Learning to Communicate – Cooperation and Interactive Systems

Kristiina Jokinen
University of Tampere and University of Helsinki

Keywords: interactive systems, dialogue management, rational communication, Ideal Cooperation

Abstract

Language and interaction have been crucial for the development of human intelligence. Also in the development of technology for complex interactive systems, understanding of human communication is necessary for furnishing the systems with intelligent interaction and learning capabilities: intelligent communication requires an ability to learn from interactions and to build internal knowledge about how to reach one’s goals. In this paper I will discuss some challenges for interactive systems in the context of information technology advancements, and argue that the systems need to be equipped with reasoning capabilities related to rational and cooperative communication. I will present the Constructive Dialogue Model approach which is inspired by Allwood’s notion of Ideal Cooperation, and aims at addressing challenges for flexible and natural dialogue management.

1. Challenges for Interactive Systems

One of the main challenges for dialogue management concerns the system being situated in a dynamic and multi-agent environment where communicative situations do not only comprise fixed interactions between users with specific knowledge and requirements, but also simultaneous interactions with different users and with other computing systems. The environment for interaction has thus expanded from a shared “closed-world” context into an uncertain world, and interaction itself requires understanding of communicative principles in order to construct shared albeit limited knowledge of the goals and intentions of the partners. An important prerequisite for the novel situations is the agent’s communicative competence, i.e. ability to cooperate and collaborate in order to coordinate interaction (e.g. turn-taking, giving feedback, and constructing the shared knowledge) and to manage the incomplete, possibly erroneous information that is to be exchanged in the course of the interaction. Such competence comprises factual knowledge about the world: activities, and entities involved, as well as attitudinal knowledge about the participants’ beliefs and attitudes. It also includes strategic knowledge learnt through the agent’s engagement in communicative situations, accumulating knowledge about interaction strategies: through positive feedback of successful actions and through lack of success of failed communicative actions, both of which suggest ways to modify one’s knowledge and behaviour in order to manage similar situations appropriately in the future.

---

1 Currently also Visiting Professor at the University of Tartu.
2 See future views on technology development e.g. in Sjöberg and Backlund (2000), and requirements for dialogue systems in Jokinen (2000).
Jokinen (2007) classifies the requirements for advanced system’s communicative competence as follows: physical feasibility of the interface, efficiency of reasoning components, natural language robustness, and conversational adequacy. The first requirement refers to the basic enablements for communication (as defined by Allwood 1976), and is related to the system’s usability and transparency from the point of view physical interaction. It includes such issues as multimodal input and output capabilities, voice quality, and ergonomy of interface devices. The efficiency of reasoning components concerns the system’s performance, and includes fast algorithms, suitable architectures as well as the system’s ability to provide appropriate responses in various problematic situations (misheard words, uncertainty of what is being referred to, etc.). Robustness of the system’s linguistic components is related to the natural language interaction, and the system needs to address linguistic variation and multilingual issues. Also the interpretation and generation of utterances require dynamic processing capabilities, and if the system is to exhibit multimodal capability, also non-verbal communication needs to be taken into account. Finally, the advanced system’s communicative competence requires conversational adequacy. This means that the system can manage rational communication with such intangible aspects as the speakers’ intentions, beliefs, and goals, and provide appropriate feedback to clear up vagueness, confusion, and misunderstanding, besides adapting its reactions to the partner’s specific needs and preferences.

2. Ideal Cooperation and Rational Agency

Following Allwood (1976), we say that the agents are engaged in Ideal Cooperation if they

a) have a joint goal,

b) consider each other cognitively,

c) consider each other ethically,

d) trust each other in (a) - (c).

The first requirement is rather obvious since the agents must cooperate on something, i.e. they must share intention to achieve a certain goal; otherwise it would make no sense to cooperate. Cooperation cannot be imposed on the agent as the whole notion of cooperation would become contradictory.

Cognitive consideration refers to the agent’s deliberation on the fulfilment of the goal. However, rational agents should not only attempt to fulfil their goals in the most reasonable way, but they should not prevent other agents from fulfilling their goals either. An important part of Ideal Cooperation is thus ethical consideration which says that it is not enough to know the efficient ways to achieve the goal, but the partner is to be treated as a rational motivated agent as well. This second, ethical dimension provides a counter-force to the epistemic rationality which deals with the agent’s knowledge of what is the effective and most rewarding way to achieve one’s goals. Ethical consideration obliges the agent not to act so as to prevent other agents from maintaining their rationality, and thus it also accounts for the agents’ seemingly irrational behaviour. The agent may e.g. choose to volunteer for a tedious task to save someone else from doing it (i.e. to increase her pain instead of pleasure), or choose a method which is inefficient but allows easier cooperation with the others (i.e. not act in a cognitively competent way). Rationality of such “irrational” actions is difficult to account for by utilitarian and selfish epistemic considerations alone; ethical consideration, however, also requires that the agent should not harm or cause pain to
Rationality can thus be seen as an effect that emerges if the agent's behaviour seems to conform to the principles of Ideal Cooperation. Moreover, acts can be rational to different degrees depending on how well the agent thinks they fit in the goal fulfilment (appropriateness) and ethics (adequacy). However, rationality describes the act's potential to advance the goal, not the actual advancement. The agent cannot know all the factors that influence the intended effects at the time of acting, and the actual effects of an act need not be exactly as intended (the answer may be evasive, an attempt to rent a car unsuccessful), although the act itself is considered rational. On the other hand, achieving the intended effects does not render an act rational either: e.g. ordering a taxi without the intention to go anywhere is considered irrational (or at least a bad joke) even though the effect of the request is achieved when the taxi arrives. Rationality is thus tied to the act’s assumed function in a larger context: rational acts are instruments in achieving some goal.

According to Allwood (1976), Ideal Cooperation also creates social normative obligations, which deal with issues such as the person's availability and contactability for information coordination, and their willingness to report on the evaluation of the information. Obligations are based on the human ability for rational coordinated interaction, and they can be understood as a level of communication through which the agents exhibit their expectations of responsiveness (responsiveness itself being a consequence of human ability for rational coordinated interaction). Obligations encode the speakers' sincerity and motivated, considerate behaviour, and they are also connected to a particular activity and the agent's role in the activity. Discussion on the principles of Ideal Cooperation with dialogue examples and references can be found e.g. in Allwood, Traum and Jokinen (2000).

3. Constructive Dialogue Model Agents

As argued above, Ideal Cooperation, first defined by Allwood (1976), is a keyword in describing the characteristics of human-human communication, and it can be considered fundamental for human-computer interaction, too. However, the system's cooperation and communicative obligations are usually hardcoded in the control structure, and the system's cooperation does not go beyond reactions that are aimed at producing the most straightforward response. Of course, depending on the task, the system can take more initiative in the dialogue and ask questions to narrow down its search space, or give additional information in anticipation of further user questions, or in resolving misunderstandings or other problematic situations, but it cannot reason on different cooperation strategies: the system is designed as a cooperative and helpful partner which provides the user with appropriate information, but does not manifest autonomous rational properties.

Practical service systems, like information-providing or booking systems, can indeed be built without reference to Ideal Cooperation, or other “higher-level” notions of communication, and be designed on the basis of the requirements of the particular task in hand. However, it is obvious that models of rationality and cooperation can improve the quality of practical systems, especially in the context of mobile information technology: the systems can be made more flexible and reactive by implementing these principles in the reasoning process rather than by hardcoding
them in the control structure. Thus cooperation is not just the system’s pre-assigned
disposition to act in a helpful way, but the system can exhibit basic communicative
capability in reasoning about appropriate responses.

My own research has dealt with these topics, and focused on the CDM approach,
Constructive Dialogue Modelling (Jokinen 2007). It has its inspiration in Allwood’s
Ideal Cooperation and rational communicative agents, and applies these notions into
dialogue agents and designing interactive systems. I developed the first version of
CDM within the EU-project PLUS (Black et al. 1991), which investigated pragmatic
aspects of human-computer dialogues, and have elaborated CDM in various recent
dialogue projects like Interact (Jokinen et al. 2002) and PUMS (Hurtig and Jokinen
2005).

According to Jokinen (2007), the main features of CDM agents are:
- The agents are rational and cooperative
- The agents exchange new information
- Mutual knowledge is constructed through interaction
- Utterances are locally planned and realized
- The agents use general conversational principles (Ideal Cooperation)

Traditionally, agents are described as capable of performing purposeful, goal-directed
actions, and their rationality appears in their ability to plan actions to reach their goals.
In communicative situations, this leads to goal-directed dialogue actions whereby the
agent expresses her intentions, while the partner tries to discover the goal and produce
a cooperative response which could help the agent to achieve her goal. Corresponding
dialogue systems exhibit system-initiative properties: the dialogue is initiated by the
user, but the dialogue steps are controlled by the system in a predefined order.
However, these goal-oriented agents tend to over-emphasise plan-based behaviour: if
the context changes, they have no particular means to evaluate rationality of their goal,
and if the adopted plan becomes inappropriate later in the dialogue, the agents still try
to achieve the goal “fanatically” (Cohen et al. 1990). In dialogue management, several
examples illustrate these unfortunate situations, from the system’s rigid applicability
to only one kind of task, to the system’s frustrating “please repeat” responses.

An alternative view sees the agent’s behaviour as fundamentally reactive: the agent
evaluates contextual changes and adapts her behaviour accordingly. In dialogue
management, reactivity usually refers to the user-initiated management style whereby
the system attempts to answer the user requests and adapt its responses to the user
needs. This approach, however, leads to under-emphasis of plan-based behaviour: the
agent may be unable to fulfil her own goals because of the constant change of
situations that require new reactions. Moreover, even if the agent can take the
initiative to pursue her own goals, the decision whether to react to the partner’s input
or follow one’s own plan needs to be made. These situations can lead to conflicts
which can only be resolved if the agent is capable of reasoning about her goals, and of
evaluating their rationality in the context.

The CDM approach advocates a middle view in which the agent is regarded as
constructive: the agent responds to the changes in the context in which the interaction
takes place, but is also capable of planning and taking initiatives to fulfil her goals.
The agent’s actions in interactive situations are usually on-line reactions to contextual
changes, but the evaluation of the changes takes place in the context of the agent’s own plans and goals (and one of the goals is the shared goal which is presupposed by Ideal Cooperation).

In CDM, dialogue management thus has a uniform basis in the communicative principles of Ideal Cooperation. While the agent’s goals are usually independently set to accomplish some real world tasks (e.g. rent a car, assemble a pump, book a flight, or just keep social company), the distribution of the knowledge and skills is such that the task goals cannot usually be successfully achieved by the agent alone: there is a need to coordinate the task and communicate about it with other agents. Consequently, through communication the agent constructs a model of how to achieve her goals and simultaneously show consideration to the partner’s goals.

The construction of a shared model takes place in two communication-related tasks: evaluating the partner’s goal, and planning an appropriate response to it. The former concerns how the partner’s goal can be accommodated in a given context, and results in strategic decisions about appropriate next goals. The latter concerns how the goal can be realized in different contexts. From the agent’s point of view, communication thus consists of an analysis of the partner's contribution, its evaluation in regard to the agent's own knowledge and intentions, and reporting the evaluation results back to the partner as the agent’s reaction. The process conforms to what Clark and Schaefer (1989) have suggested as the presentation-acceptance –cycle of communication: the speaker presents information which the partner accepts (rejects), and the partner’s acceptance (rejection) functions as a new presentation which the speaker on her turn has to react, etc. Presentation and acceptance are performed by different agents, however, whereas the analysis-evaluation-reporting –cycle describes the processing from the view-point of a single agent only. In fact, the latter can be seen as a more detailed description of Clark and Schaefer’s acceptance phase or, if we consider the reasoning process that results in the presentation of information, of their presentation phase. As for CDM, new information is exchanged in the communication cycle, and it is defined as the information that provides changes in the context. New information is the same whether looked at from the speaker’s or from the hearer’s point of view, and thus both are obliged to evaluate the exchanged new information with respect to their own knowledge in order to manage the contextual changes. The speaker evaluates new information regarding how well it carries the content she thinks contributes to the construction of the shared model, while the partner evaluates appropriateness of the new information in fulfilling what she believes is the shared model.

While the evaluation of new information is motivated by the agent’s epistemic needs to cope with the changed context, the reporting of the result of this evaluation is not. Rather, ethical consideration steps in here. As the agent is engaged in a coordinated activity with the partner, she needs to inform the partner of what is her view of the new situation, so as to allow the partner to work towards their shared goal with all the relevant information. Thus the agent’s decision of reporting the evaluation results to the partner is based on her complying with ethical consideration, although the content of the reporting falls out as a result of both cognitive and ethical considerations.

4. Learning via Interaction
A frequent problem in interactive systems is the system's knowledge of the user being restricted to a model of a standard user; this does not take into account the user's
versatile competence and accommodation to different situations. In practical dialogue systems, however, it is important to consider complex and varied interactions, too: the users’ knowledge and intentions, variation in their viewpoints and interests, cultural background, the context in which interaction takes place, etc. affect the quality of interaction and the effectiveness and usability of the system as a whole. As argued in Jokinen (2000), one of the main issues in the new information technology context is the question of knowledge acquisition. This means that the need for learning dialogue systems becomes relevant as well: the systems should learn through interaction and update their knowledge bases with novel facts.

However, a realistic on-line learning requires that the system knowledge is not modelled as fixed categories but as meaningful clusters that can change according to the system’s interactions with the users in communicative situations. The point of departure is to learn interaction patterns by experience: the model is gradually built and modified on the basis of the observed behaviour. By comparing this model with a general model, the user’s individual strategies can be distinguished once the differences between the two models become statistically significant. It is also possible to form user models by grouping together those users who have similar interaction patterns.

Previous work has explored possibilities for building speech interfaces using reinforcement learning methods (see e.g. Litman et al. (2000) and references therein). Below I will briefly sketch a model that addresses some of the basic questions concerning learning and interaction in terms of Exploratory Learning Problem (Scott and Markovitch, 1989). The basic situation consists of a set of perceptual operations, and the task is to build an (egocentric) representation of the domain to predict actions in different situations. There are different constraints such as large domain, non-deterministic changes, limited resources of time and memory, and the practical conditionals concern knowledge of applying an operation to entities, represented via operators such as:

\[
(\text{Op} (\text{Outcome}_1 \, P_1)(\text{Outcome}_2 \, P_2)\ldots(\text{Outcome}_N \, P_N))
\]

The goal is to search for a good representation using an inheritance network of classes of domain entities, and to search for informative examples of good representations among the agent’s experiences by minimizing Shannon uncertainty of the outcome in practical conditionals.

The original model has no other agent present (thus no communication with agents), and so, in order to apply the model to communicative situations, we need to change the basic assumptions so that they concern cooperative communication between rational agents. The basic requirements for linguistic communication deal with the following (cf. Allwood 1976): the agent is in contact with the partner, perceives the partner’s intention to communicate, understands the partner’s intention to communicate, and finally reacts to the partner’s intention to communicate. The communicative goal is to get a message across by maximising the agent’s and the partner’s mutual comprehensibility and by minimising any ambiguity in language expressions.

There are also certain constraints in the process: the language is possessed by a group of agents, and the agents also have limited resources. This means that the agents must cooperate, and they must adapt and learn new information. The system’s knowledge
can be expressed by “language conditionals” which are operators (dialogue acts) consisting of facts, i.e. pieces of information (P1, P2,…Pn), in a certain context (C1, C2,…Cn) as follows:

$$ (\text{DACT} ((C_1 \ P_1)(C_2 \ P_2)\ldots(C_n \ P_n)) $$

The factual and contextual information can be represented in terms of propositions which consist of entities in the world. The agent’s communicative goal can now be reduced to the goal to search for a good representation for language conditionals. Hence, learning is equal of searching for informative examples from the agent’s experience database by minimising uncertainty of the context in language conditionals, i.e. by minimizing the ambiguity and maximising mutual comprehensibility.

5. Conclusions

In this paper I have presented some challenges for interactive systems in the context of information technology advancements, and presented the Constructive Dialogue Model approach which aims at addressing these challenges from the point of view of rational and cooperative dialogue agents. The CDM approach is inspired by Allwood’s notion of Ideal Cooperation, and applies the communicative principles to interactive systems, thus aiming at flexibility and naturalness in dialogue management.

References


http://www.jrc.es/pages/iptsreport/vol49/english/MET1E496.htm