AI2Bh Communicating Knowledge

Michael Fourman
Revised by Alan Bundy

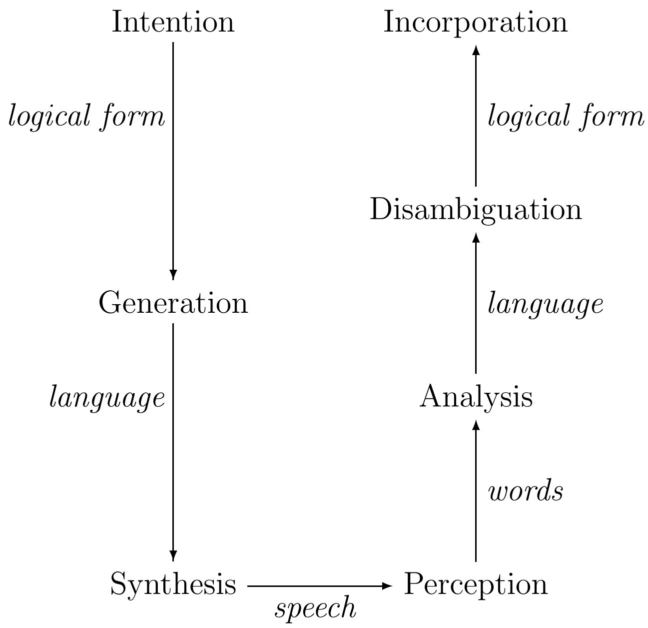
May 12, 2004

In which we represent communicative acts as actions that agents can plan to take, using modal logic to express preconditions and effects.

Our aim is not to model human communication, but to provide an architecture that will enable agents to communicate in pursuit of their goals, and to generate plans that integrate this communication with other actions. Thus, we seek simple models of communicative acts, adequate to our purpose, rather than attempting to represent faithfully the speech acts by which these models are inspired.

Communication in natural languages has many components. All accounts are based on a number of processes that use a variety of different representations of the information to be communicated.

For example, the account in Russell and Norvig has seven stages that can be pictured roughly as follows:



In this model, information is represented differently as it passes from one stage to the next. Each stage processes the information to change its representation. The most abstract representation is called *logical form*.

Our model will short-circuit this complex chain of processes by passing directly from **Intention** to **Incorporation**. We also assume that the process of **Interpretation** is trivial — the hearer accepts what it is told, without question. So the remaining problem is to account for Intention as a process that generates communicative actions in logical form.

We represent communicative acts as actions that an agent can choose or plan to do —- just as it can plan to do concrete actions like moving or grabbing.

Actions are specified by giving their parameters, preconditions, and effects. Concrete actions change the state of the world. Communicative actions change the state of knowledge of the agents participating in the communication.

We introduce communicative actions derived from three types of speech act: **Inform**, **Query**, and **Request**. In specifying these actions, we distinguish the rôles of two agents: the speaker, S, and the hearer, H.

Inform An **Inform** action has a proposition, φ as parameter — we inform someone of some fact. In our simple model, we assume that that the information given is known to the speaker, and accepted, unconditionally, by the hearer. So a precondition of the action is that $\boxed{\mathsf{K}_S}\,\varphi$, and an effect is that $\boxed{\mathsf{K}_H}\,\varphi$. Moreover, as a result of the communication the speaker knows that the hearer knows that φ, and the hearer knows that the speaker knows that the hearer knows, and so on .... We capture this by saying that the effect is to make φ common knowledge between S and H: $\boxed{\mathsf{C}_{\{S,H\}}}\,\varphi$.

$$\begin{aligned} \texttt{action:}\,& \texttt{Inform}_{S,H}(\varphi) \\ \texttt{precondition:}\,& \boxed{\mathsf{K}_S}\,\varphi \\ \texttt{add:}\,& \boxed{\mathsf{C}_{\{S,H\}}}\,\varphi \\ \texttt{delete:}\,& \end{aligned}$$

Given two agents, A, and B engaging in conversation, or dialogue, the rôles of speaker and hearer will alternate between A and B.

Query A Query asks for information:

> *Is there a smell at* $\langle 1, 1\rangle$*?*
> *Where are you now?*

In natural language, queries are often linked to answers, in the sense that the query is part of the context necessary for interpreting the answer. To the first query, an answer, *Yes*, or *No*, would suffice; an answer to the second could be simply a pair of coordinates. Modelling this aspect of human communication makes interpretation non-trivial, and would require us to complicate our architecture. We will use simple Inform actions to convey the information required in the answer. The answers to our queries might be

$\texttt{Inform}_{H,S}(\text{Smelly}(\langle 1, 1\rangle))$
$\texttt{Inform}_{H,S}(\text{At}_H(\langle 1, 1\rangle))$

We formalise the query so that its effect is to modify the hearer's goals. The aim of the speaker is to motivate the hearer to provide the required answer.

Setting goals We could model the effect of a query as directly affecting the hearer's goals — but this *ad hoc* solution would require special-purpose hooks for changing an agent's goals. Instead, we use a trick to achieve the effect of modifying the hearer's goals by providing appropriate knowledge.

We assume that each agent may have a number of goals, which are perhaps prioritised, so that in any situation, an agent chooses an action that will lead to the achievement of one of its goals. We assume that each agent has, among others, the goal of making other agents happy.

An agent A can can then effectively set a goal G for agent B by ensuring that B knows how to make A happy.

$$\boxed{\mathsf{K}_B}\Big(G \rightarrow \text{happy}(A)\Big)$$

For our simple representation of a query such as, *Is it Smelly at* $\langle 1,1\rangle$ *?*, we presume the precondition that the speaker knows that the hearer knows whether or not the Wumpus is at $\langle 1,1\rangle$,

$$\boxed{\mathsf{K}_S}\ \Big(\boxed{\mathsf{K}_H}\,\text{Smelly}(\langle 1,1\rangle) \vee \boxed{\mathsf{K}_H}\,\neg\,\text{Smelly}(\langle 1,1\rangle)\Big)$$

and that the effect is that the hearer knows that the speaker would be happy to be in such a state.

$$\boxed{\mathsf{K}_H}\ \Big(\Big(\boxed{\mathsf{K}_S}\,\text{Smelly}(\langle 1,1\rangle) \vee \boxed{\mathsf{K}_S}\,\neg\,\text{Smelly}(\langle 1,1\rangle)\Big) \rightarrow \text{happy}(S)\Big)$$

$$\begin{aligned}
\texttt{action:}&\ \texttt{Query}_{S,H}(\varphi)\\
\texttt{precondition:}&\boxed{\mathsf{K}_S}\ \Big(\boxed{\mathsf{K}_H}\,\varphi \vee \boxed{\mathsf{K}_H}\,\neg\,\varphi\Big)\\
\texttt{effect:}&\boxed{\mathsf{K}_H}\ \Big(\Big(\boxed{\mathsf{K}_S}\,\varphi \vee \boxed{\mathsf{K}_S}\,\neg\,\varphi\Big) \rightarrow \text{happy}(S)\Big)
\end{aligned}$$

For this simple query, there are two possible answers (given that we assume the hearer knows the answer). For the query, *Where are you now?* there are many possible answers, but the basic idea is the same.

The speaker knows that the hearer knows the answer. That is the speaker knows that for some (x, y), the hearer knows it is at (x, y).

$$\boxed{\mathsf{K}_S}\,\exists(x,y)\,\boxed{\mathsf{K}_H}\,\text{At}_H xo(x,y)$$

Note that although the speaker knows the hearer is somewhere,

$$\boxed{\mathsf{K}_S}\,\exists(x,y)\ \text{At}_H(x,y)$$

the speaker does not, yet, know where. To know this would make the speaker happy:

$$\exists(x,y)\,\boxed{\mathsf{K}_S}\,\text{At}_H(x,y) \rightarrow \text{happy}(S)$$

So we have another form of query

$$\begin{aligned}
\texttt{action:}&\ \texttt{Query}_{S,H}\big(?x.\varphi(x)\big)\\
\texttt{precondition:}&\boxed{\mathsf{K}_S}\,\exists x\,\boxed{\mathsf{K}_H}\,\varphi(x)\\
\texttt{effect:}&\boxed{\mathsf{K}_H}\,\Big(\exists x\,\boxed{\mathsf{K}_S}\,\varphi(x) \rightarrow \text{happy}(S)\Big)
\end{aligned}$$

Request The desired effect of a request is that the hearer should perform some action. Rather than allow requests to directly alter the hearer's plans, we again achieve this effect indirectly, by (indirectly) setting appropriate goals.

So, we will not formalise requests for action, but rather requests for effects — leaving the hearer to plan its own actions. We use the same tricks as before. This leads to a style of request reminiscent of polite society, *Lord X would be most grateful if Lady y would join him for dinner*.

To request that the hearer go to $\langle 1, 1\rangle$, we let him know that this would make the speaker happy, so we want to achieve a state where

$$\boxed{\mathsf{K}_H} \left(\mathrm{At}_H(\langle 1, 1\rangle) \rightarrow \mathrm{happy}(S)\right)$$

By now, the pattern should be clear.

$$\begin{aligned} \texttt{action:} & \texttt{Request}_{S,H}(act) \\ \texttt{precondition:} & \\ \texttt{effect:} & \boxed{\mathsf{K}_H} \left(\exists i.\ T(act, i) \rightarrow \mathrm{happy}(S)\right) \end{aligned}$$

Note the use of the T operation from Event Calculus to turn the action, *act*, into a proposition $T(act, i)$. Note also that we have not been able to provide a precondition for the `Request` operator. To do that we might use a logic of desires and intentions. The modal logic **KD** is often used for this purpose.