Using Situated Context for Language Comprehension in Cognitive Instructable Agents

Shiwali Mohan

Computer Science and Engineering
University of Michigan

July 23, 2012

Outline

Goal

Cognitive agent that acquires various kinds of knowledge through situated, interactive, instruction

- Various kinds of knowledge
  - Object recognition, categorization
  - Spatial Relationships
  - Task decomposition and execution
  - Linguistic knowledge - nouns, adjectives, preposition, verbs

Goal

Cognitive agent that acquires various kinds of knowledge through situated, interactive, instruction

- Various kinds of knowledge
  - Object recognition, categorization
  - Spatial Relationships
  - Task decomposition and execution
  - Linguistic knowledge - nouns, adjectives, preposition, verbs

- Instruction
  - Robotic agent embodied in the real world
  - Natural language instruction
    - Constrained grammar, parser available
  - Instructions accompany sensory information, guides action

Goal

Cognitive agent that acquires various kinds of knowledge through situated, interactive, instruction

- Various kinds of knowledge
  - Object recognition, categorization
  - Spatial Relationships
  - Task decomposition and execution
  - Linguistic knowledge - nouns, adjectives, preposition, verbs

- Instruction
  - Robotic agent embodied in the real world
  - Natural language instruction
    - Constrained grammar, parser available
  - Instructions accompany sensory information, guides action

- Comprehension for learning

Motivation

- Linguistic communication is contextual
    - Linguistic symbols by themselves provide limited meaning
        - *The red block is on the table.*
    - Listener uses situation, knowledge to associate meaning
    - Pronouns; ill-formed, incomplete sentences

Motivation

- Linguistic communication is contextual
    - Linguistic symbols by themselves provide limited meaning
        - *The red block is on the table.*
    - Listener uses situation, knowledge to associate meaning
    - Pronouns; ill-formed, incomplete sentences
- Theories of cognitive evolution
    - Language evolved for social coordination, collaborative tasks, learning from communication

Motivation

- Linguistic communication is contextual
  - Linguistic symbols by themselves provide limited meaning
    - *The red block is on the table.*
  - Listener uses situation, knowledge to associate meaning
  - Pronouns; ill-formed, incomplete sentences
- Theories of cognitive evolution
  - Language evolved for social coordination, collaborative tasks, learning from communication
- Context for linguistic processing
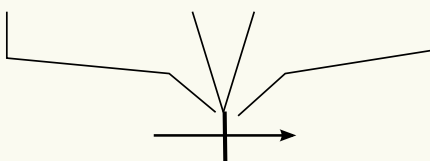
Requirements

- Referential
    - R1 Refer to objects, actions, events, spatial relationships in the environment.
    - R2 Refer to agent's experiences.
    - R3 Refer to knowledge about the world.

- Integrative
    - I1 Initiate interactions in case of incomplete information, unknown symbols.
    - I2 Initiate learning for unknown symbols.

- Incremental
    - C1 With experience expand linguistic knowledge.
    - C2 With experience expand non-linguistic knowledge.

Related Work

- Grounded Language Comprehension
  - SHRDLU (Winograd, 1972) [R1]
  - Using lambda calculus (Cantrell et al., 2010) [R1]
  - Using external knowledge base - FrameNet (Goertzel et al., 2010) [R1,R3]
  - Ground action commands (Tellex et al., 2011) [R1,R3,C1,C2]

Related Work

- Grounded Language Comprehension
    - SHRDLU (Winograd, 1972) [R1]
    - Using lambda calculus (Cantrell et al., 2010) [R1]
    - Using external knowledge base - FrameNet (Goertzel et al., 2010) [R1,R3]
    - Ground action commands (Tellex et al., 2011) [R1,R3,C1,C2]

- Grounded Language Acquisition
    - Noun acquisition from labeled pictures (Barnard et al., 2003; Gupta and Davis, 2008), computer-generated visual scenes (Roy, 2002) [C1,C2]
    - Acquisition of spatial relationships (Kollar et al., 2010) [R1,R3,C1,C2]
    - Grounding verbs in visual perception (Siskind, 2001) [C2]

Related Work

- Grounded Language Comprehension
  - SHRDLU (Winograd, 1972) [R1]
  - Using lambda calculus (Cantrell et al., 2010) [R1]
  - Using external knowledge base - FrameNet (Goertzel et al., 2010) [R1,R3]
  - Ground action commands (Tellex et al., 2011) [R1,R3,C1,C2]

- Grounded Language Acquisition
  - Noun acquisition from labeled pictures (Barnard et al., 2003; Gupta and Davis, 2008), computer-generated visual scenes (Roy, 2002) [C1,C2]
  - Acquisition of spatial relationships (Kollar et al., 2010) [R1,R3,C1,C2]
  - Grounding verbs in visual perception (Siskind, 2001) [C2]

- Learning with Interactive Instruction
  - Impasse based (Huffman and Laird, 1995) [I1,I2,C2]

Approach

- Domain
    - Table top robot, locomotive simulated robot
- Agent Design
    - Cognitive architecture, Soar
- Goal
    - Comprehend action-commands by associating words to non-linguistic context [R1,R2,R3]
        - Psycholinguistic theory: The Indexical Hypothesis
        - *Pick up the red block, Put the red block in the pantry.*
    - Leverage context to accumulate knowledge [C1,C2]
        - Learn action verbs (composition of known primitives), instructed trial
        - *Move the red block to the pantry.*
    - Embedded in a interactive instruction system.
        - not in thesis
        - Facilitates [I1, I2]
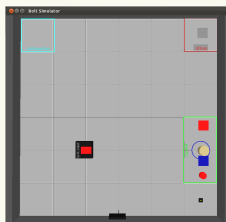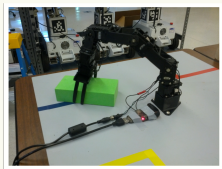
Domain Overview
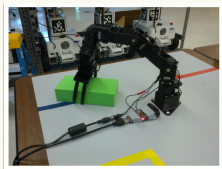
Figure: BOLT Robot

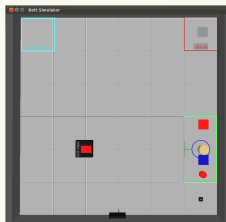Figure: BOLT++

Domain Overview

- Perception
  - Sensor: Kinect camera
  - Segments the scene into objects, locations
  - Feature vector: (*color*, *shape*, *size*) $\rightarrow$ symbolic representation
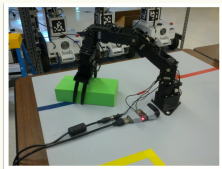

Figure: BOLT Robot

Figure: BOLT++

Domain Overview

- Perception
    - Sensor: Kinect camera
    - Segments the scene into objects, locations
    - Feature vector: (*color*, *shape*, *size*) → symbolic representation

- Controller
    - Pre-programmed behaviors
        - Object manipulation: `pick-up(object)`, `put-down(object,location)`
        - Locomotion: `goto(location)`
        - Functional: `turn-on(stove)]`, `close(pantry)` …
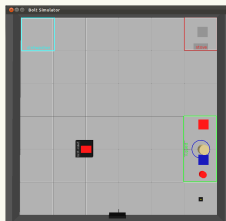

Figure: BOLT Robot

Figure: BOLT++

Domain Overview

- Perception
  - Sensor: Kinect camera
  - Segments the scene into objects, locations
  - Feature vector: (*color*, *shape*, *size*) → symbolic representation

- Controller
  - Pre-programmed behaviors
    - Object manipulation: `pick-up(object)`, `put-down(object,location)`
    - Locomotion: `goto(location)`
    - Functional: `turn-on(stove)]`, `close(pantry)` …

- Instructor Interface
  - Chat interface, *point*
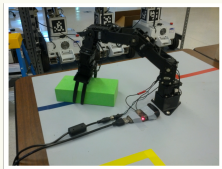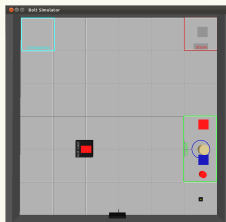 - link-grammar parser, POS tags, sentence structure

Figure: BOLT Robot

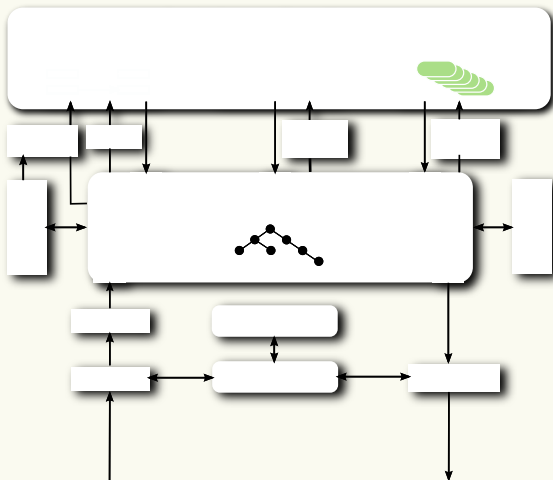Figure: BOLT++

Agent Architecture
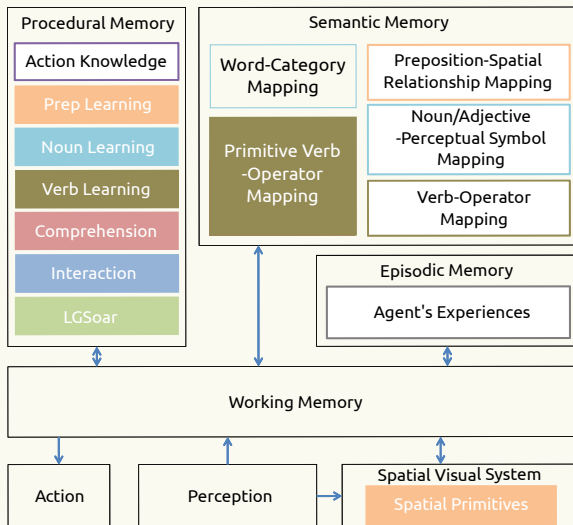

Figure: Soar (Laird, 2012)

Agent Design

Figure: Agent Overview (Mohan et al., 2012)

The Indexical Hypothesis

Glenberg and Robertson (1999)

- Sentences become meaningful by grounding their interpretation in situated action
  - *Index* words and phrases to referents.
  - Derive *affordances* from these referents.
    - *affordances* are derived from domain background knowledge
  - *Mesh* these affordances under syntactical constraints, physical constraints of the environment
  - *I painted my room orange.* v/s *I ate an orange.*

The Indexical Hypothesis

Glenberg and Robertson (1999)

- Sentences become meaningful by grounding their interpretation in situated action
    - *Index* words and phrases to referents.
    - Derive *affordances* from these referents.
        - *affordances* are derived from domain background knowledge
    - *Mesh* these affordances under syntactical constraints, physical constraints of the environment
    - *I painted my room orange.* v/s *I ate an orange.*

- In general,
    - Linguistic information specifies a scene.
    - Affordances (non-linguistic knowledge) specify the details for taking action.

Using Language for Indexing

Barsalou (1999)

Using Language for Indexing

Barsalou (1999)

- Immediate Indexing
    - *"the white screen"*
    - Participants are simultaneously embedded in the environment
    - Language is used to refer to objects and event in the current situation.
    - Used for situated examples from the current situation

Using Language for Indexing
Barsalou (1999)

- Immediate Indexing
  - *"the white screen"*
  - Participants are simultaneously embedded in the environment
  - Language is used to refer to objects and event in the current situation.
  - Used for situated examples from the current situation

- Displaced Indexing
  - *"the entrance of the Sherator Center"*
  - Referents are not currently present
  - Language is used to refer to objects and events from prior experiences with the environment
  - *shared, componential, future*

Requirement Analysis

- Indexing provides information [R1,R2,R3]
    - Index to perceptions (*immediate indexing*)
    - Index to experience, semantic knowledge (*displaced indexing*)
    - Use procedural (action-execution) knowledge (*derive affordances*)

Requirement Analysis

- Indexing provides information [R1,R2,R3]
    - Index to perceptions (*immediate indexing*)
    - Index to experience, semantic knowledge (*displaced indexing*)
    - Use procedural (action-execution) knowledge (*derive affordances*)
- Indexing facilitates interaction [I1]
    - Indexing failure (multiple indices)

Requirement Analysis

- Indexing provides information [R1,R2,R3]
    - Index to perceptions (*immediate indexing*)
    - Index to experience, semantic knowledge (*displaced indexing*)
    - Use procedural (action-execution) knowledge (*derive affordances*)
- Indexing facilitates interaction [I1]
    - Indexing failure (multiple indices)
- Indexing facilitates situated learning [I2]
    - Indexing failure (no indices)

Current Status

Outline

- Formulation
- Capabilities
    - Ambiguity Resolution
        - Noun Phrase → Situated Referent Resolution
        - Verb → Instantiated Action Resolution
    - Learning Verbs
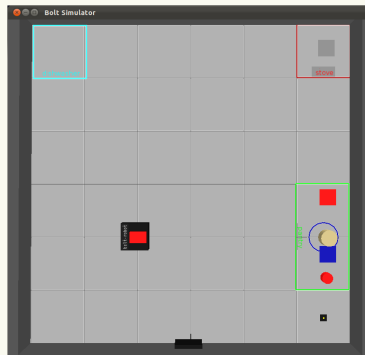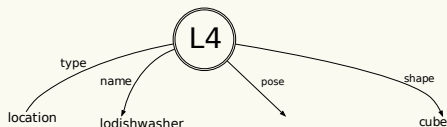   - From current scenarios

Figure: BOLT++

Formulation

Mohan and Laird (2012a)

- Background Knowledge

  (Noun/Adj:Perceptual Symbols, Preposition:Spatial Relationship)
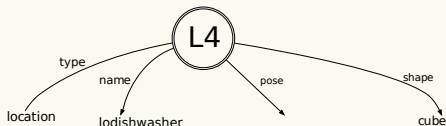
  - Domain semantic knowledge

  - (can be) Acquired from previous experiences
  - Allows the agent to communicate non-perceptible locations

Formulation

Mohan and Laird (2012a)

- Background Knowledge
  (Noun/Adj:Perceptual Symbols, Preposition:Spatial Relationship)
  - Domain semantic knowledge

    - (can be) Acquired from previous experiences
    - Allows the agent to communicate non-perceptible locations
  - Verb-action mapping

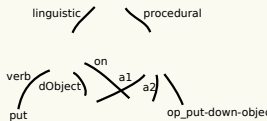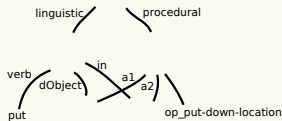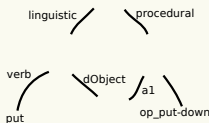   - (can be) Acquired (described later)

Formulation

Mohan and Laird (2012a)

- Background Knowledge

  (Noun/Adj:Perceptual Symbols, Preposition:Spatial Relationship)

  - Domain semantic knowledge

  - (can be) Acquired from previous experiences
  - Allows the agent to communicate non-perceptible locations
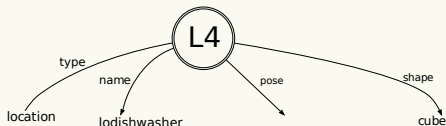
  - Verb-action mapping
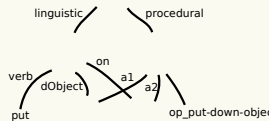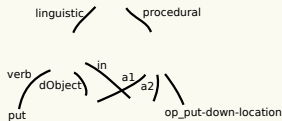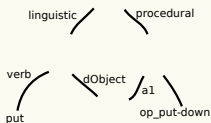

  - (can be) Acquired (described later)

- Primitive action knowledge

Formulation
Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

Formulation

Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

- Immediate Indexing
  - match description to perceptible objects [R1]
  - add all matching objects to arg-candidate set
    $CA_{dO} = \{A, C\}$
  - if nothing matches, displaced indexing

Perceptual Symbols

color: blue size: large shape: cube
name: cube pose: <x> <y> <z> <y> <p>

A B C D

Pantry

Formulation

Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

- Immediate Indexing
  - match description to perceptible objects [R1]
  - add all matching objects to arg-candidate set $CA_{dO} = \{A, C\}$
  - if nothing matches, displaced indexing
- Displaced Indexing
  - match description to semantic objects [R3]
  - add all to arg-candidate set $CA_{in} = \{dishwasher\}$

Perceptual Symbols

color: blue size: large shape: cube
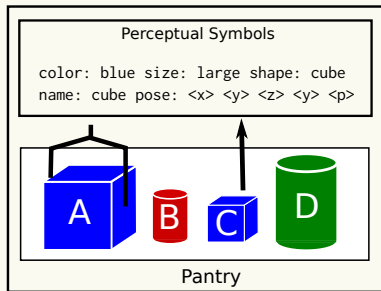name: cube pose: <x> <y> <z> <y> <p>

Pantry

Formulation
Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

- Immediate Indexing
  - match description to perceptible objects [R1]
  - add all matching objects to arg-candidate set
    $CA_{dO} = \{A, C\}$
  - if nothing matches, displaced indexing

- Displaced Indexing
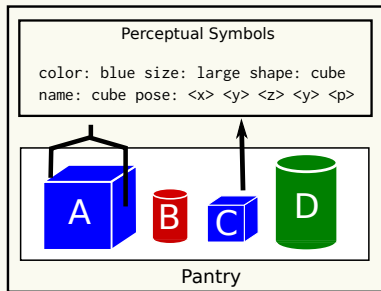 - match description to semantic objects [R3]
  - add all to arg-candidate set
    $CA_{in} = \{dishwasher\}$

- If $CA = \phi$, communicate [I1]

Formulation
Mohan and Laird (2012a)
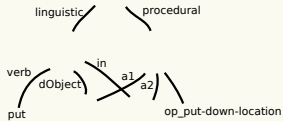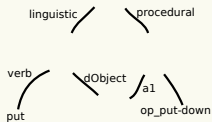
"Put a blue cube in the dishwasher"
Phase: **Index Verb**

Formulation
Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Index Verb**

- Possible Interpretations *CI*
  ```
  op_put-down-object-location [A] [dishwasher]
  op_put-down-object-location [C] [dishwasher]
  ```

Formulation
Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb

Formulation
Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - op_put-down-object-location [A] [dishwasher]

Formulation
Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - `op_put-down-object-location [A] [dishwasher]`
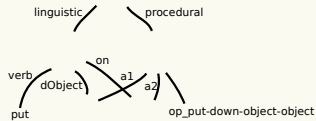- Compute the intersection
  - $CI \cap PA$

Formulation

Mohan and Laird (2012a)

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - op_put-down-object-location [A] [dishwasher]
- Compute the intersection
  - *CI ∩ PA*
- Excute action if only one element
  - Ground preposition: predicate projection, tracking

Formulation
Mohan and Laird (2012a)
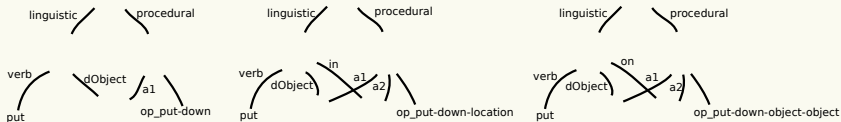
"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - op_put-down-object-location [A] [dishwasher]
- Compute the intersection
  - *CI* ∩ *PA*
- Excute action if only one element
  - Ground preposition: predicate projection, tracking
- Communicate if empty or multiple elements [I1,I2]

Situated Referent Resolution

"Put the blue large cube in the dishwasher"
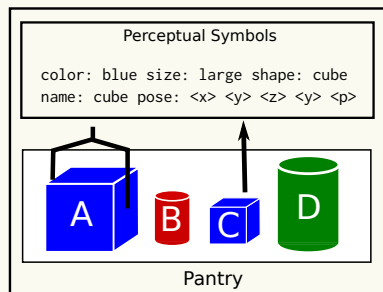
Situated Referent Resolution

"Put the blue cube in the dishwasher"
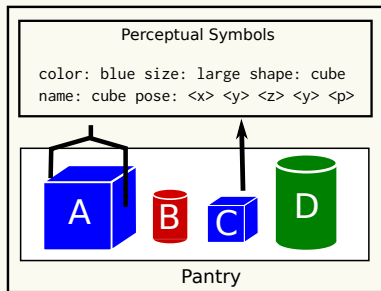
Situated Referent Resolution

"Put the object in the dishwasher"

Perceptual Symbols

color: blue size: large shape: cube
name: cube pose: <x> <y> <z> <y> <p>

Pantry

Situated Referent Resolution

- Reference resolution ambiguity
  - Constraints: linguistic, perceptual, procedural
  - Context provides information
    - Assume NP -> maps to an object
      H(NP='object'|perceptual context) =
      $-(0.25 log_2 0.25) * 4 = 2$
      H(NP='object'|procedural) = $-(1 log_2 1) = 0$

Perceptual Symbols

color: blue size: large shape: cube
name: cube pose: \<x> \<y> \<z> \<y> \<p>

A B C D

Pantry
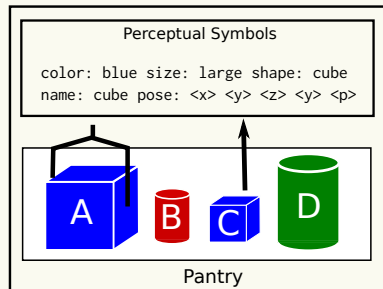
Situated Referent Resolution

- Reference resolution ambiguity
  - Constraints: linguistic, perceptual, procedural
  - Context provides information
    - Assume NP -> maps to an object
      H(NP='object'|perceptual context) =
      $-(0.25 log_2 0.25) * 4 = 2$
      H(NP='object'|procedural) = $-(1 log_2 1) = 0$

- Incomplete observations
  - Semantic knowledge of the domain

Perceptual Symbols

color: blue size: large shape: cube
name: cube pose: <x> <y> <z> <y> <p>
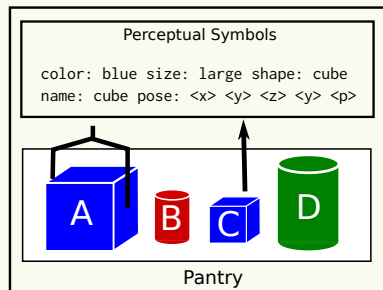
Pantry

Action Resolution
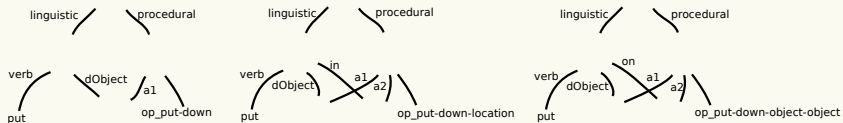
- Use argument structure of the command

"Put a blue cube in the dishwasher"

Learning Verbs in Current Scenarios

(Mohan and Laird, 2011; Mohan et al., 2012)

- Learning action verbs

  Move the red block to the pantry.

- Linguistic knowledge [C1]
  - Map verb and argument structure to action (operator) and arguments, preposition to spatial relationship
  - Acquisition
    - Reliance on language parsing system

- Learning semantic knowledge [C2]
  - Goal conditions to verb arguments, spatial relationships
  - Acquistion
    - Explicit interaction

- Learning procedural knowledge [C2]
  - Decide, execute and terminate an action in the environment
  - Acquisiton
    - Instructed trial
    - Forward projection, chunking

Future Work
Outline

- Ambiguity
  - Prepositional Phrase Attachment Resolution
- Diverse Context
  - Dialog Context
  - Comprehension by Perceptual Simulation
- Learning
  - Learning with Hypothetical Scenarios

Dialog Context for Referent Resolution

- Interaction Model (Mohan and Laird, 2012b; Mohan et al., 2012)
  maintains the context of the dialog

  Instructor: *Move the red block to the pantry.*
  Agent: *What is the next action I should take?*
  Instructor: *Pick up the red block.*

  . . .

Dialog Context for Referent Resolution

- Interaction Model (Mohan and Laird, 2012b; Mohan et al., 2012) maintains the context of the dialog

  Instructor: *Move the red block to the pantry.*
  Agent: *What is the next action I should take?*
  Instructor: *Pick up the red block.*

  . . .

- *Pick up the object. / Pick it up.*
  - Referent established by the context of the dialog.

Prepositional Phrase Attachment Resolution

Put the red large block on the left of green cylinder in the pantry.

Prepositional Phrase Attachment Resolution

Put the red large block on the left of green cylinder in the pantry.

- Multiple attachments

  Put the red large block on the left of green cylinder in the pantry.
  Put the red large block on the left of green cylinder in the pantry.

Prepositional Phrase Attachment Resolution

Put the red large block on the left of green cylinder in the pantry.

- Multiple attachments

  Put the red large block on the left of green cylinder in the pantry.
  Put the red large block on the left of green cylinder in the pantry.

- Context
  - Linguistic context does not provide enough information
  - Use situational context

Prepositional Phrase Attachment Resolution

Put the red large block on the left of green cylinder in the pantry.

- Multiple attachments

  Put the red large block on the left of green cylinder in the pantry.
  Put the red large block on the left of green cylinder in the pantry.

- Context
  - Linguistic context does not provide enough information
  - Use situational context
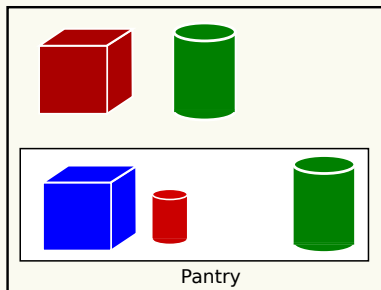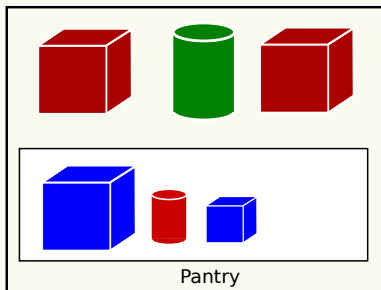
Comprehension by Perceptual Simulation

I bought an umbrella yesterday. It is purple.

Comprehension by Perceptual Simulation

I bought an umbrella yesterday. It is purple.
Does it have a handle?

Comprehension by Perceptual Simulation

I bought an umbrella yesterday. It is purple.
Does it have a handle?

- Displaced indexing: resolving to not present referents
  - Semantic memory: general knowledge about the domain (*the kitchen*)
  - Episodic memory: specific experiences (*the location you last visited*)

Comprehension by Perceptual Simulation

I bought an umbrella yesterday. It is purple.
Does it have a handle?

- Displaced indexing: resolving to not present referents
  - Semantic memory: general knowledge about the domain (*the kitchen*)
  - Episodic memory: specific experiences (*the location you last visited*)
- Mental Imagery
  - *Put the block in the location you last visited.*

Learning with Hypothetical Scenarios

To go to Kroger, take a left on to Plymouth road, drive for a mile and take a left into the complex.

Learning with Hypothetical Scenarios

To go to Kroger, take a left on to Plymouth road, drive for a mile and take a left into the complex.

- Create hypothetical learning scenarios
  Imagine a red block. Imagine a kitchen. Move the red block to the kitchen. ...

Learning with Hypothetical Scenarios

To go to Kroger, take a left on to Plymouth road, drive for a mile and take a left into the complex.

- Create hypothetical learning scenarios
  Imagine a red block. Imagine a kitchen. Move the red block to the kitchen. ...
- Prior experience with the world to imagine
  - `red block', `kitchen'

Learning with Hypothetical Scenarios

To go to Kroger, take a left on to Plymouth road, drive for a mile and take a left into the complex.

- Create hypothetical learning scenarios
  Imagine a red block. Imagine a kitchen. Move the red block to the kitchen. ...
- Prior experience with the world to imagine
  - `red block', `kitchen'
- Learn by `situating' instruction in the imagined world.

To Summarize

- Identified requirements for comprehension for situated learning
  - Referential, Integrative, Incremental
- Evaluated The Indexical Hypothesis
- Formulated for cognitive agents
  - Implemented in Soar
- Discussed the capabilities
  - Ambiguity resolution
  - Information augmentation
  - Learning

Demo

Acknowledgments

Advisor: Professor John E. Laird, University of Michigan
Graduate Student Collaborators: Aaron Mininger, James Kirk
Collaborating Organizations: APRIL at UMich, SoarTechnology
Funding Agency: Defense Advanced Research Projects Agency

Situated Comprehension

Mohan and Laird, 2012

- Contributions: formulation
- Background Knowledge
  (Noun/Adj:Perceptual Symbols, Preposition:Spatial Relationship)
  - Domain Semantic Knowledge
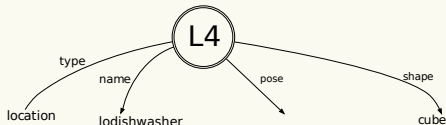
  - May be acquired from previous experiences
  - Allows the agent to communicate non-perceptible locations

Situated Comprehension

Mohan and Laird, 2012

- Contributions: formulation
- Background Knowledge
  (Noun/Adj:Perceptual Symbols, Preposition:Spatial Relationship)
  - Domain Semantic Knowledge
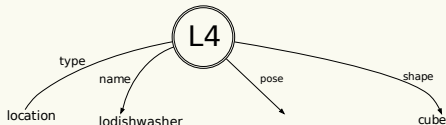
  - May be acquired from previous experiences
  - Allows the agent to communicate non-perceptible locations

- Verb-Action Mapping
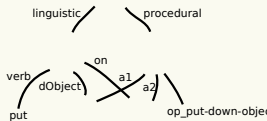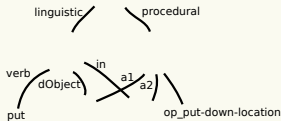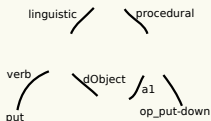
Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

- Immediate Indexing
  - match description to perceptible objects
  - add all matching objects to arg-candidate set
    $CA_{dO} = \{A, C\}$
  - if nothing matches, displaced indexing

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

- Immediate Indexing
  - match description to perceptible objects
  - add all matching objects to arg-candidate set
    $CA_{dO} = \{A, C\}$
  - if nothing matches, displaced indexing
- Displaced Indexing
  - match description to semantic objects
  - add all to arg-candidate set
    $CA_{in} = \{dishwasher\}$

Perceptual Symbols

color: blue size: large shape: cube
name: cube pose: <x> <y> <z> <y> <p>

A   B   C   D

Pantry

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Index Arguments**

- Immediate Indexing
  - match description to perceptible objects
  - add all matching objects to arg-candidate set
    $CA_{dO} = \{A, C\}$
  - if nothing matches, displaced indexing

- Displaced Indexing
  - match description to semantic objects
  - add all to arg-candidate set
    $CA_{in} = \{dishwasher\}$
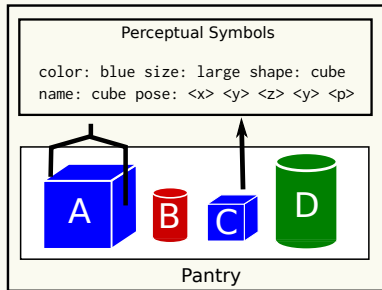
- If $CA = \phi$, communicate

Situated Comprehension

Mohan and Laird, 2012

*"**Put** a blue cube **in** the dishwasher"*
Phase: **Index Verb**
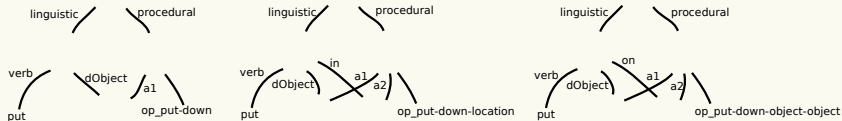
Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Index Verb**
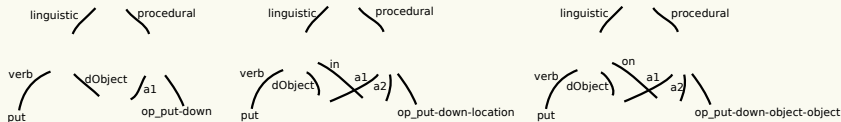
- Possible Interpretations *CI*

  op_put-down-object-location [A] [dishwasher]
  op_put-down-object-location [C] [dishwasher]

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
    - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
    - Using affordances, physical constraints
    - Soar proposals
    - `op_put-down-object-location [A] [dishwasher]`

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - Soar proposals
  - `op_put-down-object-location [A] [dishwasher]`
- Compute the intersection
  - *CI ∩ PA*

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - Soar proposals
  - op_put-down-object-location [A] [dishwasher]
- Compute the intersection
  - *CI* ∩ *PA*
- Excute action if only one element
  - Ground preposition: predicate projection, tracking

Situated Comprehension

Mohan and Laird, 2012

"Put a blue cube in the dishwasher"
Phase: **Meshing**

- Generate all possible interpretations *CI*
  - Using all candidate objects, *CA* for all arguments of the verb
- Generate all possible actions *PA*
  - Using affordances, physical constraints
  - Soar proposals
  - op_put-down-object-location [A] [dishwasher]
- Compute the intersection
  - *CI* ∩ *PA*
- Excute action if only one element
  - Ground preposition: predicate projection, tracking
- Communicate if empty or multiple elements

Linguistic Capabilities

Linguistic Capabilities

- Situated Referent Resolution
    - Using perceptual information
        - Use of most distinctive description given the perceptual state
    - Using semantic knowledge
        - mapping *the dishwasher* to a semantic object
    - Using procedural knowledge
        - Can only put down the object in gripper

Linguistic Capabilities

- Situated Referent Resolution
    - Using perceptual information
        - Use of most distinctive description given the perceptual state
    - Using semantic knowledge
        - mapping *the dishwasher* to a semantic object
    - Using procedural knowledge
        - Can only put down the object in gripper

- Situated Action Resolution
    - Using the argument structure

Required Knowledge

3 kinds of knowledge

Required Knowledge

3 kinds of knowledge

- Linguistic Knowledge
  - the verb word: *move*
  - the argument structure: *move the red block to the pantry*
  - extracted from the action command by the linguistic system
  - semantic memory

Required Knowledge

3 kinds of knowledge

- Linguistic Knowledge
  - the verb word: *move*
  - the argument structure: *move the red block to the pantry*
  - extracted from the action command by the linguistic system
  - semantic memory

- Procedural Knowledge
  - selection knowledge for breaking ties between known primitives
  - acquired through situated experience of performing the task (encoded in episodic memory)
  - procedural memory

Required Knowledge

3 kinds of knowledge

- Linguistic Knowledge
  - the verb word: *move*
  - the argument structure: *move the red block to the pantry*
  - extracted from the action command by the linguistic system
  - semantic memory

- Procedural Knowledge
  - selection knowledge for breaking ties between known primitives
  - acquired through situated experience of performing the task (encoded in episodic memory)
  - procedural memory

- Semantic Knowledge
  - goal predicates: `in{pantry, the red block}`
  - explicit, declarative description from the instructor: *"The goal is the red block is in the pantry"*.
  - semantic memory

Acquire Linguistic Mapping

[Failure: Grounded Comprehension Phase]

"Move the red cylinder to the pantry"

Pantry | Table

Garbage | Dishwasher

Procedural Memory

Action Knowledge

Prep Learning

Noun Learning

Verb Learning

Comprehension

Interaction

LGSoar

Semantic Memory

Word-Category Mapping

Preposition-Spatial Relationship Mapping

Noun/Adjective -Perceptual Symbol Mapping

Verb-Operator Mapping

Primitive Verb -Operator Mapping

Episodic Memory

Agent's Experiences

External Action

Action | Perception

Spatial Visual System

Spatial Primitives

Acquire Linguistic Mapping

[Failure: Grounded Comprehension Phase]

Acquire Linguistic Mapping

[Failure: Grounded Comprehension Phase]

Acquire Linguistic Mapping

[Failure: Grounded Comprehension Phase]

Acquire Situated Experience
[Failure: Behavior Execution Phase]

"Move the red cylinder to the pantry"

Pantry | Table

Garbage | Dishwasher

Procedural Memory
Action Knowledge
Prep Learning
Noun Learning
Verb Learning
Comprehension
Interaction
LGSoar

Semantic Memory
Word-Category Mapping
Preposition-Spatial Relationship Mapping
Primitive Verb -Operator Mapping
Noun/Adjective -Perceptual Symbol Mapping
Verb-Operator Mapping

Episodic Memory
Agent's Experiences

move ^direct-object obj1 ^to loc1
External Action

Action | Perception

Spatial Visual System
Spatial Primitives

Acquire Situated Experience

[Failure: Behavior Execution Phase]

"Move the red cylinder to the pantry"

Pantry | Table

Garbage | Dishwasher

Procedural Memory
- Action Knowledge
- Prep Learning
- Noun Learning
- Verb Learning
- Comprehension
- Interaction
- LGSoar

Semantic Memory
- Word-Category Mapping
- Preposition-Spatial Relationship Mapping
- Primitive Verb -Operator Mapping
- Noun/Adjective -Perceptual Symbol Mapping
- Verb-Operator Mapping

Episodic Memory
- Agent's Experiences

get-next-subaction
External Action

Action | Perception | Spatial Visual System
Spatial Primitives

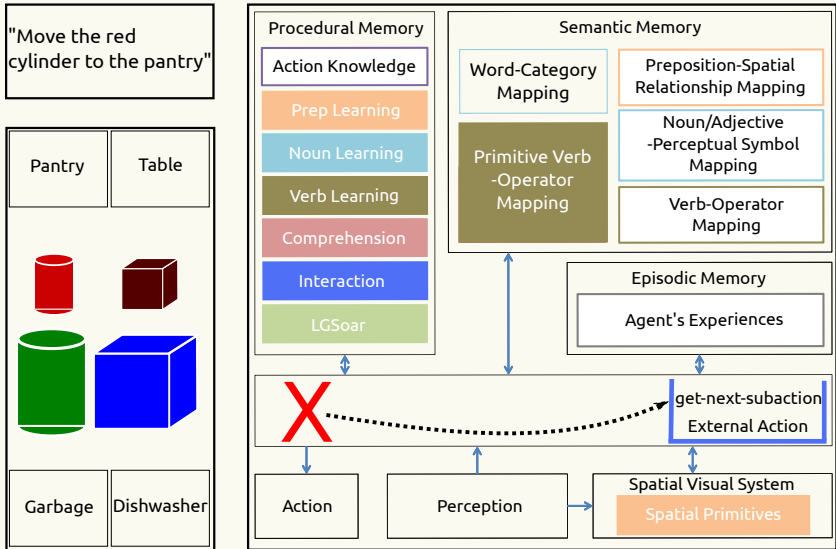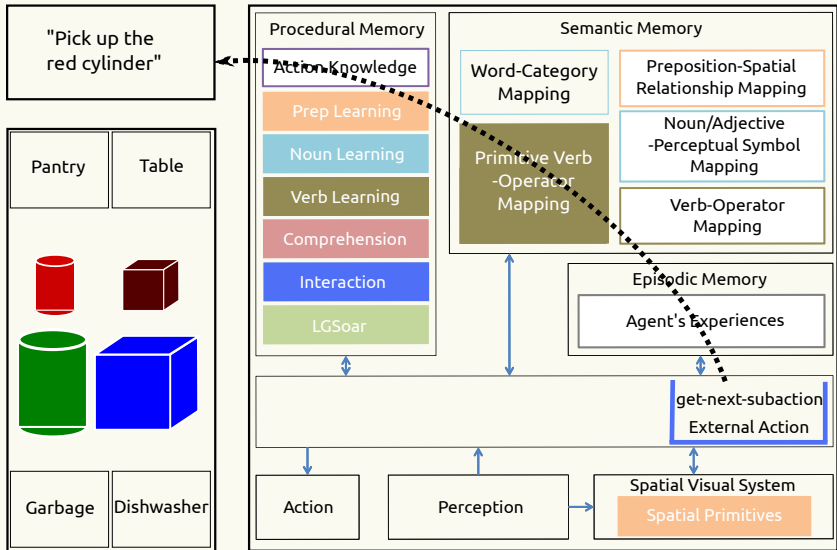# Acquire Situated Experience

[Failure: Behavior Execution Phase]

Acquire Situated Experience
[Failure: Behavior Execution Phase]

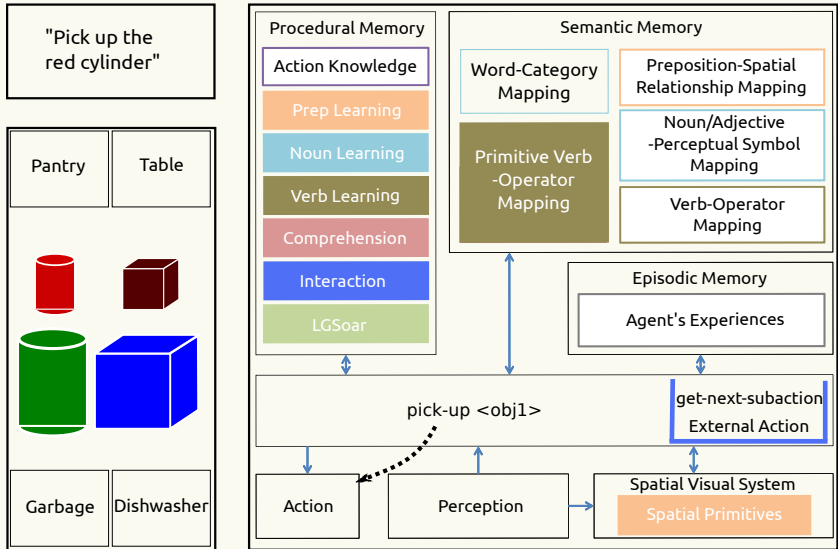"Pick up the red cylinder"

Pantry	Table

Procedural Memory
- Action Knowledge
- Prep Learning
- Noun Learning
- Verb Learning
- Comprehension
- Interaction
- LGSoar

Semantic Memory
- Word-Category Mapping
- Preposition-Spatial Relationship Mapping
- Noun/Adjective -Perceptual Symbol Mapping
- Primitive Verb -Operator Mapping
- Verb-Operator Mapping

Episodic Memory
- Agent's Experiences

pick-up <obj1>

get-next-subaction
External Action

Action	Perception

Spatial Visual System
- Spatial Primitives

Garbage	Dishwasher

Acquire Situated Experience

[Failure: Behavior Execution Phase]

"Pick up the red cylinder"

| Pantry | Table |

| Garbage | Dishwasher |

Procedural Memory
- Action Knowledge
- Prep Learning
- Noun Learning
- Verb Learning
- Comprehension
- Interaction
- LGSoar

Semantic Memory
- Word-Category Mapping
- Preposition-Spatial Relationship Mapping
- Noun/Adjective -Perceptual Symbol Mapping
- Primitive Verb -Operator Mapping
- Verb-Operator Mapping

Episodic Memory
- Agent's Experiences

pick-up <obj1>

get-next-subaction
External Action

Action

Perception

Spatial Visual System
- Spatial Primitives

Acquire Semantic knowledge
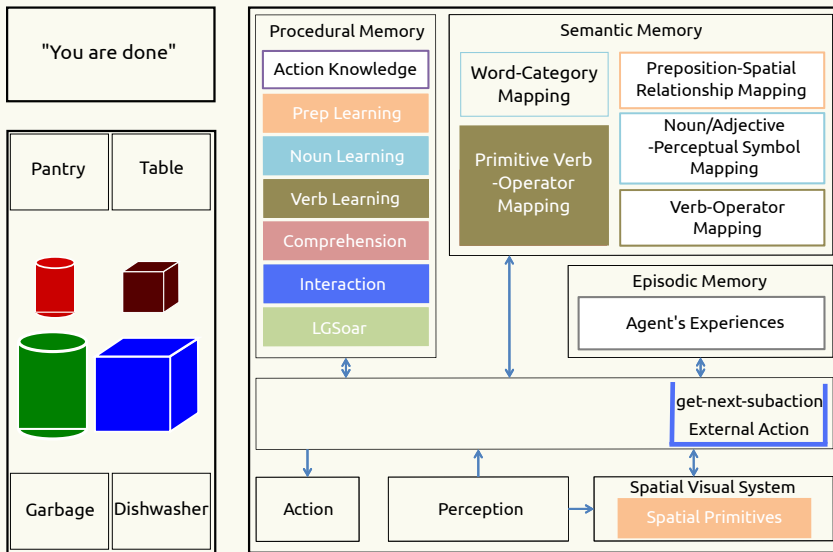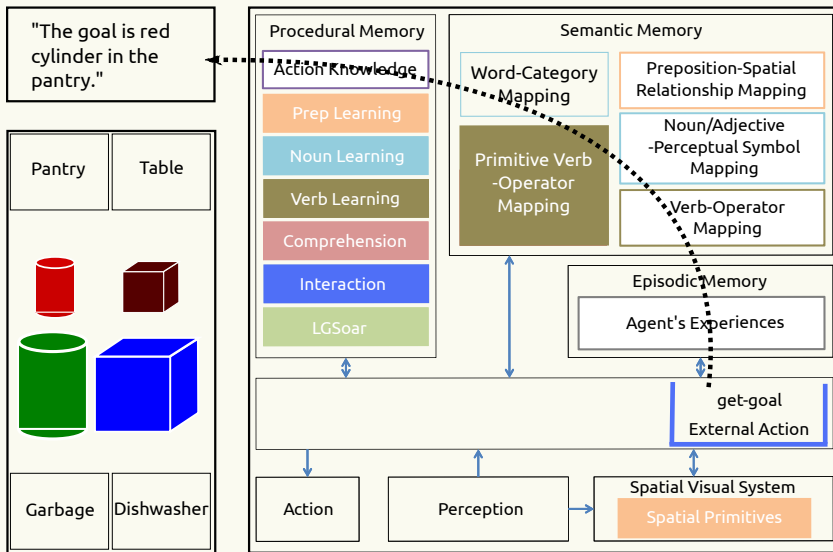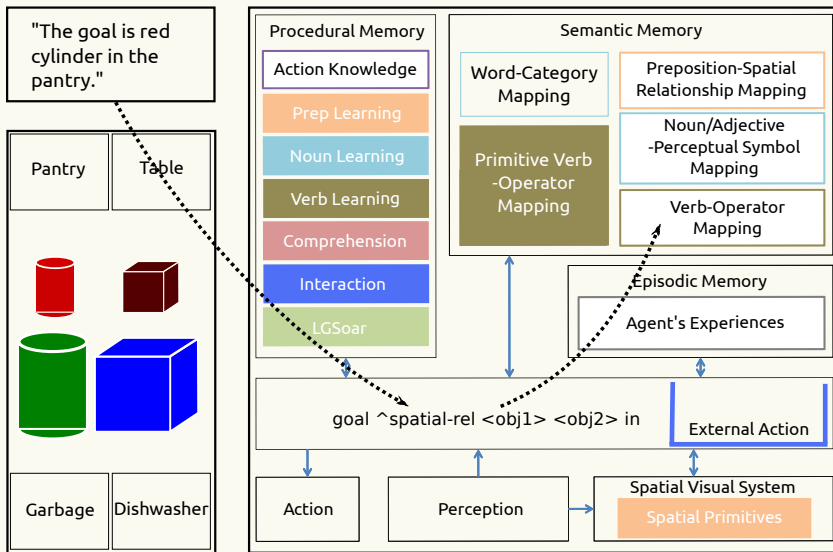
[Failure: Behavior Execution Phase]

Acquire Semantic knowledge

[Failure: Behavior Execution Phase]

Acquire Semantic knowledge
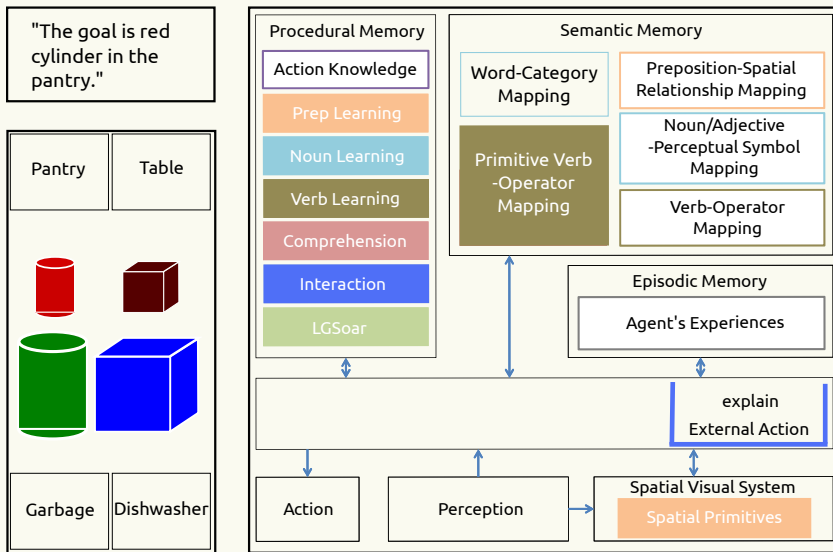
[Failure: Behavior Execution Phase]

"The goal is red cylinder in the pantry."

Pantry | Table

Garbage | Dishwasher

Procedural Memory

Action Knowledge

Prep Learning

Noun Learning

Verb Learning

Comprehension

Interaction

LGSoar

Semantic Memory

Word-Category Mapping

Preposition-Spatial Relationship Mapping

Noun/Adjective -Perceptual Symbol Mapping

Primitive Verb -Operator Mapping

Verb-Operator Mapping

Episodic Memory

Agent's Experiences

goal ^spatial-rel <obj1> <obj2> in

External Action

Action | Perception

Spatial Visual System

Spatial Primitives

Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]

"The goal is red cylinder in the pantry."

Pantry | Table

Garbage | Dishwasher

Procedural Memory

Action Knowledge

Prep Learning

Noun Learning

Verb Learning

Comprehension

Interaction

LGSoar

Semantic Memory

Word-Category Mapping

Preposition-Spatial Relationship Mapping

Primitive Verb -Operator Mapping

Noun/Adjective -Perceptual Symbol Mapping

Verb-Operator Mapping

Episodic Memory

Agent's Experiences

explain External Action

Action | Perception | Spatial Visual System

Spatial Primitives

Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]

Stage I: Retrospective Recall

Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]

Stage I: Retrospective Recall

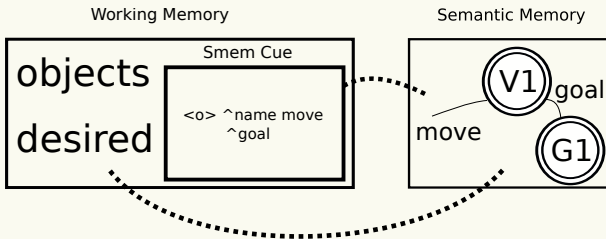- De-couple from the current state

Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]

Stage I: Retrospective Recall

- De-couple from the current state
- Recreate the `beginning' state: explanation
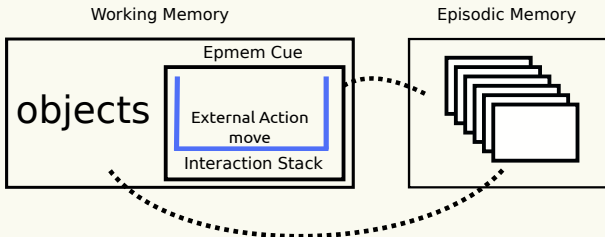
Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]

Stage I: Retrospective Recall

- De-couple from the current state
- Recreate the `beginning' state: explanation
- Assign the desired state

Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]

Stage I: Retrospective Recall

- De-couple from the current state
- Recreate the `beginning' state: explanation
- Assign the desired state
- Propose the operator to be learned in explanation
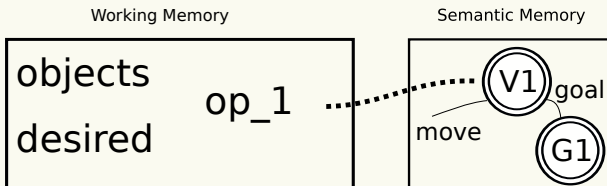
Acquire Procedural Knowledge

[Failure: Behavior Execution Phase]
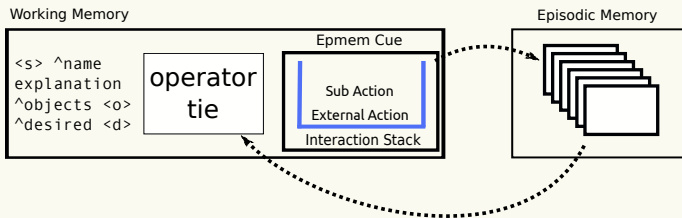
Stage II: Forward Projection

- Selection Space using Situated Experience

Acquire Procedural Knowledge
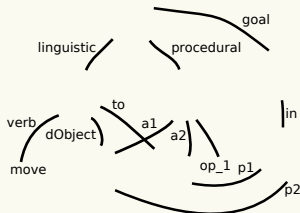
[Failure: Behavior Execution Phase]

Stage II: Forward Projection
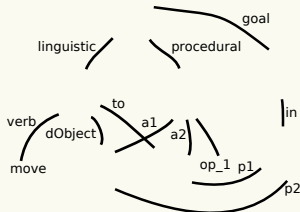
- Selection Space using Situated Experience

Acquired Knowledge

- New Verb-Operator Mapping

Acquired Knowledge

- New Verb-Operator Mapping

- Selection Chunks

```
chunk-1
<s1> ^name op_1 ^argument1 <o1> ^argument2 <l1> ^object <o1> ^object
<o2> <o2> ^category location
<s1> ^operator <op2> <op2> ^name pick-up ^argument1 <o1>
-->
<s1> ^operator <op2> >
chunk-2
<s1> ^name op_1 ^argument1 <o1> ^argument2 <l1> ^object <o1>
^grabbed <o1> ^object <o2> <o2> ^category location
<s1> ^operator <op2> <op2> ^name put-down ^argument1 <o1> ^argument2
<l1>
-->
<s1> ^operator <op2> >
```
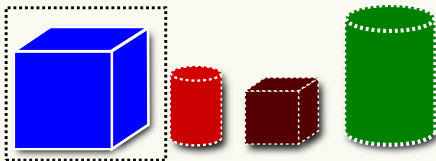
Motivating Examples

- NL exploits (non linguistic) context

Motivating Examples

- NL exploits (non linguistic) context
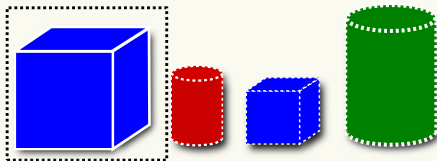

Pick up the blue cube.
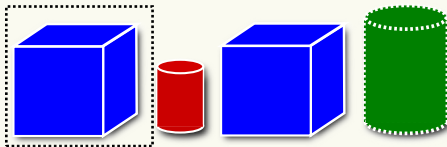
Motivating Examples

- NL exploits (non linguistic) context

Pick up the large, blue cube.

Motivating Examples
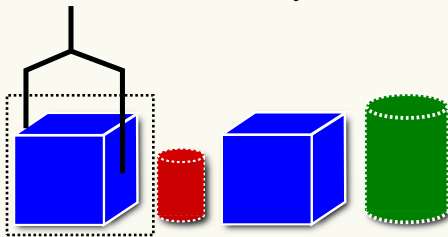
- NL exploits (non linguistic) context

Pick up the cube on the left of the red cylinder.

Motivating Examples

- NL exploits (non linguistic) context

Put down the object.

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous

Put the block down.

Put the block on the red triangle.

Put the block in the pantry.

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous

  Put the red large block to the left of green triangle in the pantry.

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous
- NL often relies of shared knowledge of the world

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous
- NL often relies of shared knowledge of the world

  I had tea yesterday. It burnt my tongue.

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous
- NL often relies of shared knowledge of the world
- NL works in incomplete observations

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous
- NL often relies of shared knowledge of the world
- NL works in incomplete observations

Get a cup from the kitchen.

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous
- NL often relies of shared knowledge of the world
- NL works in incomplete observations
- NL can be used for knowledge transfer

Motivating Examples

- NL exploits (non linguistic) context
- NL is ambiguous
- NL often relies of shared knowledge of the world
- NL works in incomplete observations
- NL can be used for knowledge transfer

  To go to Kroger, take a left on to Plymouth road, drive for a mile and take a left into the complex.