1 Purpose

I thought it would be interesting to present a quick visual image of what a conversation is about. Ideally the image would convey not only the content but also some of the structure of the conversation: what concepts are involved, and how are they connected? Who talked about what? We are used to thinking of conversations as linear, but actually they tend to slowly travel from topic to topic, sometimes bringing up new points and sometimes connecting earlier points. Both time-dependent and summarized information should be visible.

2 Visualization

The summary visualization is in the form of a “concept network.” To find out what concepts are part of the conversation, you have to have speech-to-text. This is probably the hardest from an implementation standpoint. But given that you can find some of the words, especially nouns, and can place them in time, a representative image from the Internet replaces the concept. The image thumbnails are scaled according to their frequency in the conversation. Connections between concepts can found a variety of complex ways, but the simplest might be to count their adjacency in the conversation. If two keywords are mentioned in the same sentence, they are very strongly linked, especially if this happens more than once. All the words of the conversation are therefore linked, but a cutoff can be applied based on the desired level of connectedness in the resulting graph.

The time-dependent visualization probably cannot run in real time because of the difficulty of speech-to-text, but the time can be simulated after the analysis is done. For this visualization, conversants are placed around a central circle. As they speak in (simulated) real time, the key words are converted to images and move inwards toward the center. When they reach the center, they stay for a while but shrink as new concepts enter the middle. This allows you to see the thread of conversation, as well as the “central topic” at any given time.
3 Implementation

Good libraries will probably be essential for this visualization. Sphinx is an open-source Java based speech-to-text library which boasts a 64,000 word dictionary. WordNet provides word frequency information which can be used for TF-IDF. It also has synonym and worse sense relationships which can be used to find a single representative concept word for a number of synonyms in the conversation. Finally, based on some quick tests of different public image APIs, it seems that Google’s first image result is usually pretty representative of a concept word, or is at least recognizable in the context of the conversation. Finally, for drawing, probably Processing would be a good choice since all the previous libraries have bindings in Java.

4 Example

The following is an example of a summary graph of a conversation. A test conversation was produced, but the most relevant concept words were picked arbitrarily. The images were always the first returned by a Google image search.
Next is a time slice of the same conversation might look in the (simulated) real time view.

5 Other Ideas

Mouse-overs could do some obvious things, such as revealing what words the images represent; displaying the sentence that most strongly links two concepts; displaying all the things a given conversant said; and controlling the rate of (simulated) time, or readjusting the concept graph if the layout algorithm didn’t do a good job.