Speechbuilder Tutorial
Speaker Independent; Domain Dependent

- What is a domain?
  - a vocabulary (words)
  - sentences
- How to define words?
  - English spelling and pronunciation
- How to define sentences
- Grammar
Speechbuilder

- Galaxy is the speech recognition system
- Speechbuilder is a tool to develop a domain for galaxy
- Real speech recognizers take a lot of work and detailed knowledge of all the components.
- Speechbuilder is great for prototyping
Galaxy’s Components

- Language Generation
- Speech Synthesis
- Audio
- Speech Recog.
- Language Processing
- Dialogue Management
- Context Resolution
- Database Server
- Frame Relay Server
- SpeechBuilder Server

Connections:
- Hub
- Hub to Database Server via TCP Socket
- Hub to SpeechBuilder Server via Http
- Hub to Application (cgi) via Http
- Hub to Application (Python, Java, ...) via TCP Socket

Application (Python, Java, ...)

MIT 6.893; SMA 5508  Spring 2004  Larry Rudolph  Lecture  Introduction
Speechbuilder API

- Galaxy meaning representation provided through frame relay
- Applications connect via TCP sockets
- API provided in Python, Java, Perl

Galaxy Frame Relay

TCP Socket

galaxy.frame.Frame methods:
  - getAction()
  - getAttribute(attr_name)
  - getText()
  - toString()

galaxy.server.Server methods:
  - Constructor(machine, port, ID)
  - connect()
  - processMessage(blocking)
  - disconnect()

Python class
galaxy.server.Server

Python API

Application

Python class
galaxy.frame.Frame
Grammar

What is a grammar?

- a set of terminals
  - A, B, ...
- a set of rules or productions
  - \( <\text{nt-1}> = B | <\text{nt-2}> A \)
  - \( <\text{nt-2}> = <\text{nt-1}> | \text{NULL} \)
- a sample sentence: B A A A
- \( \text{nt-1} \rightarrow \text{nt-2} A \rightarrow \text{nt-1} A \rightarrow \text{nt-2} A A \rightarrow \text{nt-1} A A \ldots \)

Can you explain this to Grandma?

- would probably use examples
Speechbuilder’s Grammar

Attributes
- think of them as: terminals
  - actually, a non-terminal that goes to a terminal

For example
- A set of terminals: lights, microwave, toaster, vcr, tv
- These are all “objects”
- So, “object” would be an attribute

Another example
- dining room, living room, kitchen
- “room” is the attribute
What does a rule look like?

- Speechbuilder calls them “actions”
- No complicated productions
- Each action is an example sentence
- Sentence contains
  - an “action” terminal
  - zero or more attributes
  - optional words
- E.g. Turn on the lights
  - “lights” is an example of an “object” attribute
  - “on” is an example of an “onoff” attribute
  - “turn” is an “action”
## Example after reduction

<table>
<thead>
<tr>
<th>All sentences for action turn</th>
<th>Action: turn</th>
<th>What gets sent to application</th>
</tr>
</thead>
<tbody>
<tr>
<td>turn all the lights off</td>
<td>action=turn&amp;frame=(object=lights, onoff=off)</td>
<td></td>
</tr>
<tr>
<td>turn off all the lights</td>
<td>action=turn&amp;frame=(object=lights, onoff=off)</td>
<td></td>
</tr>
<tr>
<td>can you turn all the lights off</td>
<td>action=turn&amp;frame=(object=lights, onoff=off)</td>
<td></td>
</tr>
<tr>
<td>can you turn off the living room lights</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
<tr>
<td>can you turn off all the lights</td>
<td>action=turn&amp;frame=(object=lights, onoff=off)</td>
<td></td>
</tr>
<tr>
<td>can you turn off the lights in the living room</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
<tr>
<td>can you turn the living room lights off</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
<tr>
<td>turn off the living room lights</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
<tr>
<td>can you turn the lights in the living room off</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
<tr>
<td>turn off the lights in the living room</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
<tr>
<td>turn the living room lights off</td>
<td>action=turn&amp;frame=(object=lights, onoff=off, room=living+room)</td>
<td></td>
</tr>
</tbody>
</table>
Domain XML example

```xml
<class name="object" type="Key">
  <entry>(television l tv) {television}</entry>
  <entry>lights</entry>
  <entry>microwave</entry>
  <entry>toaster</entry>
  <entry>v c r {VCR}</entry>
</class>
```
Domain XML example

<class name="onoff" type="Key">
  <entry>lit {on}</entry>
  <entry>off</entry>
  <entry>on</entry>
</class>
<class name="turn" type="Action">
  <entry>[can you] [please] turn all the lights off</entry>
  <entry>[can you] [please] turn off all the lights</entry>
  <entry>[can you] [please] turn off the (living room lights | lights in the living room)</entry>
  <entry>[can you] [please] turn the (living room lights | lights in the living room) off</entry>
</class>
<class name="status" type="Action">
  <entry>[can you] [please] tell me I do you know) (what | which) lights are on</entry>
  <entry>[can you] [please] tell me I do you know) if the (lights in the kitchen | kitchen lights) are on</entry>
  <entry>(is | are) the (dining room television | tv in the living room) On or Off</entry>
  <entry>(is | are) the (dining room television | tv in the living room) on</entry>
</class>
<class name="good_bye" type="Action">
  <entry>good bye</entry>
  <entry>later</entry>
</class>
<class name="room" type="Key">
  <entry>dining room</entry>
  <entry>kitchen</entry>
  <entry>living room</entry>
</class>
What happens to domain XML

- Compile the domain
  - check for errors
- Can look at reduced sentences
- DON’T click run (it will not work)
- Can download xml (if you want)
- Will start galaxy on ocha.csail.mit.edu
  - using command oxclass.cmd yes yes yes
- startup Galaudio and python on ipaq
Important stuff

- ipkg’s
  - galaudio
    - does end of sentence detection (and a little more)
    - sends waveform to galaxy
    - receives waveform from galaxy
- python classes for galaxy and xml
  - use pydoc to get documentation on these
  - need to register with frame-relay to get xml
- to modify domain (advanced)
  - modify xml of domain, compile, and restart