Using Spatial Context and Clarification Questions to Interpret Natural Language Commands

Joe Plaus
Advisor: Prof. Striegnitz
Background

- Natural Language Command Systems
- Previous Systems
  - “Understanding Natural Language Commands For Robotic Navigation and Mobile Manipulation” by Stefanie Tellex
- Explicit vs. Ambiguous Commands
- How should a system interpret ambiguity?
- Why is this important?
System Overview

1. Natural Language Command
   - Parse
   - Interpret
   - Execute

2. Visual Representation
3D Environment

- Unity Game Engine
- Room Layout
- Robot
- Moveable Objects
- C# Behavior Scripts
Parsing the Command

- Three types of commands
- Why parse the command?
- Context-Free Grammars and the Natural Language ToolKit for Python
- Semantic vs. Syntactic Parsing
- What is the result?
Parsing the Command

Basic Command: “Walk towards the green ball”

Syntactic Approach

Semantic Approach
System Overview

Natural Language Command
"Walk towards the green ball"

Visual Representation

Parse → Interpret → Execute

[NavInstr: [Action: walk] [Spacial: towards] [Destination: [Determiner: the] [Type: ball] [Color: green]]]
Interpreting the Parsed Command

Three Main Components:
- Command object
- Command Manager
- Command Processor

Navigational, Operational, or Response command?
Explicit or Ambiguous?

Where are we going?
What is the operation?
What object to manipulate?
What was the previous command?
Interpreting Explicit Commands

- Basic Commands
- Complex Commands
  - Destination or Object is determined
    - Type?
    - Color?
  - Can the user see only one object that fits the description?
  - Is it the only object of that configuration in the environment?
Interpreting Ambiguous Commands

- Example: “Pick up the box”
- Which box?
- Salience
  - Proximity
  - Field of View location
- Clarification Question
- Response Command
System Overview

Natural Language Command
"Walk to the green ball"

Parse

Interpret

Visual Representation

“Walk” to:
Objects with tag "ball"
AND
Objects of color "green"
AND
Objects that are visible

Execute

Response command

Clarification Question

[NavInstr: [Action: walk] [Spatial: towards] [Destination: [Determiner: the] [Type: ball] [Color: green]]]
Executing the Command

- Defining Robot Behavior
- Movement functions
- Operation functions
  - Trigger Animations
Testing and Evaluation

Testing:
- Basic Functionality
- How does it handle ambiguity?
- Communication between components

Evaluation:
- Functionality vs. Utility
- Is this new ability helpful?
  - System Comparison
Future Work

- Environment complexity
- Command complexity
  - Context-Free Grammar expansion
  - Multiple commands
- Additional attributes for salience calculation
Thank You