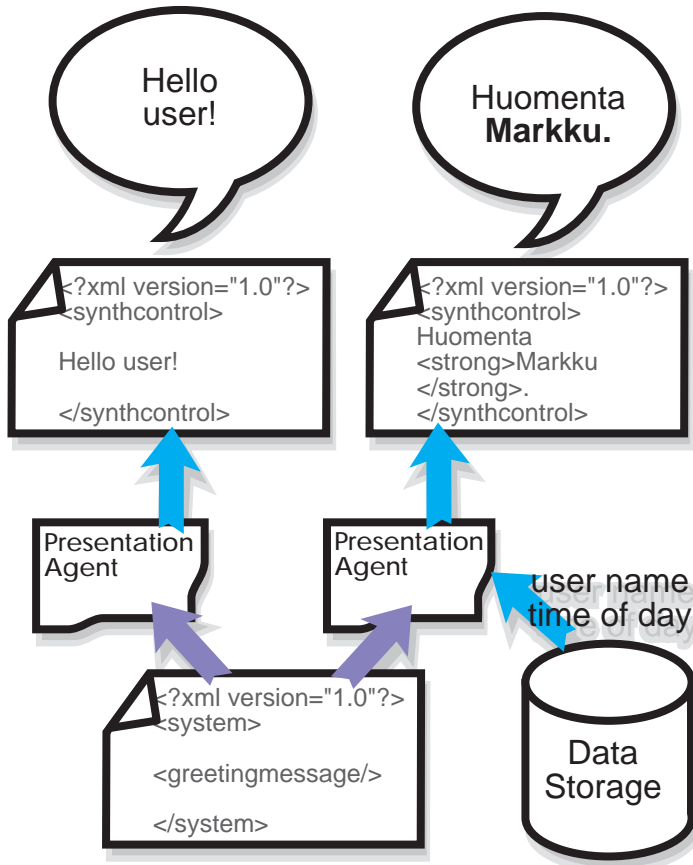




Jaakko Hakulinen, Markku Turunen
University of Tampere, Human Computer Interaction Group

INTRODUCTION

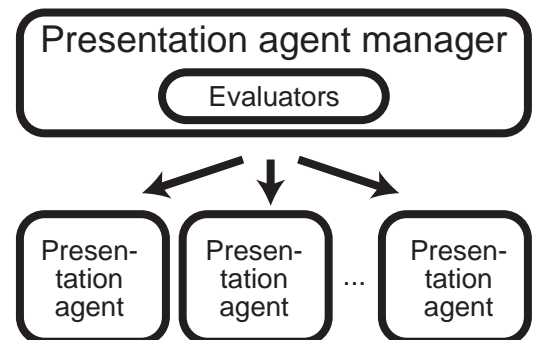
We introduce a presentation agent framework for speech applications. In this framework presentation agents are used to produce dynamic, adaptive and prosody rich speech outputs. Agents receive a very high level conceptual messages that they convert to actual speech outputs. This gives them very much freedom in the process and extends their task beyond traditional natural language generation.



A java version of this framework is implemented in our speech-only e-mail reader. We have been able to use it for example to handle multilingual issues. Our goal is to build unique computer 'voices' to make speech outputs more intelligible and pleasant for the users.

Benefits:

- Supports concept-to-synthesis paradigm**
- Supports different computer voices**
- Support for multilinguality**
- Can adapt to users**
- Can adapt to different situations**
- Can take advantage of dialogue history**



Presentation Agent Framework

In our framework outputs of speech applications are generated in software components called **presentation agents**. Typical application contains many specialized presentation agents; usually one agent is optimized for certain kind of messages. However, in order to adapt to different situations and users there can still be several agents that handle same kind of messages. Information for adaptive natural language outputs is achieved from the content of the information storage.

Presentation manager controls the agents and other components in our framework. It receives conceptual messages from dialogue management and chooses an appropriate agent to present the message. Since both situations and users differ greatly, the problem of choosing the best possible agent is not a trivial one. Presentation manager consults a set of **evaluators** to decide which agents it should choose.

Single evaluator compares the attributes of a presentation agent to the information about the current situation (obtained from information storage) and gives a score to the agent. When all evaluators have evaluated every agent, presentation manager multiplies scores to get a final score for every agent and then selects an agent with the highest score.

When the most suitable agent is found, presentation manager uses that agent to handle the message it has received from the dialog manager. The agent generates a document suitable for speech output and the document is returned to the dialogue management.

A software architecture for presentation agents framework

