

Speech Dispatcher

Hynek Hanke

Brailcom, o.p.s.

Issues with speech output for user applications

- Dependence on particular synthesizers
- Controlling synthesizers not easy
- More information on the screen that can be output by speech

What application developers need to do

- Write an entire family of drivers for different synthesizers
- Figure out how each synthesizer works
- Solve the conflicts of simultaneous speech
- Use speech synthesis

The first 3 steps are application independent and quite difficult

The last step is relatively easy (speechd-el, speechd-up)

Speech Dispatcher

- Uniform interface to speech synthesis through SSIP
- Output modules for different synthesizers
- Thinking in the terms of user interface events
- Message priority model

What a very high-level API is to graphical user interfaces, Speech Dispatcher is to applications which want to use synthesis.

Architecture and Speech Synthesis Interface Protocol (SSIP)

- Client/server architecture
- Provides a reasonable command set (no need to reinvent it)
- Switching of contexts (parameters, context pause, ...)
- Supports SSML inside messages
- (Callbacks)

Output modules

- Festival - our primary synthesizer
- Flite - fast, but with limited capabilities
- Generic module - support for simple command-lines clients (dtk-generic, epos-generic)
- Hardware synthesizers?

Message priority model

- 5 priorities with different purposes (important, message, text, notification, progress)
- Interaction of messages from one client application
- Interaction of messages from different client applications

Interface libraries

- C API: `spd_open()`, `spd_say()`, `spd_stop()`, `spd_set_voice_rate()` etc.
- Elisp API
- Common Lisp API
- Guile API
- Python API

Speech Dispatcher and Gnome-Speech

Gnome-Speech tries to solve the same issues under the Gnome Accessibility Project

- TCP/IP protocol with interface libraries vs. ORBit
- Fixed parameter set vs. variable parameter set
- More features vs. more supported synthesizers
- ...

Questions

- Speech Dispatcher and gnome-speech: Do we really need two separate projects? How to cooperate?
- Are there any other interesting projects that could use Speech Dispatcher, but they don't? If so, why? What can we improve?
- What features are missing in Speech Dispatcher as a speech output server for other projects?
- Callbacks: How to communicate them back to client applications over a socket?
- Audio output library or server: Where to find one?