

Views on Mixed-Initiative Interaction

David R. Traum

UMIACS, University of Maryland

A. V. Williams Building

College Park, MD 20742 USA

traum@cs.umd.edu

The topic of initiative is important to a greater understanding of collaboration and interaction. It is a complex characterization of several aspects of dialogue, relating to whose purposes are being concentrated on, and who is in “control”. Initiative is important for characterizing both the cooperative and competitive aspects of dialogue, including when and how to respond to the other, as well as whose agenda to follow.

The concept of initiative is closely related to similar concepts from games, such as *tempo* in chess, or *sente* in go. These zero-sum games have rigid turn-taking rules, in fact each player is allowed one move followed by one move of another. Initiative (e.g., *tempo* or *sente*) is then a second order concept: it relates to whether, given the turn, is a player free to use it to work towards his own goals, or is he “forced” to use it in reaction to the lead of the other player. A *mixed-initiative* game, would thus be one in which the initiative shifts from one player to the other throughout the course of the game. An alternative definition, applying to abstractions of games rather than actual playings, is one in which either player *could* have the initiative, according to the rules, (regardless of whether in any specific playing both parties actually did have the initiative at some time).

Such concepts can be easily applied to language situations, in which the “move” is a speech act or dialogue move. A speaker would have the initiative if the speaker had the choice as to the content of the utterance, while the other speaker would have the initiative if the speaker had to frame the utterance in response to speech by the other speaker. For some sorts of interaction, such as a question answering system, there is a clear notion of fixed initiative, when one party is responding to requests or directives of the other. Other situations could be mixed-initiative, where initiative can change through the dialogue, and is not assigned in advance.

A general dialogue situation is more complex than the above picture in two ways, though. First, the turn-taking system can be much more flexible: in general a speaker can continue with several moves or utterances, some of which are in response to the other

speaker, and some of which are not. Secondly, part of the coherence of dialogue itself, the feature that distinguishes it from a random collection of utterances is this very notion of coherence with contributions by the other speaker (e.g., see (Sperber & Wilson 1986)). Thus just in virtue of being a coherent conversation, some aspects of a speaker’s turn will be a response to the previous speaker. One of these aspects is *grounding*, establishing the common ground about what has been said (Clark & Schaefer 1989; Traum & Allen 1992; Traum 1994). Given such a flexibility of turn-lengths and contents, and a requirement to reply to some degree, we can recast the definition of initiative in more unrestricted dialogue as whether a speaker will *only* respond to the other speaker, or will instead also continue with the speaker’s own objectives.

Two questions arise now, given such a definition: how is it used by a speaker (e.g., is it just a descriptive characterization of dialogue, or can it be used by a speaker in formulating the size and content of a dialogue contribution?), and how can it be determined who has the initiative. Related to these points is the question of how one can change who has the initiative.

The question of how to tell who has the initiative can be a complex one – as complex as the general problem of deciding what an agent should say. Some (e.g., (Smith & Hipp 1994)) have used a preallocated scheme, where the initiative is decided in advance for a particular dialogue. (Walker & Whittaker 1990) based the determination of initiative on the utterance type in a turn transition, e.g., a command or question gives the speaker the initiative (they used the term *control*) through the turn of the next speaker, while a prompt gives the initiative to the other speaker. A declarative utterance gives control to the speaker unless it is in response to a question.

My own work has concentrated on the role of mental states and social attitudes. Rather than computing initiative directly from individual utterances, the computation was from the relations of mental states, and the role of utterances was to update these mental states. In (Traum & Allen 1994), a speaker who didn’t have the initiative was modeled as an agent who had an

obligation to the other agent to respond. Such an obligation could be imposed by asking a question (though also through other acts). Characterizing an agent as pre-disposed to have or not have the initiative was done by comparing the relative weighting the agent gave towards following obligations vs. acting on its own goals. A very cooperative agent (such as the TRAINS-93 system citeallen-et-al:jetai95) puts a premium on fulfilling its obligations, though the TRAINS-93 system could still get the initiative when this was not seized by the other agent. For task-oriented dialogue, the higher-level goals and expectations may also play a role, particularly when the participants have unequal abilities and roles. E.g., in the TRAINS domain, the total task of reaching a joint plan was divided into three phases: goal specification, plan proposal, and verification. The user has, by default, the initiative in the goal specification phase, since only the user knows the goal. Likewise, the system has the initiative in the verification phase, since it has privileged access to some aspects of the domain, such as how long it takes to get from city to city, or whether any obstacles exist. The middle phase of actually producing the plan was not pre-allocated by the domain itself, but the instructions (for the human-human dialogues) and the programmed behavior of the system allocated initiative in this phase to the user, as well.

There is also a range of applicability as to how to use the information about initiative. The system described in (Smith & Hipp 1994) used the notion to decide whether to respond to its own goal structure, or the one it assumed the user to have. Also, the type of utterance was based on the level of initiative, according to the styles of (Walker & Whittaker 1990). Since each dialogue participant made only one dialogue move per turn, this was an either/or choice. One problem, though, is that in directive mode, the system would completely ignore the dialogue obligations imposed by the user, and the resulting replies often had little dialogue coherence.

In the TRAINS-93 system, the direct decisions about content were made in reference to the combination of mental states, not by directly considering initiative (although, as stated above, the two are related). The deliberation procedure hard-coded obligations as more important than the system's goals, so that it would give up the initiative whenever the user tried to get it. When there were no current obligations, however, the system would take the initiative and work on it's own goals.

One place where initiative was explicitly used, however, was as a guide to turn-taking decisions, including both recognizing turn-endings and deciding when to continue speaking beyond a simple reply. Unlike the system of (Smith & Hipp 1994), the TRAINS-93 system did not have any specifications of how many moves a turn should contain. Turn-taking acts (Traum & Hinkelman 1992) were used to determine when the

turn was changing, but the system still needed to recognize the acts of the other in the perceived utterances, and needed to determine when the system should perform such acts. Some turn-taking acts were clearly recognizable through surface cues (such as intonation, continuing to speak, or asking a question), but other cases were more ambiguous, such as a plain declarative utterance. Also, from the system's point of view, once it had answered a question, should it then relinquish the turn, or proceed with a further utterance towards satisfaction of its goals, such as a followup suggestion? This decision was reached by considering who had the initiative – in the ambiguous cases, the turn should revert to the conversant with the initiative – so a speaker with the initiative should continue after a declarative utterance, while one without should not.

Some questions still remain about initiative: first, is it an essential or useful notion for a dialogue participant (especially an artificial system)? Is it important that the participant realizes who has the initiative, and act accordingly, even strategically adjusting who will hold the initiative in the future, or can everything be achieved through other means (such as reasoning about combinations of mental states or discourse coherence). For zero-sum games, it is clear that the related notions are important for successful play (though not for legal play), though in a more cooperative setting such as most task-oriented dialogue, it is less clear. Another question is how such a concept applies when the communicative setting is generalized to other modalities than speech (or speech-like text-messaging), where the notion of turn-taking is not as relevant. In such settings, where there are other communicative actions, such as drawing and pointing, and where domain action is also intermingled with conversation, there are still such notions as obligations to respond, though the means may not be as sequential, and both agents may often act simultaneously without loss of comprehensibility.

References

- Clark, H. H., and Schaefer, E. F. 1989. Contributing to discourse. *Cognitive Science* 13:259–294. Also appears as Chapter 5 in (Clark 1992).
- Clark, H. H. 1992. *Arenas of Language Use*. University of Chicago Press.
- Smith, R. W., and Hipp, D. R. 1994. *Spoken Natural Language Dialog Systems: A Practical Approach*. Oxford University Press.
- Sperber, D., and Wilson, D. 1986. *Relevance: Communication and Cognition*. Harvard University Press.
- Traum, D. R., and Allen, J. F. 1992. A speech acts approach to grounding in conversation. In *Proceedings 2nd International Conference on Spoken Language Processing (ICSLP-92)*, 137–40.
- Traum, D. R., and Allen, J. F. 1994. Discourse obligations in dialogue processing. In *Proceedings of the*

32th Annual Meeting of the Association for Computational Linguistics, 1–8.

Traum, D. R., and Hinkelman, E. A. 1992. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence* 8(3):575–599. Special Issue on Non-literal language.

Traum, D. R. 1994. *A Computational Theory of Grounding in Natural Language Conversation*. Ph.D. Dissertation, Department of Computer Science, University of Rochester. Also available as TR 545, Department of Computer Science, University of Rochester.

Walker, M. A., and Whittaker, S. 1990. Mixed initiative in dialogue: An investigation into discourse segmentation. In *Proceedings ACL-90*, 70–78.