

Summary on Towards Multimodal Dialogue Management

This paper addresses issues related not only to speech-only system but also to multimodal dialogue system.

The author indicates there are five main characteristics that a speech system may include, namely, global structure, mixed initiative, over-informativeness, contextual interpretation and failure repair.

To illustrate the system has includes these five characteristics, the system is introduced along with four basic principles.

The proposed system is consist of solutions to all components of a speech-driven system, Interpretation (Semantic techniques, e.g. how discourse is modeled?), Ontology modeling (Task techniques, e.g. what is required in order to perform a task?) and system behavior (Dialogue techniques, e.g. how to realize user goal?)

Among all interesting parts, the paper implicitly brought up an idea for the relationship between mix-initiative and failure repair. Due to the recognition (speech/intention) accuracy drops, the system may have to reduce the level of flexibility introduced by mix-initiative utterance control. The worst case may be the system doesn't read what the user says at all, the system might fall back to menu-driven utterance control.

What is multimodal system? Multimodal system is a system which allows user to interact through more than one form of communicative interface, e.g. graphical user interface and voice interface.

The issues of multimodal system are discovered around the co-operability between the GUI and VI, namely, how speech system handle non-speech input and output, how to reference to visual object, how to choose a modal to perform IO, etc.