

Current and New Directions in Discourse and Dialogue

Jan van Kuppevelt and Ronnie W. Smith (editors)
(University of Stuttgart and East Carolina University)

Dordrecht: Kluwer Academic
Publishers (Text, speech, and language
technology series, volume 22, edited by
Nancy Ide and Jean Véronis), 2003,
xii+381 pp; hardbound, ISBN
1-4020-1614-X, \$167.00, €152.00, £105.00;
paperbound, ISBN 1-4020-1615-8,
\$65.00, €59.00, £41.00

Reviewed by
Matthew Stone
Rutgers University

Van Kuppevelt and Smith's book offers a kind of archival proceedings for the Second SIGdial Workshop, which was held in 2001 in conjunction with Eurospeech. SIGdial is the Special Interest Group on Discourse and Dialogue of the Association for Computational Linguistics, and its annual workshop series, always collocated with major events in computational linguistics, has become a premier forum for empirical, formal, and computational approaches to language use. I think the breadth and energy of van Kuppevelt and Smith's collection supports an unequivocal appraisal of these meetings: If your interests touch on discourse and dialogue and your schedule and budget permit, you should be attending them. The next one will be held once again in conjunction with Eurospeech in September 2005 in Lisbon, Portugal. Go.

So should you buy this book? You might be worried about value. Inevitably, Kluwer wants \$167 from your library for the hardcover version. For that you get 12 revised versions of papers whose original versions are, in fact, archived in the ACL anthology. But before you go spooling things off to the printer, Kluwer does have a paperback version, on offer for a relatively modest \$65. Most of the chapters offer substantial additions (typically an extra methodological section) and spell out annotation standards, implementation techniques, or analytical results more thoroughly than the workshop papers. The expanded versions go a long way toward making the individual contributions more convincing and easier to replicate. In addition, the collection includes four invited new chapters, covering dialogue annotation, dialogue pragmatics, dialogue semantics, and dialogue system implementation. With all this new content, I wouldn't have hesitated to order the book for myself.

You might also be worried about the relevance of papers from a workshop from three years ago. In fact, the volume is as selective as any conference proceedings; van Kuppevelt and Smith report that the 12 contributed papers were winnowed down from 57 submissions to the workshop. However, rather than trying to present a concise and self-contained nugget of completed research, many of the most interesting papers articulate and motivate ambitious long-term agendas for dialogue research. The term *directions* in the book's title is thus faithful to its contents. The collection retains interest because it retains the adventurous and open-ended feel of a successful workshop and emphasizes these forward-looking characterizations of central problems and methods in the field. You can still use this book as a worthwhile jumping-off point for framing grant proposals or designing a graduate seminar in discourse and dialogue.

Three (invited) chapters nicely illustrate this flavor in the collection. All three chapters start from successful accounts of the dynamics of collaborative planning (Lochbaum 1998; Chu-Carroll and Carberry 2000), and take up the question of how such dynamics might be mediated by people's and systems' utterances. Jonathan Ginzburg uses corpus analysis to argue that the public contribution of an utterance follows closely from its semantics and is sharply distinguished from the speaker's more general implicatures and motives. While interlocutors' motives are always available as a metalevel topic of conversation, Ginzburg shows, only utterance content is at play in the fine-grained dynamics of clarification and grounding that allows interlocutors to achieve mutual understanding. This compelling new argument suggests that collaborative dialogue systems must signal their intentions publicly, directly in the content of their utterances. This is a challenge to semantics, as we must now represent such content in the "logical form" of discourse, perhaps along the lines being developed by Asher and Lascarides (2003). At the same time, Ginzburg's argument helps us to appreciate the conversational work that dialogue systems can do off the record, where the strategies dialogue agents use, such as indirection and politeness, can foster relationships with and trust among users (Bickmore 2003).

In another chapter, David Traum and Staffan Larsson survey their information-state approach to dialogue management. In essence, this approach offers a general knowledge representation methodology for use in characterizing dialogue context. It invites the designer to identify the meaningful distinctions in dialogue state that participants need to keep track of and to describe declaratively how utterances change it. Traum and Larsson's chapter recapitulates their earlier overview of the approach and survey of information-state systems (Larsson and Traum 2000) while clarifying their contribution to the design of reusable and sharable dialogue components. Although the chapter remains important, it might have been more instructive with a longer explanation of how and why to use the approach, as the approach continues to prove itself across an increasingly broad range of communicative behavior (Nakano et al. 2003) and increasingly varied interactions among communicators (Traum et al. 2003).

Lastly, Nate Blaylock, James Allen, and George Ferguson work to refine the moves of existing collaborative problem-solving models to the level of precision required to describe individual utterances in natural conversation. They build on previous models in factoring the course of collaboration into primitive joint steps, such as agreeing to adopt an action into a plan. (Their steps are particularly flexible because they draw on a somewhat formidable ontology to systematize the ingredients of collaborative planning and collaborative action.) They go on to characterize utterances by linking them with abstract communicative moves that specify one agent's contribution to one of these joint steps, such as initiating or completing one. Perhaps the most substantial insight here is that collaborative dialogues in dynamic domains play out differently because they must interleave planning and action. This is an increasingly important feature of current system domains from search-and-rescue to home automation. There is no going back. However, to go forward, we still need to regiment these collaborative moves to derive principled representations for agents' mental states, mutual commitments, and meanings—offering ample space for future research.

Obviously, an important theme of the collection is the interrelationships among human behavior, theoretical explanations, and system design. Two other themes of the volume deserve mention here as enduring inspiration for ongoing research in discourse and dialogue. The first theme is the challenge and promise of data annotation. For example, in the remaining invited chapter, Niels Ole Bernsen, Laila Dybkjær, and Mykola Kolodnytsky outline the requirements for annotation tools aimed at capturing the complexity of face-to-face conversation. As they outline, an ideal platform

would accommodate a range of users and uses and would support the analysis of communicative behaviors, including all aspects of their structure and interpretation, across all available modalities. This general vision, well-articulated here in one specific formulation, continues to set the agenda for tools and resources in projects like TalkBank (MacWhinney et al. 2004). Similarly, Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski document their work to create a corpus of discourse annotated with hierarchical interpreted structure in the framework of rhetorical structure theory. A set of case studies (new since the workshop submission) draw on their corpus to explore and characterize various levels of discourse: the uses of individual connectives within sentences, the expression of parallels and contrasts across multiple sentences, and the organization of documents as a whole. Annotated discourse corpora continue to be developed (Mitsakaki et al. 2004), so these strategies for mining and visualizing them will surely continue to inform their use.

The other theme is rapid development of dialogue applications. Two chapters in particular, one by Hiyun Alshawi and Shona Douglas, the other by Manny Rayner, Johan Boye, Ian Lewin, and Genevieve Gorrell (both slight revisions of original workshop papers), frame the issues in crisply insightful ways. Useful dialogue applications link domain-independent words and constructions with domain-specific responses. Configuring a dialogue system therefore requires specifications that crosscut traditional units of system design, such as grammatical resources and domain-general and domain-specific interpretive procedures. Building these specifications remains a complex task requiring enormous expertise. Rayner and colleagues aim to simplify these specifications by describing a “plug and play” architecture. System components are described in terms of their domain functionality. At the same time, they are linked with predefined hierarchical collections of linguistic resources that encapsulate the interface for that functionality. With these tools, new components can be characterized by compact specifications that can be integrated seamlessly and elegantly with an existing system. Alshawi and Douglas, by contrast, aim to learn these specifications. They start from examples pairing utterances with their domain-specific interpretations. They reconstruct the interpretation for a new utterance by directly manipulating the interpretation of a similar utterance drawn from this sample set. It testifies to the excitement of dialogue research that the basic problem of linking language to a world model admits such divergent approaches. Indeed, these approaches represent extremes of a huge design space, with alternative delineations of tasks and resources and with different demarcations between what to specify by hand, what to capture directly from data, and what to learn from annotation. This space remains an inspiration for dialogue research, including some of my own, for example (Stone et al. 2004).

Van Kuppevelt and Smith have put together an inclusive, timely, and significant collection. The book has small flaws—notably a few glitches in the typesetting of tables and of references in running text—but I find that my most severe criticism of it is simply that it is a book. In many respects, the closest parallel to it would be an issue of *AI Magazine* or *Communications of the ACM*: a selective forum with a broad audience for exciting, programmatic but precise discussion of important trends in research. This collection made me wonder whether, with its proliferation of lively SIGs and far-flung meetings, ACL might deserve an analogous series.

References

- Asher, Nicholas and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press, Cambridge.
- Bickmore, Timothy. 2003. *Relational Agents: Effecting Change through Human-Computer Relationships*. Ph.D. thesis, Massachusetts Institute of Technology.

- Chu-Carroll, Jennifer and Sandra Carberry. 2000. Conflict resolution in collaborative planning dialogues. *International Journal of Human-Computer Studies*, 53(6):969–1015.
- Larsson, Staffan and David Traum. 2000. Information state and dialogue management in the TRINDI dialogue move engine toolkit. *Natural Language Engineering*, 6:323–340.
- Lochbaum, Karen E. 1998. A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572.
- MacWhinney, Brian, Steven Bird, Christopher Cieri, and Craig Martell. 2004. TalkBank: Building an open unified multimodal database of communicative interaction. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon.
- Miltsakaki, Eleni, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2004. The Penn discourse TreeBank. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon.
- Nakano, Yukiko I., Gabe Reinstein, Tom Stocky, and Justine Cassell. 2003. Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, Sapporo, Japan.
- Stone, Matthew, Doug DeCarlo, Insuk Oh, Christian Rodriguez, Adrian Stere, Alyssa Lees, and Chris Bregler. 2004. Speaking with hands: Creating animated conversational characters from recordings of human performance. *ACM Transactions on Graphics*, 23(3):506–513.
- Traum, David, Jeff Rickel, Jonathan Gratch, and Stacy Marsella. 2003. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 441–448, Melbourne.

Matthew Stone is an assistant professor in the Computer Science Department and the Center for Cognitive Science at Rutgers, the State University of New Jersey. Stone's research explores the role of representations of pragmatic interpretation in explaining human-human dialogue and constructing conversational systems. Stone's address is 110 Frelinghuysen Road, Piscataway NJ 08854-8019; e-mail: mdstone@cs.rutgers.edu; URL: <http://www.cs.rutgers.edu/~mdstone>.