Considerate Supervisor for Multiparty Conference Call

Rahul Rajan, Cliff Chen, Peter Ming Gao, Christine Chien, Ian Lane, Ted Selker

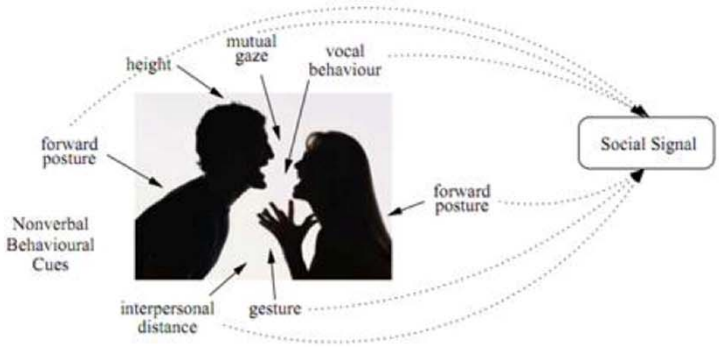
Background / Problem

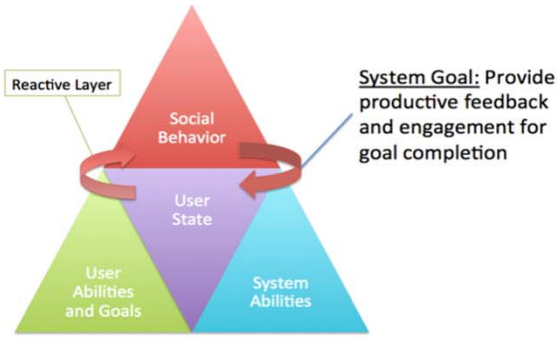
A significant amount of communication between people occurs non-verbally. Social signals such as body language and eye contact have been shown to predict behavioral outcomes in important types of social interactions. While this information is available in co-located settings, it is not for remote audio-only meetings.

Audio-only conference calls lack a significant amount of the social context available;

- Who is here?
- Who is currently speaking?

The need for meeting facilitation in distributed environments is quite pronounced, however a) there has been little work towards resolving these issues, and b) it is impractical for most organizations to have a full-time meeting facilitator.





Solution

Build a "considerate" meeting supervisor to provide non-intrusive:

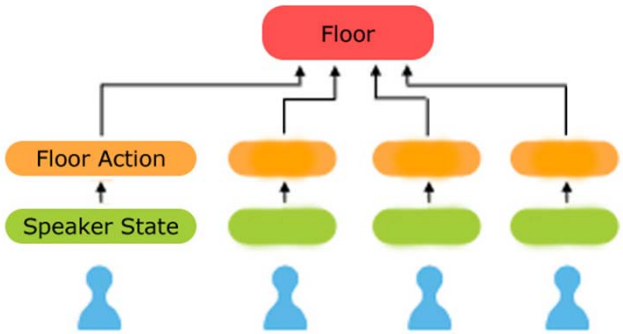
- Feedback
  - Discourage overly dominant activity
  - Mitigate collisions by allowing dormant participants the floor
  - Volume level
- Feedforward
  - Intonated entry and exit
  - Participant presence via background tracks
  - (pitch change delay?)
- Floor Status
  - Graphical User Interface
  - Speaker Activity

Approach

Using Participant States, Floor Actions and Floor States we are able to discern the social dynamics of a conversation by looking at audio cues such as:

- Total Speaking Length
- Total Successful Interruptions
- Total Unsuccessful Interruptions

- Built on top of:
  - CLAM C++ Audio Library
  - JACK Audio Connection Kit
  - Ruby on Rails Server



Future Work

- State of Dominance vs Acts of Dominance
  - Can we use a Markov model or n-gram to better predict dominance?
- 3D Audio
  - Use spatialization to take advantage of the "Cocktail Party Effect"
- User studies and experimentation