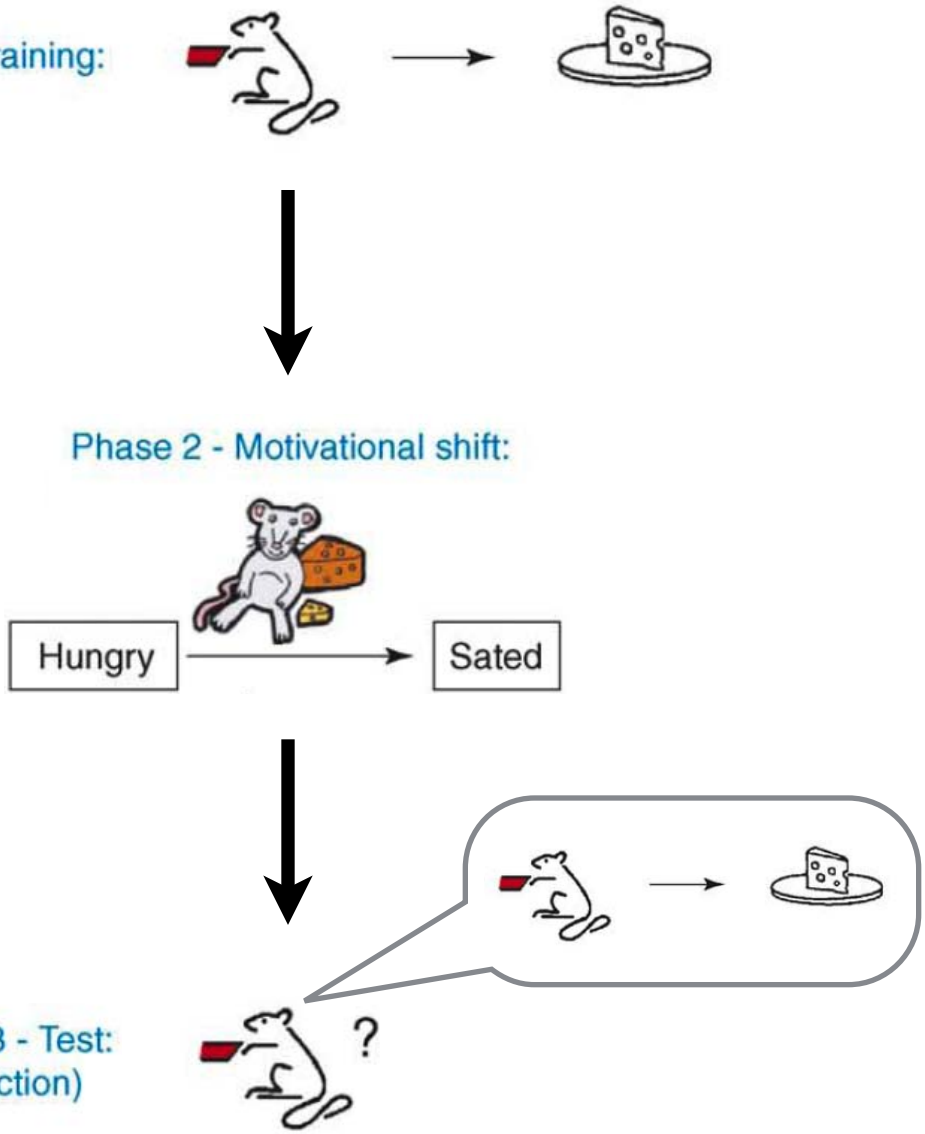
Gehring & Willoughby, *Science*, 2002



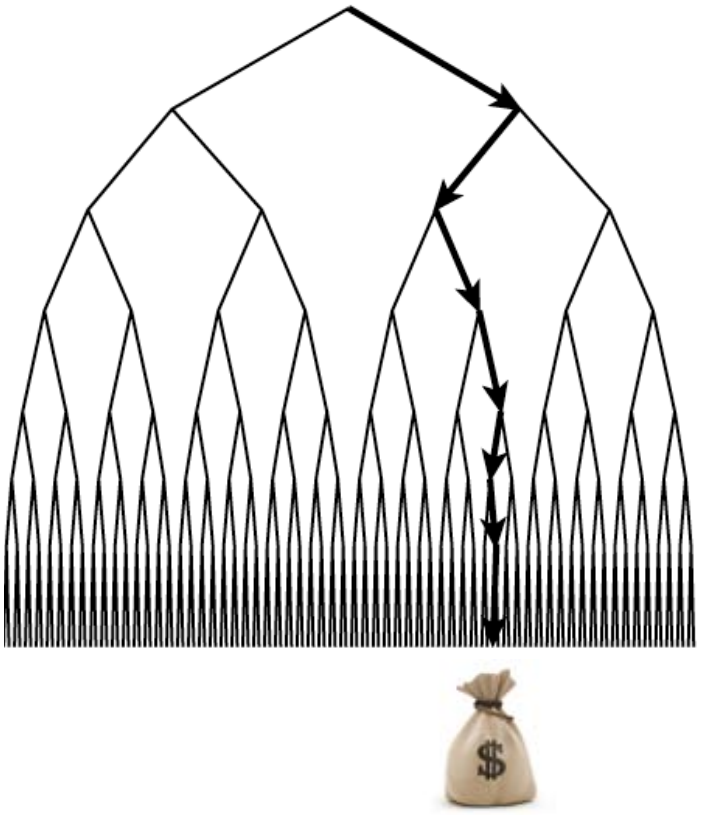
Matsumoto & Hikosaka, *Nature*, 2007



From Glasher, Daw, Dayan & O'Doherty, 2010

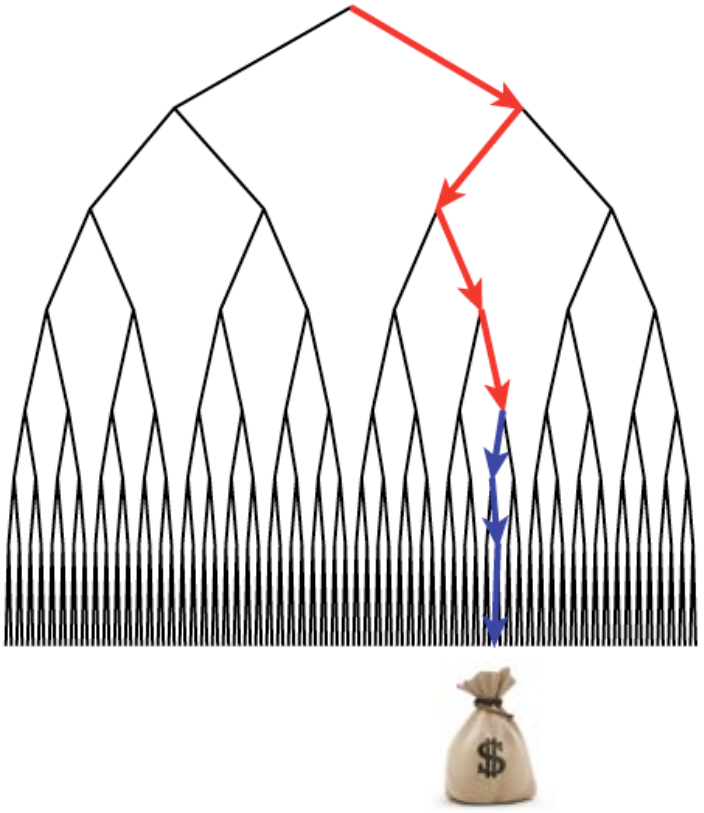


From Niv, Joel & Dayan, *TICS,* 2006 (artwork by B. Balleine)



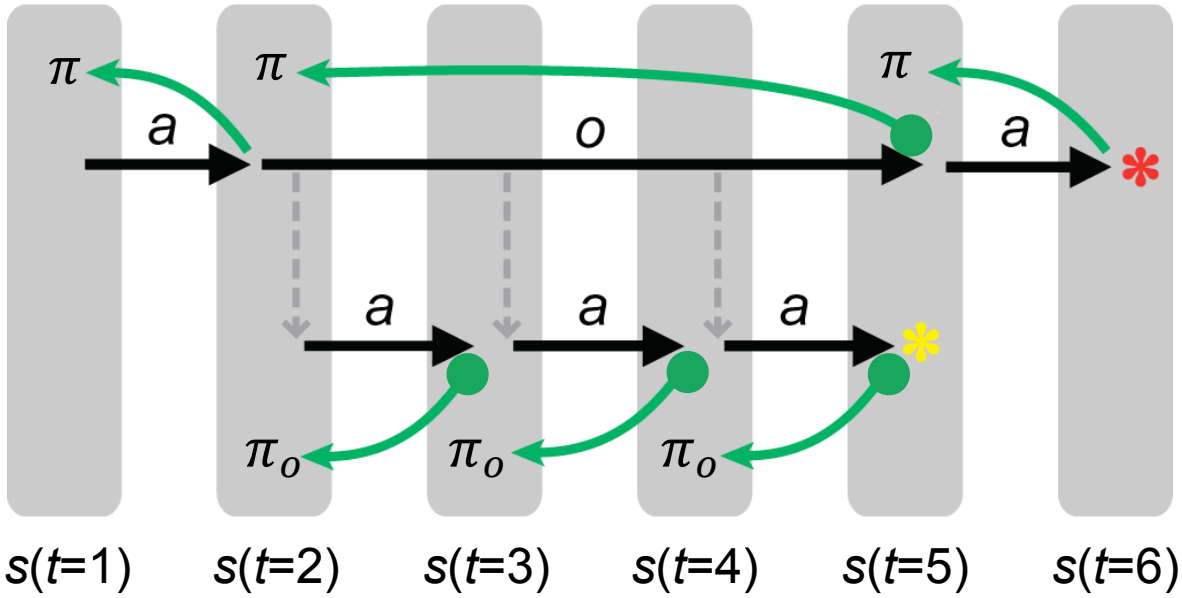
0.8%

The Curse of Dimensionality

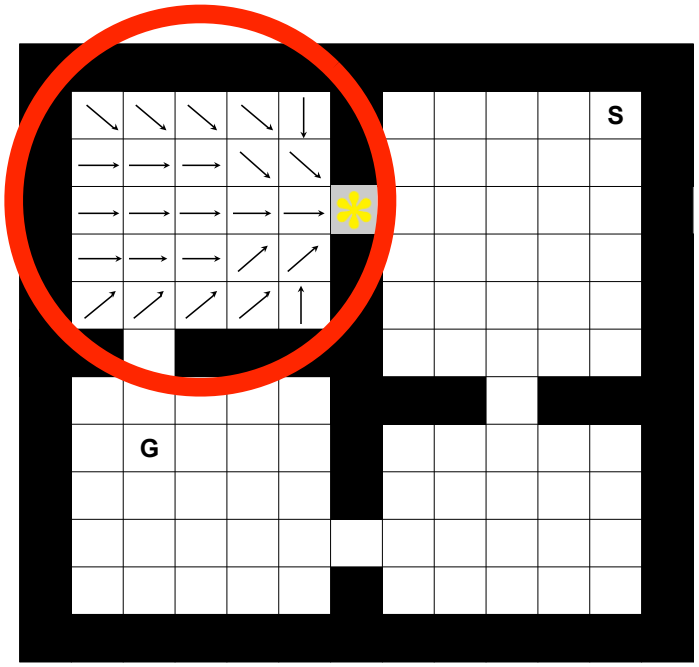


15.6%

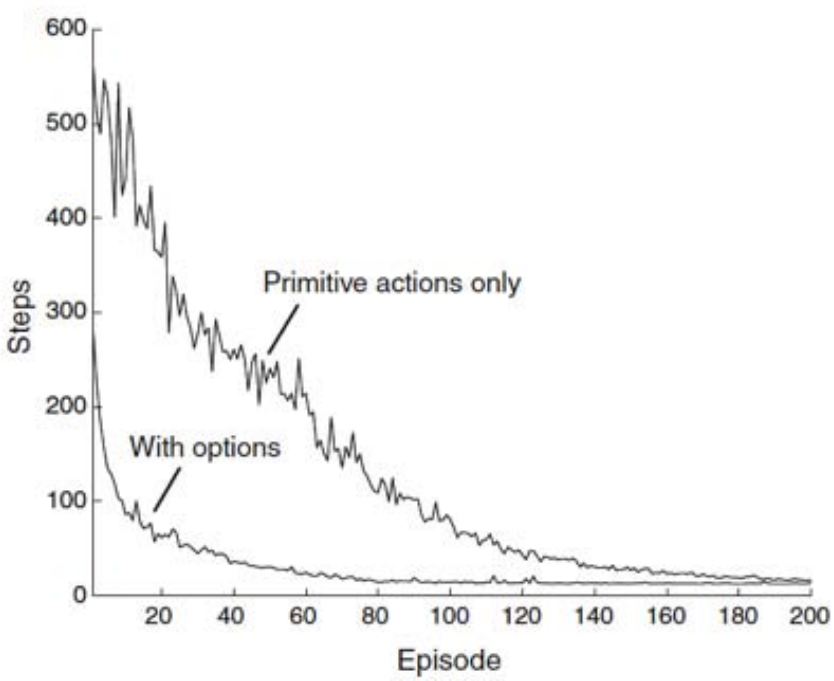
The Blessing of Abstraction



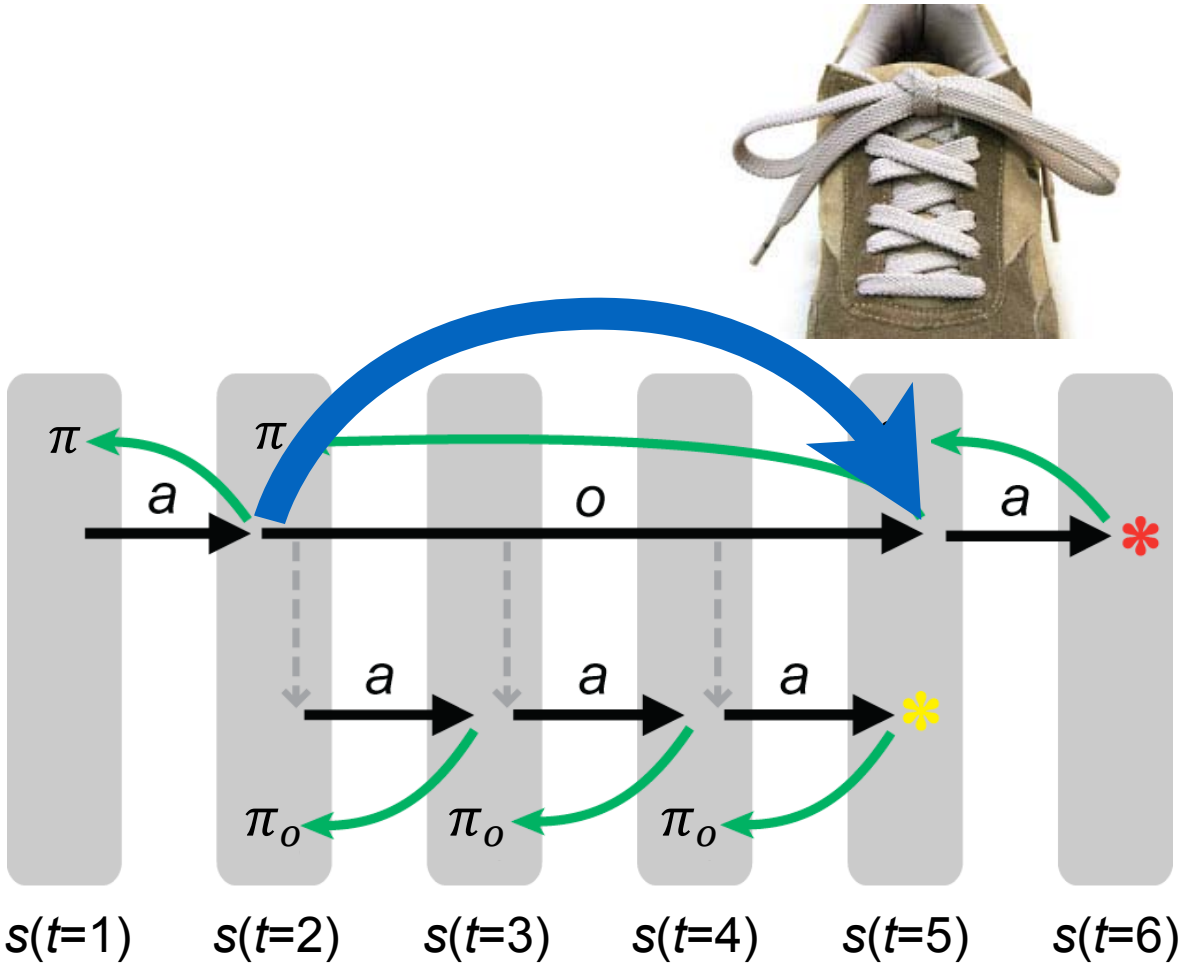
Botvinick, Niv & Barto, *Cognition*, 2009



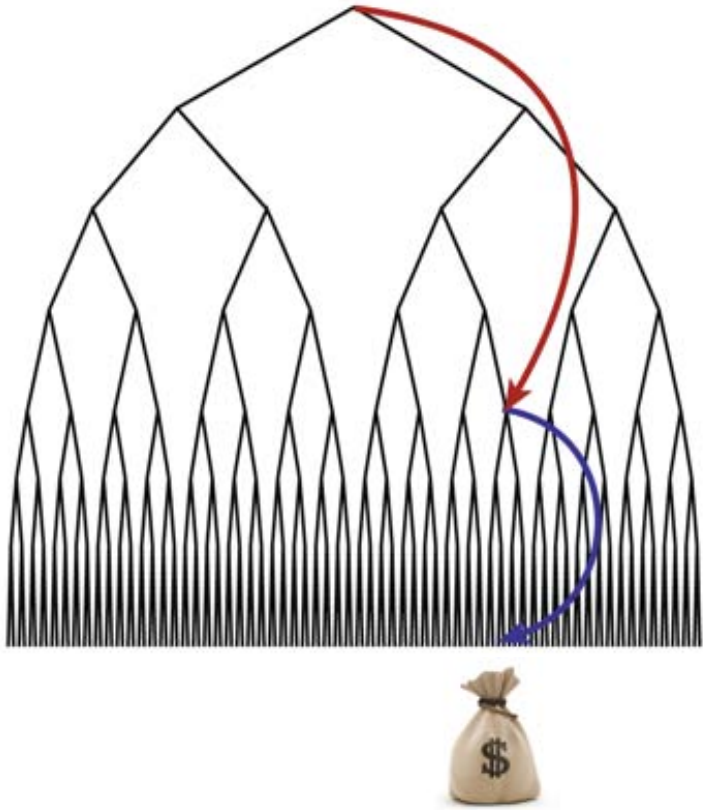
After Sutton, Precup & Singh, 1999

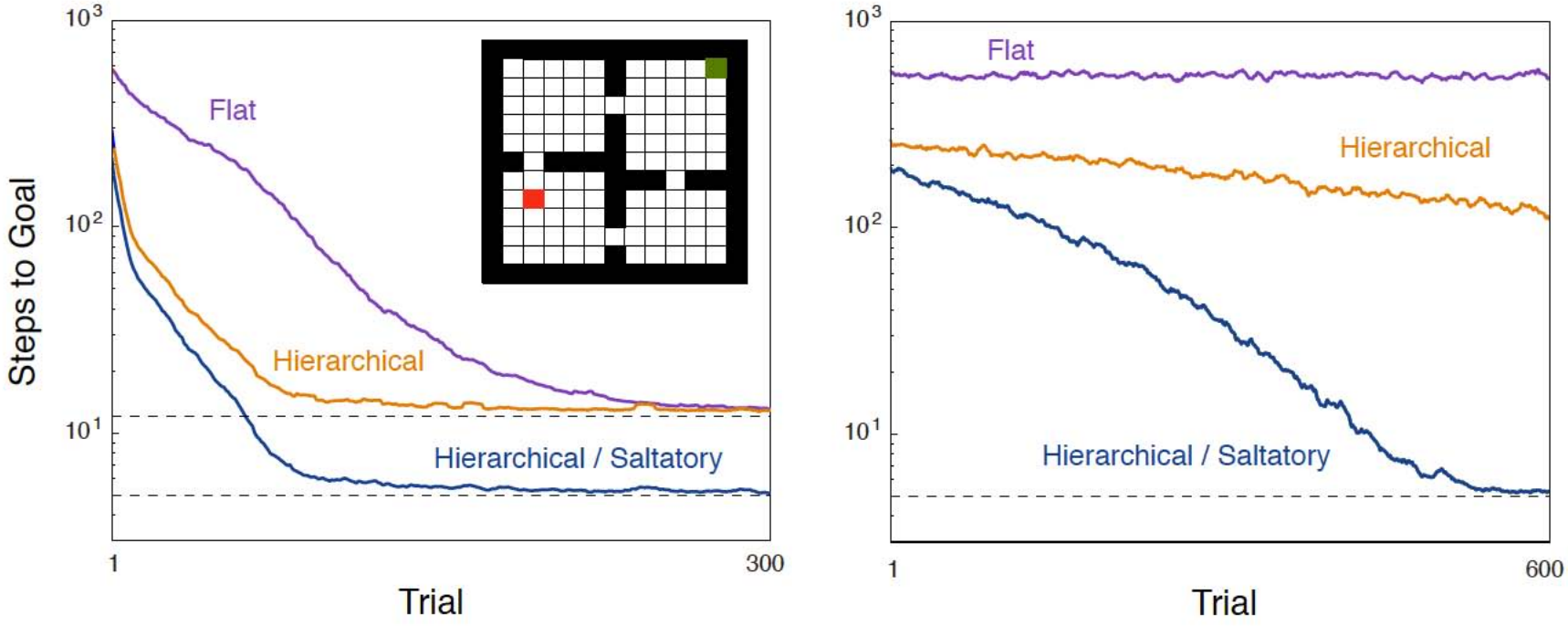


Botvinick, Niv & Barto, *Cognition*, 2009

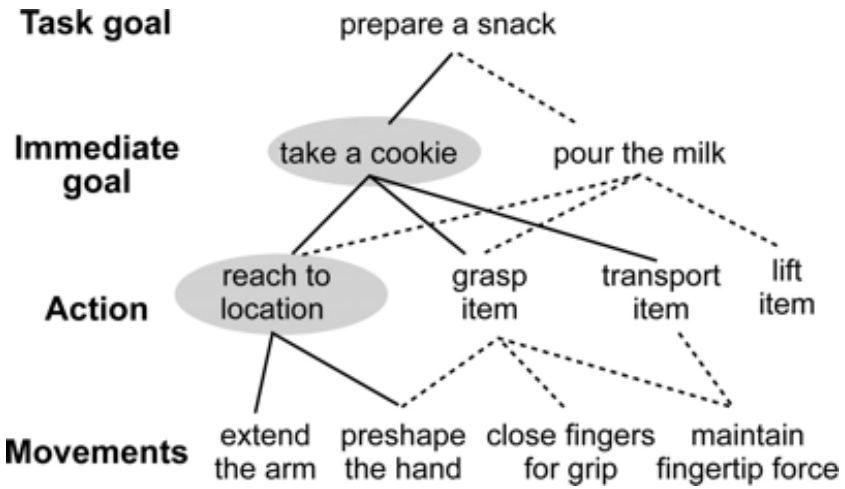


Botvinick, Niv & Barto, *Cognition*, 2009

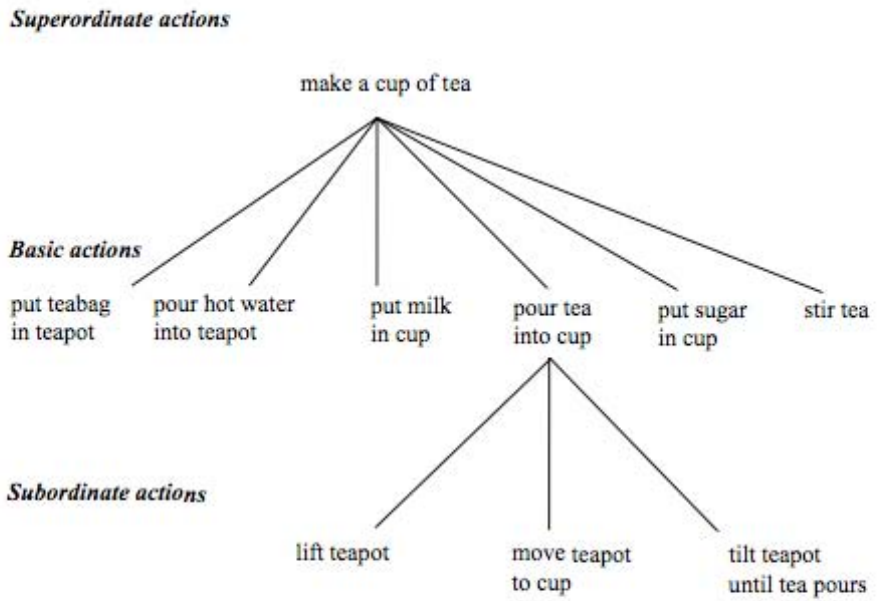




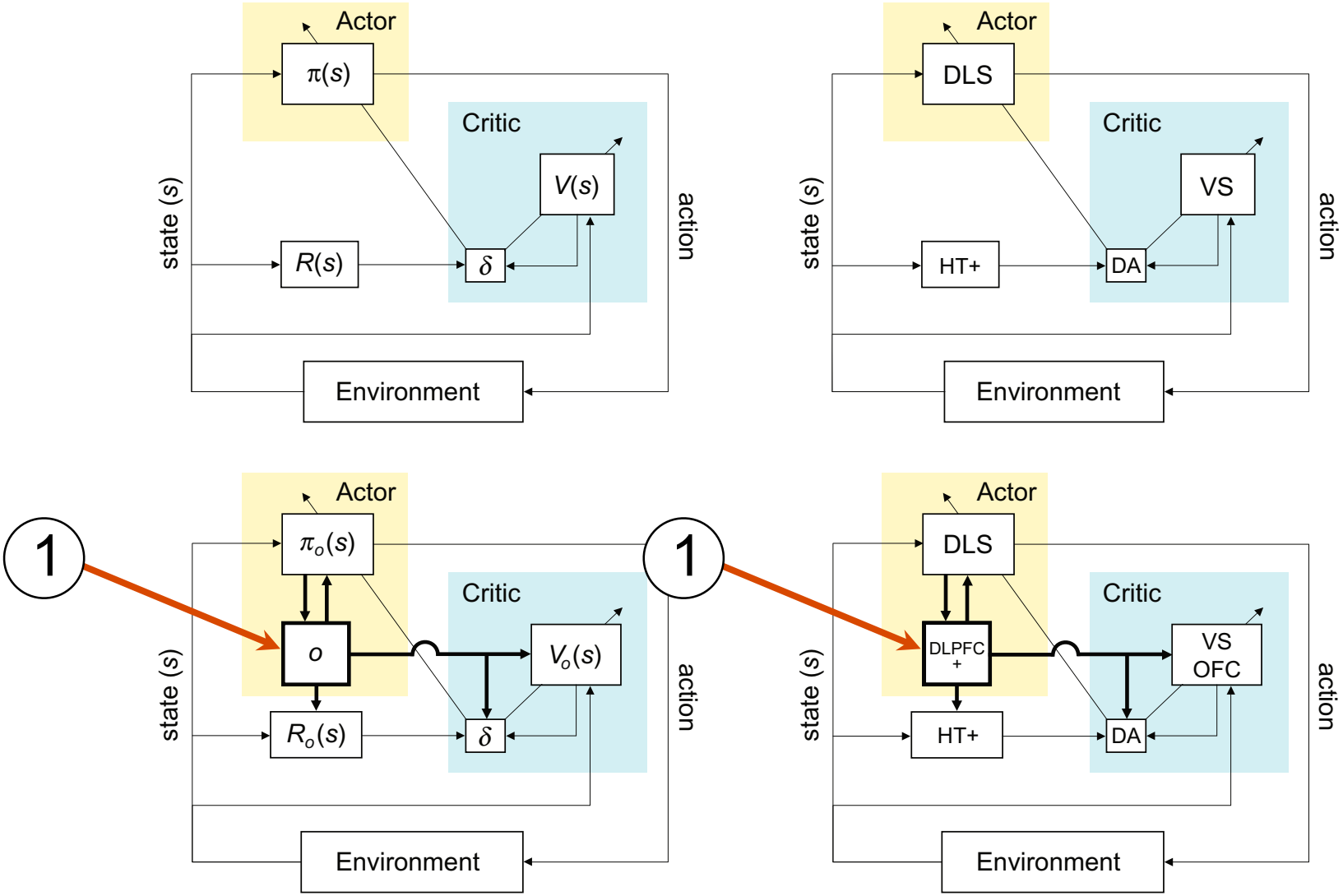
Botvinick & Weinstein, *Trans. Royal Society, 2014*



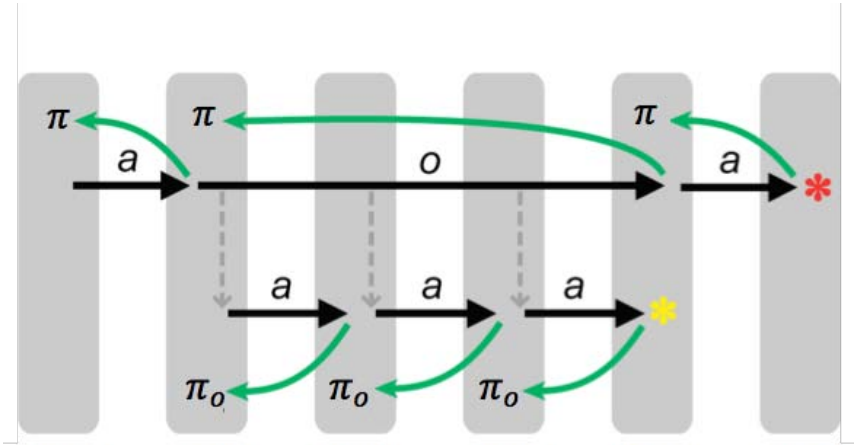
Hamilton & Grafton, *J Neurosci*, 2006

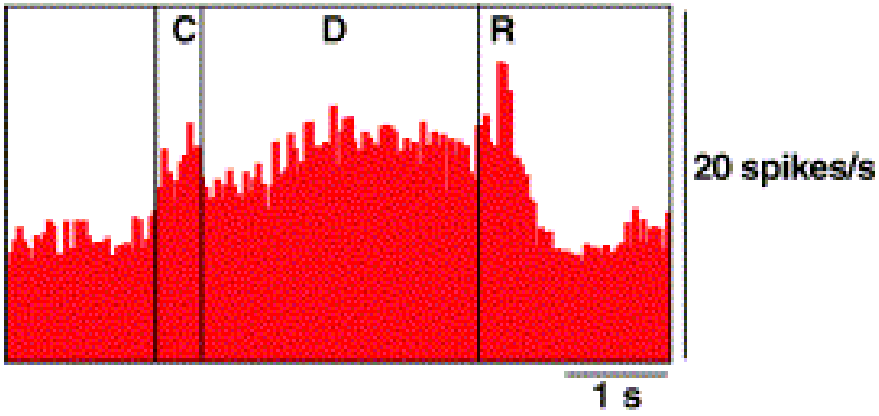


Humpheys & Forde, *Cog. Neuropsych.,* 2001

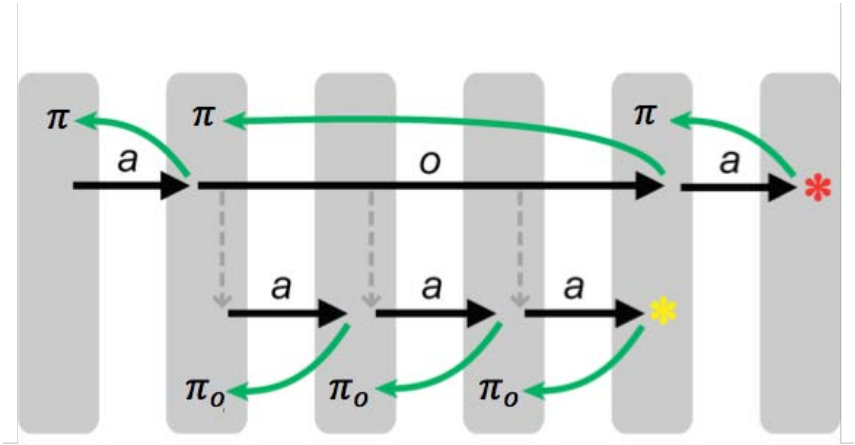


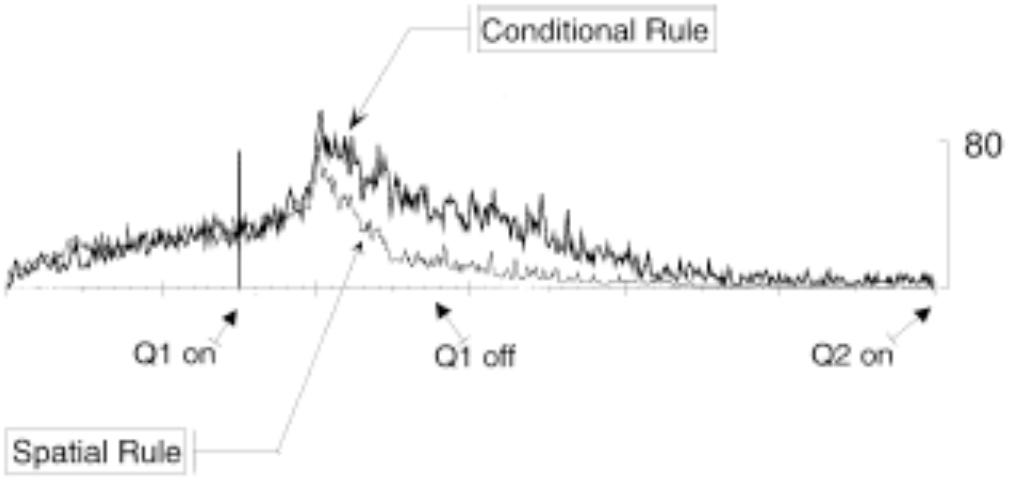
Botvinick, Niv & Barto, *Cognition*, 2009



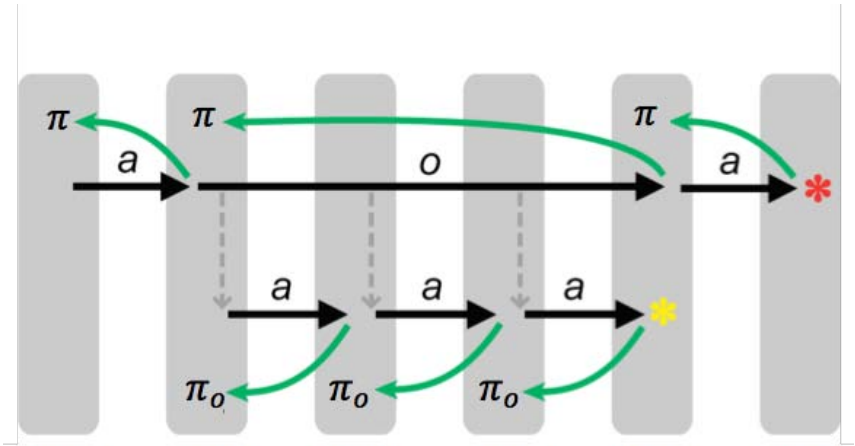


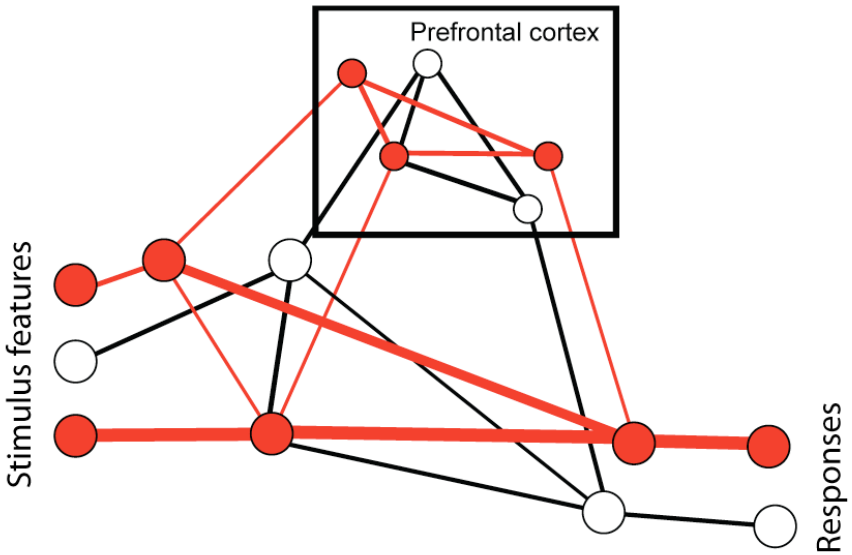
From Curtis & D'Esposito, *TICS*, 2003



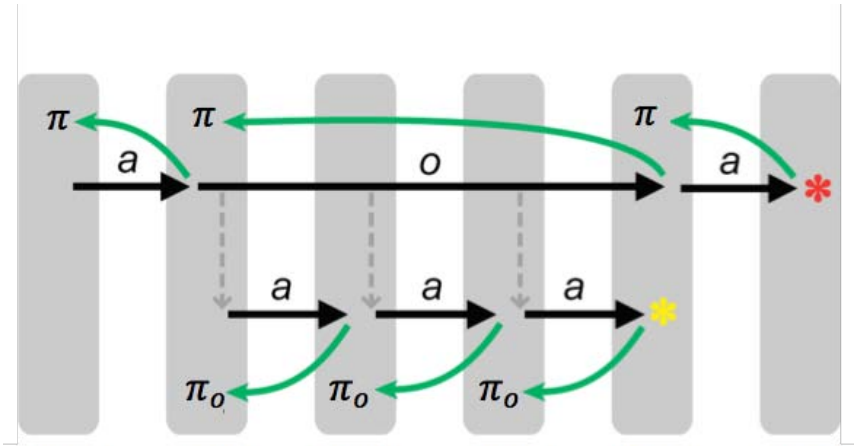


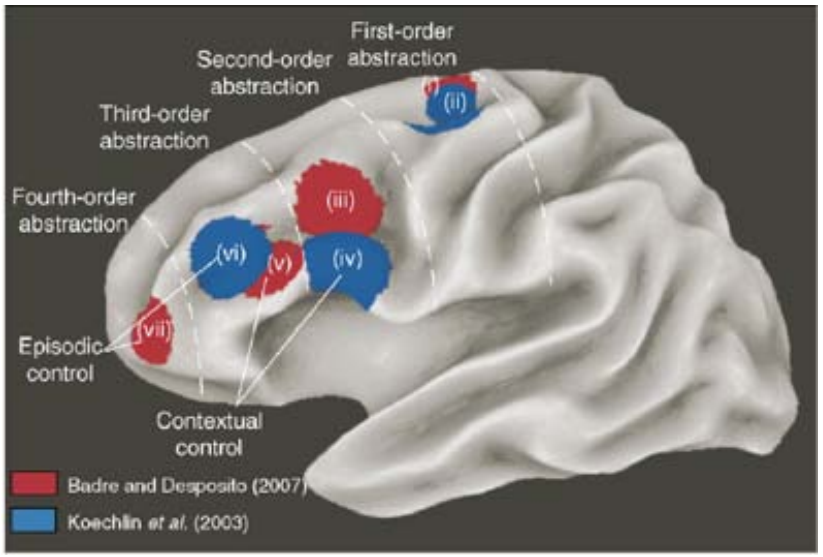
White & Wise, *Exp Br Res,* 1999



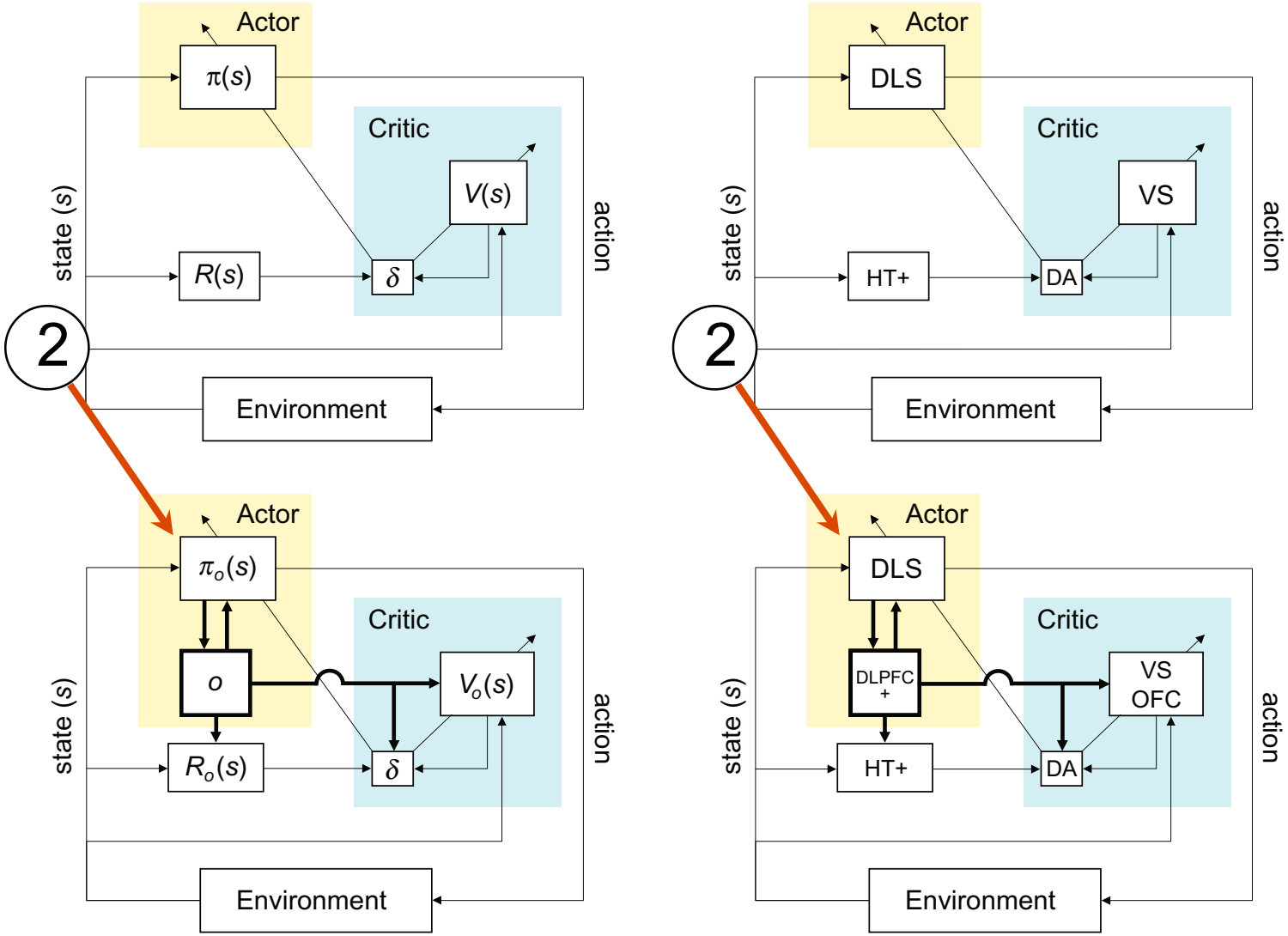


Miller & Cohen, *Ann. Rev. Neurosci*, 2001

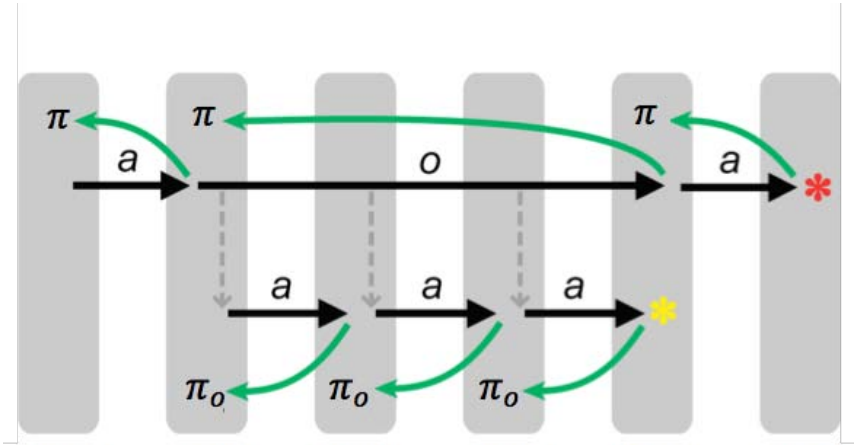


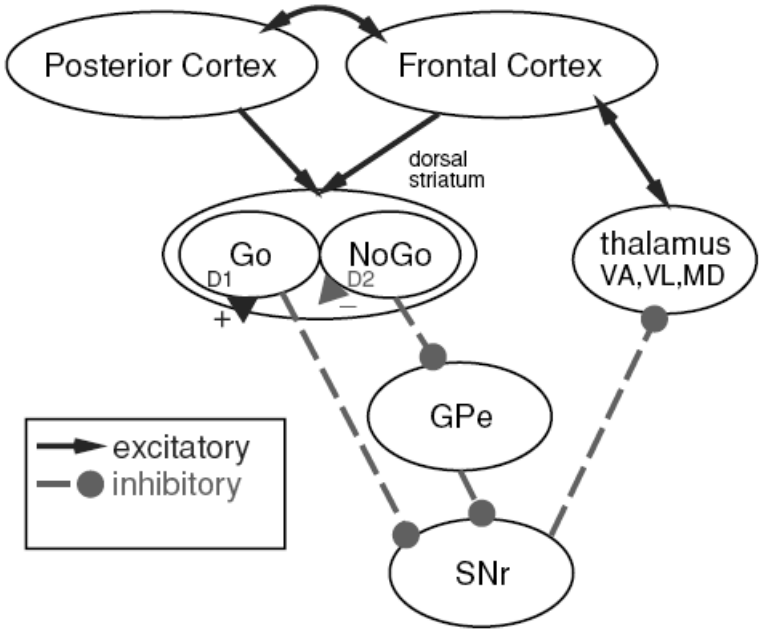


From Badre, *TICS*, 2008

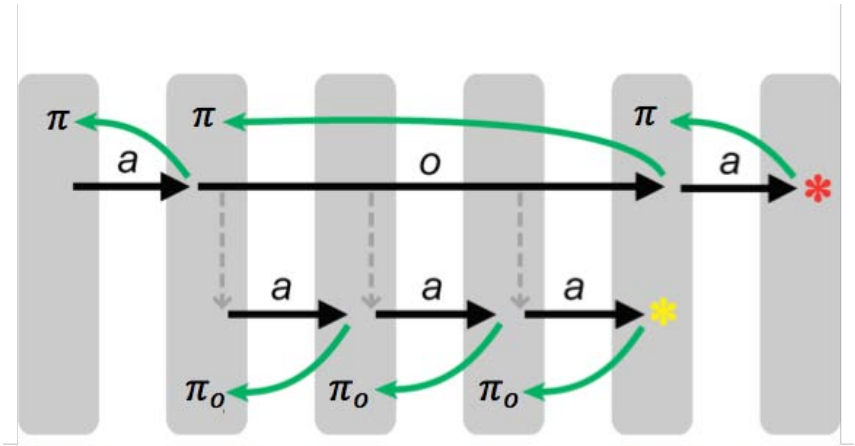


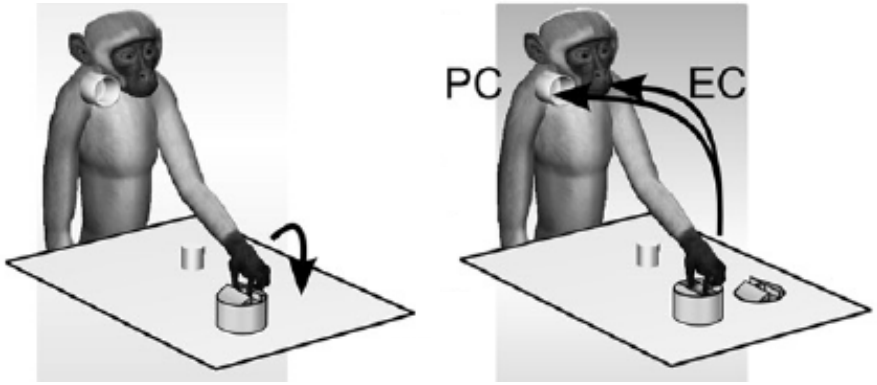
Botvinick, Niv & Barto, *Cognition*, 2009

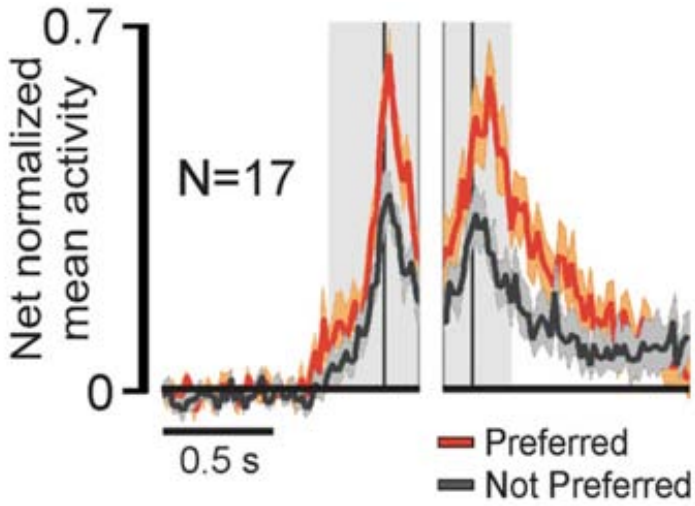




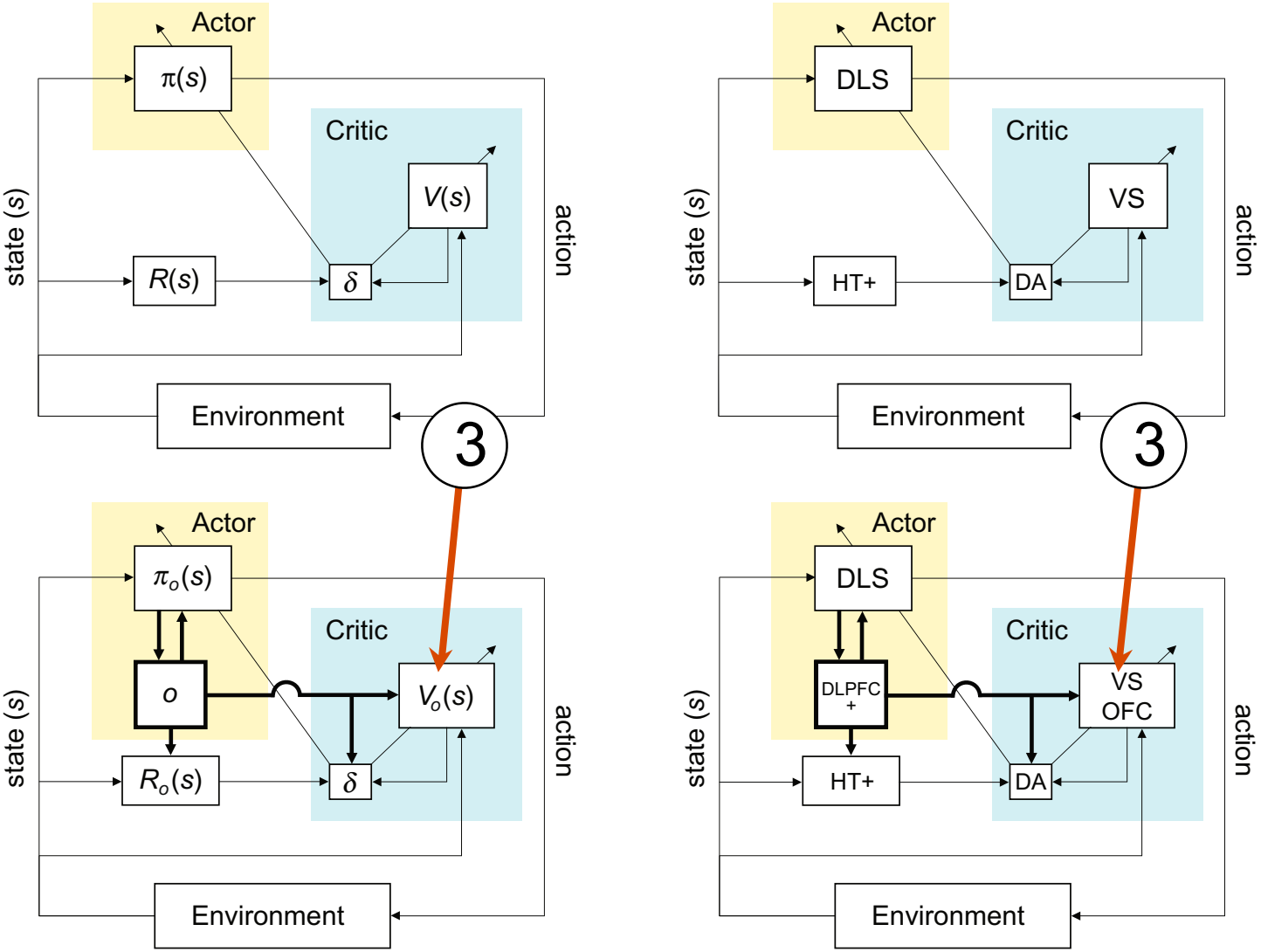
O'Reilly & Frank, *Neural Computation,* 2006



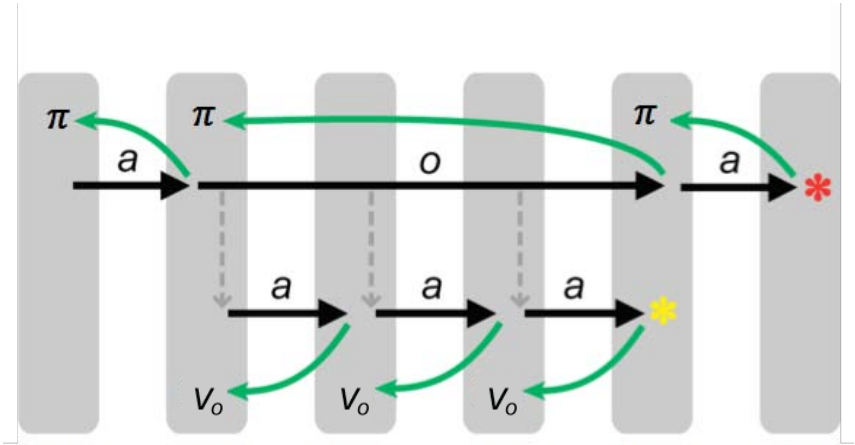


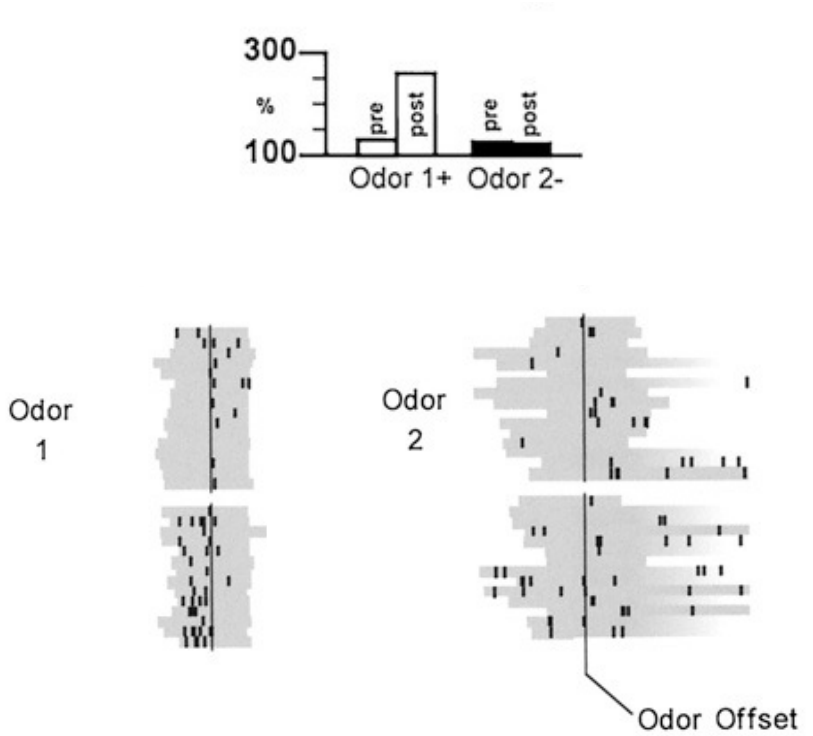


Bonini et al., *J. Neurosci.*, 2011

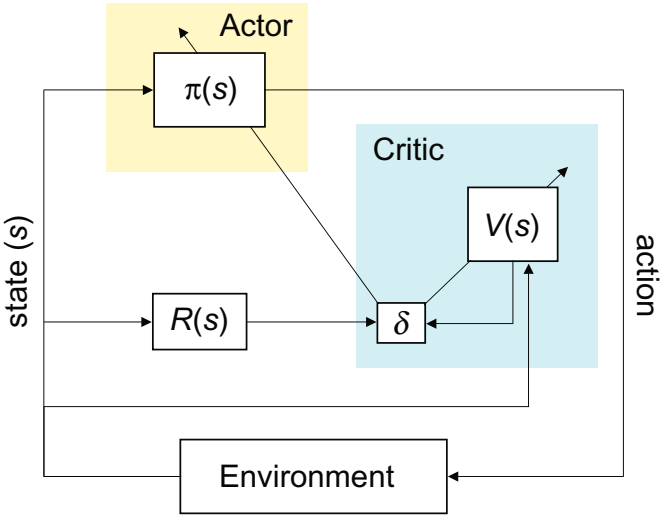


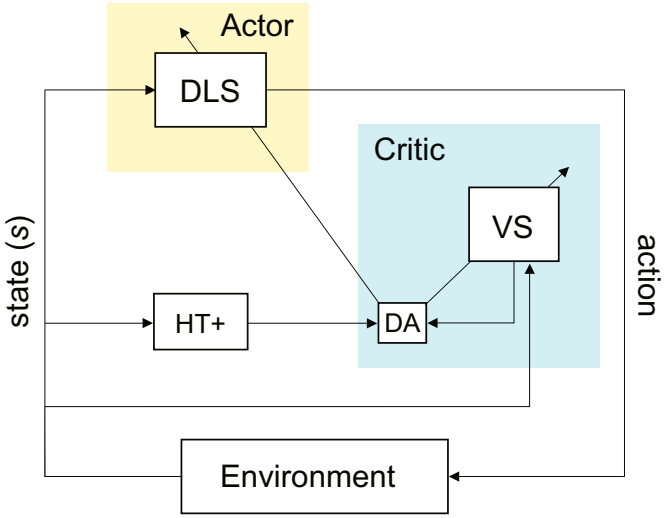
Botvinick, Niv & Barto, *Cognition*, 2009

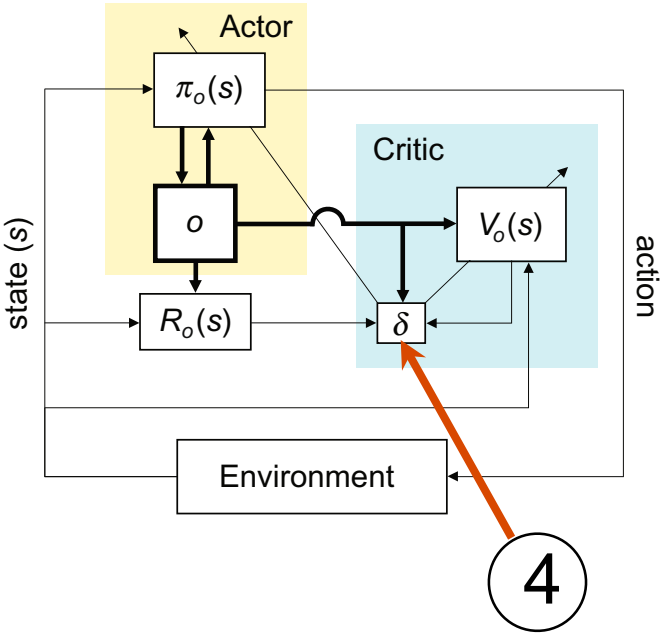


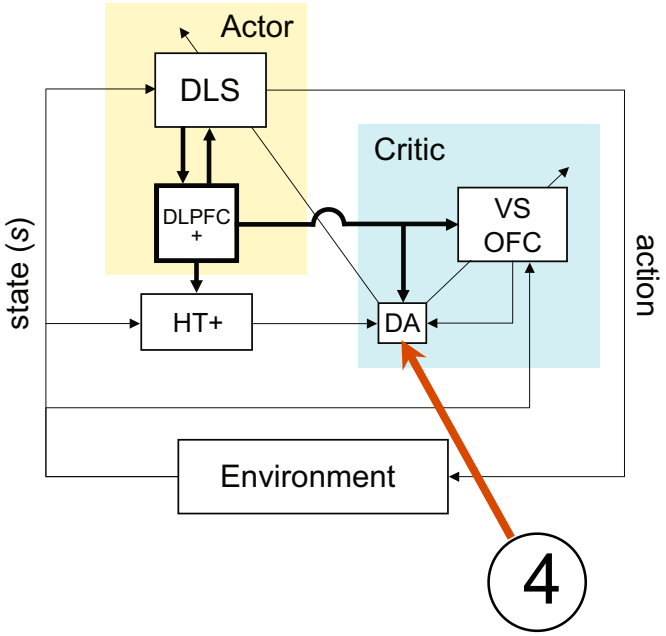


Schoenbaum, *et al. J Neurosci*. 1999

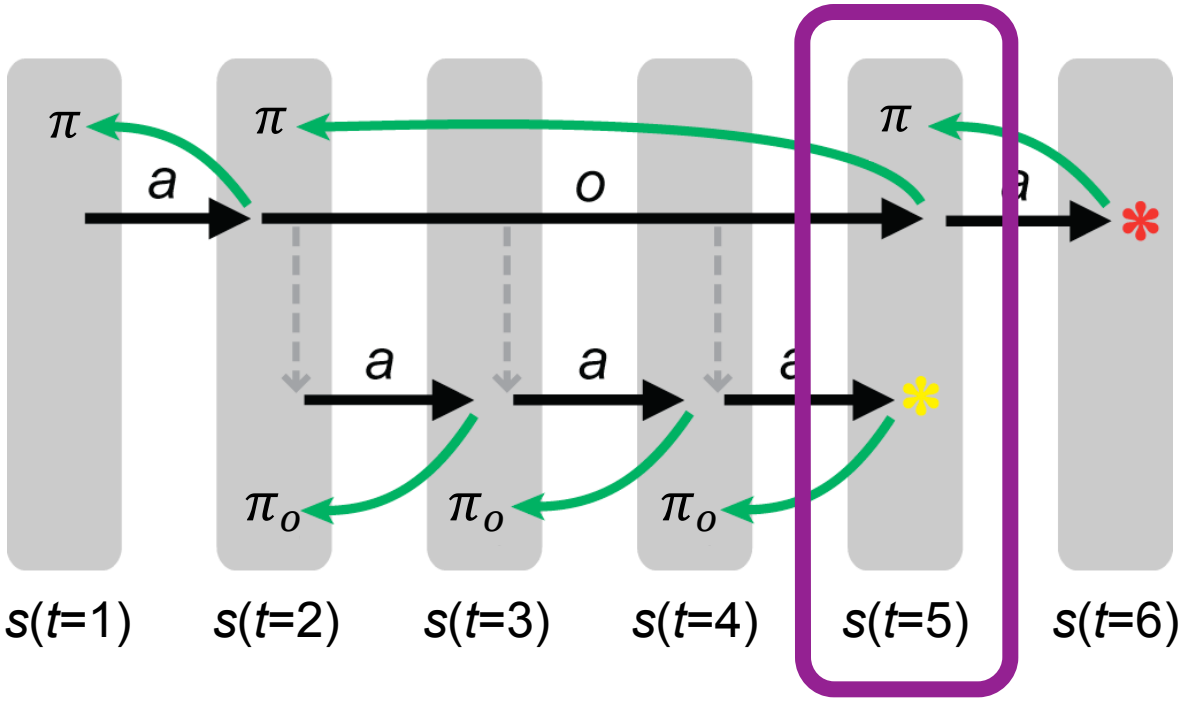






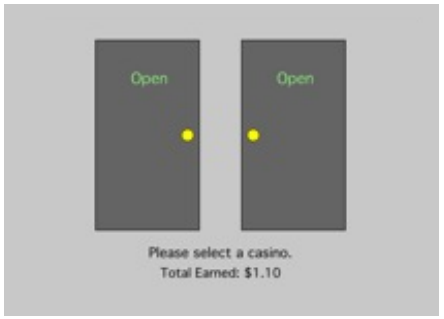


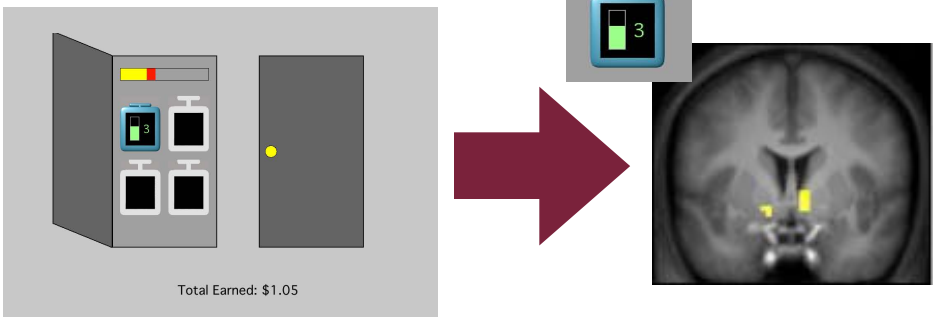
Botvinick, Niv & Barto, *Cognition*, 2009

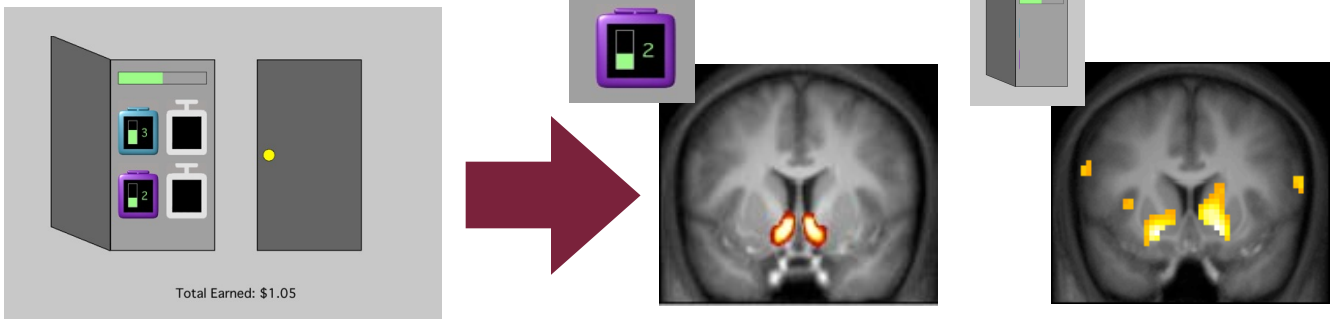

π
a
π
o
π
a
a
a
a
π_o
π_o
π_o
s(t=1)
s(t=2)
s(t=3)
s(t=4)
s(t=5)
s(t=6)

Carlos Diuk

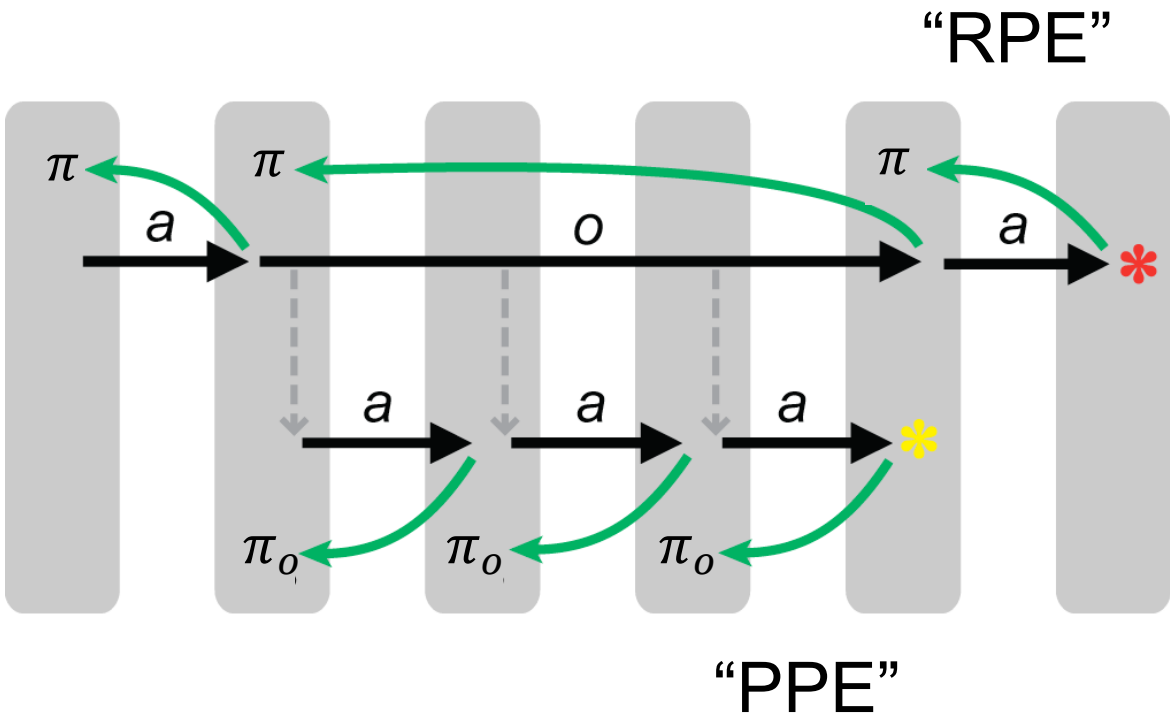
Carlos Diuk





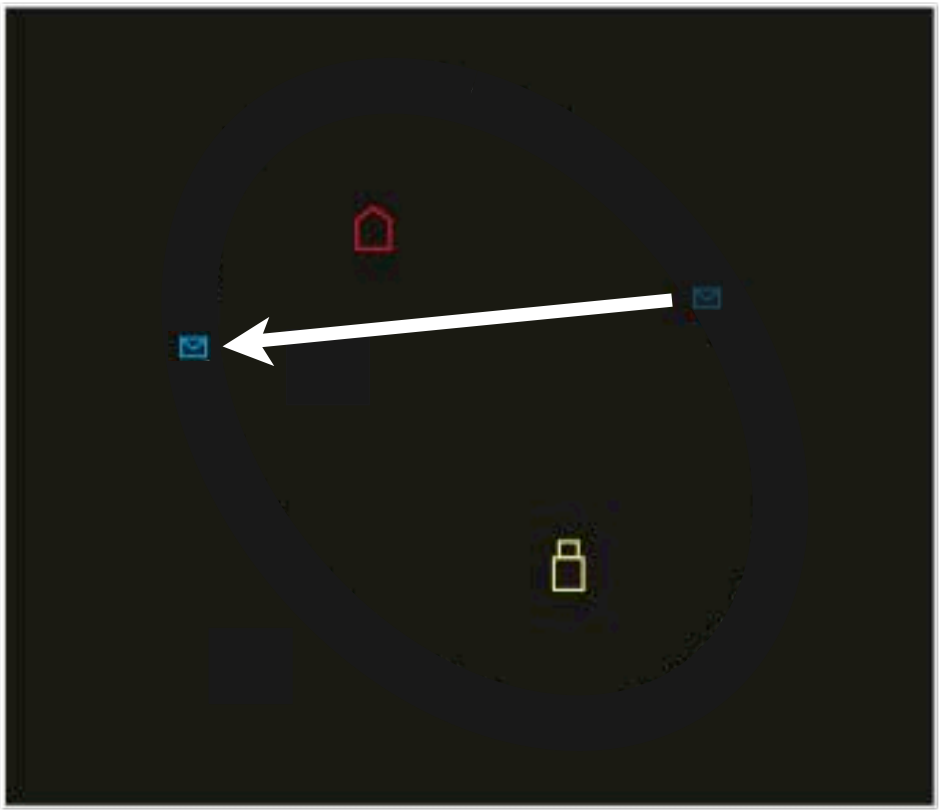


Diuk, *et al.*, *J Neurosci,* 2013


“RPE”
π
π
π
a
o
a
a
a
a
π_o
π_o
π_o
“PPE”

Jose Fernandes

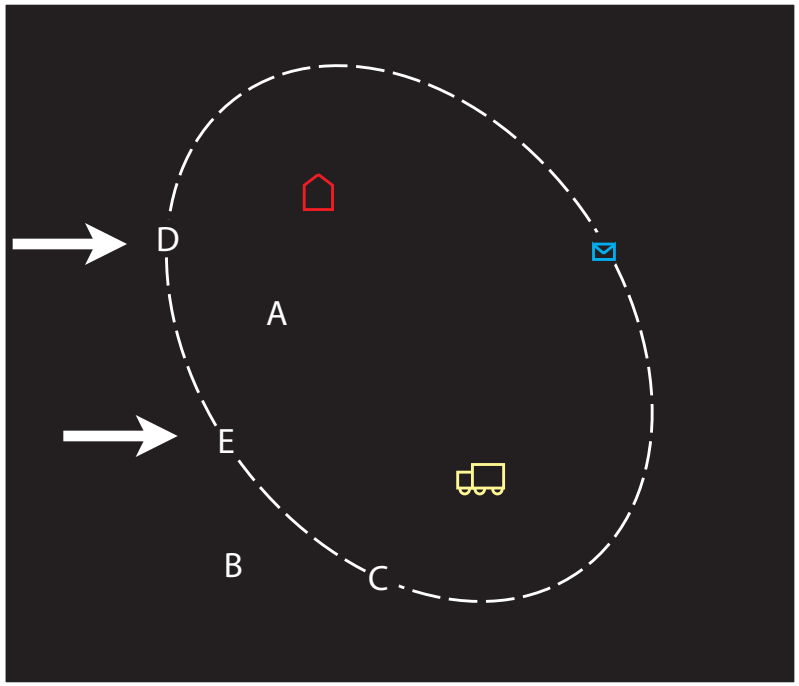
Alec Solway

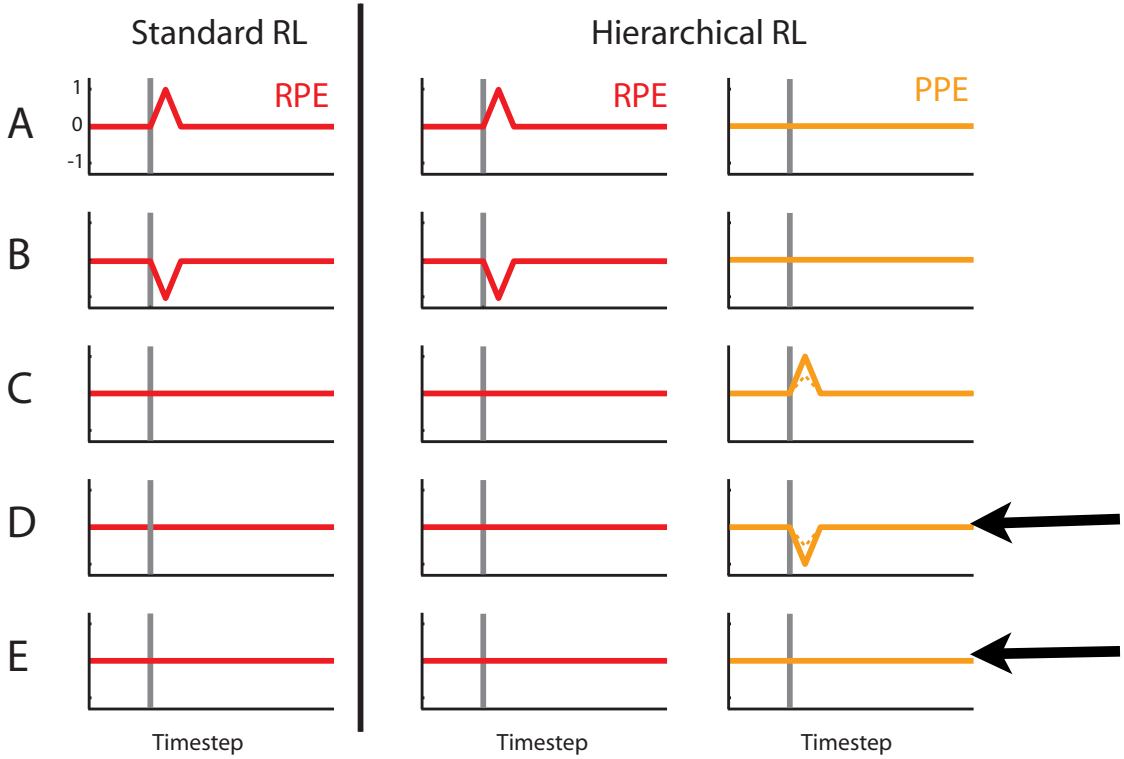


Ribas-Fernandes et al., *Neuron*, 2011


Jose Fernandes


Alec Solway

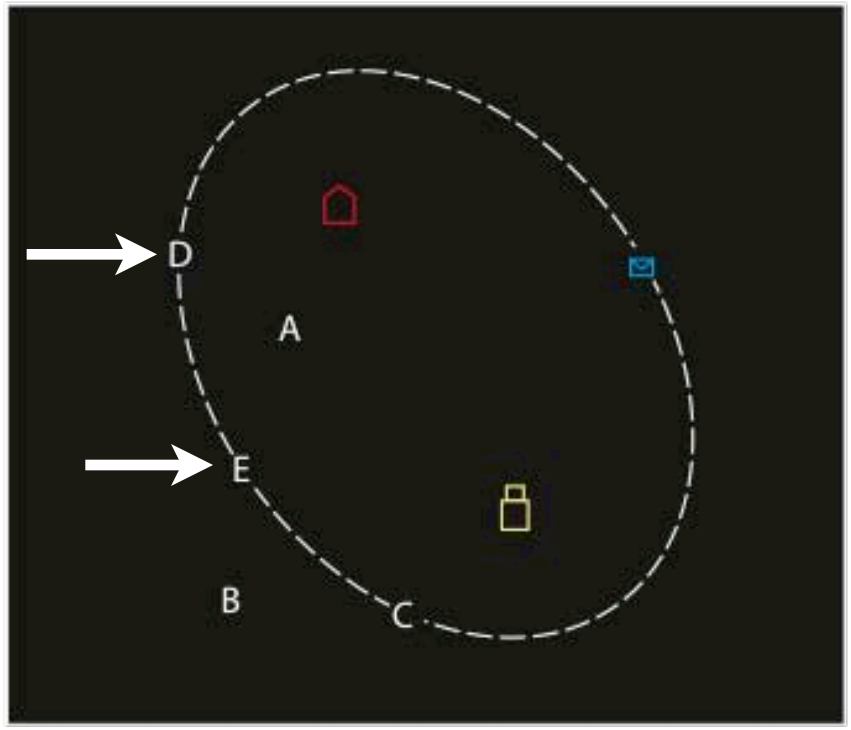


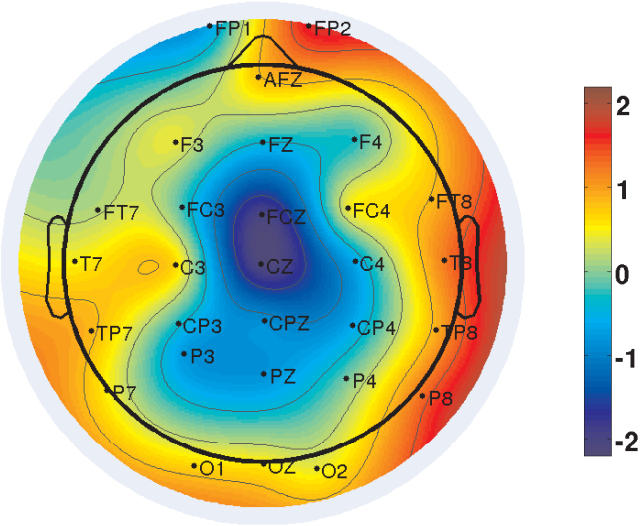


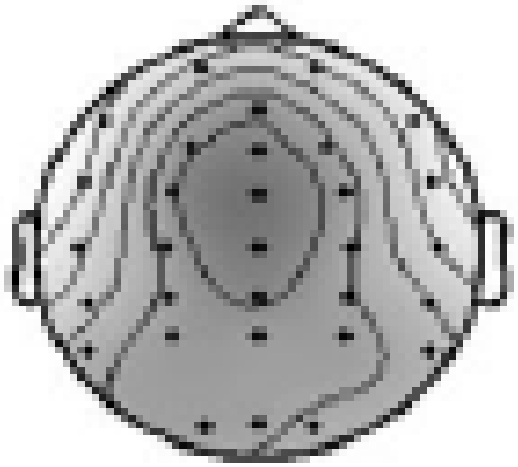
Ribas-Fernandes et al., *Neuron*, 2011

Jose Fernandes

Alec Solway





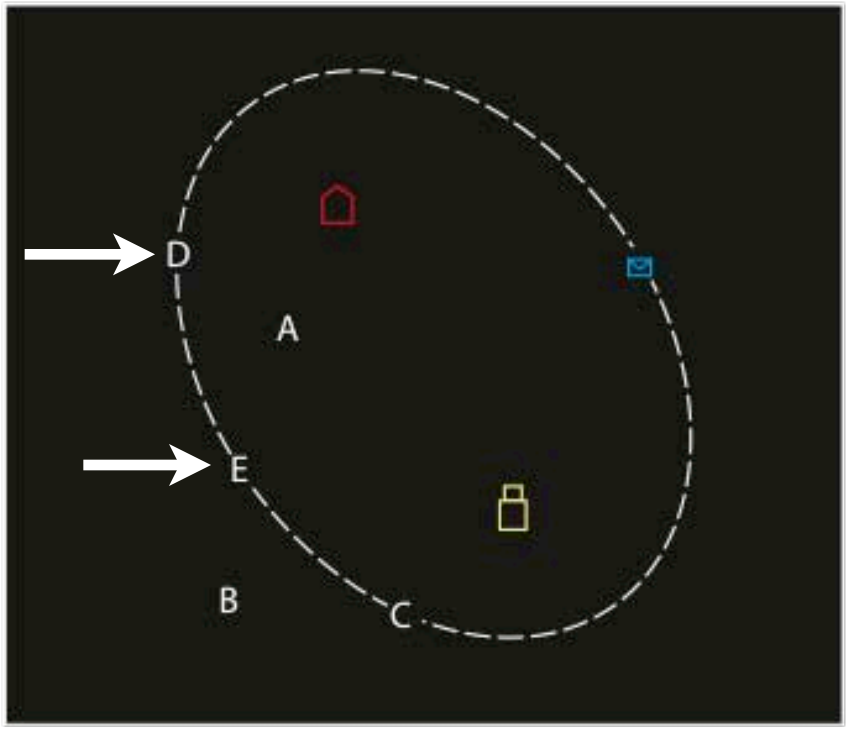


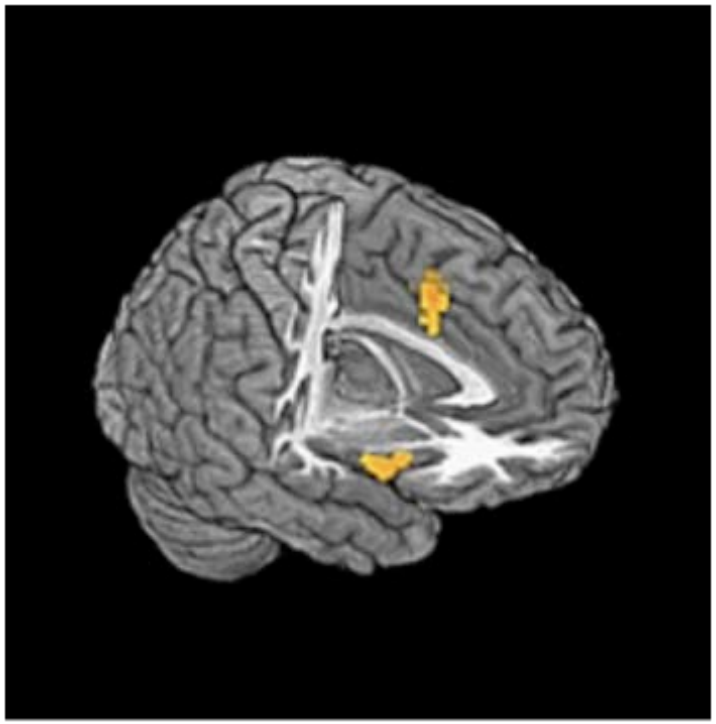
From Yeung, et al., 2005

Ribas-Fernandes et al., *Neuron*, 2011

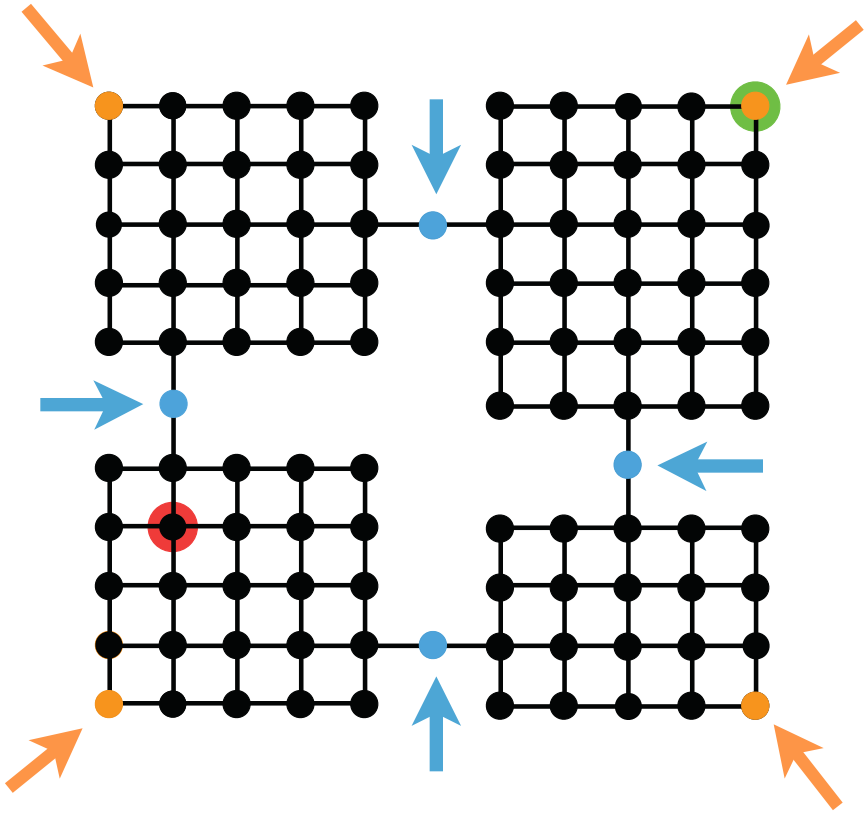

Jose Fernandes

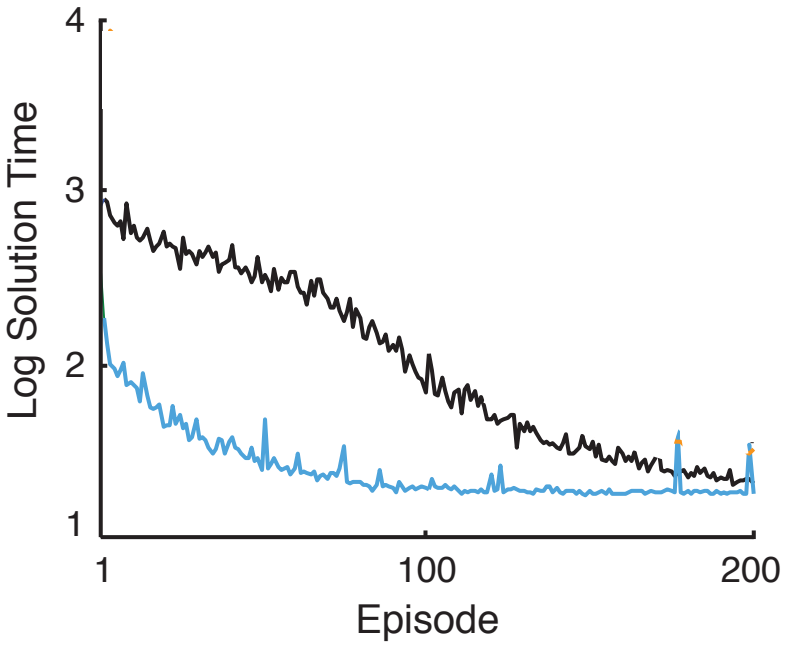

Alec Solway





Ribas-Fernandes et al., *Neuron*, 2011





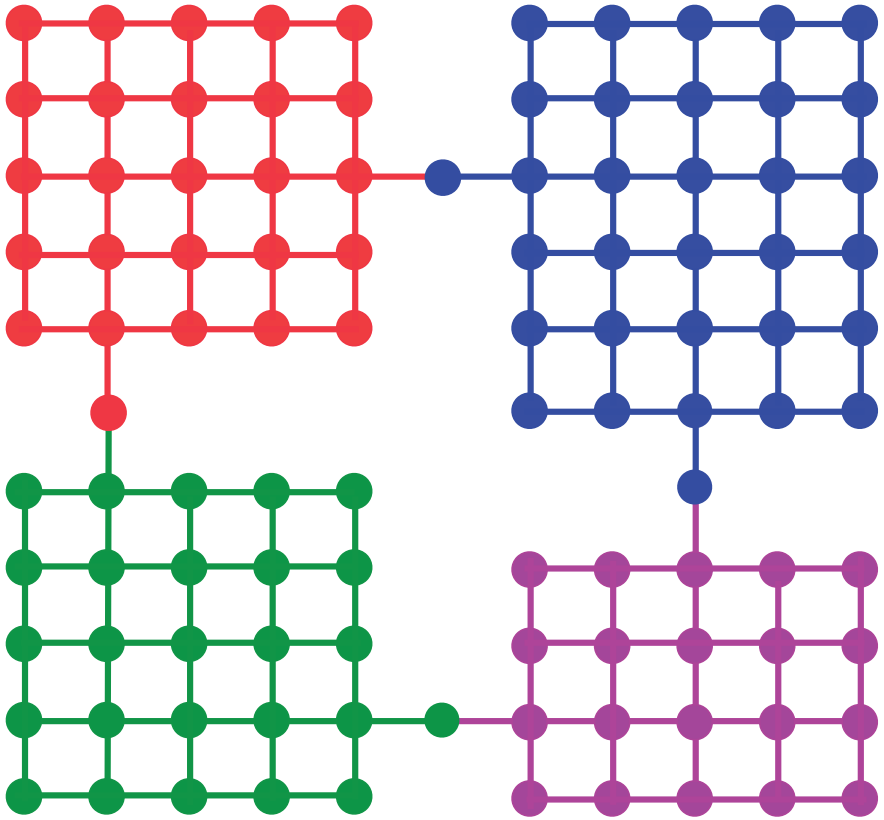
Botvinick, Niv & Barto, *Cognition*, 2009

The Burden of Abstraction

1. What should be learned?
2. Do people learn it?
3. How?


Alec Solway

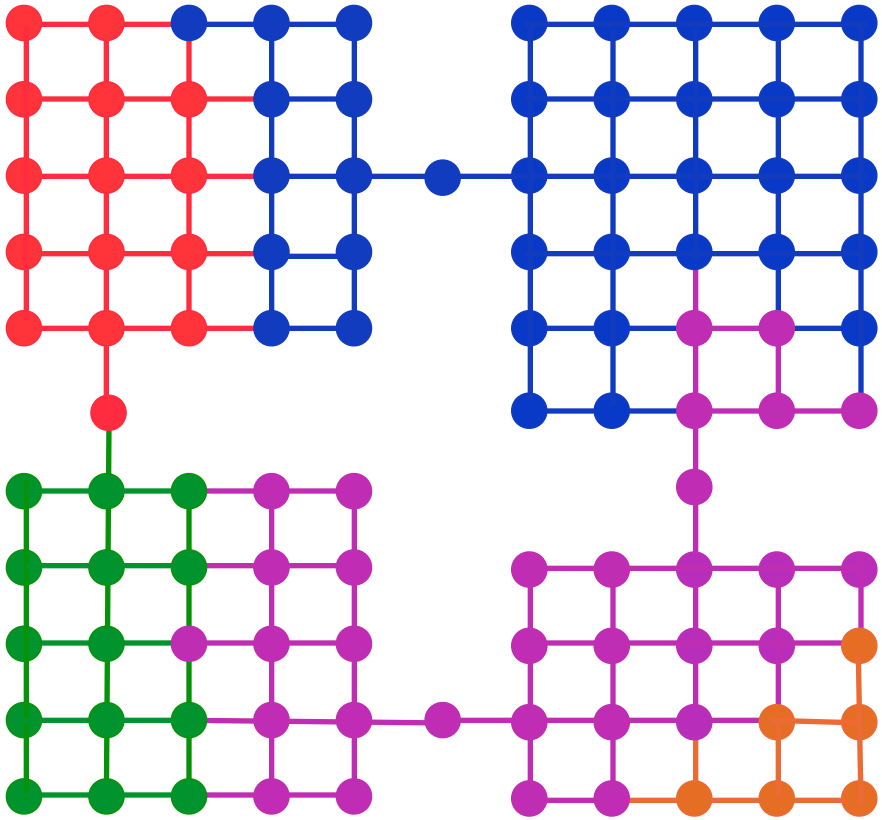

Carlos Diuk



Solway et al., *PLoS Comp. Biol., 2014*


Alec Solway

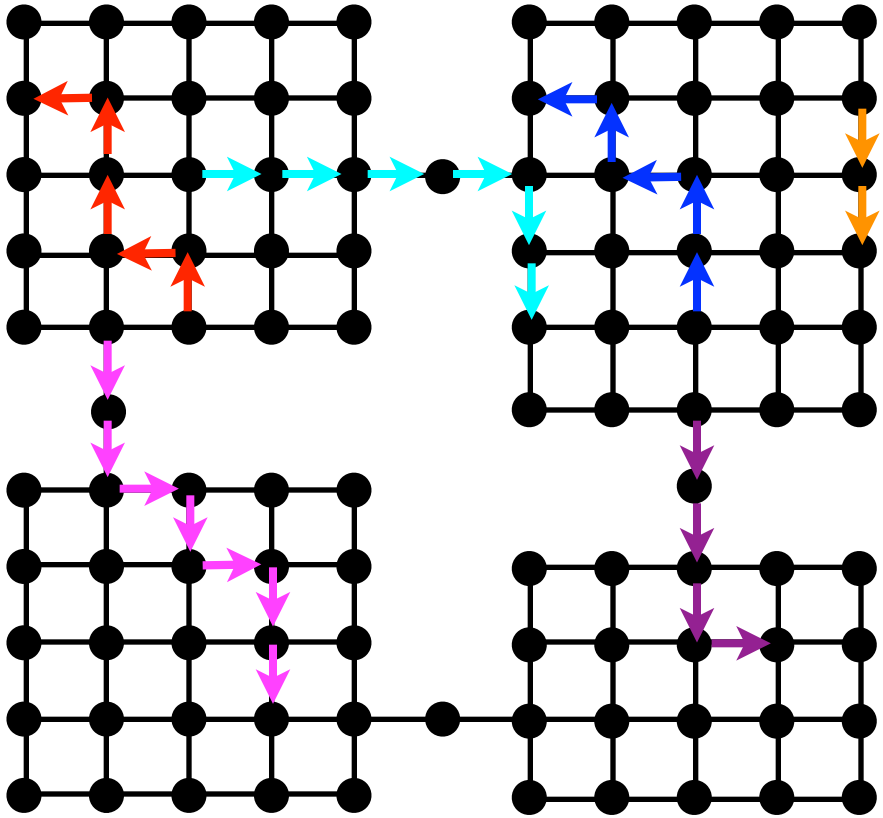

Carlos Diuk



Solway et al., *PLoS Comp. Biol., 2014*

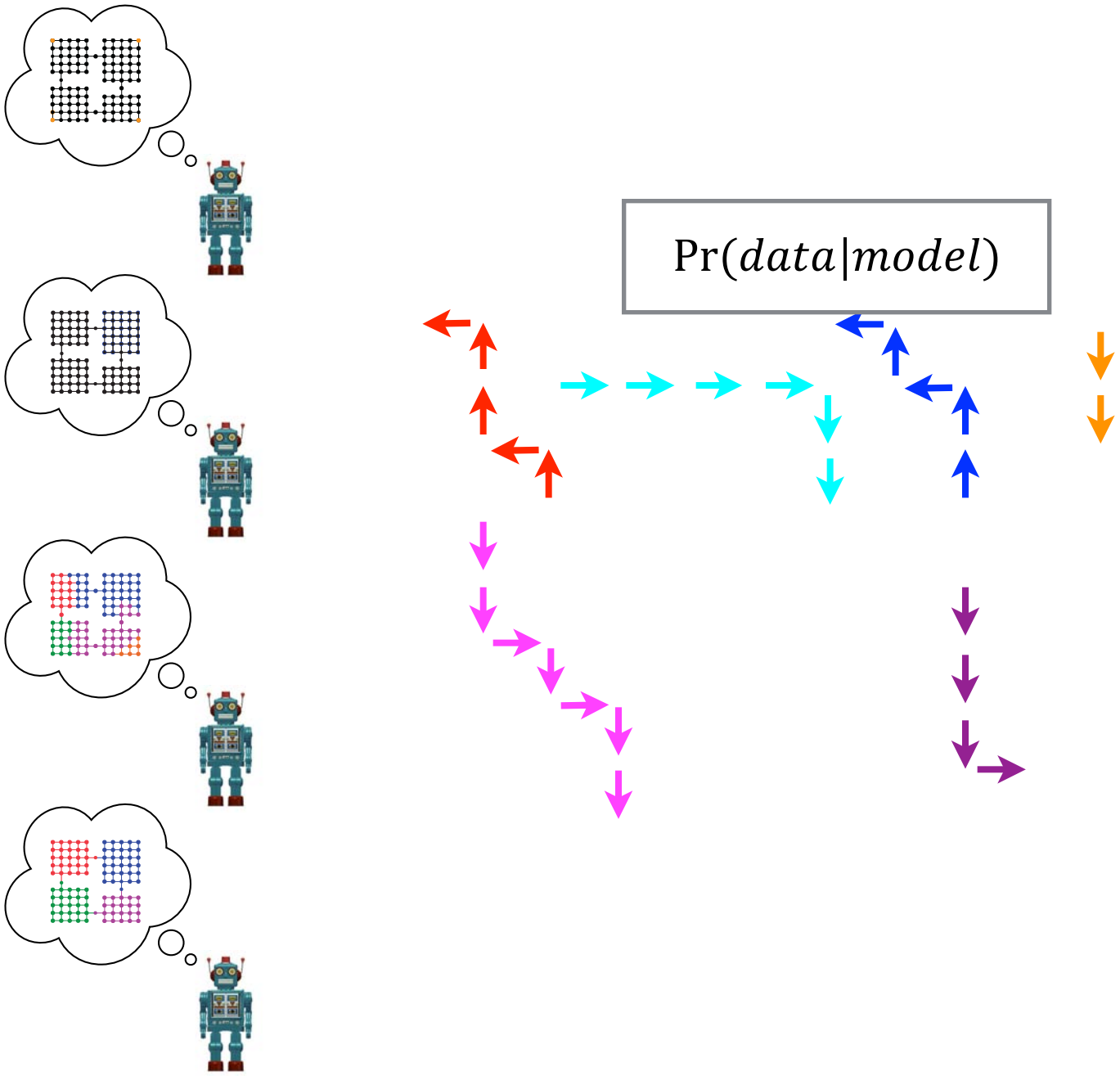

Alec Solway


Carlos Diuk

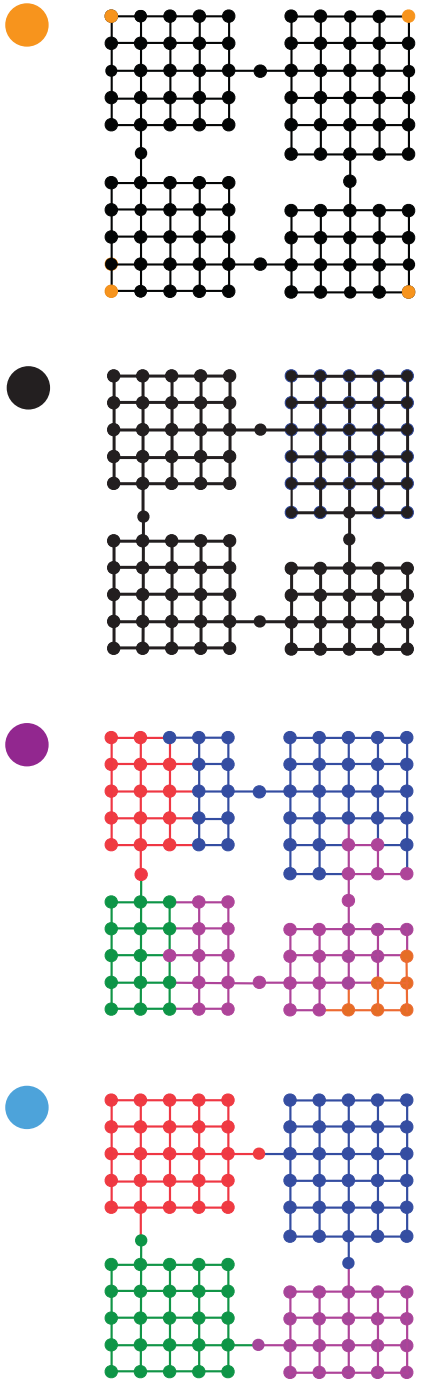


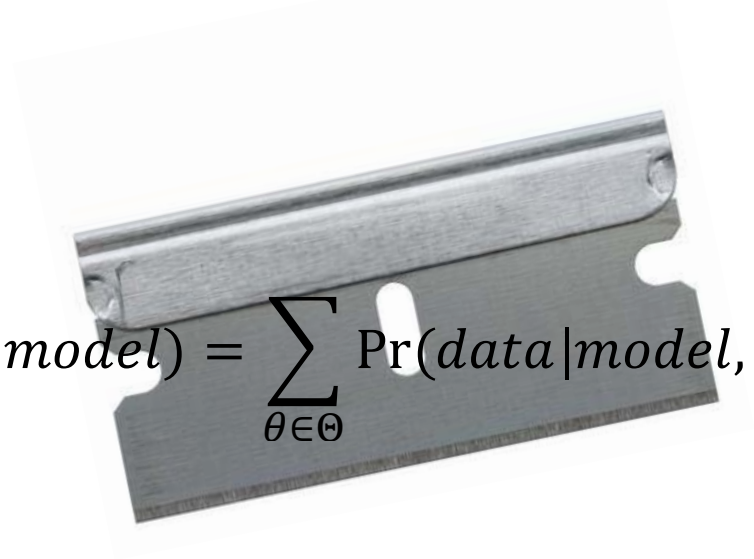
Solway et al., *PLoS Comp. Biol., 2014*

Alec Solway Carlos Diuk

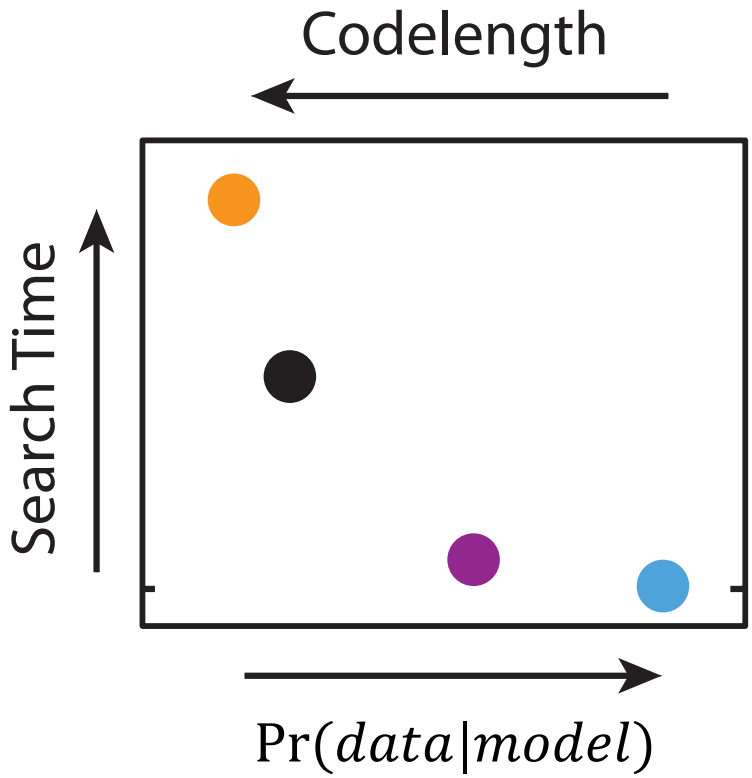


Solway et al., *PLoS Comp. Biol., 2014*

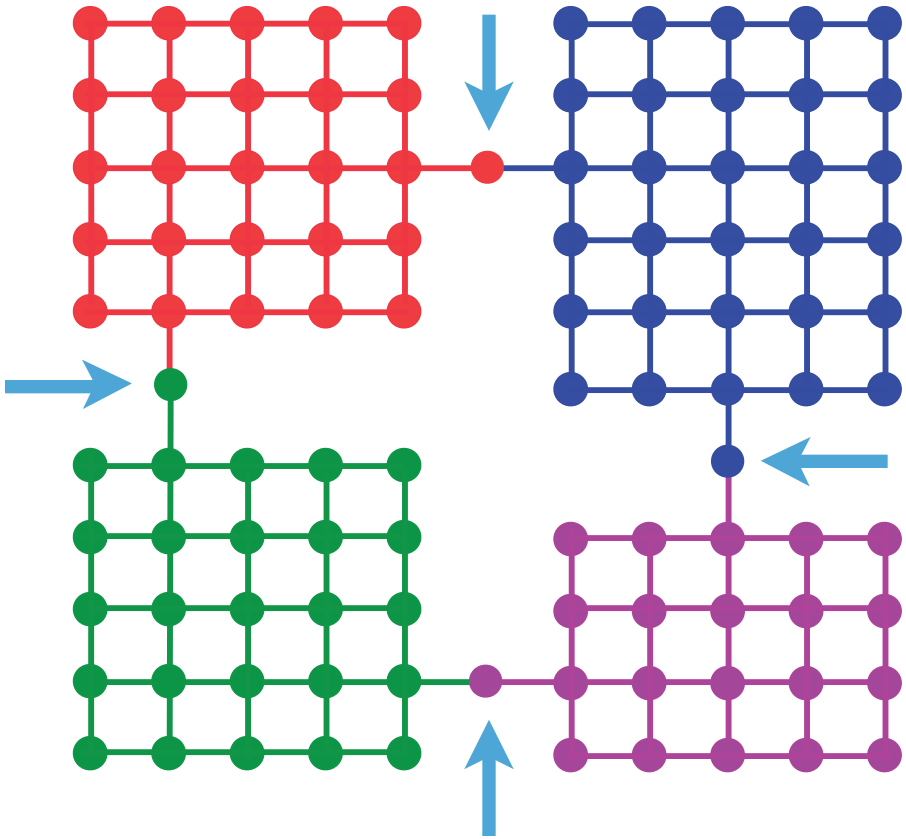




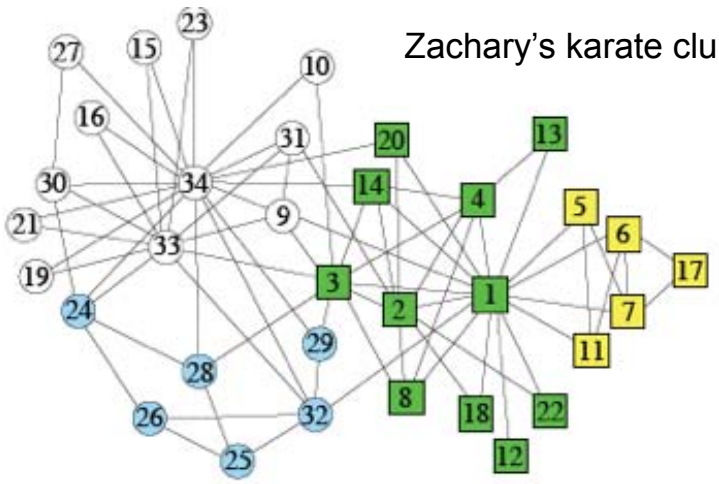
$$\Pr(data|model) = \sum_{\theta \in \Theta} \Pr(data|model, \theta) \Pr(\theta|model)$$

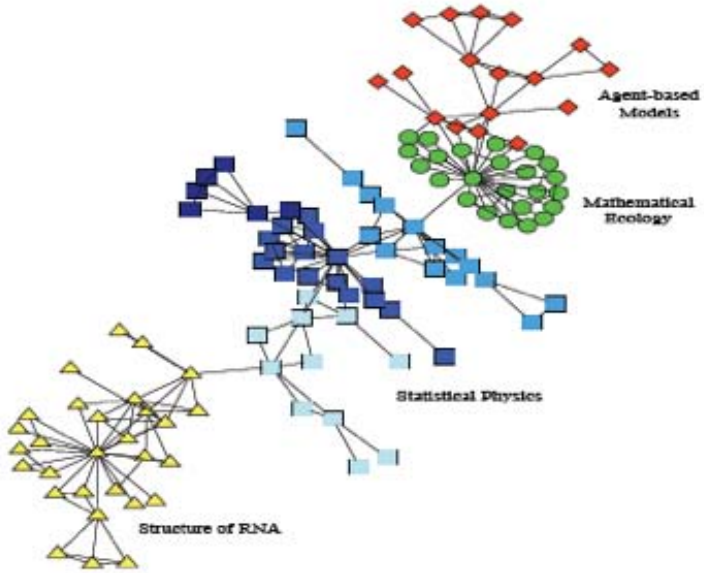


Solway et al., *PLoS Comp. Biol., 2014*

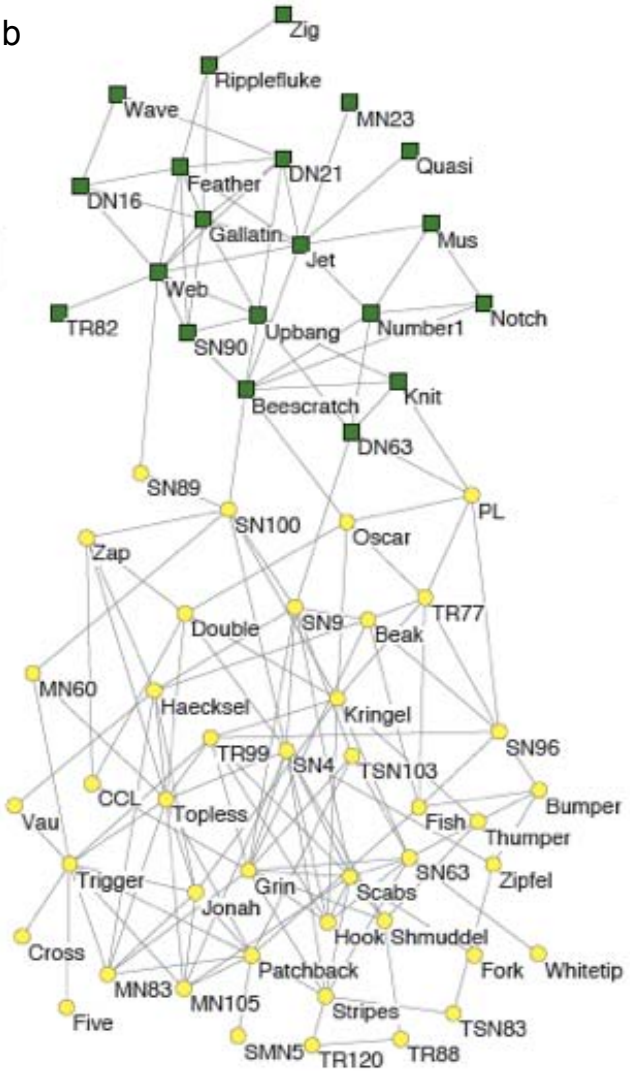


Solway et al., *PLoS Comp. Biol., 2014*



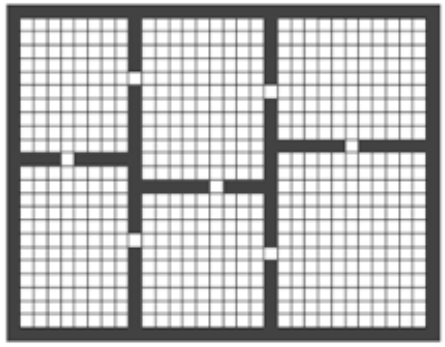


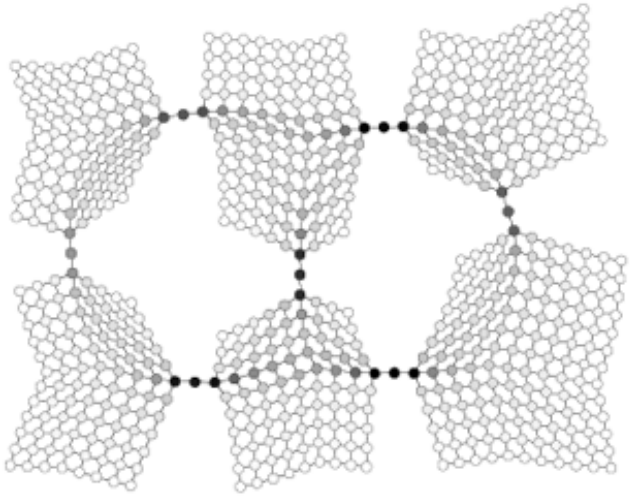
Santa Fe Institute collaborations

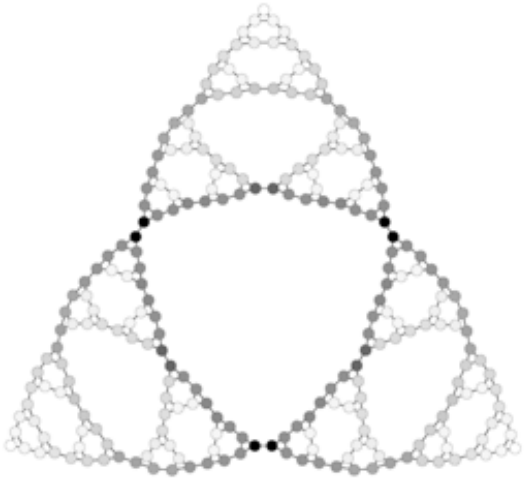


Lusseau's bottlenose dolphins

Fortunato, *Physics Reports*, 2010







Simsek, Wolfe & Barto, 2005

1. What should be learned?
2. Do people learn it?
3. How?

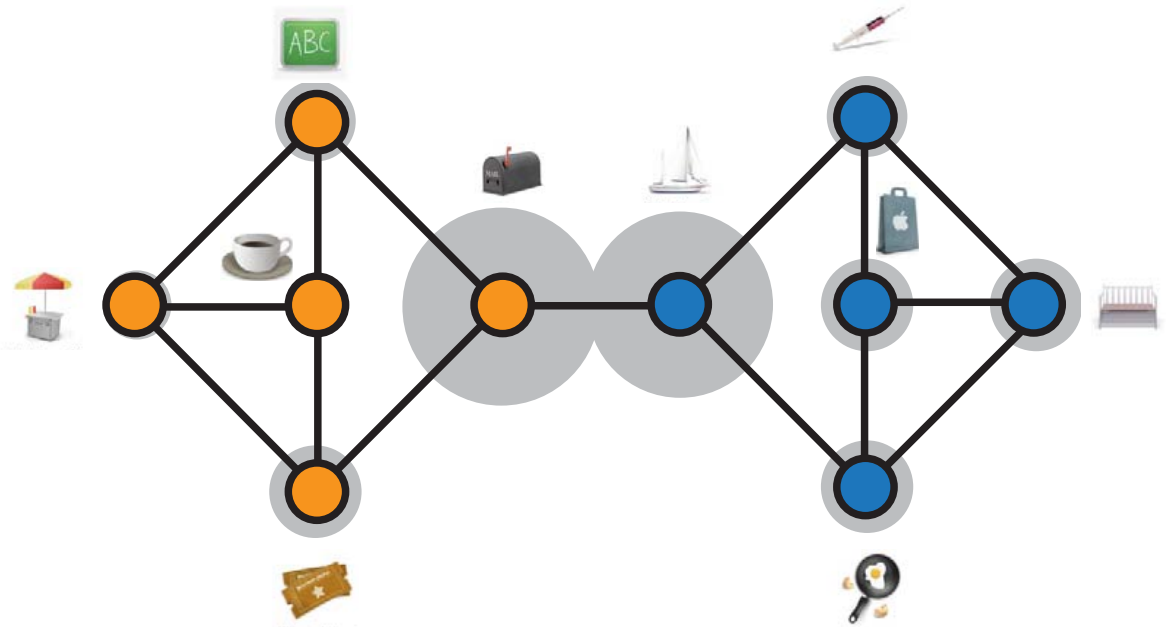

Carlos Diuk


DebbieYee

Solway et al., *PLoS Comp. Biol., 2014*


Carlos Diuk


DebbieYee



Solway et al., *PLoS Comp. Biol., 2014*

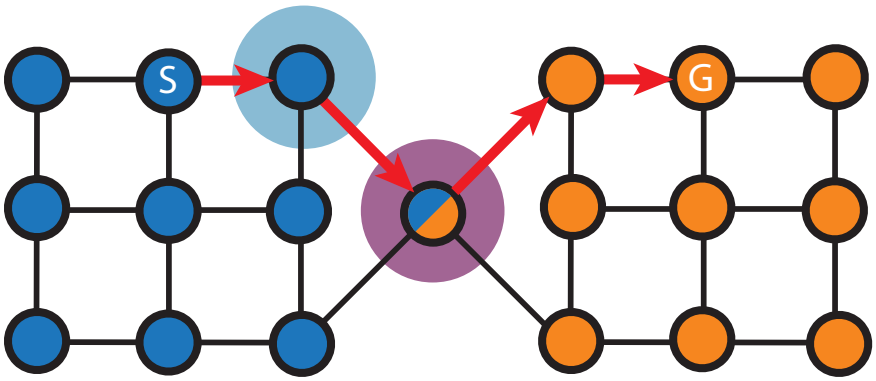

Carlos Diuk

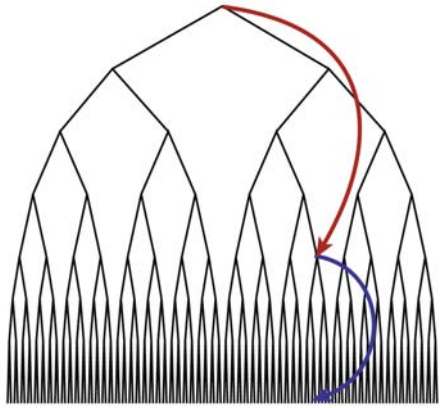

DebbieYee

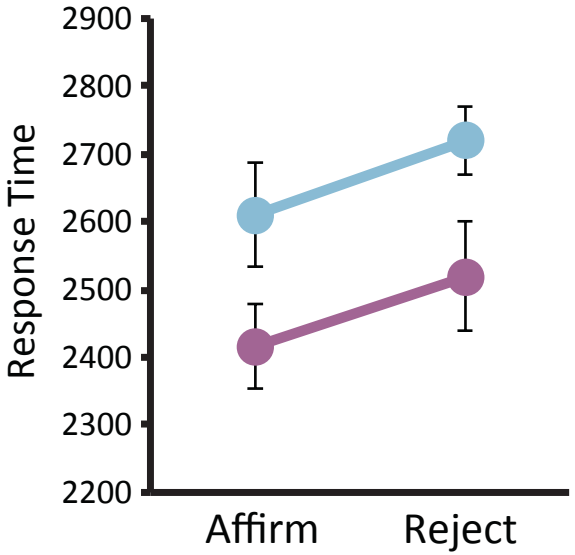
Solway et al., *PLoS Comp. Biol., 2014*


Carlos Diuk


DebbieYee

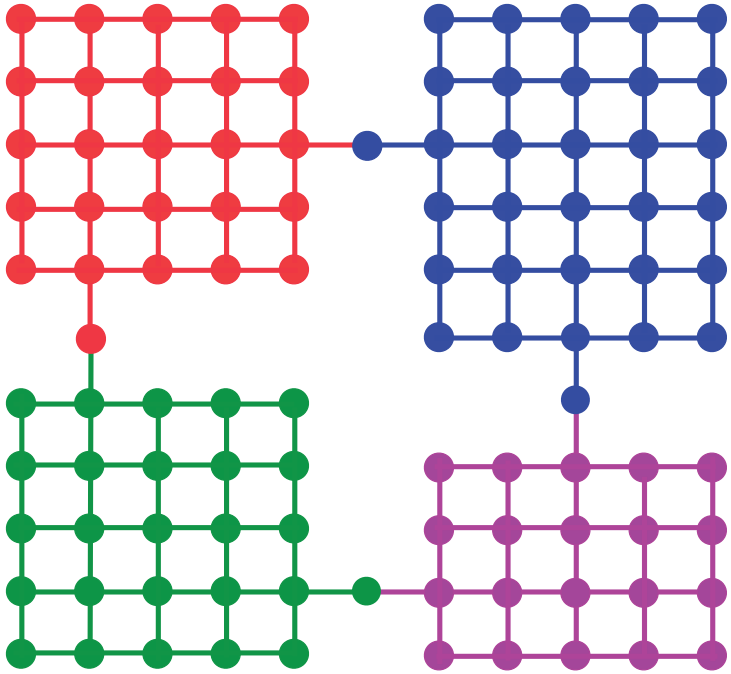






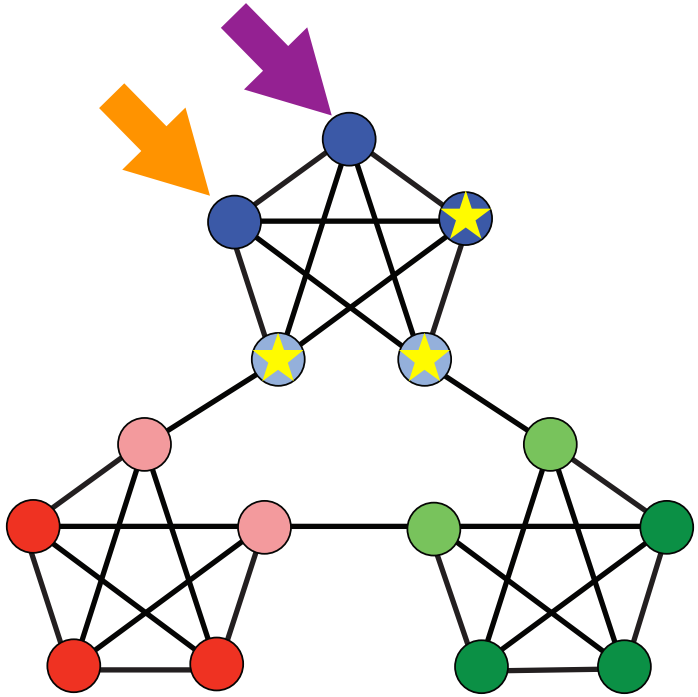
Solway et al., *PLoS Comp. Biol., 2014*

1. What should be learned?
2. Do people learn it?
3. How?

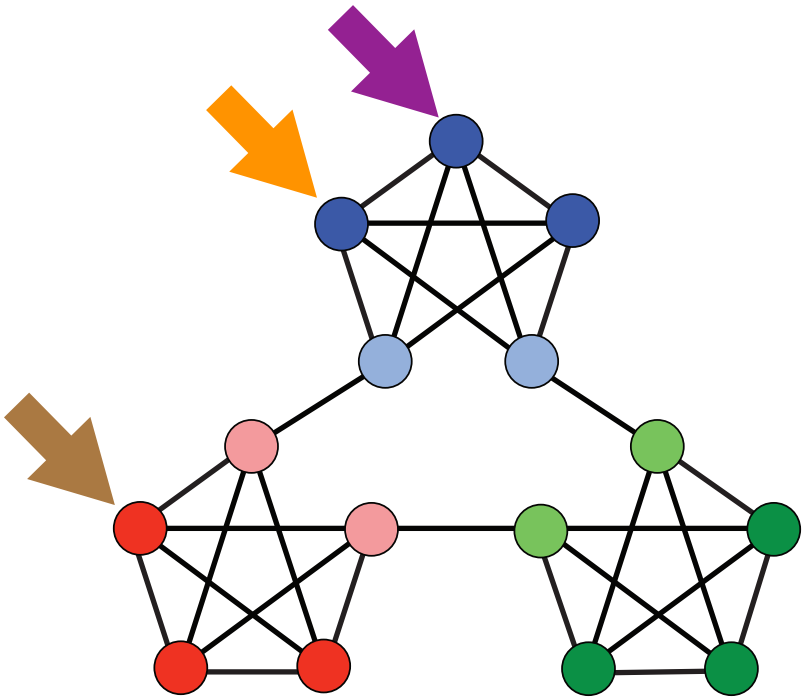


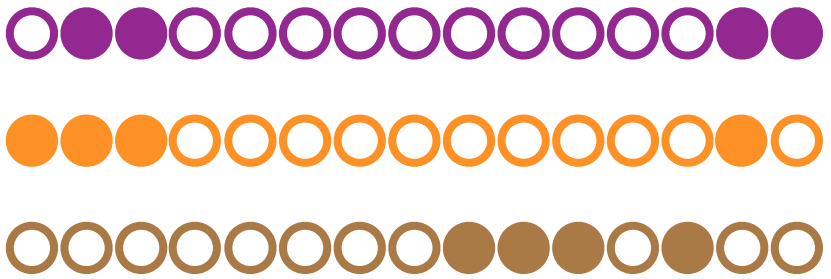
$$\Pr(data|model) = \sum_{\theta} \Pr(data|model,\theta)\Pr(\theta|model)$$

Anna Schapiro



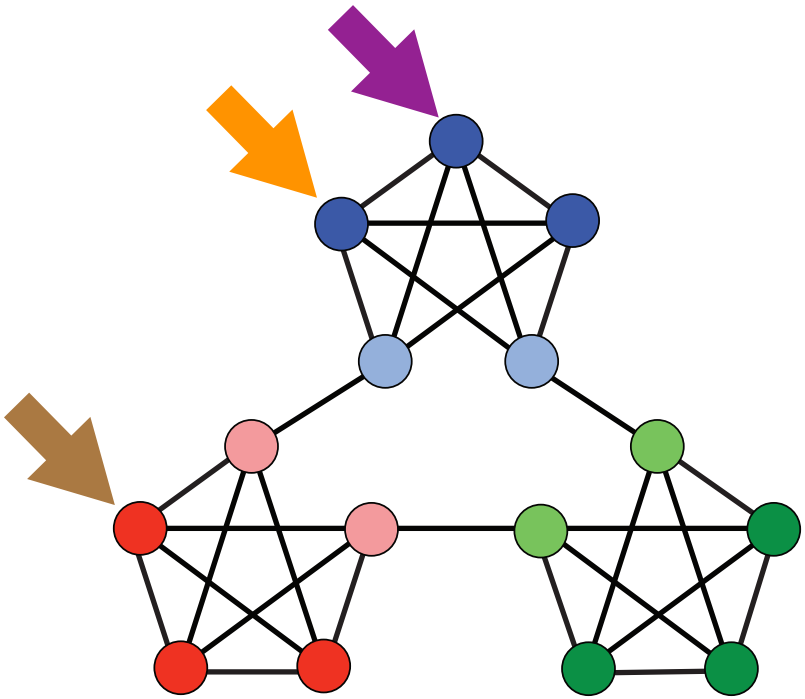
Schapiro et al., *Nature Neurosci,* 2013

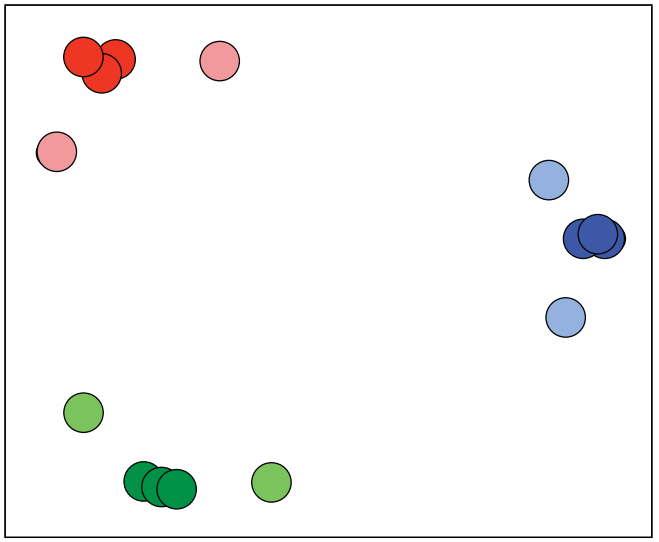




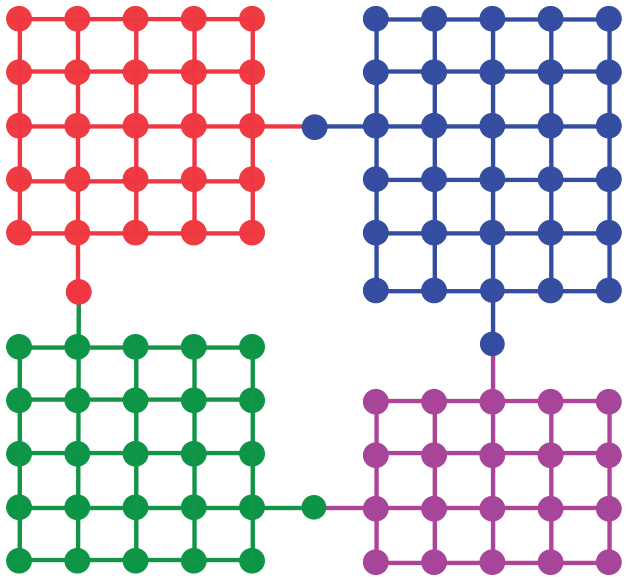
1.00	0.66	-0.36
0.66	1.00	-0.36
-0.36	-0.36	1.00

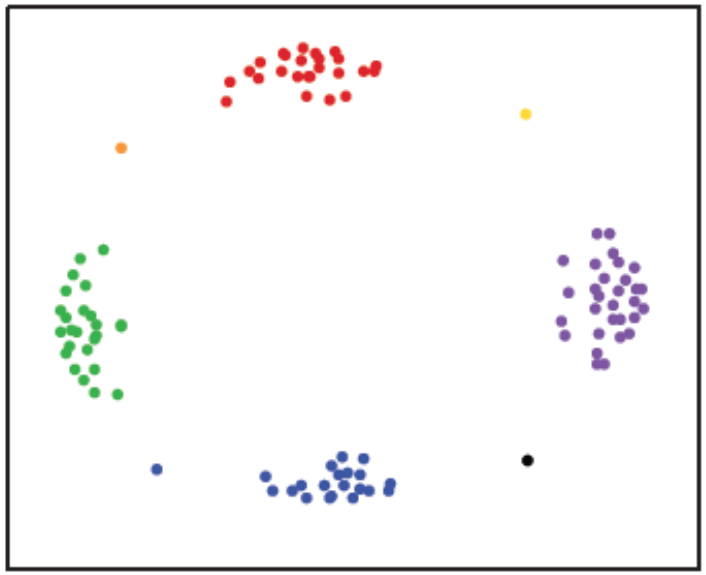
Schapiro et al., *Nature Neurosci,* 2013



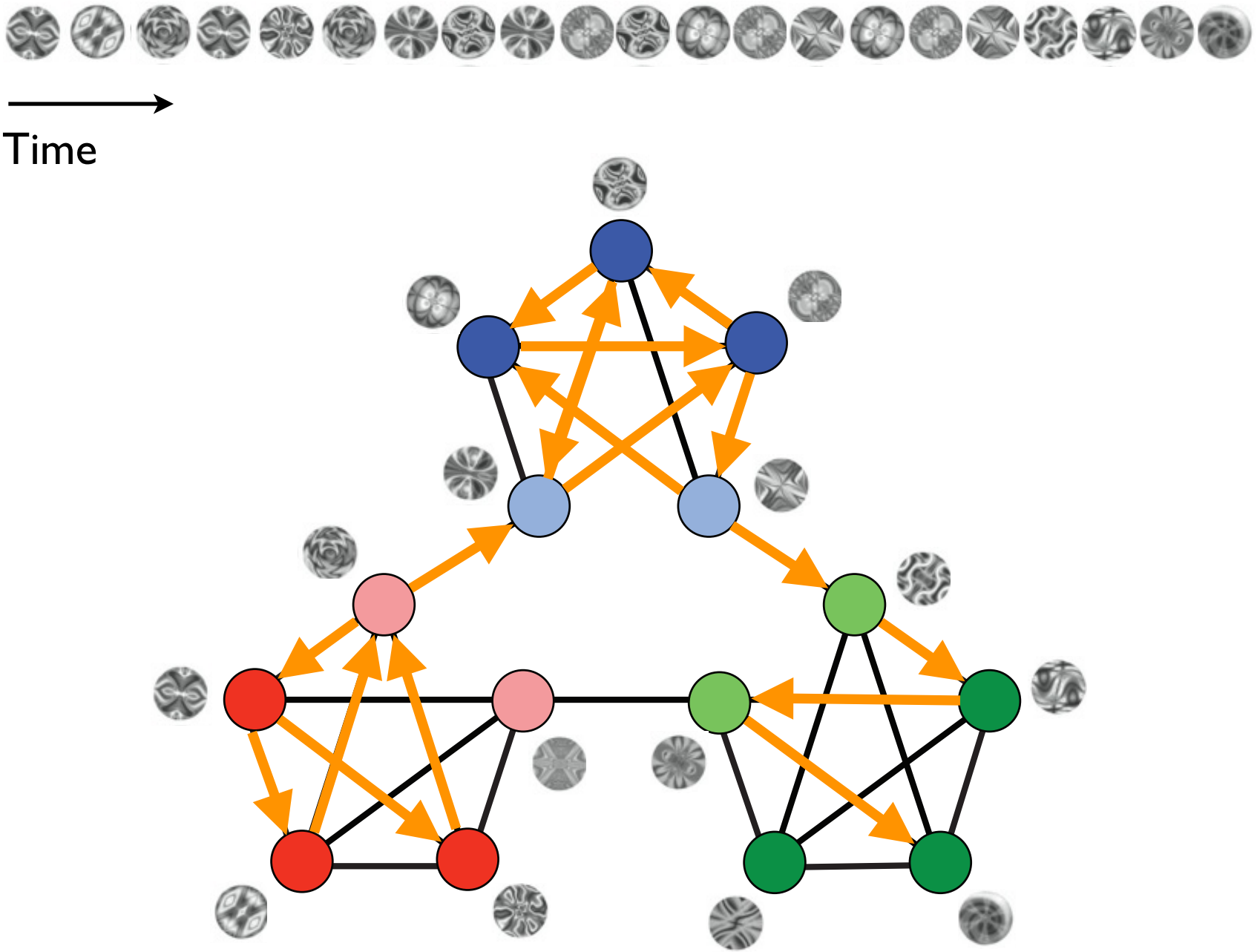


Schapiro et al., *Nature Neurosci,* 2013

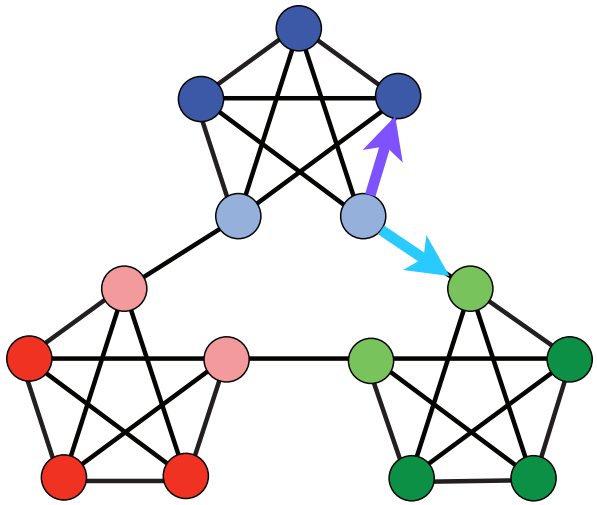


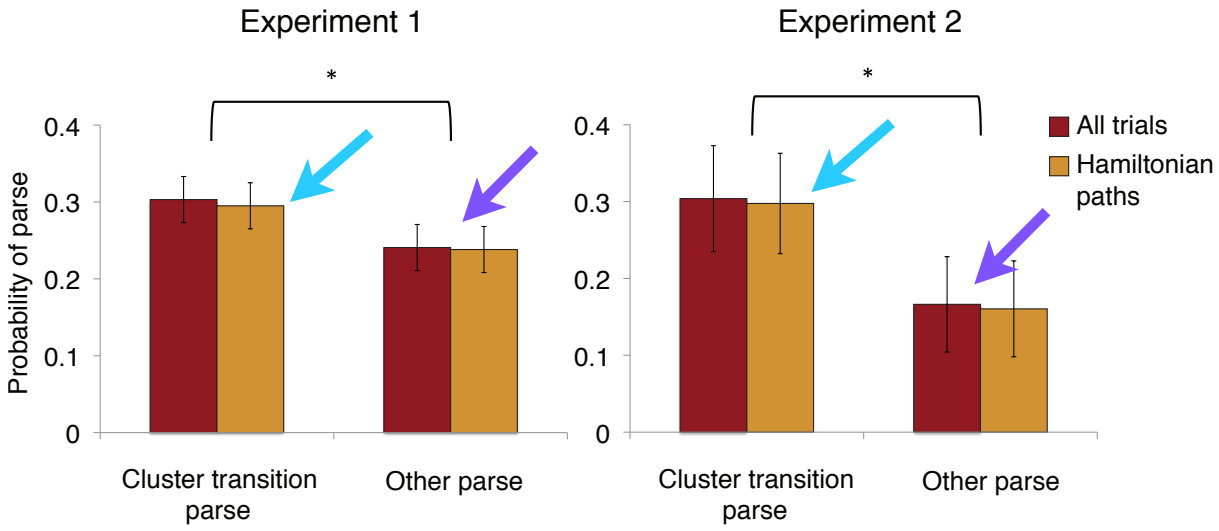


Schapiro et al., *Nature Neurosci,* 2013

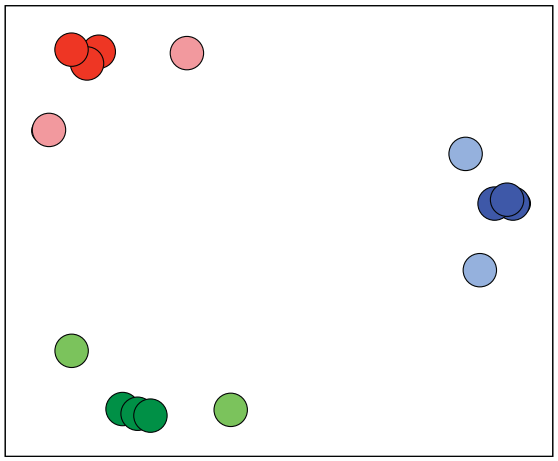


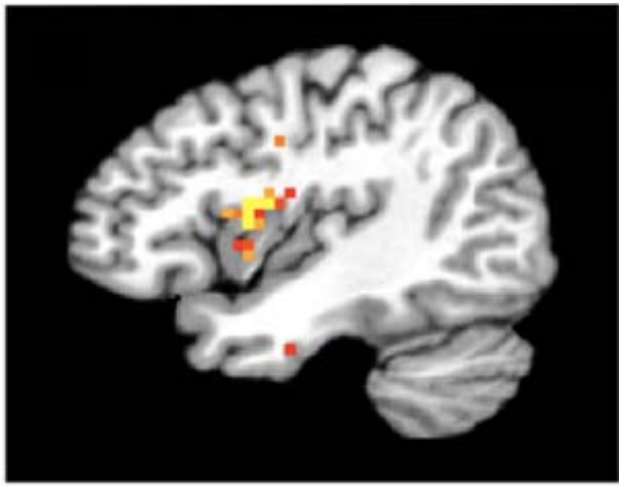
Schapiro et al., *Nature Neurosci,* 2013

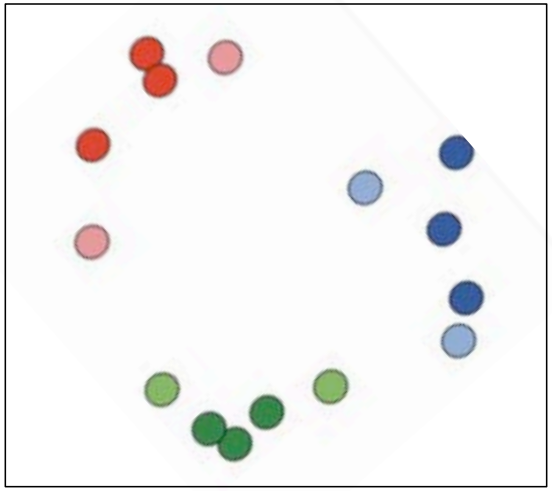




Schapiro et al., *Nature Neurosci,* 2013



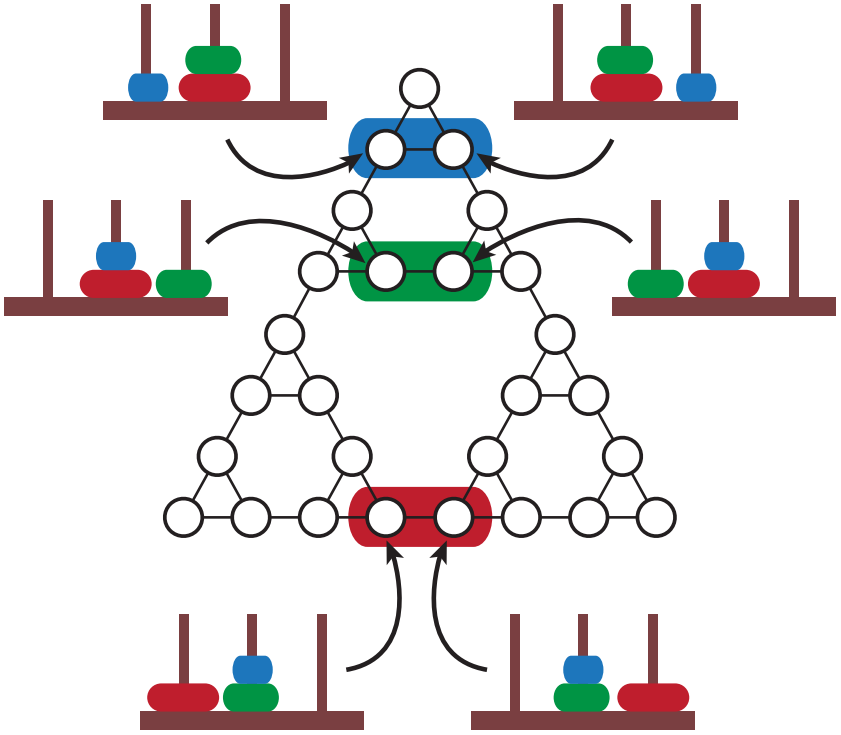


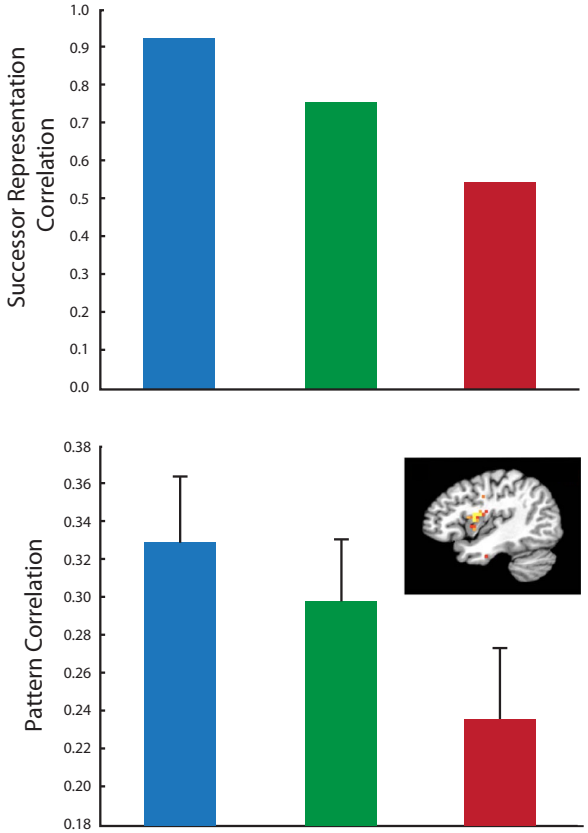


+ HC

Schapiro et al., *Nature Neurosci,* 2013

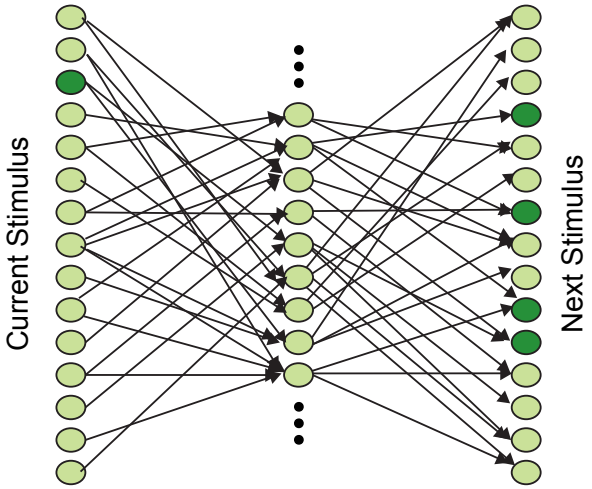
Carlos Diuk

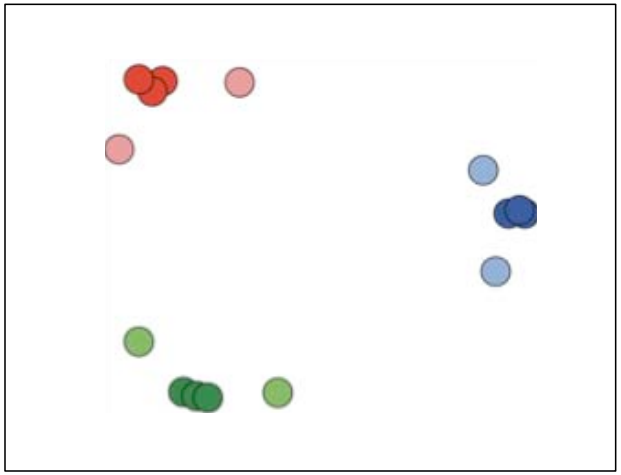


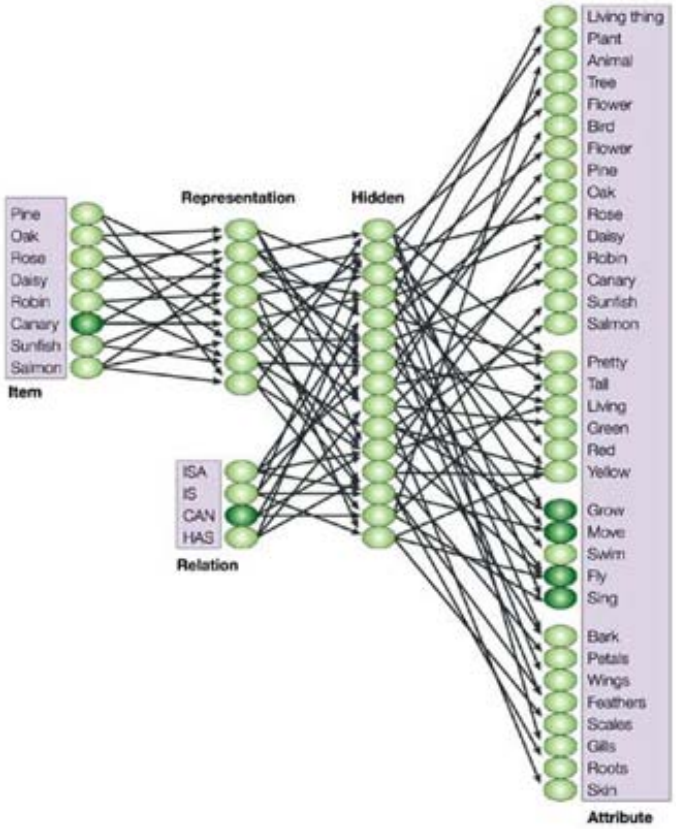


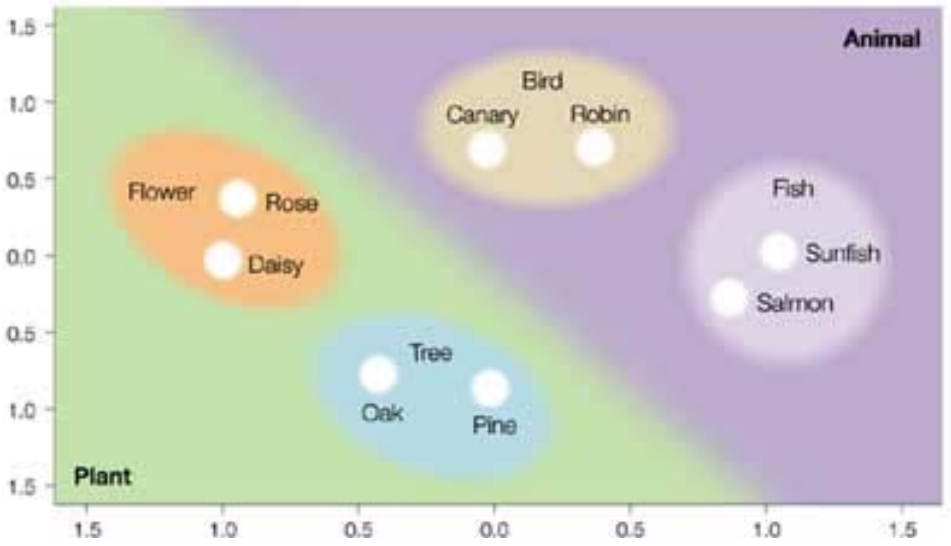
+ HC

Diuk et al., in prep.

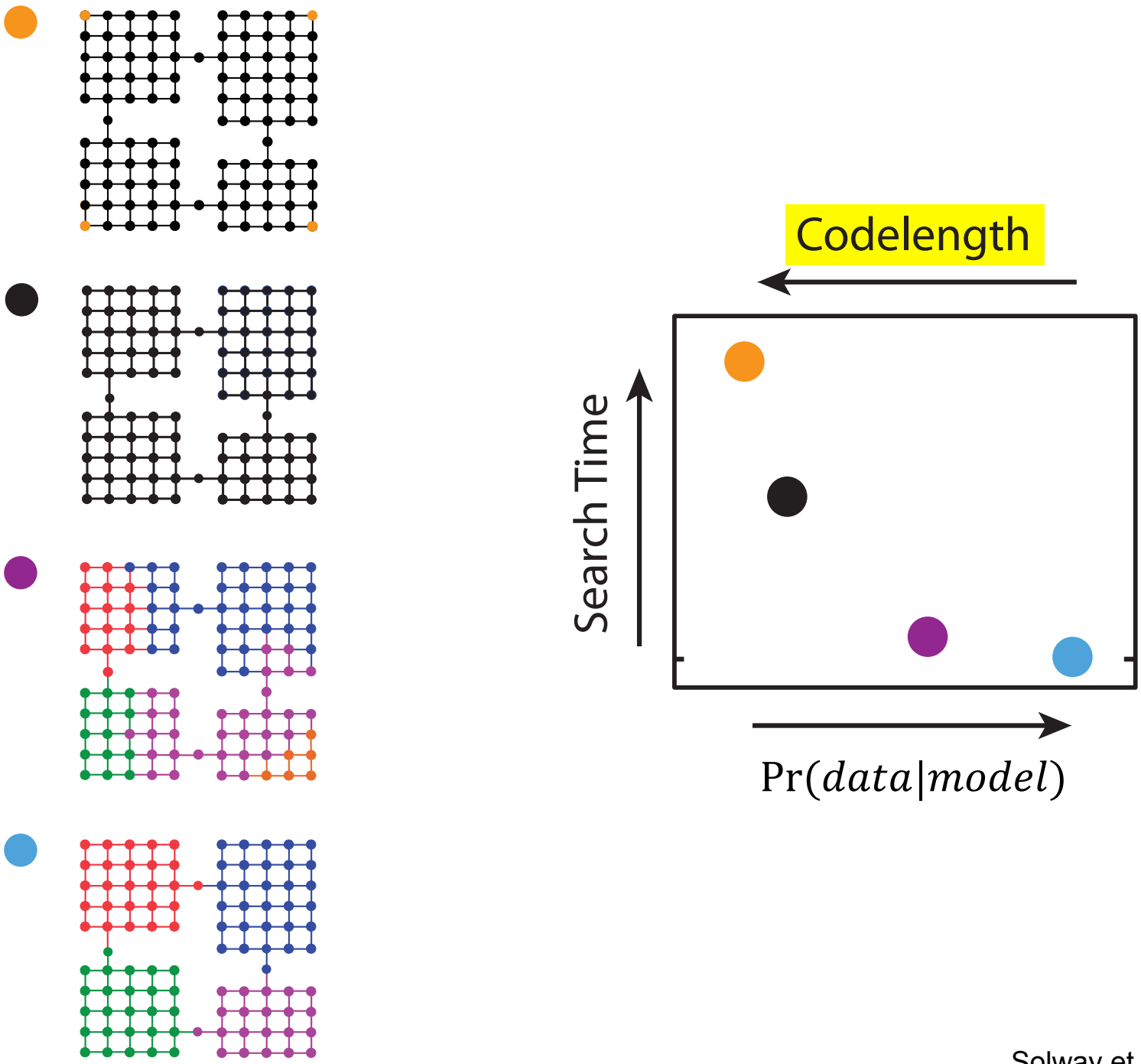




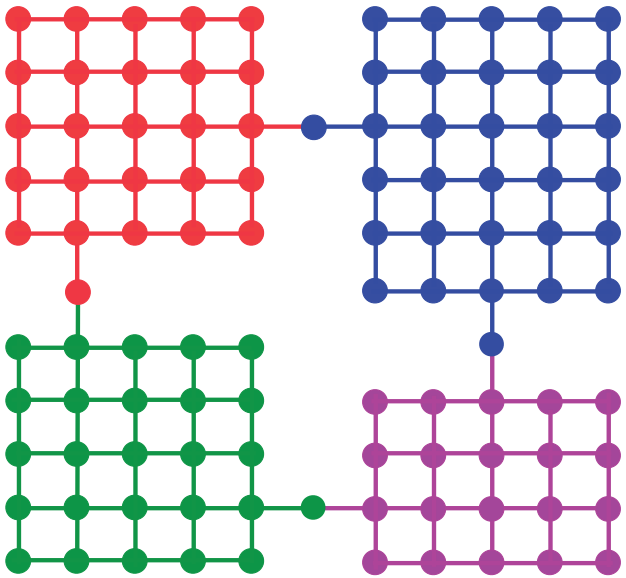


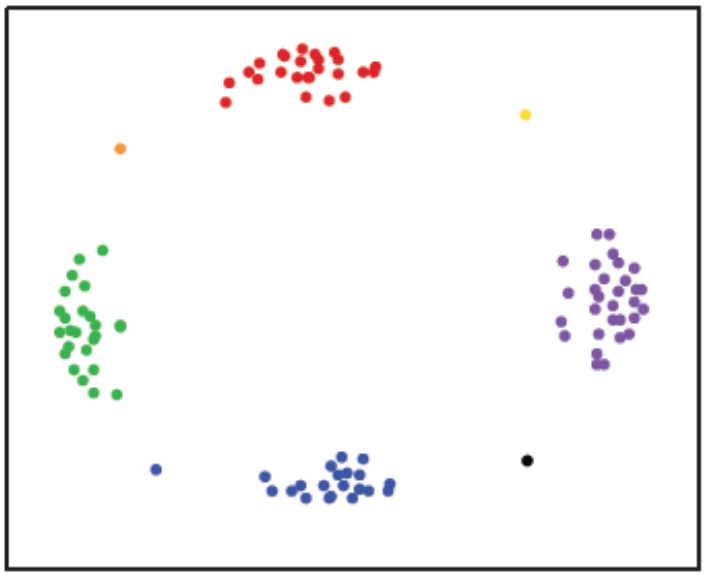


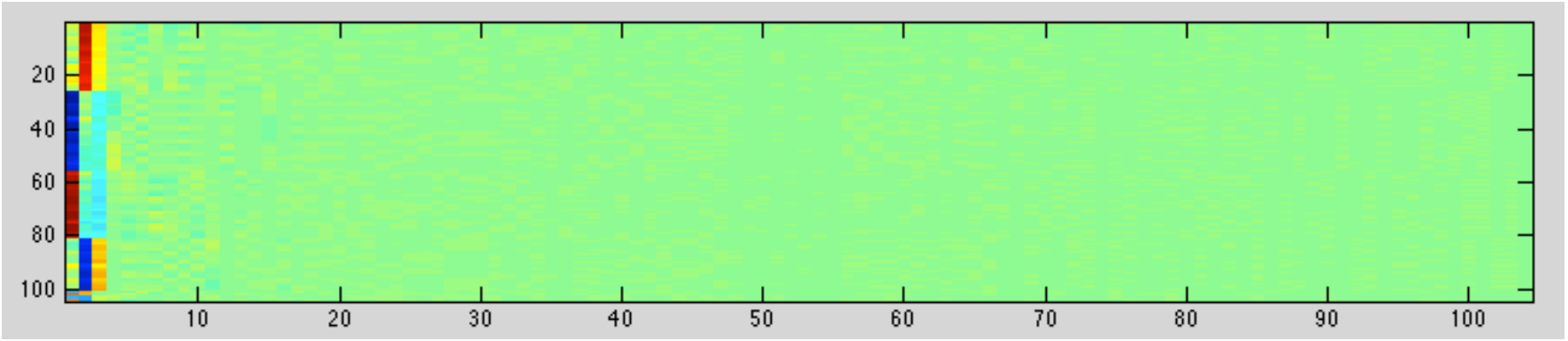
Schapiro et al., 2013.; Rogers & McClelland, 2003



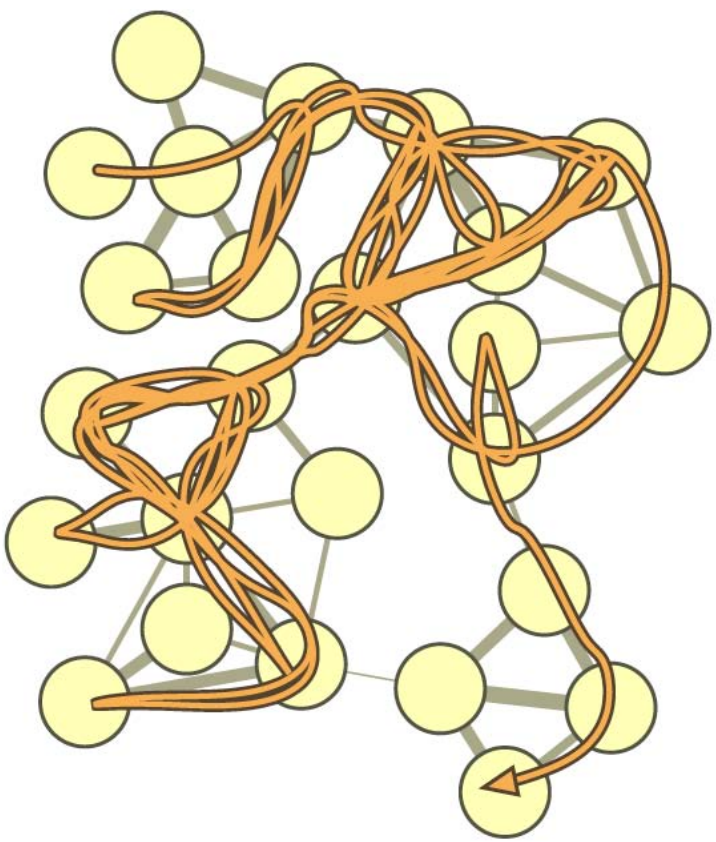
Solway et al., *PLoS Comp. Biol., 2014*

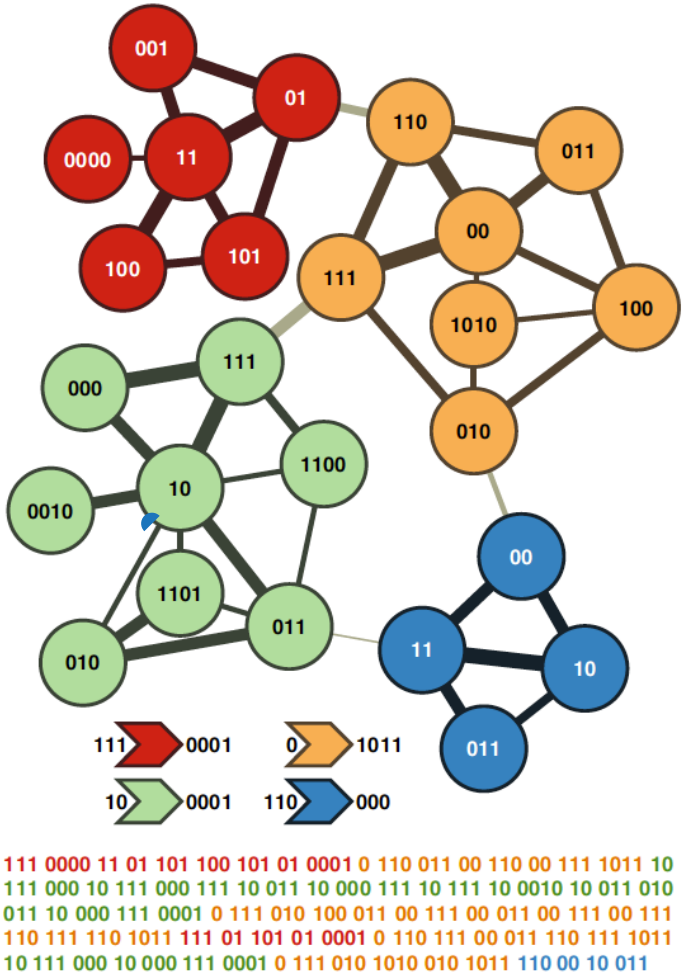






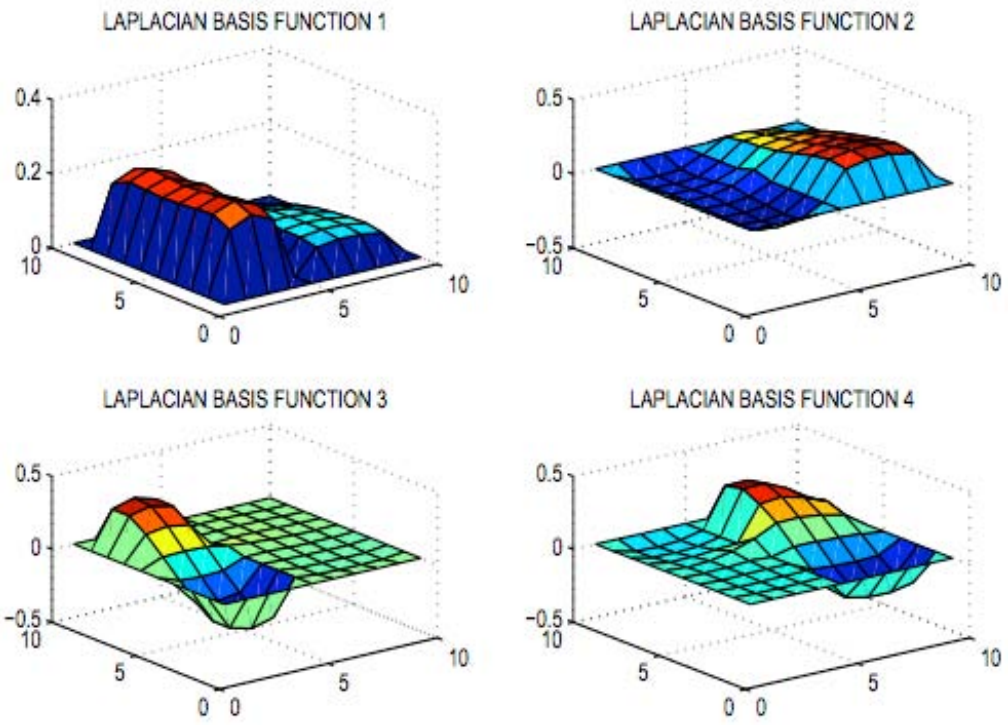
cf. Dayan, 1993



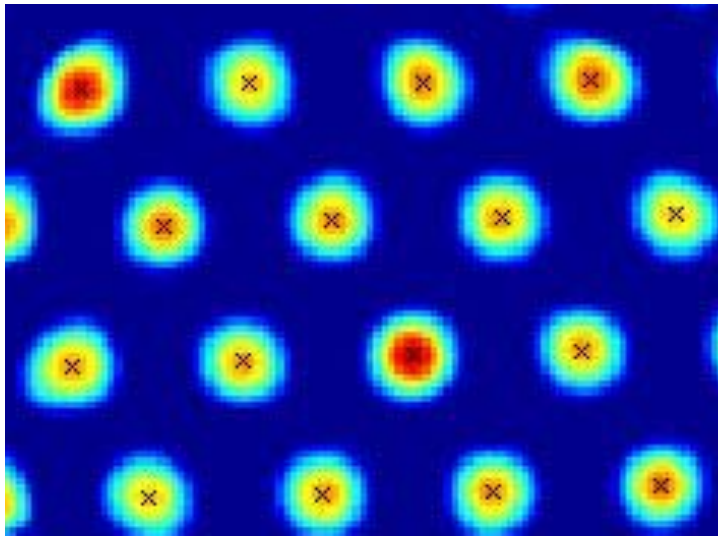


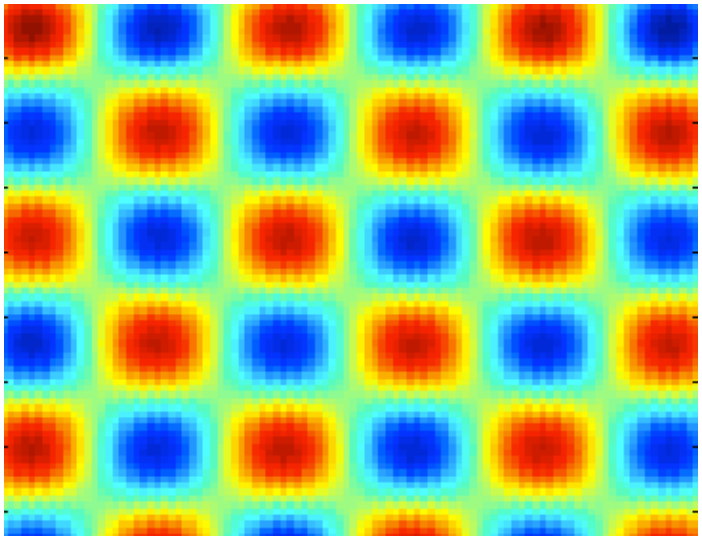
111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111 1011 10
111 000 10 111 000 111 10 011 10 000 111 10 111 10 0010 10 011 010
011 10 000 111 0001 0 111 010 100 011 00 111 00 011 00 111 00 111
110 111 110 1011 111 01 101 01 0001 0 110 111 00 011 110 111 1011
10 111 000 10 000 111 0001 0 111 010 1010 010 1011 110 00 10 011

Rosvall & Bergstrom, *PNAS,* 2008

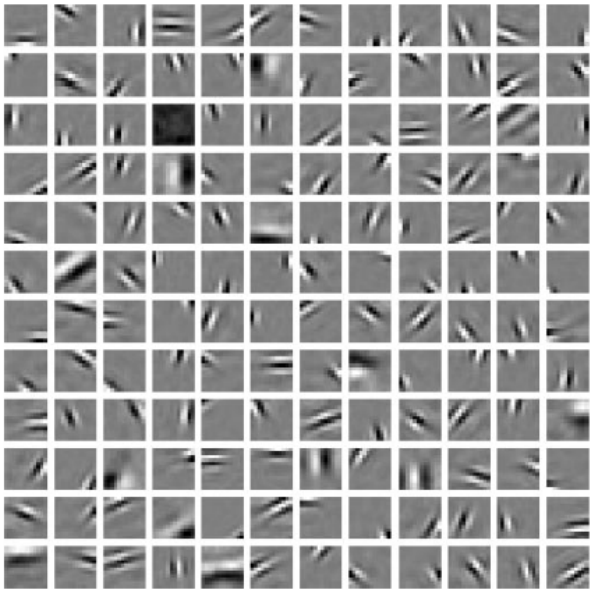


Mahadevan & Maggioni, 2005

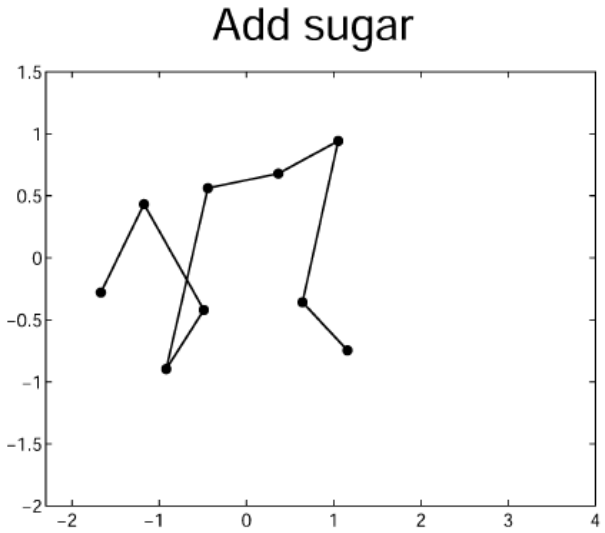


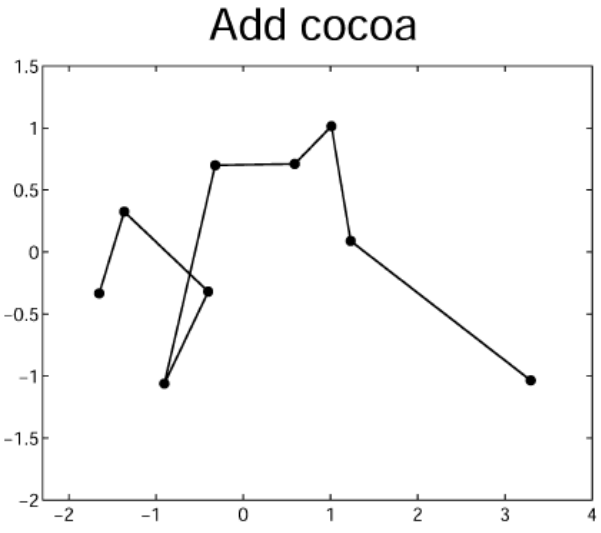


Stachenfeld, Botvinick & Gershman, *NIPS, 2014*



Olshausen & Field, *Nature,* 1996





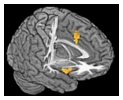
Botvinick & Plaut, *Psych Review,* 2004

Conclusions

The scaling problem in RL

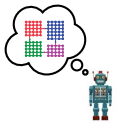
Hierarchy can help

Model-free versus model-based HRL

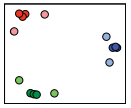


HRL in the brain

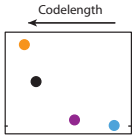
The need for good representations



Task decomposition, bottlenecks, community detection



Prospective coding and structure discovery



Hierarchy as compression

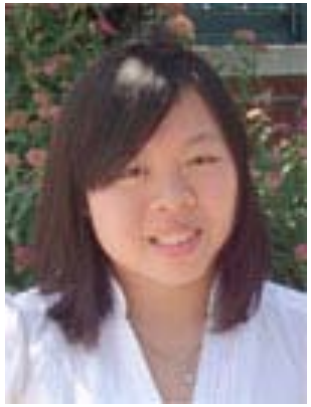



Lab Contributors

Carlos Diuk (Facebook)
Jose Ribas-Fernandes (U. Victoria)
Anna Schapiro
Alec Solway (V. Tech / UCL)
Kim Stachenfeld
Ari Weinstein
Debbie Yee (Wash. U.)

Collaborators

Andy Barto (UMass)
Yael Niv (Princeton)
Tim Rogers (Wisconsin)
Nick Turk-Browne (Princeton)

NIMH National Institute of Mental Health

National Institute of Neurological Disorders and Stroke

National Science Foundation

James S. McDonnell Foundation

John Templeton Foundation