Cognitive Architecture

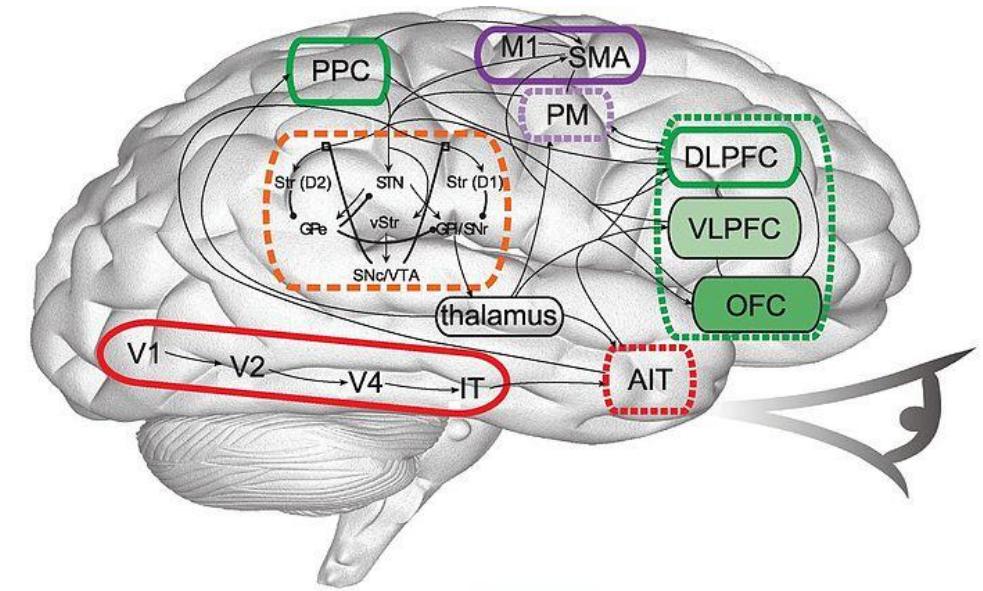
and Conversational Agents

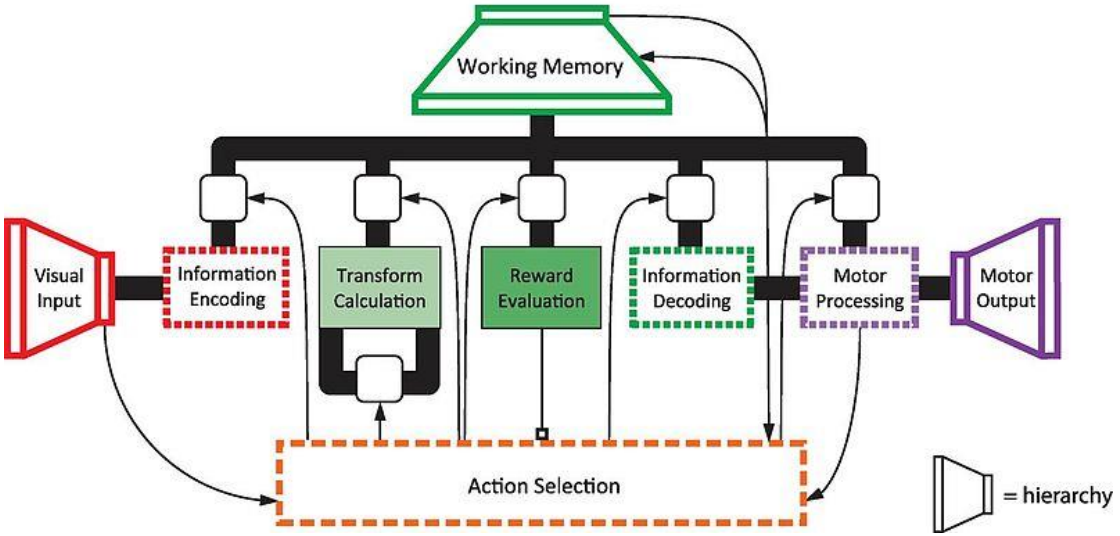
Eric Saund, Ph.D.
February, 2019

www.saund.org

Architecture

Brains and Minds





Cognitive Architecture: Outline

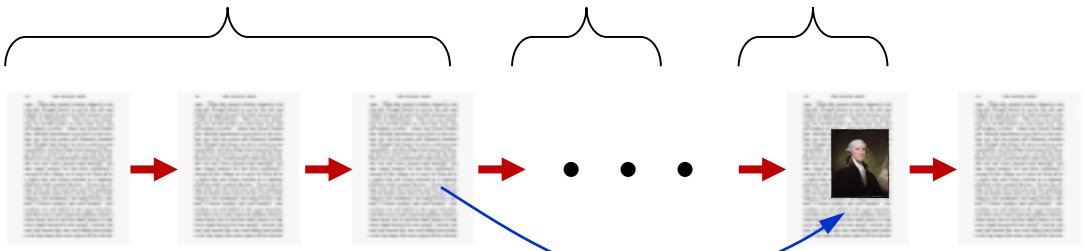
- Architecture in Information Systems
  - *"An organization of elements to achieve some information processing purpose."*
- History
  - Computational theory of mind
- The Baseline/Generic Cognitive Architecture
- Examples
  - Soar, LIDA, CopyCat
- Architecture in NN / Deep Learning Networks
- Conversational Agents
  - NLP
  - Dialog Management
  - Knowledge Graphs
  - Deep Networks

Information Architecture: A Book

Table of Contents

page numbers

logical sections



content Pages

cross-references

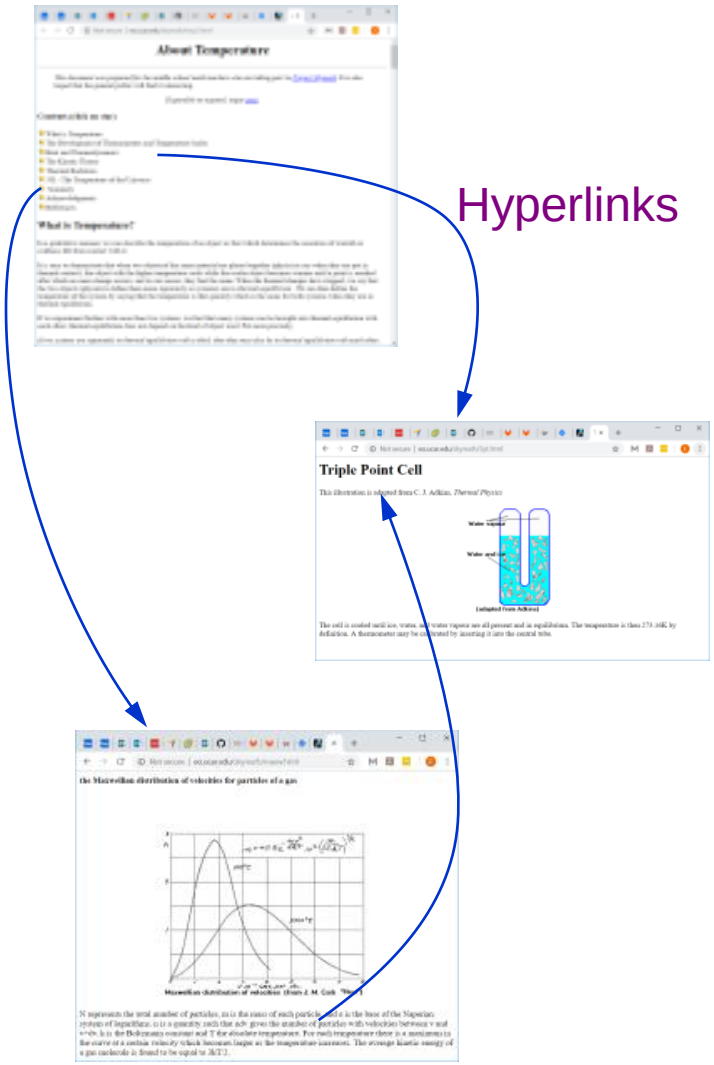
page numbers

content items

Index

Information Architecture: A Web Site

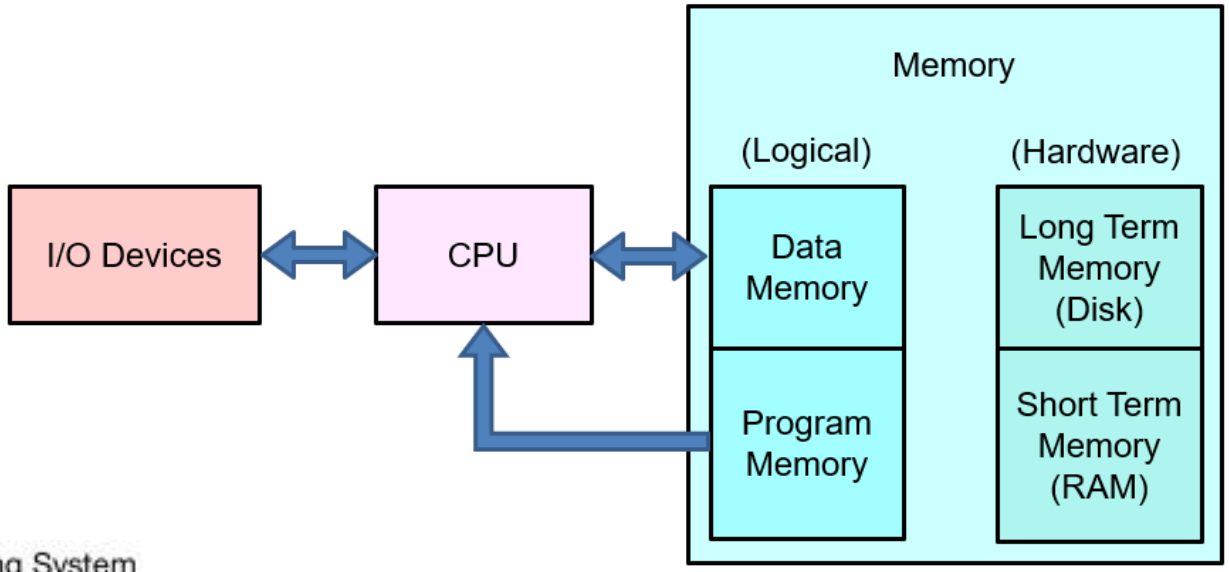
HTML1.0

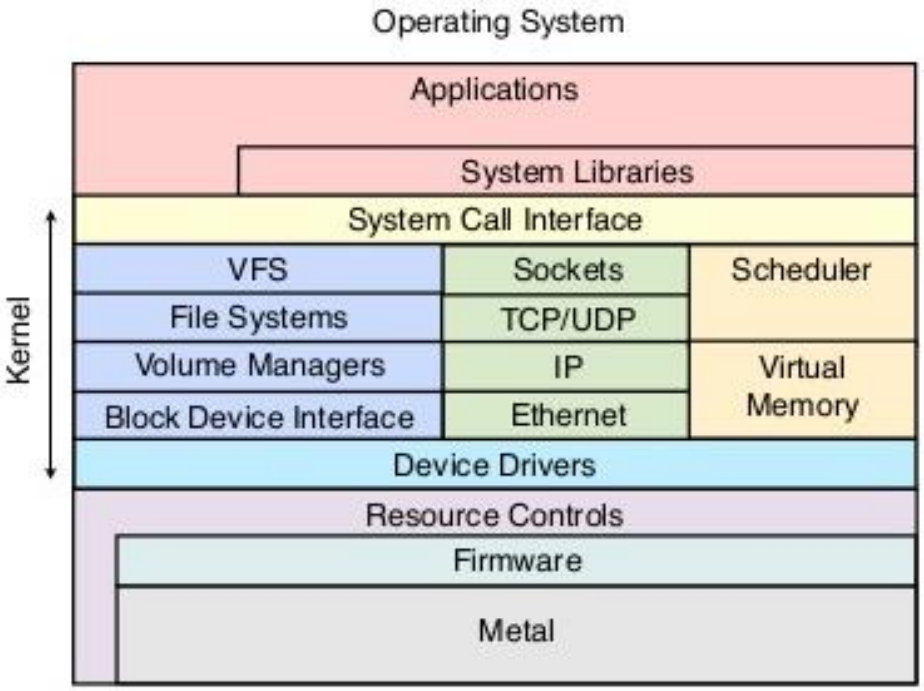


HTML5/CSS/Javascript

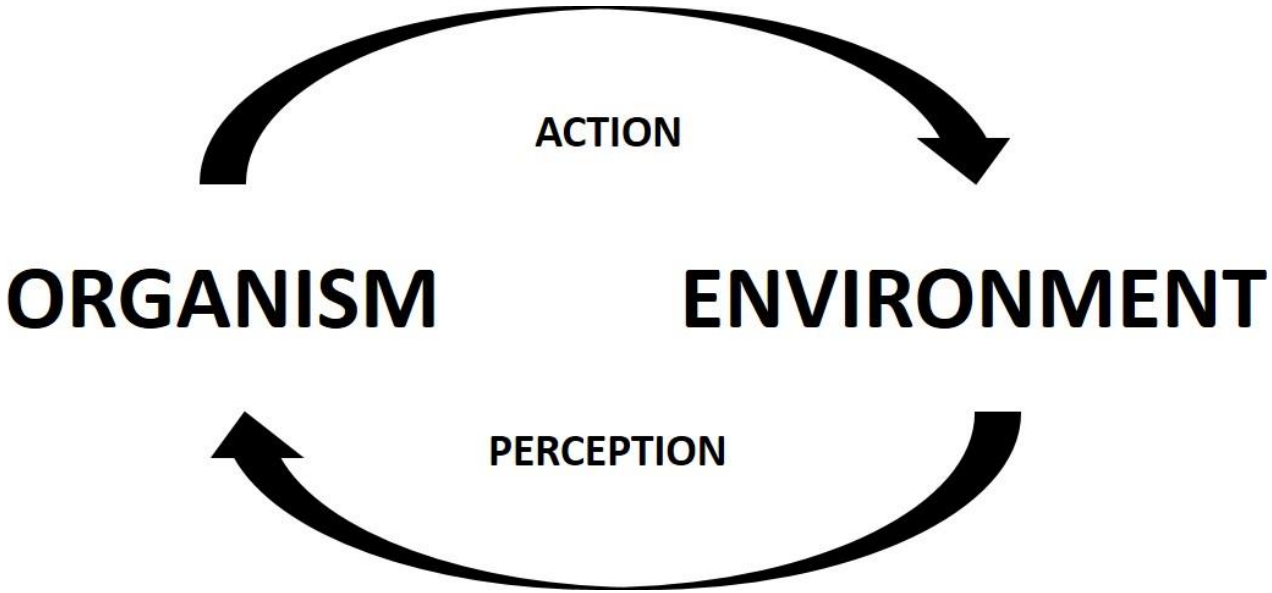
- active nav widgets
- guided tours
- search
- state
  - cookies
  - navigation history
- contingent access (passwords)
- personalization

Information Architecture: A Computer



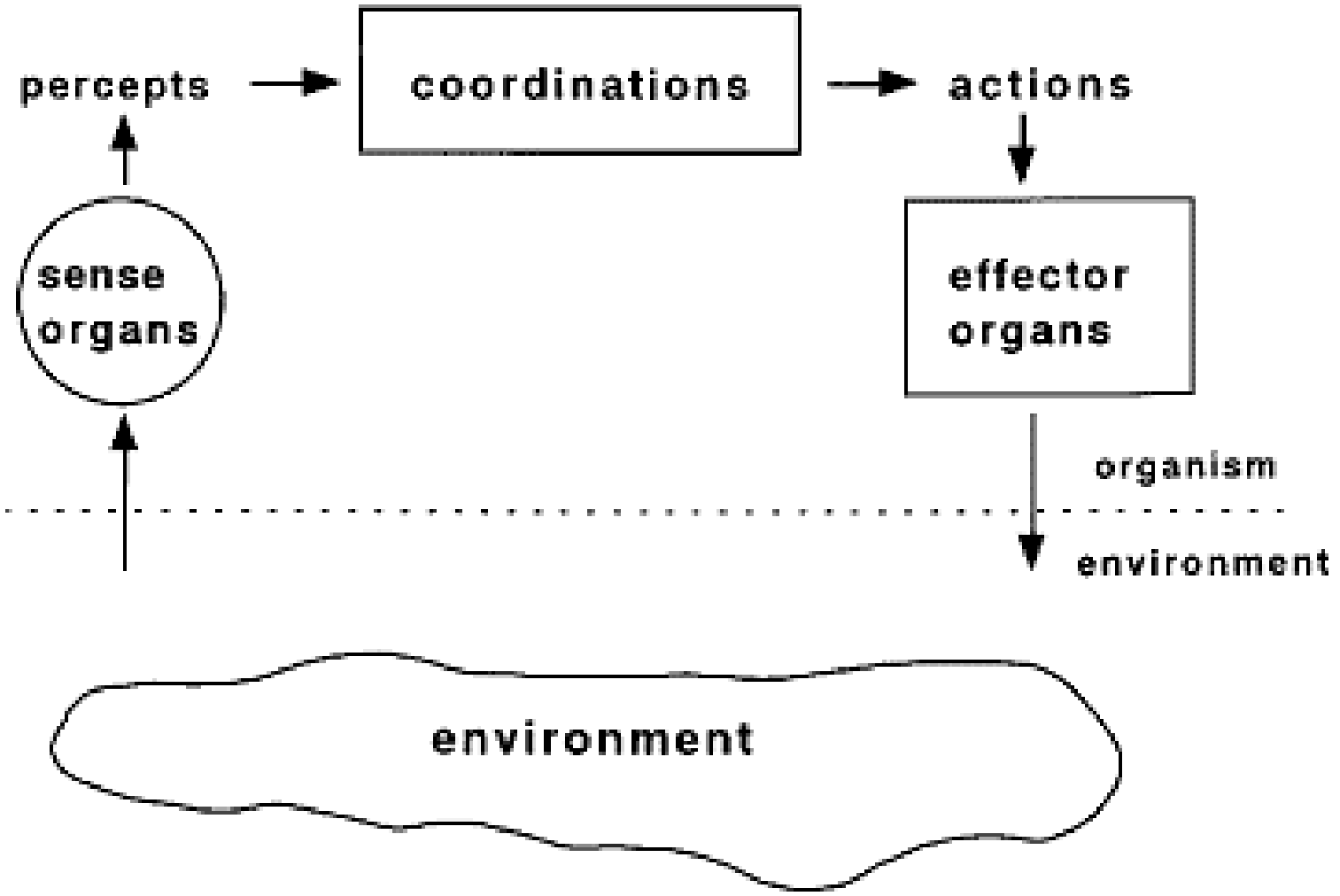


Cognitive Architecture: Basic Agent



http://psychsciencenotes.blogspot.com/2018/07/you-cannot-perceive-relational.html

Cognitive Architecture: Basic Agent



Cognitive Architecture: Historical Roots

Psychoanalytic Theories
Freud, Jung

Perceptual &
Phenomenal Psychology
Helmholtz, William James

Cybernetics
Norbert Wiener

Artificial Intelligence
McCarthy, Minsky

Computational Theory of Mind
McCulloch and Pitts

Chomsky

Newell & Simon

Theory of Computation
Turing, von Neumann

Behaviorism
Pavlov, B.F. Skinner

Guiding Metaphor

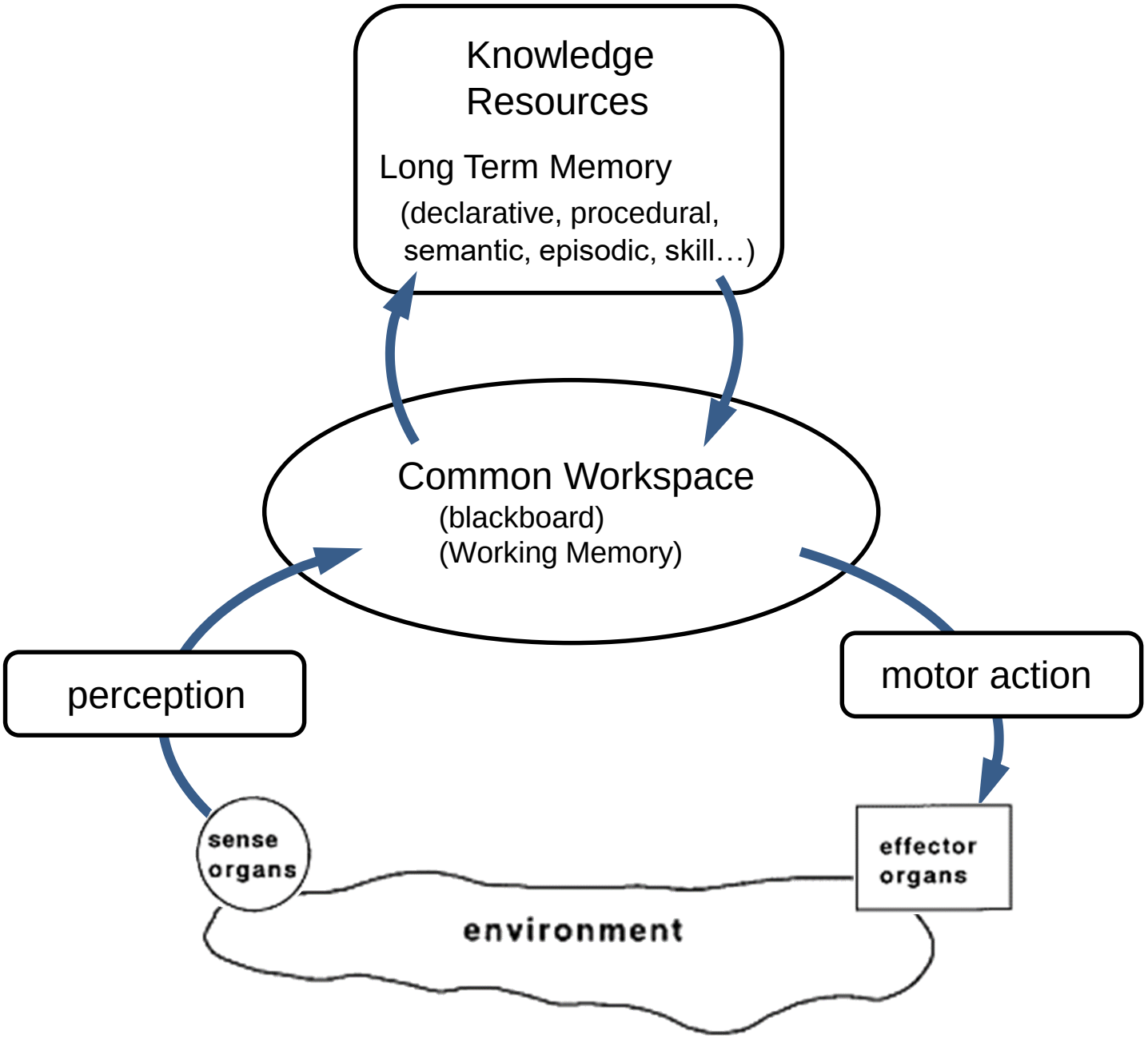
- engines and hydraulics
- signal transmission

- signal processing
- calculating machines

- computers

https://plato.stanford.edu/entries/computational-mind/

Cognitive Architecture: Standard Model



Cognitive Architecture: Big Questions

- What are the *types of content* held in the workspace?
  - percepts
  - beliefs
  - memories
  - goals
  - intentions & plans
  - emotions
  - attitudes
- What are the *representations* for state and knowledge?
  - activation patterns over fixed vectors
  - graphs of objects and relations
  - frequencies and phases of waveforms
- How is processing *controlled*?
  - automatic processes
  - conscious deliberation
  - selection of operators

Soar

(Newell, Laird,
1983 -> present)

- Definition of intelligence:
  - problem states and transitions
  - solutions found through *search* in state space
- Representation:
  - graphs of objects and relations
- Control: production system
  - Working Memory blackboard
  - procedural knowledge
  - declarative knowledge

} Long-Term Memory

Soar: Water Jug Problem Example

Definition of intelligence:

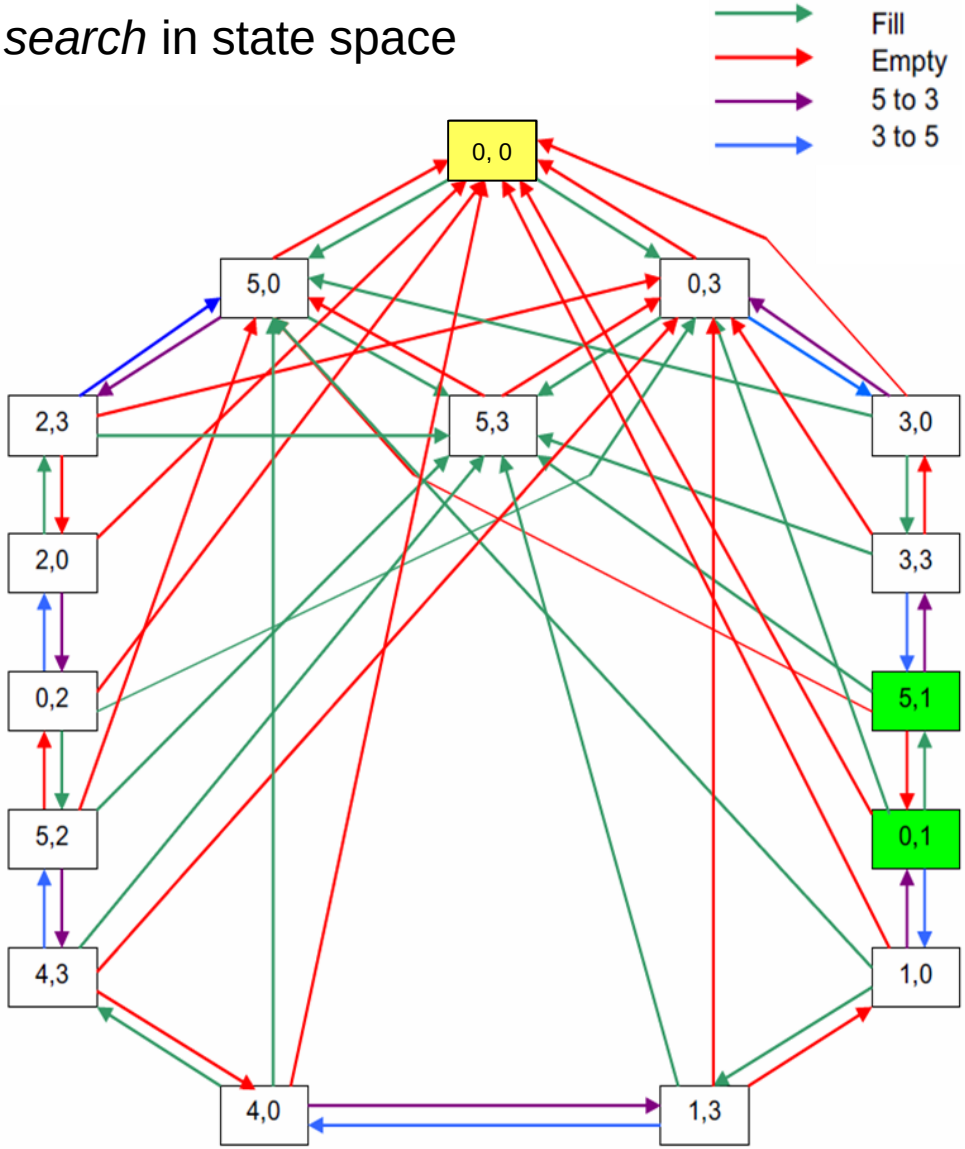
- problem states and transitions
- solutions found through *search* in state space

5 gal.

3 gal.

Start state: both jugs empty.

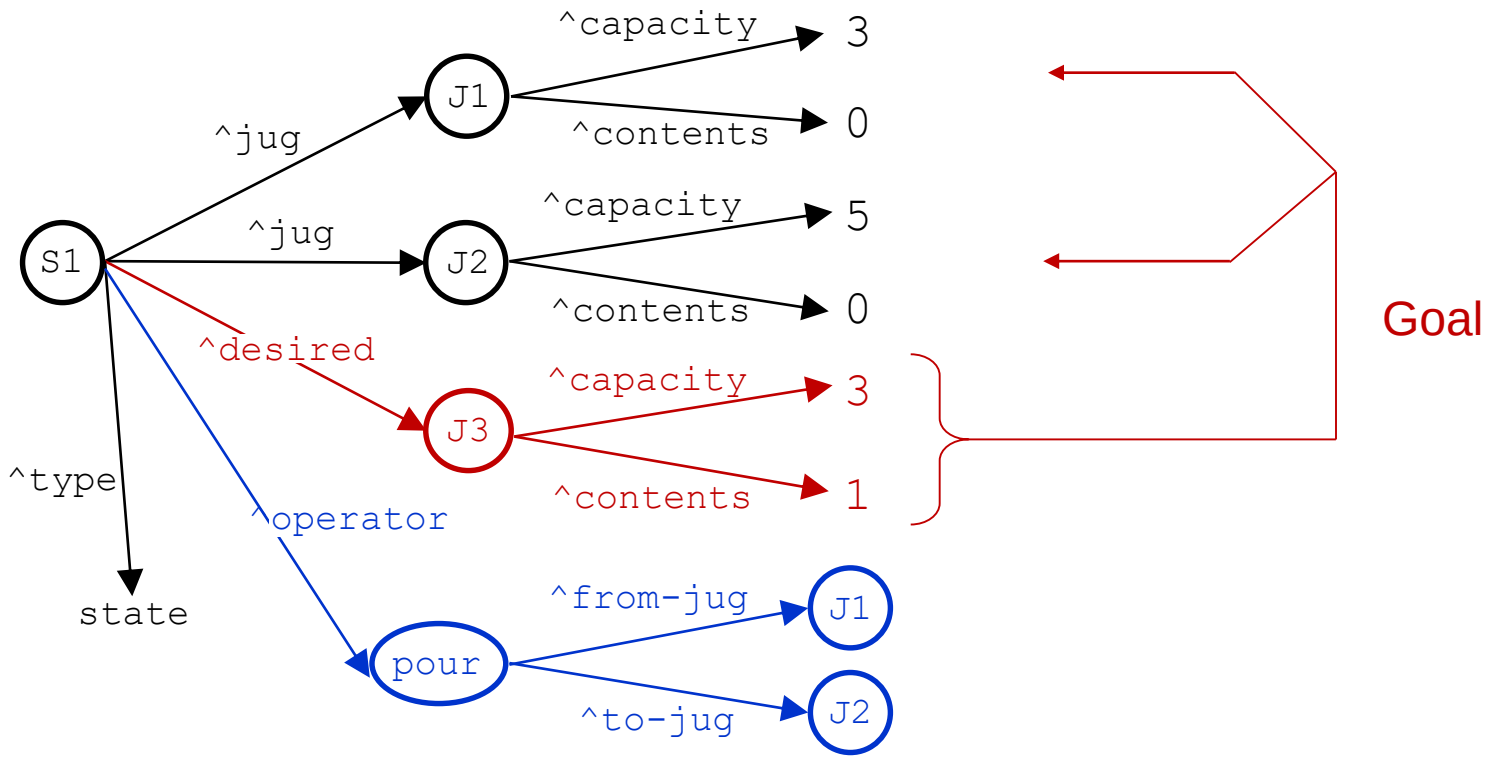
Goal state: 3-gallon jug contains 1 gallon of water.



Representation in Soar

Graph

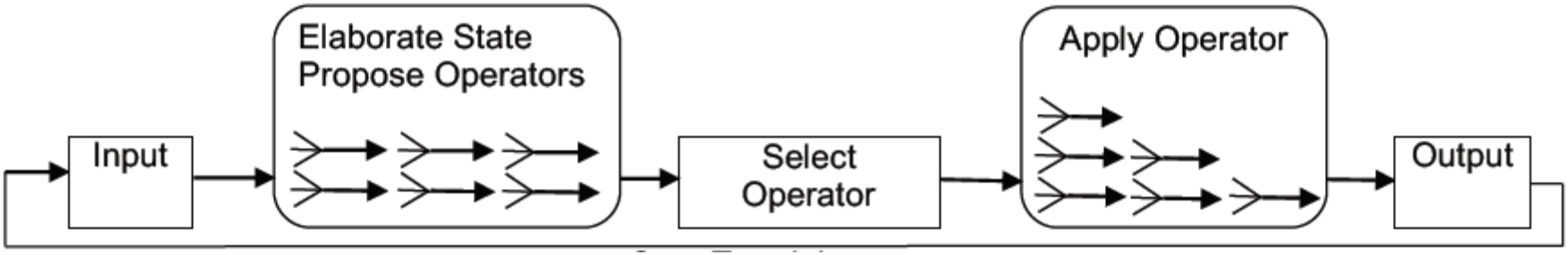
- data objects
- attributes & relations
- operators
- Working Memory (state)
- Long-Term Memory (knowledge)



Control in Soar

Production System

- Working Memory blackboard
- declarative knowledge - what
- procedural knowledge – how
  - rules
  - operators
- subgoal states



Processing Cycle

Executive Function (Psychology, Cognitive Neuroscience):

- update Working Memory from sensory and Long-Term Memory resources
- focus attention, inhibit distractors
- shift task context

Cognitive Architecture: Outline

- ✓ Architecture in Information Systems
  - "An organization of elements to achieve some information processing purpose."
- ✓ History
  - Computational theory of mind
- ✓ The Baseline/Generic Cognitive Architecture
- Examples
  - ✓ Soar, LIDA, CopyCat
- Architecture in NN / Deep Learning Networks
- Conversational Agents
  - NLP
  - Dialog Management
  - Knowledge Graphs
  - Deep Networks

Marr's Three Levels of Abstraction

David Marr: *Theoretical Neuroscience* → *Computational Intelligence*

what? why?

Example

- **Computational Theory**

  What is the computation and by what principles is it accomplished?

  Textbook:
  ToC, body text, index

- **Algorithm**

  What representations and algorithms are used to carry forth computation?

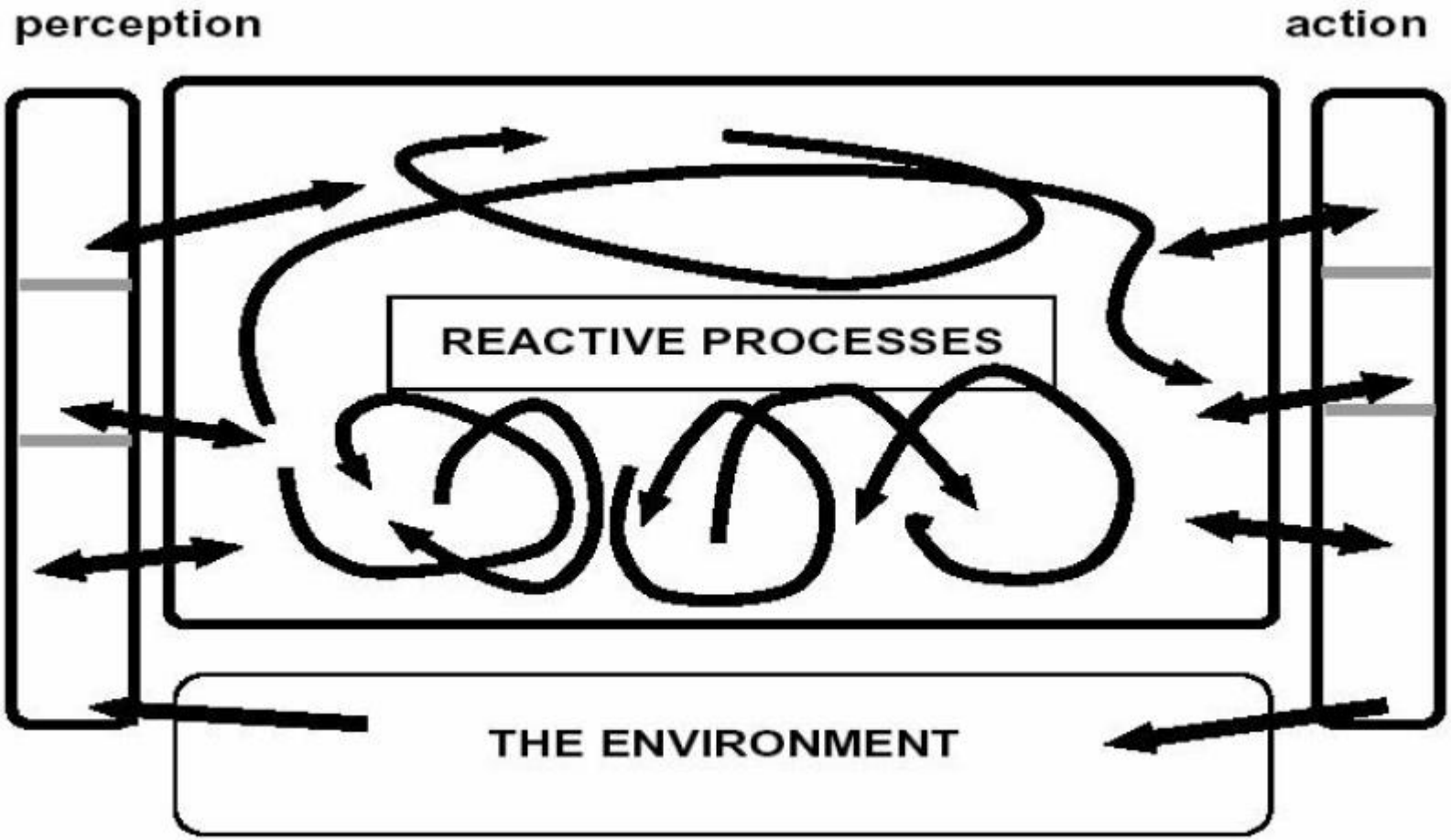
  Scrolling text roll, manual + automatic positioning

- **Implementation**

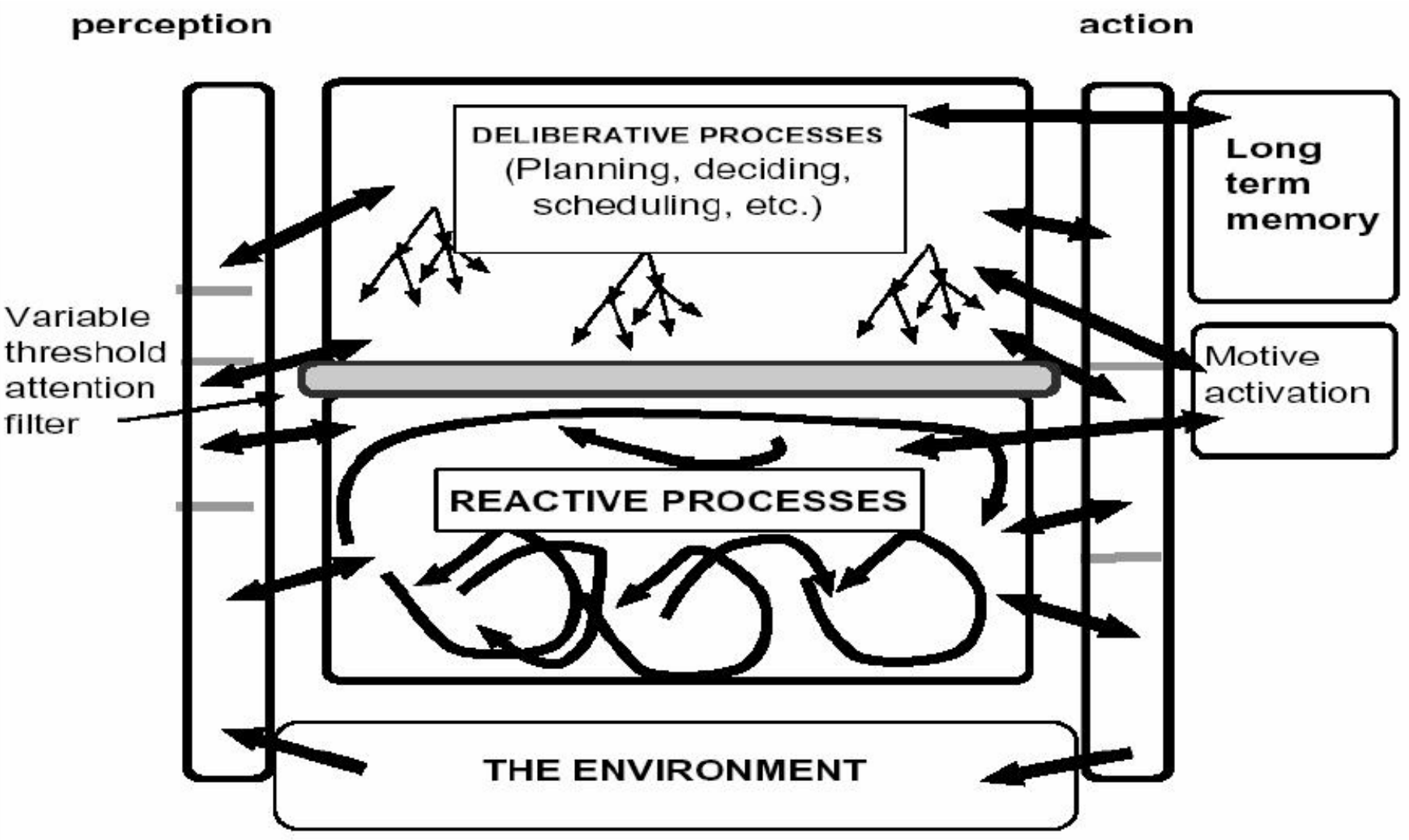
  On what physical hardware and firmware is the algorithm run?

  Projection of markings printed on acetate film, keyboard controls scrolling per ToC and index.

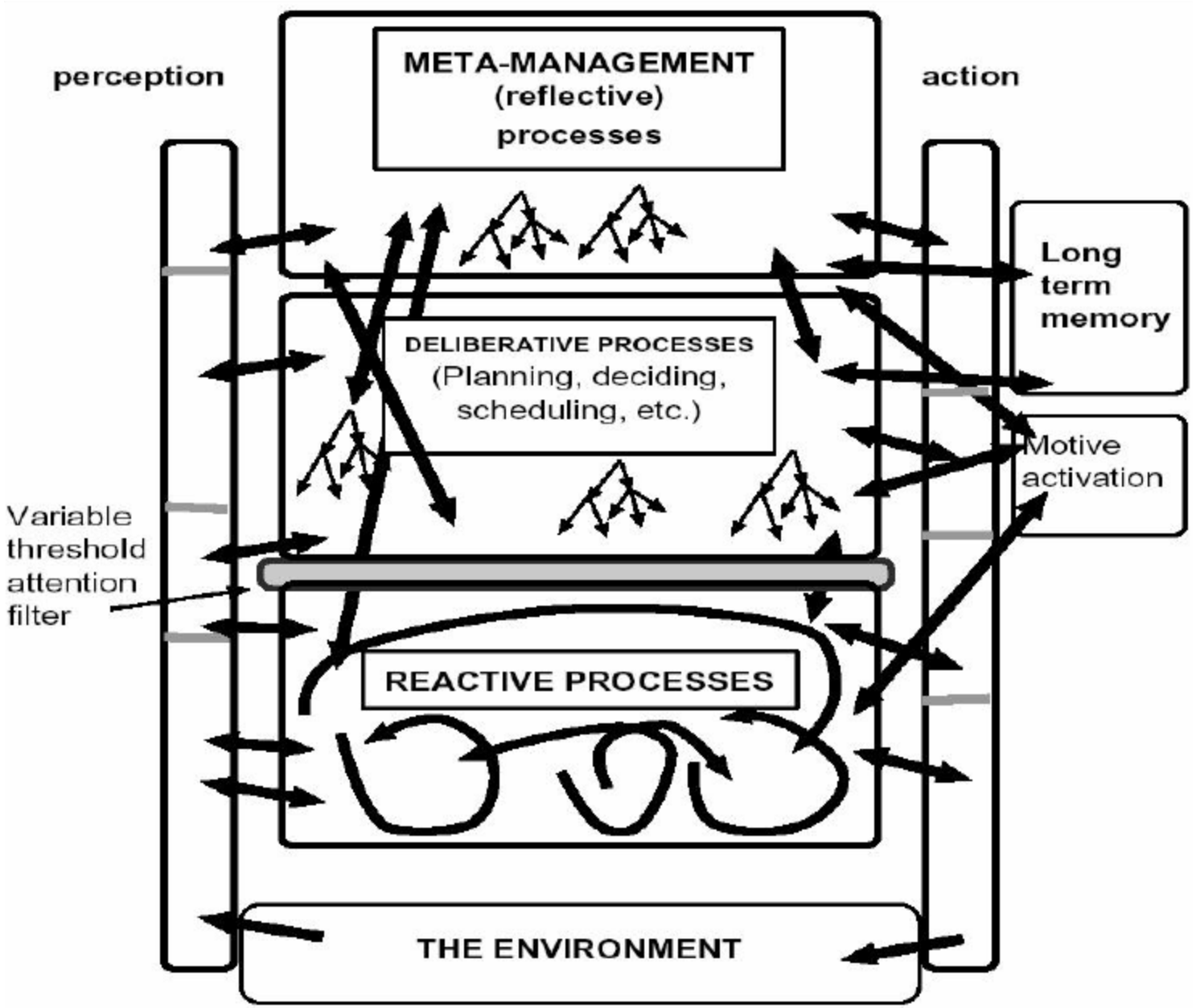
Cognitive Architecture: Reactive Agent



Cognitive Architecture: Deliberative Agent

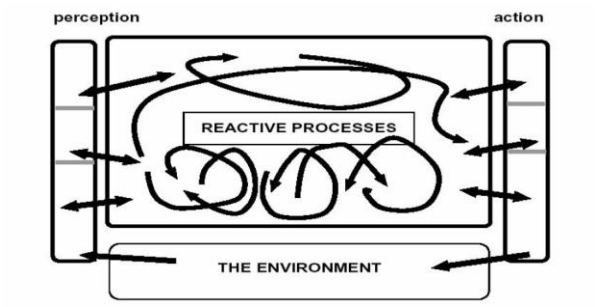


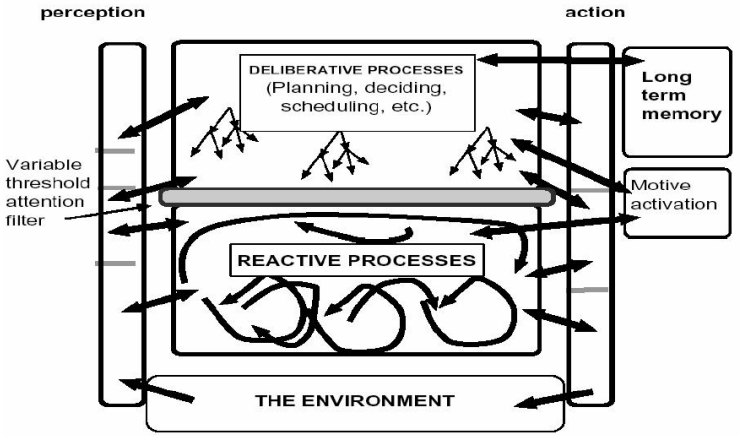
Cognitive Architecture: Reflective Agent



Aaron Sloman

Reactive vs. Deliberative





Reactive

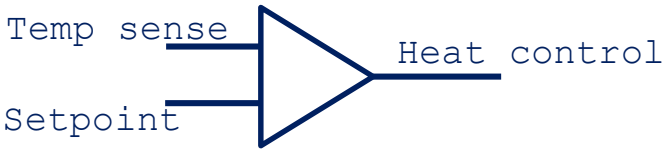
automatic & strictly determined

modest internal state

implicit representations

Kahneman System 1

Example: thermostat



Deliberative

makes choices

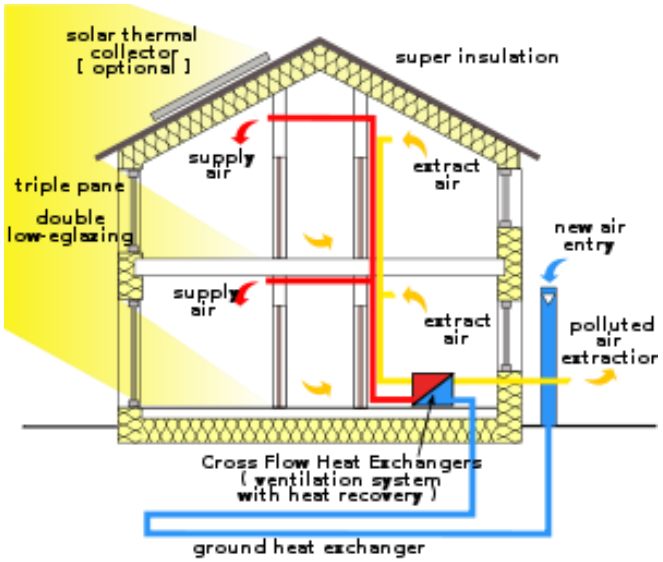
rich internal state

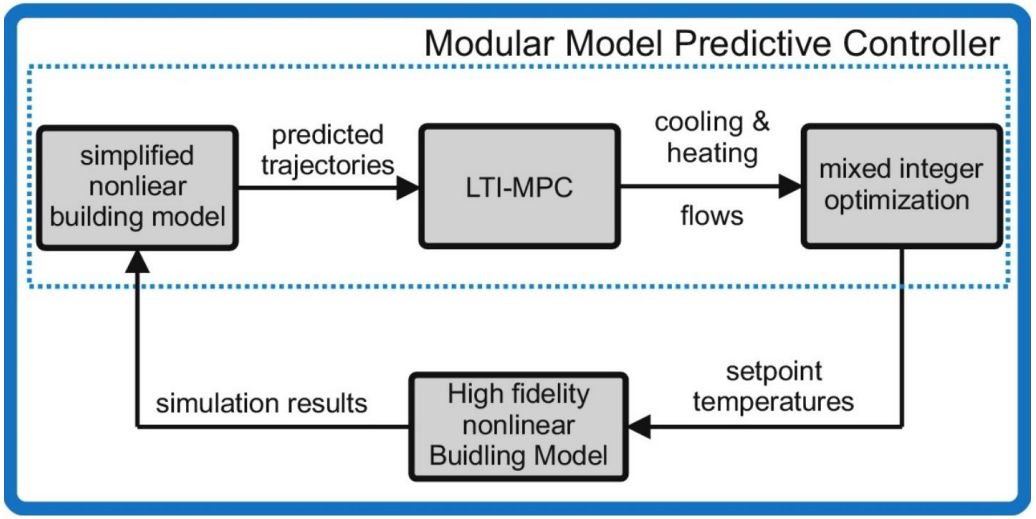
explicit world models

Kahenman System 2

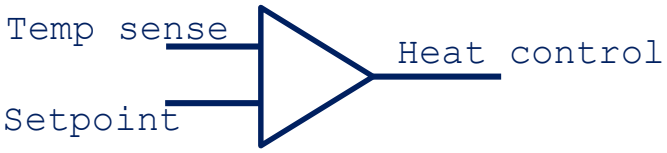
Example: building temperature management system

Reactive vs. Deliberative Building Temperature Controller



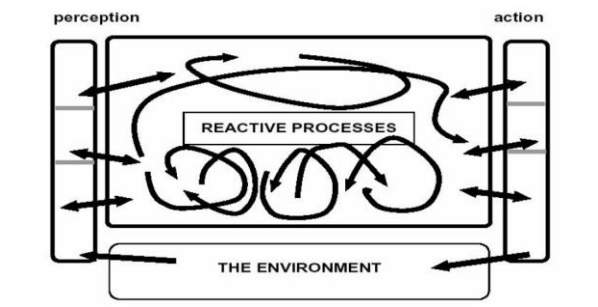


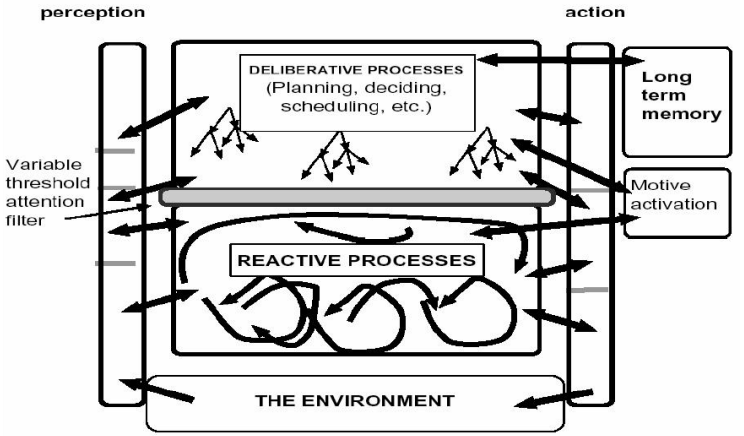
Example: thermostat



Example: building temperature management system

Reactive vs. Deliberative





Reactive

automatic & strictly determined

modest internal state

implicit representations

Deliberative

makes choices

rich internal state

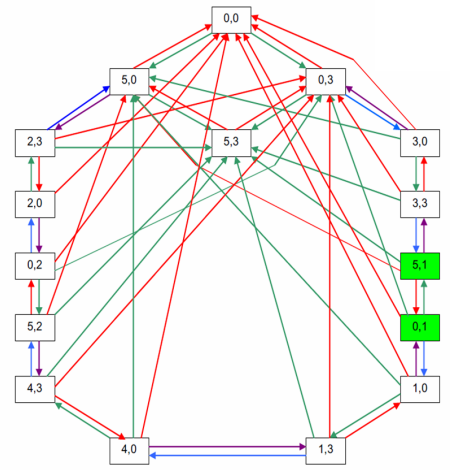
explicit world models

?

Soar

A Reactive Water Jug Solution in Soar

Deliberative search space

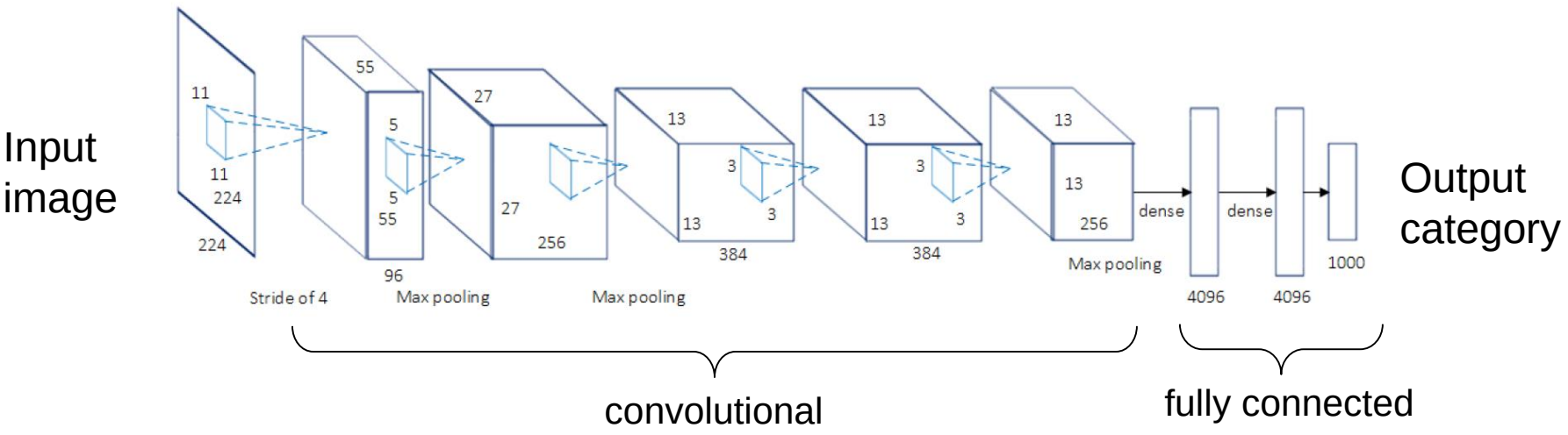


jug-5, jug-3

Reactive program rules

```
If   [0, 0]   Then   FILL(3)      =>   [0, 3]
If   [0, 3]   Then   Pour(3,5)    =>   [3, 0]
If   [3, 0]   Then   Fill(3)      =>   [3, 3]
If   [3, 3]   Then   Pour(3,5)    =>   [5, 1]   ✓
```

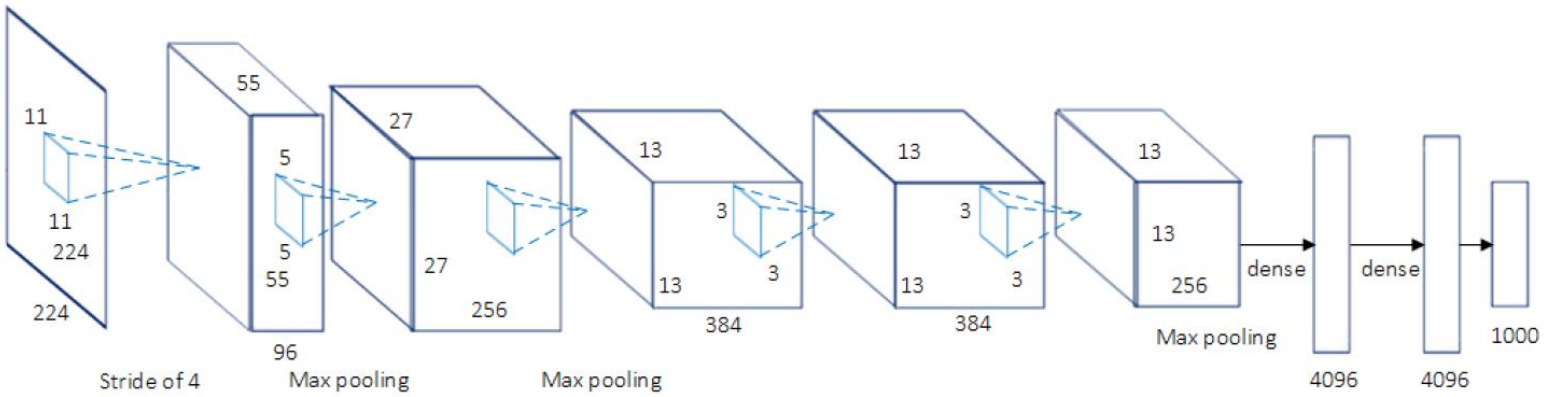
Neural Network Architectures



Alexnet

- Architectural Elements — Layer dimensions, weights, nonlinearities
- Organization — Layer connectivity
- Purpose — Function approximation

Neural Network Architectures



Reactive or deliberative?

choices?
state?
implicit/explicit

Cognitive Architecture: Fundamental Parameters

Cognitive Architecture

- Naturally Intelligent Agent
- Artificially Intelligent Agent

Where does knowledge reside?

How is state represented and utilized in decisions?

Forms of Memory

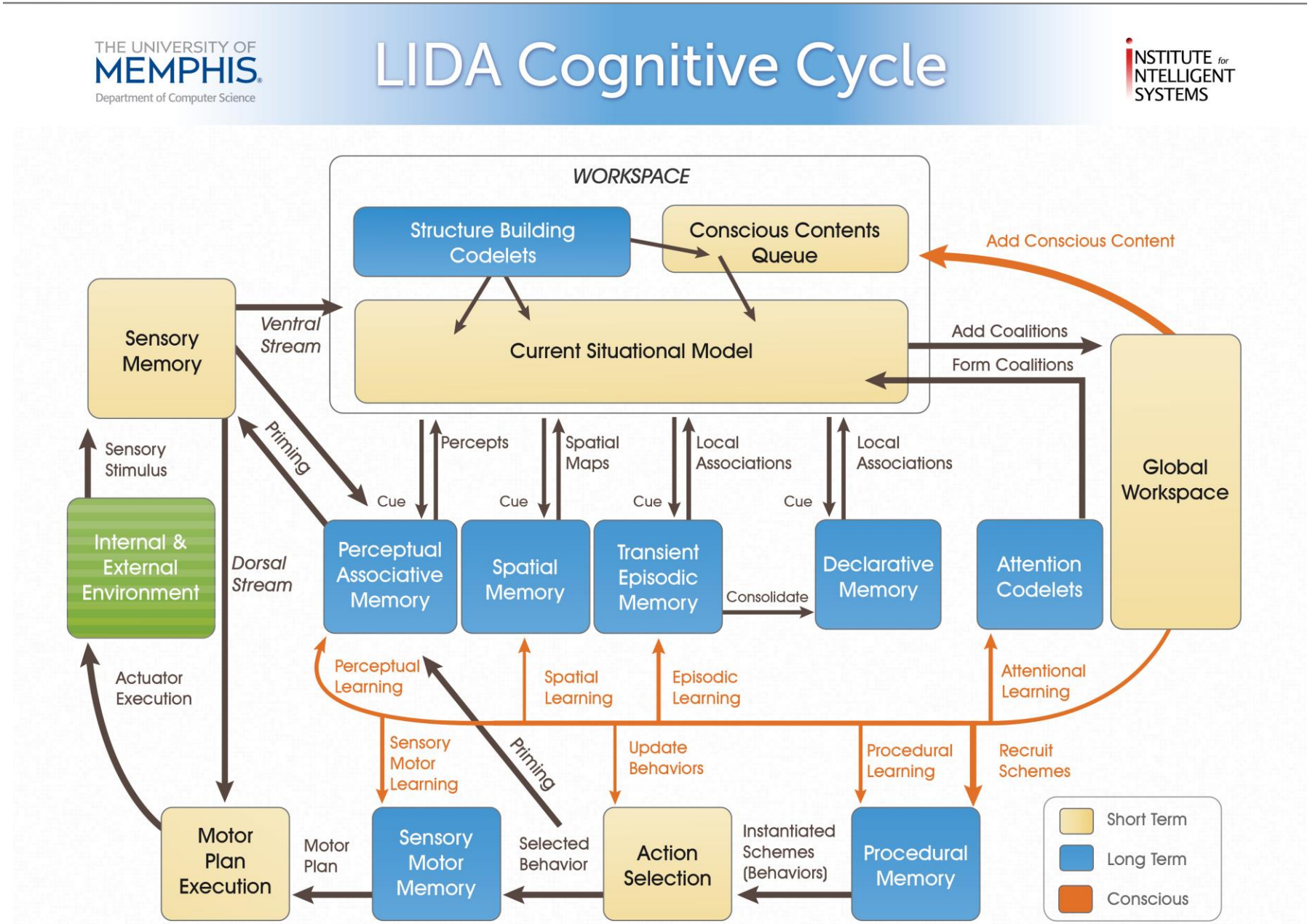
Localist vs. Distributed Representations

What is the control mechanism: what to do & think next?

Distinction between Program Control and Data

How Learning Happens

LIDA Cognitive Architecture



Stan Franklin

Copycat/Metacat

Task: Letter string analogy

$$abc \Rightarrow abd$$

$$mrrjjj \Rightarrow ?$$

Douglas Hofstadter, Melanie Mitchell, James Marshall

Copycat/Metacat

Task: Letter string analogy

$$abc \Rightarrow abd$$

$$xyz \Rightarrow ?$$

Douglas Hofstadter, Melanie Mitchell, James Marshall

Copycat Architecture

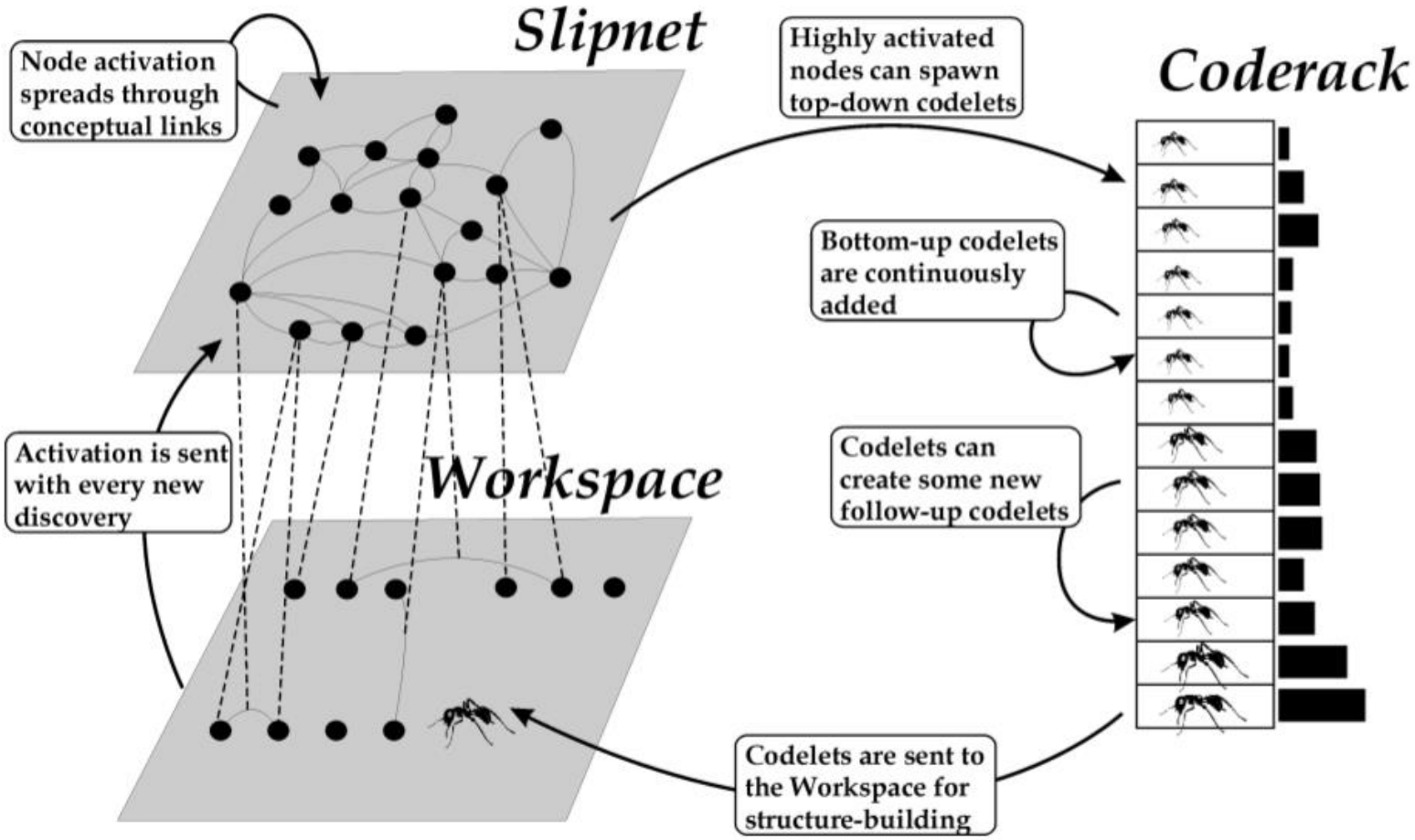
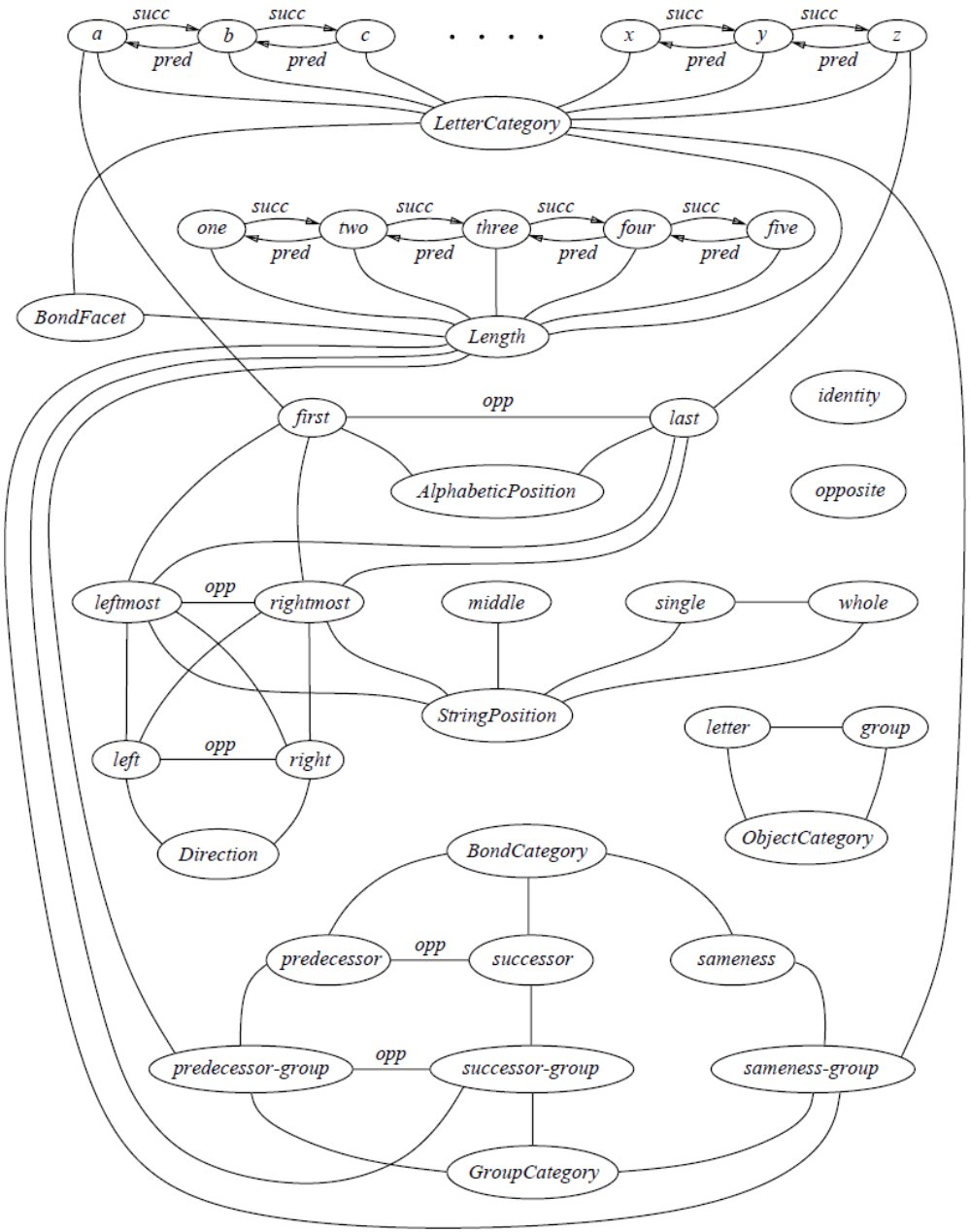


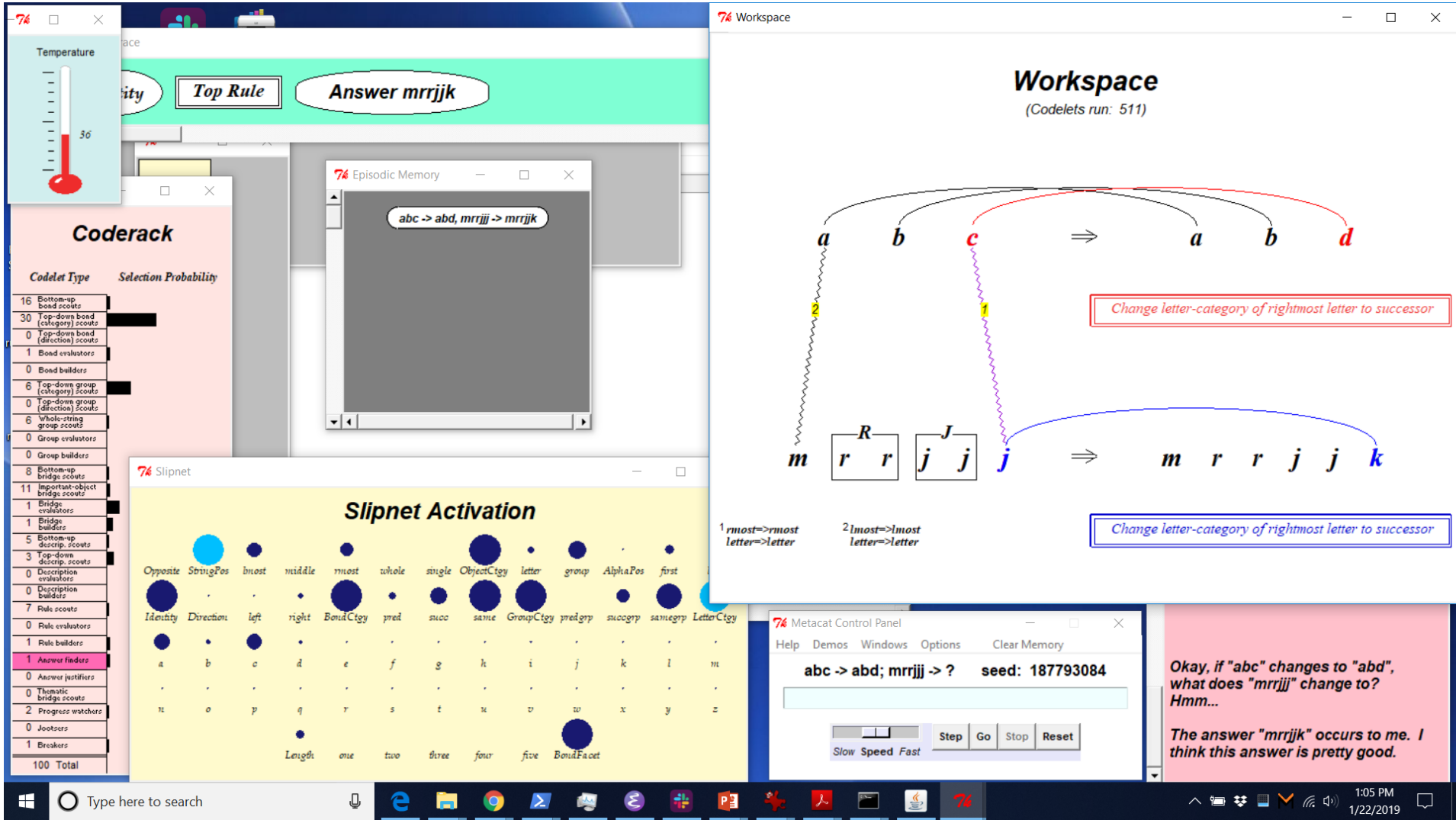
Figure 7: A feedback loop between perceptual and conceptual activity.

Joaquin Vanschoren

Copycat/Metacat Slipnet



Running Metacat



Cognitive Architecture: Outline

- Architecture in Information Systems

  "An organization of elements to achieve some information processing purpose."

- History

  Computational theory of mind

- The Baseline/Generic Cognitive Architecture

- Examples

  Soar, LIDA, CopyCat

- Architecture in NN / Deep Learning Networks

- Conversational Agents
  - NLP
  - Dialog Management
  - Knowledge graphs
  - Deep Networks

Conversational Agent

Question answering task

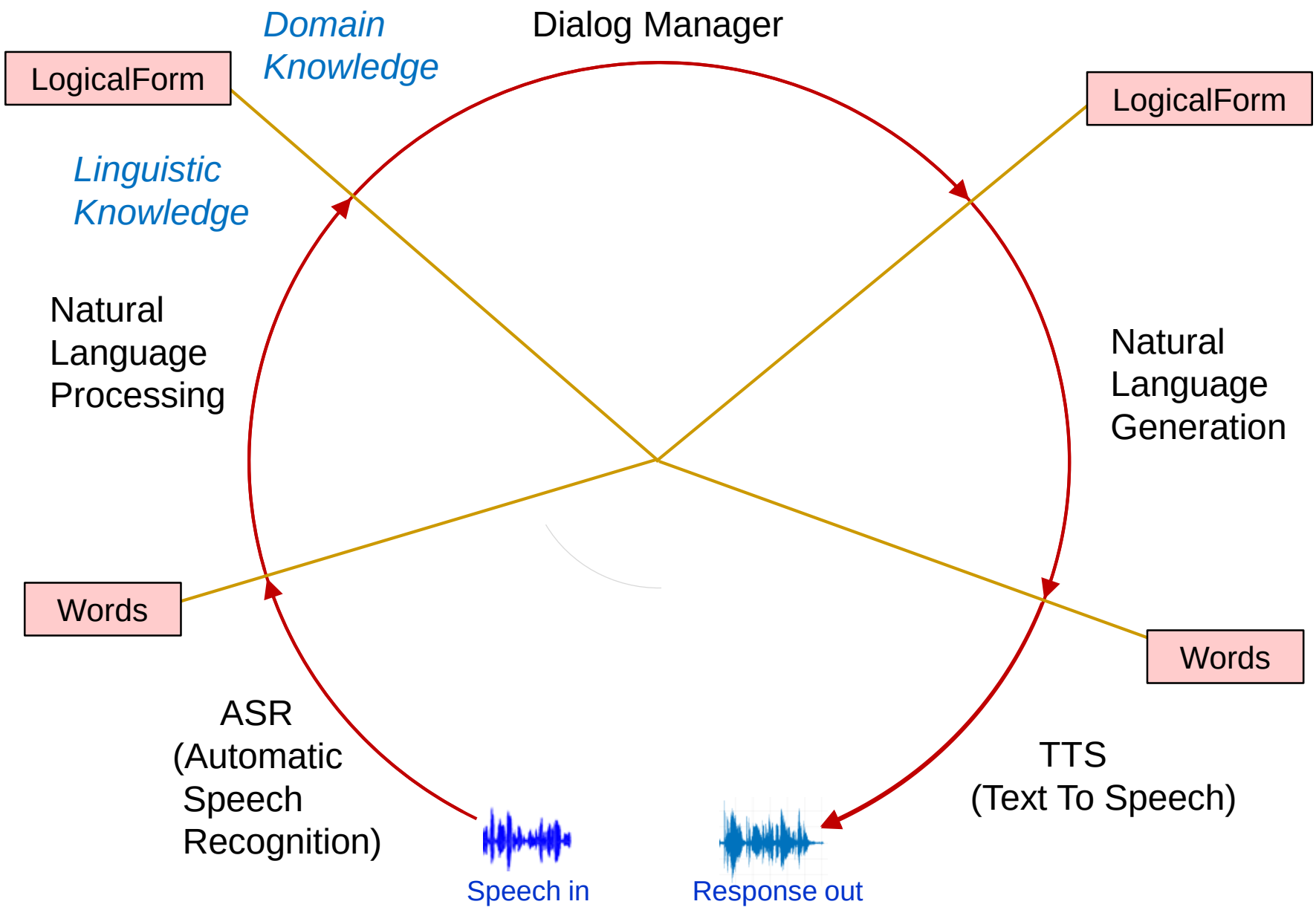
"Alexa, who won the 1934 world series?

"The Saint Louis Cardinals beat the Detroit Tigers 4-3 in the 1934 World Series."

"Alexa, what is Sam Taylor's phone number?

"Sam Taylor's telephone number in Boston is area code six-five-eight, one two three, four five six seven."

Architecture of a Conversational Agent



NLP: From Words to LogicalForms (Meaning)

"What is Sam Taylor's phone number?"

"What is the number for Sam in Boston?"

"Tell me Sam Taylor's telephone number."

"What do I dial to reach Sam Taylor?

Question-word
Command-word

Intent terms

Person-word
Location-word
Address-book-item

Entity terms

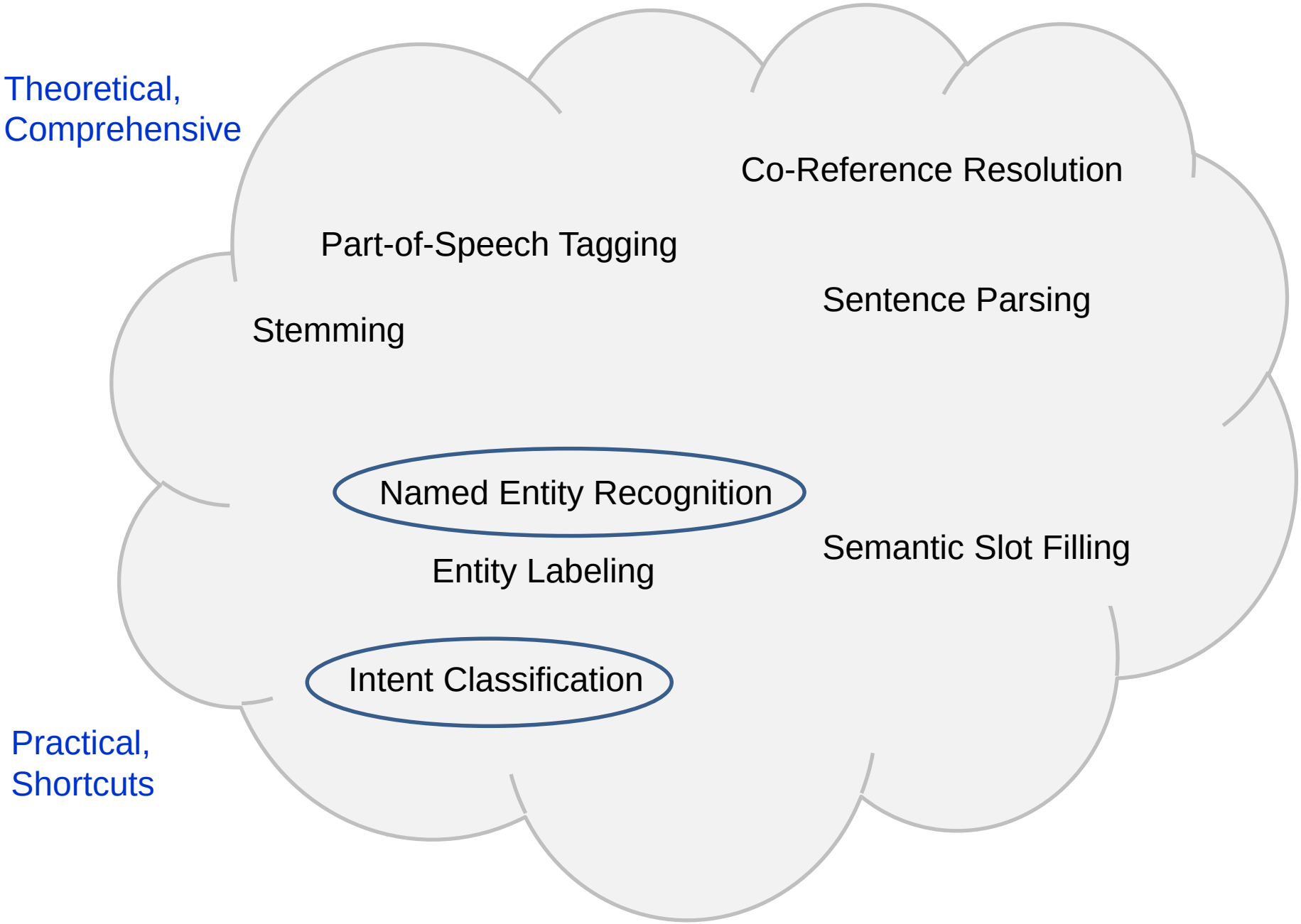
Syntactic Pattern:

{Question (WhatIs)} {Person}'s {X} ?

LogicalForm:

Query (AddressBookItem (Person (name "Sam Taylor"),
(telephoneNumber ?)))

NLP Pipeline Elements

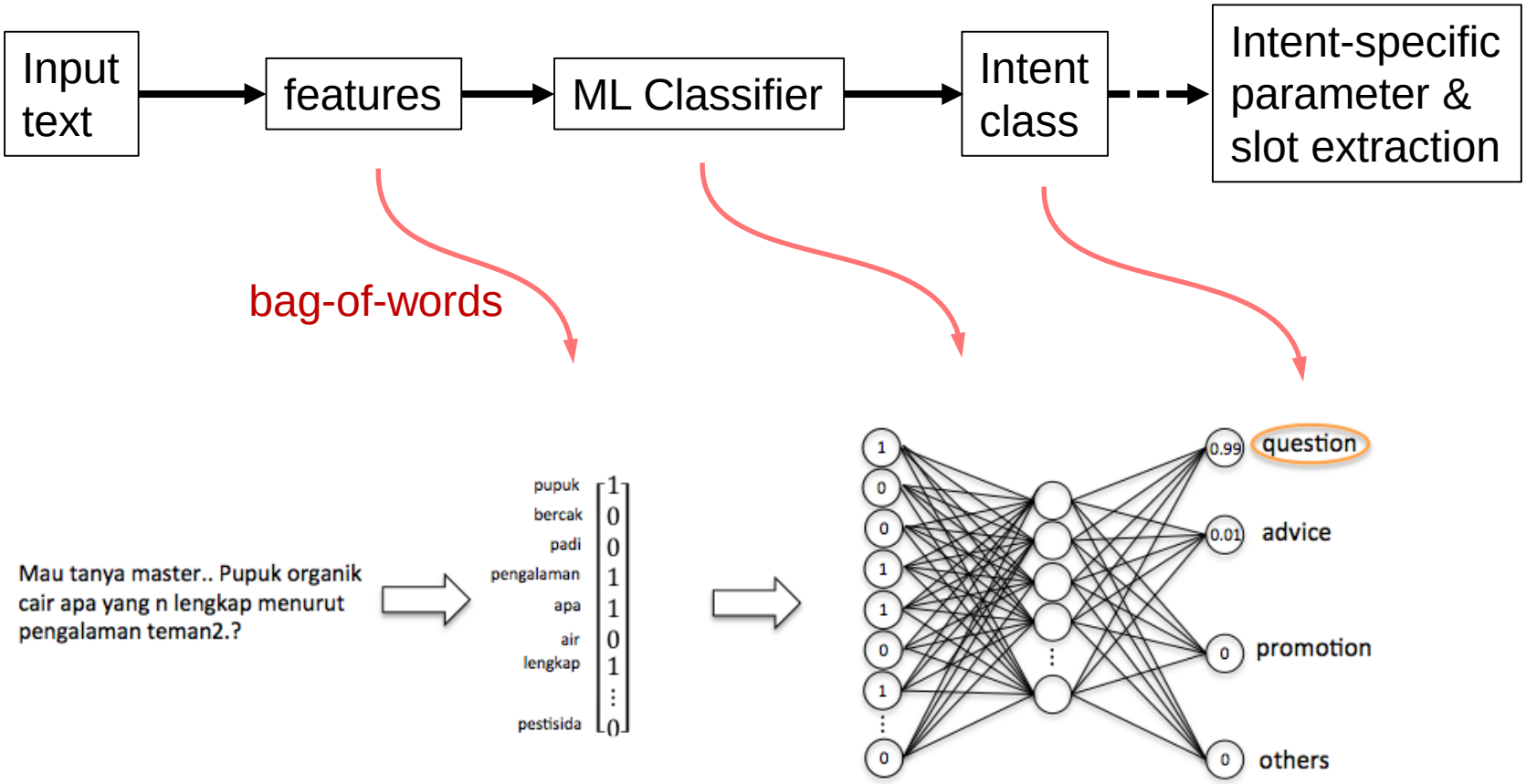


Intent Classification

Practical but Limited

What is Sam Taylor's phone number?

Data-Value-Question



Himang Sharatun

NER: Named Entity Recognition

"What is John's address in Garden Grove?"

1. common nouns:
   garden
   grove

2. but uncommon to see the bi-gram,
   ....."garden grove"...

3. City database:
   "Garden Grove, California"

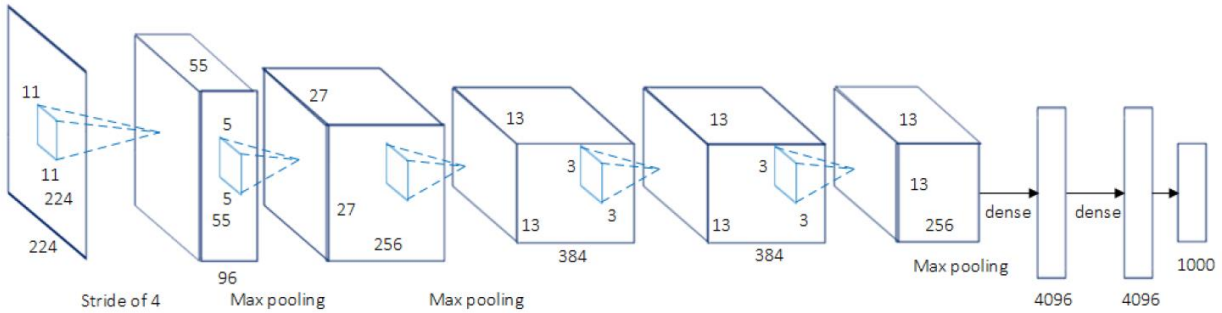
4. Lexical environment

address in { location }

- rules
- standard ML

Symbol Processing in Neural Networks

Input type:
vector



Output type:
vector

Can Neural Networks Produce Symbolic Structures from Sentences?

Input type:
Words = sequence of symbols

"What is Sam Taylor's phone number?"

Output type:
Logical Form = hierarchical tree of symbols

Query (AddressBookItem (Person (name "Sam Taylor"),
(telephoneNumber ?)))

Deep Learning Architecture for Parsing to Logical Form

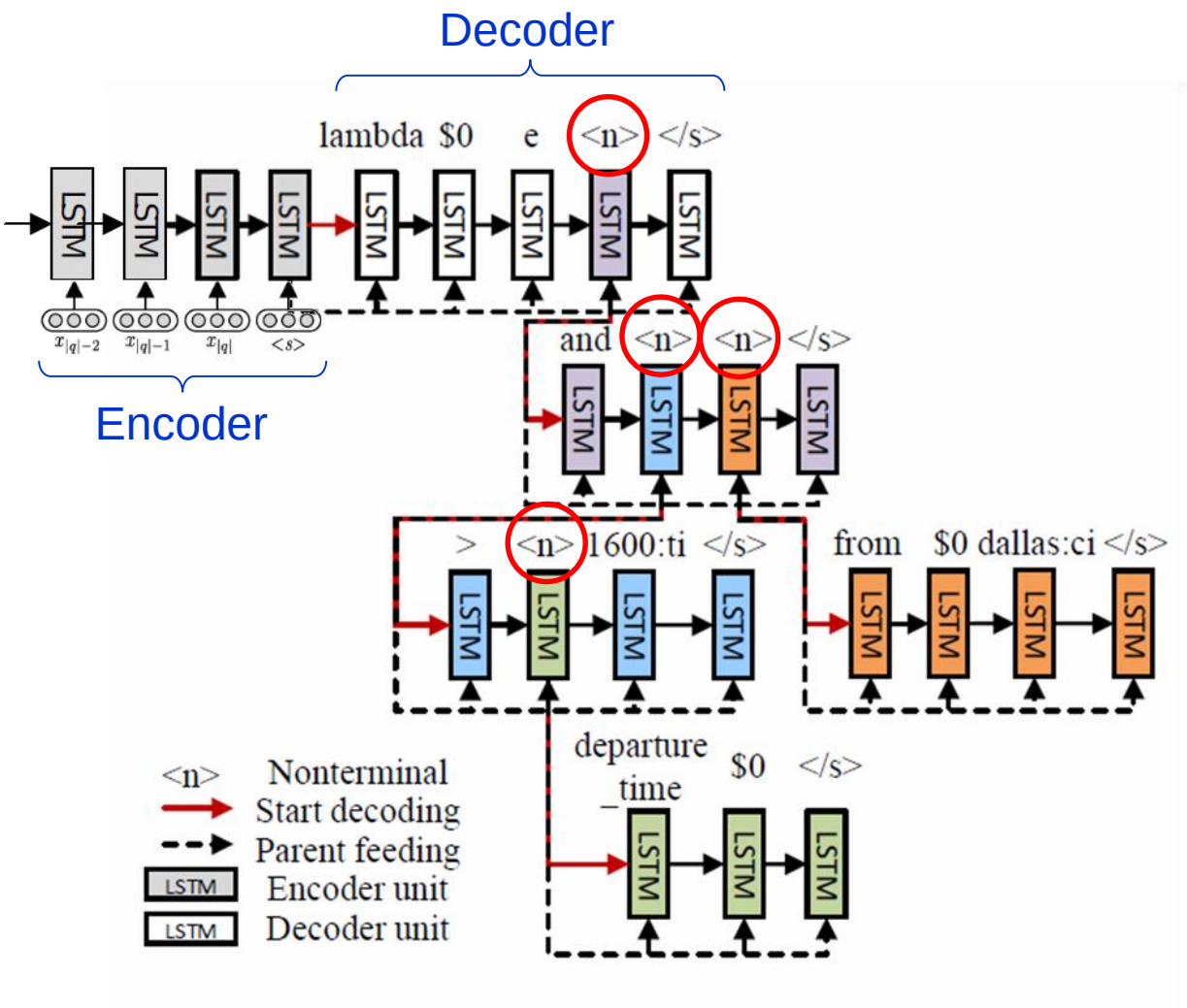


Figure 3: Sequence-to-tree (SEQ2TREE) model with a hierarchical tree decoder.

Dong and Lapata, 2016

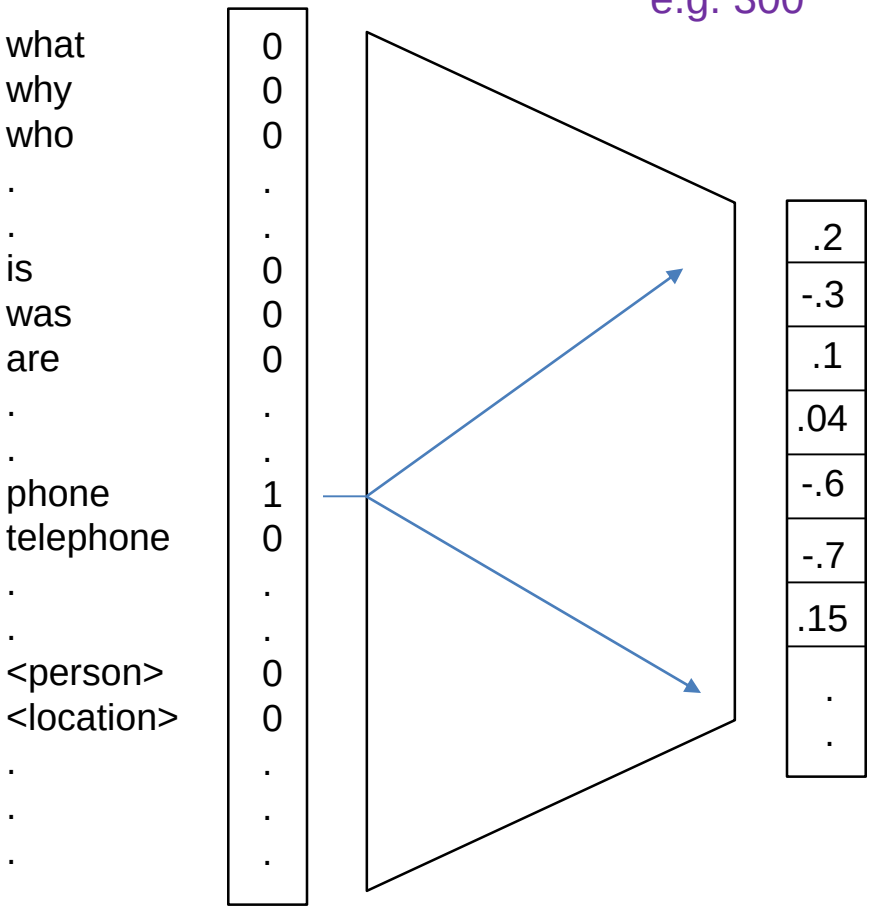
Word Vector Embedding

Purpose:

- lower dimensionality reduces data sparsity for learning
- similar words :: similar vectors leads to improved generalization

word vocabulary size
e.g. 10k

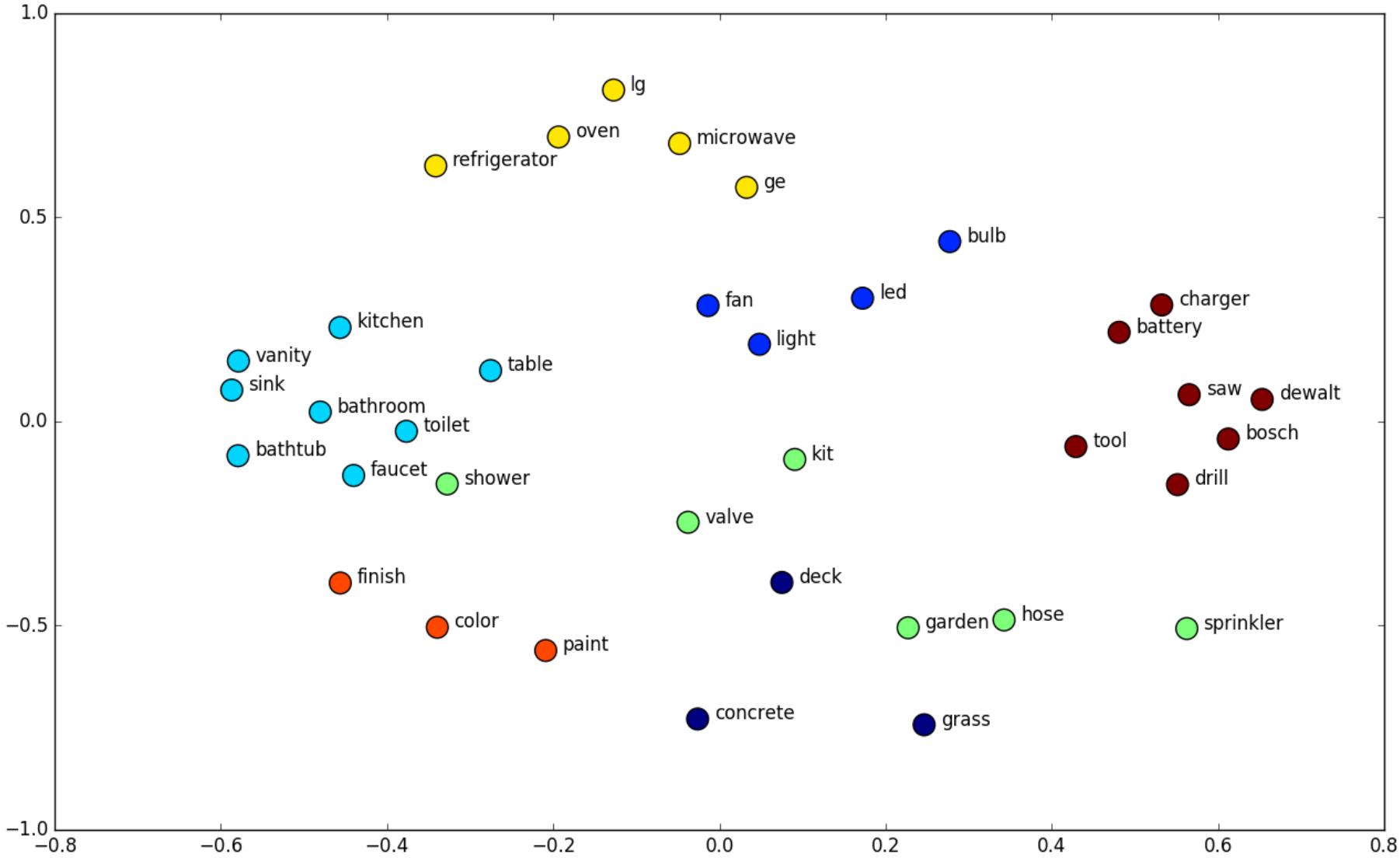
embedding vector size
e.g. 300



Word Vector Embedding

Purpose:

- lower dimensionality reduces data sparsity for learning
- similar words :: similar vectors leads to improved generalization



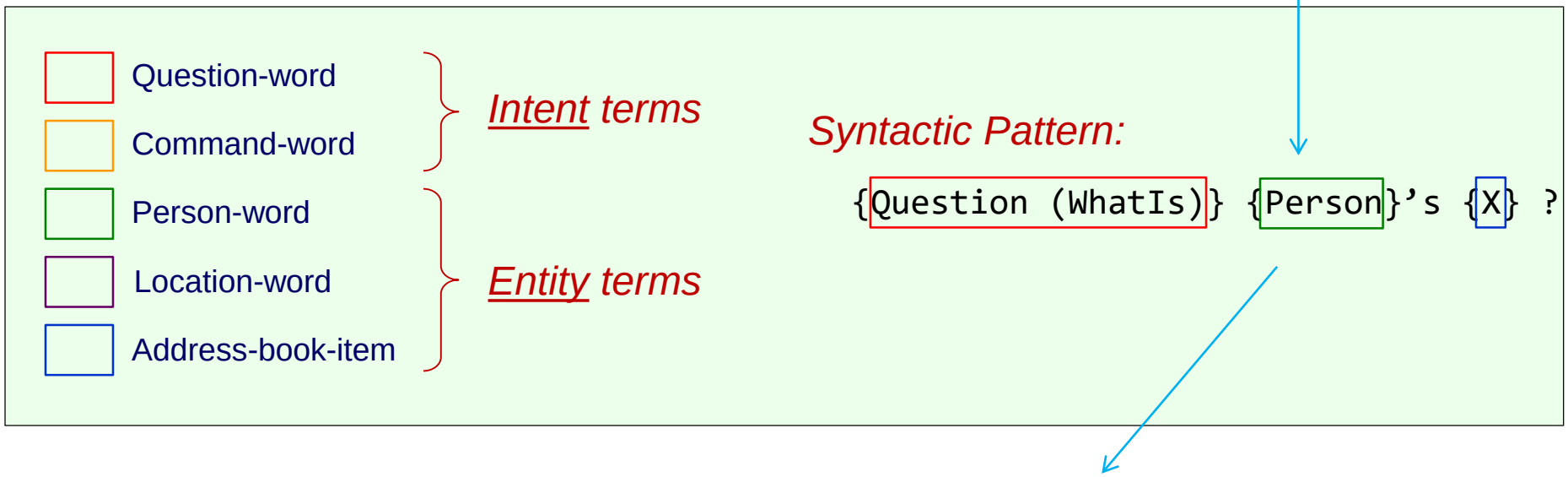
NLP: From Words to LogicalForms (Meaning)

"What is Sam Taylor's phone number?"

"What is the number for Sam in Boston?"

"Tell me Sam Taylor's telephone number."

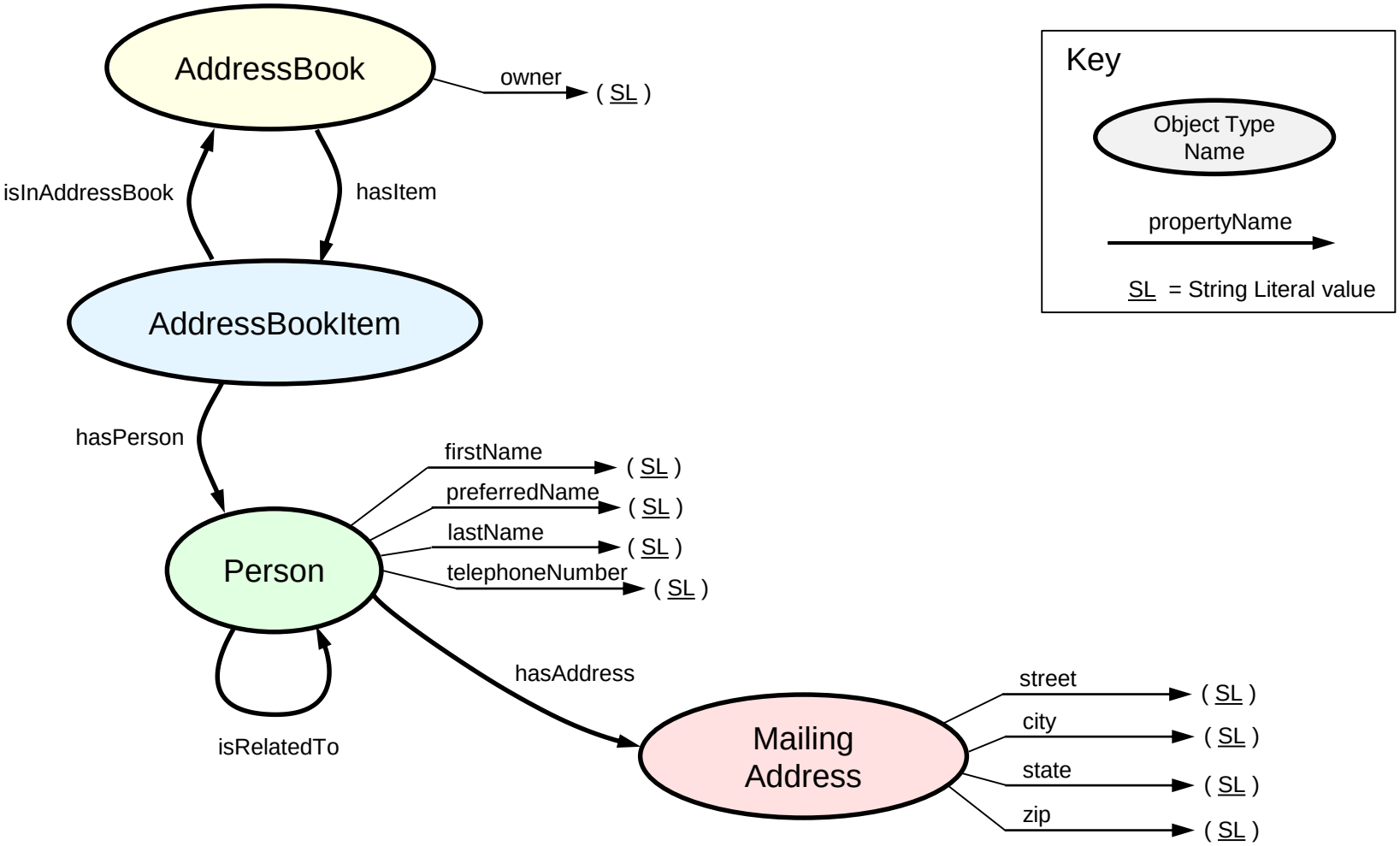
"What do I dial to reach Sam Taylor?



LogicalForm:

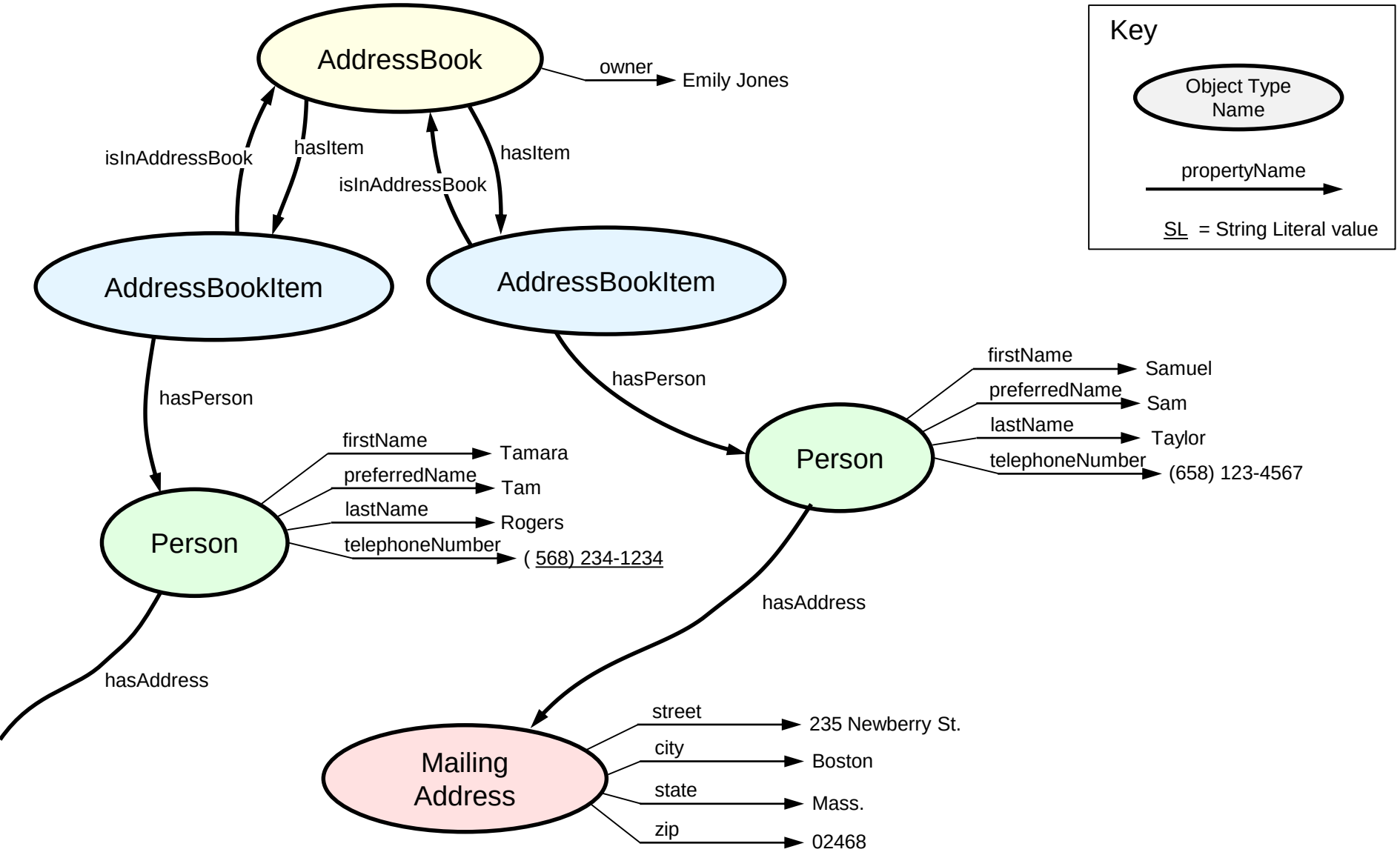
Query (AddressBookItem (Person (name "Sam Taylor"),
(telephoneNumber ?)))

Address Book Knowledge Base: Ontology Schema



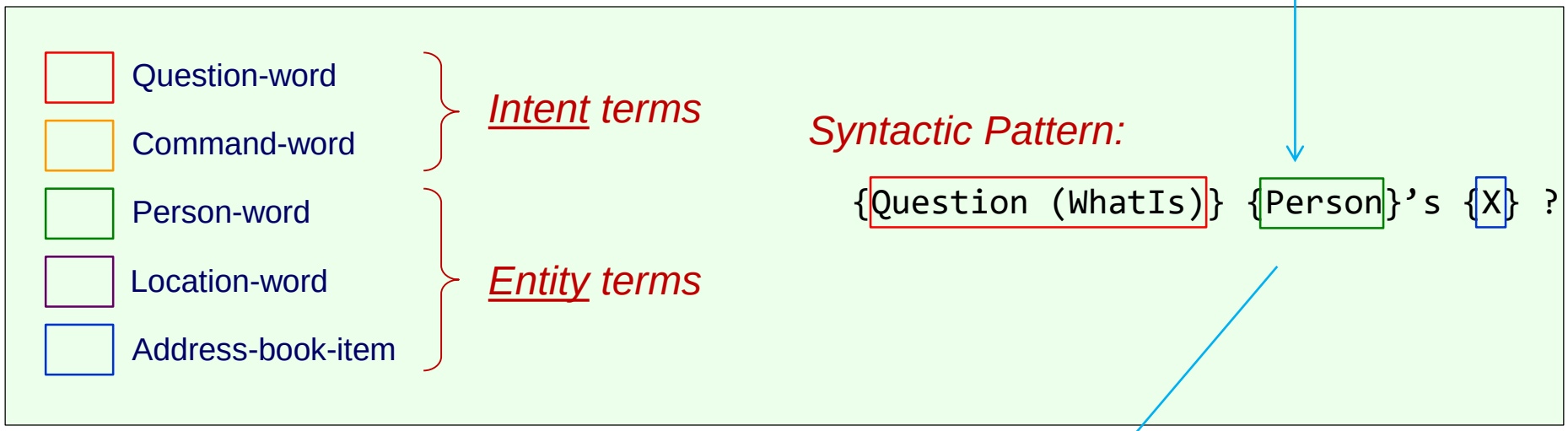
Address Book Knowledge Base: Data Example

Knowledge Graph



NLP: From Question to KB Query

"What is Sam Taylor's phone number?"



LogicalForm:

Query (AddressBookItem (Person (name "Sam Taylor"),
(telephoneNumber ?)))

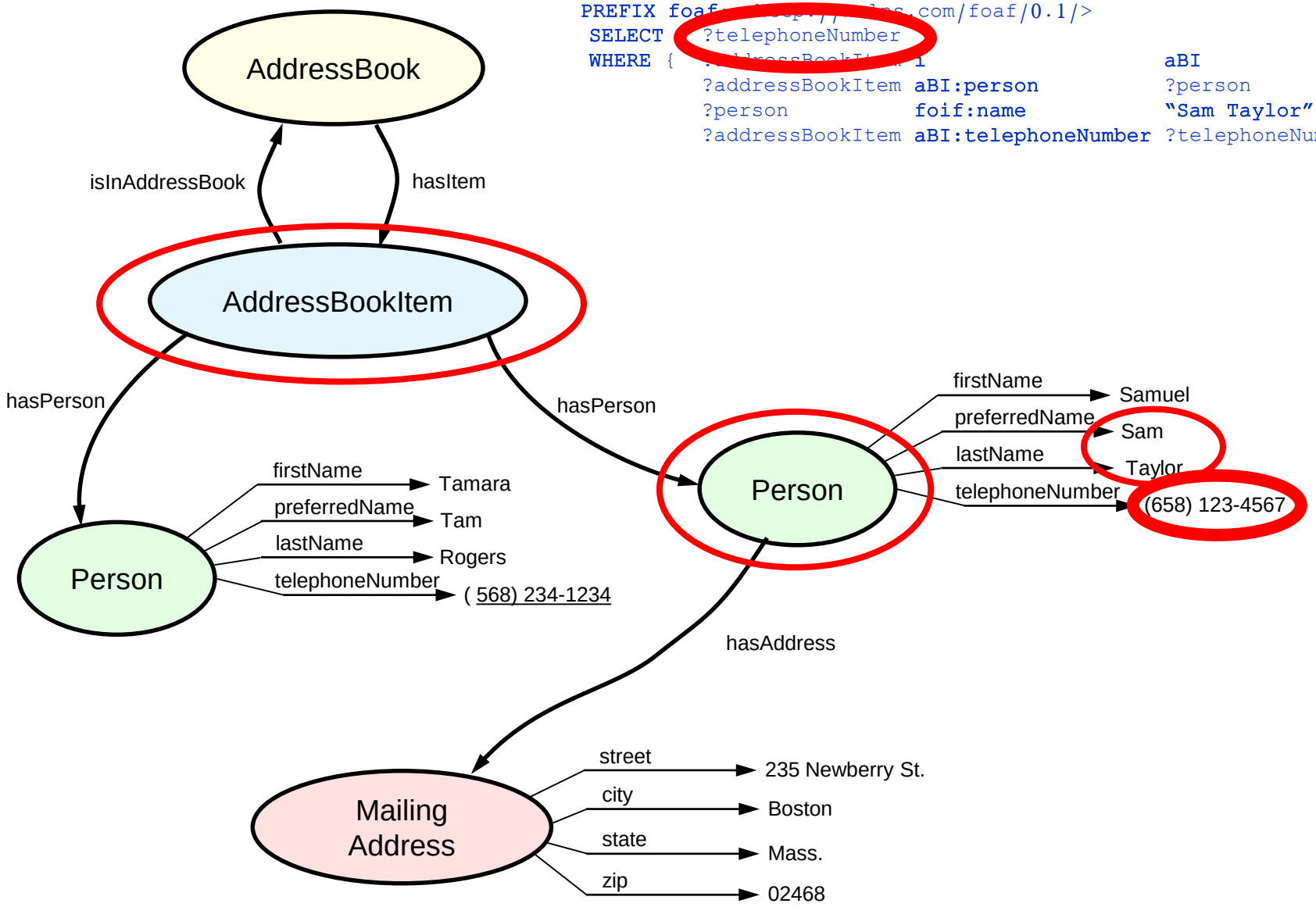
SPARQL KB Query:

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT   ?telephoneNumber
WHERE {  ?addressBookItem i                    aBI                .
         ?addressBookItem aBI:person           ?person            .
         ?person          foif:name            "Sam Taylor"       .
         ?addressBookItem aBI:telephoneNumber  ?telephoneNumber    }
```

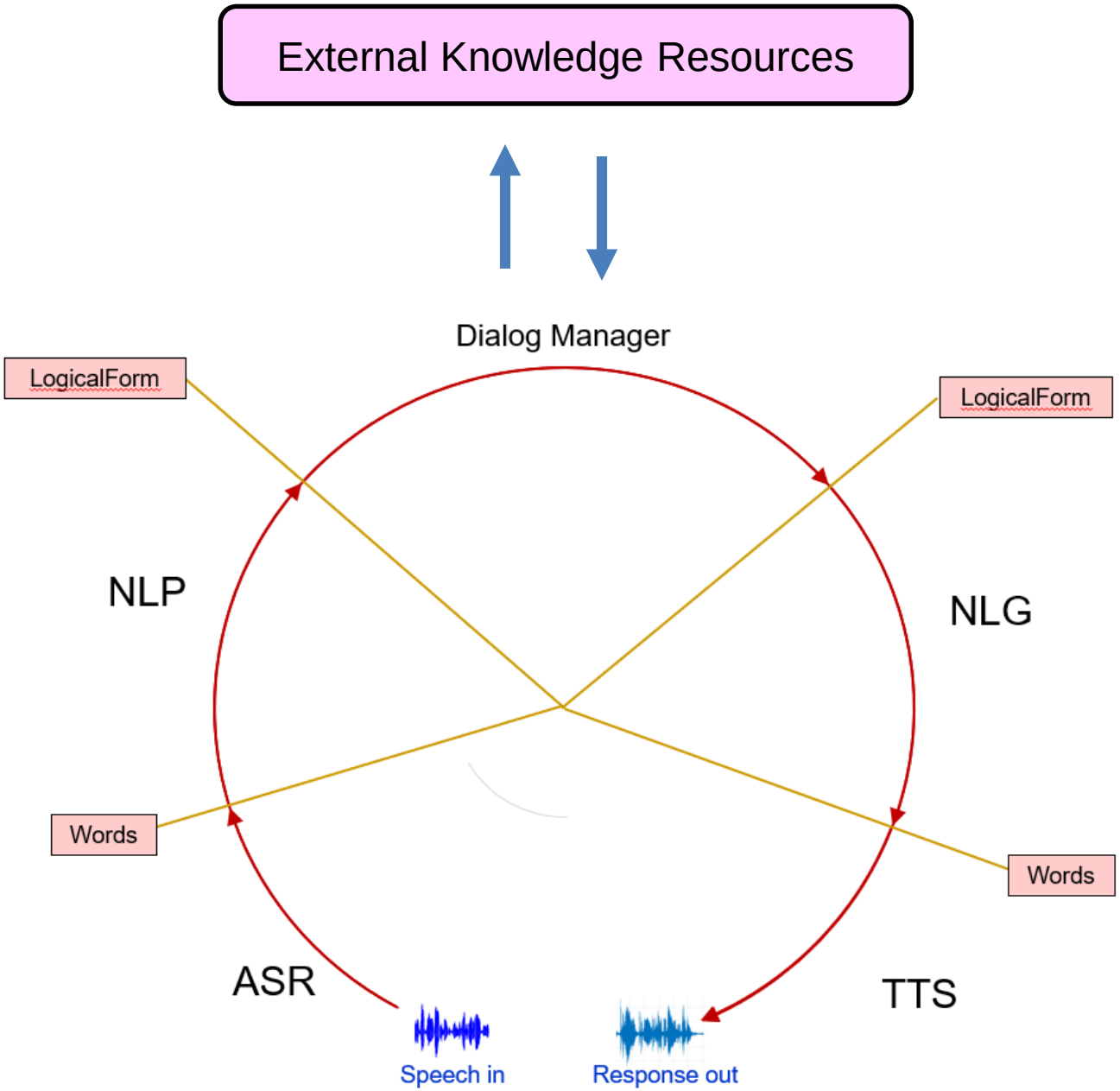
Query Match to Address Book Knowledge Base

SPARQL KB Query:

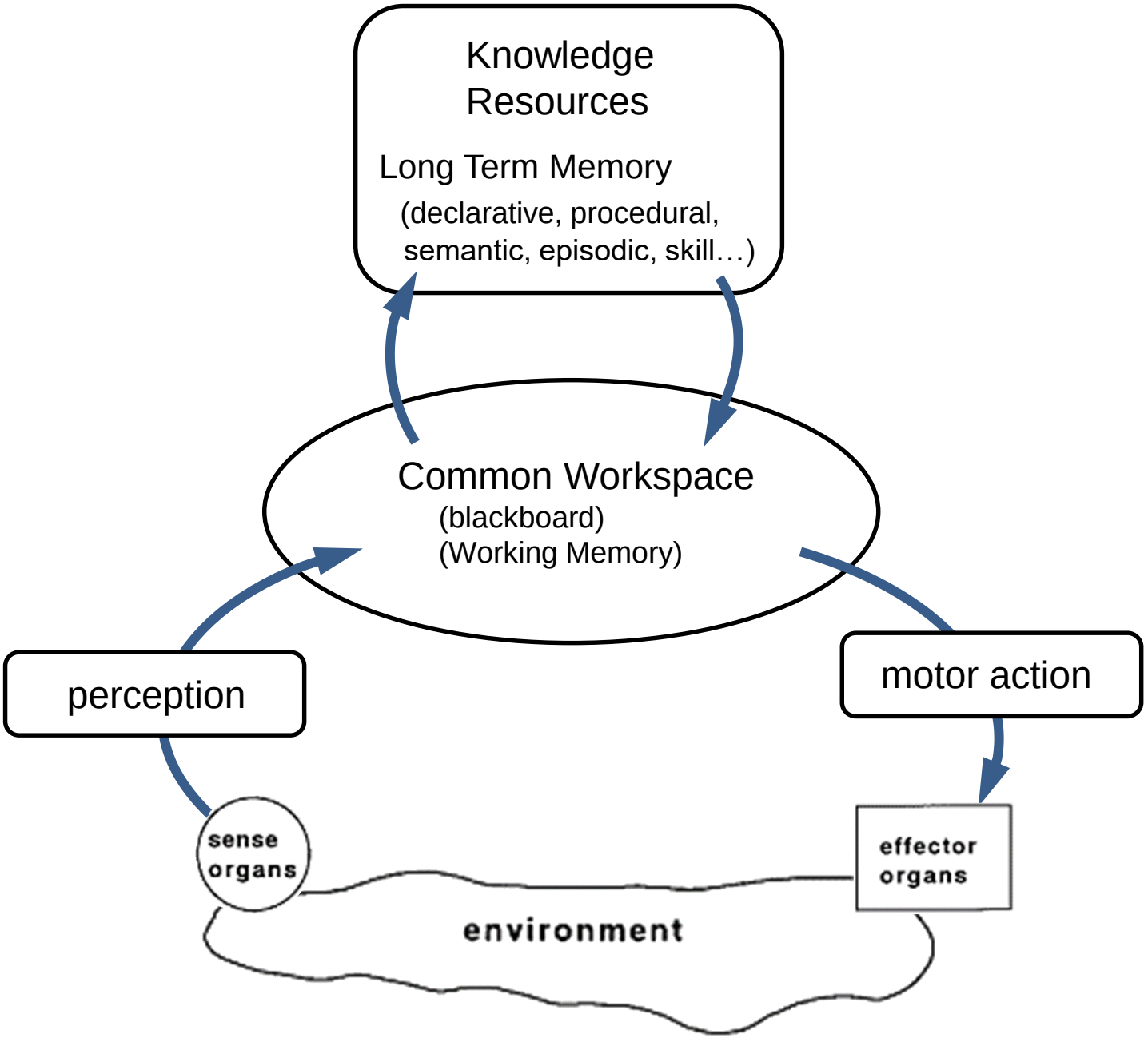
```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
 SELECT  ?telephoneNumber
 WHERE {  ?addressBookItem  rdf:type                aBI              .
          ?addressBookItem  aBI:person              ?person          .
          ?person           foif:name              "Sam Taylor"     .
          ?addressBookItem  aBI:telephoneNumber    ?telephoneNumber
```



Architecture of a Conversational Agent



Cognitive Architecture: Standard Model



Why Are Conversational Agents So Dumb?

"Alexa, who won the 1934 world series?

"The Saint Louis Cardinals beat the Detroit Tigers 4-3 in the 1934 World Series."

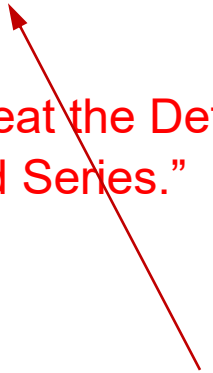
"Alexa, who was the starting pitcher for the Tigers?"

"The probable pitcher for the Tigers on Saturday February 23 has not been announced yet."

Co-Reference

"Alexa, who won the 1934 world series?

"The Saint Louis Cardinals beat the Detroit Tigers 4-3 in the 1934 World Series."



"Alexa, who was the president then?"

"This might answer your question. The president of the United States is Donald Trump."

Needed: A cognitive architecture with contextual memory.

Contextual Memory in Conversational Agents (2019)

- Remember past Intents, Entities
  - Query terms:
    - baseball
    - world series
    - 1934
  - Response terms:
    - Saint Louis Cardinals
    - Detroit Tigers
    - 4-3

Alexa
Google Home

"Who was the manager of the Cardinals?" (St. Louis Cardinals, in 1934)

- Dialog State Tracking

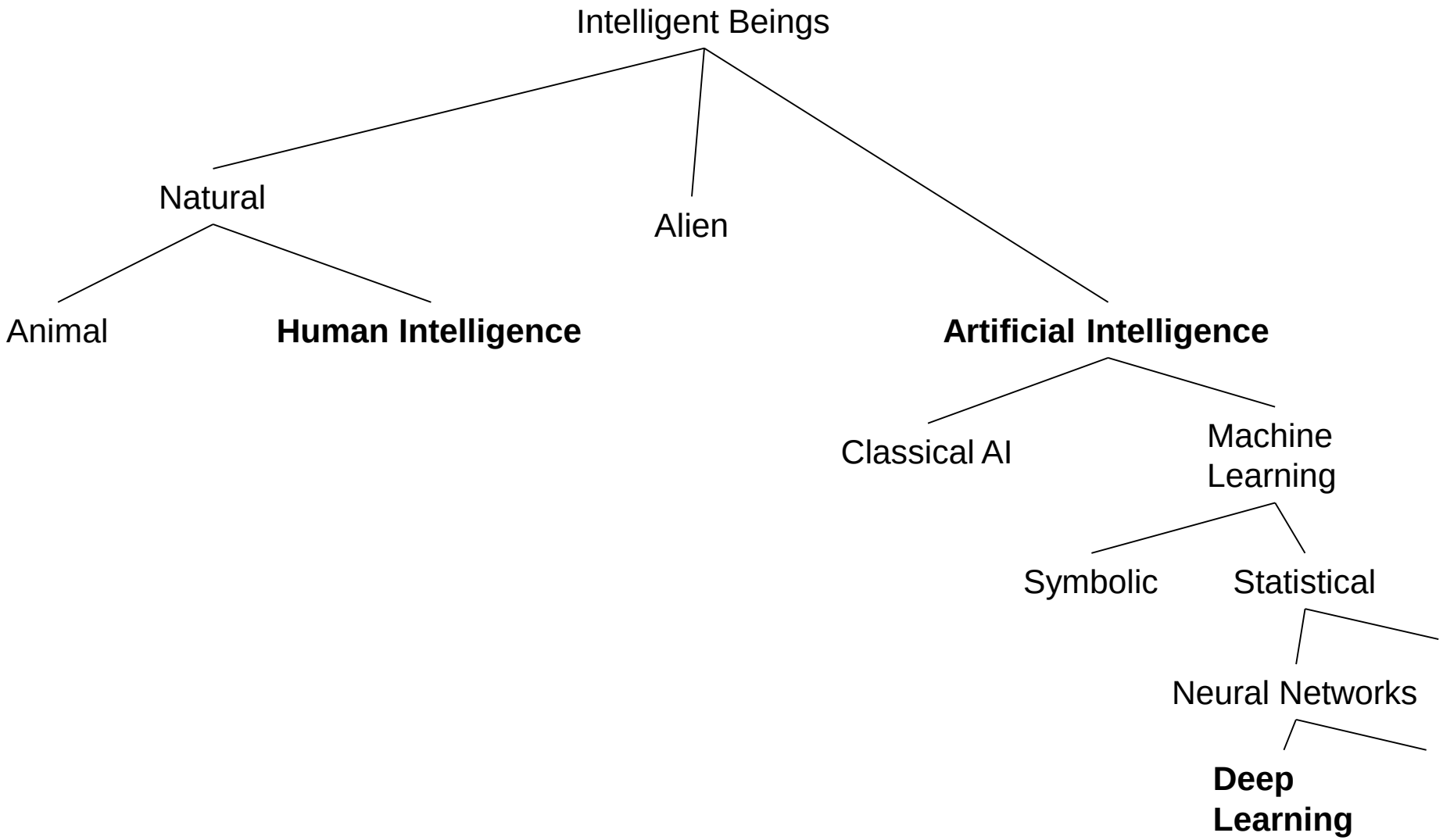
  -slot filling for restricted tasks, e.g. ordering pizza

  "Pizzabot I want a pepperoni pizza with tomato sauce"

size	?
sauce	tomato
topping1	pepperoni
topping2	X
topping3	X

"Sure, what size?"

Summary: Taxonomy



Cognitive Architecture

NN Architecture

Eric Saund

- *Research scientist in Cognitive Science and AI.*
- *Conversational Agents, Visual Perception, Cognitive Architectures.*
- *I build stuff.*

Projects

Papers

Curiosities

Links

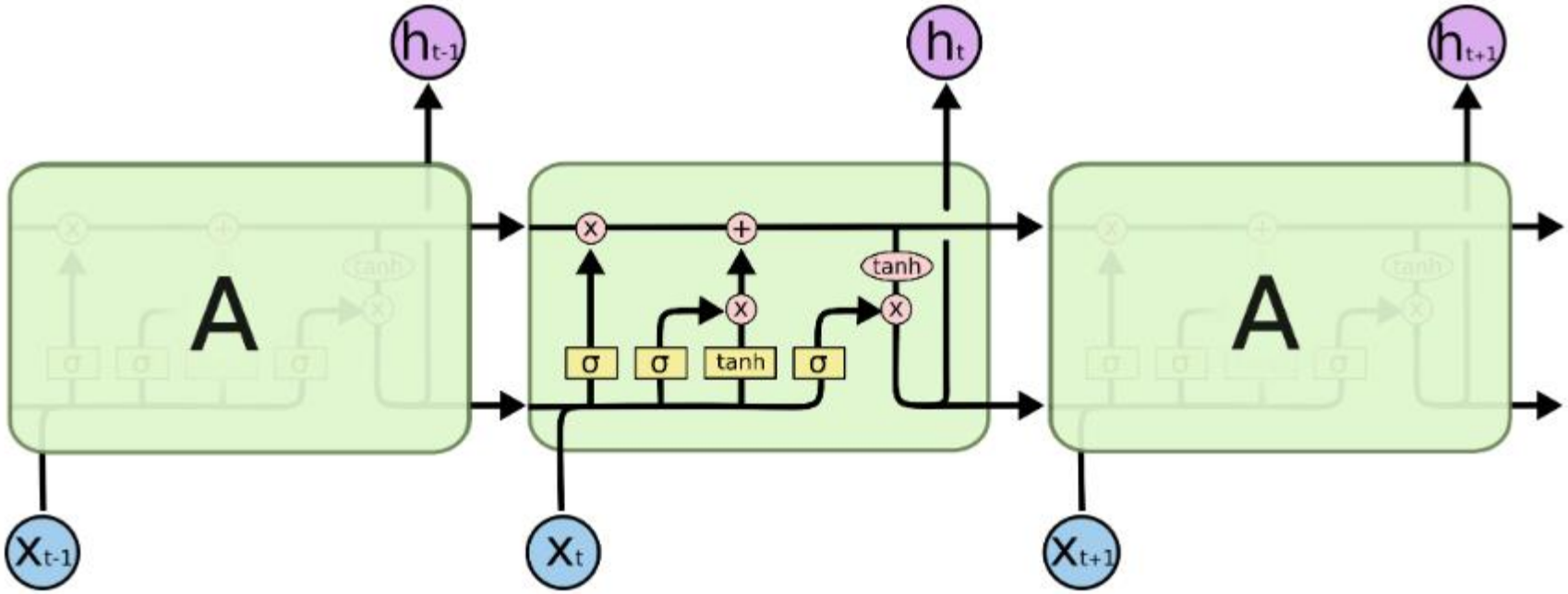
Contact

http://www.saund.org

saund@alum.mit.edu

Conversation

LSTM Module



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Deep Learning Architecture for Parsing to Frame Graph

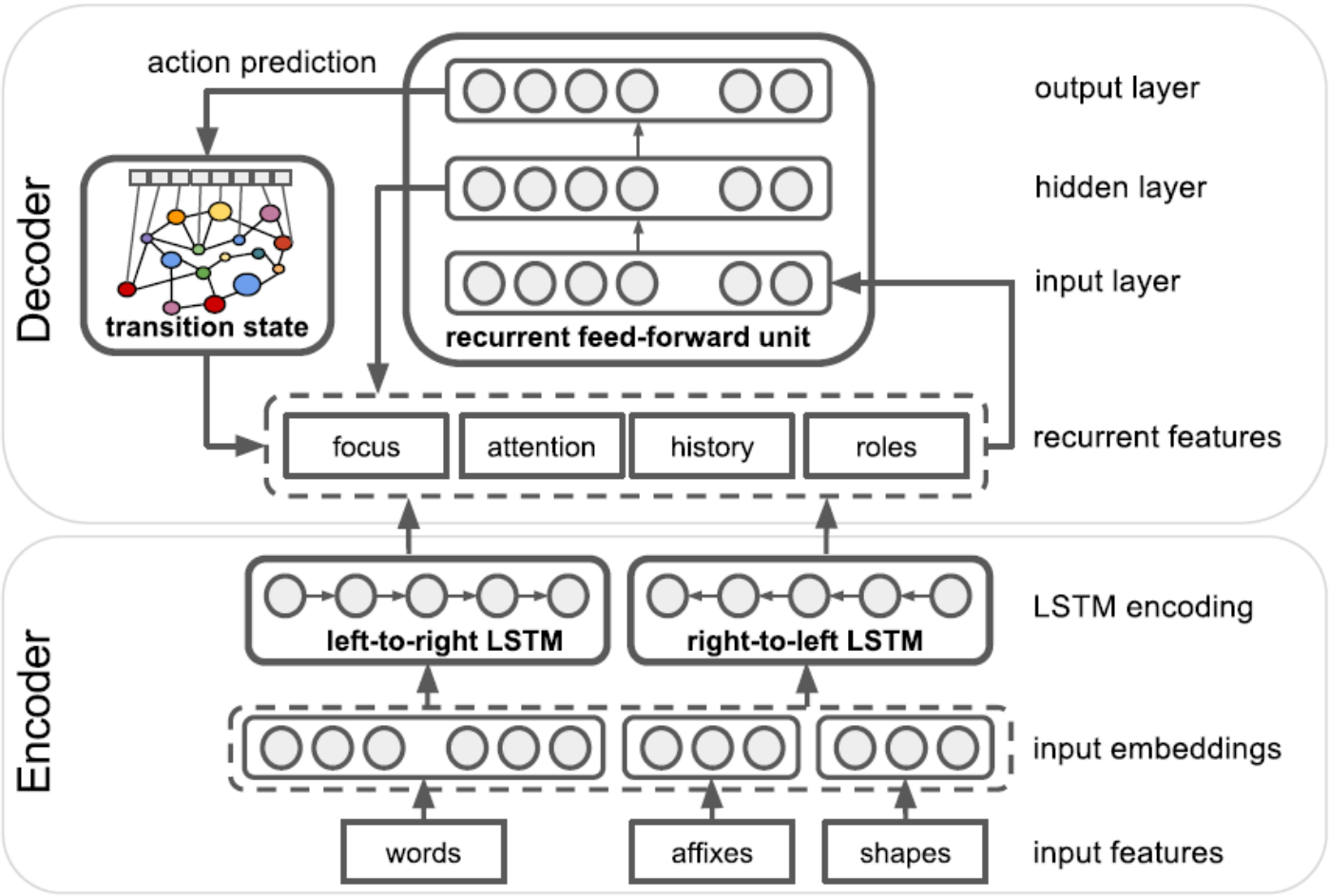


Figure 1: Neural network architecture of the SLING parser. The input is encoded by a bi-directional LSTM and fed into a recurrent feed-forward (FF) unit that proposes transition system actions. The hidden layer activations and the transition system state are combined to create the input feature vector for the next step. The FF unit is run repeatedly until the transition system has reached a final state.

https://ai.googleblog.com/2017/11/sling-natural-language-frame-semantic.html