Docket #: S04-337

A Multi-Domain Dialogue Management System

Stanford researchers have developed a robust dialogue management system for any audio controlled devices, including cell phones, computers, etc. The CSLI Dialogue Manager (DC) is a system for mediating and understanding spoken-language and multimodal (speech and graphical UI) dialogue for the purpose of supporting spoken-language and multimodal interaction with intelligent devices.

The DM supports natural interaction, including allowing users to interrupt. The DM is aware of the capabilities of the device with which it mediates interaction and the status of tasks being performed by the device, and all communication is interpreted within that context. The DM also reports on the changing status of tasks it is performing and is able to handle queries.

A central data structure of the DM for maintaining context is the Dialogue Move Tree, which is an historical account of contributions made to the dialogue from all participants. The tree structure allows the simultaneous management of multiple threads of conversation, potentially involving multiple devices and/or multiple humans. The DM also involves multi-threaded execution, allowing interruption of the interpretation and processing of a dialogue contribution, enabling speakers to interrupt each other as appropriate.

The DM is built using an open software architecture and is reconfigurable to different domains without requiring re-coding of core infrastructure, although all core processes may be easily extended or enhanced to address domain- or application-specific issues, without changing the core code-base. The DM has been built to work with different parsers, speech-recognizers, language-generation, speech-synthesis components, structured knowledge-bases, and devices, without modification behind well-defined APIs (interfaces).

Applications

- Cell phones
- Computers

Advantages

- Supports natural interaction
- Can involve multiple devices and/or humans
- Allows interruption of interpretation
- Reconfigurable to different domains without requiring re-coding of core infrastructure
- Built to work with different parsers, speech-recognizers, language-generation, speech-synthesis components, structured knowledge-bases, and devices without modification behind well-defined APIs (interfaces)

Patents

- Published Application: 20060271351
- Published Application: 20060271364
- Published Application: WO2006130314
- Issued: 8,041,570 (USA)
- Issued: 7,904,297 (USA)

Innovators

- Stanley Peters
- Lawrence Cavedon
- Institutional Work

Licensing Contact

Mona Wan

Senior Licensing Manager for Special Projects

Email