Special Lecture (406)
Spoken Language Dialog Systems
Introduction to the CSLU Speech Toolkit

Rolf Schwitter
schwitt@ics.mq.edu.au
Today's Agenda

- CSLU Speech Toolkit
- What is RAD?
- Speech Synthesis and Speech Recognition
- Objects
- Prompts
- Repair Dialog
- Variables
- Exploring the CSLU Speech Toolkit
Recap: CSLU Speech Toolkit

- Developed at the Centre for Spoken Language Understanding at the Oregon Graduate Institute in the USA.
- The CSLU toolkit has many features that are similar to commercial tools for building SLDSs.
- The toolkit includes a rapid application development environment.
- A tutorial is available at:
  http://www.cslu.ogi.edu/toolkit/docs/2.0/apps/rad/index.html
CSLU Speech Toolkit Component

- A speech recognizer
- A speech synthesizer
- A visual programming environment for dialog construction
- Facial animation (talking head)
- Waveform analysis tools
- Programming environment for Tcl
What is RAD?

• RAD stands for Rapid Application Development.
• RAD is a visual programming environment.
• You develop applications by dragging objects onto the screen.
• Objects are organized in an area on the screen called "canvas".
• By connecting objects, a set of executable instructions is formed.
Getting Started with RAD
RAD Environment

• When RAD opens, three parts will appear:
  – canvas with objects
  – talking head
  – caption
Demo: Food Ordering Service
Welcome to the Macquarie University Food Ordering Service
At any time, feel free to say help or menu for more information.
Which outlet do you want to order from; we have pastries and soft drinks,
muffins and coffee, ice cream or sandwiches.
Welcome to the muffins and coffee outlet.
Would you like a muffin or a coffee?
What sort of muffin would you like?
Do you want a blueberry, strawberry, choc chip or banana?
So that’s a banana muffin?
Would you like any coffee with your banana muffin?
What sort of coffee would you like?
You can have a short black, long black, flat white, cappuccino or an expresso.
So that’s an expresso and a banana muffin?
Please say yes or no.
Thank you for ordering from the muffins and coffee outlet.
Would you like to order from another outlet?
Thank you for ordering from the Macquarie University Food Ordering Service
Objects

Used Objects:

- Start
- Generic
- Subdialogue
- Goodbye
Speech Synthesis and Speech Recognition

Speech Recognition
(double click red port)

anything_else

Speech Synthesis
(double click icon)
Building a Simple Speech Application

- Let us build a very simple speech application — step by step:

```
Are you ready to learn?
Come on, it will be fun!
Are you ready to learn?
Okay, then proceed to the next lesson.
```
Dragging Objects onto the Canvas

- The start object appears by default.
- Left-click and drag an additional object onto the canvas:
Renaming Objects

- To rename an object, right-click on it to raise the context-menu:
Renaming Objects

- Type in the new name followed by `<Enter>`:

- Any space in a name will change to an underscore.
Adding an Output Port

• To add a red output port to a state, select "Add Port" from the context-menu:
Adding more Objects

- Add another generic object and a "goodbye" object to the canvas.
Connecting States

• To connect states, left-click and drag from a red port to a state:
Speech Synthesis

- Once the states are connected, prompts can be set for each state.
- Prompts are speech messages played to the user.
- Prompts ask a question or provide feedback to the user.
- The CSLU Speech Toolkit uses Festival as speech synthesis system:
  
  http://www.cstr.ed.ac.uk/projects/festival/
Setting Prompts

• Double clicking on the icon, opens the prompt dialog for a state:

> RAD prompt: key_question

<table>
<thead>
<tr>
<th>TTS</th>
<th>Markup</th>
<th>Recorded</th>
<th>Recognition</th>
<th>DTMF</th>
<th>Misc</th>
</tr>
</thead>
</table>

Prompt

Are you ready to learn?

> Rec
Setting Prompts

• Define the next prompt:

RAD prompt: not_ready_to_learn

Come on, it will be fun!
Setting Prompts

• Define the "goodbye" prompt:

RAD prompt: goodbye

- TTS
- Markup
- Recorded
- On Exit

Prompt:
Okay, then proceed to the next lesson.

- Rec
Speech Recognition

- The red ports on the generic objects are the location where speech recognition is configured in RAD.
- If the user speaks the proper words for a particular port, then the program flow will continue through that port.
- RAD only recognises words from a predefined list of words.
- RAD does not support continuous style speech recognition.
Configuring Speech Recognition

- Remember the prompt for "key_question" has already been defined.
- Double clicking on the left red port, opens the vocabulary dialog.
- Add the words yes, ... to the vocabulary and click "Update Empty".

![Diagram showing vocabulary dialog with words and pronunciation entries.]
Configuring Speech Recognition

• Now the right red port of the same state needs to be configured.
• Add the following words to the vocabulary: no, not yet, *any.
• When *any is available, the state will always branch right if recognition was uncertain.
Running the Application

• By clicking on build and then on run the application is executed:

• The following prompts are played during the execution:

  Are you ready to learn?
  Come on, it will be fun!
  Are you ready to learn?
  Okay, then proceed to the next lesson.
Calibrating the Microphone

- If your microphone appears not to be working, go to Preferences > Audio tab and click on Calibrate and then follow the instructions.
Assume a very simple application that contains only the prompt
- Who is the president of the United States?
and the name
  - George W. Bush
in the exit port for this state.

With this configuration the state will only exit if this name is recognized.

RAD can be configured to invoke a repair dialog if it cannot recognize what was spoken.
Repair Dialog
Repair Dialog

- To enable repair, select Preferences from the File menu.
- Open the preference dialog and select Repair.
Default Repair Dialog

- A repair dialog is a special type of subdialog.
- It steps in, if recognition fails.
- First, the default repair will respond with "Sorry" and then re-attempt the prompt and response.
- Second, the default repair will respond with "I still don’t understand" and then re-attempt the prompt and response.
- Third, the default repair will repeat "I still don’t understand" and then halt the application.
Default Repair Dialog

• You can view the default repair be selecting it from the View menu:
Default Repair Dialog
Variables

- You can set a variable, assign it a value, and use it in a prompt.
- Therefore, you use an action object and name it "set_variable".
### Setting Variables

- Double click on the "set_variable" state to open an action dialog.
- Set the Tcl variable `topic` to "Language Technology":

```
set topic "Language Technology"
```
Using Variables

- In the "use_variable" prompt, place *Welcome to $topic*.
- At runtime the value for $topic will be substituted in the prompt:
Exploring the CSLU Speech Toolkit

- The goal of the practical part is to make experiences with the toolkit:
  - Build the application "ready to learn" discussed in this lecture.
  - Check out the default repair dialog of the toolkit.
  - Try to understand how TCL variables are used in the toolkit.