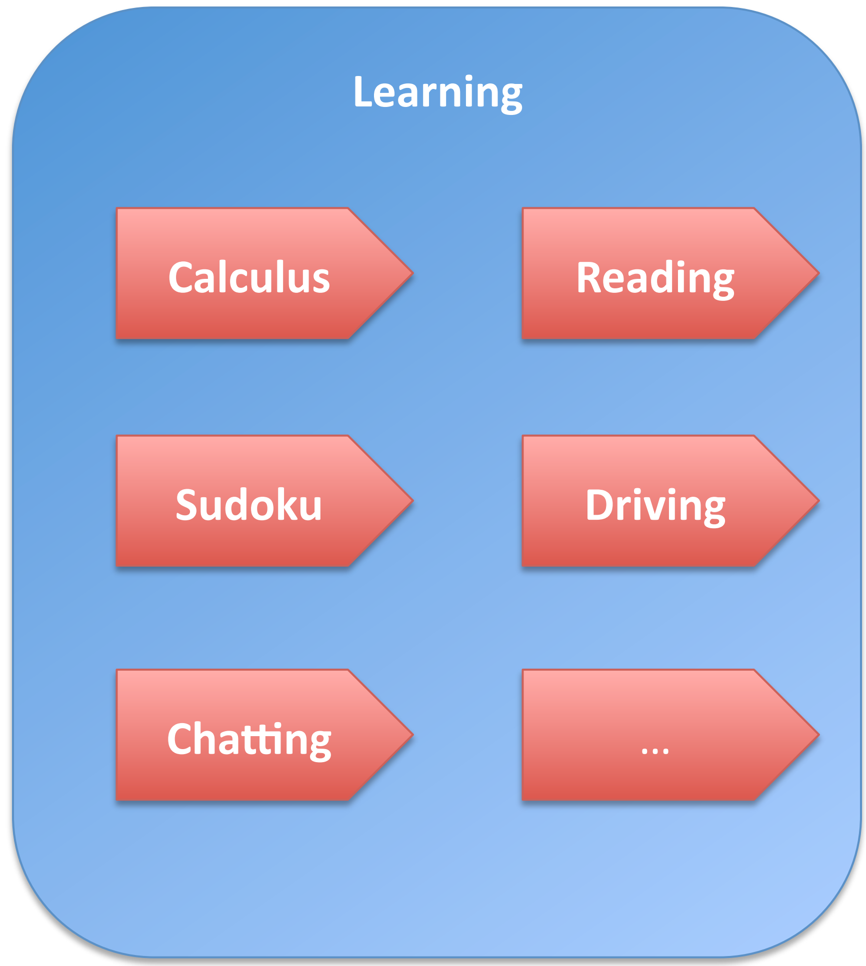
# Long-Term Declarative Memory for Generally Intelligent Agents

## Thesis Proposal

Nate Derbinsky

# Generally Intelligent Agents

- Autonomous
- Continually embedded in a diverse, dynamic environment
- Long-living
  - Months, Years
- Multiple, complex tasks



# Long-Term Memory Systems

Class of mechanism to cope with dynamic, partially-observable environment

- **Encodes** experience
- **Stores** internally
- Supports **retrieval**

Human memory

- Biased, error-prone
- Continually able to encode new experience
- Lends to improved performance with greater task experience

# Long-Term Memory (LTM) for Artificial General Intelligence (AGI)

AGI Constraints -> LTM Requirements [Laird & Wray, 2010]

- No concurrent solution
- Insufficient comprehensiveness of learning
- Opportunities for cross-fertilization

| | Incremental Learning | Comprehensive Learning | Diverse Representation | Scale Efficiently | Effective Access | Task Independence |
|---|---|---|---|---|---|---|
| Cognitive Modeling/ Architecture | ☑ | ☐ | ☑ | | ☑ | ☑ |
| Case-Based Reasoning | ☐ | ☐ | ☐ | ☐ | ☑ | |
| Information Retrieval/ Databases | | ☐ | ☑ | ☑ | ☐ | ☑ |
| Knowledge Representation Reasoning | ☐ | | ☑ | ☐ | ☑ | ☑ |

# Our Proposal

| | Incremental Learning | Comprehensive Learning | Diverse Representation | Scale Efficiently | Effective Access | Task Independence |
|---|---|---|---|---|---|---|
| Cognitive Architecture | ☑ | ↑ | ☑ | ↑ | ☑ | ☑ |

Explore…

– 2 long-term declarative memory systems

– 2 research questions

- Encoding/Storage, Retrieval

# Proposal: Memory Systems

Psychological and computational evidence for the functional necessity of dissociated memory systems

[Derbinsky & Laird, 2010]


“Know”

**Semantic**

Context A


owledge


“Remember”

**Episodic**

# Proposal: Research Questions

Encoding/Storage

What agent experience should a task-independent memory system *encode* and *store*?

Retrieval

How can a task-independent memory system *retrieve* the most useful knowledge?

# Outline

Introduction

Prior Work
- Episodic
- Semantic

Future Work
- Evaluation Strategy
- Proposed Extensions
- Timeline

# Prior Work

**Focus**

Understand the efficiency challenges in extending the Soar cognitive architecture with basic episodic and semantic memory functionality

Requirements

- Task-independence
- Expressive representation
- Scales to large knowledge
  - Boundedness, 50-100ms

**Outline**

- Human "definition"
- Functional benefits
- Related work
- Architectural integration
- Contributions
- Scaling evaluation

# Episodic Memory in Humans

Long-term, contextualized store of specific events [Tulving, 1983]

What you "remember" vs. what you "know"

Properties

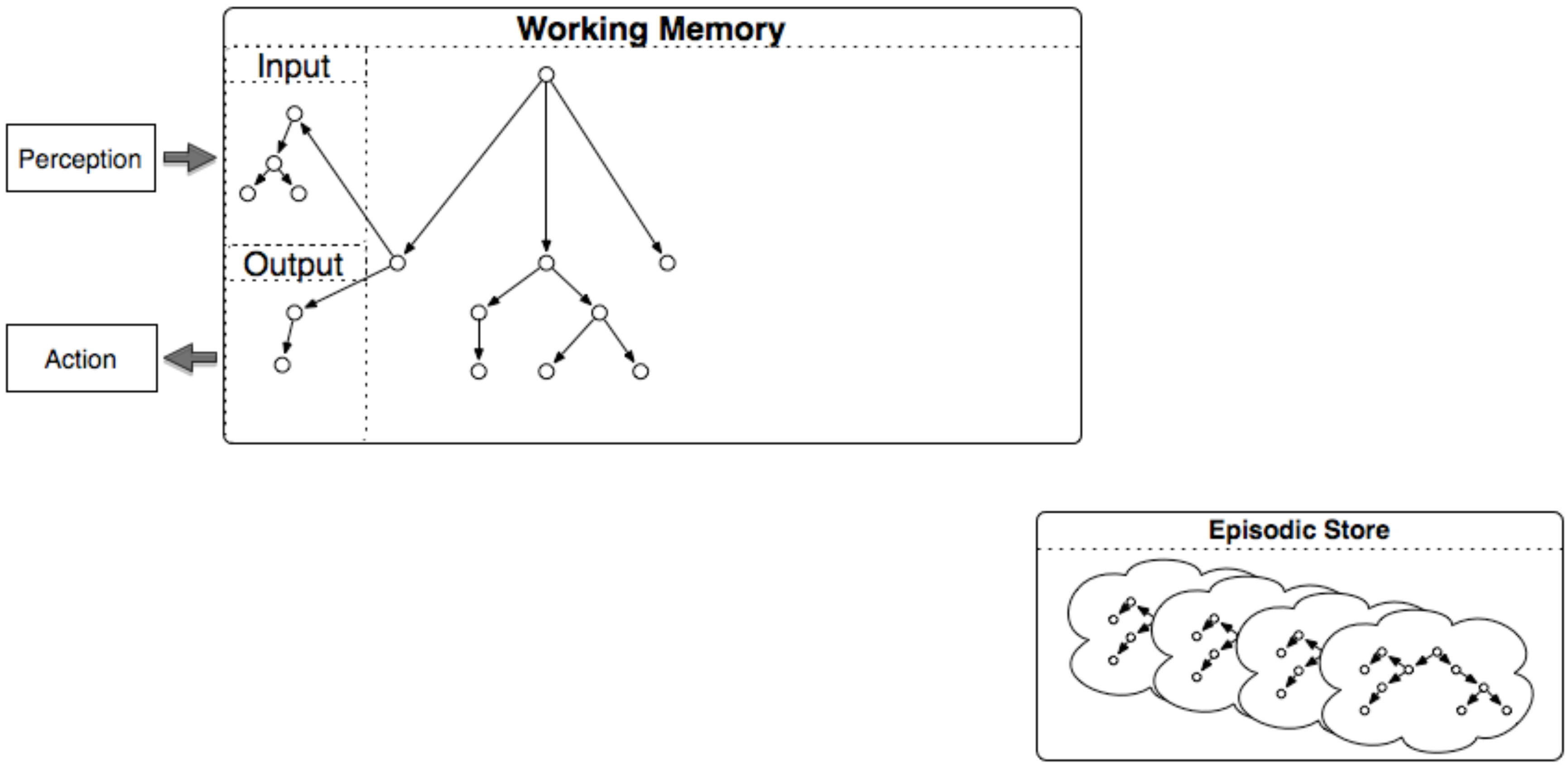
- Autobiographical
- Task-independent
- Automatic
- Autonoetic
- Temporally indexed

# EpMem: Functional Benefits

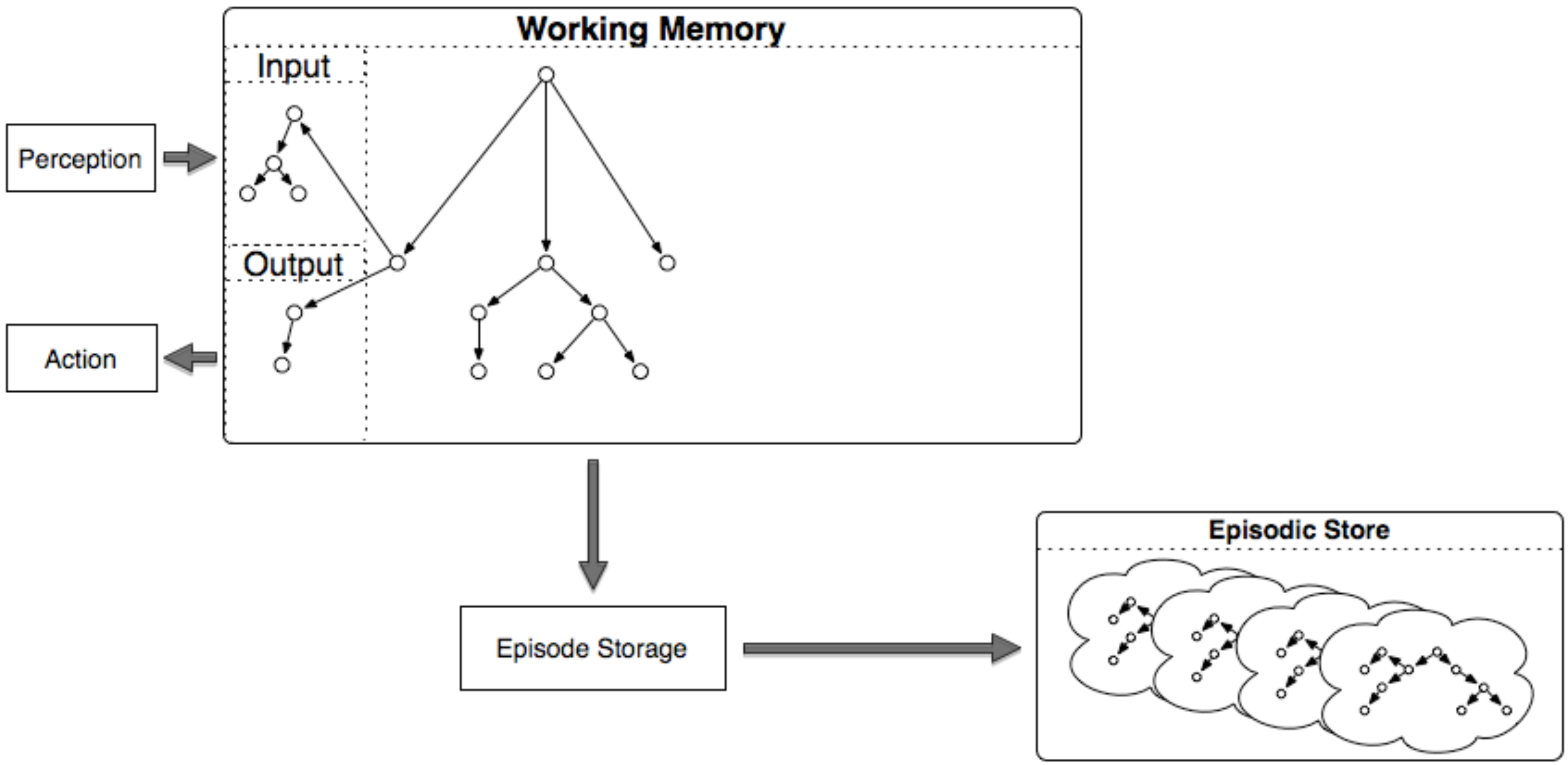
Supports enhanced situational awareness, reasoning, and learning via numerous general capabilities

- Virtual Sensing [Nuxoll & Laird, 2007]
  - Expands agent sensing beyond immediate perception via access to details of past situations
- Action Modeling [Laird et al., 2010; Xu & Laird, 2010]
  - Informs predictions about the result of actions in present or future situations based upon prior experience

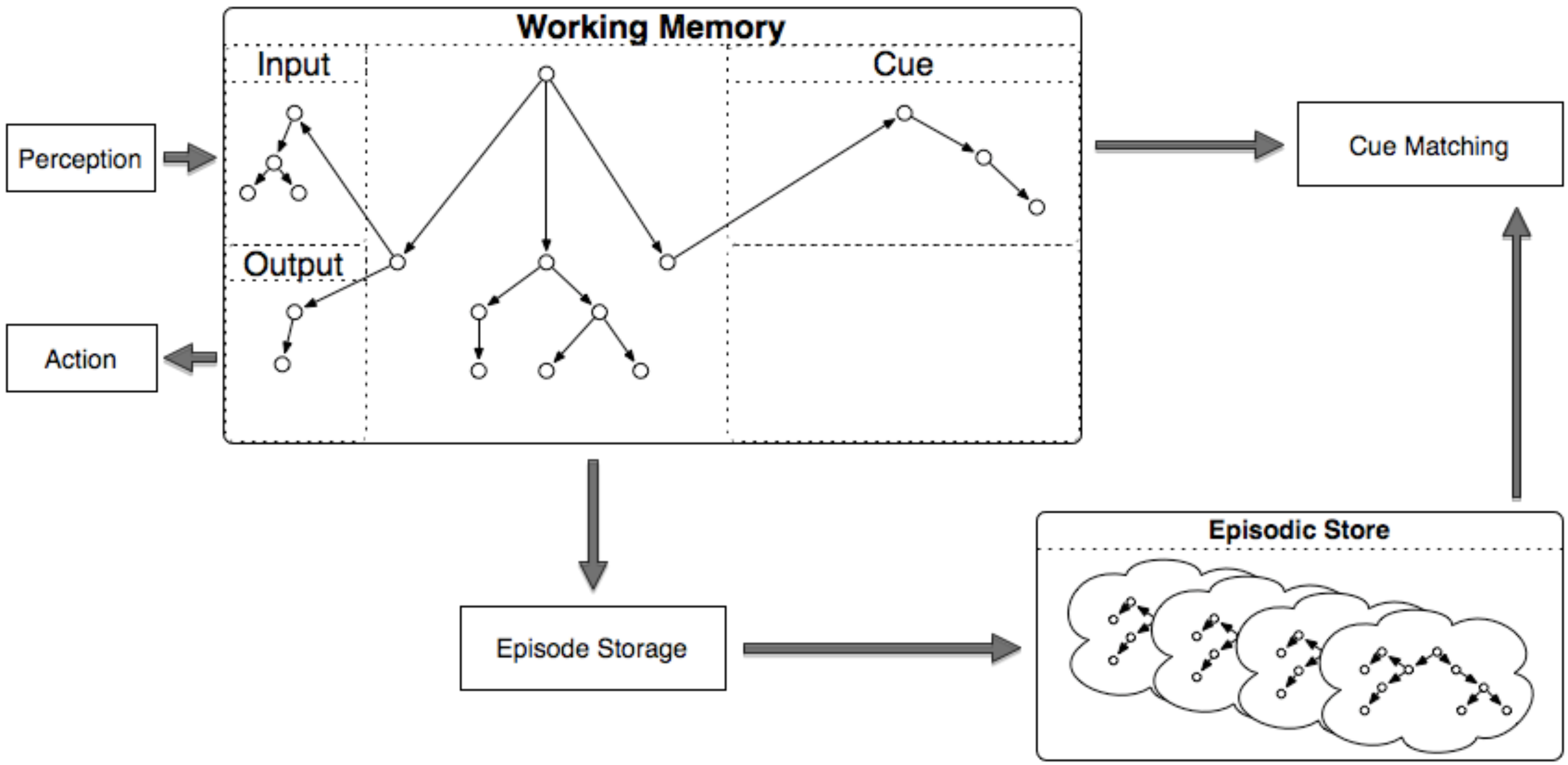
# EpMem: Architectural Integration



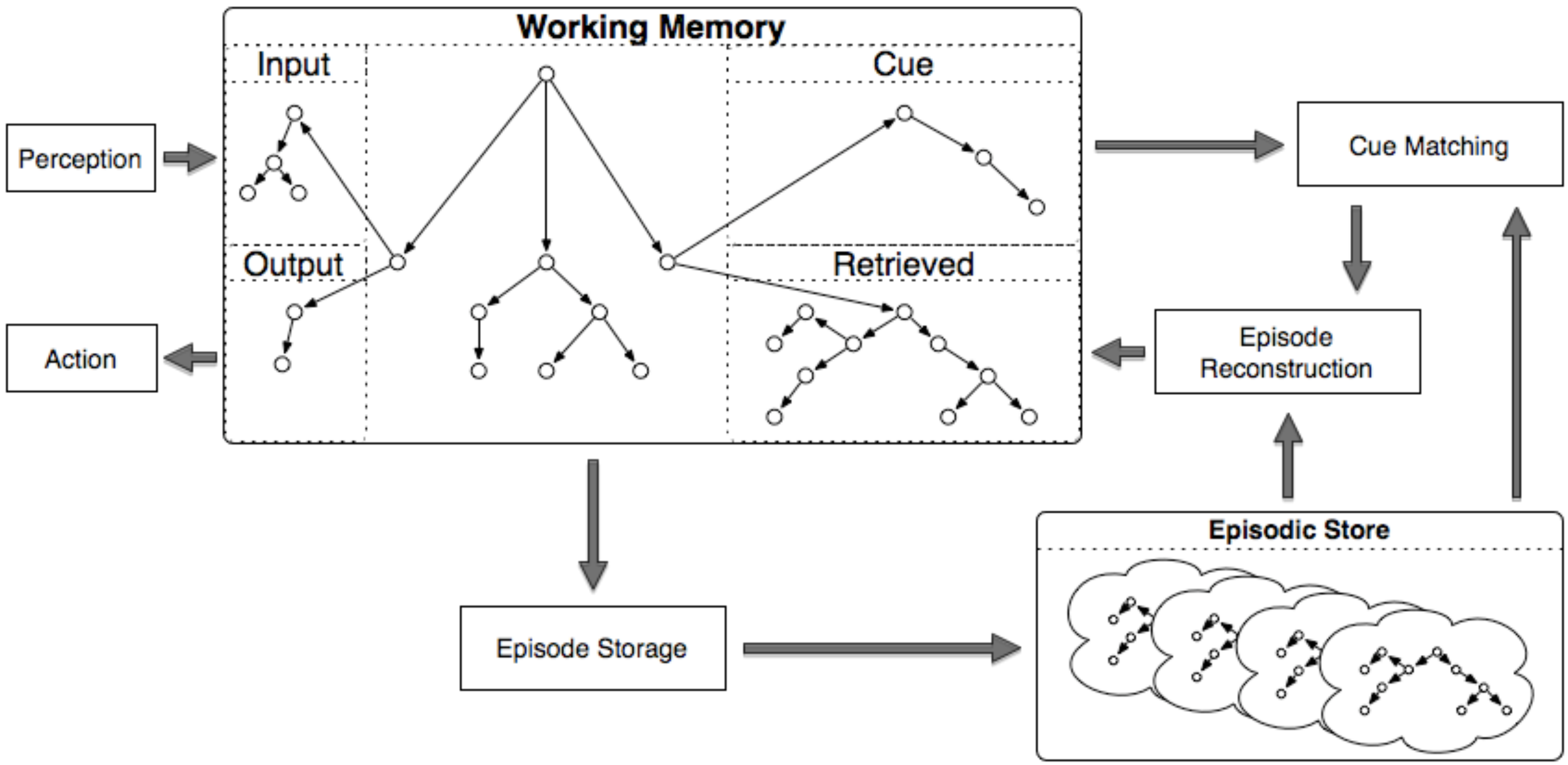
# EpMem: Architectural Integration



# EpMem: Architectural Integration



# EpMem: Architectural Integration



# EpMem: Contributions

Developed novel data structures and algorithms to support basic functionality in graph-based, task-independent episodic memory systems

[Derbinsky & Laird, 2009]

Faithful storage and reconstruction of episodes

- Exploited structural re-use & temporal contiguity

Cue matching

- Qualitative nearest-neighbor, biased by recency
- Two-stage matching strategy

# EpMem: Scaling Evaluation

Stressful domain

> 2500 features

~ 70-90% inputs change (30-100)

**1 million** episodes (~hours-days of real-time)

- 10 trials
- Commodity hardware

| Storage | Cue Matching* | Reconstruction** | Total |
|---|---|---|---|
| 2.68ms<br>625-1620MB<br>(0.64-1.66KB/ep) | 57.6ms | 22.65ms | **82.93ms** |

* 15 cues

** 50 random episodes

# Semantic Memory in Humans

Long-term store of facts independent of original context

- What you "know" vs. what you "remember"

Computational models account for human performance in numerous activities

- Categorization, task switching, linguistics, …

# SMem: Functional Benefits

Supports enhanced situational awareness, reasoning, and learning via access to large stores of general knowledge about the world

| | |
|---|---|
| **Lexical** | Word meanings, synonyms, … |
| **Mathematical** | Arithmetic facts, function/relation definitions, …. |
| **Geographical** | Capitals, bodies of water, … |
| **Historical** | Wars, discoveries, reigns of power, … |
| **Ontological** | Biology, technology, art, … |
| **Commonsensical** | “Tables typically have four legs” |
| … | |

# SMem: Large Stores

Complex tasks require access to large stores of knowledge

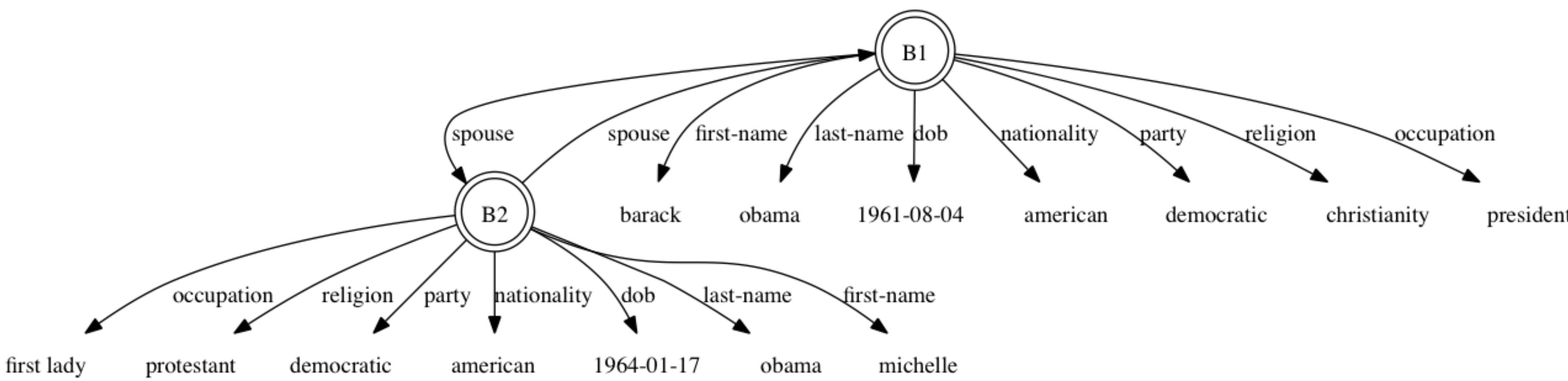
| SUMO | WordNet | Cyc |
|---|---|---|
| Ontology | Lexicon | "Common Sense" |
| 4.5K classes<br>250K facts | 212K word senses<br>820K facts | 500K concepts<br>5M facts |

# SMem: Architectural Integration

Hierarchical store of concepts and associated features

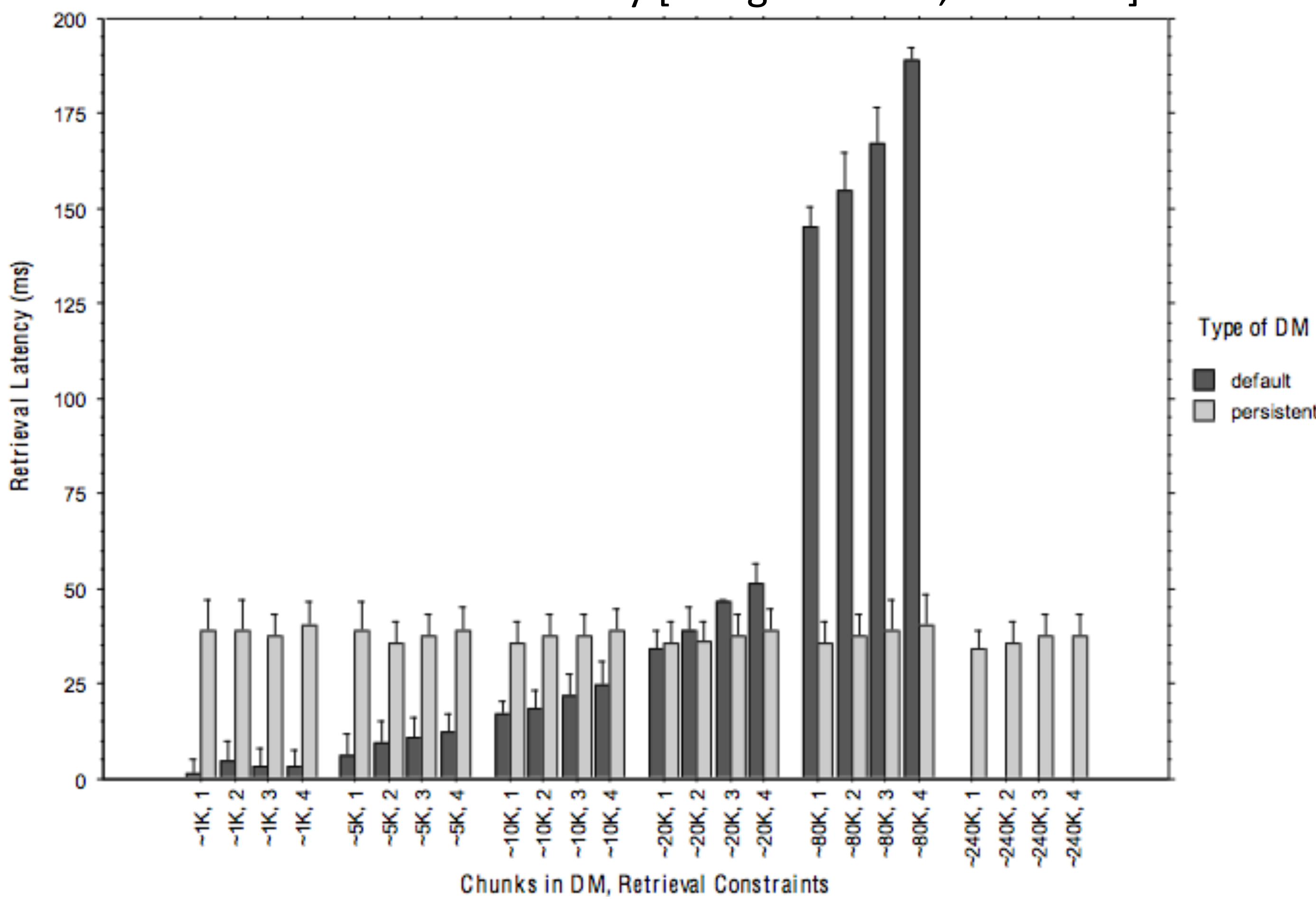
Supports…

- Associative cue: given feature subset, retrieve node
- Concept expansion: given node, retrieve features



# SMem: Related Work (1)

ACT-R DM Retrieval Latency [Douglass et al., ICCM '09]



# SMem: Related Work (2)

| | Douglass et al., '09 |
|---|---|
| **Problem Formulation** | Empirical |
| **Methods & Analysis** | System Dependent |
| **Implementation** | PostgreSQL+ACT-R |
| **Matching** | Symbolic* |
| **Evaluation** | WN-Lexical*, 240K chunks |

# SMem: Contributions

| | Douglass et al., '09 | Derbinsky et al., '10 |
|---|---|---|
| **Problem Formulation** | Empirical | Empirical<br>Formal |
| **Methods & Analysis** | System Dependent | System Independent |
| **Implementation** | PostgreSQL+ACT-R | SQLite+Soar |
| **Matching** | Symbolic* | Symbolic |
| **Evaluation** | WN-Lexical*, 240K chunks | WN-Lexical, 820K chunks<br>Synthetic, 3.6M chunks |

**100x faster retrievals on a comparable set of cues**
**scaling to a 3x larger semantic store**

# Outline

Introduction

Prior Work
- Episodic
- Semantic

Future Work
- Evaluation Strategy
- Proposed Extensions
- Timeline

# Future Work

Prior work focused on understand the <u>efficiency</u> challenges in extending the Soar cognitive architecture with <u>basic</u> episodic and semantic memory functionality

We now propose to extend functionality, as guided by our core research questions, while maintaining efficiency

# Q1: Encoding/Storage

What aspects of agent experience should a task-independent memory system *encode* and *store* …

such as to functionally support performance across a variety of tasks…

while maintaining reactivity in complex, dynamic environments?

# Q2: Retrieval

*Max left his iPhone at the **bank***


What task-independent regularities of agent experience…

can efficiently supplement impoverished cues…

to improve the expected utility of *retrieved* memories?

# Future Work: Extensions

| | Episodic | Semantic |
|---|---|---|
| **Q1: Encoding/Storage** | X1 | X3 |
| **Q2: Retrieval** | X2 | X4 |

# Evaluation

No accepted benchmarks or metrics for comparing task-independent memory systems for generally intelligent agents

**Strategy**

- Focused computational benchmarks
- Apposite model comparisons
- Thematic complex domain: cognitive robotics

**Metrics**

- Computational
  - Space, time
  - Analytical, empirical
  - Maximum, average
- Task Performance
  - Quality
  - Time-to-completion

# X1: Episodic Encoding

**Problem**

Autonomous agents are exposed to large amounts of information

- Experiential
- Conceptual

How can episodic memory improve space requirements over long lifetimes while maintaining useful retrievals?

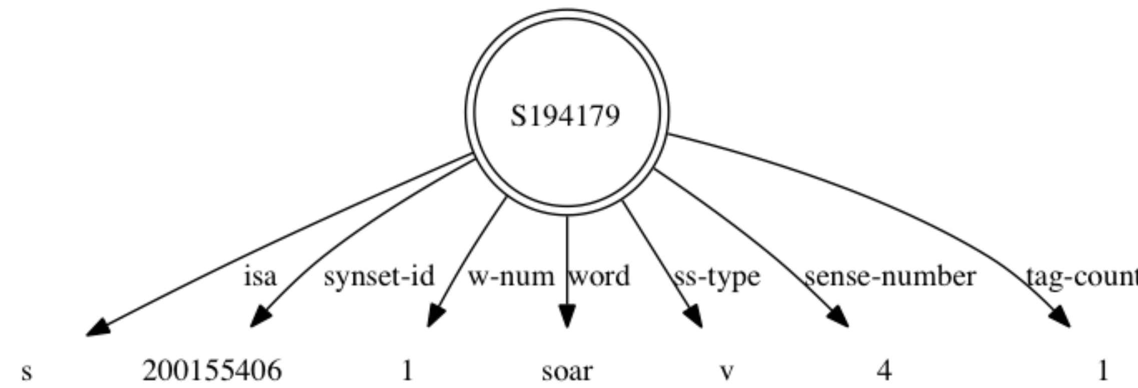
**Approach**

Explore tradeoffs in policy of not encoding in episodic memory the substructure and relations of semantic concepts

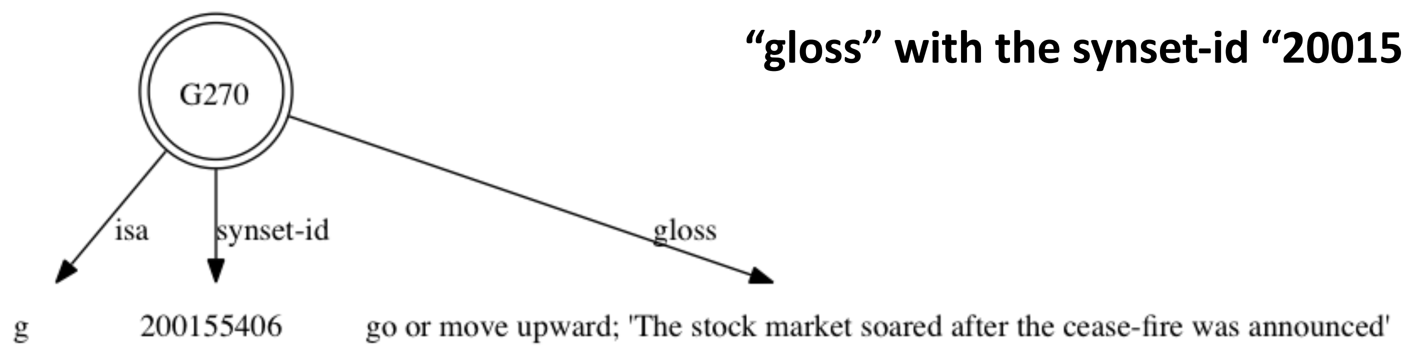
- Reduces storage
- Sacrifices fidelity
- May require semantic retrieval to reconstruct episodes

# Example Semantic Knowledge: "soar"

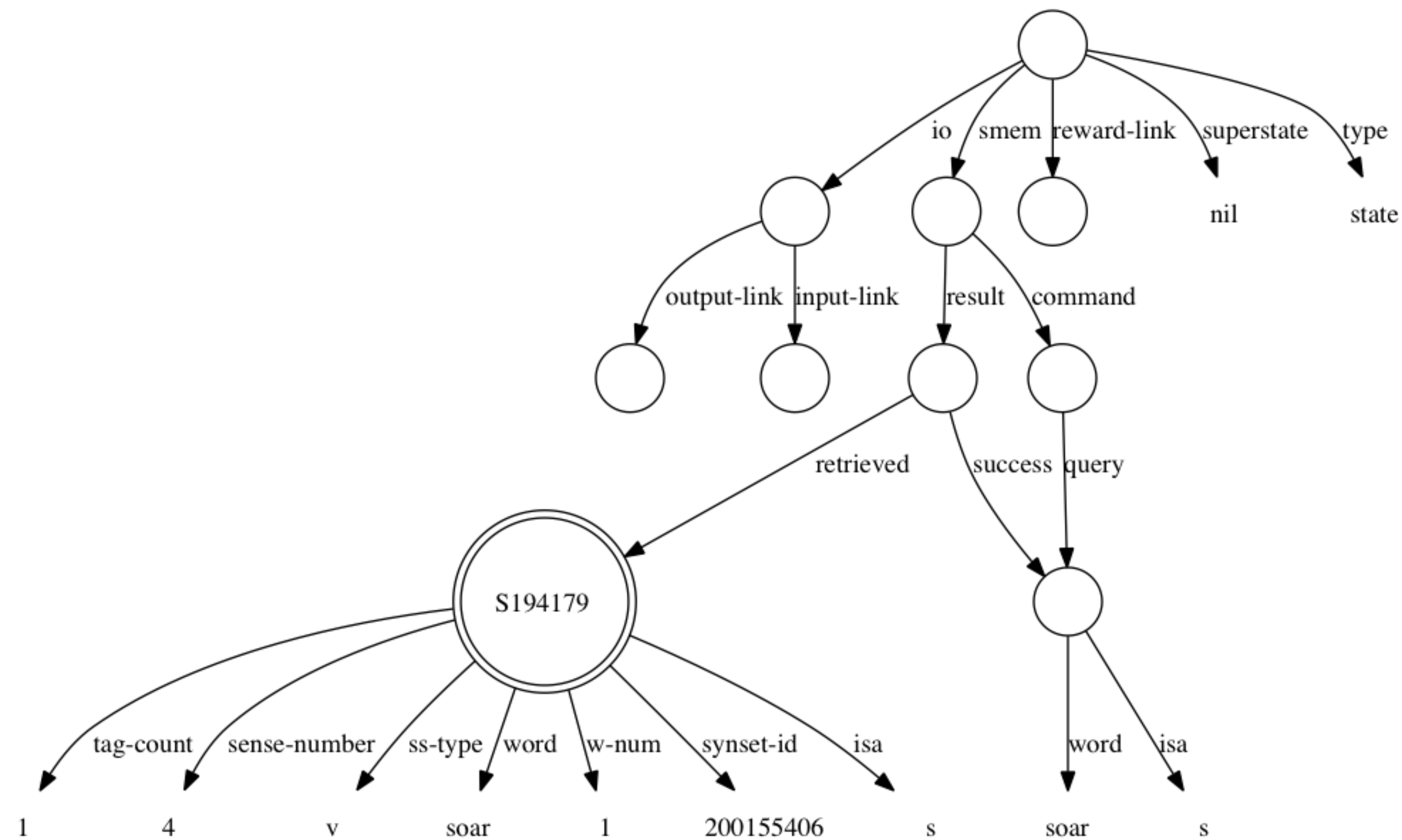
**"sense" of the word "soar"**




**"gloss" with the synset-id "200155406"**

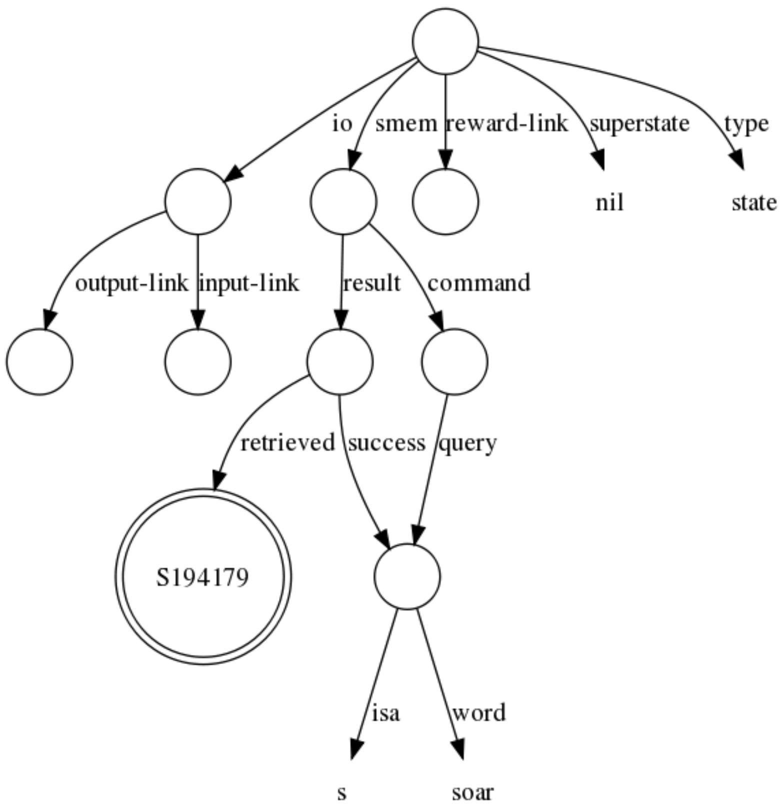


# Example Episode

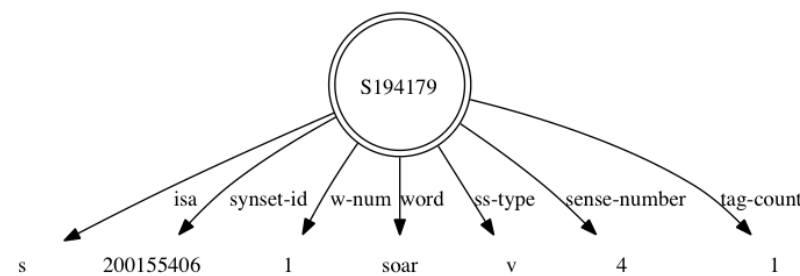


# Pruned Episodic Encoding

**Episodic**




**Semantic**



# X2: Episodic Retrieval

**Problem**

Current match metrics

- Cue element cardinality
- Recency

Given an under-specified query, can additional sources of knowledge improve retrieval quality while scaling to large bodies of knowledge?

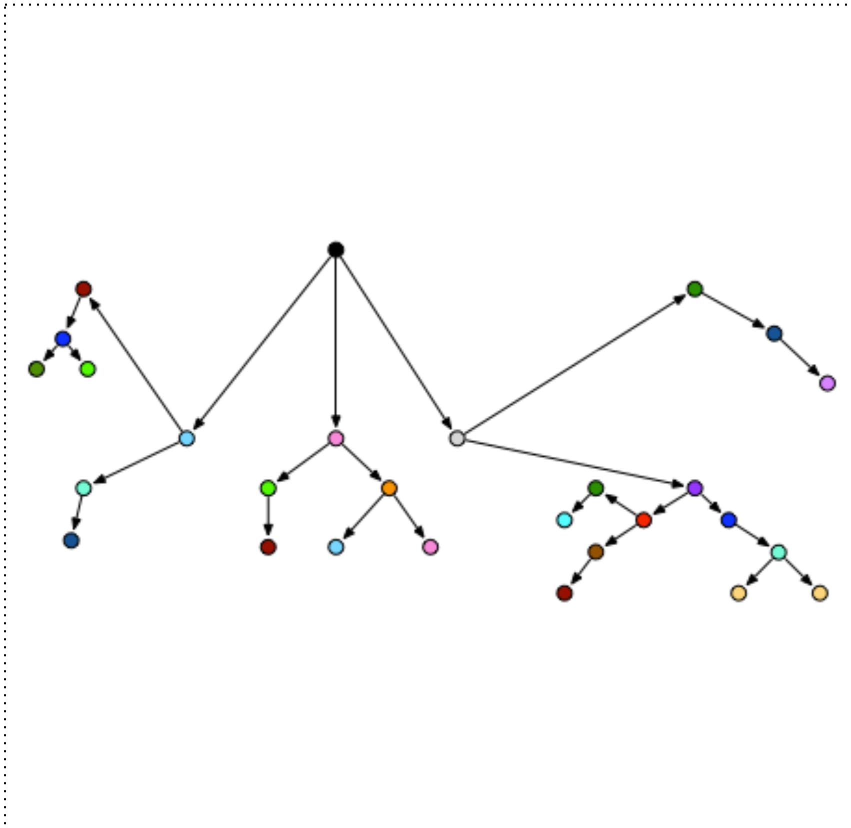
**Approach**

Explore methods of *efficiently* incorporating bias at two granularities

- Episode
- Element

# Episodic Bias Granularity

**Element**



**Episode**

# X3: Semantic Encoding

**Problem**

Given conceptual information about the world, semantic retrievals support numerous cognitive functions

Linguistics, communication, inference, …

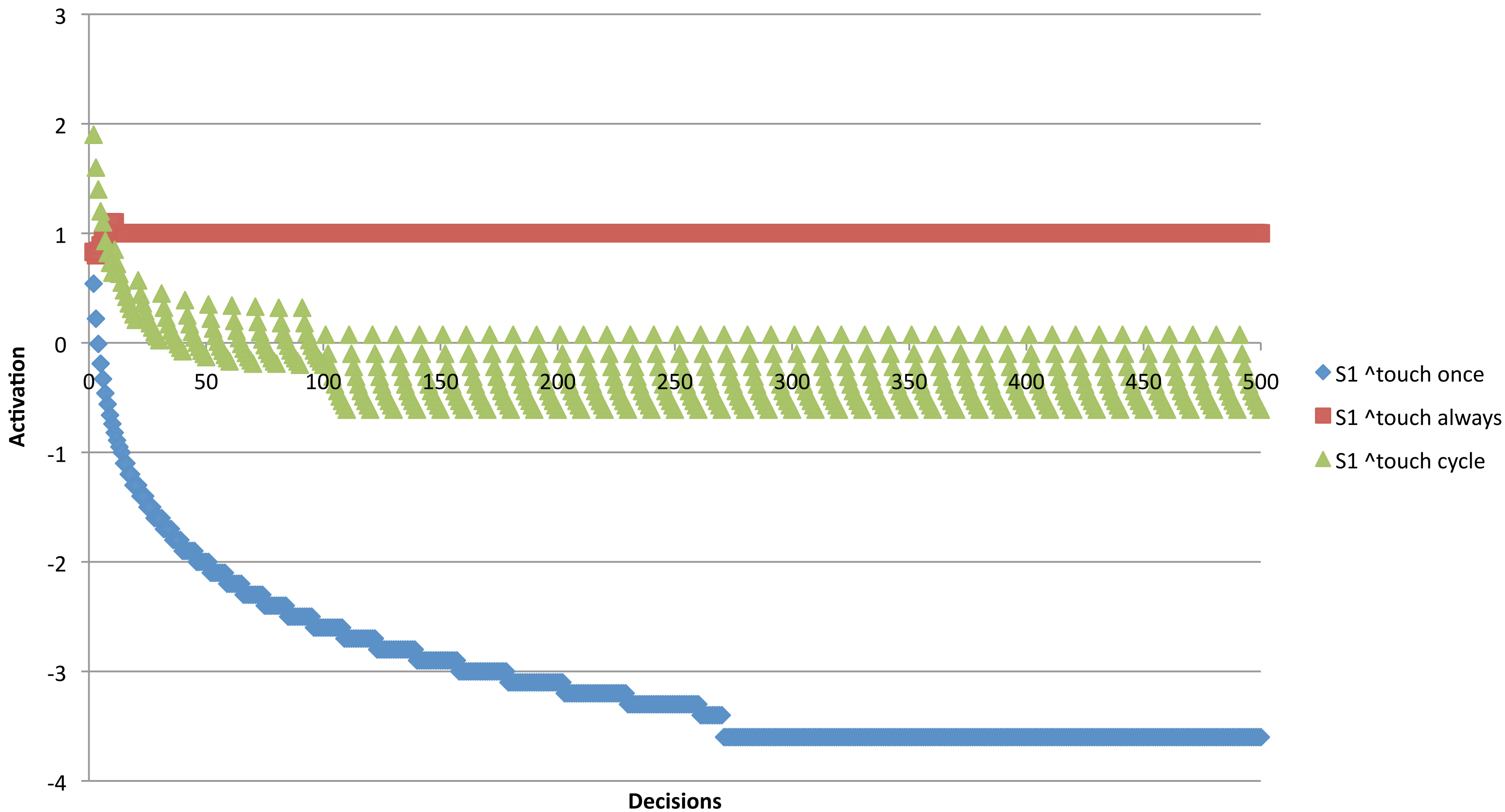
How should a generally intelligent agent incrementally acquire this knowledge over a long lifetime?

**Approach**

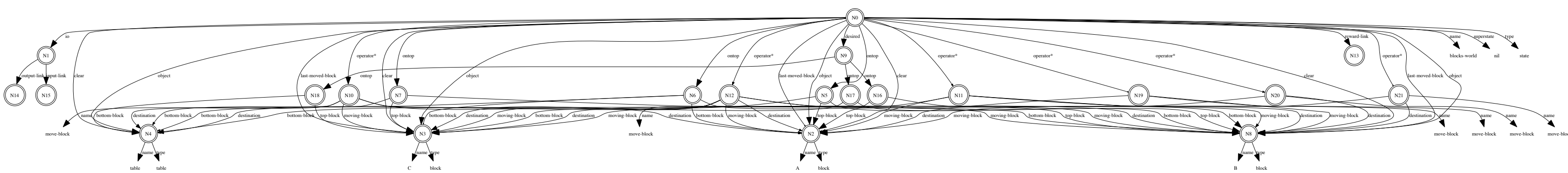
Evaluate a small space of automatic encoding policies, spanning two pruning heuristics

- Situational focus
- Structural stability

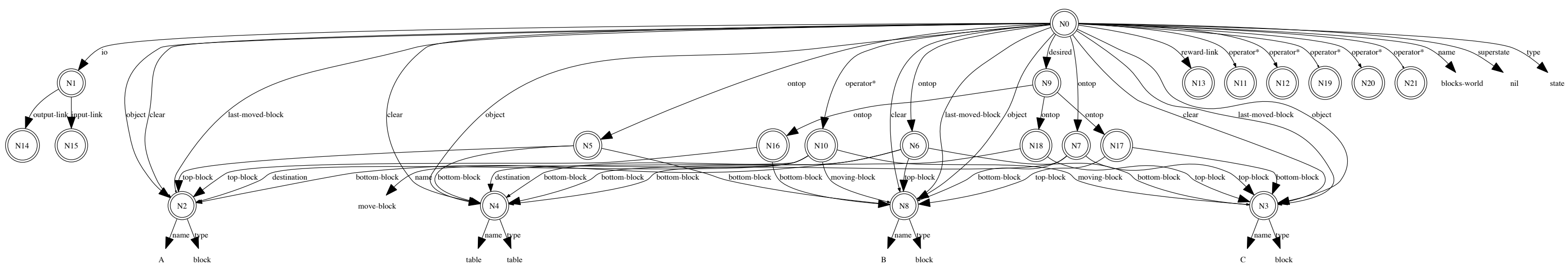
# Situational Focus: Working Memory Activation



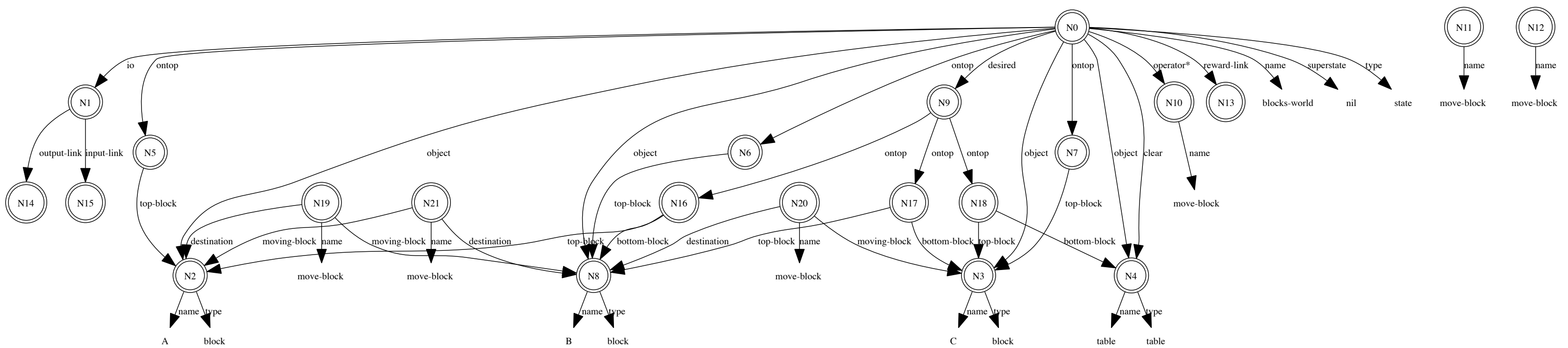
# Structural Stability (1)



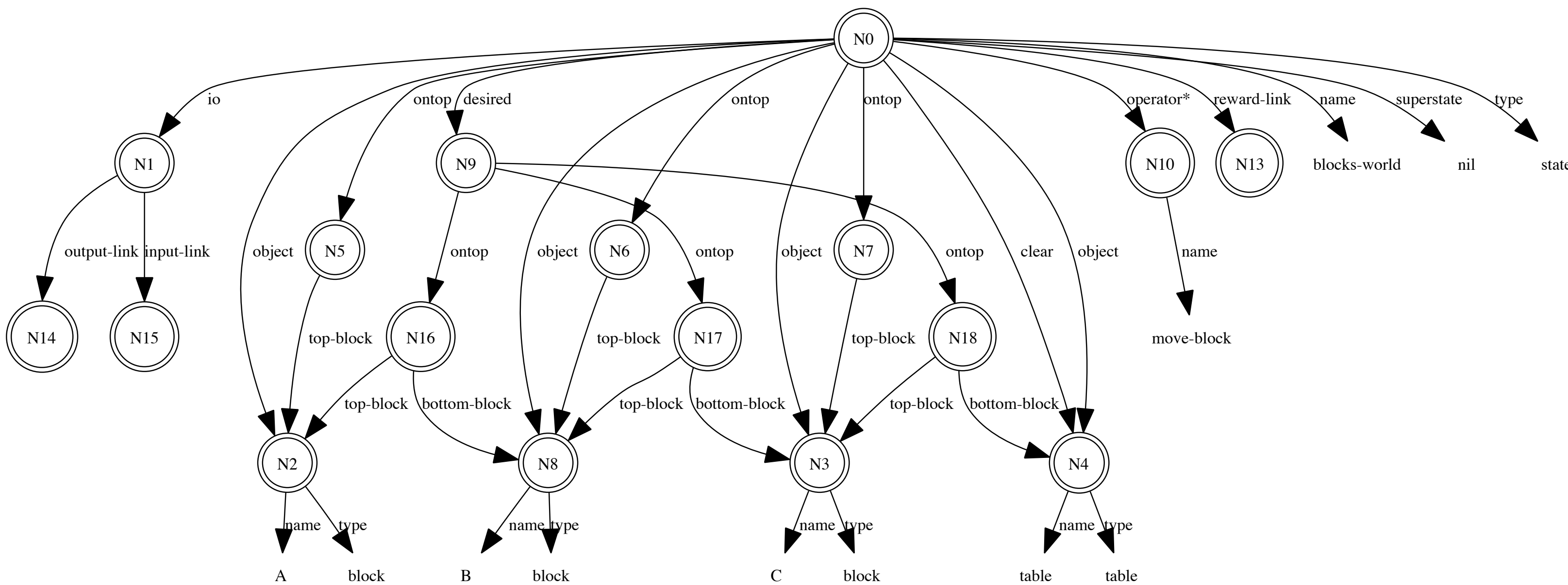
# Structural Stability (2)



# Structural Stability (3)



# Structural Stability (4)



# X4: Semantic Retrieval

**Problem**

For underspecified cues, rational analysis (Anderson, 1990) suggests semantic memory must be sensitive to statistical environmental regularities

- Retrieval history
- Context

Current models cannot scale to even moderate sized knowledge stores

**Approach**

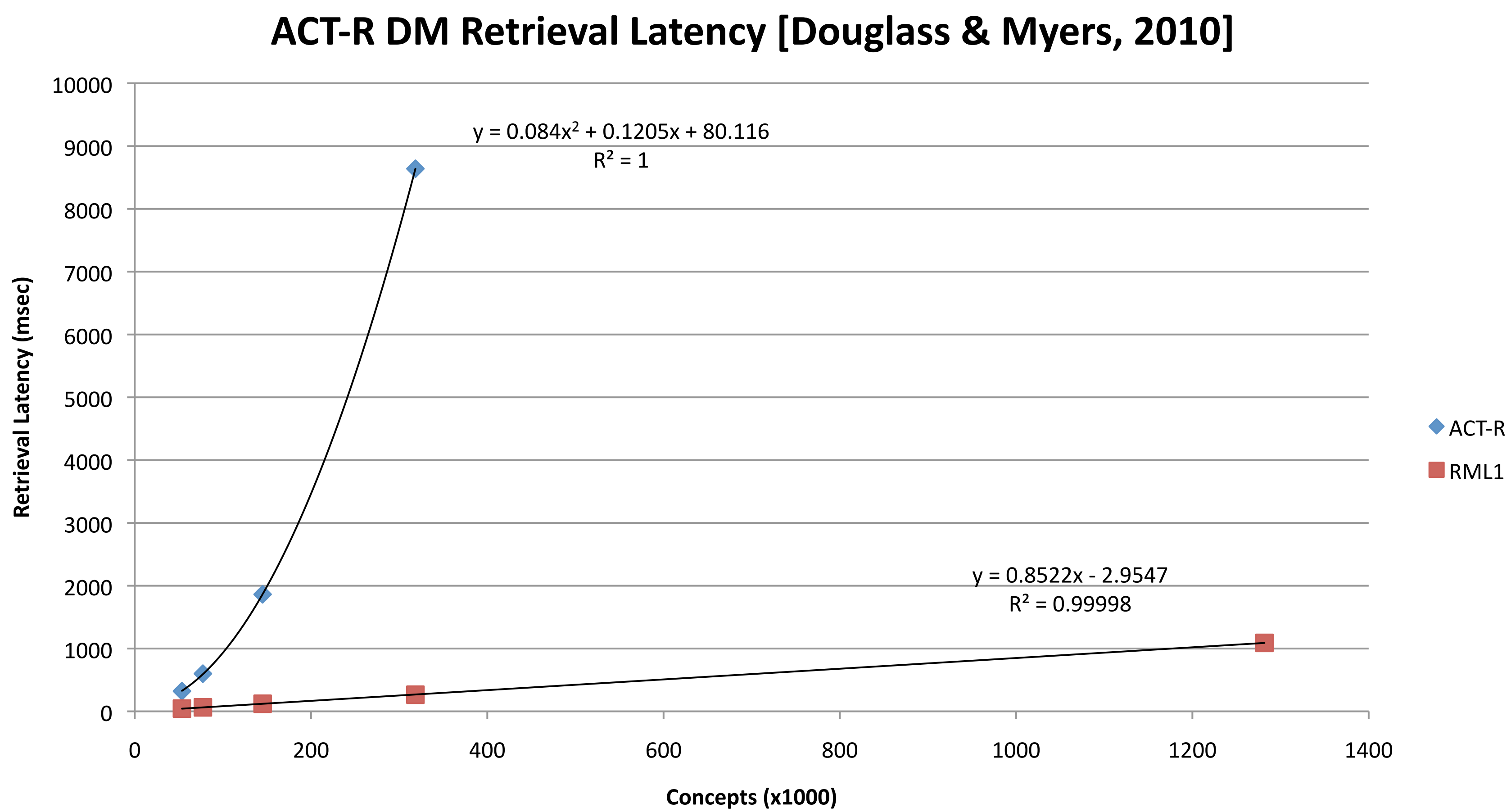
Implement and evaluate parallel and approximate forms of activation bias

Evaluate Word Sense Disambiguation model degradation as focused task

# Semantic Bias

*Max left his iPhone at*
*the dock near* ***the bank***

# Semantic Bias Latency



# Future Work: Expected Contributions

| | Incremental Learning | **Comprehensive Learning** | Diverse Representation | **Scale Efficiently** | Effective Access | Task Independence |
|---|---|---|---|---|---|---|
| Cognitive Architecture | ☑ | ↑ | ☑ | ↑ | ☑ | ☑ |

Mechanism space
- Implemented software
- Data structures, algorithms
- Computational analysis

Functional agent demonstrations
- High-level cognitive capabilities
- Useful across numerous tasks, long lifetimes

Modeling constraint
- AGI Constraints -> LTM Requirements -> Scientific Exploration

# Future Work: Interactions

| | Episodic | Semantic |
|---|---|---|
| **Q1: Encoding/Storage** | X1 | X3 |
| **Q2: Retrieval** | X2 | X4 |

# Future Work: Timeline

**October 2010 – January 2011**

– Semantic Retrieval (X4)

**January 2011 – April 2011**

– Semantic Encoding (X3)

**April 2011 – September 2011**

– Episodic Encoding (X1)
– Episodic Retrieval (X2)

**October 2011 – March 2012**

– Thesis data analysis, writing, and defense

Questions?

# THANK YOU!