#### Jared Sylvester 15 July 2013

# The Neural Executive

#### Can Gated Attractor Networks Account for Cognitive Control?



Computer Science



# GALIS

#### Gated Attractors Learning Instruction Sequences

## **Artificial Intelligence?**

"AI is the study of how to make computers do things at which, at the moment, people are better."



– E. Rich & K. Knight, *Artificial Intelligence*, 1983

**Deep Blue**, 1997

**Watson**, 2011



## **Artificial Intelligence?**

"AI is the attempt to make computers do what they do in the movies."

– Astro Teller



## **Artificial Intelligence?**

"AI is the attempt to make computers do what they do in the movies."

– Astro Teller



## Motivation

- Gap between neural & symbolic AI systems
   Symbolic: planning, goals, rules, deduction, ....
   Neural: perception, motor control, pattern rec, ...
- Neural systems are very "hard-wired"
  - Behavior is often baked into architecture
  - New problems require entirely new systems



GALIS is designed to address both of these



# **Cognitive Control**

- ♦ Managing & recruiting other cognitive processes.
  - ◊ "Executive functions"
  - *e.g.* maintaining working memory, shifting attention, selecting responses, setting goals, inhibiting irrelevant stimuli...



## GALIS' Three Hypotheses

- Cortex is a distributed net of interacting regions
  - ♦ Functional segregation & function integration
- ♦ Regions ≈ attractor nets
   ♦ Can process sequences
- Regions "gate" each other as well as passing input
  - *i.e.* modulate flow between other regions





Adapted from: Vértes, et al. "Simple models of human brain functional networks." PNAS 2012.

### Working Memory: Sequential Attractor Nets









time

### Working Memory: Sequential Attractor Nets



# Adding Cognitive Control

Multiple networks

 Linked with gated connections
 Two memory layers
 Original one for external stimuli
 New memory for task procedures
 Using the same paradigm
 Task memory controls gating

















[ make tea | ? ] [ make tea | boil water ]



17

### n-Back

♦ Given sequence of inputs:

...does most recent input match input *n* steps ago?

Must maintain sequence in WM; make judgments



- ♦ GALIS model learns n=1,2,3,4,5
  - Learns all five without knowing which it will perform
  - Version determined by input patterns only

#### **GALIS** Architecture





### Human Comparison: Accuracy



### Human Comparison: Response Time



### Changing *n*



### **Visuospatial Architecture**



### **Visuospatial Architecture**



## Reconcile the Neural and the Symbolic

#### Neural

- Hopfield nets
- Non-linear
   interactions
   between layers
- Hebbian learning



### Symbolic

- Discrete attractors
- ♦ Gating { open/closed excite/inhibit update/maintain
- Stored programs



## Instruction vs. Construction

#### Behavior based on memory contents not just architecture



Bletchley Park Bombe, 1940



von Neumann & ENIAC, 1946

◇ Can "program" a neural net
 ◇ Now programs are hand-crafted by modeler
 ◇ Store → improve → learn ab initio

#### Thank you

www.cs.umd.edu/~jared/

Sylvester, Reggia, Weems & Bunting. "Controlling Working Memory with Learned Instructions." *Neural Networks* 41. 2013.