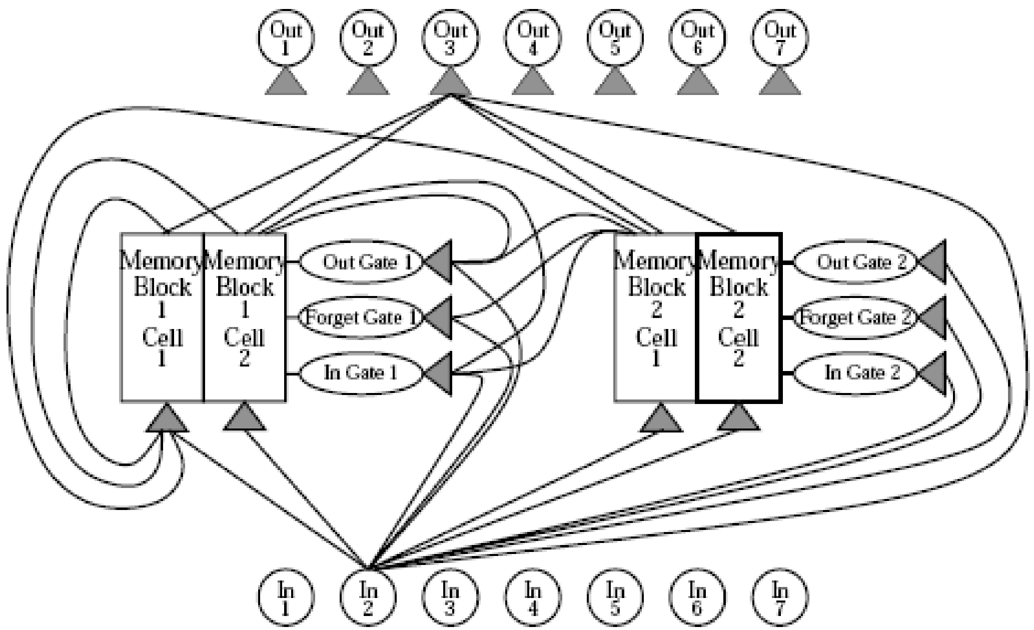
Introduction

- Humans and animals exhibit remarkable cognitive flexibility, i.e. the ability to perform new tasks quickly and flexibly.
  - Key role for prefrontal cortical areas (PFC), the striatum and neuromodulators
- Components of competence (Braver, Cohen, Frank, Fusi, O'Reilly, et al):
  - Rule learning, matching, generation (hippocampus, cortex, striatum)
  - Storage of variables – PFC working memory
  - Rapid, input/rule-dependent updating of variables – (striatal) gating
  - Use of working memory to control gating, inference, selection, attention et al
- Expect shared structure
  - Basis functions of control

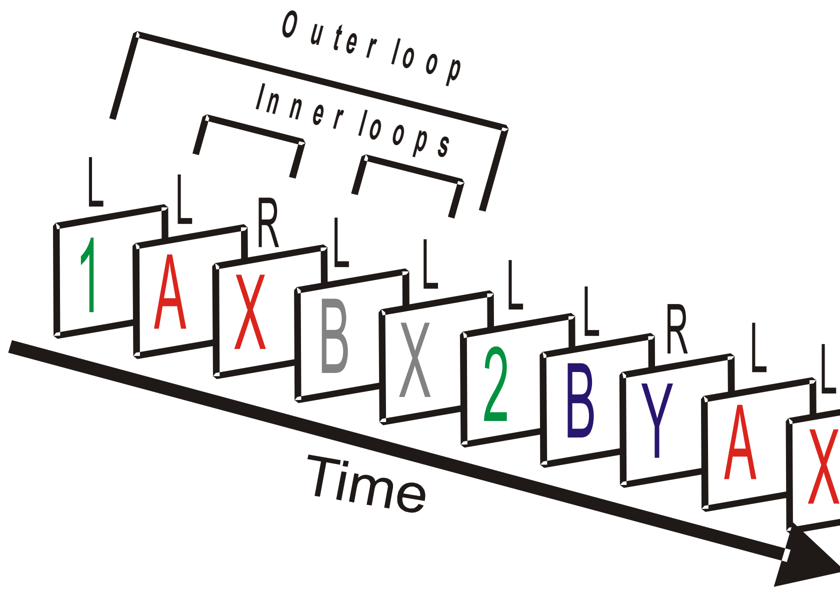
  Idea:
  - Targeted learning through shaping can acquire reusable structure
- Tested using a hierarchically structured memory task

Long-Short Term Memory (LSTM) Network



- Gating architecture proposed by Hochreiter & Schmidhuber
- LSTM is a 3 layer recurrent network with “dedicated” activation based memory cells and sigmoidal activation functions
- Gradient based supervised learning algorithm
- Learns selective, conditional gating
- Easily modularised by adding / enabling individual memory blocks

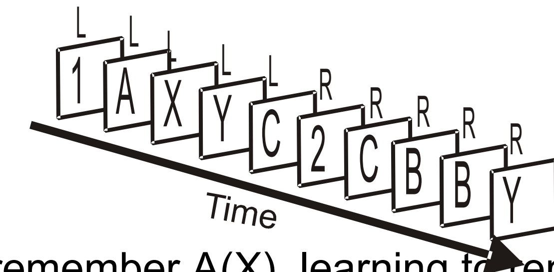
12-AX Task



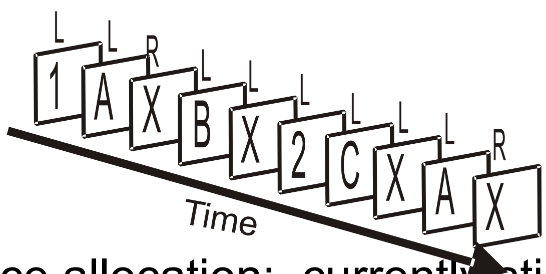
- Sequential, hierarchical decision making task
- Outer loop: present "1" or "2", followed by 1 -4 inner loops
- Inner loop: A, B, C followed by X or Y
- Target sequence: AX in context 1, BY in context 2
- Allows task variations
  - Create a family of similarly structured tasks.
- First proposed and modeled by O'Reilly et al using their PBWM framework
  - PFC-controlled striatal gating for subroutines
  - Stripe-specific error signal
- Used 12-AX to show the effects of shaping in an LSTM network

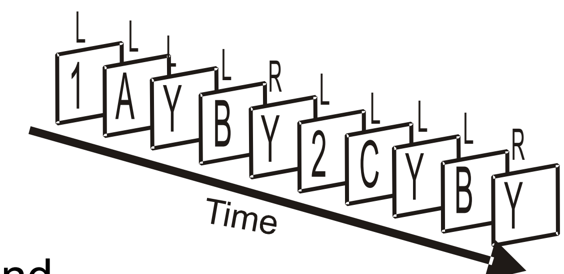
Simple Shaping

- 5 stage hierarchical process.
  - Stage 1 - 3: learning to gate 1 and 2, with increasing length (e.g. 4,12,40)

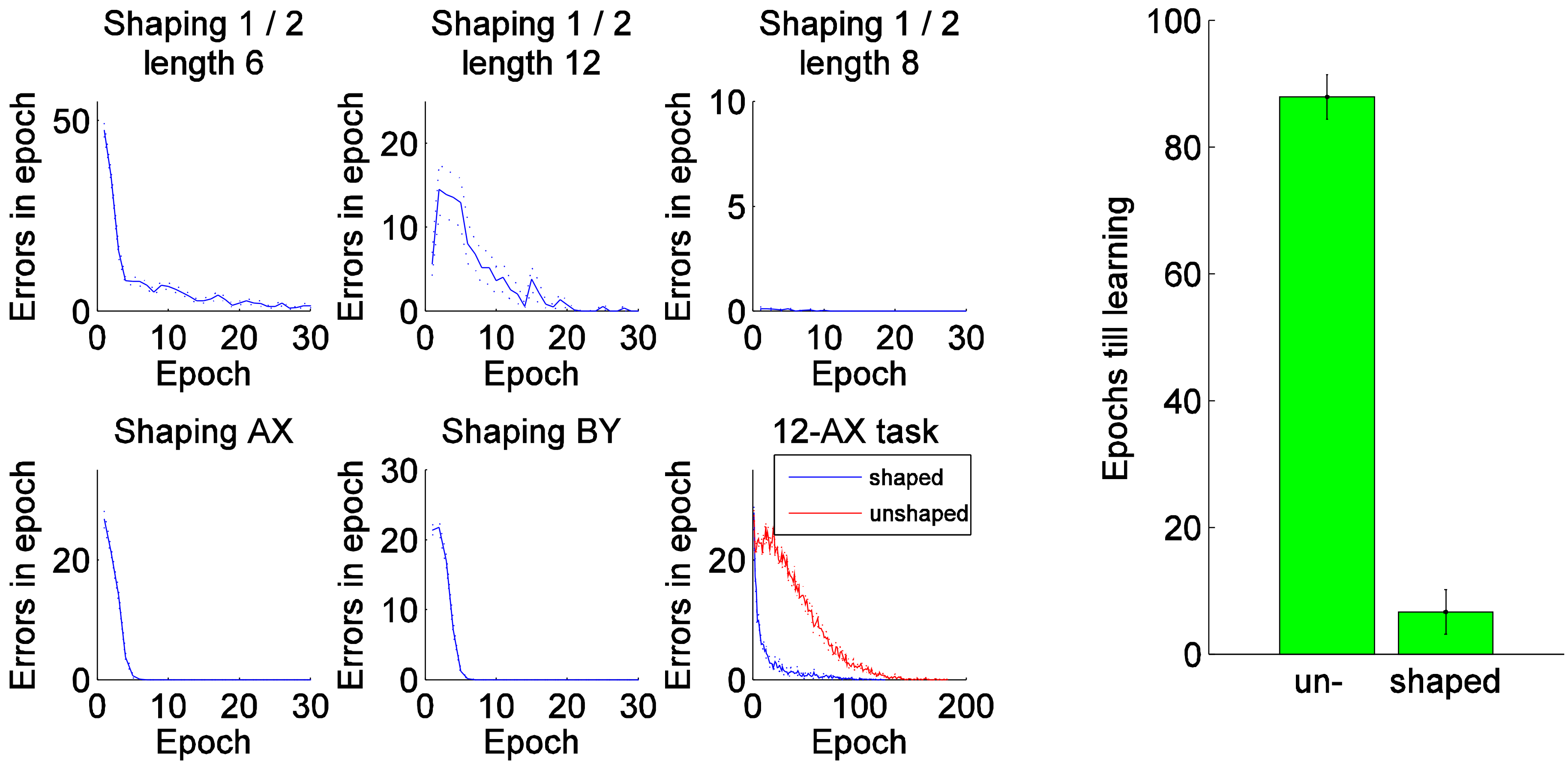


  - Stage 2 / 3: learning to remember A(X), learning to remember B(Y)





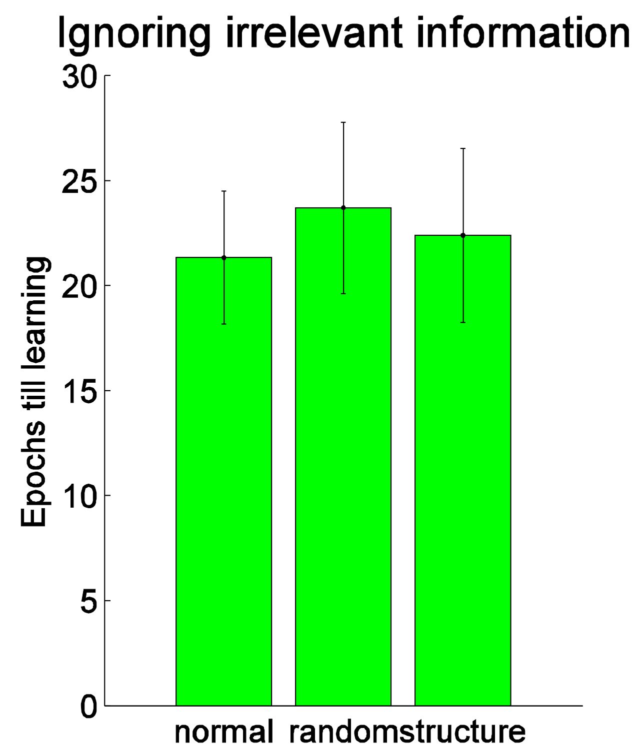
- Resource allocation: currently still done by hand
  - Only a subset of memory blocks is enabled and learning
- Full 12-AX task: all cells are enabled



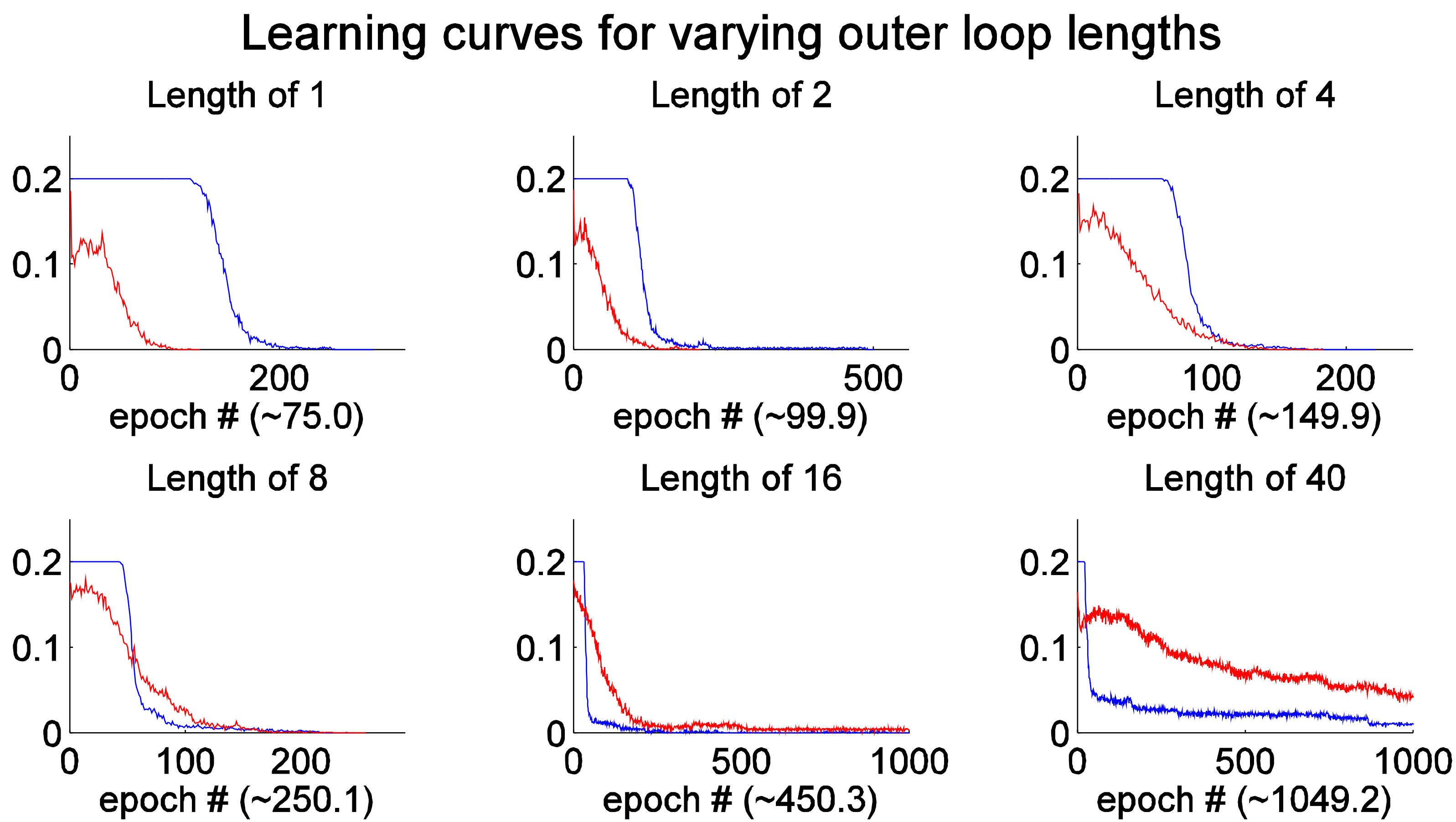
- Results: significantly faster learning due to available structure.
  - Each subtask is rapidly learned
  - Tasks combine well, sometimes requires less than ten epochs
- Including shaping time: still comparable (see temporal complexity)
- Shaping can lead to cleaner results. (see rule abstraction)

Distractibility by Irrelevant Knowledge

- Flexibility requires availability of many basis functions.
  - Need to distinguish between relevant and irrelevant basis functions
- Two types of irrelevance:
  - Correlated / similar conditions
  - Uncorrelated / widely differing
- Two controls:
  - Random structure: By adding additional randomised memory cells
  - Similar structure: Additional memory cells are trained to strongly respond to target sequence (“AB” “AZ” and differentiate A from B)
- Additional structure does not interfere



Temporal Complexity



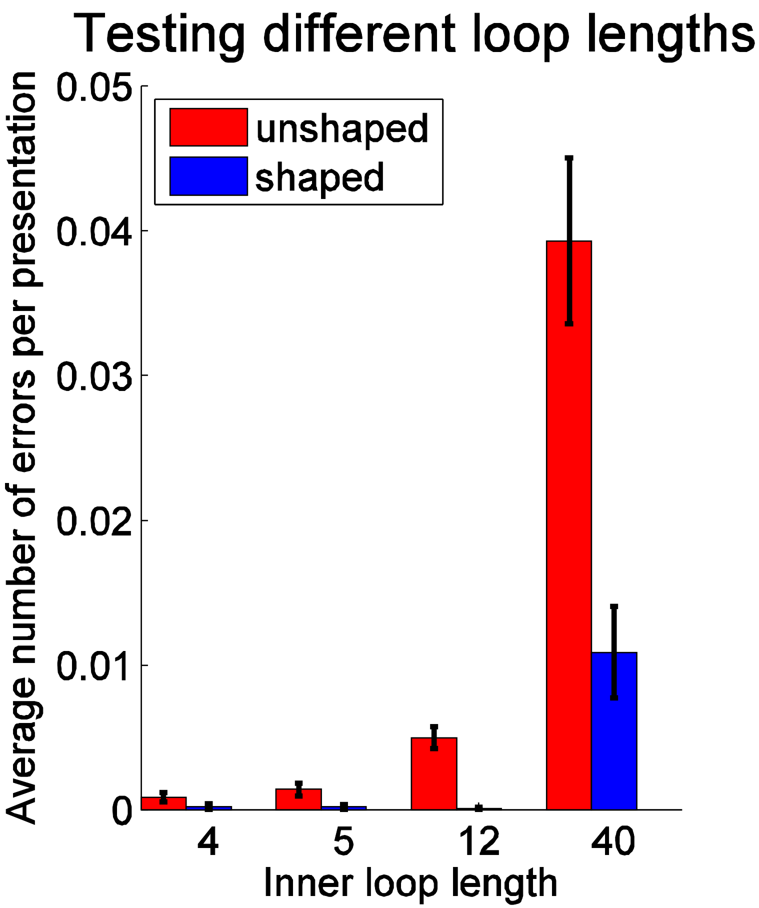
Temporal Complexity

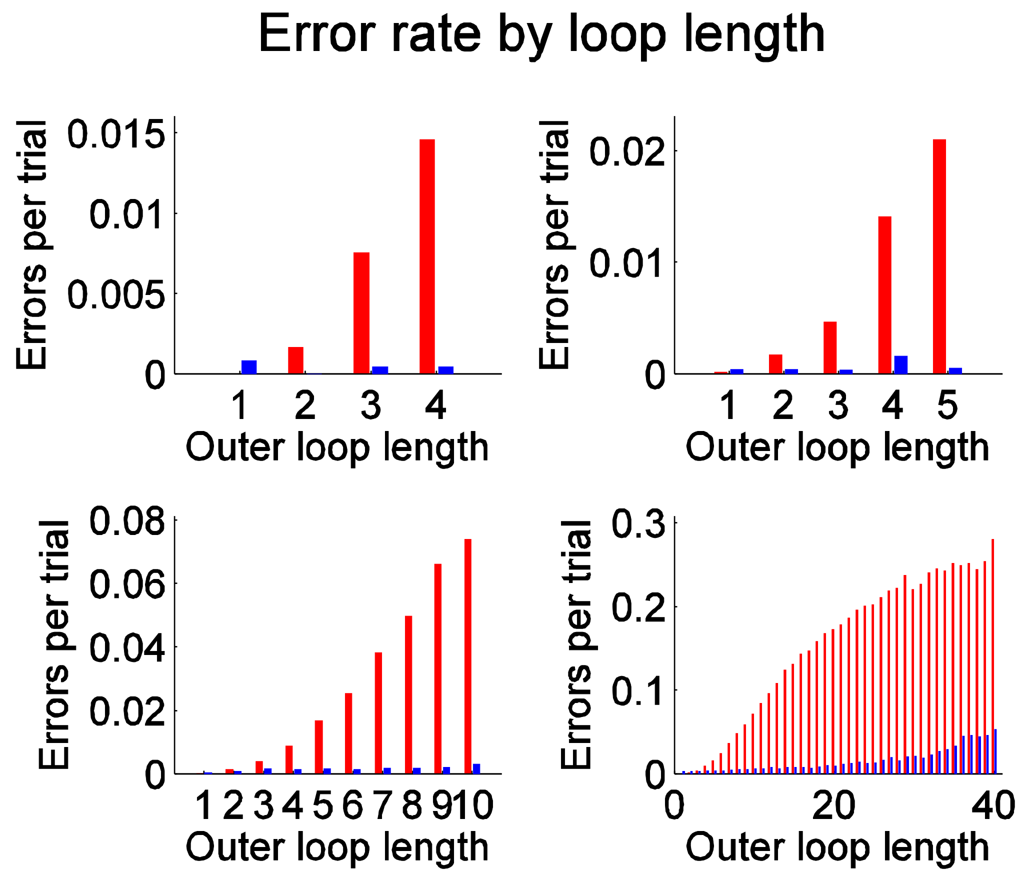
- Behavioral relevant correlations between temporally distal events are often difficult to learn (temporal credit assignment)
- Shaping creates internal representations of relevant events.
  - Strengthens correlations
  - Hence expect shaping to become increasingly favorable with increased complexity.
- Simulated complexity by altering the outer loop length
  - Vary maximum loop length between 1 and 40
  - All other conditions are otherwise unaltered
- Results confirm prediction:
  - Shaping is not necessary for very simple tasks
  - Greatly improves learning for temporally complex tasks

Rule Abstraction

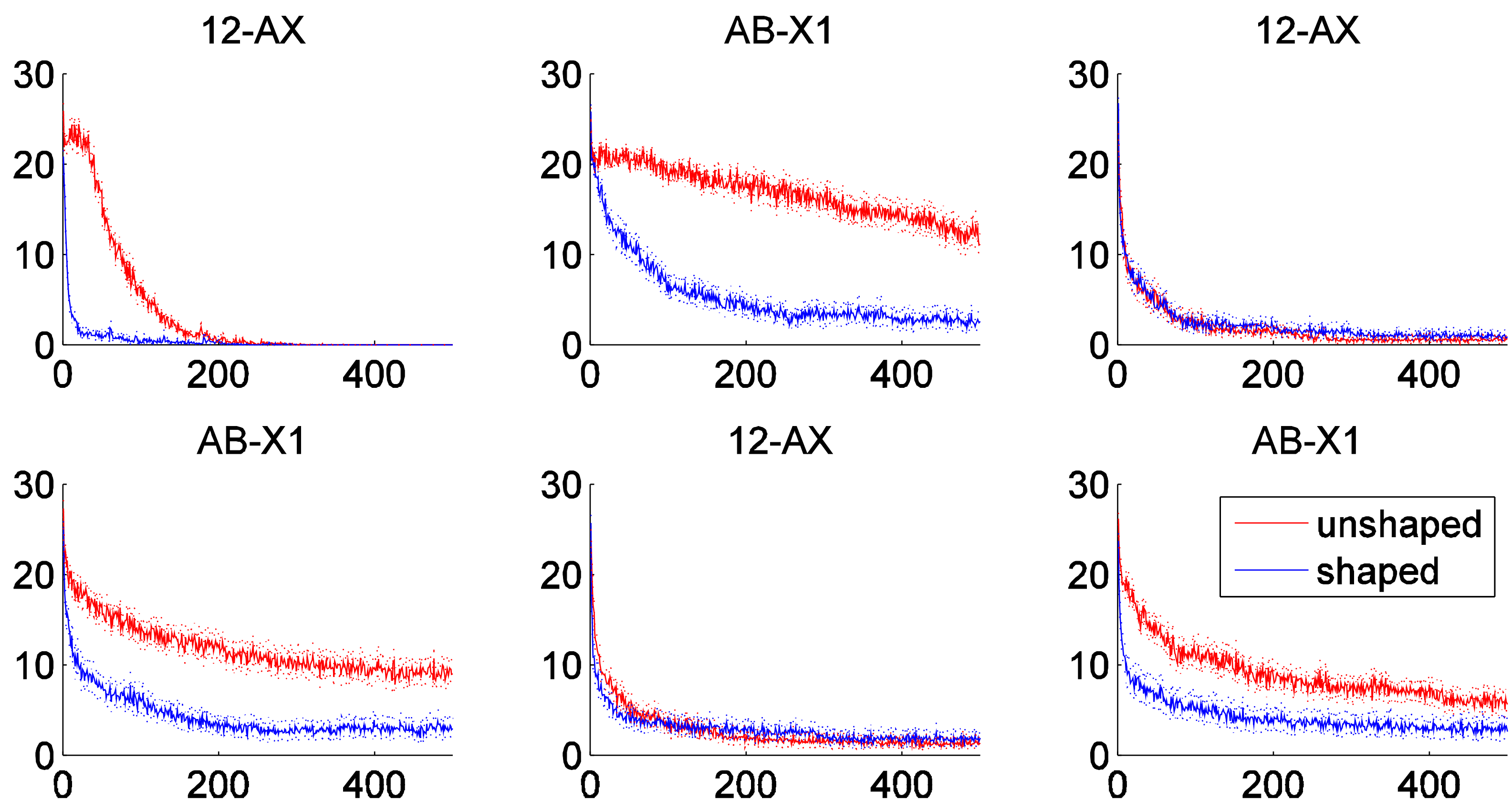
- Question: How well are rules abstracted to their essential structure?
  - Train on base 12-AX task
  - Test with decreasing context frequency (increased loop length)
    - Abstract rules underlying the task variations are identical
    - Should perform perfectly
- Hypothesis: Due to its separately learned components, the shaped network is capable of capturing rules more abstractly.
- Results: shaped network outperforms monolithic learning
  - Performance degrades rapidly for unshaped network
  - Even well trained lengths of 3 and 4 cause slight troubles
  - Four fold improvement at longest length
- Correlation between loop length and errors
  - Unshaped network: strong correlations
  - Shaped network: close to no correlations up to length of 20, weak afterwards

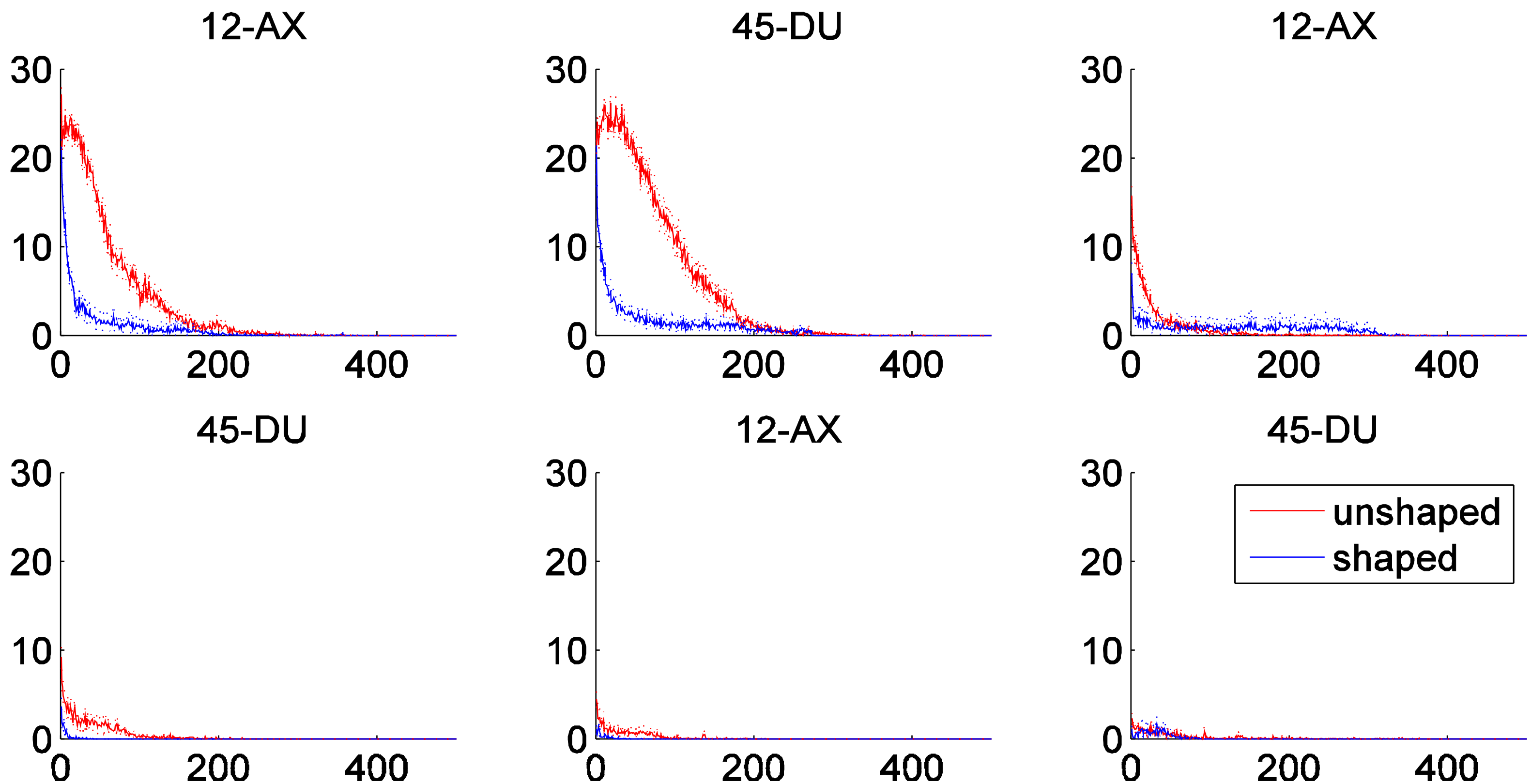
Rule Abstraction





Flexible Representations





- Flexibility requires rapid adaptation to task changes
- Reversal learning: (AB-X1 task)
  - Similar inputs, changed reward contingency
  - Shaped network: performs better as only needs to remap basis functions
  - Unshaped network: problems with repeated reversals
- Task shifting: (45-DU task)
  - Due to different input symbols little interference
  - Results are similar to normal learning for both networks

Conclusion

- We have compared learning a sequential decision making task between
  - Unstructured naïve networks
  - Networks previously exposed to relevant but simpler elements
- Shaping is mostly beneficial in learning
  - Quickly recombines relevant components, ignores others
  - Greatly reduces learning time
- Shaping can reduce the difficulty with long temporal sequences
  - Convert temporal credit assignment to structural assignment
- It allows for quicker task switching
  - Encourages a more modular representation in the network
- Animal training has always used shaping as a method
  - Should consider training procedure more while modeling learning
  - Look into the ability to reuse prior knowledge, transfer learning
  - Rather than increase complexity of monolithic learner

Future Work

- Resource allocation:
  - Need a principled way of achieving (based on uncertainty?)
    - E.g. mixture of experts, ART…
  - Currently done by hand (homunculus)
- Reinforcement Learning:
  - Extend learning from supervised to more natural reinforcement learning
- Variable substitution
  - Currently neither network can truly represent general rules
    - i.e. apply identical rules to different inputs
  - Need a variable substitution mechanism
- Bilinear matching
  - Computational basis functions
  - Match to inputs and state provided by PFC
  - Actions control WM gating
  - Interaction between PFC, Striatum and Hippocampus
- Human and animal performance on 12-AX task
  - What are the constraints given by natural learners?

References

1. O'Reilly, Randall C. and Frank, Michael J. Making Working Memory Work: A Computational model of learning in Prefrontal Cortex and Basal Ganglia*, Neural Computation,* 18 (2), 2005
2. Gers, Felix A. and Schmidhuber, Juergen and Cummins, Fred, Learning to forget: Continual prediction with LSTM*, Neural Computation,* 12 (10), 2000
3. Braver, Todd S. and Cohen, Jonathan D. On the control of control: The role of dopamine in regulating prefrontal function and working memory. *Control of cognitive processes: Attention and performance XVIII*, 2000
4. Cohen, Jonathan D. and Braver, Todd S. and O'Reilly, Randall C. A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges, *Philosophical Transactions of Royal Society (London) B*, 351, 1996
5. Durstewitz, Daniel and Seamans, Jeremy K. and Sejnowski, Terrence J. Dopamine-Mediated Stabilization of Delay-Period Activity in a Network Model of Prefrontal Cortex, *J. Neurophysiol*, 83, 2000
6. Jordan, Michael I and Jacobs, Robert A. Hierarchical mixtures of experts and the EM algorithm, *Neural Computation*, 6, 1994
7. Carpenter, G. A. and Grossberg, S. Adaptive Resonance Theory, 1994

Acknowledgments

Support from the Gatsby Charitable Foundation