BOLT PROJECT OVERVIEW

John E. Laird

Soar Workshop

PIs:
Bob Bechtel, Soar Technology
John Laird, University of Michigan
Edwin Olson, University of Michigan

DARPA BOLT, Activity E

(Broad Operational Language Translation)

"Grounded Language Acquisition: research in deep semantic language acquisition using robotic visual and tactile information as input for experiential learning of objects actions and learning of objects, actions, and consequences."

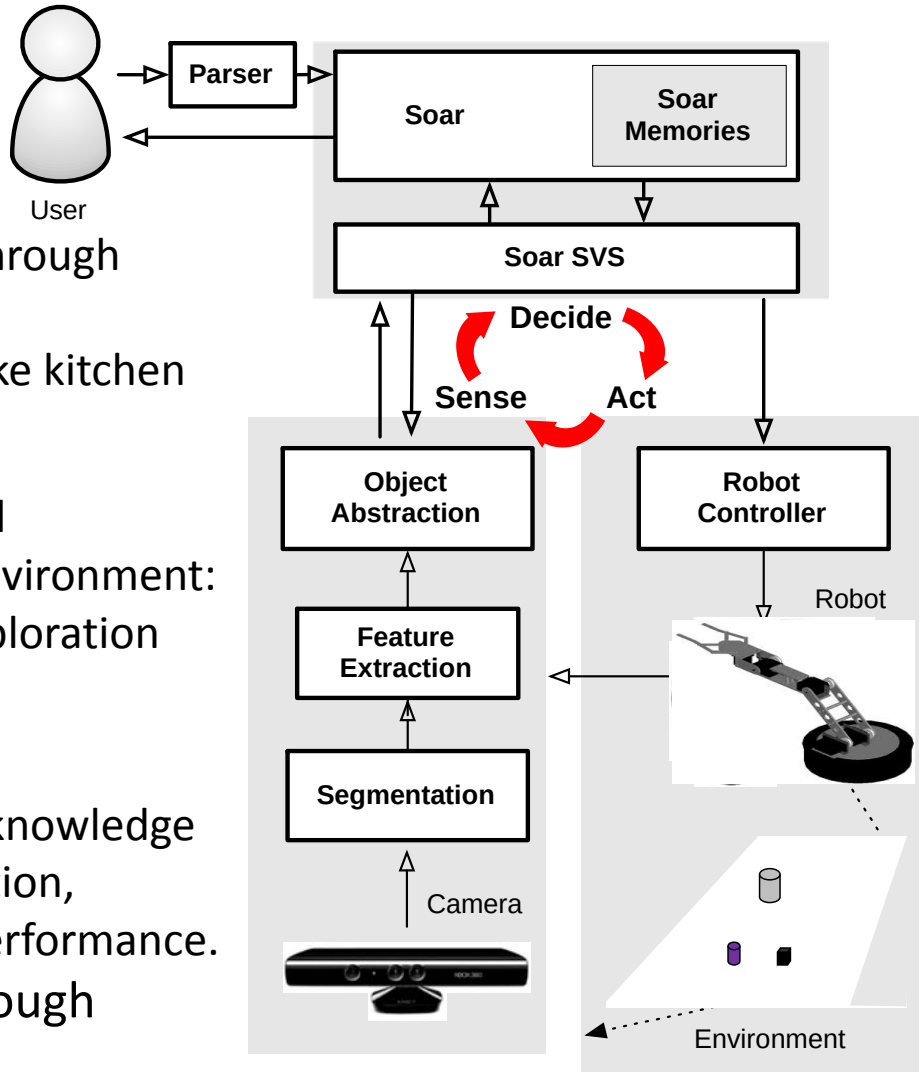
- Five-year project
- Five funded groups
  - Soar Tech/Michigan, Arizona, Berkeley, MIT, Rochester
  - Wide variety of approaches being pursued
  - We are only one using a cognitive architecture
- Why are we doing it?
  - Another step in integrating Soar with real world.
  - Another step in using and learning natural language.
  - Another step in using instruction to build up knowledge.

Our Research Goal

- *Fast* and *robust* language learning that is *grounded* in ongoing experience.
  - Learn adjectives/nouns, prepositions, verbs in a real-world robotic environment
  - Learn during task performance – dynamically extending language
- Approach:
  - Situated Interactive Instruction
  - Supplemented by supervised, unsupervised, reinforcement, and analytic learning mechanisms
  - In a cognitive architecture (Soar!)
  - Learning across production, episodic, and semantic memories.

Approach: Situated Interactive Instruction

- *Instruction-based learning*
  - Human guides language acquisition through language
  - Initial task is object movement in a fake kitchen
- *Situated* Instruction
  - Instruction grounded in the real world
  - Agent learns while interacting with environment: guided by instruction and through exploration
- *Interactive* Instruction
  - Agent asks for help when it needs it
  - Human corrects and extends agent's knowledge
  - Human can aid with language acquisition, instruction interpretation, and task performance.
- Understanding is demonstrated through performance



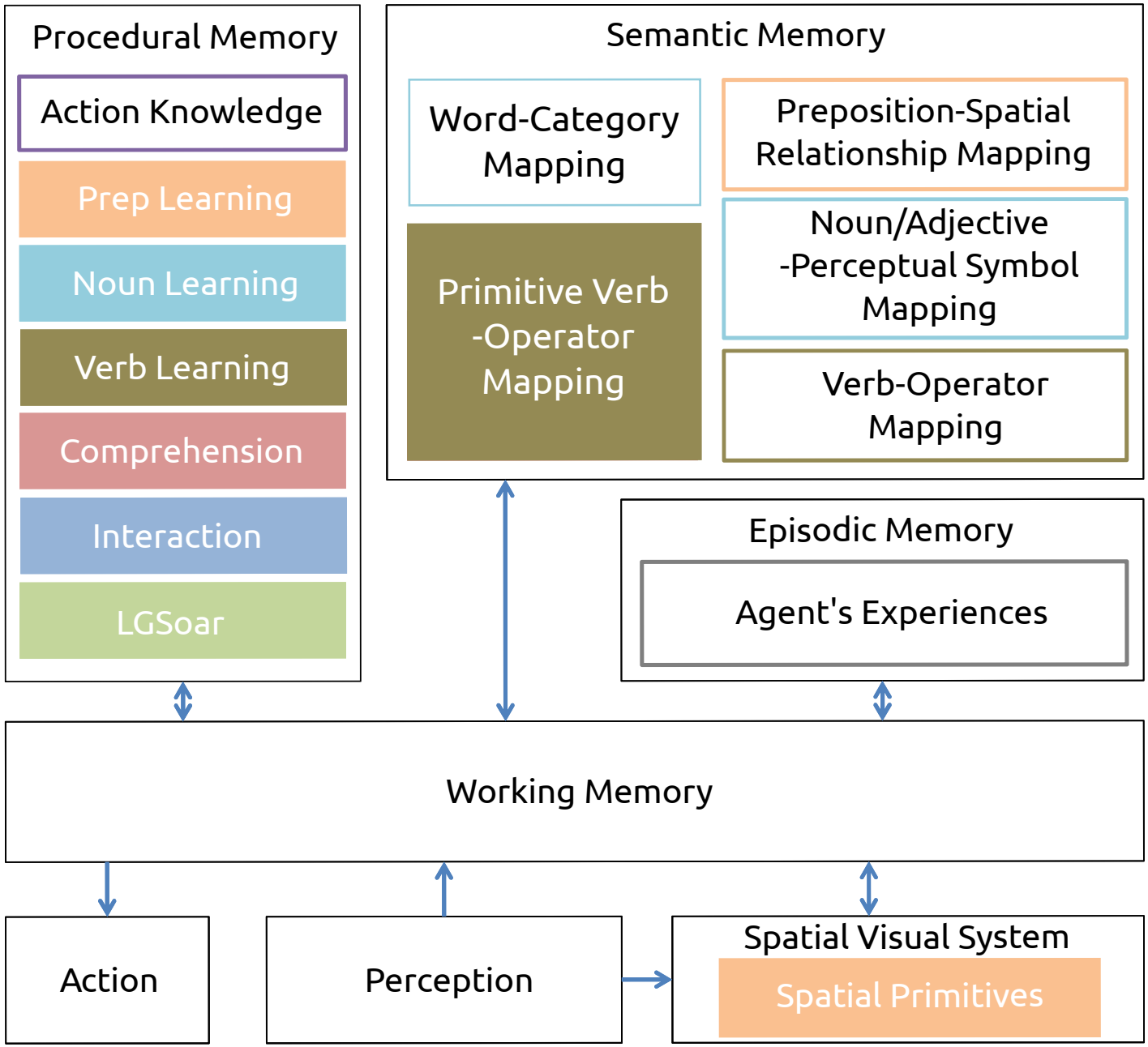
Environment and Agent Interface

- Test Environment
  - Tabletop with robot arm at center
  - Identified regions
  - Foam blocks of various shapes and colors
- System physical capabilities (and in simulation)
  - Visual sensing:
    - Color camera and Kinect. XYZ location, color, size, shape
  - Arm can cover about 330° and has max range of about one foot
- Primitive commands – point-to, pick-up(ObjID), place-at(X,Y,Z)
- Basic syntactic knowledge (sentence structure) is built in

Simple Instruction Examples

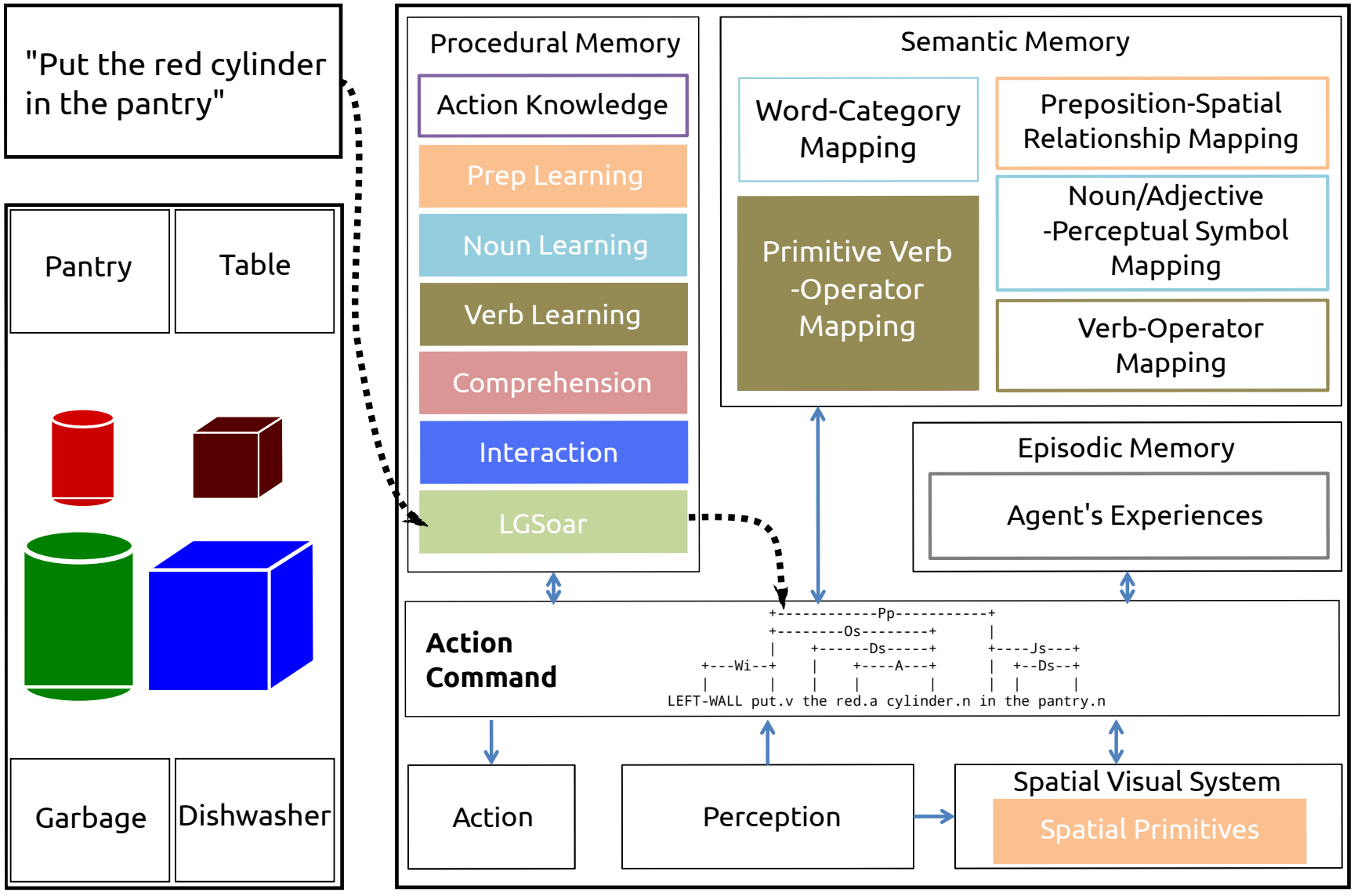
- Learn new nouns and adjectives
  - "This is a red triangle"
- Learn new prepositions that map to spatial relations
  - "The red triangle is right-of *the* blue sphere"
- Demonstrate what is has learned
  - "Describe the scene."
- Execute primitive verb using learned features
  - "Put the red cylinder in the pantry"
- Learn new verb that is composition of primitive verbs
  - "Move the red triangle to the pantry"

Agent Overview



Phase I: Syntactical Processing

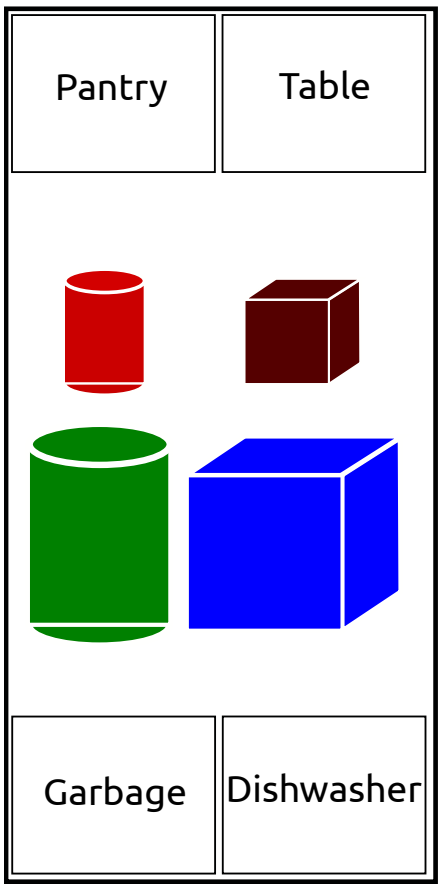
Language Parsing and Sentence Categorization - Sam Wintermute

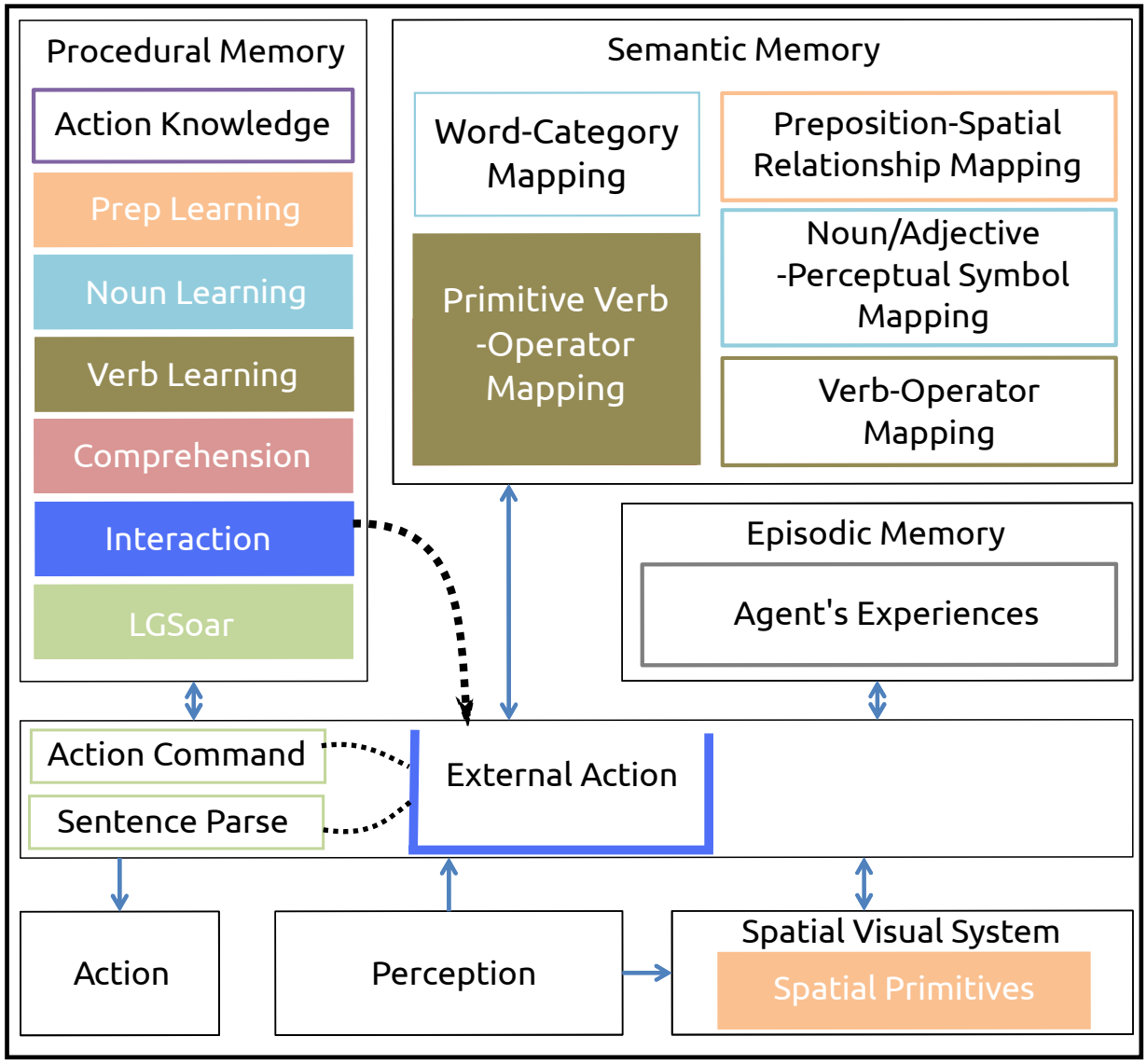


Phase II: Interaction Managment

Interaction Model - Shiwali Mohan

"Put the red cylinder in the pantry"

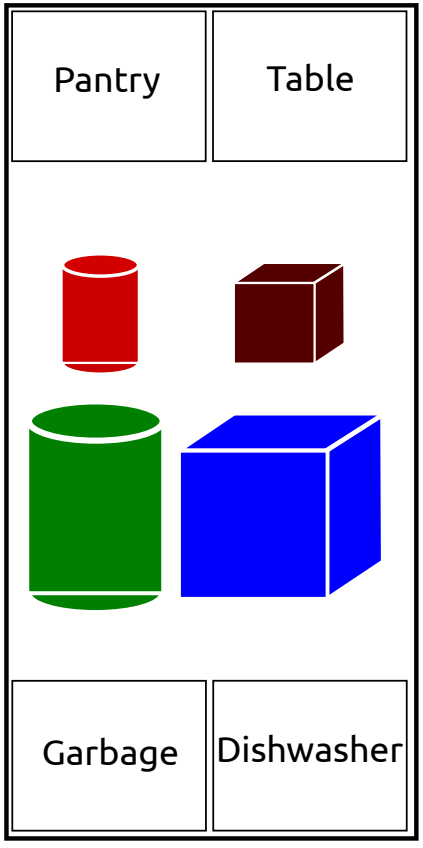


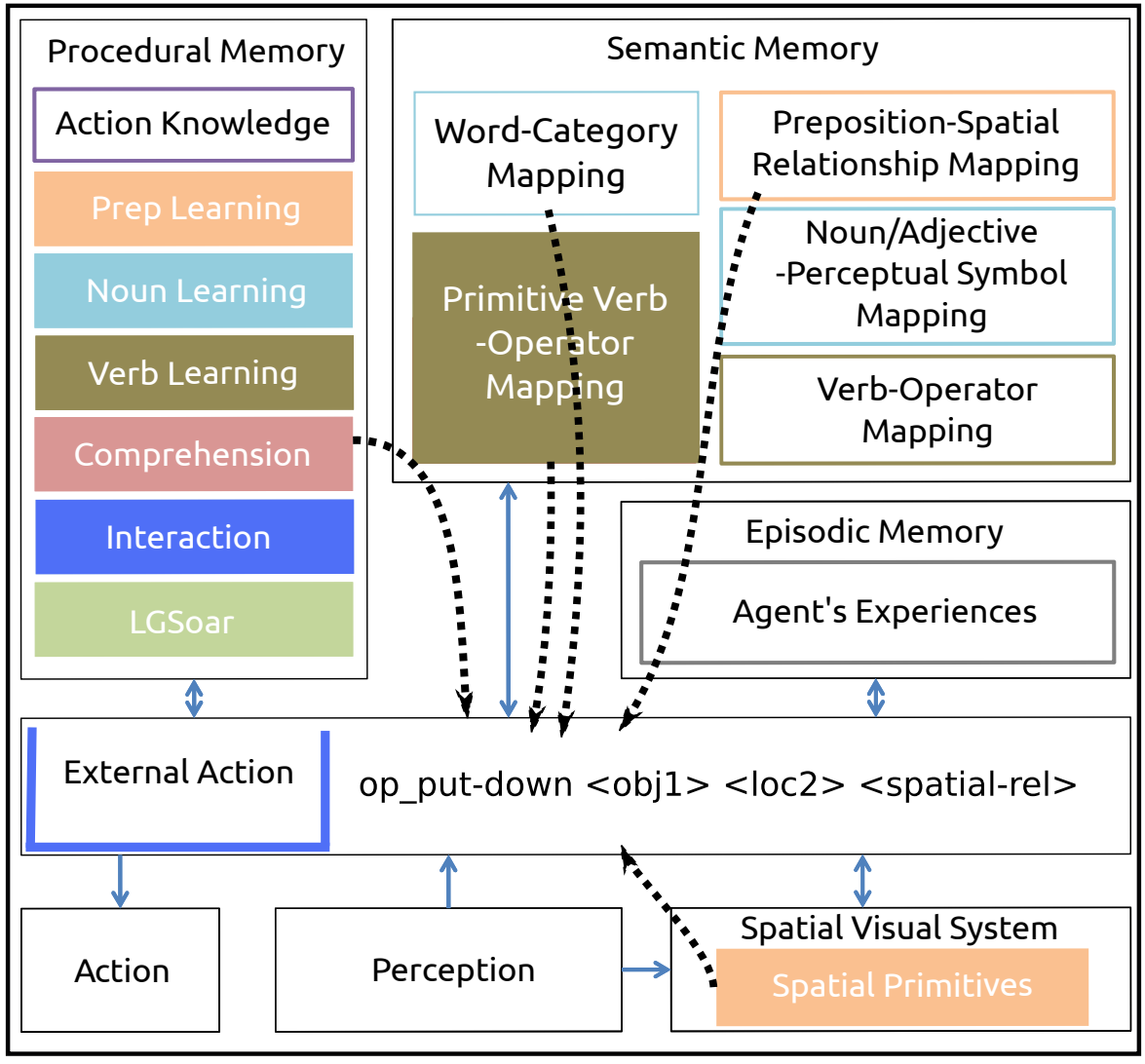


Phase III: Grounded Comprehension

Noun/Adjective Comprehension and Learning - Aaron Mininger
Preposition Comprehension Learning - James Kirk
Situated Comprehension of Action Commands - Shiwali Mohan

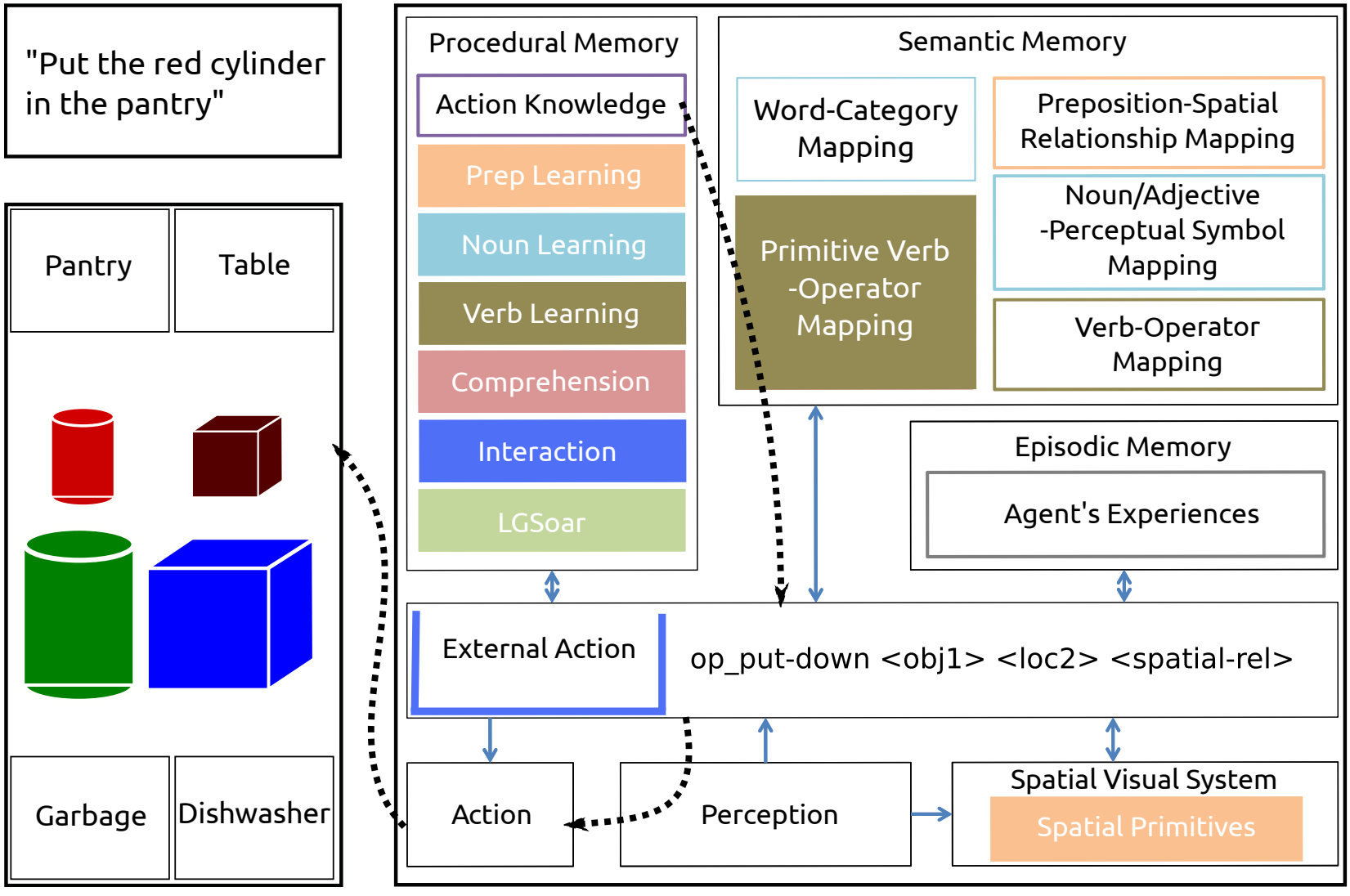
"Put the red cylinder in the pantry"





Phase IV: Behavior Execution

Verb Learning - Shiwali Mohan



Demonstration Tonight!