Int. J. Human-Computer Studies (2001) **54,** 351–405
doi:10.1006/ijhc.2000.0450
Available online at http://www.idealibrary.com on IDEAL®

An agent-based architecture for multimodal interaction

CATHOLIJN M. JONKER, JAN TREUR AND WOUTER C. A. WIJNGAARDS†

Department of Artificial Intelligence, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands. URL: http://www.cs.vu.nl/ ~ {jonker,treur,wouterw} email: {jonker,treur,wouterw}@cs.vu.nl

(Received 20 October 2000, and accepted in revised form 31 October 2000, published electronically 15 February 2001)

In this paper, an executable generic process model is proposed for combined verbal and non-verbal communication processes and their interaction. The agent-based architecture can be used to create multimodal interaction. The generic process model has been designed, implemented and used to simulate different types of interaction between verbal and non-verbal communication processes: a non-verbal communication process can add and modify content to a verbal communication process, but can also provide a protocol for the (control of the) verbal communication process. With respect to the communication protocol both stimulus–response behaviour and deliberative behaviour have been modelled and simulated. The semantics of the model has been formalized by three-levelled partial temporal models, covering both the material and mental proceses and their relations.

KEYWORDS: intentional models; agent modelling; non-verbal communication; formalization; human–computer interaction; semantic levels; meta-levels.

1. Introduction

Communication is often modelled exclusively as a process of information exchange, abstracting from the physical realization of the communication process. This approach is based on the more general perspective that a strict separation between the *mental* aspect (symbolic processes) and the *material* aspect (physical processes) has advantages in modelling. Following this perspective, the communication process is not embedded in the physical environment. In particular, for the case of embodied non-verbal communication, for example on the basis of body language or gestures, a strict separation of communication processes from interaction at the material level can be rather artificial. On the other hand, describing communication processes only in physical terms may be too restrictive as well. In the model of communication presented in this paper both the material level and the level of information, as well as the interaction between the levels are taken into account in one generic process model.

At the symbolic level a distinction can be made between information representing the subject matter of the communication, and information that refers to the *process* of

†Parts of the material of this paper were presented (in preliminary form) in WECC'98 and ACL'99.

1071-5819/01/030351 + 55 $35.00/0

communication, for example, information on the communication protocol that is followed, or the current stage of the communication process within the protocol. Thus, in the model three semantic or representation levels are introduced.

- The material level (for the physical world).
- The symbolic object level (for symbolic manipulation of content information on the physical world)
- The symbolic meta-level (for symbolic manipulation of the dynamics and reflective aspects of the agents).

Each of these three levels uses representations of structures of one of the other levels: the symbolic object level uses representations of the information from the material world, the symbolic meta-level uses representations of the information from the symbolic object level, and the material level uses representations of the information from the symbolic object level. The three levels of representation and the forms of interaction between verbal and non-verbal communication are detailed in Section 2. It is shown how the generic process model can be applied to model both verbal communication and non-verbal communication, as well as different types of interaction between these forms of communication, in one semantic framework. One example application of the model is presented in which the communication protocol of a verbal communication process is modelled as reflex-based non-verbal communication (one gesture triggers steps in the verbal communication and other gestures by stimulus–response behaviour). In a second example model, the non-verbal communication in the communication protocol is guided by conscious deliberation. A third model addresses goal-directed conscious deliberation.

Next, in Section 3 an executable generic process model is introduced in which a number of processes within a communication process are distinguished. This process model is in accordance with the semantics given in Section 2.

The generic process model proposed in this paper can be used to create multimodal interactive systems, such that apart from the possibility of using a verbal language to interact with the system, also non-verbal interaction is possible. By using the components and knowledge types of the architecture, it becomes possible for the system to handle the user integrating non-verbal cues in his or her communication and to return integrated non-verbal cues to the user as well. This makes for much more intuitive interaction, bringing the machine closer to the user by allowing a broader scope of interaction. The agent-based architecture proposed here can thus be used to easily create a system capable of verbal and non-verbal interaction.

An interactive prototype has been built of the example communication process in this paper. It uses icons to display the agents' situation. The reasoning model used inside the simulated agents can be selected. Also, the user can take on the role of an agent, interacting with the other agent, the interactive demo is presented in Section 4.

The behaviour displayed by the model is described in Section 5. The reflex-based, conscious deliberative and goal-directed conscious deliberative models are described separately. In Section 6, the results are evaluated. The three levels are explained in more detail in Section 7; a process semantics is defined by means of multi-levelled traces based on triples of three-valued states, formalized by partial models (Engelfriet & Treur, 1995). Section 8 summarizes and compares this approach with other work. In Appendix A, a trace of the process is given, demonstrating the stimulus–response case. In Appendix B, the trace using conscious reasoning is given and in Appendix C the trace using goal-directed conscious reasoning is given.

2. Distinguishing and relating verbal and non-verbal aspects

To create an appropriate model we first need to understand verbal and non-verbal communication processes and their interactions. In order to do this, in Section 2.1 we first put both types of communication on a common ground by giving a semantic model.

During real-life communication processes, several types of communication play a role, among them are verbal and non-verbal communication. Furthermore, combining verbal and non-verbal communication can sometimes be done in a stimulus–response manner in contrast to consciously. In Section 2.2, different types of interactions between verbal and non-verbal communication are distinguished and explained. Also, an example of communication process is introduced, which will be reused later on. Section 2.3 proceeds by relating verbal and non-verbal processes, showing how differences between reflex-based (also called direct stimulus–response) reactions and conscious reactions on non-verbal communication can be modelled.

2.1. SEMANTIC LEVELS IN COMMUNICATION

The semantic model is built by identifying the semantic levels involved, making the representation relations between them explicit, and by determining the state transitions at the different levels during the communication process. Three semantic levels are used in the semantic model of communication presented in this paper. The first is the *material level*: of the world itself, the physical reality. The second level is the level of symbolic representation of the world state: the *symbolic object level*. Within the agents, symbols are used to represent information regarding the world state. The third semantic level is the symbolic level where the agent reasons about aspects of its own state, e.g. its own knowledge and actions it intends to perform in the world: the *symbolic meta-level*. Symbolic expressions at this level do not represent information about world states, but instead are an explicit representation of information about the state of the process, of an agent's related mental aspects, such as its state of knowledge, its goals and its intentions.

2.2. A SIMPLE EXAMPLE COMMUNICATION PROCESS

All examples in this paper are about a conference lecture, which is finishing. The chair person, agent A, puts up a slide expressing where to find tea and coffee. A thirsty person in the audience, agent B, interprets this information. However, the communication may be affected by an event in the material world, for example, somebody erroneously standing between the projector and the screen. A trace is shown in Table 1, the expressions used at the symbolic meta-level are explained in Table 2.

In the example, agent A has a representation of the world information that pot 2 contains tea, represented at the symbolic object level by the symbolic expression `contains(pot2, tea)`. By upward reflection to the symbolic meta-level it establishes that it has the positive belief that pot 2 contains tea. The agent A reasons at the symbolic meta-level and concludes that this world information should be communicated to agent B.

Table 1

Multi-levelled trace of an example communication process

External world material level	Symbolic object level	Symbolic meta-level
Agent A		
pot 2 contains tea	contains(pot2, tea)	belief(contains(pot2, tea),pos)
		has_verbal_material_rep(contains(pot2, tea),pos) at_position(pattern3,p0),pos)
		has_effect(put_slide3, at_position(pattern3,p0),pos)
		to_be_observed(I:INFO_ELEMENT)
		to_be_communicated(contains(pot2, tea),pos)
		to_be_achieved(at_position(pattern3,p0),pos)
		to_be_performed(put_slide3)
slide 3 at projector		
pattern3 at p0		
Agent B		
pot2 contains tea		to_be_observed(I:INFO_ELEMENT)
slide 3 at projector		has_verbal_material_rep(contains(pot2, tea),pos, at_position(pattern3,p0),pos)
pattern3 at p0		
		observation_result(at_position(pattern3,p0),pos)
	at_position(pattern3,p0)	
		belief(at_position(pattern3,p0),pos)
		has_been_communicated(contains(pot2, tea),pos)
	contains(pot2, tea)	
		belief(contains(pot2, tea),pos)

TABLE 2
Explanations for the expressions at the symbolic meta-level

Symbol	Explanation
`belief`	The beliefs of the agent. Refers to world information and a sign. The sign (`pos` or `neg`) indicates if the information is `true` of `false`.
`has_effect`	Denotes that an action is capable of bringing about some state in the world, given as the statement that becomes true or false in the world.
`to_be_observed`	The information that the agent focuses on and observes in the world.
`to_be_communicated`	The information that is determined to need communication to the other agent. Includes sign. It has been kept simple in this model, but could be extended to include a recipient agent.
`to_be_achieved`	A goal to be achieved. The goal is stated as a state of the world that must be brought about. The sign indicates if the statement should hold or not.
`to_be_performed`	Refers to the action(s) to be performed by the actuators of the agent.
`has_verbal_material_rep`	The first information element has the second information element as a verbal material representation. Thus, the first is an interpretation of the second.
`observation_result`	The information received by the sensors of the agent. Includes sign of the information.
`has_been_communicated`	The information that has been received by communication. Includes sign of the information.

Using knowledge of the meaning that can be associated to certain material configurations, it discovers that if at position p0 in the material world pattern 3 is visible then this world situation represents that pot 2 contains tea (e.g. a written text on a visible place). Moreover, still reasoning at the symbolic meta-level, it finds out that an action "put slide 3" exists, which has as an effect that pattern 3 is at position p0. Therefore, it concludes at the symbolic meta-level that the action "put slide 3" has to be performed. The action is performed, and the intended effect is realized in the world state at the material level. In the example, the two agents are assumed to have a common ontology on the world including the names of all objects in the world, like pot 2, pattern 3 and the names of the positions.

Agent B performs the observation that pattern 3 is at position p0 (which provides information at the symbolic meta-level, namely the meta-fact that this has been observed), and represents the information acquired at the symbolic object level by `at_position(pattern3, p0)` (the agent B's world model). Note that agent B cannot observe directly the world information that pot 2 contains tea or that slide 3 is on the projector, but it can observe that pattern 3 is at position p0. Knowing at the symbolic meta-level that to this world situation the interpretation "pot 2 contains tea" can be associated, it now concludes at the symbolic meta-level that it has been communicated that pot 2 contains tea. This information is then stored by B at the symbolic object level in its representation of the world state. Note that after this process, the representation of the world state at the symbolic object level includes information that was acquired by observation (pattern 3 is at position p0), and also information that was not obtainable by observation, but acquired by the communication process (pot 2 contains tea).

This example communication process can be described by tracing the states and state transitions at the different levels; see Table 1. In this table, each cell describes a state, and state transitions are indicated by a line separating the two states in the transition. Time goes from top to bottom. In the table, only the relevant new information elements are represented. The first part of the table gives the state of the external world (first column), and the states of the symbolic object level and meta-level of agent A (second and third column). The second part of the table gives the same for agent B. The first part of the table "happens" before the second part of the table.

2.3. TYPES OF INTERACTION BETWEEN VERBAL AND NON-VERBAL COMMUNICATION

The example communication process in Section 2.2 shows only verbal communication. The verbal type of communication is the type that is normally exclusively considered in research in multi-agent systems in which all agents are software agents. However, in real-life situations, often the non-verbal communication is as important as verbal communication. In this section, non-verbal communication is addressed and classified into three different types of interaction with verbal communication. In the classification one of the distinguishing criteria is the nature of the processing of the communication: based on reflex reactions, or on conscious reactions. For each of the types of interaction between verbal and non-verbal communication an example is presented.

The following kinds of interaction between non-verbal and verbal communication are distinguished.

A. Interaction of non-verbal communication with the *content* of verbal communication.
   1. Non-verbal communication provides information additional to the content information transferred by the verbal communication.
      - The subject of the non-verbal communication has no connection with the subject of the verbal communication.
      - The subject of the non-verbal communication is related to the subject of the verbal communication.
   2. Non-verbal communication affects the interpretation of the contents of the verbal communication; modified interpretation.

B. Interaction of non-verbal communication with the *process* of verbal communication.
   1. The verbal communication process is affected by reflex-based reactions to the non-verbal communication.
   2. The verbal communication process is affected by conscious reactions to the non-verbal communication.

Notice that non-verbal communication of type A will lead to conscious reactions of the recipient; as the interpretation of observations as being communicated information is a conscious process. Combinations of the different types of interaction can occur during one communication process. In the examples it is assumed that the agents share a common ontology for world information. Simple examples of the different types of non-verbal communication are.

A. *Interaction of non-verbal communication with the content of verbal communication.*

1. *Additional information*
   (a) *No connection.* Agent A communicates verbally to B that tea can be found in pot 2. Agent B observes that agent A smiles to him and concludes that agent A recognizes him. The observation does not influence the communication process concerned with where the tea can be found. Furthermore, agent B does not change his interpretation of the verbal communication on account of noticing that agent A recognizes him.
   (b) *Related.* Agent A communicates verbally to B that tea can be found in pot 2. During the communication agent A points to position p3. Agent B observes the direction that agent A is pointing in and concludes that agent A is telling it that tea can be found in pot 2 that can be found at position p3.

2. *Modified interpretation.* Agent A communicates verbally to B that **fresh** tea can be found in pot 2. However, agent A makes a face that indicates she is disgusted at the moment the verbal communication process takes place. Agent B combines this non-verbal communication with the verbal one and, therefore, interprets the communication as follows: tea can be found in pot 2, but it is definitely not fresh. Modification of the interpretation of the verbal communication appeared to be necessary, based on the non-verbal part of the communication.

B. *Interaction of non-verbal communication with the process of verbal communication.* Agent A initiates the communication process by walking to position p1. She notices that agent B is looking at her and she starts her communication to agent B that tea can be found in pot 2, by putting the correct slide (slide 3) on the projector. However, after performing this action agent A observes that agent B is looking in another direction; in reaction (either by reflex or consciously) she breaks off the communication process by removing the slide, and (to get attention) starts tapping the microphone. Agent B observes the noise of the microphone and (either by reflex or consciously) reacts by looking at agent A with interest. Agent A waits until she observes that agent B is looking interested, and then resumes the verbal communication process by putting the slide on the projector again. In such a case, the information transferred by verbal communication is not affected by the non-verbal communication, but (the control of) the process of communication is.

In Vongkasem and Chaib-draa (2000) this would correspond to the first level of joint action, the *behaviour* level, where the sender gets the receiver to attend to the message. As well as level 2 of joint action, the *signal* level. In this level, the sender verifies that the communication is well received by the receiver.

An example communication process in which a combination of types of interaction occurs is the following.

Example 1. (*The complete tea story*)

1. Agent A wants to communicate to agent B that non-fresh tea can be found in pot 2 that is located at position p3.

2. Agent A figures out how to perform the communication. She does not have a (verbal) slide that reflects all that she wants to communicate; she only has a slide that says that fresh tea can be found in pot 2. The rest of the communication will, therefore, have to be non-verbal. She finds the following solution: she will put the slide on the projector, point at position p3 and pull a disgusted face at the same time.
3. Agent A attracts the attention of her audience (agent B) by going to position p1.
4. Agent B observes this movement and responds by looking interested.
5. Agent A observes that agent B is looking interested and performs the prepared actions.
6. However, in the mean time, agent B's attention is distracted by a noise from outside. As a reaction to the noise from outside agent B looks away from agent A and stops looking interested.
7. Agent A notices that agent B no longer is looking in her direction. Therefore, she removes the slide, stops pointing, and reverts her face to a neutral expression. Furthermore, in order to attract agent B's attention again, she taps the microphone.
8. Agent B observes the noise inside the room and towards the front of the room (i.e. in the direction of agent A) and looks interested.
9. Agent A waits until agent B looks interested again, and then communicates verbally to agent B that fresh tea can be found in pot 2 (she puts slide 3 on the projector). Agent A makes a face that indicates she is disgusted at the moment the verbal communication takes place. At the same time agent A points to position p3.
10. Agent B observes:
    (a) The pattern on the projection screen that is caused by the slide.
    (b) That agent A is pointing towards position p3.
    (c) That agent A has a disgusted face.
11. Agent B concludes:
    (a) Tea can be found in pot 2: the interpretation of this part of the verbal communication of the pattern caused by the slide is not affected by any of the non-verbal communication of agent A.
    (b) The tea can be found at position p3: this additional information comes from interpreting the pointing gesture of agent A.
    (c) The tea is definitely not fresh: this interpretation is based on modification of the contents of the verbal communication (fresh tea) because of the non-verbal communication (disgusted face of agent A) and the knowledge that tea has a taste.
12. At the same time agent A looks questioningly towards agent B.
13. Agent B observes the questioning face of agent A and puts his thumb up.
14. Agent A observes that agent B's thumb is up and walks away from position p1.

3. The generic process model for communication

3.1. THE PROCESS MODEL FOR COMMUNICATION

Within a communication process as described in Section 2 a number of more specific symbolic processes can be distinguished.

- Observation generation.
- Information interpretation.

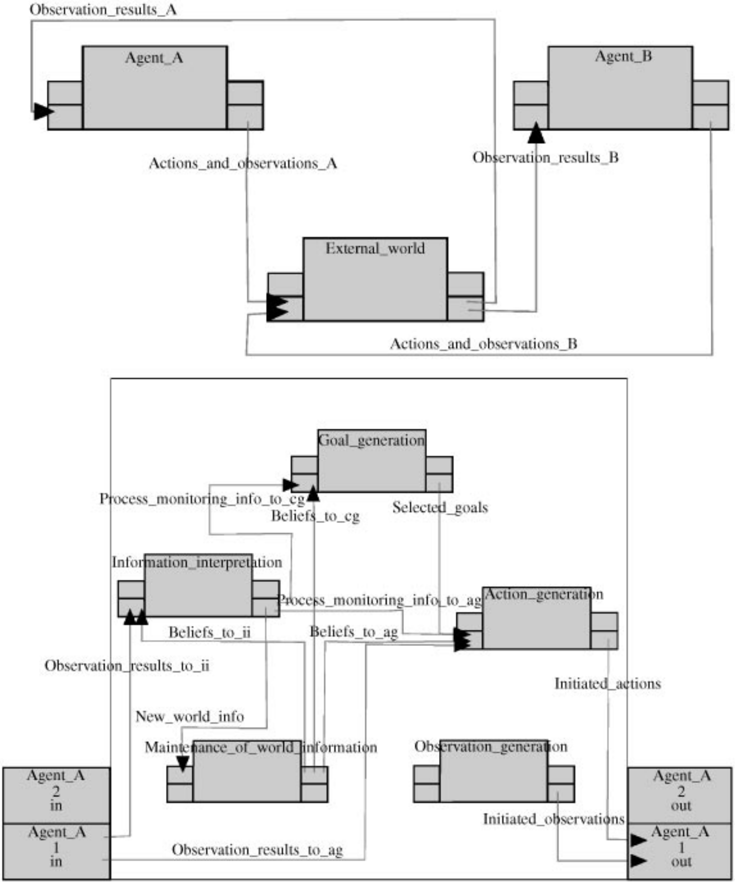


FIGURE 1. The two highest process abstraction levels.

- Goal generation.
- Action generation.
- Maintenance of world information.

Together with the physical processes in the material world that realize action and observation execution, these processes are used as building blocks to compose an executable generic process model of communication. Within the DESIRE approach each of the processes is modelled by a component (the boxes in Figure 1). Each component processes information as soon as it becomes available. The information is processed by applying knowledge rules to the input information, deriving conclusions that are the output information of the component. Each component conceptually runs in parallel, but on a single processor the component activations are serialized. The interaction is modelled by information links (the arrows in Figure 1). By default information links transfer information from the source to the destination when it becomes available. Agent B is identical in composition to agent A.

The component observation_generation contains statements of facts, rules without preconditions. These state what the agent wishes to observe in the world, this is a subset of all possible information elements. The truth value of these elements is of interest to the agent, for example to_be_observed(at_position(agent_A,p1)) means that the agent is interested in knowing if agent A is at p1 or not. Conceptually, this set of to_be_observed facts indicates the focus of attention of the agent, i.e. what it is looking at in the world. And although the agent maybe looking at something this does not mean it will receive the requested observation result. In the example, agent B is initiating observation of the projection screen and agent A observes agent B. And although agent B wants to observe the speakers position, at a later moment, when agent B is looking_away, it no longer receives observation results on the speakers position.

The information_interpretation component interprets facts. It examines the observation results and determines whether these new results should be added to the set of beliefs of the agent. Also it inspects beliefs on the world to check for additional meaning of these world configurations. First, the beliefs that have additional meanings are characterized in as being communications by a particular modality, before the communications in the different modalities are combined to finally conclude the communicated information. This information is then added to the agent's beliefs. Notice that this makes information interpretation a conscious reasoning process. In the example both agents trust their senses, and accordingly store all observed information in their belief set.

The agent's beliefs are the model of the world that the agent currently ascribes to, and it must be stored somewhere. The maintenance_of_world_information component takes care of this. The component takes care of storing the currently believed world model, possibly applying some knowledge rules to augment and extrapolate the world knowledge.

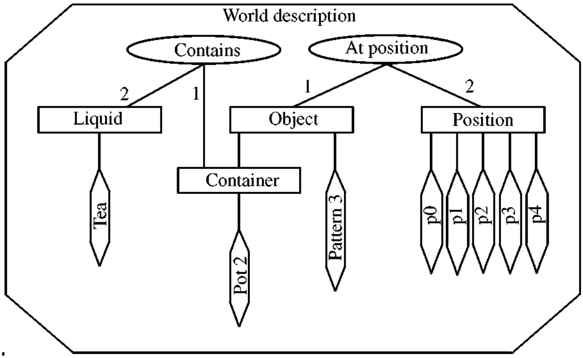
The component goal_generation is a component operating at a high level of abstraction. The component takes as input goals on which information should be communicated, the current status of the communication process and the beliefs of the agent, using the available knowledge to generate intentions for changes to the world. In order to select world configurations that will communicate certain information it selects intentions to_be_achieved by first assigning modalities to the information elements, then selecting intentions to bring about representations of these information elements. When reasoning in a goal-directed manner, the necessary intentions are selected by this component. The information elements to be communicated to the other agent are fixed in the example; some other apparatus should generate the content of the communication.

The agent needs to be capable of selecting actions to perform in the world. The action_generation component does this. Firstly, it takes the to_be_achieved world situations, the intentions that are transferred from goal_generation, and selects appropriate actions for them. Secondly, it takes the observation_results and applies the stimulus–response rules to them, generating the reflex actions. Thirdly, conscious reasoning takes place in the action_generation component. The to_be_performed facts are transferred to the output of the agent to be executed in the physical or simulated, world by the external_world component.

The component external_world takes the observations and the actions initiated by the agents and produces the observation results for the agents. It maintains the state of the world and takes care of the changes to the world state. The physical manifestations of the agents are assumed to be in the external_world component. Thus, a robot's body would reside as an object in the external world, changing state by receiving actions, returning sensor information that is relayed as observation results. A software agent's screen presence and program code is also part of the external world, the state of the software system being conveyed as observation results. Both these examples put the agent body in the external world, and the agent "mind" in the agent_A and agent_B components. For the example communication process the world stores the world situation and has knowledge on the effect of actions, applying it to simulate the effect of the actions. The relevant information elements in the world situation are known completely, i.e in the simulation nothing is unknown. The observation_results facts are derived from this stored world situation, the requested observations and knowledge on the effect of actions on possible observations. The knowledge used by the external_world component in order to return simulated results is not further detailed in this paper.

3.1.1. Description at the symbolic object level. A world state at the material level can be formally described by a two-valued model (assuming that the world in principle can be described in a deterministic manner) for the following language.

```
Information type
 world_description
 sorts
  OBJECT, CONTAINER,
  LIQUID, POSITION
 subsorts
  CONTAINER: OBJECT
 objects
  pattern3: OBJECT;
  pot2: CONTAINER;
  p0, p1, p2, p3, p4: POSITION;
  tea: LIQUID
 relations
  contains: CONTAINER * LIQUID;
  at_position: OBJECT * POSITION;
end information type
```

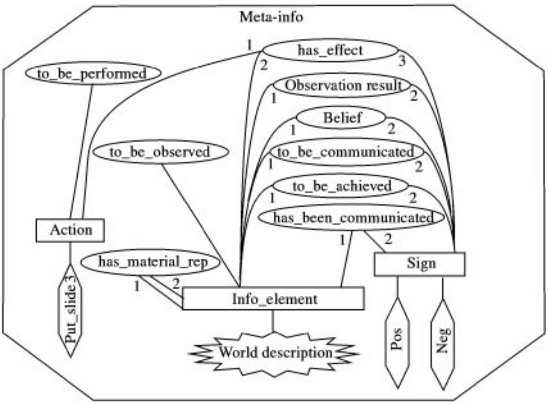


Although the formal definitions of information types is given in Section 7, Definition 1, a short intuitive explanation can suffice here. The information type world_description defines a set of atoms. Examples of the possible atoms are contains (pot2, tea),
at_position(pattern3, p0) and
at_position(pot2, p3). A more detailed explanation of the visual knowledge representation format can be found in Jonker, Kremer, Leeuwen, Pan and Treur (1999).

A state of the symbolic object level can be formally described by a three-valued model for the same language. The third truth value is used to express that some fact is not represented (e.g. not known to the agent).

3.1.2. Description at the symbolic metalevel. The symbolic meta-level can be formally described by three-valued models for the following language.

```
Information type meta-info
 sorts
  INFO_ELEMENT, ACTION, SIGN;
 Subsorts
  INFO_ELEMENT;
 meta-description
  world_description:
    INFO_ELEMENT;
 objects
  pos, neg: SIGN;
  put_slide3: ACTION;
 relations
  to_be_observed: INFO_ELEMENT;
  observation_result:
    INFO_ELEMENT * SIGN;
  belief: INFO_ELEMENT * SIGN;
  to_be_performed: ACTION;
  has effect: ACTION * INFO_ELEMENT * SIGN;
  to_be_communicated: INFO_ELEMENT * SIGN;
  to_be_achieved: INFO_ELEMENT * SIGN;
  has_been_communicated: INFO_ELEMENT * SIGN;
  has_material_rep: INFO_ELEMENT * INFO_ELEMENT;
end information type
```



The meta-description used in the above information-type transforms all ground atoms specified by information type `world_description` into ground terms of sort `INFO ELEMENT`. Examples of atoms in `meta-info` are
`to_be_observed(at_position (pot2,p0))`,
`observation_result(at_position (pot2,p0),neg)`,
`belief (at_position(pot2,p0),neg)` and
`to_be_performed(put_slide3)`. The relations are explained in Table 2. Formalizing information states as partial models makes it possible to also model the reasoning behaviour of common inference mechanisms, such as chaining or unit-resolution, in terms of all ground literal conclusions that have been derived up to a certain moment in time: the third truth value unknown is used for information that has not (yet) been derived in the current state.

The `information_interpretation` component reasons about the meaning of material configurations in the world. But, both `has_verbal_material_representation` and `has_non_verbal_material_representation` seem to relate not meta-level concepts, but materials and meanings. How a material property can be related to an interpretation of this property is shown in Figure 2. The expression `has_material_rep(contains(pot2,tea), at_position(pattern3,p0))` is an element of the symbolic meta-level, and denotes that the interpretation of the presence of `pattern3` is the teapot `pot2` contains tea. This element at the symbolic meta-level refers to two elements at the symbolic object level. Firstly, it refers to the world situation being represented: the `contains(pot2,tea)` element. Secondly, the element `at_position(pattern3,p0)` is referred to as the situation in the world that has an interpretation.

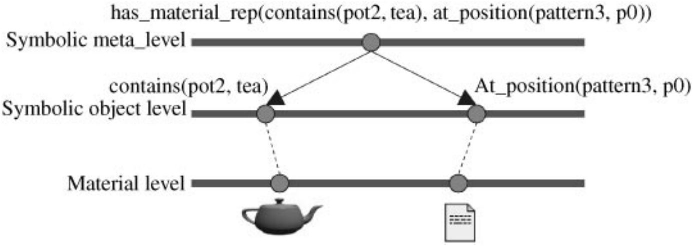

has_material_rep(contains(pot2, tea), at_position(pattern3, p0))
Symbolic meta_level
contains(pot2, tea)
At_position(pattern3, p0)
Symbolic object level
Material level

FIGURE 2. Representation relations used in information interpretation.

These two elements at the symbolic object level, in turn, are symbolic notations for facts true in the state of the world. The `contains(pot2, tea)` element is the symbolic notation for the teapot `pot2` containing tea in the actual world. The `at_position-(pattern3, p0)` element is the symbolic notation for the physical circumstance of `pattern3` being present at location p0. So, the elements of the symbolic object level represent material configurations in the world and the elements at the symbolic meta-level refer to the elements of the symbolic object level.

The reasoning of the agent is specified using a number of generic rules using the information types given above. For example, in the component `information interpretation` the following knowledge is used.

```
if belief(R:INFO_ELEMENT, pos)
  and has_verbal_material_representation(C:INFO_ELEMENT, pos,
    R:INFO_ELEMENT, pos)
then has_been_communicated_by_modality(C: INFO_ELEMENT, pos, verbal)
```

This rule describes a meta-reasoning transition; it determines the content information that is communicated. It realizes the state change depicted in Table 1, from the 4th to the 5th row in the table depicting agent B's process. The other knowledge bases are specified in a similar way. In the model all components and information links in agents A and B process information when the input information changes, possibly in parallel.

In Brazier, Jonker and Treur (2000) a compositional generic agent model GAM is proposed. The model introduced here fits in the generic agent model. The `maintenance_of_world_information` component is directly taken from the GAM. The `goal_generation` component presented here is actually a subcomponent of the `agent_interaction_management` component of the GAM, as the content to be communicated is decided upon by other parts of `agent_interaction_management` or is transmitted from `cooperation_management`. The `world_interaction_management` component from the GAM is split here in three components, actually subcomponents of `world_interaction_management`. The components are `observation_determination`, `action_generation` and `information_interpretation`. These (sub-)components of the GAM are put directly into the agent component in this paper for ease of exposition.

3.2. MODELLING THE EXAMPLE PROTOCOL

The next question is how much interpretation is needed to decide upon some action. The above example allows for three possible models.

1. Agent A observes that agent B is looking away and directly taps the microphone (a form a direct stimulus–response response behaviour, also called behaviour by reflex).
2. Agent A observes that agent B is looking away, interprets this information as a belief of the form "Agent B is not paying attention", and on the basis of this belief decides to tap the microphone.
3. As in 2. Agent A generates the belief that Agent B is not paying attention, and realizes that she needs to attract his attention (as a goal), and decides to tap the microphone (as an action to realize the goal). This is a form of goal-directed behaviour based on conscious interpretations.

The generic process model described in Section 3.1 has been used to design and implement specific models for each of these three cases. An integral part of the communication process is the interpretation of the content information, which is assumed to be a conscious process. The knowledge to interpret information is the same for both agents and is used within the component `information interpretation`. The definitions of the notation used for the information are given in Section 3.1.

component `information interpretation`

(reflex-based and conscious)

Generic knowledge:

```
if belief(R:INFO_ELEMENT,SR:SIGN)
  and has_verbal_material_representation(C:INFO_ELEMENT,SC:SIGN,
    R:INFO_ELEMENT, SR:SIGN)
then has_been_communicated_by_modality(C:INFO_ELEMENT, S:SIGN, verbal);

if belief(R:INFO_ELEMENT, SR:SIGN)
  and has_non_verbal_material_representation(C:INFO_ELEMENT,S:SIGN,
    R:INFO_ELEMENT, SR:SIGN)
then has_been_communicated_by_modality(C:INFO_ELEMENT, S:SIGN, non_
  verbal);

if has_been_communicated_by_modality(C:INFO_ELEMENT, S:SIGN, verbal)
  and not concerns_taste(C:INFO_ELEMENT)
then has_been_communicated(C:INFO_ELEMENT, S:SIGN);

if has been communicated_by_modality(C:INFO_ELEMENT, S:SIGN, non_verbal)
then has_been_communicated(C:INFO_ELEMENT, S:SIGN);
```

Domain-specific knowledge:

```
if has_been_communicated_by_modality(C:INFO_ELEMENT, pos, verbal)
  and concerns_taste(C:INFO_ELEMENT)
  and has_been_communicated_by_modality(tastes_good, neg, non_verbal)
then has_been_communicated(C:INFO_ELEMENT, neg);
```

Component `maintenance of world information` is used by the agent to maintain a (what he hopes to be correct) representation of the current state of the external world. It uses the following knowledge, if the agent behaves consciously.

component `maintenance of world information`

```
(conscious)
```

Domain-specific knowledge:

```
if has_property(A:AGENT, looks_interested)
then has_property(A:AGENT, paying_attention);

if not has_property(A:AGENT, looks_interested)
then not has_property(A:AGENT, paying_attention);
```

Within the component `goal_generation` the agent determines the goals that should be achieved if some information is to be communicated. Because of lack of space only the knowledge of agent A is given (that of agent B is similar).

component `goal generation`

(reflex-based and conscious)

generic part to choose modality and material representation (*similar knowledge is used for non-verbal communication*):

```
if to_be_communicated(C:INFO_ELEMENT, S:SIGN)
  and has_verbal_material_representation(C:INFO_ELEMENT, S:SIGN,
  R:INFO_ELEMENT, SR:SIGN)
then to_be_communicated_by_modality(C:INFO_ELEMENT, S: SIGN, verbal);

if to_be_communicated_by_modality(C:INFO_ELEMENT, S:SIGN, verbal)
  and has_verbal_material_representation(C:INFO_ELEMENT, S:SIGN,
  R:INFO_ELEMENT, SR:SIGN)
then to_be_achieved(R:INFO_ELEMENT, SR:SIGN);
```

Domain-specific part to combine verbal and non-verbal communication:

```
if to_be_communicated(C:INFO_ELEMENT, neg)
  and has_verbal_material_representation(C:INFO_ELEMENT, pos,
  RV: INFO_ELEMENT, SV:SIGN)
  and concerns_taste(C:INFO_ELEMENT)
  and has_non_verbal_material_representation(tastes_good, neg,
    RNV:INFO_ELEMENT, SN:SIGN)
then to_be_communicated_by_modality(C:INFO_ELEMENT, pos, verbal)
  and to_be_communicated_by_modality(tastes_good, neg, non_verbal);
```

In this case, the difference between reflex-based and conscious behaviour can be easily modelled. The reflex-based agent uses the knowledge within component `action_generation` in which all actions to be performed are guarded by conditions on available observation results. This means that as soon as these observation results become available the agent can react according to the goals set by (in this case) component `goal_generation`. A conscious agent does not react directly on observation results, it first interprets this information and decides what information it wants to believe. The conscious agent then reacts on its beliefs and not directly on its observation results. Therefore, the knowledge of the conscious agent with respect to action generation can be modelled by taking the following knowledge base and changing every relation `observation_result` into the relation `belief`. Because of lack of space only the knowledge of agent A is given (that of agent B is similar).

component action generation
(reflex-based)

Generic communication action knowledge:

```
if to_be_achieved(G:INFO_ELEMENT, S:SIGN)
  and possible_action_effect(A:ACTION; G:INFO_ELEMENT, S:SIGN)
  and observation_result(has_property(agent_B, looks_interested), pos)
  and observation_result(at_position(agent_A, p1), pos)
  and observation_result(has_property(agent_A, disgusted_face), neg)
  and observation_result(has_property(agent_A, questioning_face), neg)
then to_be_performed(A:ACTION);
```

Domain-specific knowledge:

```
if observation_result(there_is_noise_inside, pos)
  and observation_result(has_property(agent_B, looks_away), neg)
then to_be_performed(stop_tapping);

if to_be_achieved(G:INFO_ELEMENT, SG:SIGN)
  and observation_result(has_property(agent_B, looks_away), pos)
  and observation_result(at_position(agent_A, p1), pos)
  and observation_result(has_property(agent_A, disgusted_face), pos)
then to_be_performed(tap_microphone)
  and to_be_performed(remove(S:SLIDE))
  and to_be_performed(do_not_point)
  and to_be_performed(pull_no_disgusted_face);

if observation_result(has_property(agent_B, looks_interested), pos)
  and to_be_achieved(G:INFO_ELEMENT, S:SIGN)
  and observation_result(G:INFO_ELEMENT, S:SIGN)
  and observation_result(has_property(agent_A, questioning_face), neg)
then to_be_performed(pull_questioning_face);

if to_be_achieved(G:INFO_ELEMENT, S:SIGN)
  and observation_result(at_position(agent_A, p1), neg)
then to_be_performed(go_to(p1));

if observation_result(has_property(agent_B, thumb_up), pos)
  and observation_result(at_position(agent_A, p1), pos)
then to_be_performed(go_to(p2))
  and to_be_performed(pull_no_questioning_face);
```

Both agents always have knowledge in their symbolic meta-level information states:

Agent A always has the following at its symbolic meta-level.

```
has_verbal_material_representation(contains(pot2, tea), pos,   at_position
  (pattern_3, p0), pos)
has_verbal_material_representation (has_property(tea, fresh), pos, at_posi-
  tion(pattern_3, p0), pos)
concerns_taste(has_property (tea, fresh))
has non_verbal_material_representation(tastes_good, neg, has_property
  (agent_A, disgusted_face), pos)
has_non_verbal_material_representation (at_position (pot2, p3),
  pos, has_property(agent_A, points_to(p3)), pos)
possible_action_effect(go_to(p1), at_position(agent_A, p1), pos)
possible_action_effect(pull_disgusted_face, has_property(agent_A,
  disgusted_face), pos)
possible_action_effect(point_to(p3), has_property(agent_A,
  points_to(p3)), pos)
```

```
possible_action_effect(show(slide_3),at_position(pattern_3,p0),pos)
possible_action_effect(tap_microphone,there_is_noise_inside,pos)
possible_action_effect(pull_questioning_face,
  has_property(agent_A,questioning_face),pos)
has  non_verbal_material_representation(communication_succeeded,pos,
has_property(agent_B,thumb_up),pos)
```

Agent B always has the following knowledge at its symbolic meta-level.

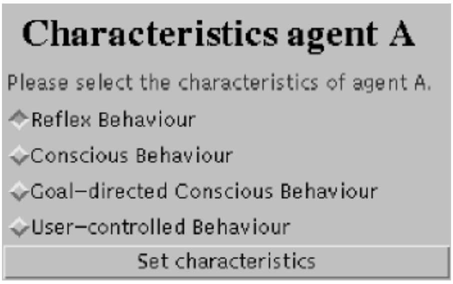
```
possible_action_effect(look_interested, has_property(agent_B,
    looks_interested),pos)
has_verbal_material_representation(has_property(tea,fresh),pos,
  at_position(pattern_3,p0),pos)
has_verbal_material_representation(contains(pot2,tea),pos,at_position
  (pattern_3,p0),pos)
has non_verbal_material_representation(tastes_good,neg,
  has_property(agent_A, disgusted_face),pos,
has non_verbal_material_representation(at_position(pot2,p3),pos,
  has_property(agent_A,points_to(p3)),pos)
not concerns_taste(contains(pot2,tea))
concerns_taste(has_property(tea,fresh))
has non_verbal_material_representation(message_understood,pos,
  has_property(agent_B,thumb_up),pos)
possible_action_effect(put_thumb_up, has_property(agent_B, thumb_up),pos)
```

4. Implementation

An interactive demo has been developed which allows users to hold a conversation with an artificial agent that incorporates both verbal and non-verbal aspects and interaction between these aspects. The user can make each agent behave differently, see Figure 3. The desired type of behaviour can be selected for both agents A and B.

It is also possible to select user-controlled behaviour. The user is then regularly prompted to select zero or more actions using a dialogue window pictured in Figure 4. In this way, the user can play the role of an agent and have a communication with the other agent. Thus, the user communicates with the other agent in the same way that agents communicate between themselves, and the agent sees the user embodied as another agent to interact with.

In this manner of human–computer interaction the user adapts and communicates in a language the agent will understand. Given smart enough agents, this should not be



Characteristics agent A

Please select the characteristics of agent A.

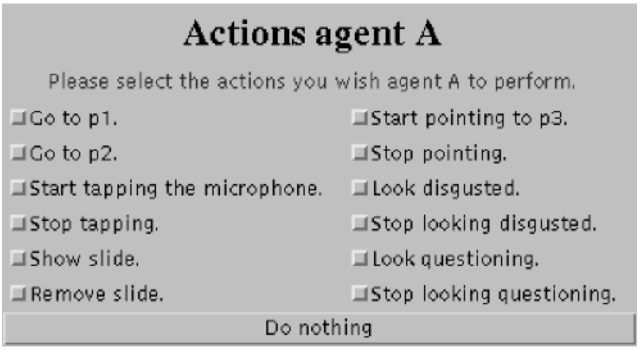
◇ Reflex Behaviour

◇ Conscious Behaviour

◇ Goal-directed Conscious Behaviour

◇ User-controlled Behaviour

FIGURE 3. Sample screen shot from the implementation. A window is shown that allows the user to select the characteristics of an agent, its behaviour will change accordingly.



Actions agent A

Please select the actions you wish agent A to perform.

□ Go to p1.
□ Go to p2.
□ Start tapping the microphone.
□ Stop tapping.
□ Show slide.
□ Remove slide.
□ Start pointing to p3.
□ Stop pointing.
□ Look disgusted.
□ Stop looking disgusted.
□ Look questioning.
□ Stop looking questioning.

FIGURE 4. Sample screen shot from the implementation. This window is shown when the user is playing the role of Agent A and can fill in zero or more actions to do.

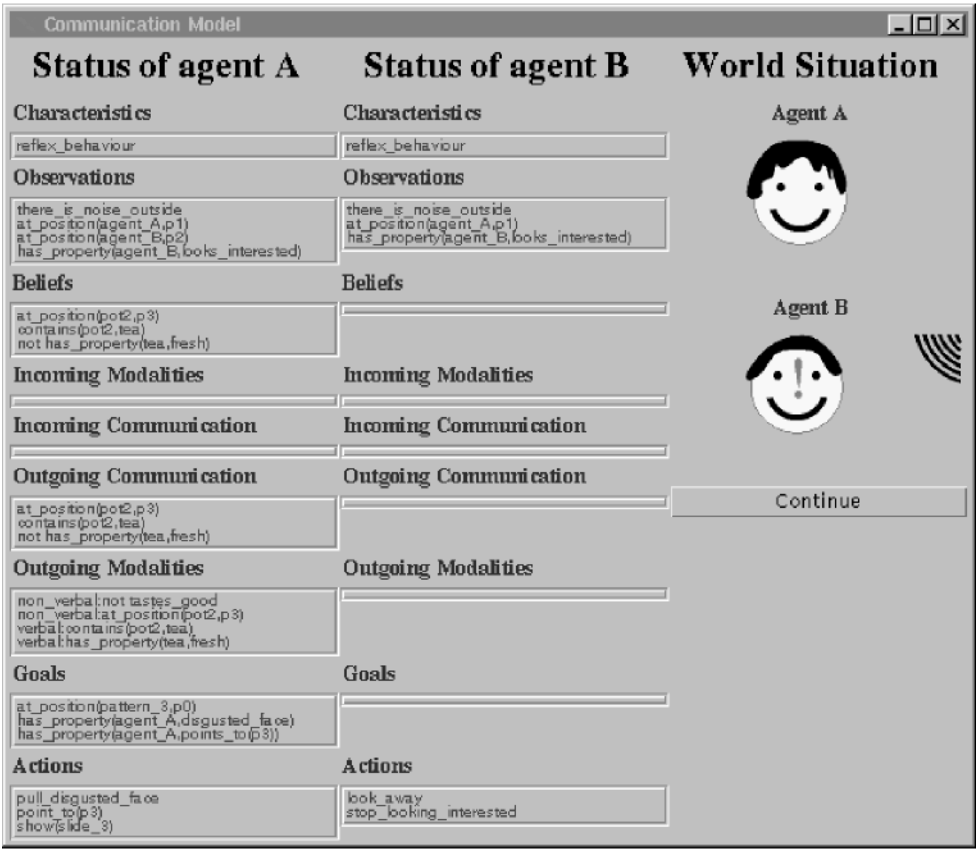
a problem to novice users. The real strength is the representation of the user as an agent in the system. This type of agent is called an *avatar*. But, the user-controlled agent in this model is not just a representation of the user in the system. Indeed, the user can only observe what the agent observes and act according to the capabilities of the agent. The user is playing the role of the agent, taking manual control of the actions the agent should perform.

The current state of affairs is also displayed, the status of both agents as well as the state of the world is given. In Figure 5, a situation is shown where agent A has received the attention of agent B and thus wants to start communicating, but agent B has just decided to look towards an interesting sound. For the agents the selected characteristics, the observed world facts, the set of beliefs, the incoming modalities, the incoming communication, the outgoing communication, the outgoing modalities, the goals and the selected actions to be performed are shown. In the communication modalities the communicated facts are listed by modality. The interpretation of these communicated facts is listed in the incoming communication and outgoing communication boxes.

The world situation is depicted using icons of the agents. The exclamation mark is used to represent focused attention. The sound that just appeared near agent B is shown as a set of concentric lines. There are different icons for every situation.

5. Behaviour displayed by the model

Following are three subsections describing the behaviour of the model, when the agents behave in a stimulus–response (Section 5.1), conscious (Section (5.2) and in a conscious goal-directed (Section 5.3) manner. Irrelevant and repeated facts have been omitted for brevity within the listings of the trace, see Appendices A–C. Only true facts and revisions are shown. For the meta-level relations `belief`, `observation_result` and `currently` the negatives, i.e. `belief(` … `, neg)`, have been left out, except for `has_property (tea, fresh)`. The label `revised` in front of a fact means that the fact is retracted by the revision system, thus its truth value has changed from `true` to `unknown`. When a fact was `true` and is changed to `false`, usually by newly transferred information, it is labelled `inverted`. The numbers of the rows correspond roughly with the same situations in each



Status of agent A

Characteristics

```
reflex_behaviour
```

Observations

```
there_is_noise_outside
at_position(agent_A,p1)
at_position(agent_B,p2)
has_property(agent_B,looks_interested)
```

Beliefs

```
at_position(pot2,p3)
contains(pot2,tea)
not has_property(tea,fresh)
```

Incoming Modalities

Incoming Communication

Outgoing Communication

```
at_position(pot2,p3)
contains(pot2,tea)
not has_property(tea,fresh)
```

Outgoing Modalities

```
non_verbal:not tastes_good
non_verbal:at_position(pot2,p3)
verbal:contains(pot2,tea)
verbal:has_property(tea,fresh)
```

Goals

```
at_position(pattern_3,p0)
has_property(agent_A,disgusted_face)
has_property(agent_A,points_to(p3))
```

Actions

```
pull_disgusted_face
point_to(p3)
show(slide_3)
```

Status of agent B

Characteristics

```
reflex_behaviour
```

Observations

```
there_is_noise_outside
at_position(agent_A,p1)
has_property(agent_B,looks_interested)
```

Beliefs

Incoming Modalities

Incoming Communication

Outgoing Communication

Outgoing Modalities

Goals

Actions

```
look_away
stop_looking_interested
```

FIGURE 5. Screen shot. The state of affairs is reported in the status window, which displays the current state of the agents and depicts the world situation using icons.

of the traces, and are taken from Example 1 in Section 2.3. Events 5 and 6 of the complete tea story, see Example 1, happen simultaneously, as do events 10 and 12, followed directly by 11a and 11b.

5.1. A TRACE FOLLOWING THE COMMUNICATION PROTOCOL BY STIMULUS–RESPONSE BEHAVIOUR

In this process, both agents follow the protocol by stimulus–response behaviour. For a full listing of the trace see Appendix A. The starting situation is that the world has a pot2 at p3 containing old tea, B is looking forward, uninterestedly, agent A is at p2. Agent A knows that pot2 at p2 contains old tea, in the symbolic object level. At the symbolic meta-level agent A believes these facts and wants to communicate these facts. The `goal_generation` component reasons using this information at the symbolic meta-level with knowledge rules as described in Section 4.2 deriving that certain information should be communicated using certain modalities. The results are in row 2. Also the goals which must be achieved to materialize the communication are derived. The observation_generation components of both agents derive where they want to look at, these to_be_observed facts are transferred to the external world by the links. These facts stay the same throughout the simulation, although modifiers are applied by looking away. They are not shown in the traces.

These goals are transmitted by the link called selected_goals from the goal_generation component to the action_generation component. The action_generation component reasons at the symbolic meta-level in a stimulus–response manner, deriving that the action go_to(p1) should be performed, as can be seen in row 3. The link initiated_actions then transports the to_be_performed(go_to(p1)) fact to the output of agent A. It is then transported by the actions_and_observations_A link to the external_world component.

The actions and observations now happen in the external world, and after some time new observation results become available, these are transported by the links observation_results_A and B to the respective agents. Inside agent B the observation_result information is transferred, among others, by the observation_results_to_ag link to the action_generation component. This component then reasons in a stimulus–response manner, from the observations to the actions to be done, in this case that agent B should start looking interested as somebody has just stepped up to the microphone stand, see row 4. The observation results are also transferred in a similar fashion to the action_generation component of agent A. The previous observation result that agent A is at p2 is retracted, and the observation result that agent A is at p1 is made true. Since this invalidates one or more premises of the knowledge rule which derived the action to go to p1 in row 3, this conclusion is no longer valid. The run-time truth maintenance system of DESIRE keeps track of this, and retracts the no longer needed conclusion. The statement to_be_performed(go_to(p1)) becomes unknown, and this new truth-value is further propagated by the links.

5.1.1. The first attempt at communication. At row 5/6 the world changes. Agent B starts to look interested, however, also some distracting noise outside appears. Both agents observe this, yet whilst agent A sees that B is paying attention and wants to start communicating, agent B is distracted by the noise outside and looks away. The component action_generation in agent A derives the applicable actions to achieve the goals specified by goal_generation, and the beliefs still present on the input of action_generation, as the situation seems ok to start communicating. In the action_generation component of agent B the previous action to_be_performed(look_interested) is retracted, and new conclusions are reached: to_be_performed(look away) and to_be_performed (stop_looking_interested). The actions of both agents are performed.

As this happens, agent B is looking away and no longer observing agent A. Agent A, however, does see what is going on. It derives that communication should be stopped, and the attention of agent B should be attracted, by tapping the microphone.

In row 8a agent B observes that there is noise inside, caused by tapping the microphone by agent A in row 7, and concludes in the component action_generated that it should perform look_towards_front. In row 8b agent B has looked back, and since it is reasoning in a stimulus response manner did not remember that agent A was in the speakers position, but now he observes this fact again. Agent B concludes he should be looking interested. Agent A observes that B is looking back again, and this fact retracts the original to_be_performed(tap_microphone) and causes the derivation of to_be_performed(stop_tapping). As agent A in row 9 observes that B is paying attention again, by observation_result(has_property(agent_B, looks_interested), pos), it starts the communication actions again.

5.1.2. Successful communication interpretation. In row 10/11/12 the agents observe that A is communicating. Note that agent A also observes that it is communicating, by its observation of its own activities. Agent A concludes that since the results of the communication actions are now visible it can stop doing so, also it decides to pull a questioning face, asking if the communication has succeeded. Agent B observes the communicative aspects of A, the component information_interpretation concludes that they should be stored and they are sent by the link new_world_info to the maintenance_of_world_information component. This link performs a downward reflection, the atom new_world_info (has_property(agent_A, disgusted_face), pos) at the symbolic meta-level causes the positive assumption of has_property(agent_A, disgusted_face) at the symbolic object level. The component maintenance_of_world_information reasons trivially at the symbolic object level, concluding all its input facts, thus remembering the world state. The links going out from the maintenance_of_world_information component then perform upward reflection, where among others the symbolic object level fact that at_position(pattern_3, p0) is true is first transformed to a built-in type, true(at_position(pattern_3, p0)), this type is then transformed to a belief(at_position(pattern_3, p0), pos) at the inputs of the receiving components. The information_interpretation component also receives the new beliefs. It concludes that these new beliefs mean something and several facts have been communicated using the modalities; the has_been_communicated_by_modality facts in row 10/11/12. It then integrates these partial communicative results and concludes that several facts have been communicated; the has_been_communicated facts. These facts are also sent to be stored in the world model of agent B.

At row 13 the communicated information becomes available in agent B as beliefs. Meanwhile agent B observes that agent A is looking questioningly. Since it has received information, it decides to put up a thumb in response.

In row 14 agent B's thumb is now up, and agent A observes this, decoding the message in the component information_interpretation as a communication by a non_verbal modality. Realizing that this message is not modified by other communicated facts the has_been_communicated(communication_succeeded, pos) can be derived. The action_generation component of agent_A now derives that it can look normal again and return to position p2. In the next row, row 15, agent A arrives at p2 thus ending this run of the prototype implementation.

5.2. A TRACE FOLLOWING THE COMMUNICATION PROTOCOL CONSCIOUSLY

In this process, both agents follow the communication protocol in a conscious manner. For a full listing of the trace see Appendix B. The presentation of the trace starts at a moment when the component external world considers the world to be in a certain state, see the material level in row 1, agent A maintains the situation in the world to be as stated in the symbolic object level column of row 1 and the corresponding beliefs at the symbolic meta-level are present, as shown in row 1. The component `goal_generation` has just generated the information to be communicated, as shown in row 1 as well. All components and information links are *awake*. This means that they will process as soon as the input information changes. Components will derive new or retract old and invalid inferences, information links will transfer the information from the source to the destination of the link.

As shown in row 2, `goal_generation` derives that, using the knowledge presented in Section 4.2, the information to be communicated should be communicated using certain modalities, and concludes goals to be achieved. These all appear on the output of the component, changing the source of the link selected goals, which transfers only the goals to be achieved to `action_generation`. This component then processes the new information, as shown in row 3, using the knowledge presented in Section 4.2. The output is transferred to the component `external_world`, where it is processed and the results of the action will be incorporated in the observation results returned to the agents, the observation results are shown in the symbolic meta-level of row 4. The components `information_interpretation` select the newly observed information to be stored in the components `maintenance of world_information`, as denoted in the symbolic object level column of row 4. The information then becomes a belief, also shown in row 4. The component `goal_generation` of agent B then derives based on the new beliefs that it should look interested. The component `action_generation` derives the corresponding action to be performed.

5.2.1. The first attempt at communication. The `external_world` incorporates the changes, as in the material level of row 5/6. However, suddenly something unspecified causes noise outside to appear at the material level. So whilst agent A observes B looking interested, and ignores the noise, continuing to derive communication actions, by `action generation`, agent B is distracted, and looks away, no longer looking interested at the speaker.

In row 7, the actions have had their effect on the world situation, agent A observes the state of agent B, stores this at the symbolic object level and based on its beliefs decides to alert agent B by tapping the microphone. Meanwhile, agent B is looking away and is oblivious to the communication attempted by agent A. Row 8 denotes that the sound outside has disappeared, the effects of the previous actions in the material level. Agent B observes this, stores, and believes it and then decides to look towards the speaker again, remembering that agent A was at p1, it immediately concludes it should also be looking interested, as agent A is at p1, the microphone stand.

After that, in row 9, agent A observes that agent B is looking away, stores these belief in its world model and continues to derive it should stop communicating and start attracting attention. Agent B observes, and believes that the results of the previous actions are now present in the world. Thus, it revises the `to_be_performed` actions from row 8.

5.2.2. Successful communication interpretation. The row labelled 10/11/12 denotes agent A observing the success of its actions, and agent B understanding the communication.

Agent A observes and believes the success of the previous actions and revises the intentions to do so. It also pulls a questioning face, asking for feedback. Agent B observes the meangingful situations, and believes them. It then starts interpreting them by first deriving the `has_been_communicated_by_modality` facts before integrating them into `has_been_communicated` facts. These are also stored in B's beliefs, appearing at the symbolic object level of agent B and later in the symbolic meta-level of agent B. (The dotted lines denote passage of time.)

Agent A displays a questioning face at row 13, and whilst agent B revises its beliefs and actions, agent B answers. Agent B reacts by deciding `to_be_performed(put_thumb_up)`. In row 14, agent A observes, believes and interprets this, concluding the communication was successful. As this conclusion is reached, the conclusions from row 1 about the information to be communicated are now revised. The `to_be_communicated_by_modality` further derived from the `to_be_communicated` information are also revised. Agent A then goes to p2, ending the simulation.

5.3. A TRACE FOLLOWING THE COMMUNICATION PROTOCOL IN A GOAL-DIRECTED MANNER

In this process, both agents follow the communication protocol in a goal-directed, conscious manner. For a full listing of the trace see Appendix C. The process starts like the other traces, with agents A and B at `p2`, agent A believes the old tea is in `pot2` at `p3`, and agent A wants to communicate this. However, in row 2 no `to_be_achieved` goals are derived for starting the communication, since the protocol derivations now happen in a goal-directed manner in the component `goal_generation` and prevent goals to accomplish the direct communication, since first agent A needs to get B's attention. At row 3 agent A observes and believes that it is at `p2`, not `p1`, and derives the goal `to_be_achieved(at_position(agent_A, p1), pos)` in the component `goal_generation`. It continues in the component `action_generation` to derive the means to accomplish the goal, `to_be_performed(go_to_(p1))`.

As agent A arrives at `p1`, in row 4, it believes this and revises old observation results and beliefs. Also the goal is retracted, and thus all necessary actions, the single `go_to(p1)` in this case, are also retracted. Agent B now observes agent A, believes this new information, and derives the goal to look interested, finishing with the action to be done.

5.3.1. The first attempt at communication. In row 5/6, agent B looks interested and also in this trace suddenly a noise occurs outside, distracting agent B. Agent A observes agent B paying attention, and ignoring the noise, derives the goals to start the communication now. It proceeds with deriving the communication actions. Agent B observes the noise and itself looking interested. It decides proceed with following the goals of looking uninterested and away at the distraction. It concludes with the actions to achieve the goals, revising the old goal and action.

Agent A observes the problem in row 7. It observes the results of the communicative actions, and that B is looking away. It then believes these facts. From these beliefs it formulates the goals to remove the communication actions, and attract B's attention yet again. First, the goal is formulated that B should be looking back, this in then made more specific by determining that there should be a noise inside the room. The goals which were previously derived are now retracted, followed by their actions. The actions to accomplish the current goals are derived too.

As in row 8, agent A observes the communicative action results disappearing in the world, agent B observes the noise inside the room and decides on looking back. Remembering that agent A is at p1 it immediately derives the goal to be looking interested as it looks back at p1. Also both agents note that the distraction disappears. In row 9, agent B observes the results of its action and revises it goals and actions accordingly. Agent A observes that the conditions are right for communication to be attempted, and formulates goals to accomplish this. The actions are derived too.

5.3.2. Successful communication interpretation. In row 10/11/12, whose numbering refers to dots 10–12 from the description of Example 1, agent B interprets the communication attempt. It observes the state of the world, and succeeds in believing this. It then interprets these beliefs in terms of their modality and what they mean, combing them into new information to be stored. It then stores these new beliefs, below the dotted line, in the component `maintenance_of_world_information`. Note that it also formulates the goal that it should be looking interested, as someone is at the speakers chair and there are no distractions, but this goal does not lead to a `to_be_performed` action as it is already fulfilled. Meanwhile, agent A observes its success of executing actions and revises its facts accordingly. Agent A also concludes that the communication attempt is now visible, and that it should pull a questioning face by generating the necessary goals. The corresponding actions are concluded.

Agent A observes that it is looking questioningly and that the communicative information elements are now no longer visible in the world in row 13. It believes these observations and revises the goals and actions based on these beliefs. Agent B also observes the removal of the communicative elements, and believes and revises based on this. It also derives that `to_be_communicated(message_understood)` should be done. The modality is decided upon as `non_verbal`, and the representation of the message is made a goal to pursue. The corresponding action is selected. The previous intermediate conclusions for the interpretation of the incoming communication attempt are now revised, as there are no communicative elements in the world anymore, yet the concluded upon beliefs are still maintained in the `maintenance_of_world_information` component.

The communication attempt is noticed by agent A in row 14, where it observes, believes and then interprets the fact of B having this `thumb_up`. As it concludes the communication has succeeded it decides to go to p2 and revises the `to_be_communicated` and `to_be_communicated_by_modality` elements about the tea. It also concludes it can pursue the goal of not having a questioning face. The corresponding actions are then derived. At the same time agent B observes that its thumb is now up and believes this. It then revises the action to fulfil the goal, as it has been successful already, but not the goal, since A is still looking questioningly. As agent A then reaches p2 the simulation ends.

6. Result evaluation

In the previous section, the example communication process is followed in three different manners. In all cases, the communication succeeds in the end. One might infer from this

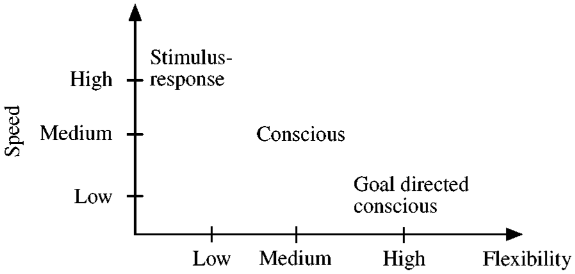


FIGURE 6. The speed of derivation vs. the possible quality of the output for the different cognitive models.

that the different methods to control the agent behaviour are equal. They are not. They differ in the flexibility of the behaviour that can be produced, and in the speed with which the output can be produced, see Figure 6.

The model needs to be capable of generating responses in time, especially when this time is very limited. An example would be a situation where our agent observes certain cues, such as slightly raised eyebrows in the current speaker, which indicate that the agent can become speaker, yet another agent also wants to become speaker. A quick response with the right signals will make the agent the current speaker, yet a slow response will make the agent have to wait. So some types of behaviour need to be fast.

Also flexibility is important. When confronted with a new situation, the agent must be able to cope with it. For example, when the available slides cannot convey the desired message, it has to be able to adapt and work around the problem. Therefore, the agent must be capable of flexibility in generating a response.

The speed of computation can be divided into these classes.

- *High*: stimulus–response—within the model only the symbolic object level is used. Only the observation_generation component and action_generation component perform any derivation.
- *Medium*: conscious—also the symbolic meta-level is used, and the information_interpretation component and the maintain_world_information component are used as well.
- *Low*: goal-directed—the symbolic meta-level is used, and the goal_generation component is used as well.

One point is that although the processes in the different traces are run with different methods determining the behaviour, some processes are done in the same manner each time. The interpretation of facts is a symbolic meta-level reasoning process that is done in a conscious manner in every trace. This is because the interpretation of elements of the world as having some extra meaning, then expanding the set of facts about the world with these derived meanings is by its very nature a conscious process. The existence of a belief set is assumed. This is augmented. Doing such a thing with observation results—which would be necessary when one wanted to interpret facts in a stimulus–response manner—would then augment these observation results with the newly derived meanings. Yet these new facts are not "observation results" they are additional information about the state of the world, which we usually term "beliefs". Therefore, stimulus–response world interpretation is not an option. A goal-directed interpretation process is possible, but it would entail that the agent is looking for a specific meaning in the world, as opposed to checking all of them.

Also, this architecture relies on the fact that the inputs and outputs are at a sufficiently high level of abstraction. This means that a sort of pattern-matching has already been performed on the raw sensor data; matching concepts to inputs. Thus, in the case of observing the projection screen, the raw sensor data of blotches of different colours has already been made into a declarative representation, e.g. `at_position(pattern_3, p0)`. In the same vein, the outputs of, for example `to_be_performed(go_to(p1))`, are taken as commands for lower level software to try to accomplish these tasks, for example to turn on motors, avoid obstacles and steer towards p1.

The flexibility of the derivation is judged by the power and expressivity of the rules that can be used in a certain model. The following gradations of flexibility can be discriminated from the previous traces:

- *Low:* stimulus–response—only object-level reasoning is done, using only the currently observed facts. Each rule can handle only the situations where the antecedent is observed and then do some set of actions.
- *Medium:* conscious—can reason at the symbolic meta-level to resolve a situation, using observed facts as well as remembered beliefs. Each rule can have an antecedent that builds upon the conclusions of previous rules, concluding perhaps again intermediate results. It can be applicable in a wide variety of situations, and be applied again on conclusions based on its own conclusions—recursion. Thus, rules can generate an analysis of the situation, and use the results of other rules. The derivation is much more flexible than in the stimulus–response case.
- *High:* goal-directed—can solve more difficult problems, by meta-level reasoning and by generating plans and (sub)goals using beliefs. Rules can use observations, beliefs and intermediate conclusions derived by conscious deliberation. Rules can conclude goals to achieve, affecting actions not just in the now, but providing consistent focus over time. Subgoals needed in order to reach an overarching goal can be planned. The rules can specify the broader goals to reach, instead of only immediate actions to perform. This makes the derivation more flexible than a derivation using conscious deliberation.

An example of something that is derived quickest in a stimulus–response manner: consider an agent pulling its hand away from a hot stove. The stimulus–response reasoning for this goes along the lines of `if hand_too_hot` then `retract_hand`, or it can even be done by the lower level (hardware) systems. This clearly takes much less time than first interpreting the `hand_too_hot` and updating the belief set before continuing with deriving the action as in the conscious case. In the goal-directed case additional subgoals may even need to be derived.

For example, the typical back-channel communications may be hard-wired as stimulus–response rules in the agent, as could be done for agent B putting up his thumb in response to the questioning face of agent A. Between humans this sort of thing is very common, consisting of small movements of the eyebrows and glancing towards or away of the speaker and such, to indicate interest and attention. People nodding affirmatively whilst listening to someone are another example. In our example model, the situation where agent B looks attentive when A is at the speakers position is an occurrence. In software agents sending acknowledgements that observation was successful would be a prime candidate. Deriving these communicative facts in a conscious or goal-directed conscious manner, while possible, would be slower.

Of course, there is a trade-off for the speed, the flexibility is reduced in the stimulus–response case. Consider, for example, row 13 in the example traces, where agent B decides to put up his thumb to acknowledge the receipt of new information. This is easily derivable in the conscious and goal-directed reasoning methods, a questioning face and the receipt of new information lead to concluding the thumb-up action. The stimulus–response case, in our example, also put the thumb up. Yet, since the arrival of new information is not an observation result, and a type of meta-level process knowledge the stimulus–response rule cannot take it into account. It merely put the thumb up because it saw a questioning face, and querying the internal state of the agent is not possible, thus this stimulus–response case cannot signal a thumb down in case nothing has been received. The response has been quick, but meaningless. The conscious and goal-directed conscious reasoning processes are capable of handling this, thus providing better quality reasoning than the stimulus–response case.

Goal-directed conscious reasoning may provide even more flexible reasoning. An example is when unexpected things occur. Conscious reasoning, like any algorithm, can handle it if this had been specified. Goal-directed reasoning can be more flexible using goals and subgoals. This flexibility has a penalty, again in speed, as the tree of goals and subgoals has to be searched. For example, consider an agent that can operate a conveyor belt and wants to go to a certain place. The conscious reasoning part may well be programmed to have the agent move using its wheels to the desired place. So would the goal-directed part state that the subgoal is attainable by using its wheels. But if the wheels of the agent break down, by having a flat tire, the conscious reasoning process may come to a grinding halt, giving no alternatives, unless this failure was explicitly catered for. The goal-directed conscious reasoning process may in this case be able to reason that the wheels are inoperable, but since it happens to be standing on top of the conveyor belt, and this conveyor belt will move an object, and this is the direction it wants to move in, it should activate the conveyor belt. In this case, the goal-directed reasoning can combine subgoals usually used for reasoning about moving objects on the conveyor belt to achieve goals for itself—as it happens to be standing on the belt. A possible drawback is that these emergent effects can also cause unforeseen problems as well as unforeseen advantage taking of opportunities. In this way, goal-directed reasoning can with less specification take advantage of synergies in goals.

These different types of reasoning, stimulus–response, conscious and goal directed, allow an agent to both react at its top speed when necessary, be able to handle common situations and have the capability to behave in a most thoughtful, flexible manner. These types of reasoning will need to be used simultaneously and operate at different semantic levels. In the next section, a formalization of the semantic levels and the transitions possible in the semantic model are delineated.

7. Semantic formalization of the model

In the semantic formalization a state-based semantics is chosen. Each of the states is formalized by a partial (or three-valued: with truth values *true*, *false* and *undefined*) model (Blamey, 1986; Langholm, 1988). The signature (the lexicon) for these partial models is partitioned according to the three semantic levels. To define semantics of the whole communication process, partial temporal models are used (Treur, 1994; Engelfriet & Treur, 1995). Within this approach the semantics of a process is formalized by a set of (alternative) partial temporal models, i.e. sequences of partial models. For an example of such a partial temporal model, see Table 1. Within each of the states in such a temporal model the three semantic levels can be distinguished. Representation relations are defined in the sense that symbolic expressions at the object level (such as `contains(pot2, tea))` refer to the state of the material level, symbolic expressions at the meta-level (such as `belief(contains(pot2, tea), pos))` refer to states of the object level, and the material level configurations (such as pattern 3) refer to object-level information (that pot 2 contains tea).

Only the following types of transitions between states are allowed in these temporal partial models:

Single-level transitions

- *World change transition*

  A change in the state of the material level
- *Object reasoning transition*

  A change in the state of the symbolic object level
- *Meta-reasoning transition*

  A change in the state of the symbolic meta-level

Level interaction transitions

- *Upward reflection transition*

  A change of the meta-level state under influence of the object-level state.
- *Downward reflection transition*

  A change of the object-level state under influence of the meta-level state.
- *Action execution transition*

  A change of the material-level state under influence of the meta-level state.
- *Observation transition*

  A change of the meta-level state under influence of the material-level state.

A levelled transition is a transition of the overall (three-level) state which is induced by a transition of one of the types distinguished above.

The transition types depicted in the first part of the trace are, subsequently the following.

- Meta-reasoning; three times:
  - — deciding that `contains(pot2, tea)` has to be communicated,
  - — selecting the goal `at_position(pattern3, p0)` to be achieved,
  - — determining the action `put_slide3` to achieve the goal
- Action execution.

The transition types depicted in the second part of the trace are, respectively the following.

- Observation (of `at_position(pattern3, p0))`.
- Downward reflection (including `at_position(pattern3, p0)` in B's world model).

- Upward reflection (identifying at_position(pattern3, p0) as a belief).
- Meta-reasoning (interpreting at_position(pattern3, p0) as a communication of the information contains(pot2, tea)).
- Downward reflection (including contains(pot2, tea) in B's world model).
- Upward reflection (identifying contains(pot2, tea) as a belief).

The formal definitions are given below. The elements used to describe the states (the ground atoms) are expressed in a language defined by an information type.

Definition 1 (*Information state*): An *information type* Σ is a structure of symbols defining a set of *ground atoms* $\mathtt{At}(\Sigma)$. An *information state* for an information type Σ is a mapping M: $\mathtt{At}(\Sigma) \rightarrow \{0, 1, u\}$ from the set of ground atoms $\mathtt{At}(\Sigma)$ to the set of truth values $\{0, 1, u\}$; i.e. a (*partial*) *model*. The set of all information states of information type Σ is denoted by $\mathtt{IS}(\Sigma)$. An information state M: $\mathtt{At}(\Sigma) \rightarrow \{\mathtt{0},\ \mathtt{1},\ u\}$ is called a *two-valued* information state if $M(a) \in \{\mathtt{0},\ \mathtt{1}\}$ for all $a \in \mathtt{At}(\Sigma)$. The set of two-valued information states for Σ is denoted by $\mathtt{IS}_2(\Sigma)$.

An example of a structure that defines an information type is a tuple of (sub-)sorts, constants, functions, and predicates of an order-sorted predicate logic. Each of the three levels has such an information type, see Sections 3.1.

Definition 2 (*Transition*): A *transition between information states* is a pair of partial models; i.e. an element $< S, S' >$ (also denoted by $S \rightarrow S'$) of $\mathtt{IS}(\Sigma) \times \mathtt{IS}(\Sigma)$. A *transition relation* is a set of these transitions, i.e. a relation on $\mathtt{IS}(\Sigma) \times \mathtt{IS}(\Sigma)$.

Behaviour is the result of transitions from one information state to another. If a transition relation is functional then it specifies deterministic behaviour. By applying transitions in succession, sequences of states are constructed. These sequences, also called traces (and interpreted as temporal models), formally describe behaviour.

Definition 3 (*Trace and temporal model*): A *trace* or *partial temporal model* of information type Σ is a sequence of information states $(M^t)_{t \in N}$ in $\mathtt{IS}(\Sigma)$. The set of all partial temporal models is denoted by $\mathtt{IS}(\Sigma)^{\mathtt{N}}$, or $\mathtt{Traces}(\Sigma)$.

A set of partial temporal models is a declarative description of the semantics of the behaviour of a process; each temporal model can be seen as one of the alternatives for the behaviour. Next, these notions are applied to the three levels distinguished in a communication process.

Definition 4 (*Levelled information state*): The *set of levelled information states* of the whole process is defined by $\mathtt{IS} = \mathtt{IS}_2(\Sigma^{\mathtt{mat}}) \times \mathtt{IS}(\Sigma^{\mathtt{obj}}) \times \mathtt{IS}(\Sigma^{\mathtt{meta}})$.

Here, $\mathtt{IS}_2(\Sigma^{\mathtt{mat}})$ represents the state of the material world, $\mathtt{IS}(\Sigma^{\mathtt{obj}})$ contains the state of the symbolic object level and $\mathtt{IS}(\Sigma^{\mathtt{meta}})$ contains the symbolic meta-level state. For the example communication process, $\Sigma^{\mathtt{obj}}$ = world_description. Similarly, $\Sigma^{\mathtt{meta}}$ = meta-info. The world is simulated by the external_world component using the same world model as for the agents, $\Sigma^{\mathtt{mat}}$ = world_description.

Levelled transitions and traces adhere to the levelled structure of the states: a levelled transition describes a levelled information state that changes in time. Following the levelled structure, only some types of transitions are allowed. For each of the levels a transition limited to this level (leaving untouched the other levels) is possible: a *world change*, an *object-level reasoning* step, or *meta-level reasoning* step. Two examples of transitions involving interaction between levels are *upward reflection* (information from the symbolic object level is lifted and incorporated in the symbolic meta-level), *downward reflection* (information from the symbolic meta-level influences the information at the symbolic object level). Other examples of transitions involving interaction between levels are *observation* (material level influences symbolic meta-level information), and *action execution* (symbolic meta-level information influences the material level). The following definition postulates that only these types of transitions are possible.

Definition 5 (*Levelled transition*):
(1) The following types of transitions are defined:

world change transition	$\mathtt{IS}_2(\Sigma^{\mathtt{mat}}) \rightarrow \mathtt{IS}_2(\Sigma^{\mathtt{mat}})$
object reasoning transition	$\mathtt{IS}(\Sigma^{\mathtt{obj}}) \rightarrow \mathtt{IS}(\Sigma^{\mathtt{obj}})$
meta-reasoning transition	$\mathtt{IS}(\Sigma^{\mathtt{meta}}) \rightarrow \mathtt{IS}(\Sigma^{\mathtt{meta}})$
upward reflection transition	$\mathtt{IS}(\Sigma^{\mathtt{obj}}) \times \mathtt{IS}(\Sigma^{\mathtt{meta}}) \rightarrow \mathtt{IS}(\Sigma^{\mathtt{meta}})$
downward reflection transition	$\mathtt{IS}(\Sigma^{\mathtt{obj}}) \times \mathtt{IS}(\Sigma^{\mathtt{meta}}) \rightarrow \mathtt{IS}(\Sigma^{\mathtt{obj}})$
action execution transition	$\mathtt{IS}_2(\Sigma^{\mathtt{mat}}) \times \mathtt{IS}(\Sigma^{\mathtt{meta}}) \rightarrow \mathtt{IS}_2(\Sigma^{\mathtt{mat}})$
observation transition	$\mathtt{IS}_2(\Sigma^{\mathtt{mat}}) \times \mathtt{IS}(\Sigma^{\mathtt{meta}}) \rightarrow \mathtt{IS}(\Sigma^{\mathtt{meta}})$

(2) A *levelled transition* is a transition: $\mathtt{IS} \rightarrow \mathtt{IS}$ which is based on a transition of one of the types defined in (1).

The model as described in Section 3 will only make these types of transitions. The maintenance_of_world_information component makes object-reasoning transitions, the other components in the agent make meta-reasoning transitions. The links between the agents and the external_world make trivial identity transitions of meta-reasoning. Inside the agents the links are all of this type too, except for the links connecting to the maintenance_of_world_information component. The new_world_info link makes downward reflection transitions, changing the model of the world at the symbolic object level by applying meta-level facts stating newly discovered information. The beliefs_to_ag, beliefs_to_cg and beliefs_to_ii links make upward reflection transitions. The upward reflections allow the facts that are stored in the believed world model to be included at the symbolic meta-level as beliefs.

In the simulation world change transitions, action execution transitions and observation transitions are simulated using object-reasoning transitions in the external_world component. If there would be no simulation, the world change transitions would happen in the physical world itself, observation transitions would occur when sensor data are collected and action execution transitions are effected by sending the actions to the robot actuators. This would also be taken care of by the external_world component.

Definition 6 (*Levelled trace*): (1) A *levelled trace* is a sequence of information states $(M^t)_{t \in \mathrm{N}}$ in IS. The set of all levelled traces is denoted by $\mathtt{IS}^{\mathrm{N}}$, or Traces.

(2) An element $(M^t)_{t \in \mathrm{N}} \in$ `Traces` is called *coherent* if for all time points t the step from M^t to M^{t+1} is defined in accordance with a levelled transition. The set of coherent levelled traces forms a subset `CTraces` of `Traces`.

Note that in Table 1 a coherent levelled trace is depicted. It is possible and sometimes necessary to define more constraints on the transitions. For example, physical laws for the material level, or if an observation transition leads to meta-level information `observation_result(a,pos)`, then `a` is true in the current world state, or if an object-reasoning transition adds information to the object level, then this information is in the deductive closure of the object-level knowledge (consisting of the object-level knowledge base and the information from the current object-level information state). Also the traces in Sections 6.1–6.3 are coherent traces.

8. Discussion

In the area of agents, communication processes play an important role. In this paper, a semantic model has been proposed for combined verbal and non-verbal communication processes and their interaction. The semantic model distinguishes three semantic or representation levels; it has been formalized on the basis of three-levelled partial temporal models (Treur, 1994; Engelfriet & Treur, 1995). These partial temporal models formalize both the material and mental processes and their relations. The relations between the levels can be defined as (partly circular) representation relations, as in Jonker and Treur (1997).

Moreover, using the compositional design method DESIRE (Brazier, Jonker & Treur, 1998), an executable generic process model has been designed that can be used to simulate different types of combinations of verbal and non-verbal communication processes. In an example, it has been shown how a non-verbal communication process can be modelled that adds and modifies content to a verbal communication process, but also provides a protocol for the (control of the) verbal communication process. In this example, different types of behaviour with respect to the communication protocol have been modelled: stimulus–response behaviour and variants of deliberate behaviour. The first type of behaviour has the advantage that it is very direct, but the second type of behaviour may be more flexible in unforeseen circumstances. The distinction made between stimulus–response behaviour and deliberate conscious reasoning in communication is also made in Chaib-draa and Levesque (1994). Three levels of cognitive control are distinguished, corresponding to the stimulus–response level, the conscious level and a heuristic-based level in between these two. The heuristic-based level, used in familiar, although not entirely routine, situations, can be expressed in the proposed model by rules guarded with beliefs and concluding actions.

Different types of interaction between verbal and non-verbal communication have been distinguished. The non-verbal communication can interact with the *content* of the verbal communication (type A). It then provides either additional information (type A1), with a subject unrelated (A1a) or related (A1b) to the verbal content, or the non-verbal communication affects the interpretation (type A2) of the verbal communication. Also, the non-verbal communication can interact with the *process* of the verbal communication (type B). The verbal communication process can be modified by stimulus–response reactions (B1) or by conscious deliberative reactions (B2). The agent-based interaction model presented is capable of handling all these types of interaction between verbal and non-verbal communication.

An interactive demo implementing the agent-based interaction model has been developed which allows the user to view the process step by step. The user can also take matters into his or her own hand and select the types of behaviour for each agent, or select specific actions by hand, taking the role of one or both agents.

The integration of non-verbal gestures together with verbal communication is also done in Bos, Huls and Claassen (1994). Their EDWARD system allows human–machine interaction with full multimodality, as does our model when the user takes the role of an agent. Their approach concentrates on using natural language, paying attention to context and referring expressions. The EDWARD system can generate and recognize non-verbally added additional information (i.e. type A1, see Section 2.3). Most of these concern the disambiguation of referring expressions, type A1b, providing information related to the verbal content. They conclude that multimodal human–computer interaction brings increased ease of use and higher cooperation closer to realization.

The manner in which in the example the non-verbal communication process modifies the meaning of the verbal communication can be used as a more general approach to model irony; see, e.g. Perrault (1990) for another approach to irony, based on an application of default logic to speech act theory.

The generic model for verbal and non-verbal communication presented here makes the communication process vulnerable to the effects of (other) causal patterns in the material world. This may occur as a negative aspect, as often communication is assumed ideal. However, in situations in which this assumption is not fulfilled, the approach introduced gives the possibility to explicitly model the interaction between the communication process and other material processes, and to take into account causality within the communication process, situated in the other processes; see also Dretske (1981). Moreover, this approach can be a starting point for simulations of the development of communication from an evolutionary perspective on communication, e.g. see Hauser (1996).

A possible modification of the model presented in this paper, is that if a particular action is derived very often given a set of observations, a stimulus–response rule can be added to cope with the often-occurring situation efficiently. Or, if goal-directed reasoning given some set of beliefs consistently concludes a certain goal, and the set of beliefs occurs often, a consciously deliberate rule may be generated to achieve the goal whenever the situation and beliefs present themselves. In this manner, an agent would learn to compute actions more efficiently in often occurring situations. To avoid an overabundance of rules at the conscious or stimulus–response response level, their number may be limited, and older ones may be forgotten. Also, the goal-directed deliberate reasoning needs to be able to turn off a certain pre-canned response, over-riding it with another response. Learning and forgetting inference rules is an interesting topic for further research.

A substantial part of research on communication (e.g. see Cohen, Morgan & Pollack, 1990) takes speech act theory (Austin, 1962; Searle, 1969) as a point of departure. Agent communication languages such as KQML (Finin, Labrou & Mayfield, 1997) have been proposed. Semantics of such languages is still an issue to be investigated further (Cohen

& Levesque, 1997). Speech acts are used to designate the effect an agent wishes to bring about in the world by communicating. Sometimes they are labelled as "promise" or "threaten", or more precisely "put receiver clear about employer–employee relationship, and use this to force decision" or "appeal to receiver to reconsider decision in light of new information". Speech act theory treats these things in great detail. The theory, however, only considers the content of the communication, abstaining from the material realization. The limitation is that these approaches focus on communication at a verbal or symbolic level, and do not cover non-verbal communication processes in which the physical aspects play an important role. In conclusion, speech act theory aids us when determining the content of any communication, but when the material level is important (as in non-verbal interaction), it is less applicable.

In Cassell *et al.* (1994), a model is presented that does integrate verbal and non-verbal communication, however, the model is restricted to only gestures (with the arms). The gestures are modelled in a naturalistic fashion, as the authors take great care to generate realistic verbal and non-verbal behaviour. The interaction is limited to only one kind, type A1 (see Section 4.1), where the non-verbal information content complements the verbal content. Several types of complementary gestures are distinguished, but the overall model is less refined and a formal semantics is not presented.

In comparison to Goldsmith and Spires (2000), at the sender's side, our processes `goal_ generation` and `action_generation_information` correspond to their 'goal formulation' and 'message formulation'. Moreover, at the receiver's side, our processes `observation_generation` and `interpretation` correspond to their 'message deconstruction', and 'reception'. In principle, our model could be extended to also include processes for their 'motivation' and 'elaboration' and 'message expectation'. A main difference is that our model explicitly treats verbal and non-verbal communication.

References

Austin, J. L. (1962). *How do Things with Words.* London: Oxford University Press.

Blamey, S. (1986). Partial logic. In D. Gabbay & F. Günthner, Eds. *Handbook of Philosophical Logic*, Vol. III, pp. 1–70. Dordrecht: Reidel.

Bos, E., Huls, C. & Claassen, W. (1994). EDWARD: full integration of language and action in a multimodal user interface. *International Journal of Human–Computer Studies*, **40,** 473–495.

Bradshaw, J. M., Ed. (1997). *Software Agents.* Cambridge, MA: AAAI Press/MIT Press.

Brazier, F. M. T., Jonker, C. M. & Treur, J. (1998). Principles of compositional multi-agent system development. In J. Cuena, Ed. *Proceedings of the 15th IFIP World Computer Congress, WCC'98, Conference on Information Technology and Knowledge Systems, IT&KNOWS'98*, pp. 347–360. IOS Press, 2001 (in press).

Brazier, F. M. T., Jonker, C. M. & Treur, J. (2000). Compositional design and reuse of a generic agent model. *Applied Artificial Intelligence Journal*, **14,** 491–538.

Cassell, J., Steedman, M., Badler, N. I., Pelachaud, C., Stone, M., Douville, B., Prevost, S. & Achorn, B. (1994). *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, Georgia Institute of Technology, Atlanta, USA.

Chaib-Draa, B. & Levesque, P. (1994). Hierarchical model and communication by signs, signals and symbols in multi-agent environments. In J. W. Perram & J. P. Mueller, Eds. *Distributed Software Agents and Applications*: *6th European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, MAAMAW'94, Lecture Notes in Computer Science, Vol. 1069. Berlin: Springer-Verlag. Also in: *Journal of Experimental and Theoretic Artificial Intelligence*, **8,** 7–20, 1996.

COHEN. P. R. & LEVESQUE, H. J. (1997). *Communicative actions for artificial agents.* In J. M. BRADSHAW, Ed. *Software Agents*, pp. 419–436. Cambridge, MA: AAAI Press/MIT Press.

COHEN, P. R., MORGAN, J. & POLLACK, M., Eds. (1990). *Intentions in Communication.* Cambridge, MA: MIT Press.

DRETSKE, F. I. (1981). *Knowledge and the Flow of Information.* Oxford: Basil Blackwell.

ENGELFRIET, J. & TREUR, J. (1995). Temporal theories of reasoning. *Journal of Applied Non-Classical Logics*, **5,** 239–261.

FININ, T., LABROU, Y. & MAYFIELD, J. (1997). KQML as an agent communication language. In J. M. BRADSHAW, Ed. *Software Agents*, pp. 291–316. Cambridge, MA: AAAI Press/MIT Press.

GOLDSMITH, S. & SPIRES, S. (2000). Agent communication using distributed metaobjects. In F. DIGNUM, B. CHAIB-DRAA & H. WEIGAND, Eds. *Proceedings of the Agent Communication Languages Workshop 1999,* ACL'99. pp. 161–169.

HAUSER, M. D. (1996). *The Evolution of Communication.* Cambridge, MA: MIT Press.

JONKER, C. M., KREMER, R., LEEUWEN VAN, P., PAN, D. & TREUR, J. (1999). Mapping visual to textual representation of knowledge in DESIRE. *Proceedings of the 12th International Conference on Industrial and Engineering Applications of AI and Expert Systems*, IEA/AIE'99. Lecture Notes in AI. Berlin: Springer-Verlag.

JONKER, C. M. & TREUR, J. (1997). Modelling an agent's mind and matter. *Proceedings of the 8th European Workshop on Modelling Autonomous Agents in a Multi-Agent World.* MAAMAW'97, Lecture Notes in AI, Vol. 1237, pp. 210–233. Berlin: Springer-Verlag.

LANGHOLM, T. (1988). Partiality, truth and persistence. CSL1 Lecture Notes, No. 15. Stanford University, Stanford.

PERRAULT, C. R. (1990). An application of default logic to speech act theory. In P. R. COHEN, J. MORGAN & M. POLLOCK, Eds. *Intentions in Communications*, pp. 161–186. Cambridge, MA: MIT Press.

SEARL, J. R. (1969). *Speech Acts*: *an Essay in the Philosophy of Languages.* New York: Cambridge University Press.

TREUR, J. (1994). Temporal semantics of meta-level architectures for dynamic control of reasoning. In L. FRIBOURG & F. TURINI, Eds. *Logic Program Synthesis and Transformation-Meta-Programming in Logic*, *Proceedings of the Fourth International Workshop on Meta-Programming in Logic, META'94.* Lecture Notes in Computer Science, Vol. 883. pp. 353–376. Berlin: Springer-Verlag.

VONGKASEM, L. & CHAIB-DRAA, B. (2000). ACL as a joint project between participants: a preliminary report. In: E. DIGNUM & M. GREAVES, Eds. *Issues in Agent Communication. Lecture Notes in Artificial Intelligence*, Vol. 1916, pp. 235–248. Berlin: Springer-Verlag.

Paper accepted for publication by Editor, Brian Gaines

Appendix A: A trace following the communication protocol by stimulus–response behaviour (see Section 5.1)

	External world	Agents	
	Material level	Symbolic object level	Symbolic meta-level

1	Pot 2 contains tea.
	Pot 2 is at p3.
	The tea is old.
	Agent B is looking towards the front.
	Agent B is not looking interested.
	Agent A and B are at p2.
		Agent A:
		at_position(pot2, p3)
		contains(pot2, tea),
		not has_property(tea, fresh)
			Agent A:
			to_be_communicated(has_property(tea, fresh), neg)
			to_be_communicated(contains(pot2, tea), pos)
			to_be_communicated(at_position(pot2, p3), pos)
			belief(at_position(pot2, p3), pos)
			belief(contains(pot2, tea), pos)
			belief(has_property(tea, fresh), neg)

2			Agent A:
			to_be_communicated_by_modality(contains(pot2, tea), pos, verbal)
			to_be_communicated_by_modality(has_property(tea, fresh), pos, verbal)
			to_be_communicated_by_modality(tastes_good, neg, non_verbal)
			to_be_communicated_by_modality(at_position(pot2, p3), pos, non_verbal)
			to_be_achieved(at_position(pattern_3, p0), pos)
			to_be_achieved(has_property(agent_A, points_to(p3)), pos)
			to_be_achieved(has_property(agent_A, disgusted_face), pos)

3			Agent A:
			observation_result(at_position(agent_A, p2), pos)
			observation_result(at_position(agent_B, p2), pos)
			to_be_performed(go_to(p1))

4	Agent A is at p1.		Agent A: observation_result(at_position(agent_A, p1), pos) revised observation_result(at_position (agent_A, p2), pos) revised to_be_performed(go_to(p1)) Agent B: observation_result(at_position(agent_A, p1), pos) to_be_performed(look_interested)

Appendix A: *continued*

	External world	Agents

5/6	Agent B looks interested.
	Noise outside.
		Agent A:
		observation_result(there_is_noise_outside, pos)
		observation_result(has_property(agent_B, looks_interested), pos)
		to_be_performed(pull_disgusted_face)
		to_be_performed(point_to(p3))
		to_be_performed(show(slide_3))
		Agent B:
		observation_result(there_is_noise_outside, pos)
		observation_result(has_property(agent_B, looks_interested), pos)
		revised to_be_performed(look_interested)
		to_be_performed(look_away)
		to_be_performed(stop_looking_interested)

7	Pattern3 is at p0. Agent A has a disgusted expression. Agent A points to p3. Agent B is looking away. Agent B is not looking interested.	Agent A: observation_result(at_position(pattern_3, p0), pos) observation_result(has_property(agent_A, disgusted_face), pos) observation_result(has_property(agent_A, points_to(p3)), pos) observation_result(has_property(agent_B, looks_away), pos) to_be_performed(do_not_point) to_be_performed(pull_no_disgusted_face) to_be_performed(tap_microphone) to_be_performed(remove(slide_3)) revised observation_result(has_property (agent_B, looks_interested), pos) revised to_be_performed(pull_disgusted_face) revised to_be_performed(point_to(p3)) revised to_be_performed(show(slide_3)) Agent B: observation_result(has_property(agent_B, looks_away), pos) revised observation_result(has_property (agent_B, looks_interested), pos) revised observation_result(at_position (agent_A, p1), pos) revised to_be_performed(look_away) revised to_be_performed(stop_looking_interested)

8a	Pattern3 is not at p0. Agent A has no disgusted expression. Agent A is not pointing to p3. Noise inside. No noise outside.	Agent A: at_position(pattern_3,p0)	Agent A: belief(at_position(pattern_3,p0),pos) observation_result(there_is_noise_inside,pos) revised to_be_performed(do_not_point) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised to_be_performed(tap_microphone) revised to_be_performed(remove(slide_3)) revised observation_result(has_property(agent_A,disgusted_face),pos) revised observation_result(there_is_noise_outside,pos) revised to_be_performed(pull_no_disgusted_face) revised observation_result(at_position(pattern_3,p0),pos) Agent B: observation_result(there_is_noise_inside,pos) to_be_performed(look_towards_front) revised observation_result(there_is_noise_outside,pos)
8b	Agent B is looking towards the front.	Agent A: not at_position(pattern_3,p0)	Agent A: inverted belief(at_position(pattern_3,p0),pos) to_be_performed(stop_tapping) revised observation_result(has_property(agent_B,looks_away),pos) Agent B: observation_result(at_position(agent_A,p1),pos) to_be_performed(look_interested) revised observation_result(has_property(agent_B,looks_away),pos) revised to_be_performed(look_towards_front)
9	No noise inside. Agent_B is looking interested.		Agent A: observation_result(has_property(agent_B,looks_interested),pos) to_be_performed(pull_disgusted_face) to_be_performed(point_to(p3)) to_be_performed(show(slide_3)) revised to_be_performed(stop_tapping) revised observation_result(there_is_noise_inside,pos) Agent B: observation_result(has_property(agent_B,looks_interested),pos) revised to_be_performed(look_interested) revised observation_result(there_is_noise_inside,pos)

Appendix A: *continued*

	External world	Agents	

10/
11/
12
	Pattern 3 is at p0.
	Agent A has disgusted expression.
	Agent A points to p3.
			Agent A:
			observation_result(at_position(pattern_3,p0),pos)
			observation_result(has_property(agent_A,disgusted_face),pos)
			observation_result(has_property(agent_A,points_to(p3)),pos)
			to_be_performed(do_not_point)
			to_be_performed(pull_no_disgusted_face)
			to_be_performed(pull_questioning_face)
			to_be_performed(remove(slide_3))
			revised to_be_performed(pull_disgusted_face)
			revised to_be_performed(point_to(p3))
			revised to_be_performed(show(slide_3))
			Agent B:
			observation_result(has_property(agent_A,disgusted_face),pos)
			observation_result(has_property(agent_A,points_to(p3)),pos)
			observation_result(at_position(pattern_3,p0),pos)

		Agent A: at_position(pattern_3,p0)	Agent A: belief(at_position(pattern_3,p0),pos)

		Agent B:
		has_property(agent_A, disgusted_face)
		has_property(agent_A, points_to(p3)
		at_position(pattern_3,p0)
			Agent B:
			belief(has_property(agent_A,disgusted_face),pos)
			belief(has_property(agent_A,points_to(p3)),pos)
			belief(at_position(pattern_3,p0),pos)
			has_been_communicated_by_modality(has_property(tea,fresh),pos,verbal)
			has_been_communicated_by_modality(contains(pot2,tea),pos,verbal)
			has_been_communicated_by_modality(tastes_good,neg,non_verbal)
			has_been_communicated_by_modality(at_position(pot2,p3),pos,non_verbal)
			has_been_communicated(contains(pot2,tea),pos)
			has_been_communicated(tastes_good,neg)
			has_been_communicated(at_position(pot2,p3),pos)
			has_been_communicated(has_property(tea,fresh),neg)

		Agent B:
		contains(pot2,tea)
		at_position(pot2,p3)
		not has_property(tea, fresh)
			Agent B:
			belief(at_position(pot2,p3),pos)
			belief(contains(pot2,tea),pos)
			belief(has_property(tea,fresh),neg)

13	pattern 3 is not at p0. Agent A has no disgusted expression. Agent A is not pointing to p3. Agent A has a questioning face.	Agent A: not at_position(pattern_3,p0) Agent B: not at_position(pattern_3,p0) not has_property(agent_A, disgusted_face) not has_property(agent_A, points_to(p3))	Agent A: observation_result(has_property(agent_A,questioning_face),pos) revised to_be_performed(do_not_point) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised to_be_performed(pull_questioning_face) revised to_be_performed(remove(slide_3)) revised observation_result(has_property(agent_A,disgusted_face),pos) revised to_be_performed(pull_no_disgusted_face) revised observation_result(at_position(pattern_3,p0),pos) Agent B: observation_result(has_property(agent_A,questioning_face),pos) to_be_performed(put_thumb_up) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised observation_result(has_property(agent_A,disgusted_face),pos) revised observation_result(at_position(pattern_3,p0),pos) Agent A: inverted belief(at_position(pattern_3,p0),pos) Agent B: inverted belief(at_position(pattern_3,p0),pos) inverted belief(has_property(agent_A,disgusted_face),pos) inverted belief(has_property(agent_A,points_to(p3)),pos) revised has_been_communicated(at_position(pot2,p3),pos) revised has_been_communicated_by_modality(at_position(pot2,p3),pos,non_verbal) revised has_been_communicated(has_property(tea,fresh),neg) revised has_been_communicated(tastes_good,neg) revised

Appendix A: *continued*

	External world		Agents
			has_been_communicated_by_modality(has_property(tea, fresh), pos, verbal) revised has_been_communicated_by_modality(contains(pot2, tea), pos, verbal) revised has_been_communicated(contains(pot2, tea), pos) revised has_been_communicated_by_modality(tastes_good, neg, non_verbal)

14	Agent_B has his thumb up.		Agent A:
			observation_result(has_property(agent_B, thumb_up), pos)
			to_be_performed(pull_no_questioning_face)
			to_be_performed(go_to(p2))
			Agent B:
			observation_result(has_property(agent_B, thumb_up), pos)
			revised to_be_performed(put_thumb_up)

		Agent A: has_property(agent_B, thumb_up)	Agent A: belief(has_property(agent_B, thumb_up), pos) has_been_communicated(communication_succeeded, pos) has_been_communicated_by_modality(communication_succeeded, pos, non_verbal revised to_be_communicated(at_position(pot2, p3), pos) revised to_be_communicated_by_modality(has_property(tea, fresh), pos, verbal) revised to_been_communicated_by_modality(tastes_good, neg, non_verbal) revised to_be_achieved(at_position(pattern_3, p0), pos) revised to_be_achieved(has_property(agent_A, disgusted_face), pos) revised to_be_communicated(has_property(tea, fresh), neg) revised to_be_communicated_by_modality(contains(pot2, tea), pos, verbal) revised to_be_achieved(has_property(agent_A, points_to(p3)), pos) revised to_be_communicated_by_modality(at_position(pot2, p3), pos, non_verbal) revised to_be_communicated(contains(pot2, tea), pos)
15	Agent A is at p2.		

Appendix B: A trace following the communication protocol consciously

(see Section 5.2)

External world		Agents	
	Material level	Symbolic object level	Symbolic meta-level

1	Pot2 contains tea.
	Pot 2 is at p3.
	The tea is old.
	Agent B is looking towards the front.
	Agent B is not looking interested.
	Agent A and B are at p2.
		Agent A:
		at_position(pot2,p3)
		contains(pot2, tea)
		not has_property(tea, fresh)
		not has_property(agent_B, looks_away)
			Agent A:
			to_be_communicated(has_property(tea, fresh), neg)
			to_be_communicated(contains(pot2, tea), pos)
			to_be_communicated(at_position(pot2, p3), pos)
			belief(at_position(pot2, p3), pos)
			belief(contains(pot2, tea), pos)
			belief(has_property(tea, fresh), neg)
			belief(has_property(agent_B, looks_away), neg)

2			Agent A:
			to_be_communicated_by_modality(contains(pot2, tea), pos, verbal)
			to_be_communicated_by_modality(has_property(tea, fresh), pos, verbal)
			to_be_communicated_by_modality(tastes_good, neg, non_verbal)
			to_be_communicated_by_modality(at_position(pot2, p3), pos, non_verbal)
			to_be_achieved(at_position(pattern_3, p0), pos)
			to_be_achieved(has_property(agent_A, points_to(p3)), pos)
			to_be_achieved(has_property(agent_A, disgusted_face), pos)

3		Agent A:
		at_position(agent_A, p2)
		at_position(agent_B, p2)
			Agent A:
			observation_result(at_position(agent_A, p2), pos)
			observation_result(at_position(agent_B, p2), pos)
			belief(at_position(agent_A, p2), pos)
			belief(at_position(agent_B, p2), pos)
			to_be_performed(go_to(p1))

4	Agent A is at p1.	Agent A: at_position(agent_A, p1) not at_position(agent_A, p2)	Agent A: observation_result(at_position(agent_A, p1), pos) revised observation_result(at_position(agent_A, p2), pos) belief(at_position(agent_A, p1), pos) inverted belief(at_position(agent_A, p2), pos) revised to_be_performed(go_to(p1))

Appendix B: *continued*

External world		Agents	

		Agent B:
		at_position(agent_A,p1)
			Agent B:
			observation_result(at_position(agent_A,p1),pos)
			belief(at_position(agent_A,p1),pos)
			to_be_performed(look_interested)

5/6	Agent B looks interested.
	Outside noise.
		Agent A:
		
		there_is_noise_outside
		has_property(agent_B, looks_interested)
		has_property(agent_B, paying_attention)
			Agent A:
			
			observation_result(there_is_noise_outside,pos)
			observation_result(has_property(agent_B, looks_interested),pos)
			belief(there _is_noise_outside,pos)
			belief(has_property(agent_B, looks_interested),pos)
			belief(has_property(agent_B,paying_attention),pos)
			to_be_performed(pull_disgusted_face)
			to_be_performed(point_to(p3))
			to_be_performed(show(slide_3))

		Agent B:
		
		there_is_noise_outside
		has_property(agent_B, looks_interested)
			Agent B:
			
			observation_result(there_is_noise_outside,pos)
			observation_result(has_property(agent_B,looks_interested),pos)
			belief(there_is_noise_outside,pos)
			belief(has_property(agent_B,looks_interested),pos)
			to_be_performed(look_away)
			to_be_performed(stop_looking_interested)
			revised to_be_performed(look_interested)

7	Pattern3 is at p0.
	Agent A has a disgusted expression
	Agent A points to p3.
	Agent B is looking away.
	Agent B is not looking interested.
		Agent A:
		
		not has_property(agent_B, looks_interested)
		not has_property (agent_B,paying_attention)
		has_property(agentB, looks_away)
			Agent A:
			
			observation_result(at_position(pattern_3,p0),pos)
			observation_result(has_property(agent_A,disgusted_face),pos)
			observation_result(has_property(agent_A,points_to(p3)),pos)
			observation_result(has_property(agent_B, looks_away),pos)
			revised observation_result(has_property(agent_B, looks_interested),pos)
			belief(at_position(pattern_3,p0),pos)
			belief(has_property(agent_A,disgusted_face),pos)
			belief(has_property(agent_A,points_to(p3)),pos)

		at_position(pattern_3,p0) has_property(agentA, disgusted_face has_property(agent_A, points_to(p3)) Agent B: has_property(agent_B,looks_away not has_property(agent_B,looks_interested)	belief(has_property(agent_B,looks_away),pos) inverted belief(has_property(agent_B,looks_interested),pos) inverted belief(has_property(agent_B,paying_attention),pos) to_be_performed(do_not_point) to_be_performed(pull_no_disgusted_face) to_be_performed(tap_microphone) to_be_performed(remove(slide_3)) revised to_be_performed(pull_disgusted_face) revised to_be_performed(point_to(p3)) revised to_be_performed(show(slide_3)) Agent B: observation_result(has_property(agent_B,looks_away),pos) revised observation_result(has_property(agent_B,looks_interested),pos) revised observation_result(at_position(agent_A,p1),pos) belief(has_property(agent_B,looks_away),pos) inverted belief(has_property(agent_B,looks_interested),pos) revised to_be_performed(look_away) revised to_be_performed(stop_looking_interested)
8	Pattern3 is not at p0. Agent A has no disgusted expression. Agent A is not pointing to p3. Noise inside. No noise outside.	Agent A: there_is_noise_inside not there_is_noise_outside not at_position (pattern_3,p0) not has_property (agent_A,disgusted_face) not has_property (agent_A,points_to(p3)) Agent B: not there_is_noise_outside there_is_noise_inside	Agent A: observation_result(there_is_noise_inside,pos) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised observation_result(has_property(agent_A,disgusted_face),pos) revised observation_result(there_is_noise_outside,pos) revised observation_result(at_position(pattern_3,p0),pos) belief(there_is_noise_inside,pos) inverted belief(there_is_noise_outside,pos) inverted belief(at_position(pattern_3,p0),pos) inverted belief(has_property(agent_A,disgusted_face),pos) inverted belief(has_property(agent_A,points_to(p3)),pos) revised to_be_performed(do_not_point) revised to_be_performed(tap_microphone) revised to_be_performed(remove(slide_3)) revised to_be_performed(pull_no_disgusted_face)

Appendix B: *continued*

External world		Agents	

			Agent B:
			
			observation_result(there_is_noise_inside, pos)
			revised observation_result(there_is_noise_outside, pos)
			belief(there_is_noise_inside, pos)
			inverted belief(there_is_noise_outside, pos)
			to_be_performed(look_towards_front)
			to_be_performed(look_interested)

9	Agent_B is looking interested.
	Agent B is looking towards the front.
		Agent A:
		
		has_property(agent_B, looks_interested)
		not has_property (agent_B,, looks_away)
		has_property(agent_B, paying_attention)
		
		Agent B:
		
		not has_property (agent_B, looks_away), pos
		has_property(agent_B, looks_interested), pos
			Agent A:
			
			observation_result(has_property(agent_B, looks_interested), pos)
			revised observation_result(has_property(agent_B, looks_away), pos)
			inverted belief(has_property(agent_B, looks_away), pos)
			belief(has_property(agent_B, looks_interested), pos)
			belief(has_property(agent_B, paying_attention), pos)
			to_be_performed(pull_disgusted_face)
			to_be_performed(stop_tapping)
			to_be_performed(point_to(p3))
			to_be_performed(show(slide_3))
			
			Agent B:
			
			observation_result(at_position(agent_A, p1), pos)
			observation_result(has_property(agent_B, looks_interested), pos)
			revised observation_result(has_property(agent_B, looks_away), pos)
			inverted belief(has_property(agent_B, looks_away), pos)
			belief(has_property(agent_B, looks_interested), pos)
			revised to_be_performed(look_interested)
			revised to_be_performed(look_towards_front)

10/
11/
12
	No noise inside.
	Pattern3 is at p0.
	Agent A has a disgusted expression.
	Agent A points to p3.
		Agent A:
		
		not here_is_noise_inside
		at_position(pattern_3, p0)
		has_property(agent_A, disgusted_face)
		has_property(agent_A, points_to(p3))
		
		Agent B:
		
		not there_is_noise_inside
		has_property(agent_A, disgusted_face)
		has_property(agent_A, points_to(p3))
		at_position(pattern_3, p0)
			Agent A:
			
			observation_result(at_position(at_position(pattern_3, p0), pos)
			observation_result(has_property(agent_A, disgusted_face), pos)
			observation_result(has_property(agent_A, points_to(p3)), pos)
			revised observation_result(there_is_noise_inside, pos)
			inverted belief(there_is_noise_inside, pos)
			belief(at_position(pattern_3, p0), pos)
			belief(has_property(agent_A, disgusted_face), pos)
			belief(has_property(agent_A, points_to(p3)), pos)
			to_be_performed(do_not_point)
			to_be_performed(pull_no_disgusted_face)
			to_be_performed(pull_questioning_face)
			to_be_performed(remove(slide_3))
			revised to_be_performed(pull_disgusted_face)
			revised to_be_performed(stop_tapping)
			revised to_be_performed(point_to(p3))
			revised to_be_performed(show(slide_3))

		contains(pot2, tea)
		at_position(pot2, p3)
		not has_property(tea, fresh)
			Agent B:
			
			observation_result(at_position(pattern_3, p0), pos)
			observation_result(has_property(agent_A, disgusted_face), pos)
			observation_result(has_property(agent_A, points_to(p3)), pos)
			revised observation_result(there_is_noise_inside, pos)
			inverted belief(there_is_noise_inside, pos)
			belief(at_position(pattern_3, p0), pos)
			belief(has_property(agent_A, disgusted_face), pos)
			belief(has_property(agent_A, points_to(p3)), pos)
			has_been_communicated_by_modality(has_property(tea, fresh), pos, verbal)
			has_been_communicated_by_modality(contains(pot2, tea), pos, verbal)
			has_been_communicated_by_modality(tastes_good, neg, non_verbal)
			has_been_communicated_by_modality(at_position(pot2, p3), pos, non_verbal)
			has_been_communicated(contains(pot2, tea), pos)
			has_been_communicated(tastes_good, neg)
			has_been_communicated(at_position(pot2, p3), pos)
			has_been_communicated(has_property(tea, fresh), neg)
			
			belief(contains(pot2, tea), pos)
			belief(at_position(pot2, p3), pos)
			belief(has_property(tea, fresh), neg)

Appendix B: *continued*

	External world	Agents	
13	Agent A has a questioning face. Pattern3 is not at p0. Agent A has no disgusted expression. Agent A is not pointing to p3.	Agent A: not at_position (pattern_3,p0) not has_property (agent_A,disgusted_face) not has_property (agent_A,points_to(p3)) has_property(agent_A, questioning_face) Agent B: not at_position (pattern_3,p0) not has_property (agent_A,disgusted_face) not has_property agent_A,points_to(p3)) has_property(agent_A, questioning_face)	Agent A: observation_result(has_property(agent_A,questioning_face),pos) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised observation_result(has_property(agent_A,disgusted_face),pos) revised observation_result(at_position(pattern_3,p0),pos) inverted belief(at_position(pattern_3,p0),pos) inverted belief(has_property(agent_A,disgusted_face),pos) inverted belief(has_property(agent_A,points_to(p3)),pos) belief(has_property(agent_A,questioning_face),pos) revised to_be_performed(do_not_point) revised to_be_performed(pull_questioning_face) revised to_be_performed(remove(slide_3) revised to_be_performed(pull_no_disgusted_face) Agent B: observation_result(has_property(agent_A,questioning_face),pos) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised observation_result(has_property(agent_A,disgusted_face),pos) revised observation_result(at_position(pattern_3,p0),pos) inverted belief(at_position(pattern_3,p0),pos) inverted belief(has_property(agent_A,disgusted_face),pos) inverted belief(has_property(agent_A,points_to(p3)),pos) belief(has_property(agent_A,questioning_face),pos to_be_performed(put_thumb_up) revised has_been_communicated(at_position(pot2,p3),pos) revised has_been_communicated_by_modality(at_position(pot2,p3),pos,non_verbal) revised has_been_communicated(has_property(tea, fresh),neg) revised has_been_communicated(tastes_good,neg) revised has_been_communicated_by_modality(has_property(tea, fresh),pos,verbal) revised has_been_communicated_by_modality(contains(pot2,tea),pos,verbal) revised has_been_communicated(contains(pot2,tea),pos) revised has_been_communicated_by_modality(tastes_good,neg,non_verbal)

14	Agent B has his thumb up.	Agent A: has_property(agent_B, thumb_up) Agent B: has_property(agent_B, thumb_up)	Agent A: observation_result(has_property(agent_B, thumb_up), pos) belief(has_property(agent_B, thumb_up), pos) has_been_communicated_by_modality(communication_succeeded, pos, non_verbal) has_been_communicated(communication_succeeded, pos) to_be_performed(pull_no_questioning_face) to_be_performed(go_to(p2)) revised to_be_communicated(at_position(pot2, p3), pos) revised to_be_communicated_by_modality(has_property(tea, fresh), pos, verbal) revised to_be_communicated_by_modality(tastes_good, neg, non_verbal) revised to_be_achieved(at_position(pattern_3, p0), pos) revised to_be_achieved(has_property(agent_A, disgusted_face), pos) revised to_be_communicated(has_property(tea, fresh), neg) revised to_be_communicated_by_modality(contains(pot2, tea), pos, verbal) revised to_be_achieved(has_property(agent_A, points_to(p3)), pos) revised to_be_communicated_by_modality(at_position(pot2, p3), pos, non_verbal) revised to_be_communicated(contains(pot2, tea), pos) Agent B: observation_result(has_property(agent_B, thumb_up), pos) revised to_be_performed(put_thumb_up)
15	Agent A is at p2.		

Appendix C: A trace following the communication protocol in a goal-directed manner

(see Section 5.3)

External world		Agents	
	Material level	Symbolic object level	Symbolic meta-level

1	Pot 2 contains tea.
	Pot 2 is at p3.
	The tea is old.
	Agent B is looking towards the front.
	Agent B is not looking interested.
	Agent A and B are at p2.
		Agent A:
		at_position(pot2, p3)
		contains(pot2, tea)
		not has_property(tea, fresh)
		not has_property(agent_B, looks_away)
			Agent A:
			to_be_communicated(has_property(tea, fresh), neg)
			to_be_communicated(contains(pot2, tea), pos)
			to_be_communicated(at_position(pot2, p3), pos)
			belief(at_position(pot2, p3), pos)
			belief(contains(pot2, tea), pos)
			belief(has_property(tea, fresh), neg)

2			Agent A:
			to_be_communicated_by_modality(contains(pot2, tea), pos, verbal)
			to_be_communicated_by_modality(has_property(tea, fresh), pos, verbal)
			to_be_communicated_by_modality(tastes_good, neg, non_verbal)
			to_be_communicated_by_modality(at_position(pot2, p3), pos, non_verbal)

3		Agent A:
		at_position(agent_A, p2)
		at_position(agent_B, p2)
			Agent A:
			observation_result(at_position(agent_A, p2), pos)
			observation_result(at_position(agent_B, p2), pos)
			belief(at_position(agent_A, p2), pos)
			belief(at_position(agent_B, p2), pos)
			to_be_achieved(at_position(agent_A, p1), pos)
			to_be_performed(go_to(p1))

4	Agent A is at p1.	Agent A: at_position(agent_A, p1) not at_position(agent_A, p2) Agent B: at_position(agent_A, p1)	Agent A: observation_result(at_position(agent_A, p1), pos) revised observation_result(at_position(agent_A, p2), pos) belief(at_position(agent_A, p1), pos) inverted belief(at_position(agent_A, p2), pos) revised to_be_achieved(at_position(agent_A, p1), pos) revised to_be_performed(go_to(p1))

			Agent B:
			observation_result(at_position(agent_A, p1), pos)
			belief(at_position(agent_A, p1), pos)
			to_be_achieved(has_property(agent_B, looks_interested), pos)
			to_be_performed(look_interested)

5/6	Agent B looks interested. Outside noise.	Agent A:
		there_is_noise_outside
		has_property(agent_B, looks_interested)
		has_property(agent_B, paying_attention)
		Agent B:
		there_is_noise_outside
		has_property(agent_B, looks_interested)
			Agent A:
			observation_result(there_is_noise_outside, pos)
			observation_result(has_property(agent_B, looks_interested), pos)
			belief(there_is_noise_outside, pos)
			belief(has_property(agent_B, looks_interested), pos)
			belief(has_property(agent_B, paying_attention), pos)
			to_be_achieved(at_position(pattern_3, p0), pos)
			to_be_achieved(has_property(agent_A, points_to(p3)), pos)
			to_be_achieved(has_property(agent_A, disgusted_face), pos)
			to_be_performed(pull_disgusted_face)
			to_be_performed(point_to(p3))
			to_be_performed(show(slide_3))
			Agent B:
			observation_result(there_is_noise_outside, pos)
			observation_result(has_property(agent_B, looks_interested), pos)
			belief(there_is_noise_outside, pos)
			belief(has_property(agent_B, looks_interested), pos)
			to_be_achieved(has_property(agent_B, looks_away), pos)
			to_be_achieved(has_property(agent_B, looks_interested), neg)
			to_be_performed(look_away)
			to_be_performed(stop_looking_interested)
			revised to_be_performed(look_interested)
			revised to_be_achieved(has_property(agent_B, looks_interested), pos)

Appendix C: *continued*

	External world		Agents

7	Pattern 3 is at p0.
	Agent A has a disgusted expression.
	Agent A points to p3.
	Agent B is looking away.
	Agent B is not looking interested.
		Agent A:
		not has_property (agent_B, looks_interested)
		not has_property (agent_B, paying_attention)
		has_property(agent_B, looks_away)
		Agent B:
		not has_property (agent_B, looks_interested)
		has_property(agent_B, looks_away)
			Agent A:
			observation_result(at_position(pattern_3, p0), pos)
			observation_result(has_property(agent_A, disgusted_face), pos)
			observation_result(has_property(agent_A, points_to(p3)), pos)
			observation_result(has_property(agent_B, looks_away), pos)
			revised observation_result(has_property(agent_B, looks_interested), pos)
			belief(at_position(pattern_3, p0), pos)
			belief(has_property(agent_A, disgusted_face), pos)
			belief(has_property(agent_A, points_to(p3)), pos)
			belief(has_property(agent_B, looks_away), pos)
			inverted belief(has_property(agent_B, looks_interested), pos)
			inverted belief(has_property(agent_B, paying_attention), pos)
			to_be_achieved(has_property(agent_B, looks_away), neg)
			to_be_achieved(there_is_noise_inside, pos)
			to_be_achieved(has_property(agent_A, disgusted_face), neg)
			to_be_achieved(has_property(agent_A, points_to(p3)), neg)
			to_be_achieved(at_position(pattern_3, p0), neg)
			to_be_performed(do_not_point)
			to_be_performed(pull_no_disgusted_face)
			to_be_performed(remove(slide_3))
			to_be_performed(tap_microphone)
			revised to_be_performed(pull_disgusted_face)
			revised to_be_performed(point_top(p3))
			revised to_be_performed(show(slide_3))
			revised to_be_achieved(at_position(pattern_3, p0), pos)
			revised to_be_achieved(has_property(agent_A, disgusted_face), pos)
			revised to_be_achieved(has_property(agent_A, points_to(p3)), pos)
			Agent B:
			observation_result(has_property(agent_B, looks_away), pos)
			revised observation_result(has_property(agent_B, looks_interested), pos)
			revised observation_result(at_position(agent_A, p1), pos)
			belief(has_property(agent_B, looks_away), pos)
			inverted belief(has_property(agent_B, looks_interested), pos)
			revised to_be_performed(look_away)
			revised to_be_performed(stop_looking_interested)

8	Pattern 3 is not at p0. Agent A has no disgusted expression. Agent A is not pointing to p3. Noise inside. No noise outside.	Agent A: not there_is_noise_outside there_is_noise_inside not has_property (agent_A,disgusted_face) not has_property (agent_A,points_to(p3) not at_position(pattern_3, p0) Agent B: not there_is_noise_outside there_is_noise_inside	Agent A: observation_result(there_is_noise_inside,pos) revised observation_result(has_property(agent_A,points_to(p3)),pos) revised observation_result(has_property(agent_A,disgusted_face),pos) revised observation_result(there_is_noise_outside,pos) revised observation_result(at_position(pattern_3,p0),pos) belief(there_is_noise_inside,pos) inverted belief(there_is_noise_outside,pos) inverted belief(at_position(pattern_3,p0),pos) inverted belief(has_property(agent_A,disgusted_face),pos) inverted belief(has_property(agent_A,points_to(p3)),pos) revised to_be_performed(do_not_point) revised to_be_performed(tap_microphone) revised to_be_performed(remove(slide_3)) revised to_be_performed(pull_no_disgusted_face) Agent B: observation_result(there_is_noise_inside,pos) revised observation_result(there_is_noise_outside,pos) belief(there_is_noise_inside,pos) inverted belief(there_is_noise_outside,,pos) to_be_achieved(has_property(agent_B,looks_away),neg) to_be_achieved(has_property(agent_B,looks_interested),pos) to_be_performed(look_interested) to_be_performed(look_towards_front) revised to_be_achieved(has_property(agent_B,looks_away),pos) revised to_be_achieved(has_property(agent_B,looks_interested),neg)

9	Agent_B is looking interested.
	Agent B is looking towards the front.
		Agent A:
		
		has_property(agent_B, looks_interested)
		has_property(agent_B, paying_attention)
		not has_property (agent_B,looks_away)
			Agent A:
			
			observation_result(has_property(agent_B,looks_interested),pos)
			revised observation_result(has_property(agent_B,looks_away),pos)
			inverted belief(has_property(agent_B,looks_away),pos)
			belief(has_property(agent_B,looks_interested),pos)
			belief(has_property(agent_B,paying_attention),pos)
			to_be_achieved(there_is_noise_inside,neg)
			to_be_achieved(at_position(pattern_3,p0),pos)
			to_be_achieved(has_property(agent_A,disgusted_face),pos)
			to_be_achieved(has_property(agent_A,points_to(p3)),pos)
			to_be_performed(pull_disgusted_face)

Appendix C: *continued*

External world		Agents	
		Agent B:	to_be_performed(stop_tapping)
			to_be_performed(point_to(p3))
		has_property(agent_B,	to_be_performed(show(slide_3))
		looks_interested)	revised to_be_achieved(has_property(agent_B, looks_away), neg)
		not has_property	revised to_be_achieved(at_position(pattern_3, p0), neg)
		(agent_B, looks_away)	revised to_be_achieved(has_property(agent_A, disgusted_face), neg)
			revised to_be_achieved(there_is_noise_inside, pos)
			revised to_be_achieved(has_property(agent_A, points_to(p3)), neg)
			Agent B:
			observation_result(at_position(agent_A, p1), pos)
			observation_result(has_property(agent_B, looks_interested), pos)
			revised observation_result(has_property(agent_B, looks_away), pos)
			inverted belief(has_property(agent_B, looks_away), pos)
			belief(has_property(agent_B, looks_interested), pos)
			revised to_be_achieved(has_property(agent_B, looks_away), neg)
			revised to_be_performed(look_interested)
			revised to_be_performed(look_towards_front)
10/	No noise inside.	Agent A:	Agent A:
11/	Pattern3 is at p0.		
12	Agent A has a	not there_is_noise_inside	observation_result(at_position(pattern_3, p0), pos)
	disgusted expression.	has_property(agent_A,	observation_result(has_property(agent_A, disgusted_face), pos)
	Agent A points to p3.	disgusted_face)	observation_result(has_property(agent_A, points_to(p3)), pos)
		has_property(agent_A,	revised observation_result(there_is_noise_inside, pos)
		points_to(p3)	inverted belief(there_is_noise_inside, pos)
		at_position(pattern_3, p0)	belief(at_position(pattern_3, p0), pos)
			belief(has_property(agent_A, disgusted_face), pos)
		Agent B:	belief(has_property(agent_A, points_to(p3)), pos)
			to_be_achieved(at_position(pattern_3, p0), neg)
		not there_is_noise_inside	to_be_achieved(has_property(agent_A, disgusted_face), neg)
		has_property(agent_A,	to_be_achieved(has_property(agent_A, questioning_face), pos)
		disgusted_face)	to_be_achieved(has_property(agent_A, points_to(p3)), neg)
		has_property(agent_A,	to_be_performed(do_not_point)
		points_to(p3)	to_be_performed(pull_no_disgusted_face)
		at_position(pattern_3, p0)	to_be_performed(pull_questioning_face)

	to_be_performed(remove(slide_3))
contains(pot2, tea)	revised to_be_performed(pull_disgusted_face)
at_position(pot2,p3)	revised to_be_performed(stop_tapping)
not has_property(tea, fresh)	revised to_be_performed(point_to(p3))
	revised to_be_performed(show(slide_3))
	revised to_be_achieved(at_position(pattern_3,p0),pos)
	revised to_be_achieved(has_property(agent_A,disgusted_face),pos)
	revised to_be_achieved(there_is_noise_inside,neg)
	revised to_be_achieved(has_property(agent_A,points_to(p3)),pos)
	Agent B:
	observation_result(at_position(pattern_3,p0),pos)
	observation_result(has_property(agent_A,disgusted_face),pos)
	observation_result(has_property(agent_A,points_to(p3)),pos)
	revised observation_result(there_is_noise_inside,pos)
	inverted belief(there_is_noise_inside,pos)
	belief(at_position(pattern_3,p0),pos)
	belief(has_property(agent_A,disgusted_face),pos)
	belief(has_property(agent_A,points_to(p3)),pos)
	to_be_achieved(has_property(agent_B, looks_interested),pos)
	has_been_communicated_by_modality(has_property(tea, fresh),pos, verbal)
	has_been_communicated_by_modality(contains(pot2, tea),pos, verbal)
	has_been_communicated_by_modality(tastes_good, neg, non_verbal)
	has_been_communicated_by_modality(at_position(pot2,p3),pos, non_verbal)
	has_been_communicated(contains(pot2, tea),pos)
	has_been_communicated(tastes_good, neg)
	has_been_communicated(at_position(pot2,p3),pos)
	has_been_communicated(has_property(tea, fresh), neg)
	belief(at_position(pot2,p3),pos)
	belief(contains(pot2, tea),pos)
	belief(has_property(tea, fresh), neg)

Appendix C: *continued*

	External world		Agents
13	Pattern 3 is not at p0. Agent A has no disgusted expression. Agent A is not pointing to p3. Agent A has a questioning face.	Agent A: not has_property (agent_A, disgusted_face) not has_property (agent_A, points_to(p3) not at_position(pattern_ 3, p0) has_property(agent_A, questioning_face) Agent B: not has_property (agent_A, disgusted_face) not has_property (agent_A, points_to(p3) not at_position(pattern_ 3, p0) has_property(agent_A, questioning_face)	Agent A: observation_result(has_property(agent_A, questioning_face), pos) revised observation_result(has_property(agent_A, points_to(p3)), pos) revised observation_result(has_property(agent_A, disgusted_face), pos) revised observation_result(at_position(pattern_3, p0), pos) inverted belief(at_position(pattern_3, p0), pos) inverted belief(has_property(agent_A, disgusted_face), pos) inverted belief(has_property(agent_A, points_to(p3)), pos) belief(has_property(agent_A, questioning_face), pos) revised to_be_performed(do_not_point) revised to_be_achieved(has_property(agent_A, questioning_face), pos) revised to_be_performed(pull_questioning_face) revised to_be_achieved(at_position(pattern_3, p0), neg) revised to_be_achieved(has_property(agent_A, disgusted_face), neg) revised to_be_performed(remove(slide_3)) revised to_be_achieved(has_property(agent_A, points_to(p3)), neg) revised to_be_performed(pull_no_disgusted_face) Agent B: observation_result(has_property(agent_A, questioning_face), pos) revised observation_result(has_property(agent_A, points_to(p3)), pos) revised observation_result(has_property(agent_A, disgusted_face), pos) revised observation_result(at_position(pattern_3, p0), pos) inverted belief(at_position(pattern_3, p0), pos) inverted belief(has_property(agent_A, disgusted_face), pos) inverted belief(has_property(agent_A, points_to(p3)), pos) belief(has_property(agent_A, questioning_face), pos) to_be_communicated(message_understood, pos) to_be_communicated_by_modality(message_understood, pos, non_verbal) to_be_achieved(has_property(agent_B, thumb_up), pos) to_be_performed(put_thumb_up) revised has_been_communicated(at_position(pot2, p3), pos) revised

			has_been_communicated_by_modality(at_position(pot2,p3),pos,non_verbal) revised has_been_communicated(has_property(tea, fresh),neg) revised has_been_communicated(tastes_good,neg) revised has_been_communicated_by_modality(has_property(tea, fresh),pos,verbal) revised has_been_communicated_by_modality(contains(pot2, tea),pos,verbal) revised has_been_communicated(contains(pot2, tea),pos) revised has_been_communicated_by_modality(tastes_good,neg,non_verbal)

14	Agent B has his thumb up.	Agent A:
		has_property(agent_B, thumb_up)
		Agent B:
		has_property(agent_B, thumb_up)
			Agent A:
			observation_result(has_property(agent_B, thumb_up),pos)
			belief(has_property(agent_B, thumb_up),pos)
			has_been_communicated_by_modality(communication_succeeded,pos, non_verbal)
			has_been_communicated(communication_succeeded,pos)
			to_be_achieved(at_position(agent_A,p2),pos)
			to_be_achieved(has_property(agent_A,questioning_face),neg)
			to_be_performed(pull_no_questioning_face)
			to_be_performed(go_to(p2))
			revised to_be_communicated(at_position(pot2,p3),pos)
			revised to_be_communicated_by_modality(has_property(tea, fresh),pos,verbal)
			revised to_be_communicated_by_modality(tastes_good,neg,non_verbal)
			revised to_be_communicated(has_property(tea, fresh),neg)
			revised to_be_communicated_by_modality(contains(pot2, tea),pos,verbal)
			revised to_be_communicated_by_modality(at_position(pot2,p3),pos,non_verbal)
			revised to_be_communicated(contains(pot2, tea),pos)
			Agent B:
			observation_result(has_property(agent_B, thumb_up),pos)
			belief(has_property(agent_B, thumb_up),pos)
			revised to_be_performed(put_thumb_up)

15	Agent A is at p2.		