An Executable Model of the Interaction between Verbal and Non-Verbal Communication

Catholijn M. Jonker, Jan Treur, Wouter C.A. Wijngaards

Vrije Universiteit Amsterdam
Department of Artificial Intelligence
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
URL: http://www.cs.vu.nl /~{jonker,treur,wouterw} Email: {jonker,treur,wouterw}@cs.vu.nl
Tel: +31 20 44477 {43,63,50 resp.} Fax: +31 20 4447653

Abstract

In this paper an executable generic process model is proposed for combined verbal and non-verbal communication processes and their interaction. The model has been formalised by three-levelled partial temporal models, covering both the material and mental processes and their relations. 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 that adds and modifies content to a verbal communication process, but also provides a protocol for the (control of the) verbal communication process. With respect to the communication protocol stimulus-response behaviour and deliberate behaviour have been modelled and simulated.

1 Introduction

Communication is often modelled exclusively as a process of information exchange, abstracting from the physical realisation of the communication process. This approach is based on the more general perspective that a strict separation between the level of *symbols* (symbolic processes) and the *material* level (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 communication, for example, information on the communication protocol that is followed, or the current stage of the communication process within the protocol. Thus, three semantic or representation levels are introduced:

- the material level (for the physical world),
- the symbolic object level (for symbolic manipulation of information on the physical world), and
- 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 material world, the symbolic meta-level uses representations of the symbolic object level, and the material level uses representations of the symbolic object level. The three levels are explained in more detail in Section 2; process semantics is defined by means of multi-levelled traces based on triples of three-valued states, formalised by partial models (Engelfriet and Treur, 1995).

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. In Section 4 different types of interaction between non-verbal communication and verbal communication processes are identified; 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 model is presented in which the communication protocol of a verbal communication process is modelled as reflex-based non-verbal communication (one gesture triggers another one by stimulus-response behaviour), and an alternative model in which the non-verbal communication in the communication protocol is guided by conscious deliberation. In section 5 the implementation of the model is described. Section 6 summarizes and compares this approach with other work. In Appendix A a trace of the process is given in symbolic notation, demonstrating the stimulus-response case.

2 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.

2.1 A simple example communication process

All examples in this paper are about a 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. 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. 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 realised 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 Figure 1. In this figure 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 figure only the relevant new information elements are represented. The first 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 table gives the same for agent B. The first table 'happens' before the second table.

External World *material level*	*Agent A* *symbolic object level*	 *symbolic meta-level*
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		

External World *material level*	*Agent B* *symbolic object level*	 *symbolic meta-level*
pot 2 contains tea slide 3 at projector pattern3 at p0		to_be_observed(I:INFO_ELEMENT) has_verbal_material_rep(contains(pot2,tea),pos, at_position(pattern3,p0),pos)
		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)

Figure 1. Multi-levelled trace of an example communication process

2.2 Formalisation of the semantic model

In the semantic formalisation a state-based semantics is chosen. Each of the states is formalised 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 and Treur, 1995). Within this approach the semantics of a process is formalised by a set of (alternative) partial temporal models, i.e., sequences of partial models. For an example of such a partial temporal model, see Figure 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 two subsequent 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:

- 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,

- 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)

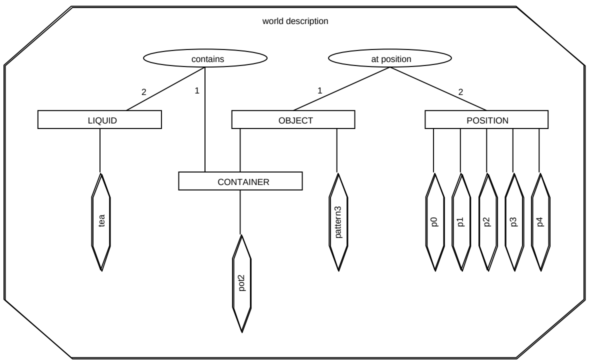
An *information type* Σ is a structure of symbols defining a set of *ground atoms* $\mathrm{At}(\Sigma)$. An *information state* for an information type Σ is a mapping $M : \mathrm{At}(\Sigma) \rightarrow \{0, 1, u\}$ from the set of ground atoms $\mathrm{At}(\Sigma)$ to the set of truth values $\{0,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

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.

Description of material level and 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

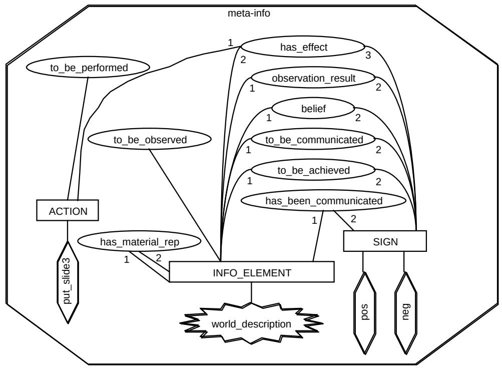


A state of the symbolic object level can be formally described by a three-valued model for the same language. The third truth value is used to express that some fact is not represented (e.g., not known to the agent).

Description of symbolic meta-level

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. Formalising 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.

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 $IS(\Sigma) \times IS(\Sigma)$. A *transition relation* is a set of these transitions, i.e., a relation on $IS(\Sigma) \times IS(\Sigma)$.

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 $IS(\Sigma)$. The set of all partial temporal models is denoted by $IS(\Sigma)^{\mathbf{N}}$, or $Traces(\Sigma)$.

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: $IS = IS(\Sigma^{mat}) \times IS(\Sigma^{obj}) \times IS(\Sigma^{meta})$

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 a *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)

a) The following types of transitions are defined:

- *world change transition* $IS_2(\Sigma^{mat}) \rightarrow IS_2(\Sigma^{mat})$
- *object reasoning transition* $IS(\Sigma^{obj}) \rightarrow IS(\Sigma^{obj})$
- *meta-reasoning transition* $IS(\Sigma^{meta}) \rightarrow IS(\Sigma^{meta})$
- *upward reflection transition* $IS(\Sigma^{obj}) \times IS(\Sigma^{meta}) \rightarrow IS(\Sigma^{meta})$
- *downward reflection transition* $IS(\Sigma^{obj}) \times IS(\Sigma^{meta}) \rightarrow IS(\Sigma^{obj})$
- *action execution transition* $IS_2(\Sigma^{mat}) \times IS(\Sigma^{meta}) \rightarrow IS_2(\Sigma^{mat})$
- *observation transition* $IS_2(\Sigma^{mat}) \times IS(\Sigma^{meta}) \rightarrow IS(\Sigma^{meta})$

b) A *levelled transition* is a transition: $IS \rightarrow IS$ which is based on a transition of one of the types defined in a).

Definition 6 (levelled trace)

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

b) An element $(M^t)_{t \in \mathbf{N}} \in$ 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 Figure 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).

3 A 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
- communication initiation
- action generation
- maintenance of world information

Together with the physical processes in the material world that realise 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 2). The interaction is modelled by information links (the arrows in Figure 2). Agent B is identical in composition to agent A.

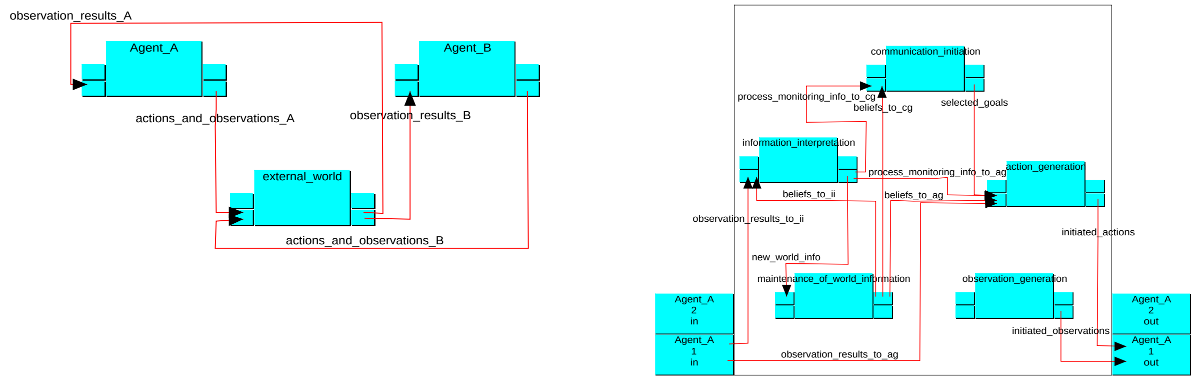


Figure 2. The two highest process abstraction levels

The component observation generation determines the observations the agent initiates. In the example, agent B is initiating observation of the projection screen and agent A, and agent A observes agent B. The component information interpretation produces the information conveyed in the observation results, using specific interpretation knowledge (which involves, among others, knowledge on material representation of information). The component maintain world information constructs and maintains the agent's world model. The component communication initiation makes the decision to communicate and does some preparation of the communication; it uses knowledge about when to communicate information, and the selection of goals in relation to the material representation of information. The component action generation produces the actions to be performed, using knowledge of the effect of actions.

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 reasoning of the agent is specified using a number of generic rules; 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 realises the state change depicted in Figure 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.

4 Combining verbal and non-verbal communication

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 4.1 different types of interaction between verbal and non-verbal communication are distinguished and explained. In Section 4.2 it is shown how differences between reflex-based (also called direct stimulus-response) reactions and conscious reactions on non-verbal communication can be modelled.

4.1 Types of interaction between verbal and non-verbal communication

The example communication process in Section 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 of 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 the verbal communication:
   1. the verbal communication process is affected by reflex-based reactions to the non-verbal communication
   2. the verbal communication process is affec ted 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 recognises him. This 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 recognises 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.

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 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. and 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.

4.2 Modelling a communication protocol: by stimulus-response or by deliberate behaviour

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 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 realises that she needs to attract his attention (as a goal), and decides to tap the microphone (as an action to realise the goal. This is a form of goal directed behaviour based on conscious interpretations.

The generic process model described in Section 3 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.

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 communication initiation 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 **communication initiation**

(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 communication initiation. 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);

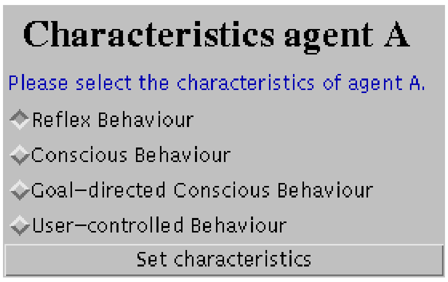
Both agents always have knowledge in their symbolic meta-level information states:

Agent A always has the following knowledge 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_position(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)

5 Implementation

A demo has been developed which allows the user to view the process step-by-step. They can also take matters into their own hand and select the types of behavior for each agent, or even select specific actions by hand, taking the role of one or both agents. In Figure 3 example screen shots from the interactive demo are shown.



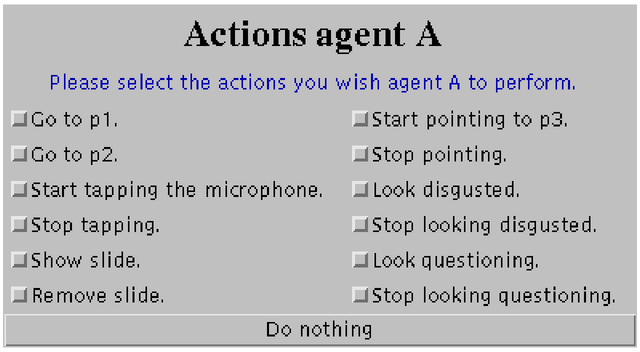


Figure 3. Sample screen shots from the interactive demo.

6 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 formalised on the basis of three-levelled partial temporal models (Treur, 1994; Engelfriet and Treur, 1995). These partial temporal models formalise 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 development method DESIRE (Brazier, Dunin-Keplicz, Jennings, and Treur, 1997), 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.

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

References

Austin, J.L. (1962). How to do things with words. Oxford University Press. London.

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

Bradshaw, J.M. (ed.) (1997). Software Agents. AAAI Press/MIT Press.

Brazier, F.M.T., Dunin-Keplicz, B.M., Jennings, N.R. and Treur, J. (1995) Formal Specification of Multi-Agent Systems: a Real World Case, In: Lesser, V. (Ed.), *Proceedings First International Conference on Multi-Agent Systems, ICMAS'95*, MIT Press, pp. 25-32. Extended version in: Huhns, M. and Singh, M. (Eds.), *International Journal of Co-operative Information Systems, IJCIS* vol. 6 (1), special issue on Formal Methods in Co-operative Information Systems: Multi-Agent Systems, pp. 67-94.

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

Chaib-draa, B., and Levesque, P. (1994) Hierarchical model and communication by signs, signals and symbols in multi-agent environments. In: J.W. Perram and 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, Springer Verlag. Also in: Journal of Experimental and Theoretic Artificial Intelligence, vol. 8, pp. 7-20, 1996.

Cohen, P.R., J. Morgan, M. Pollack (eds.), Intentions in Communication. MIT Press, 1990

Cohen, P.R., Levesque, H.J. (1997). Communicative Actions for Artificial Agents. In: (Bradshaw, 1997), pp. 419-436

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, 1995, pp. 239-261.

Finin, T., Labrou, Y., Mayfield, J., (1997). KQML as an Agent Communication Language. In: (Bradshaw, 1997), pp. 291-316.

Hauser, M.D. (1996). The Evolution of Communication. MIT Press.

Jonker, C.M., and J. Treur (1997). Modelling an Agent's Mind and Matter. In: Proceedings of the 8th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'97, Lecture Notes in AI, vol. 1237, Springer Verlag, pp. 210-233.

Langholm, T. (1988). Partiality, Truth and Persistence. *CSLI Lecture Notes, No. 15*. Stanford University, Stanford.

Perrault, C.R., An Application of Default Logic to Speech Act Theory. In: (Cohen, Morgan, and Pollack, 1990), pp. 161-186

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

Treur, J. (1994), Temporal Semantics of Meta-Level Architectures for Dynamic Control of Reasoning. In: L. Fribourg and F. Turini (ed.), Logic Program Synthesis and Transformation-Meta-Programming in Logic, Proceedings of the Fourth International Workshop on Meta-Programming in Logic, META'94. Springer Verlag, Lecture Notes in Computer Science, vol. 883, 1994. pp. 353-376.

Appendix A A trace following the communication protocol by stimulus-response behaviour

In this process both agents follow the protocol by stimulus-response behaviour.

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

5/6	Agent B looks interested. Noise outside.	*Agent* B: there_is_noise_outside	*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)) *Agent* B: observation_result(there_is_noise_outside, pos) 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: observaton_result(has_property(agent_B, looks_interested), neg) observaton_result(has_property(agent_B, looks_away), pos) to_be_performed(do_not_point) to_be_performed(pull_no_disgusted_face) to_be_performed(remove(slide_3)) to_be_performed(tap_microphone)
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* B: observation_result(there_is_noise_outside, neg) observation_result(there_is_noise_inside, pos) to_be_performed(look_towards_front) to_be_performed(look_interested)
9	Noise inside. Agent_B is looking interested. Agent B is looking towards the front.		*Agent* A: observation_result(has_property(agent_B, looks_away), neg) observation_result(has_property(agent_B, looks_interested), pos) to_be_performed(stop_tapping) to_be_performed(pull_disgusted_face) to_be_performed(point_to(p3)) to_be_performed(show(slide_3))
10/ 11/ 12	No noise inside. Pattern3 is at p0. Agent A has disgusted expression. Agent A points to p3.	*Agent* B: has_property(agent_A, disgusted_face) has_property(agent_A, points_to(p3) at_position(pattern_3, p0)	*Agent* A: observation_result(there_is_noise_inside, neg) observation_result(has_property(agent_A, disgusted_face), pos) observation_result(has_property(agent_A, points_to(p3)), pos) to_be_performed(pull_questioning_face) *Agent* B: observation_result(there_is_noise_inside, neg) 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) 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)
13	Agent A has a questioning face.	*Agent* B: contains(pot2, tea) at_position(pot2, p3) not has_property(tea, fresh)	*Agent* B: observation_result(has_property(agent_A, questioning_face), pos) observation_result(has_property(agent_A, looks_interested), pos) to_be_performed(put_thumb_up)
14	Agent_B has his thumb up.		*Agent* A: observation_result(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(go_to(p2))
15	Agent A is at p2.		