Interaction with a Virtual World through an Information State Dialog System

Nigel Kilmer

Architecture

- Initially using TRINDIKIT
  - ... but ran into a lot of issues getting it to work

- Switched to VoiceXML
  - Allowed for a simpler, custom information state model
  - Back-end coding in Javascript instead of Prolog

World Description

- The world contains objects, which have:
  - a type (table, box, ball)
  - a color
  - an optional positional relationship with another object
    (the ball is in the box)

- Represented as an array of Javascript objects
  - elements of each object are 'type', 'color', 'index', 'relPrep', 'relObj'
    - 'relPrep' is the preposition that defines the optional positional relationship
    - 'relObj' is a reference to the other object in the positional relationship

Information State

- The information state contains:

  - a plan (list of actions for the agent to perform) with elements containing:
    - an action
    - arguments for the action (direct object, indirect object, etc.)
  - information about the previous utterance by the user:
    - the dialog act (request, question)
    - the requested action
    - the queried property
    - arguments for the action or query

VoiceXML Architecture

- Application is contained in two forms:

  - "agentMove" is responsible for interpreting the current information state
    - Constructs new plan elements based on user's previous utterance
    - Performs the highest priority element on the plan

  - "userMove" is responsible for interpreting user utterances
    - Determines the dialog act, requested action / queried property, arguments
    - Sets appropriate values in the information state

Example

A: "There is a brown table. There is a red box on the brown table. There is a green ball in the red box."

U: "Remove the red box."

A: "I have removed the red box."

U: "Tell me about the world."

A: "There is a brown table. There is a green ball on the brown table."

U: "Remove the ball."

A: "I have removed the green ball."

U: "Describe the table."

A: "There is a brown table."

Next Steps

- Add question handling to the system
- Add requests for:
  - recoloring
  - repositioning
  - adding objects
- Ambiguity resolution
- Handling erroneous assumptions by the user
  - **U:** "Remove the red box."
  - **A:** "There is no red box. Did you mean the blue box?"
  - **U:** "Yes."
  - **A:** "I have removed the blue box."

Questions?

Hmmm...?