Constructing Meaning in Small Increments

Peter Lindes (plindes@umich.edu) University of Michigan

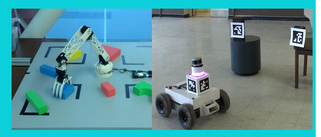
CogSci 2020

Overview

Human Comprehension

- Functionality
  - Integrated in a robotic agent
  - Produce an actionable messages in real time
  - Generality
- Processing
  - One sentence at a time
  - Incremental, word-by-word
  - Immediate interpretation
- Mechanisms
  - Use general cognitive abilities
  - Implementation in a cognitive architecture

Rosie Embodiment



Learns games

Learns fetch and deliver tasks

```
TEACHER: The name of the puzzle is blocks-world.
Please setup the puzzle.
TEACHER: Load init-blocksworld.
Ok.
TEACHER: Ok.
Please describe the actions, goals, and failure conditions.
TEACHER: You can move a clear block onto a clear object.
Please describe the meaning of 'clear' in this context.
TEACHER: If a location is not below an object then it is clear.
Ok, I've learned the meaning of 'clear' for this context.
I've learned the action.
TEACHER: Load final-blocksworld.
Ok.
TEACHER: The goal is that a red block is on a green block and the red block
is below an orange block.
Please setup the goal state.
TEACHER: Ok.
I've learned the goal.
TEACHER: Done.
I've learned blocks-world. Should I try to solve the puzzle?
TEACHER: Yes.
That was easy.
```

Computational Questions

- How is knowledge of language represented?
  - How to map form → meaning?
  - How to make small, composable units?
- How can meaning be built piece-by-piece?
  - How to compose to give generality and creativity?
  - How to provide immediate interpretation?
- How can general cognitive mechanisms do this?
  - Where is knowledge of language stored?
  - How is it retrieved?
  - How are ambiguities resolved?

The Lucia Comprehender

Knowledge of Language

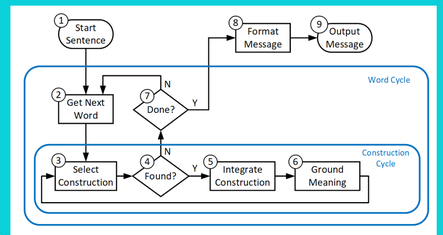
```
construction TransitiveCommand
  subcase of Imperative
  constructional
    constituents
      verb: ActionVerb
      object: RefExpr
  meaning: ActOnIt
    constraints
      self.m.action <--> verb.m
      self.m.object <--> object.m
```

```
schema Action
  roles
    action
    direction
    location

schema ActOnIt
  subcase of Action
  roles
    object
```

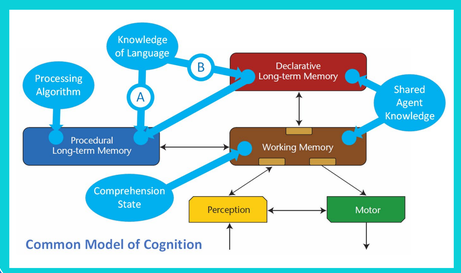
- Embodied Construction Grammar (ECG): a formal, declarative representation
  - Constructions: map form to meaning
  - Schemas: represent meaning structures
  - Compositional hierarchy: pieces to combine
  - Class hierarchy: semantic precision, syntactic flexibility
  - Unification constraints: semantic composition
- Custom ECG grammar: covers Rosie's sentences for Interactive Task Learning (ITL)

Processing Algorithm



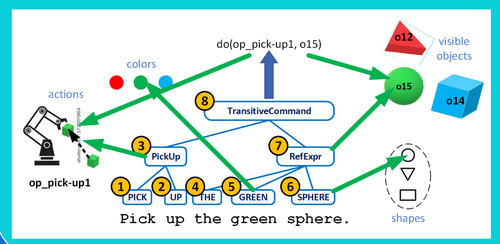
- Sentence meaning structure built incrementally
- One form-meaning pair (construction) at a time
- Construction Cycle:
  - Selection of next construction to apply
  - Procedures build and integrate the form-meaning pair
  - Resulting meaning is grounded to agent's knowledge
- Lexical constructions: build word meanings
- Composite constructions: compose larger "chunks"
- Local selection and repair to resolve ambiguities
- An output message formatted from each complete sentence meaning

Mechanisms



- Modeled on the Common Model of Cognition (CMC)
  - Declarative working memory (WM)
  - Procedural long-term memory (p-LTM)
  - Declarative long-term memory (d-LTM)
  - Cognitive cycle, ~50ms
  - Implementation in the Soar cognitive architecture
- Dynamic comprehension state in WM
- Two implementations:
  - A: ECG translated directly to p-LTM
  - B: ECG -> d-LTM, then p-LTM form learned by processing
- Construction selection:
  - A: by fixed procedures built from ECG
  - B: retrievals from d-LTM biased by context

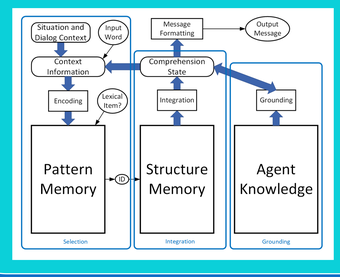
Example Parse



- Each numbered node is a construction
- Constructions are built as numbered
- Lexical items represent word meanings
- Composite items build larger phrases
- Green arrows are grounding to agent's knowledge
  - Actions and shapes in long-term memory
  - Visible objects in short-term memory
- Final meaning is formatted as an output message

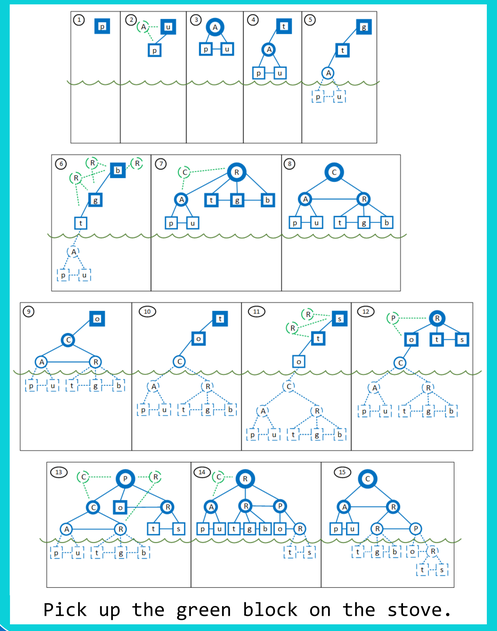
The Comprehension Process

A Construction Cycle



- Each construction cycle has three phases:
  - *Selection*: all available knowledge is applied to choose which construction to apply next
  - *Integration*: the selected construction and its meaning are instantiated and attached into the comprehension state
  - *Grounding*: the meaning is connected to the agent's knowledge in short- and long-term memories
- There may be several construction cycles per word
- Grounding provides *immediate interpretation*
- When the sentence is complete, the full sentence meaning is *formatted* to produce an actionable output message
- This process provides *end-to-end comprehension* in simulated real time

15 Construction Cycles for One Sentence



Key

- A Action verb
- C Command
- R Referring expression
- P Prepositional phrase
- x Word construction

Threshold of accessibility

Border width indicates distance from root

Observations

- Construction nodes in working memory form a tree extending from a root node
- Constructions are combined into larger chunks
  3 p u → A 8 A R → C
- Once created, a construction's depth may change
  A Levels 1 → 2 → 3 → 4 → 2 → 3 → 4 → 5 → 4 → 3 → 2
- Attention spreads from the root node down to deeper levels in the tree
- There is a threshold of accessibility
- Only construction nodes above the threshold are accessible for combining into larger chunks
- Constructions at levels 4 and deeper are below the threshold of accessibility, but are not forgotten
- Local repairs are sometimes performed by discarding a construction and re-using its constituents
  In 13 → 14, C is discarded so its R constituent can be combined with the P to form a larger R.

Conclusions

Hypotheses on Human Comprehension

- Knowledge of language is stored as small, composable units of form → meaning mapping
- Meaning is constructed in these small increments
  - Each new unit added is grounded to the agent's knowledge
  - Grounded meetings provide immediate interpretation
- Working memory holds the current state modeled as a tree of nodes
  - As the tree grows, nodes may go in and out of accessibility
  - Attention spreads through the tree from its root
  - Only the first 3-4 levels from the root are accessible for selecting the next construction to apply
  - There may be a limit, ~5-7?, beyond which a node is lost

Questions for Future Work

- How can we improve the model?
  - Optimize selection of constructions in the architecture?
  - Refine the modeling of timing details?
- Can this model help explain human parsing difficulties and linguistic universals in grammar?
- How can we model learning knowledge of language from experience?
- What human experiments can test the model?
  - Relate the model to human brain measurements?
  - Can we observe construction cycles in humans?
  - Test hypotheses about working memory?
- How might this model be represented with neurons?