Leveraging Cognitive Context for Language Processing in Soar

6/28/2012

SOARTECH

Modeling human reasoning.
Enhancing human performance.

A Black-box Robot Architecture

```
→ [ Parsing ] → [ Semantics ] →        [ Object
                → [ Vision ] →           Grounding ] → [ Cognition ] → [ Action ] →
```
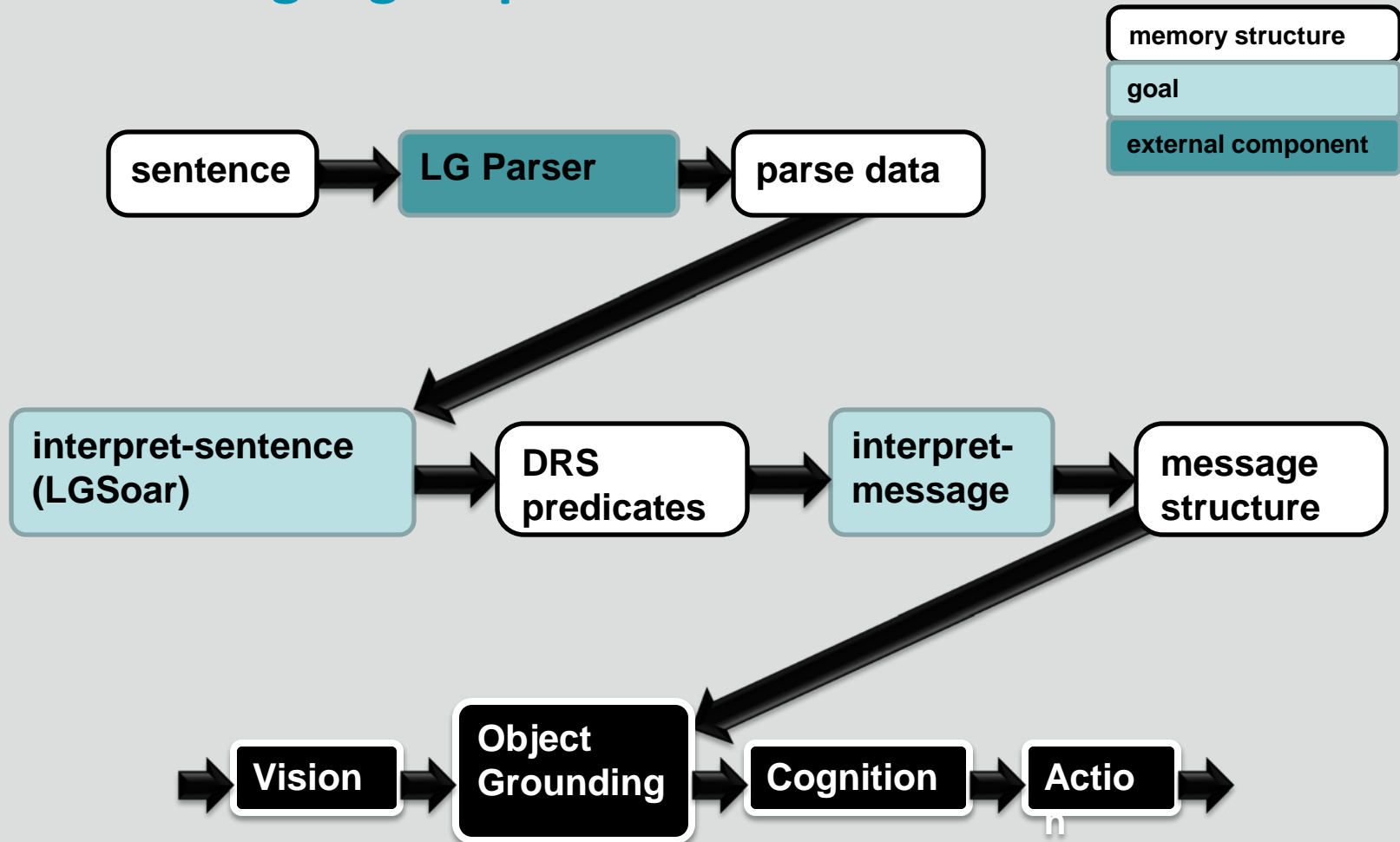
- Problem: Parsing and Semantics are underdetermined by language
- Approach: Use cognitive and environmental context to help
- Cognitive context
  - What is the purpose of the current dialog event?
  - What is the overall dialog about?
  - What is known about the words being used?
- Environmental context
  - What objects exist nearby?
  - What are the affordances of the nearby objects?

- Focus here: Using cognitive context to improve parsing

Basic Pipeline: Components

- Link Grammar (LG) parser
  - Domain-independent syntax parsing
  - Existing off-the-shelf system
  - Lots of language knowledge
- LGSoar
  - Domain-independent semantics
  - Existing research system
  - Minimal language knowledge
- Message Interpretation
  - Domain-specific semantics
  - New component
  - Minimal language knowledge
- Vision, Grounding, Action, etc..
  - See other talks

Soar Technology, Inc. Proprietary

Soar Language Pipeline

memory structure

goal

external component

sentence → LG Parser → parse data

interpret-sentence (LGSoar) → DRS predicates → interpret-message → message structure

Vision → Object Grounding → Cognition → Action

Soar Technology, Inc. Proprietary

Example: Object Message

Input   "The triangle is red."

LG Parser

```
        +--Ds-+--Ss----+-Pa-+
        |     |        |    |
       the triangle.n is.v red.a
```

LGSoar

```
       DEF(N4)
       triangle(N4)
       red(N4)
```

Message Interpretation

```
       (I6 ^object-message I5)
         (I5 ^object O37)
            (O37 ^specifier DEF)
            (O37 ^word red)
            (O37 ^word triangle)
```

Soar Technology, Inc. Proprietary

Example: Object Message

Input	"This is a red triangle"

LG Parser

```
              +-----Ost-----+
              |   +----Ds----+
        +-Ss*b+   |     +---A--+
        |     |   |     |      |
      this.p is.v a red.a triangle.n
```

LGSoar

```
INDEF(N4)
this(N4)
triangle(N4)
red(N4)
```

Message Interpretation

```
(I6 ^object-message I5)
  (I5 ^object O37)
    (O37 ^specifier this)
    (O37 ^word red)
    (O37 ^word triangle)
```

Soar Technology, Inc. Proprietary

Example: Object Message

Input   "The red thing is a triangle."

LG Parser

```
        +-------------------Xp---------------------+
        +--------Wd--------+                       |
        |     +-----Ds----+        +--Ost--+       |
        |     |    +---A--+--Ss--+  +-Ds-+         |
        |     |    |      |      |  |    |         |
    LEFT-WALL the red.a thing.n is.v a triangle.n .
```

LGSoar

```
INDEF(N5)
DEF(N5)
red(N5)
thing(N5)
triangle(N5)
```

Message Interpretation

```
(I6 ^object-message I5)
  (I5 ^object O38)
    (O38 ^specifier DEF)
    (O38 ^word triangle)
    (O38 ^word thing)
    (O38 ^word red)
```

Example: Relation Question

Input "What is on top of the big blue block?"

LG Parser

```
    +-----------------------------Xp----------------------------+
    |                               +----------Js----------+     |
    |                               |     +--------Ds--------+   |
    |                +-----Pp----+   |     +------A------+     |
    +---Wq--+Ss*w+   +ID+IDV+    |   |           +---A--+     |
    |       |    |   |   |   |   |   |           |      |     |
  LEFT-WALL what is.v on top of the big.a blue.a block.n ?
```

LGSoar
```
block(N8)
DEF(N8)
big(N8)
blue(N8)
what(N4)
on-top-of(N4,N8)
```

MI
```
(I9 ^relation-question I5)
  (I5 ^question-word what)
  (I5 ^relation R23)
    (R23 ^p1.object O51)
      (O51 ^specifier none)
      (O51 ^word UNKNOWN)
    (R23 ^p2.object O57)
      (O57 ^specifier DEF)
      (O57 ^word blue)
      (O57 ^word big)
      (O57 ^word block)
    (R23 ^word on-top-of)
```

## Problem: Ambiguous Prepositional Phrase Attachment

- Agent is told "Point at the circle on the table."
- Link Grammar parser returns two possible parses:
  - Point (at the circle) (on the table)
  - Point (at (the circle on the table))
- First (incorrect) parse is presented as the best

- Agent may know things about the verb "point"
  - It doesn't usually have an attached preposition
- Even if the agent hasn't learned that, heuristics should be explicit Soar knowledge (not implicit parser knowledge)

Prepositional Phrase Attachment Solution

- System was modified to put all valid parses in Working Memory
- Select-parse subgoal was added to choose a single parse
  - Scores from the parser are used to rank parses
  - If there is an alternative to the highest-ranked parse with a different pp attachment, a conflict is identified
  - Heuristic rules resolve conflicts, losing sentence is removed from consideration
    - Future work: semantic memory verb knowledge
  - Attachments are resolved iteratively left to right in the sentence
    - Rough approximation of incremental parsing

Soar Language Pipeline

Soar Technology, Inc. Proprietary   6/28/2012

Soar Language Pipeline

Soar Technology, Inc. Proprietary
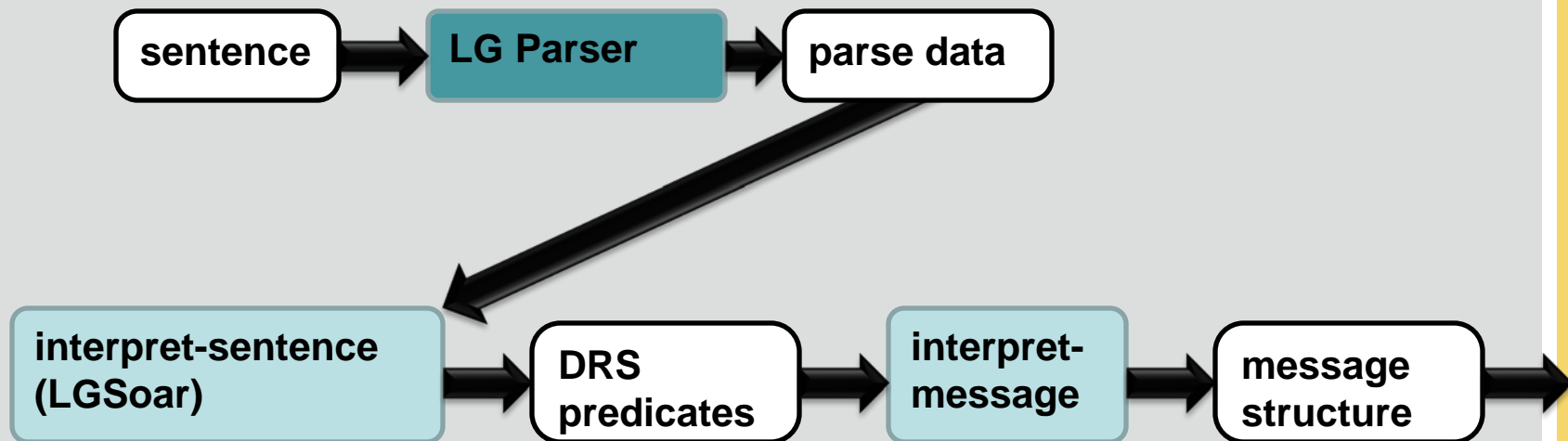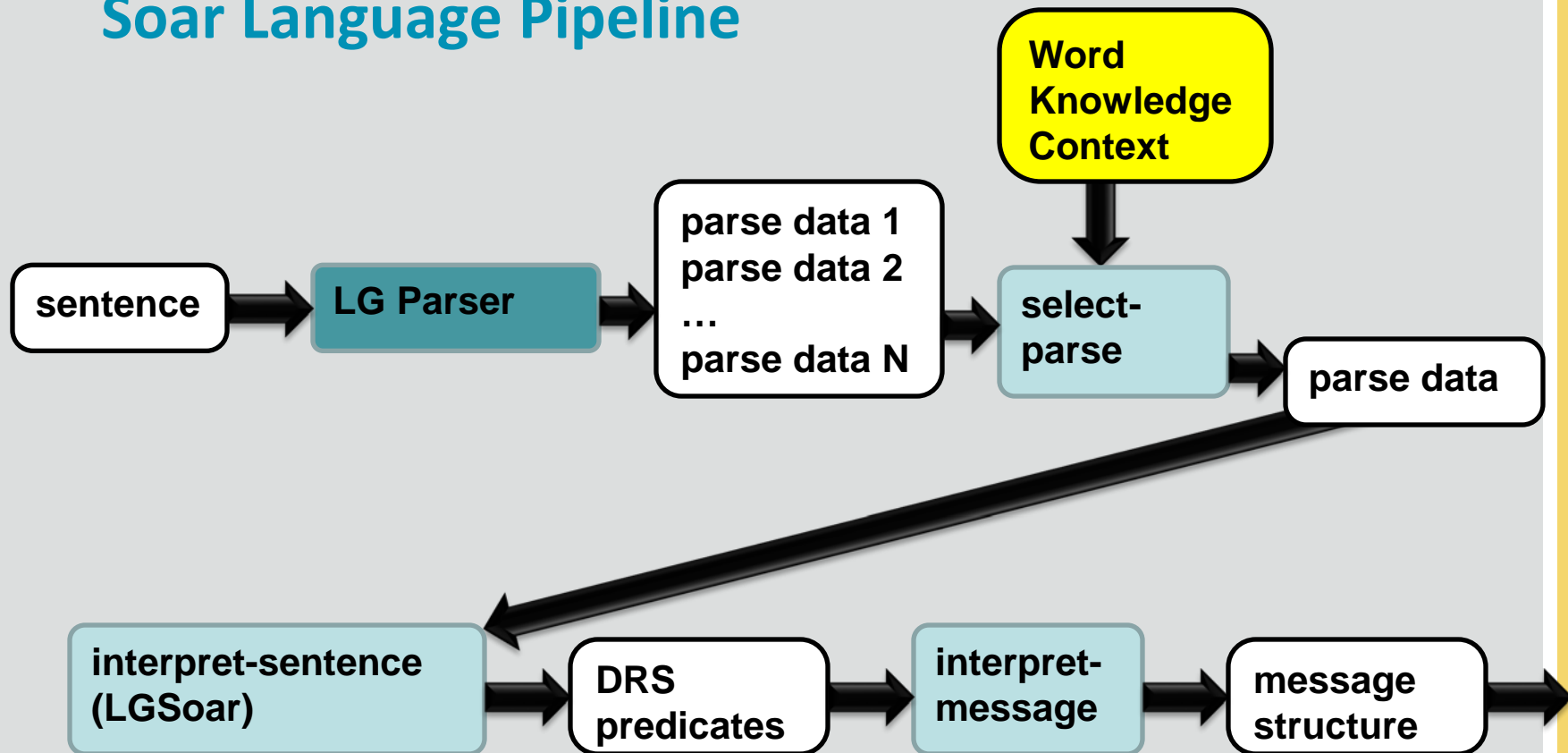
Problem: Implicit Quoting

- Robot is learning the word "circle", used to describe an object
- Robot knows perceptual categories like shape, color, and size
- Dialog:
  - Human: "This is a circle."
  - Robot: "What category does circle belong to?"
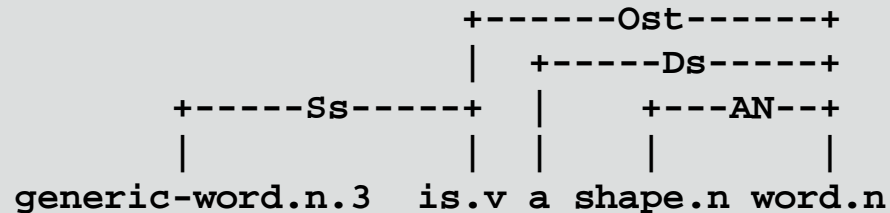 - Human: "Circle is a shape word."

- LG Parser cannot handle this:

```
                    +----------Os----------+
                    |              +-----Ds-----+
          +----Wi---+              |     +---AN--+
          |         |              |     |       |
    LEFT-WALL circle.v [is] a shape.n word.n
```
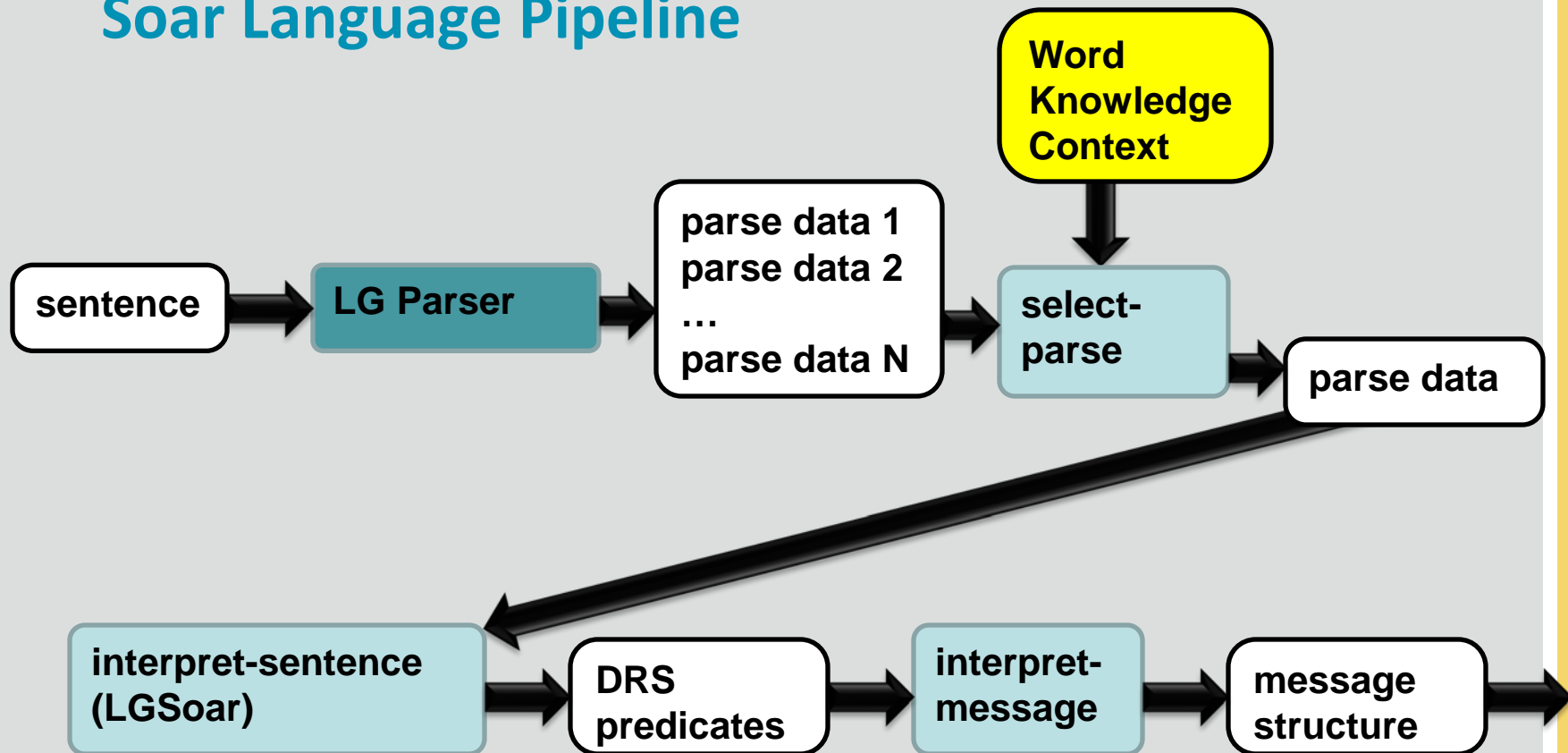
Implicit Quoting Solution

- The agent knows that the word itself is under discussion, this information should be used to help the parser

- preprocess-sentence subgoal was added to let Soar control parser input
- Link Grammar dictionaries were augmented with generic words
- If a word is under discussion, it is replaced with a generic that can be linked as the subject of the sentence (a mass noun)
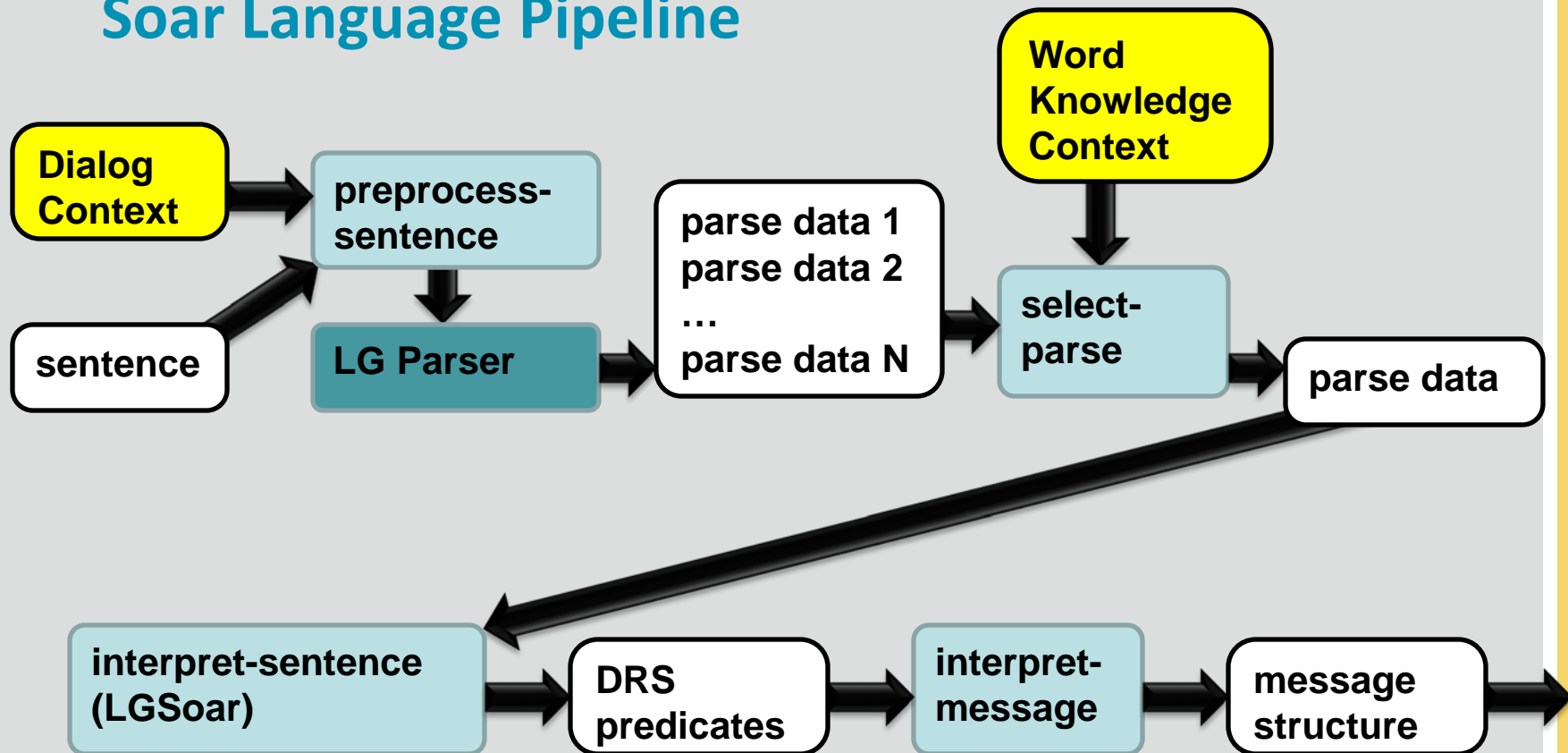
```
                              +------Ost------+
                              |   +-----Ds-----+
            +-----Ss-----+    |      +---AN--+
            |            |    |  |   |       |
     generic-word.n.3  is.v a shape.n word.n
```

- Later, Soar rules re-substitute the real word
- Capability will also be used to support learning of parser knowledge

Soar Language Pipeline

sentence → LG Parser → parse data 1 / parse data 2 / … / parse data N → select-parse → parse data

Word Knowledge Context → select-parse

parse data → interpret-sentence (LGSoar) → DRS predicates → interpret-message → message structure

6/28/2012

Soar Language Pipeline

Soar Technology, Inc. Proprietary

6/28/2012

Problem: Noun phrases

- Natural interaction involves many utterances that aren't sentences
- In our domain, plain noun phrases are common:
  - Human: Pick up the red square.
  - Robot: Which red square do you mean?
  - Human: The red square on the table.

- The parser is cannot handle this:

```
      +------------------Xp------------------+
      +-----Wd----+              +---Js---+   |
      |      +-DD-+---Sp--+-MVp-+  +--Ds-+   |
      |      |    |       |     |  |     |   |
  LEFT-WALL the red.a square.v on the table.n .
```

- Other phrases have no valid parses

Soar Technology, Inc. Proprietary   6/28/2012

Noun Phrase Solution

- Link Grammar parser was modified to handle noun phrases
  - Parses prefixed by a generic verb are always considered along with standard parses
- The select-parse subgoal was augmented to bias noun phrase parses based on dialog context
  - Normally, bias towards full-sentence parses
  - If the agent just asked a "which" question, bias towards noun phrases

```
    +-------------------Xp-------------------+
    +-----Wd----+                +---Js---+  |
    |        +-DD-+---Sp--+-MVp-+  +--Ds-+   |
    |        |    |       |     |  |     |   |
 LEFT-WALL the red.a square.v on the table.n .


              +----------Os----------+
              |             +-----Ds-----+      +---Js---+
    +------Wi-----+         |      +---A---+--Mp-+  +--Ds-+
    |             |         |      |    |     |  |    |
 LEFT-WALL NOUN-PHRASE-WALL the red.a square.n on the table.n
```
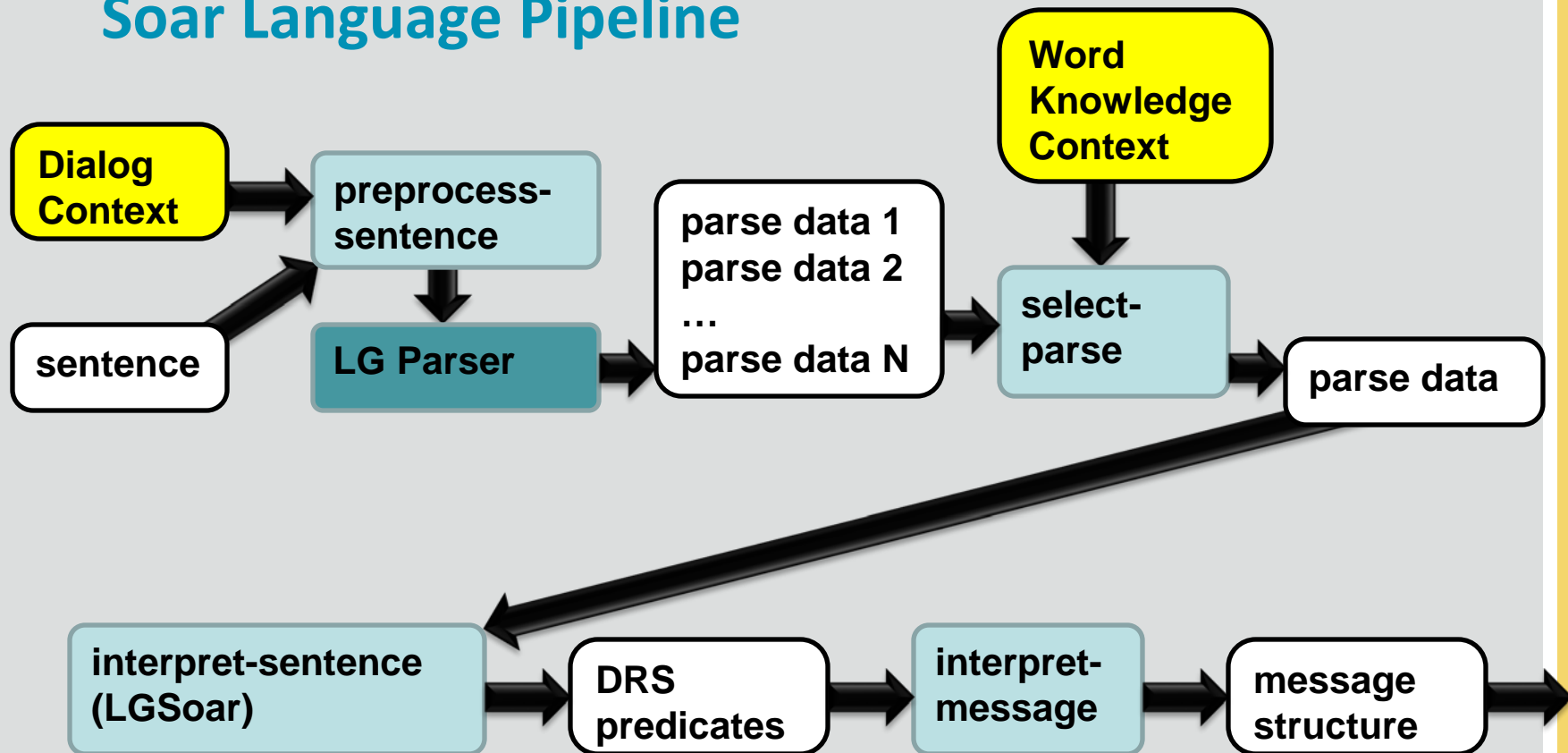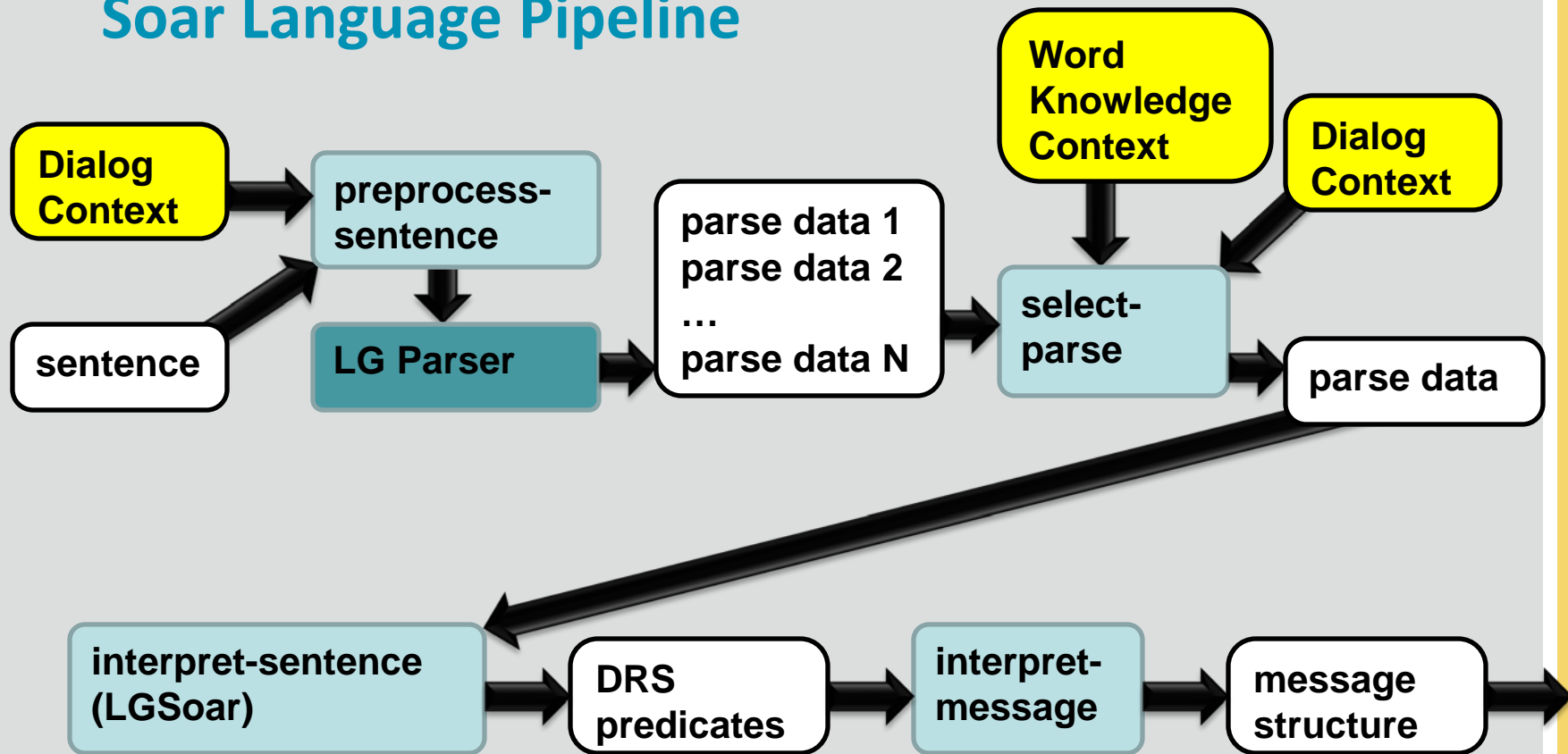
Soar Language Pipeline

Soar Technology, Inc. Proprietary 6/28/2012

Soar Language Pipeline

Soar Technology, Inc. Proprietary

6/28/2012

Future Work

- Prepositional phrase improvements
  - Integrate with verb learning system to use knowledge about verbs for prepositional phrase attachment
  - Use grounded environment context to resolve pp attachments, based on feedback from systems further down the line
- Learning of parser knowledge
  - Addition of new words to vocabulary

Soar Technology, Inc. Proprietary
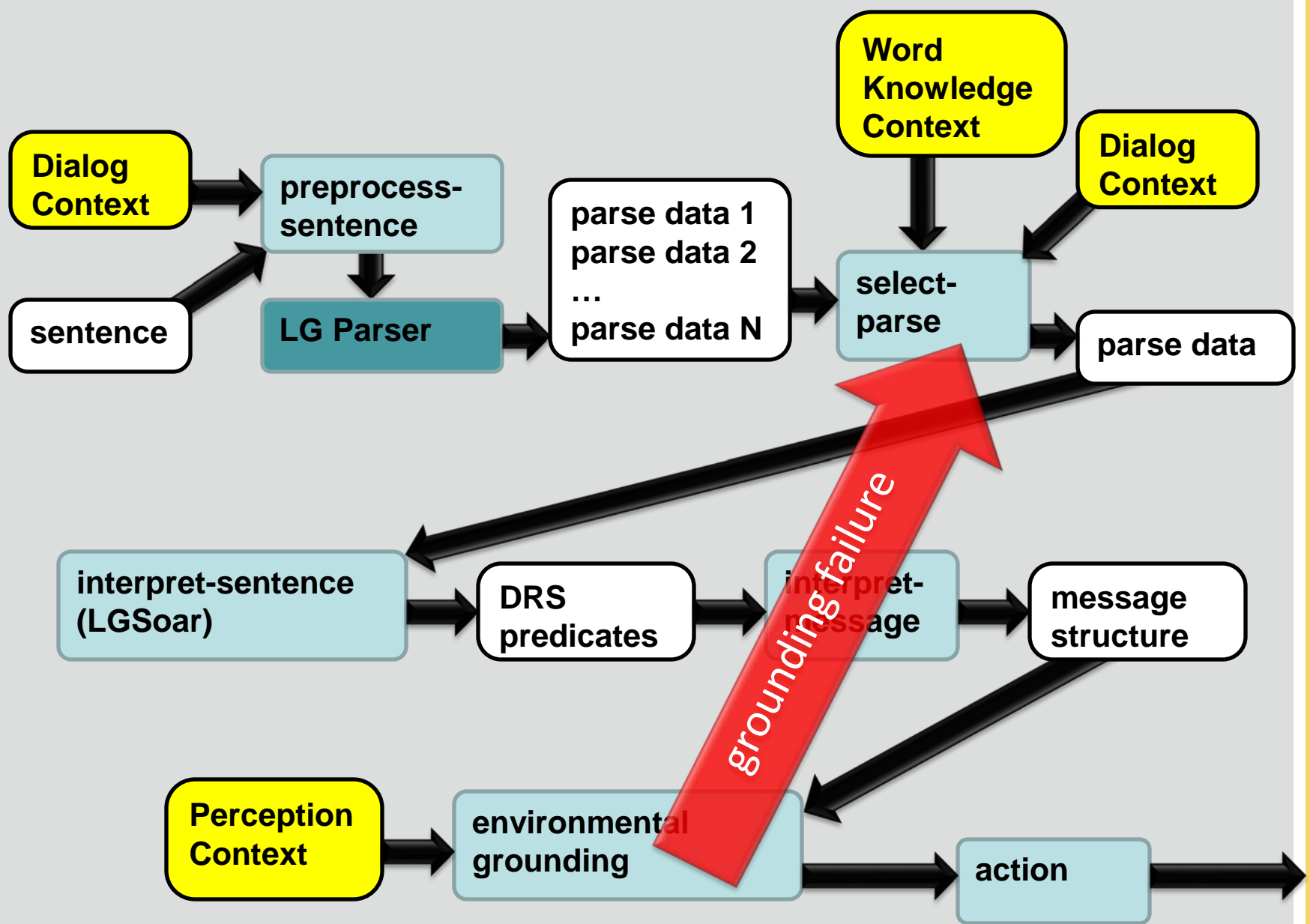
Nuggets and Coal

- Nuggets
    - Link Grammar/LGSoar systems bring in lots of existing work
    - So far, no real problems in giving Soar more control over the external parser
    - Using a functioning system focuses the research on practical issues
- Coal
    - Not incremental, the entire sentence is parsed before other stages begin
        - This may introduce computational issues (hundreds of parses in WM)
        - Unclear if this is more than a computational issue
    - Lots of coordination between research groups is necessary, especially for environmental context