

## Persistent Conversation: A Dialog Between Design and Research

Thomas Erickson  
*IBM T. J. Watson Research Center*  
*snowfall@acm.org*

Susan Herring  
*Program in Linguistics*  
*University of Texas at Arlington*  
*susan@ling.uta.edu*

Conversation is one of the most important activities in which humans engage. It is through conversation that we portray ourselves, advancing our personal goals, taking credit, sharing blame, and accomplishing other social ends. Conversation is also a medium for knowledge creation and validation, decision making, negotiation, and conducting business. Although often taken for granted, the mechanisms underlying conversation are highly complex, involving coordination of speaker change and other precise timing phenomena, tracking of reference across extended exchanges, avoidance of and repair of misunderstanding, and constructing appropriate inferences in context.

In ordinary face-to-face conversation, non-verbal cues facilitate these complex tasks. People express their attention and level of understanding via nods, eye contact, facial expressions, yawns, fidgeting and the like. Valuable cues are also provided by prosody: tone of voice, pitch, loudness, and pausing can signal speaker affect, turn boundaries, and cross-turn coherence.

When conversation takes place in the digital medium, however, such cues are typically absent. In text-based computer-mediated communication (CMC), the audience we are used to seeing and hearing becomes effectively invisible and silent. Moreover, messaging systems impose their own constraints on turn-taking. It isn't surprising, therefore, that digital conversations are cumbersome, particularly among more than a few participants, and often lack coherence, tending towards drift and dissolution.

The changes wrought by the digital medium are not all negative, however. The persistence of CMC expands conversation beyond those within earshot, making it accessible to people in other places and at later times. CMC may be more structured, or more amorphous, than an oral exchange; it may have the formality of published text or the informality of chat. The persistence of such conversations enables them to be stored as archives, which can then be used in a variety of new ways. These new uses raise many

questions, from the ways in which the persistence of conversation reshapes everyday work practice, to the social and ethical implications of persistent text.

### The Minitrack

This minitrack explores persistent conversation, the transposition of ordinarily ephemeral conversation into the potentially persistent digital medium. The phenomena of interest include conversations carried out using email, mailing lists, news groups, bulletin board systems, textual and graphic MUDs, chat clients, structured conversation systems, document annotation systems, etc. The persistence of such conversations as computerized records, although variable in duration and ease of user access, gives them the potential to be searched, browsed, replayed, annotated, visualized, restructured, and recontextualized, thus making possible a variety of new uses and practices.

The aim of the minitrack is to bring together researchers who analyze existing computer-mediated conversational practices and sites, with designers who propose, implement, or deploy new types of conversational systems. By bringing together participants from such diverse areas as anthropology, computer-mediated communication, HCI, interaction design, linguistics, psychology, rhetoric, and sociology, we hope that the work of each might inform the others, suggesting new questions, methods, perspectives, and design approaches relating to the theme of persistent conversation.

### Session 1

The first session begins with David Millen's examination of an online community of journalists who communicate over a private mailing list. Millen focuses on the group's use of its on-line conversation archives, and documents the awareness and use of the archives by com-

munity members, as well as evidence for the development of social norms for using the archives. Millen also describes the use of conversation maps, an analytical tool for visualizing the 30 months worth of data in question.

Next, Nonnecke and Preece report on the phenomenon of lurking — the tendency of a participant in a communications forum to read but not to respond. They interviewed 10 online community participants, not preselected for lurking, and found that all participants reported lurking in at least one of the communities in which they were involved. Nonnecke and Preece discuss some of the motivations for lurking, the ways in which these are entwined with the persistence of the conversation, and some of the implications of lurking for systems design.

The first session concludes with an examination of hate speech on the internet by Rosemarie Coste. Coste describes the use of the internet to present, disseminate, and discuss white supremacist David Duke's book, *My Awakening*. Of particular interest is Duke's supporters' use of seemingly neutral fora (such as Amazon.com's online book reviews) to present their message to a larger audience and avoid censorship by hate filters. Coste concludes by suggesting that tools which would allow readers to recontextualize isolated messages might be of value in these circumstances.

## Session 2

The second session focuses on ways of dealing with very large collections of messages. First, John Paolillo explores the use of techniques from corpus linguistics to visualize and present Usenet discussions. By extracting features from the messages and subjecting them to common factor analysis, he shows that it is possible to create message spaces in which messages are organized by language, topic of discussion, or frequency of interaction.

The second paper of the session, by Warren Sack, also addresses Usenet message archives. Sack takes analytical

representations — semantic networks and social networks — and recasts them as interfaces to very large scale conversations. He illustrates this approach with NB, an implemented system which allows very large scale conversations to be browsed using an interface which combines social and semantic network representations with indications of message topic and thread structure.

## Session 3

Session 3 begins with a paper by Shankar, VanKleek and Smith which describes a novel approach to designing a synchronous chat system. Drawing on studies of face-to-face conversation, Shankar, et al. argue that rhythm is an important component of conversation which is not supported by today's systems. They describe their Fugue system, which uses representational ideas from musical scores to graphically portray temporal and rhythmic characteristics of online textual conversations.

Next, Steven Thorne describes a two-level theoretical approach to studying CMC that is rooted in activity theory. Thorne demonstrates this approach by studying the use of a MOO environment aimed at supporting language learning. Drawing on both interview data and transcript analysis, he argues that it is not sufficient simply to look at the discourse within the CMC system, but that activity in the wider cultural context (including previous Internet experience) must also be taken into account.

The session concludes with a paper by Marko Balabanovic that describes the design of multimedia chronicles. Multimedia chronicles are persistent audio narrations that weave through and refer to an array of documents, URLs, photographs, and other digital objects. Visually, a multimedia chronicle consists of graphic representations of the audio narratives, deictic references, and the digital objects referenced. As these are browsed and added on to over time, they can come to form a rich organizational memory.