Mixed-Initiative Interaction for Learning with Instruction

Shiwali Mohan and John E. Laird

June 20, 2012
Outline

1 Interaction in BOLT

2 Interaction Model: Requirements

3 Interaction Model: Design Details

4 Conclusions
Why is an Interaction Module needed?

- Learning from mixed-initiative communication
  - *Interaction module as a dialog manager*
Why is an Interaction Module needed?

- Learning from mixed-initiative communication
  - *Interaction module as a dialog manager*
- Many different capabilities
  - Linguistic processing
  - Task execution
  - Learning (semantic, procedural, perceptual)
  - *Interaction module as a scheduler*
Why is an Interaction Module needed?

- Learning from mixed-initiative communication
  - *Interaction module as a dialog manager*
- Many different capabilities
  - Linguistic processing
  - Task execution
  - Learning (semantic, procedural, perceptual)
  - *Interaction module as a scheduler*
- Interruptions in processing
  - Vision system is not perfect
  - Actions are non deterministic
  - *Interaction module for task management*
Why is an Interaction Module needed?

- Learning from mixed-initiative communication
  - *Interaction module as a dialog manager*
- Many different capabilities
  - Linguistic processing
  - Task execution
  - Learning (semantic, procedural, perceptual)
  - *Interaction module as a scheduler*
- Interruptions in processing
  - Vision system is not perfect
  - Actions are non deterministic
  - *Interaction module for task management*
- Implementation
  - State of interaction is maintained as an ‘interaction stack’
    - maintained in working memory, not state stack
  - Communication, learning, actions change the state of interaction through interaction operators
Interaction Management

"Put the red cylinder in the pantry"

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

Episodic Memory
- Agent's Experiences

Action Command
- Sentence Parse

External Action

Spatial Visual System
- Spatial Primitives
Requirements for a Mixed Initiative Interaction Model

• Integrative: Combine dialog, linguistic processing, planning, execution and learning.

• Mixed-Initiative: Both the instructor and the agent should be able to assume control of the interactions.

• Instructor is able to provide situated examples.

• Agent is able to pose queries.

• Contextual: The model should provide useful context for instructor's utterance.

• Agent uses dialog context for interpretation (syntactical, semantic, pragmatic comprehension).

• Task Relevant: Agent's utterances should be informed by its decision processes, knowledge, learning.

• Temporal: The model and the sequence of interactions should inform agent's learning.

• Agent is able to learn from temporally delayed information.
Requirements for a Mixed Initiative Interaction Model

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
Requirements for a Mixed Initiative Interaction Model

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Instructor is able to provide *situated* examples
  - Agent is able to pose queries
- **Contextual**: The model should provide useful context for instructor’s utterance.
  - Agent uses dialog context for interpretation (syntactical, semantic, pragmatic comprehension)
- **Task Relevant**: Agent’s utterances should be informed by its decision processes, knowledge, learning.
- **Temporal**: The model and the sequence of interactions should inform agent’s learning.
  - Agent is able to learn from temporally delayed information
Requirements for a Mixed Initiative Interaction Model

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.

- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Instructor is able to provide *situated* examples
  - Agent is able to pose queries

- **Contextual**: The model should provide *useful* context for instructor’s utterance.
  - Agent uses dialog context for interpretation
    (syntactical, semantic, pragmatic comprehension)
Requirements for a Mixed Initiative Interaction Model

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.

- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Instructor is able to provide *situated* examples
  - Agent is able to pose queries

- **Contextual**: The model should provide *useful* context for instructor’s utterance.
  - Agent uses dialog context for interpretation (syntactical, semantic, pragmatic comprehension)

- **Task Relevant**: Agent’s utterances should be informed by its decision processes, knowledge, learning.
Requirements for a Mixed Initiative Interaction Model

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Instructor is able to provide *situated* examples
  - Agent is able to pose queries
- **Contextual**: The model should provide *useful* context for instructor’s utterance.
  - Agent uses dialog context for interpretation
    (syntactical, semantic, pragmatic comprehension)
- **Task Relevant**: Agent’s utterances should be informed by its decision processes, knowledge, learning.
- **Temporal**: The model and the sequence of interactions should inform agent’s learning.
  - Agent is able to learn from temporally delayed information
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Events
change the state of interaction

- Action events
  - change in the environment state
- Dialog events
  - instructor/agent elicitation
- Learning events
  - change in agent’s knowledge

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>dialog-event</td>
<td>Human: Move the red block to the pantry.</td>
</tr>
<tr>
<td>dialog-event</td>
<td>Agent: I do not know how to perform this action.</td>
</tr>
<tr>
<td>dialog-event</td>
<td>Human: Pick up the red block.</td>
</tr>
<tr>
<td>action-event</td>
<td>Agent: (picks up the red block)</td>
</tr>
<tr>
<td>dialog-event</td>
<td>Agent: What do I do next?</td>
</tr>
<tr>
<td>dialog-event</td>
<td>Human: Put the object in the pantry.</td>
</tr>
<tr>
<td>action-event</td>
<td>Agent: (puts the red block in the pantry)</td>
</tr>
<tr>
<td>dialog-event</td>
<td>Agent: What do I do next?</td>
</tr>
<tr>
<td>dialog-event</td>
<td>Human: You are done.</td>
</tr>
<tr>
<td>learning-event</td>
<td>Agent: (learn with forward projection)</td>
</tr>
</tbody>
</table>
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events
purpose, satisfaction

dialog-event Human: Move the red block to the pantry.
dialog-event Agent: I do not know how to perform this action.
dialog-event Human: Pick up the red block.
action-event Agent: (picks up the red block)
dialog-event Agent: What do I do next?
dialog-event Human: Put the object in the pantry.
action-event Agent: (puts the red block in the pantry)
dialog-event Agent: What do I do next?
......
dialog-event Human: You are done.
learning-event Agent: (learn with forward projection)
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events
*purpose, satisfaction*

- A question-answer sequence

```
linguistic segment

Human: Move the red block to the pantry.
Agent: I do not know how to perform this action.
Human: Pick up the red block.
Agent: (picks up the red block)
Agent: What do I do next?
Human: Put the object in the pantry.
Agent: (puts the red block in the pantry)
Agent: What do I do next?
.....
Human: You are done.
Agent: (learn with forward projection)
```
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events

*purpose, satisfaction*

- A question-answer sequence
- A command-action sequence

---

mixed segment

Human: Move the red block to the pantry.
Agent: I do not know how to perform this action.
Human: Pick up the red block.
Agent: (picks up the red block)
Agent: What do I do next?
Human: Put the object in the pantry.
Agent: (puts the red block in the pantry)
Agent: What do I do next?
.....
Human: You are done.
Agent: (learn with forward projection)
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events

- A question-answer sequence
- A command-action sequence
- Hierarchical

Human: Move the red block to the pantry.
Agent: I do not know how to perform this action.
Human: Pick up the red block.
Agent: (picks up the red block)
Agent: What do I do next?
Human: Put the object in the pantry.
Agent: (puts the red block in the pantry)
Agent: What do I do next?
.....
Human: You are done.
Agent: (learn with forward projection)
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events
purpose, satisfaction

• Heuristically determined purpose, satisfaction
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events
*purpose, satisfaction*

- Heuristically determined purpose, satisfaction
  - Domain based heuristics
    - action-command: *purpose* - external action; *satisfaction* - successful action/indication of successful action
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Segment
contiguous block of events

- Heuristically determined purpose, satisfaction
  - Domain based heuristics
    - action-command: *purpose* - external action; *satisfaction* - successful action/indication of successful action
  - Learning based heuristics
    - learning composite action dominates primitive action execution
Interaction Model
adapted from Rich and Sidner (1998)
based on Collaborative Discourse Theory (Grosz and Sidner, 1986)

Interaction Stack
contiguous block of events
purpose, satisfaction

- Represents the state of dialog between the instructor and the agent
- A stack of open segments (purpose has not been achieved)
- The top segment determines the current focus of dialog.

<table>
<thead>
<tr>
<th>get-next-subaction</th>
<th>external action</th>
<th>move</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human: Move the red block to the pantry.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent: I do not know how to perform this action.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human: Pick up the red block.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent: (picks up the red block)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent: What do I do next?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human: Put the object in the pantry.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent: (puts the red block in the pantry)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent: What do I do next?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.....</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human: You are done.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent: (learn with forward projection)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Requirement Analysis

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
  - Utterances, actions, learning as events
Requirement Analysis

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
  - Utterances, actions, learning as events
- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Anyone can initiate a segment.
Requirement Analysis

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
  - Utterances, actions, learning as events
- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Anyone can initiate a segment.
- **Contextual**: The model should provide *useful* context for instructor’s utterance.
  - Domain/learning specific heuristics
requirement analysis

• **integrated**: combine dialog, linguistic processing, planning, execution and learning.
  • Utterances, actions, learning as events

• **mixed-initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  • Anyone can initiate a segment.

• **contextual**: The model should provide *useful* context for instructor’s utterance.
  • Domain/learning specific heuristics

• **task relevant**: Agent’s utterances should be informed by its decision processes, knowledge, learning.
  • Impasse driven
Requirement Analysis

- **Integrative**: Combine dialog, linguistic processing, planning, execution and learning.
  - Utterances, actions, learning as events

- **Mixed-Initiative**: Both the instructor and the agent should be able to assume control of the interactions.
  - Anyone can initiate a segment.

- **Contextual**: The model should provide *useful* context for instructor’s utterance.
  - Domain/learning specific heuristics

- **Task Relevant**: Agent’s utterances should be informed by its decision processes, knowledge, learning.
  - Impasse driven

- **Temporal**: The model and the sequence of interactions should inform agent’s learning.
  - Episodic memory encodes changes in interaction state.
Nuggets and Coal

- Nuggets
  - The interaction model integrates well with other modules
    - in deployment
  - Progress from Huffamn and Laird (1995)
    - Did not allow instructor initiated instructions.

- Coal
  - Limited understanding of ‘initiative’.
  - Intentions are heuristically derived
    - Hard in complex scenarios