Running head: Collective Representational Content for Shared Extended Mind

Collective Representational Content for Shared Extended Mind

Tibor Bosse

Vrije Universiteit Amsterdam

Catholijn M. Jonker

Radboud Universiteit Nijmegen

Martijn C. Schut

Vrije Universiteit Amsterdam

Jan Treur

Vrije Universiteit Amsterdam

Contact information:
        Vrije Universiteit Amsterdam
        Department of Artificial Intelligence
        Room T3.20a
        De Boelelaan 1081a
        1081 HV Amsterdam
        The Netherlands
        Email: tbosse@cs.vu.nl
        Office: +31 20 598 7750
        Fax: +31 20 598 7653

Abstract

Some types of species exploit the external environment to support their cognitive processes, in the sense of patterns created in the environment that function as external mental states and serve as an extension to their mind. In the case of social species the creation and exploitation of such patterns can be shared, thus obtaining a form of shared mind or collective intelligence. This paper explores this shared extended mind principle for social species in more detail. The focus is on the notion of representational content in such cases. Proposals are put forward and formalised to define collective representational content for such shared external mental states. Two case studies in domains in which shared extended mind plays an important role are used as illustration. The first case study addresses the domain of social ant behaviour. The second case study addresses the domain of human communication via the environment. For both cases simulations are described, representation relations are specified and are verified against the simulated traces.

Collective Representational Content for Shared Extended Mind

1. Introduction

Behaviour is often not only supported by internal mental structures and cognitive processes, but also by processes based on patterns created in the external environment that serve as external mental structures; cf. (Clark, 1997, 2001; Clark and Chalmers, 1998, Dennett, 1996). Such a pattern in the environment is often called an *extended mind*. Examples of extended mind are the use of 'to do lists' and 'lists of desiderata'. Having written these down externally (e.g., on paper, in your diary, in your organizer or computer) makes it unnecessary to have an internal memory about all the items. Thus internal mental processing can be kept less complex. Other examples of the use of extended mind are doing mathematics or arithmetic, where external (symbolic, graphical, material) representations are used; e.g., (Bosse et al., 2003). In (Menary, 2004) a collection of papers can be found based on presentations at the conference 'The Extended Mind: The Very Idea' that took place in 2001. Clark (2001) points at the roles played by both internal and external representations in describing cognitive processes: 'Internal representations will, almost certainly, feature in this story. But so will external representations, …' (Clark, 2001, p. 134). From another, developmental angle, also Griffiths and Stotz (2000) endorse the importance of using both internal and external representations; they speak of 'a larger representational environment which extends beyond the skin', and claim that 'culture makes humans as much as the reverse' (Griffiths and Stotz, 2000, p. 45).

Allowing mental states, which are in the external world and thus accessible for any agent around, opens the possibility that other agents also start to use them. Indeed, not only in the individual, single agent case, but also in the social, multi-agent case the extended mind principle can be observed, e.g., one individual creating a pattern in the environment, and one or more other

individuals taking this pattern into account in their behaviour. For the human case, examples can be found everywhere, varying from roads, and traffic signs to books or other media, and to many other kinds of cultural achievements. Also in (Scheele, 2002) it is claimed that part of the total team knowledge in distributed tasks (such as air traffic control) comprises external memory in the form of artefacts. In this multi-agent case the extended mind principle serves as a way to build a form of social or collective intelligence, that goes beyond (and may even not require) social intelligence based on direct one-to-one communication.

Especially in the case of social species external mental states created by one individual can be exploited by another individual, or, more general, the creation and maintenance, as well as the exploitation of external mental states can be activities in which a number of individuals participate. For example, presenting slides on a paper with multiple authors to an audience. In such cases the external mental states cross, and in a sense break up, the borders between the individuals and become *shared extended mental states*. Another interesting and currently often studied example of collective intelligence is the intelligence shown by *stigmergy*. Stigmergy was defined originally as the indirect communication taking place among individuals in social insect societies (e.g., ant colonies), see (Grassé, 1959; Bonabeau, 1999; Bonabeau et al., 1999). Indeed, in this case the external world is exploited as an extended mind by using pheromones. While they walk, ants drop pheromones on the ground. The same or other ants sense these pheromones and follow the route in the direction of the strongest sensing. Pheromones are not persistent for long times; therefore such routes can vary over time. Currently, in the domain of computer science, the notion of stigmergy is used to solve many complex problems, e.g. concerning optimisation, coordination, or self-organisation.

In the literature on Philosophy of Mind, there is an ongoing discussion about the exact definitions of 'mind' and 'shared extended mind' (e.g., Clark and Chalmers, 1998; Tollefsen, 2006). Although none of these authors provides a complete definition, a number of criteria for shared extended mind are commonly accepted:

- The environment participates in the agents' mental processes

- The agents' internal mental processes are simplified

- The agents have a more intensive interaction with the world

- The agents depend on the external world in the sense that they delegate some of their mental representations and capabilities to it

To this discussion, we want to add two novel questions. A first question is whether an agents' explicit intention to create the shared extended mind is a necessary requirement. As opposed to the mainstream view in the field, in the present paper this requirement is dropped, i.e., the one(s) 'creating' the shared extended mind do(es) not need to be aware of this. This means that agents with limited internal cognitive processes can nevertheless contribute to the emergence of a complex structure that can be described as a 'mind'. For example, we consider the pheromone mechanism used by ants for foraging similar to other common examples of the extended mind (computer, notepad, and so on). See Section 8 for an elaborate discussion on this topic. A second question with respect to the definition of shared extended mind is whether the mind needs to be useful for the agents that create it. Also this criterion is not considered necessary in the current paper. This means that we also allow cases where the shared extended mind may be disadvantageous for the agent that creates it. For example, in case of a predator-prey relationship, the traces that the prey leaves in the environment may be seen as a shared extended mind for the predators: they give information about the location of the prey, although this is completely

against the preys interest. Tackling these kinds of examples may contribute to a more precise definition of shared extended mind. A possible approach in this respect is to define a classification of different categories of shared extended mind. This option will be explored in future work.

In (Bosse et al., 2005) the shared extended mind principle is worked out in more detail. The paper focusses on formal analysis and formalisation of the dynamic properties of the processes involved, both at the local level (the basic mechanisms) and the global level (the emerging properties of the whole), and their relationships. A case study in social ant behaviour in which shared extended mind plays an important role is used as illustration. In the current paper, as an extension to (Bosse et al., 2005), the notion of *representational content* is analysed for mental processes based on the shared extended mind principle. The analysis of notions of representational content of *internal* mental state properties is well-known in the literature on Cognitive Science and Philosophy of Mind. In a nutshell, the question in this literature is 'what does it mean for an agent to have a mental state', or 'what information does the mental state represent'? Usually this question is answered by taking a relevant internal mental state property m and identifying a *representation relation* that indicates in which way m relates to properties in the external world or the agent's interaction with the external world; cf. (Bickhard, 1993; Jacob, 1997, Kim, 1996, pp. 184-210). For the case of extended mind an extension of the analysis of notions of representational content to *external* state properties is needed. Moreover, for the case of external mental state properties that are *shared*, a notion of *collective* representational content is needed (in contrast to a notion of representational content for a single agent). As a result, the question to be answered then becomes 'what information does a shared extended mental state

(e.g. a heap of pheromones) represent for the group'? This is one of the main questions to be answered in this paper.

Thus, by addressing examples such as ant colonies and modelling them from an extended mind perspective, a number of challenging new issues on cognitive modelling and representational content are encountered:

- How to define representational content for an *external* mental state property

- How to handle *decay* of a mental state property

- How can *joint* creation of a *shared* mental state property be modelled

- What is an appropriate notion of *collective* representational content of a shared external mental state property

- How can representational content be defined in a case where a behavioural choice depends on *a number of mental state properties*

In this paper these questions are addressed. To this end the shared extended mind principle is analysed in more detail, and a formalisation is provided of its dynamics. It is discussed in particular how a notion of collective representational content for a shared external mental state property can be formulated. In the literature notions of representational content are usually restricted to internal mental states of one individual. The notion of collective representational content developed here extends this in two manners: (1) for external instead of internal mental states, and (2) for groups of individuals instead of single individuals. The proposals put forward are evaluated in two case studies of social behaviour based on shared extended mind. First, as an example of an *unintentionally* created shared extended mind (by species with limited cognitive capabilities), a case study of a simple ant colony is addressed. Next, as an example of an *intentionally* created shared extended mind (by species with more complex cognitive

capabilities), a case study is addressed involving a person that presents slides to an audience. The analysis of these case studies comprises multi-agent simulation based on identified local dynamic properties, identification of dynamic properties that describe collective representational content of shared extended mind states, and verification of these dynamic properties.

2. State Properties and Dynamic Properties

Dynamics will be described in the next section as evolution of states over time. The notion of state as used here is characterised on the basis of an ontology defining a set of physical and/or mental (state) properties that do or do not hold at a certain point in time. For example, the internal state property 'the agent A has pain', or the external world state property 'the environmental temperature is $7°$ C', may be expressed in terms of different ontologies. To formalise state property descriptions, an ontology is specified as a finite set of sorts, constants within these sorts, and relations and functions over these sorts. The example properties mentioned above then can be defined by nullary predicates (or proposition symbols) such as pain, or by using n-ary predicates (with n≥1) like has_temperature(environment, 7). For a given ontology Ont, the propositional language signature consisting of all *state ground atoms* (or *atomic state properties*) based on Ont is denoted by APROP(Ont). The *state properties* based on a certain ontology Ont are formalised by the propositions that can be made (using conjunction, negation, disjunction, implication) from the ground atoms. A *state* S is an indication of which atomic state properties are true and which are false, i.e., a mapping S: APROP(Ont) → {true, false}.

To describe the internal and external dynamics of the agent, explicit reference is made to time. Dynamic properties can be formulated that relate a state at one point in time to a state at another point in time. A simple example is the following dynamic property specification for belief creation based on observation:

'at any point in time t1 if the agent observes at t1 that it is raining,

then there exists a point in time t2 after t1 such that at t2 the agent believes that it is raining'.

To express such dynamic properties, and other, more sophisticated ones, the temporal trace language TTL is used; cf. (Jonker et al., 2003). To express dynamic properties in a precise manner a language is used in which explicit references can be made to time points and traces. Here trace or trajectory over an ontology Ont is a time-indexed sequence of states over Ont. The sorted predicate logic temporal trace language TTL is built on atoms referring to, e.g., traces, time and state properties. For example, 'in the output state of A in trace γ at time t property p holds' is formalised by state(γ, t, output(A)) |= p. Here |= is a predicate symbol in the language, usually used in infix notation, which is comparable to the Holds-predicate in situation calculus. Dynamic properties are expressed by temporal statements built using the usual logical connectives and quantification (for example, over traces, time and state properties). For example the following dynamic property is expressed:

'in any trace γ, if at any point in time t1 the agent A observes that it is raining,

then there exists a point in time t2 after t1 such that at t2 in the trace the agent A believes that it is raining'.

In formalised form:

$\forall$t1   [state(γ, t1, input(A)) |= agent_observes_itsraining $\Rightarrow$

$\exists$t2 $\geq$ t1   state(γ, t2, internal(A)) |= belief_itsraining   ]

Language abstractions by introducing new (definable) predicates for complex expressions are possible and supported.

A simpler temporal language has been used to specify simulation models. This language (the LEADSTO language) offers the possibility to model direct temporal dependencies between two state properties in successive states. This executable format is defined as follows. Let α and β be state properties of the form 'conjunction of atoms or negations of atoms', and e, f, g, h non-negative real numbers. In the LEADSTO language $\alpha \twoheadrightarrow_{e, f, g, h} \beta$, means:

If   state property α *holds for a certain time interval with duration g,*

then after some delay (between e and f) state property β *will hold for a certain time interval of length h.*

For a precise definition of the LEADSTO format in terms of the language TTL, see (Jonker et al., 2003). A specification of dynamic properties in LEADSTO format has as advantages that it is executable and that it can often easily be depicted graphically.

3. Representation for Shared Extended Mind

Originally, in the literature on Cognitive Science and Philosophy of Mind, the concept of representational content is applicable to internal (mental) states of agents (Bickhard, 1993; Jacob, 1997; Jonker and Treur, 2003; Kim, 1996, pp. 191-193, 200-202). As mentioned earlier, the common idea is that the occurrence of the internal (mental) state property m at a specific point in time is related (by a representation relation) to the occurrence of other state properties, at the same or at different time points. Such a representation relation then describes in a precise manner what the internal state property m represents. To define a representation relation, the *causal-correlational* approach is often discussed in the literature in Philosophy of Mind. However, this approach has a number of severe limitations and problems (e.g., the conjunction or transitivity problem, the disjunction problem, and the dynamics problem); cf. Kim (1996), Jacob (1997). Two approaches that are considered to be more promising are the *interactivist* approach (Bickhard, 1993; Jonker and Treur, 2003) and the *relational specification* approach (Kim, 1996). As the causal-correlational approach is too limited for the case addressed here, this paper will concentrate on the latter two approaches. For the interactivist approach, a representation relation relates the occurrence of an internal state property to sets of past and future interaction traces. The relational specification approach to representational content is based on a specification of how a representation relation relates the occurrence of an internal state property to properties of states distant in space and time; cf. (Kim, 1996, pp. 200-202). As mentioned in the Introduction,

one of the goals of this paper is to apply these approaches to *shared extended* mental states instead of *internal* mental states of a *single* agent. Thus, it will be explored for shared extended mental states (such as 'a heap of pheromones' or 'a slide on an overhead projector') what information they represent for a group of agents'.

Suppose p is an external state property used by a collection of agents in their shared extended mind, for example, as an external belief. At a certain point in time this mental state property was created by performing an action a_1 (or maybe a collection of actions) by one or more agents to bring about p in the external world. This situation is depicted schematically in Figure 1. Here, the circles indicate state properties, the arrows indicate causal temporal relationships, and the dotted rectangles indicate (different) agents[1]. As can be seen in the figure, the chain of events can be followed further back, from action a_1 to internal mental state m_1, then to observation o_1, and finally to external world state q. Likewise, the chain of events can be followed in the direction of the future. Thus, given the created occurrence of p, at a later point in time any agent can observe this external state property (by observation o_2) and take it into account in determining its behaviour. Subsequently, this observation of p may lead to internal mental state m_2, then to action a_2, and finally to external world state r. For a representation relation, which indicates representational content for such a mental state property p several possibilities are considered:

- a representation relation relating the occurrence of p to one or more events in the past (*backward*)
- a representation relation relating the occurrence of p to behaviour in the future (*forward*)

Moreover, for each category, the representation relation can be described by referring to:

- external world state properties (e.g., using the *relational specification* approach)
- observation state properties for the agent (e.g., using the *interactivist* approach)

- internal mental state properties for the agent (e.g., using the *relational specification* approach)

- action state properties for the agent (e.g., using the *interactivist* approach)

So, eight types of approaches (2x4) to representational content are distinguished. The different options are illustrated by Figure 2 (backward case) and Figure 3 (forward case). For example, Figure 2a gives an example of a *backward* representation relation following the *relational specification* approach. Here, the relation is backward because the presence of p is related only to events in the past, and it is according to the relational specification approach because it involves only external world properties. In the next section it is shown how the different approaches can be applied in a concrete case study.

In principle, to define the representational content of a (shared extended) mental state in a precise manner, a combination of a backward and a forward representation relation can be used (i.e., combining one of the pictures in Figure 2 with one of the pictures in Figure 3). However, throughout this paper, the backward and forward case will be treated separately.

4. Ants Case Study

In this section, the idea of collective representational content will be illustrated first for species with limited cognitive processes. This is done by means of a case study in the domain of ants.

To facilitate understanding, two separate variants of the case study are distinguished. This distinction depends on the nature of extended mental state property p:

- *the qualitative case*. Here p may be the result of the action of one agent (e.g., p is 'the presence of pheromone'). Therefore, it has a binary nature: it is either true or false.

- *the quantitative case*. Here p may be the result of actions of multiple agents. Here p has a certain degree or level (e.g., p is 'the presence of a certain accumulated level of pheromone'); in decisions levels for a number of such state properties p may be taken into account

First, in Section 4.1, a domain description for the case study is provided. Next Section 4.2 addresses the qualitative case, and Section 4.3 addresses the quantitative case. For each case a number of the different types of representation relations in Figure 2 and 3 will be shown.

4.1. Domain Description

For the ants case study, the world in which the ants live is described by a labeled graph as depicted in Figure 4. Locations are indicated by A, B,…, and edges by e1, e2,… The ants move from location to location via edges; while passing an edge, pheromones are dropped. The objective of the ants is to find food and bring this back to the nest. In this example there is only one nest (at location A) and one food source (at location F).

The example concerns multiple agents (the ants), each of which has input (to observe) and output (for moving and dropping pheromones) states, and a physical body which is at certain positions over time, but no internal mental state properties (they are assumed to act purely by stimulus-response behaviour). An overview of the formalisation of the state properties of this example is shown in Table 1. In these state properties, a is a variable that stands for ant, l for location, e for edge, and i for pheromone level. Note that in some of the state properties the direction of an ant is incorporated (e.g., ant a is at location l coming from e, ant a is at edge e to l2 coming from location l1). This direction is meant to relate to the orientation of the ant's body in space, which is a genuine state property; but for convenience this is expressed by referring to the past or future states involved.

In the next sections, it will be explored for a number of the different types of representation relations shown in Figure 2 and 3 how they work out. This will be done first for the qualitative case (Section 4.2) and then for the more complicated quantitative case (Section 4.3). Although in theory eight different representation relations can be specified for each case, only half of them are worked out in detail. In particular, for each case we address one backward relation according to the interactivist approach, one backward relation according to the relational specification approach, one forward relation according to the interactivist approach, and one forward relation according to the relational specification approach (see Table 2 for an overview). The other combinations can be modelled in a similar manner.

4.2. The Qualitative Case

In this section representational content is addressed for the qualitative case. This means that an external state property p (e.g., the presence of pheromone) is the result of the action of one agent (e.g., dropping the pheromone).

4.2.1. Backward Interactivist Approach

Looking backward, for the qualitative case the preceding state is the action a_1 by an arbitrary agent, to bring about p (see Figure 1). This action a_1 is an interaction state property of the agent. Thus, for the interactivist approach a representation relation can be specified by temporal relationships between p (the presence of the pheromone at a certain edge), and a_1 (the action of dropping this pheromone). In an informal notation, this representation relation looks as follows:

If   at some time point in the past an agent dropped pheromone at edge e,
then after that time point the pheromone was present at edge e.

If   the pheromone is present at edge e,
then at some time point in the past an agent dropped it at e.

Although this relation would qualify as a correct representation relation according to the interactivist approach (see Figure 2d), it is rather trivial (almost tautological), and therefore not very informative. To obtain a more informative notion of representational content, the chain of processes leading to the interaction state property can be followed further back. In fact, one step back, the action of dropping pheromone at an edge was performed because the agent observed that it was present at that edge (assuming that the ants perform stimulus-response behaviour without involvement of complex internal states). Such observations are also interaction states. Thus, for the interactivist approach another (more informative) representation relation can be specified by temporal relationships between p (the presence of the pheromone at a certain edge), and o1 (the observation of being present at this edge). In an informal notation, this representation relation looks as follows:

> If   at some time point in the past an agent observed that it was present at edge e,
> then after that time point the pheromone was present at edge e.

> If   the pheromone is present at edge e,
> then at some time point in the past an agent observed that it was present at e.

Note that this situation corresponds to the example depicted in Figure 2b: the representation relation relates the external world state property to an observation state property in the past. A formalisation is as follows:

> $\forall$t1 $\forall$l $\forall$l1 $\forall$e $\forall$a [state(γ, t1) |= observes(a, is_at_edge_from_to(e, l, l1))
> $\Rightarrow$ $\exists$t2>t1 state(γ, t2) |= pheromone_at(e)]

> $\forall$t2 $\forall$e [state(γ, t2) |= pheromone_at(e)
> $\Rightarrow$ $\exists$a, l, l1, t1<t2 state(γ, t1) |= observes(a, is_at_edge_from_to(e, l, l1))]

Note here that the sharing of the external mental state property is expressed by using explicit agent names in the language and quantification over (multiple) agents (using variable a). In the

'traditional' case of a representation relation for a (non-shared) extended mind of a single agent, no explicit reference to the agent itself would be needed.

4.2.2. Backward Relational Specification Approach

As mentioned above, the action of dropping pheromone can be related to the agent's observations for being at a certain edge. However, these observations concern observations of certain state properties of the external world. Therefore, the chain of processes in history can be followed one step further, arriving eventually at other external world state properties. These external world state properties will be used for the representation relation conform the relational specification approach. It may be clear that if complex internal processes come between, such a representation relation can become complicated. However, if the complexity of the agent's internal processes is kept relatively simple (as is one of the claims accompanying the extended mind principle), this amounts in a feasible approach.

For the relational specification approach a representation relation can be specified by temporal relationships between the presence of the pheromone (at a certain edge), and other state properties in the past or future. Although the relational specification approach as such does not explicitly exclude the use of state properties related to input and output of the agent, in our approach below the state properties will be limited to external world state properties. As the mental state property itself also is an external world state property, this implies that temporal relationships are provided only between external world state properties. The pheromone being present at edge e is temporally related to the existence of a state at some time point in the past, namely an agent's presence at e:

If   at some time point in the past an agent was present at e,
then after that time point the pheromone was present at edge e.

If   the pheromone is present at edge e,

then at some time point in the past an agent was present at e.

This situation corresponds to the example depicted in Figure 2a: the representation relation relates the external world state property to another external world state property in the past. A formalisation is as follows:

∀t1 ∀l ∀l1 ∀e ∀a [state(γ, t1) |= is_at_edge_from_to(a, e, l, l1)
        ⇒ ∃t2>t1 state(γ, t2) |= pheromone_at(e)]

∀t2 ∀e [state(γ, t2) |= pheromone_at(e)
        ⇒ ∃a, l, l1, t1<t2 state(γ, t1) |= is_at_edge_from_to(a, e, l, l1)]

4.2.3. Forward Interactivist Approach

Looking forward, in general the first step is to relate the extended mind state property p to the observation o2 of it by an agent (under certain circumstances c). However, again the chain of processes can be followed further (possibly through this agent's internal processes) to the agent's actions (for the interactivist approach) and their effects on the external world (for the relational specification approach).

For the example, an agent's action based on its observation of the pheromone is that it heads for the direction of the pheromone. So, according to the interactivist approach, the representation relation relates the occurrence of the pheromone (at edge e) to the conditional (with condition that it observes the location) fact that the agent heads for the direction of e. The pheromone being present at edge e is temporally related to a conditional statement about the future, namely if an agent later observes the location, coming from any direction e', then he will head for direction e:

If   the pheromone is present at edge e1,
then   if at some time point in the future, an agent observes a location l, connected to e1,
             coming from any direction e2≠e1,
             then the next direction he will choose is e1.

If   a time point t1 exist such that

	at t1 an agent observes a location l (connected to e1), coming from any direction e2 ≠e1,
and	if at any time point t2 ≥ t1 an agent observes this location l coming from any direction e3≠e1,
	then the next direction he will choose is e1,
then	at t1 the pheromone is present at direction e1.

This situation corresponds to the example depicted in Figure 3d: the representation relation relates the external world state property to an action state property in the future. A formalisation is as follows:

∀t1 ∀l ∀l1 ∀e1 [state(γ, t1) |= pheromone_at(e1) ⇒
  ∀t2>t1 ∀e2, a
    [e2 ≠ e1 & state(γ, t2) |= connected_to_via(l, l1, e1) &
     state(γ, t2) |= observes(a, is_at_location_from(l, e2)) ⇒
       ∃t3>t2 state(γ, t3) |= to_be_performed(a, go_to_edge_from_to(e1, l, l1)) &
         [∀t4 t2<t4<t3 ⇒ observes(a, is_at_location_from(l, e2))]]]

∀t1 ∀l ∀e1 [∃a, e2   e2 ≠ e1 &
  state(γ, t1) |= observes(a, is_at_location_from(l, e2)) &
    [∀t2≥t1 ∀a, e3 [e3 ≠ e1 & state(γ, t2) |= observes(a, is_at_location_from(l, e3)) ⇒
      ∃t3>t2 ∃l1 state(γ, t3) |= to_be_performed(a, go_to_edge_from_to(e1, l, l1)) &
        [∀t4 t2<t4<t3 ⇒ observes(a, is_at_location_from(l, e3))]]]
          ⇒ state(γ, t1) |= pheromone_at(e1)]

4.2.4. Forward Relational Specification Approach

The effect of an agent's action based on its observation of the pheromone is that it is at the direction of the pheromone. So, according to the relational specification approach the representation relation relates the occurrence of the pheromone (at edge e) to the conditional (with condition that it is at the location) fact that the agent arrives at edge e. The pheromone being present at edge e is temporally related to a conditional statement about the future, namely if an agent arrives at the location, coming from any direction e', then later he will be at edge e:

If	the pheromone is present at edge e1,
then	if at some time point in the future, an agent arrives at a location l, connected to e1,

coming from any direction e2≠e1,

    then the next edge he will be at is e1.

If   a time point t1 exist such that

    at t1 an agent arrives at a location l (connected to e1), coming from any direction e2 ≠e1,

 and   if at any time point t2 ≥ t1 an agent arrives at this location l coming from any direction e3≠e1,

    then the next edge he will be at is e1,

then   at t1 the pheromone is present at direction e1.

This situation corresponds to the example depicted in Figure 3a: the representation relation relates the external world state property to another external world state property in the future. A formalisation is as follows:

$\forall$t1 $\forall$l $\forall$l1 $\forall$e1 [state(γ, t1) |= pheromone_at(e1) $\Rightarrow$

 $\forall$t2>t1 $\forall$e2, a

  [e2 $\neq$ e1 & state(γ, t2) |= connected_to_via(l, l1, e1) &

   state(γ, t2) |= is_at_location_from(a, l, e2) $\Rightarrow$

    $\exists$t3>t2 state(γ, t3) |= is_at_edge_from_to(a, e1, l, l1) &

     [$\forall$t4 t2<t4<t3$\Rightarrow$ is_at_location_from(a, l, e2)]]]

$\forall$t1 $\forall$l $\forall$e1 [$\exists$a, e2 e2 $\neq$ e1 &

 state(γ, t1) |= is_at_location_from(a, l, e2) &

  [$\forall$t2$\geq$t1 $\forall$a, e3 [e3 $\neq$ e1 & state(γ, t2) |= is_at_location_from(a, l, e3) $\Rightarrow$

   $\exists$t3>t2 $\exists$l1 state(γ, t3) |= is_at_edge_from_to(a, e1, l, l1) &

    [$\forall$t4 t2<t4<t3 $\Rightarrow$ is_at_location_from(a, l, e3)]]]

    $\Rightarrow$ state(γ, t1) |= pheromone_at(e1)]

4.3. The Quantitative Case

The quantitative, accumulating case allows us to consider certain levels of a mental state property p; in this case a mental state property is involved that is parameterised by a number: it has the form p(r), where r is a number, denoting that p has level r. This differs from the above in that now the following aspects have to be modeled: (1) joint creation of p: multiple agents together bring about a certain level of p, each contributing a part of the level, (2) by decay, levels

may decrease over time, (3) behaviour may be based on a number of state properties with different levels, taking into account their relative values, e.g., by determining the highest level of them. For the ants example, for each choice point multiple directions are possible, each with a different pheromone level; the choice is made for the direction with the highest pheromone level (ignoring the direction the ant just came from).

4.3.1. Backward Interactivist Approach

To address the backward quantitative case (i.e., the case of joint creation of a mental state property), the representation relation is analogous to the one described in Section 4.2, but now involves not the presence of one agent at one past time point, but a summation over multiple agents at different time points. Moreover a decay rate r with $0 < r < 1$ is used to indicate that after each time unit only a fraction r is left.

For the ants example in mathematical terms the following property is expressed (according to the interactivist approach, Figure 2b):

There is an amount v of pheromone at edge e, if and only if there is a history such that at time point 0 there was $ph(0, e)$ pheromone at e, and for each time point k from 0 to t a number $dr(k, e)$ of ants observed being present at e, and $v = ph(0, e) * r^t + \Sigma_{k=0}^{t} dr(t-k, e) * r^k$

A formalisation of this property in the logical language TTL is as follows:

$\forall t\ \forall e\ \forall v\ \ state(\gamma, t) \models pheromones\_at(e, v) \Leftrightarrow$
  $\Sigma_{k=0}^{t} \Sigma_{a=ant1}^{ants} case([\ \exists l, l1\ state(\gamma, k) \models$
    $observes(a, is\_at\_edge\_from\_to(e, l, l1))], 1, 0) * r^{t-k} = v$

Here for any formula f, the expression $case(f, v1, v2)$ indicates the value $v1$ if f is true, and $v2$ otherwise.

4.3.2. Backward Relational Specification Approach

Using the relational specification approach, the only difference is that the ants' observations of being present at the edge are replaced by their presence at the edge (see Figure 2a):

There is an amount v of pheromone at edge e, if and only if there is a history such that at time point 0 there was ph(0, e) pheromone at e, and for each time point k from 0 to t a number dr(k, e) of ants was present at e, and v = ph(0, e) * r^t + $\Sigma_{k=0}^{t}$ dr(t-k, e) *r^k

A formalisation of this property in the logical language TTL is as follows:

$\forall t\ \forall e\ \forall v$ state(γ, t) |= pheromones_at(e, v) $\Leftrightarrow$
  $\Sigma_{k=0}^{t}\ \Sigma_{a=ant1}^{ants}$ case([$\exists l,l1$ state(γ, k) |=
    is_at_edge_from_to(a, e, l, l1)], 1, 0) * r^{t-k} = v

4.3.3. Forward Interactivist Approach

The forward quantitative case involves a behavioural choice that depends on the relative levels of multiple mental state properties. This makes that at each choice point the representational content of the level of one mental state property is not independent of the level of the other mental state properties involved at the same choice point. Therefore it is only possible to provide representational content for the combined mental state property involving all mental state properties involved in the behavioural choice.

For the ants example the following property is specified according to the interactivist approach (see Figure 3d):

If at time t1 the amount of pheromone at edge e1 (connected to location 1) is maximal
with respect to the amount of pheromone at all other edges connected to that location 1,
except the edge that brought the ant to the location,
   then, if an ant observes that location l at time t1,
        then the next direction the ant will choose at some time t2 > t1 is e1.

If at time t1 an ant observes location 1 and
     for every ant observing that location 1 at time t1,
       the next direction it will choose at some time t2 > t1 is e1,
   then the amount of pheromone at edge e1 is maximal with respect to the amount of pheromone
   at all other edges connected to that location l, except the edge that brought the ant to the location.

A formalisation of this property in TTL is as follows:

$\forall$t1,l,l1,e1,e2,i1
    [e1≠e2 &

state(γ, t1) |= connected_to_via(l, l1, e1) &

state(γ, t1) |= pheromones_at(e1, i1) &

[∀l2≠l1, e3≠e2 [state(γ, t1) |= connected_to_via(l, l2, e3) ⇒

    ∃i2 [0≤i2<i1 & state(γ, t1) |= pheromones_at(e3, i2)]]

⇒ ∀a [state(γ, t1) |= observes(a, is_at_location_from(l, e2)) ⇒

    ∃t2>t1 state(γ, t2) |= to_be_performed(a, go_to_edge_from_to(e1, l, l1)) &

    [∀t3 t1<t3<t2 ⇒ observes(a, is_at_location_from(l, e2))]]]]

∀t1, l,l1,e1,e2

    [e1≠e2 &

    state(γ, t1) |= connected_to_via(l, l1, e1) &

    ∃a state(γ, t1) |= observes(a, is_at_location_from(l, e2)) &

    ∀a [state(γ, t1) |= observes(a, is_at_location_from(l, e2)) ⇒

        ∃t2>t1 state(γ, t2) |= to_be_performed(a, go_to_edge_from_to(e1, l, l1)) &

        [∀t3 t1<t3<t2 ⇒ observes(a, is_at_location_from(l, e2))]]]

    ⇒ ∃i1 [state(γ, t1) |= pheromones_at(e1, i1) &

        [∀l2≠l1, e3≠e2 [state(γ, t1) |= connected_to_via(l, l2, e3)

            ⇒ ∃i2 [0≤i2≤i1 & state(γ, t1) |= pheromones_at(e3, i2)]]]]

4.3.4. Forward Relational Specification Approach

Likewise, according to the relational specification approach the following property is specified (see Figure 3a):

If at time t1 the amount of pheromone at edge e1 (connected to location 1) is maximal

with respect to the amount of pheromone at all other edges connected to that location l,

except the edge that brought the ant to the location,

    then, if an ant is at that location l at time t1,

        then the next edge the ant will be at some time t2 > t1 is e1.

If at time t1 an ant is at location 1 and

    for every ant arriving at that location 1 at time t1,

      the next edge it will be at some time t2 > t1 is e1,

then the amount of pheromone at edge e1 is maximal with respect to the amount of pheromone

at all other edges connected to that location l, except the edge that brought the ant to the location.

A formalisation of this property in TTL is as follows:

∀t1,l,l1,e1,e2,i1

    [e1≠e2 &

    state(γ, t1) |= connected_to_via(l, l1, e1) &

    state(γ, t1) |= pheromones_at(e1, i1) &

    [∀l2≠l1, e3≠e2 [state(γ, t1) |= connected_to_via(l, l2, e3) ⇒

        ∃i2 [0≤i2<i1 & state(γ, t1) |= pheromones_at(e3, i2)]]

    ⇒ ∀a [state(γ, t1) |= is_at_location_from(a, l, e2) ⇒

        ∃t2>t1 state(γ, t2) |= is_at_edge_from_to(a, e1, l, l1) &

        [∀t3 t1<t3<t2 ⇒ is_at_location_from(a, l, e2)]]]]

∀t1, l,l1,e1,e2

    [e1≠e2 &

    state(γ, t1) |= connected_to_via(l, l1, e1) &

    ∃a   state(γ, t1) |= is_at_location_from(a, l, e2) &

    ∀a [state(γ, t1) |= is_at_location_from(a, l, e2)   ⇒

        ∃t2>t1 state(γ, t2) |= is_at_edge_from_to(a, e1, l, l1) &

        [∀t3 t1<t3<t2 ⇒ is_at_location_from(a, l, e2)]]

    ⇒ ∃i1 [state(γ, t1) |= pheromones_at(e1, i1) &

      [∀l2≠l1, e3≠e2 [state(γ, t1) |= connected_to_via(l, l2, e3)

        ⇒ ∃i2 [0≤i2≤i1 & state(γ, t1) |= pheromones_at(e3, i2)]]]]

5. Simulation and Verification - Ants

5.1. A Simulation Model of the Ants Scenario

In (Bosse et al., 2005) a simulation model of an ant society is specified in which shared extended mind plays an important role. This model is based on local dynamic properties, expressing the basic mechanisms of the process. In this section, a selection of these local properties is presented, and a resulting simulation trace is shown. In the next section it will be explained how the representation relations specified earlier can be verified against such simulation traces. Again, a is a variable that stands for ant, l for location, e for edge, and i for pheromone level.

LP5 (Selection of Edge)

This property models (part of) the edge selection mechanism of the ants. It expresses that, when an ant observes that it is at location l, and there are two edges connected to that location, then the ant goes to the edge with the highest amount of pheromones. Formalisation:

observes(a, is_at_location_from(l, e0)) and neighbours(l, 3) and connected_to_via(l, l1, e1) and observes(a, pheromones_at(e1, i1)) and connected_to_via(l, l2, e2) and observes(a, pheromones_at(e2, i2)) and e0 ≠ e1 and e0 ≠ e2 and e1 ≠ e2 and i1 > i2 •→» to_be_performed(a, go_to_edge_from_to(e1, l1))

Note that this property represents simple stimulus-response behaviour: observations in the external world directly lead to actions. In case an ant arrives at a location where there are two edges with an equal amount of pheromones, its selection is based on the attractive_direction_at[2] predicate (see also the complete set of local properties in Appendix A).

LP9 (Dropping of Pheromones)

This property expresses that, if an ant observes that it is at an edge e from a location l to a location l1, then it will drop pheromones at this edge e. Formalisation:

observes(a, is_at_edge_from_to(e, l, l1)) •→» to_be_performed(a, drop_pheromones_at_edge_from(e, l))

LP13 (Increment of Pheromones)

This property models (part of) the increment of the number of pheromones at an edge as a result of ants dropping pheromones. It expresses that, if an ant drops pheromones at edge e, and no other ants drop pheromones at this edge, then the new number of pheromones at e becomes i*decay+incr. Here, i is the old number of pheromones, decay is the decay factor, and incr is the amount of pheromones dropped. Formalisation:

to_be_performed(a1, drop_pheromones_at_edge_from(e, l1)) and ∀l2 not to_be_performed(a2, drop_pheromones_at_edge_from(e, l2)) and ∀l3 not to_be_performed(a3, drop_pheromones_at_edge_from(e, l3)) and a1 ≠ a2 and a1 ≠ a3 and a2 ≠ a3 and pheromones_at(e, i) •→» pheromones_at(e, i*decay+incr)

LP14 (Collecting of Food)

This property expresses that, if an ant observes that it is at location F (the food source), then it will pick up some food. Formalisation:

observes(a, is_at_location_from(l, e)) and food_location(l) •→» to_be_performed(a, pick_up_food)

LP18 (Decay of Pheromones)

This property expresses that, if the old amount of pheromones at an edge is i, and there is no ant dropping any pheromones at this edge, then the new amount of pheromones at e will be i*decay. Formalisation:

pheromones_at(e, i) and ∀a,l not to_be_performed(a, drop_pheromones_at_edge_from(e, l)) •↠
pheromones_at(e, i*decay)

A special software environment has been created to enable the simulation of executable
models. Based on an input consisting of dynamic properties in LEADSTO format, the software
environment generates simulation traces. An example of such a trace can be seen in Figure 5.
Time is on the horizontal axis, the state properties are on the vertical axis. A dark box on top of
the line indicates that the property is true during that time period, and a lighter box below the line
indicates that the property is false. This trace is based on all local properties identified.

For the sake of readability, in the example situation depicted in Figure 5, only three ants are
involved. However, similar experiments have been performed with a population of 50 ants. Since
the abstract way of modelling used for the simulation is not computationally expensive, also
these simulations took no more than 30 seconds.

As can be seen in Figure 5 there are two ants (ant1 and ant2) that start their search for food
immediately, whereas ant3 comes into play a bit later, at time point 3. When ant1 and ant2 start
their search, none of the locations contain any pheromones yet, so basically they have a free
choice where to go. In the current example, ant1 selects a rather long route to the food source
(via locations A-B-C-D-E-F), whilst ant2 chooses a shorter route (A-G-H-F). Note that, in the
current model, a fixed route preference (via the attractiveness predicate) has been assigned to
each ant for the case there are no pheromones yet. After that, at time point 3, ant3 starts its search
for food. At that moment, there are trails of pheromones leading to both locations B and G, but
these trails contain exactly the same number of pheromones. Thus, ant3 also has a free choice
among location B and G, and chooses in this case to go to B. Meanwhile, at time point 18, ant2
has arrived at the food source (location F). Since it is the first to discover this location, the only
present trail leading back to the nest, is its own trail. Thus ant2 will return home via its own trail.

Next, when ant1 discovers the food source (at time point 31), it will notice that there is a trail leading back that is stronger than its own trail (since ant2 has already walked there twice: back and forth, not too long ago). As a result, it will follow this trail and will keep following ant2 forever. Something similar holds for ant3. The first time that it reaches the food source, ant3 will still follow its own trail, but some time later (from time point 63) it will also follow the other two ants. To conclude, eventually the shortest of both routes is shown to remain, whilst the other route evaporates. Other simulations, in particular for small ant populations, show that it is important that the decay parameter of the pheromones is not too high. Otherwise, the trail leading to the nest has evaporated before the first ant has returned, and all ants get lost!

5.2. Verification for the Ants Scenario

In addition to the simulation software, a software environment has been developed that enables to check dynamic properties specified in TTL against simulation traces. This software environment takes a dynamic property and one or more (empirical or simulated) traces as input, and checks whether the dynamic property holds for the traces. Using this environment, the formal representation relations presented in Section 4 have been automatically checked against traces like the one depicted in Section 5.1. For example, when checking the following property:

$$\forall t1 \; \forall l \; \forall l1 \; \forall e \; \forall a \; [\; state(\gamma, t1) \models is\_at\_edge\_from\_to(a, e, l, l1)$$
$$\Rightarrow \exists t2 > t1 \; state(\gamma, t2) \models pheromone\_at(e) \;]$$

the software simply verifies whether it is always the case that, if an agent is at a certain edge, then at a later time point there is pheromone at that edge. The duration of these checks varied from 1 to 10 seconds, depending on the complexity of the formula (in particular, the backward representation relation has a quite complex structure, since it involves reference to a large number of events in the history). All these checks turned out to be successful, which validates (for the given traces at least) our choice for the representational content of the shared extended

mental state property pheromones_at(e, v). However, note that these checks are only an empirical validation, they are no exhaustive proof as, e.g., model checking is. Currently, the possibilities are explored to combine TTL with existing model checking techniques.

In the process of verifying properties, the specification can be iteratively revised leading to a better specification. For example, the forward representational relations initially did not contain the condition "except the edge that brought the ant to the location" (formalised by the expression e1≠e2; see, e.g., Section 4.3.3). By means of the automated checks, such errors can easily be detected, and recovered. Additionally, open questions may be answered during the verification process. E.g., what is a suitable pheromone decay rate at which ants still accomplish the foraging sufficiently good?

In addition to simulated traces, the checking software allows to check dynamic properties against other types of traces as well. In the future, the representation relations specified in this paper will be checked against traces resulting from other types of ants simulations, and possibly against empirical traces.

6. Slide Case Study

The approach to collective representational content put forward in this paper can be applied in different cases, varying from simple organisms to more complex organisms, such as human beings. The ants case study shows an example in which the internal cognitive processes are simple: the ants are assumed to have purely reactive behaviour (stimulus-response). In this section, in a different type of example it is shown how more complex internal cognitive processes can be taken into account.

In Section 6.1, an example scenario is sketched. In Section 6.2, it is shown how collective representational content can be defined for the example. To illustrate the example, two of the

different types of representation relations shown in Figure 2 and 3 are worked out: Section 6.2.1 addresses a backward relation according to the relational specification approach, and Section 6.2.2 addresses a forward relation according to the relational specification approach. Again, the other combinations can be modelled in a similar manner.

6.1. An Example Scenario

The example, in an adapted and simplified form taken from (Jonker, Treur and Wijngaards, 2001), is about a conference session, which is finishing. The session chairperson, agent A, puts up a slide on the overhead projector, expressing where to find tea and coffee. The persons in the audience, among which agent B, interpret the information available in the projection on the screen. An empirical trace is shown in Table 3, the state properties used are explained in Table 4. In the example, the agents are assumed to have a common ontology on the world including the names of all objects in the world, like the pot for tea, the pattern on the screen, and the names of positions.

In the example, the following world state properties hold, and persist. They express that pot 2 contains tea and that an overhead projector is present:

    contains(pot2,tea)
    is_present(projector)

The scenario is as follows. Agent A observes the world, represented by

    to_be_observed_by(I:INFO_ELEMENT, a)

and as a result obtains information that pot 2 contains tea and that a projector is present, represented, respectively, by:

    observation_result(contains(pot2,tea), pos, a)
    observation_result(is_present(projector), pos, a)

After this it creates the positive beliefs that pot 2 contains tea and that a projector is present:

```
belief(contains(pot2, tea), pos, a)
belief(is_present(projector),pos, a)
```

Based on the belief about the tea, and the agent's characteristic that it is willing to provide information about this to other agents, represented by

```
information_provision_proactive_for(a, contains(pot2, tea))
```

the agent A reasons and concludes that it is desirable that the information about tea is available as a belief to the agents in the audience:

```
desire(belief(contains(pot2,tea), pos, b), a)
```

It is assumed that agent A also has available a belief that to a certain material configuration, namely pattern 3 at position p0 (the screen), the information can be associated that pot 2 contains tea, represented by

```
belief(has_material_rep(contains(pot2,tea), pos, at_position(pattern3,p0), pos), pos, a)
```

This shows that its desire to communicate the information about the tea will be fulfilled if at position p0 in the material world pattern 3 is present. Therefore it generates the intention to bring this about in one way or the other:

```
intention(achieve(at_position(pattern3,p0),pos), a)
```

Moreover, it has beliefs available that an action 'put slide 3' exists which has as an effect that pattern 3 is at position p0 and as opportunity that an overhead projector is present:

```
belief(has_effect(put_slide3,  at_position(pattern3,p0), pos), pos, a)
belief(has_opportunity(put_slide3, is_present(projector), pos), pos, a)
```

Moreover, it has the belief available that indeed the opportunity for this action that a projector is present holds in the world state:

```
belief(is_present(projector), pos, a)
```

Therefore it concludes that the action 'put slide 3' is to be performed:

```
action_initiation(put_slide3, a)
```

This action is performed, and the intended effect is realised in the external world state:

at_position(pattern3, p0)

This effect, the world state property 'pattern 3 is at position p0' is considered an extended mental state for agent A but also for the agents in the audience.

Next it is described how an agent in the audience interacts with this external world state. Agent B (as just one of the agents in the audience) observes the world, represented by

to_be_observed_by(I:INFO_ELEMENT, b)

and as a result obtains information that pattern 3 is at p0:

observation_result(at_position(pattern3, p0), pos, b)

Based on this observation it generates the belief that pattern 3 is at position p0:

belief(at_position(pattern3, p0), pos, b)

Note that agent B is not able to 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. Having the belief (like agent A) that to this world situation the interpretation 'pot 2 contains tea' can be associated, i.e.,

belief(has_material_rep(contains(pot2,tea), pos, at_position(pattern3,p0), pos), pos, b)

it now generates the belief that pot 2 contains tea:

belief(contains(pot2,tea), pos, b)

Note that after this process, the agent B's belief state includes both information that was acquired by observation (pattern 3 is at position p0, which by itself is not of use anymore), and information that was not obtainable for B by direct observation, namely that pot 2 contains tea, which will be useful in guiding the agent's behaviour during the break. This is the information that was acquired via the shared extended mind state 'pattern 3 is at position p0'.

This example scenario of the use of a shared extended mind state is summarised in Table 3 by tracing the states. Time goes from top to bottom. In the table only the relevant information

elements are represented. Notice that the beliefs about the relation between 'pattern 3 at position p0' a 'pot 2 contains tea', and about the opportunity and effect of action 'put slide 3' are persistent beliefs that are available throughout the whole period, but are shown only when taken into account by the agent. The same holds for the information provision proactiveness characteristic of agent A.

The first part of the table gives the state of the external world (first column), and the internal states of the agent A (second column). The second part of the table gives the same for agent B. The first part of the table takes place before the second part of the table.

6.2. Collective Representational Content for the Example

The shared extended mind state considered in this example is the state that pattern 3 is at position p0. The representational content for this state can be relationally specified as before in the following manner.

6.2.1. Backward Relational Specification Approach

For the backward case, the internal state of agent A is involved, in particular its desire to communicate the information about the tea:

> If at some time point t1 an agent a has the desire that another agent b has the belief that pot 2 contains tea,
> and the projector is present,
> then at some later time point t2 pattern 3 will be present at p0.

> If at some time point t2 pattern 3 is present at p0,
> then an agent a exists that at an earlier time point t1 had the desire that
> another agent b has the belief that pot 2 contains tea,
> and the projector was present at t1.

Note that this situation corresponds to the example depicted in Figure 2c: the representation relation relates the external world state property to an internal state property in the past. A formalisation is as follows:

∀γ:TRACE, t1:TIME, a:AGENT, b:AUDIENCE_AGENT
    state(γ, t1) |= desire(belief(contains(pot2, tea), pos, b), a) &
    state(γ, t1) |= is_present(projector) ⇒
        ∃t2:TIME>t1:TIME state(γ, t2) |= at_position(pattern3, p0)

∀γ:TRACE, t2:TIME
    state(γ, t2) |= at_position(pattern3, p0) ⇒
        ∃t1:TIME<t2:TIME, a:AGENT, b:AUDIENCE_AGENT
        state(γ, t1) |= desire(belief(contains(pot2, tea), pos, b), a)
        state(γ, t1) |= is_present(projector)

Note that in this case it is assumed that the agent has some (persistent) beliefs about relevant world knowledge. For example, it beliefs that the information that pot 2 contains tea may be materially represented by pattern 3 at position p0. Without this assumption, such beliefs have to be included in the formalisation as well, yielding the following (slightly more complicated) expressions:

∀γ:TRACE, t1:TIME, a:AGENT, b:AUDIENCE_AGENT
    state(γ, t1) |= desire(belief(contains(pot2, tea), pos, b), a) &
    state(γ, t1) |= belief(has_material_rep(contains(pot2,tea), pos,
        at_position(pattern3,p0), pos), pos, a) &
    state(γ, t1) |= belief(has_effect(put_slide3, at_position(pattern3,p0), pos), pos, a) &
    state(γ, t1) |= belief(has_opportunity(put_slide3, is_present(projector), pos), pos, a) &
    state(γ, t1) |= is_present(projector) ⇒
        ∃t2:TIME>t1:TIME state(γ, t2) |= at_position(pattern3, p0)

∀γ:TRACE, t2:TIME
    state(γ, t2) |= at_position(pattern3, p0) ⇒
        ∃t1:TIME<t2:TIME, a:AGENT, b:AUDIENCE_AGENT
        state(γ, t1) |= desire(belief(contains(pot2, tea), pos, b), a)

& state(γ, t1) |= belief(has_material_rep(contains(pot2, tea), pos,

at_position(pattern3,p0), pos), pos, a) &

state(γ, t1) |= belief(has_effect(put_slide3, at_position(pattern3,p0), pos), pos, a) &

state(γ, t1) |= belief(has_opportunity(put_slide3, is_present(projector), pos), pos, a) &

state(γ, t1) |= is_present(projector)

6.2.2. Forward Relational Specification Approach

For the forward case, the internal state of an agent in the audience (e.g., agent B) is relevant,

in particular its belief about the tea:

If at some time point t1 pattern 3 is present at p0, then for all agents in the audience
there will be a later time point t2 on which they have the belief that pot 2 contains tea.

If at some time point t2 an agent in the audience has the belief that pot 2 contains tea,
then at an earlier time point t1 pattern 3 was present at p0.

Note that this situation corresponds to the example depicted in Figure 3c: the representation

relation relates the external world state property to an internal state property in the future. A

formalisation is as follows:

∀γ:TRACE, t1:TIME
    state(γ, t1) |= at_position(pattern3, p0) ⇒
            ∀a: AUDIENCE_AGENT ∃t2:TIME>t1:TIME
                [state(γ, t2) |= belief(contains(pot2, tea), pos, a)]

∀γ:TRACE, ∀a: AUDIENCE_AGENT, t2:TIME
    state(γ, t2) |= belief(contains(pot2, tea), pos, a)] ⇒
            ∃t1:TIME<t2:TIME
                state(γ, t1) |= at_position(pattern3, p0)

7. Simulation and Verification - Slide

Similar to the ants example, also for the slide example a simulation model has been made,

based on which a number of traces have been generated, and the representation relations have

been checked against the traces.

7.1. A Simulation Model of the Slide Scenario

The scenario from agent A's observations to agent B's belief has been modelled in an executable manner by means of the LEADSTO language. A number of the local dynamic properties that have been used for the model are shown below. See Appendix B for the complete set of local properties.

LP6 (Belief Generation)

This property expresses that, if an agent receives an observation result about certain information, it will believe this information. Formalisation:

observation_result(i, s, a) •→» belief(i, s, a)

LP7 (Desire Generation)

This property expresses that, if an agent beliefs something, and it is willing to share this type of information with others, it will have the desire that all other agents have the same belief. Formalisation:

belief(i, s, a) ∧ information_provision_proactive_for(a, i) •→» ∀b [desire(belief(i, s, b), a)]

LP8 (Intention Generation)

This property expresses that, if an agent desires that other agents belief something, and it beliefs that this information can be materially represented by some pattern, then it will have the intention to create this pattern. Formalisation:

desire(belief(i1, s1, b), a) ∧ belief(has_material_rep(i1, s1, i2, s2), pos, a) •→» intention(achieve(i2, s2), a)

LP9 (Action Initiation)

This property expresses that, if an agent has the intention to create a pattern, and it beliefs that an action ac exists which results in that pattern and for which there is an opportunity, and the pattern is not present yet, then the agent will initiate that action ac. Formalisation:

intention(achieve(i1, s1), a) ∧ belief(has_effect(ac, i1, s1), pos, a) ∧ belief(has_opportunity(ac, i2, s2), pos, a) ∧ belief(i2, s2, a) ∧ not i1 •→» action_initiation(ac, a)

An example trace that was generated on the based of these properties is shown in Figure 6. As the figure shows, initially (at time point 4) only agent A beliefs that pot 2 contains tea. However, it then (at time point 10) initiates the action 'put slide 3', which results in the presence of pattern 3 at position p0 (from time point 12). As a result, agent B and C observe this, so that eventually all agents belief that pot 2 contains tea.

7.2. Verification for the Slide Scenario

Like in Section 5.2, also for the slide case the representational content specification have been checked against simulated traces. Again, all checks eventually turned out to be successful, which validates (for the given traces at least) our choice for the representational content of the shared extended mental state property at_position(pattern3, p0).

Also for the slide example the automated checks turned out useful to detect some initial errors in the representation relations. For example, initially the distinction between AGENT and AUDIENCE_AGENT was not made (i.e., we only used AGENT). However, for the case of the forward representational content (see Section 6.2.2) this resulted in the expression that a belief about the tea is *always* caused by an observation of pattern 3. Obviously, this is incorrect, since agent A's belief about the tea is caused by observation of the tea itself, and *not* by observation of pattern 3. Making a distinction between AGENT and AUDIENCE_AGENT helped to solve this problem.

8. Discussion

In the previous sections, the shared extended mind principle has been applied in two case studies. First, in Section 4 and 5, to illustrate the case of an *unintentionally* created shared extended mind (by species with limited cognitive capabilities) the ants example was addressed. Next, in Section 6 and 7, to illustrate the case of an *intentionally* created shared extended mind (by species with more complex cognitive capabilities) the slide example was addressed. For the latter case, there will not be much discussion about why this example counts as extended mind; the example satisfies all usual criteria for extended mind (e.g., Clark and Chalmers, 1998). However, the former case needs some more explanation. Therefore, this section aims to provide an argument why an ant colony can be interpreted as having a shared extended mind as well.

Moreover, it is discussed whether it is legitimate to speak about a collective representational content in this case.

Historically, the idea of an extended mind used in this paper continues on the idea of active externalism by Clark and Chalmers (1998), that is based on the active role of the environment in driving cognitive processes. The central notion is that the individual brain performs some operations, while others are delegated to manipulations of external media. The authors build on the ideas of Kirsh and Maglio (1994) of *epistemic* actions that alter the world so as to aid and augment cognitive processes such as recognition and search. This in contrast with merely *pragmatic* actions, that alter the world because some physical change is desirable for its own sake (e.g., putting cement into a hole in a dam). Applying these notions to the use of pheromones by ants, the ant can be considered linked with external matter in a two-way interaction. The ant drops pheromones (for its own use and that of others) and detects pheromones that help it in its route taking. In other words, the ant and its pheromone enhanced environment form a coupled system that can be seen as a cognitive system in its own right. If the external component, the pheromones, are removed, the system's behavioural competence decreases: the ants will have more problems in finding the shortest paths. Furthermore, the use of pheromones cannot be considered pragmatic actions, for the presence of pheromones is not desirable for its own sake, i.e., independent of the presence of ants. Therefore, one can conclude that ants engage in active externalism. This, by itself, is not enough to conclude that ants and the pheromones in its environment together form an extended mind.

The notion of mind is irrevocably linked to the notion of mental states (e.g., experiences, beliefs, desires, emotions). Researchers such as Clark and Chalmers (1998), and Tollefsen (2006) in this volume argue that such mental states can be based on features of the external world just as

they might on features of the brain. A recurrent example is the use of a notebook instead of the memory function of the brain that underlies beliefs. The transferral to ants and pheromones are obvious: pheromones are used by the individual ant as indicators of interesting locations (e.g., food, nest). To accept the combination of internal mental states with features of the external world as extended mental states, a number of criteria should be met (see also the Introduction). Clark and Chalmers (1998) state that some fundamental criteria are a high degree of *trust*, *reliance*, and *accessibility*. According to Tollefsen (2006), Clark and Chalmers (1998) also added that the information contained in the resource must have been previously endorsed by the subject. We agree that if the information mysteriously appeared we would be less inclined to accord it the status of a belief. Nevertheless, we did not find this formulation as an additional criterion for the extended mental state in (Clark and Chalmers, 1998). In our opinion this fourth criterion is covered by the criteria trust and reliance. It is irrelevant in what manner the information came to be in the extended mental state, the point is whether or not the behaviour is based on this information. The behaviour will only be effective if the information is reliable, and the behaviour will only occur if the information is trusted (implicitly or explicitly).

To return to the ant example, all three criteria (trust, reliance, and accessibility) are met. The ant places implicit trust in the pheromones. The reliability of pheromones is an interesting point, since pheromones degrade over time. Therefore, the reliability of pheromones depends on the frequency with which pheromone trails are travelled. The reliability further depends on the evaporative properties of pheromones that enable the map of the world as presented by pheromone trails adapts over time. Blocked and other ineffective routes will lose their pheromone trails over time by the fact that ants only maintain effective routes with pheromones. The reliability of pheromones for the individual ant is therefore not guaranteed, but of high

enough quality as shown by the effectiveness of ant colonies in the evolution. The accessibility to the pheromones is constant, except for such strange situations as being picked up and dropped at a place not having pheromones. For the ant the number of types of mental states might not be impressive, but in our view a "small" mind is still a mind. From this last remark another discussion might be triggered, i.e., that of representation.

The discussion about representation relations and whether or not some part of a mental state has a representational content has led to interesting debates, see (Haugeland 1991; Kosslyn, 1994; Clark, 1997). Especially, the question of the form that mental representation takes in non-linguistic creatures, such as human infants and non-human animals is of interest for the work presented in this paper. In (Haugeland, 1991), three requirements for representational content of a "non-extended" mind are defined. In (Clark, 1997), it is discussed how these requirements are translated to the case of an extended mind. Clark (1997)'s line of reasoning provides an interesting point of departure for discussions of collective representational content. The three requirements from (Haugeland, 1991), translated for the extended mind case, are as follows. A system uses external representation in case:

- The system must coordinate its behaviours with some environmental features F that are not always reliably present to the system.

- The system copes by having some other external features F' (instead of the aforementioned environmental features F) that guide the behaviour instead of F.

- The external features F' are part of a more general representational scheme that allows for a variety of related representational states.

Haugeland (1991) writes:

'That which stands in for something else in this way is a representation; that which it stands for is its content; and its standing in for that content is representing it.' (Haugeland, 1991)

In our notation: F' represents content F. Clark provides examples of situations in which, for the internal representation, the three requirements together are too strict. Especially points 2 and 3 cause problems that transfer to the above requirements for external representation with respect to the extended mind. It is not our intention to repeat the debate for the internal mind and its representations for the extended mind and its external representations. However, for the examples in this paper it still needs to be determined to what extend the correlations can be called representations. For this discussion, part of Clark's reasoning is essential. For the internal representation Clark argues that:

'Nor is the mere existence of a reliable, and even nonaccidental, correlation between some inner state and some bodily or environmental parameter sufficient to establish representational status. The mere fact of correlation does not count so much as the nature and complexity of the correlation and the fact that a system in some way consumes or exploits a whole body of such correlations for their specific semantic contents. It is thus important that the system uses the correlations in a way that suggests that the system of inner states has the function of carrying specific types of information. … It may be a static structure or a temporally extended process. It may be local or highly distributed. It may be very accurate or woefully inaccurate. What counts is that it is supposed to carry a certain type of information and that its role relative to other inner systems and relative to the production of behaviour is precisely to bear such information.' (Clark, 1997)

Consider the example of the ants and the use of pheromones. For the individual ant (i.e., the qualitative case of Section 4.2), the reasoning of Clark holds neatly: pheromones are supposed to carry information (e.g., the way to the nest, the way to food) and the whole behaviour of the ant is dependent on this supposition. On the other hand, it is quite unclear whether in this case the correlation between pheromones and directional information has the necessary complexity, and whether the ant's use of pheromones is systematic enough to call the relation between pheromones and ants representational. To establish this once and for all, a quantified measurement of the required complexity and a precise enough characterisation of the required nature of the correlation is needed. So far, such a quantification and characterisation is not available, see (Haselager et al., 2003). Although the representational status of pheromones is still

under debate, pheromones do satisfy to the notion of *adaptive hookup*. Clark (1997) defines

adaptive hookup as:

'Adaptive hookup goes beyond mere causal correlation insofar it requires that the inner states of the system are supposed (by evolution, design, or learning) to coordinate its behaviours with specific environmental contingencies. Representation talk gets its foothold, I suggest, when we confront inner states that, in addition, exhibit a systematic kind of coordination with a whole space of environmental contingencies… Adaptive hookup thus phases gradually into genuine internal representation as the hookup's complexity and systematicity increase.' (Clark, 1997)

Of course "internal" first needs to be read as "external" in order to transfer this definition to

the discussion about pheromones and ants. Millikan (1996, 2001) describes the dance of the

honey bee as a representation of where the nectar is and where the watching bees are to go. The

pheromones used by ants and the dance of the honey bees are both external structures. Like the

dance of the honey bee, an ant's pheromones also have a double function: they represent in

which direction a goal can be found (e.g., nest, or food), and to ants that detect the pheromones it

conveys this same information. Such representations are said to be *pushmi-pullyu*

representations (which roughly correspond to the adaptive hookup of Clark), and far removed

from human desires, beliefs, and such. Millikan's discussion of the map that the honey bee

maintains of the surroundings of its hive allow us to more precisely position the pheromones of

the ants on the transition from adaptive hookup to genuine external representations. Millikan

writes:

"Using a map you can be guided directly from one place to another regardless of whether you have travelled any part of the route before. Thus a bee, when transported by any route to any location in its territory, knows how to fly directly home, or to another location it has in mind, as soon as it has taken its bearing. The bee knows how to make shortcuts." (Millikan, 2001, p.8)

The individual ant, as argued above, cannot entirely copy this feat of finding its way. For

example, if the ant is positioned at a location not frequently visited by other ants, it cannot use

the map in the way a bee can, in order to determine a shortcut to its desired location. Still, in

reality ants do make shortcuts. How this is possible? The answer can be found in the properties

of pheromones, and the fact that the ant is not alone. Thus, as a group the colony of ants produces a map of its territory that enables any ant in its territory to travel along the shortest paths between all major locations in the territory.

By this argument, a gentle transition can be made from the discussion of individual ants having an extended mind (i.e., the qualitative case of Section 4.2), to the colony of ants having a shared (or collective) extended mind (i.e., the quantitative case of Section 4.3). An agreed upon exact definition of a shared extended mind is not available in the literature. Clark and Chalmers (1998) consider socially extended cognition a reasonable option. They write:

> 'What is central is a high degree of trust, reliance, and accessibility. In other social relationships these criteria may not be so clearly fulfilled, but they might nevertheless be fulfilled in specific domains. For example, the waiter at my favorite restaurant might act as a repository of my beliefs about my favorite meals (this might even be construed as a case of extended desire). In other cases, one's beliefs might be embodied in one's secretary, one's accountant, or one's collaborator.' (Clark and Chalmers, 2005)

Tollefsen (2006) states that when minds extend to encompass other minds, a *coupled system* is formed. In this manner, Tollefsen makes more explicit what Clark and Chalmers (1998) hint at with the idea of socially extended cognition. She allows the mind to extend beyond the skin to encompass non-biological artefacts, and other biological agents as resources in the environment. Can we think of a colony of ants having a mind, that is the collective mind of the ants involved? Wilson (2004) provides a first test by differentiating a collective mind from a social manifestation. Social manifestation is the fact that individuals have properties that are manifest only when those individuals form part of a group of a certain type. On the contrary, he defines a collective mind with the fact that a group has properties, including mental states, that are not reducible to the individual states of the individuals. Irrespective of whether or not one judges an ant to have a mind or not, Wilson's differentiation is of interest. A colony of ants has a social manifestation: the individual ants follow a shortest path from one location in the territory to

another. In fact, this observation first led entomologists to thinking that the individual ant maintains a map of the territory. Given the discovery of pheromones, a more parsimonious model of ants arose in which no individual ant has such a map, but that colony of ants does maintain such a map in the form of pheromones that each individual can follow. Therefore, the colony has the property of such a map, the individual has not. Thus, the colony of individuals and pheromones together form one collective mind, or maybe more succinctly formulated: a shared extended mind.

9. Conclusion

The extended mind perspective introduces a high-level conceptualisation of agent-environment interaction processes. By modelling the ants example and the slide example from an extended mind perspective, the following challenging issues on cognitive modelling and representational content were encountered:

1. How to define representational content for an external mental state property

2. How to handle decay of a mental state property

3. How can joint creation of a shared mental state property be modelled

4. What is an appropriate notion of collective representational content of a shared external mental state property

5. How can representational content be defined in a case where a behavioural choice depends on a number of mental state properties

These questions were addressed in this paper. For example, modelling joint creation of mental state properties (3.) was made possible by using relative or leveled mental state properties, parameterised by numbers. Each contribution to such a mental state property was modelled by addition to the level indicated by the number. Collective representational content

(4.) from a looking backward perspective was defined by taking into account histories of such contributions. Collective representational content from a forward perspective was defined taking into account multiple parameterised mental state properties, corresponding to the alternatives for behavioural choices, with their relative weights. In this case it is not possible to define representational content for just one of these mental state properties, but it is possible to define it for their combination or conjunction (5.).

The high-level conceptualisation has successfully been formalised and analysed in a logical manner. The formalisation enables simulation and automated checking of dynamic properties of traces or sets of traces, in particular of the representation relations.

The two case studies considered in this paper have some interesting differences and commonalities. In the slide example, the individual agents are assumed to have complex internal cognitive processes. Therefore, it is very natural to speak about a 'shared mind' in that case, since the pattern created in the external world (i.e., a slide on an overhead projector) represents a shared mental state of the group (i.e., a belief that a pot contains tea). However, in the ants example, the internal processes of the individual agents are assumed to be very limited. The ants do not really 'understand' what they are doing. Nevertheless, as a result of their behaviour, a structure (i.e., a pattern of pheromones) emerges that can be described as a shared 'mind', as shown in the previous section. For example, the external world state property 'pheromone is present at edge e' can be described using a mental notion such as 'the group believes that e is a relevant direction'. This is in line with the famous claim (often used as a reaction to Searle's Chinese Room Argument, Searle (1980, 1984)) that 'intelligence' of complex systems is often an emerging property of the whole, not of the individual parts; e.g., Dennett (1991), pp. 435-440.

Considering related work, there is a large body of literature that has some connection to the issues addressed in this paper, both in the area of Cognitive Science and beyond. From a broad perspective, the issues have been investigated for years in disciplines such as psychology, Computer-Supported Cooperative Work (CSCW) and Human-Computer Interaction (HCI). Two main examples are the concepts of Distributed Cognition and Activity Theory. *Distributed Cognition* (Hutchins, 1991; Salomon, 1993; Hollans et al., 2000) is a branch of Cognitive Science that proposes that human knowledge and cognition are not confined to the individual. Instead, it is distributed by placing memories, facts, or knowledge on the objects, individuals, and tools in our environment. Distributed cognition it is a useful approach for (re)designing social aspects of cognition by putting emphasis on the individual and their environment. The approach views a system as a set of representations, and models the interchange of information between these representations. These representations can be either in the mental space of the participants or external representations available in the environment. *Activity Theory*, invented by A.N. Leontiev (1981), is a Soviet psychological theory, based on the idea that cognition is distributed among individuals and part of the environment. Activity Theory became one of the major psychological theories in the Soviet Union, being used widely in areas such as the education of disabled children and the design of equipment control panels. See (Nardi, 1996) for a collection of papers about Activity Theory applied in different contexts.

From a narrow perspective, recently other researchers in Cognitive Science have also tried to define criteria for the notion of a collective mind, consisting of multiple extended minds (e.g. Clark and Chalmers (1998), and Tollefsen (2006), see Section 8). Tollefsen works out the thesis that the mind is not bounded by skin and bones, making way for the concept of a collective mind. She includes multiple thought experiments that illustrate how to extend the *coupled system* to

cover not only individual-artifact (computers, calculators, etcetera), but also individual-individual. An example thought experiment concerns a husband and wife, where he is a disorganised professor and she reminds him of appointments, meetings, and so on. Together they form a coupled system; hence they have a collective mind. Our work described here considers a similar notion of a collective (or: shared extended) mind, but with some important differences. Firstly, we explicitly address the issue of representational content (Bickhard, 1993) in going from a single extended mind to a collective mind. Secondly, the thought experiments consider only one-to-one interactions (husband-wife), whereas our experiments consider many-to-many interactions.

For future research, it is planned to make the distinction between extended mind states and other external world states more concrete. In addition, the approach will be applied to several other cases of extended mind. For example, can the work be related to AI planning representations, traffic control, knowledge representation of negotiation, and to the concept of "shared knowledge" in knowledge management?

References

Bickhard, M.H. (1993). Representational Content in Humans and Machines. *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 5, pp. 285-333.

Bonabeau, E. (1999). Editor's Introduction: Stigmergy. In: *Artificial Life*, Vol. 5, issue 2, pp. 95-96.

Bonabeau, J. Dorigo, M. and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, New York.

Bosse, T., Jonker, C.M., Schut, M.C., and Treur, J. (2005). Simulation and Analysis of Shared Extended Mind. In: Davidsson, P., Gasser, L., Logan, B., and Takadama, K. (eds.), *Proceedings of the First Joint Workshop on Multi-Agent and Multi-Agent-Based Simulation, MAMABS'04*. Lecture Notes in AI, vol. 3415. Springer Verlag, pp. 248-264.

Bosse, T., Jonker, C.M., and Treur, J. (2003). Simulation and analysis of controlled multi-representational reasoning processes. *Proc. of the Fifth International Conference on Cognitive Modelling, ICCM'03*. Universitats-Verlag Bamberg, pp. 27-32.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. MIT Press.

Clark, A. (2001). Reasons, Robots and the Extended Mind. In: *Mind & Language*, vol. 16, pp. 121-145.

Clark, A., and Chalmers, D. (1998). The Extended Mind. In: *Analysis,* vol. 58, pp. 7-19.

Dennett, D.C. (1991). *Consciousness Explained*, Penguin, London.

Dennett, D.C. (1996). *Kinds of Mind: Towards an Understanding of Consciousness*, New York: Basic Books.

Grassé, P.-P. (1959). La Reconstruction du nid et les Coordinations Inter-Individuelles chez Bellicositermes Natalensis et Cubitermes sp. La théorie de la Stigmergie: Essai d'interpretation du Comportement des Termites Constructeurs. In : *Insectes Sociaux*, vol. 6., pp. 41-81.

Griffiths, P. and Stotz, K. (2000). How the mind grows: a developmental perspective on the biology of cognition. *Synthese*, vol.122, pp. 29--51.

Haselager, P., Groot, A. de, and Rappard, H. van. (2003). Representationalism vs. Anti-representationalism: a debate for the sake of appearance. In: *Philosophical Psychology*, vol. 16, no. 1.

Haugeland, J. (1991). Representational Genera. In: W. Ramsey, S.P. Stich, and D.E. Rumelhart (eds.), *Philosophy and Connectionist Theory*, Hillsdale, N.J.: Erlbaum, pp. 61-78.

Hollan, J., Hutchins, E., and Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. In: *ACM Transactions on Computer-Human Interaction*, Vol. 7, Issue 2, Special issue on human-computer interaction in the new millennium, Part 2, pp. 174-196.

Hutchins, E. (1991). The social organisation of distributed cognition. In: Resnick, L. (ed.) *Perspectives on Socially Shared Cognition*. Washington, DC: American Psychological Association, pp. 238-287.

Jacob, P. (1997). *What Minds Can Do: Intentionality in a Non-Intentional World*. Cambridge University Press, Cambridge.

Jonker, C.M., and Treur, J. (2003). A Temporal-Interactivist Perspective on the Dynamics of Mental States. *Cognitive Systems Research Journal*, vol. 4, pp. 137-155.

Jonker, C.M., Treur, J., and Wijngaards, W.C.A. (2003). A Temporal Modelling Environment for Internally Grounded Beliefs, Desires and Intentions. *Cognitive Systems Research Journal*, vol. 4, pp. 191-210.

Kim, J. (1996). *Philosophy of Mind*. Westview Press.

Kirsh, D. and Maglio, P. (1994). On distinguishing epistemic from pragmatic action. In: *Cognitive Science*, vol. 18, no.4, pp. 513-549.

Kosslyn S. (1994). *Image and Brain*. Cambridge, MA: MIT Press.

Leontiev, A.N. (1981). *Problems in the development of the mind*. Moscow: Progress.

Menary, R. (ed.) (2004). *The Extended Mind,* Papers presented at the Conference *The Extended Mind - The Very Idea: Philosophical Perspectives on Situated and Embodied Cognition*, University of Hertfordshire, 2001. John Benjamins.

Millikan, R.G. (1996). Pushmi-pullyu Representations. In: Tomberlin, J. (ed.), *Philosophical Perspectives*, vol. IX, Ridgeview Publishing, pp. 185-200. Reprinted in: May, L., and Friedman, M. (eds.), *Mind and Morals*, MIT Press, pp. 145-161.

Millikan, R.G. (2001). Kantian Reflections on Animal Minds. In: *A Priori*, http://www.apriori.canterbury.ac.nz.

Nardi, B. A. (ed.) (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction*. Cambridge, MA: MIT Press.

Salomon, G. (1993). Editor's introduction. In: G. Salomon (ed.), *Distributed Cognitions: Psychological and Educational Considerations* (pp. xi-xxi). NY: Cambridge University Press.

Scheele, M. (2002). Team Action: A Matter of Distribution. Distributed Cognitive systems and the Collective Intentionality they Exhibit. *Proc. of the Third International Conference on Collective Intentionality*.

Searle, J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3, pp. 417-424.

Searle, J. (1984). *Minds, Brains, and Science*. Cambridge: Harvard University Press.

Tollefsen, D.P. (2006). From Extended Mind to Collective Mind. *Cognitive Systems Research Journal* (this volume).

Wilson, R. (2004). *Boundaries of the mind*. Cambridge University Press.

Author Note

Table 1

State Properties used in the Ants Scenario.

	body positions in world:
pheromone is present at edge e (only used in qualitative case)	pheromone_at(e)
pheromone level at edge e is i	pheromones_at(e, i)
ant a is at location l coming from e	is_at_location_from(a, l, e)
ant a is at edge e to l2 coming from location l1	is_at_edge_from_to(a, e, l1, l2)
ant a is carrying food	is_carrying_food(a)
	world state properties:
edge e connects location l1 and l2	connected_to_via(l1, l2, e)
location l is the nest location	nest_location(l)
location l is the food location	food_location(l)
location l has i neighbours	neighbours(l, i)
edge e is most attractive for ant a coming from location l	attractive_direction_at(a, l, e)
	input state properties:
ant a observes that it is at location l coming from edge e	observes(a, is_at_location_from(l, e))
ant a observes that it is at edge e to l2 coming from location l1	observes(a, is_at_edge_from_to(e, l1, l2))
ant a observes that edge e has pheromone level i	observes(a, pheromones_at(e, i))
	output state properties:
ant a initiates action to go to edge e to l2 coming from location l1	to_be_performed(a, go_to_edge_from_to(e, l1, l2))
ant a initiates action to go to location l coming from edge e	to_be_performed(a, go_to_location_from(l, e))
ant a initiates action to drop pheromones at edge e coming from location l	to_be_performed(a, drop_pheromones_at_edge_from(e, l))
ant a initiates action to pick up food	to_be_performed(a, pick_up_food)
ant a initiates action to drop food	to_be_performed(a, drop_food)

Table 2

Different types of Representation Relations.

	backward interactivist approach	backward relational specification approach	forward interactivist approach	forward relational specification approach
qualitative case	4.2.1	4.2.2	4.2.3	4.2.4
quantitative case	4.3.1	4.3.2	4.3.3	4.3.4

Table 3

Empirical Trace of the Slide Scenario.

External World	Agent A
pot 2 contains tea projector is present	to_be_observed_by(I:INFO_ELEMENT, a)
	observation_result(contains(pot2,tea), pos, a) observation_result(is_present(projector), pos, a)
	belief(contains(pot2,tea), pos, a) belief(is_present(projector),pos, a) information_provision_proactive_for(a, contains(pot2, tea))
	desire(belief(contains(pot2,tea), pos, b), a) belief(has_material_rep(contains(pot2,tea), pos, at_position(pattern3,p0), pos), pos, a)
	intention(achieve(at_position(pattern3,p0),pos), a)

	belief(has_effect(put_slide3,
	   at_position(pattern3,p0), pos), pos, a)
	belief(has_opportunity(put_slide3,
	   is_present(projector), pos), pos, a)

	action_initiation(put_slide3, a)
slide 3 at projector pattern 3 at p0	

External World	Agent B
pot 2 contains tea slide 3 at projector pattern 3 at p0	to_be_observed_by(I:INFO_ELEMENT, b)
	observation_result(at_position(pattern3,p0), pos, b)
	belief(at_position(pattern3,p0), pos, b) belief(has_material_rep(contains(pot2,tea), pos, at_position(pattern3,p0), pos), pos, b)
	belief(contains(pot2,tea), pos, b)

Table 4

State Properties used in the Slide Scenario.

Formalisation	Explanation
to_be_observed_by	The information that the agent focuses on and observes in the world.
observation_result	The information received by the sensors of the agent, including the sign of the information. The sign (pos or neg) indicates if the information is true or false.
belief	The beliefs of the agent. Refers to world information and a sign.
desire	A desire of the agent.
intention	An intention of the agent.
has_effect	Denotes that an action is capable of bringing about some state in the world, given as the state that becomes true or false in the world.
has_opportunity	Denotes that an action has a condition (a world state property) that indicates when there is an opportunity for the action.
action_initiation	Indicates that the agent initiates a specified action.
has_ 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.
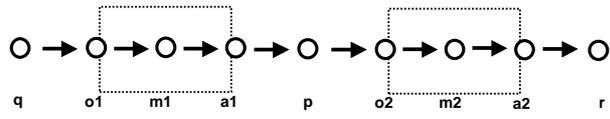
Figure 1: Processes involved in the Creation and Utilisation of Shared Extended Mind.

a) Reference to External World State
(e.g. using relational specification approach)

b) Reference to Observation State
(e.g. using interactivist approach)

c) Reference to Internal State
(e.g. using relational specification approach)

d) Reference to Action State
(e.g. using interactivist approach)
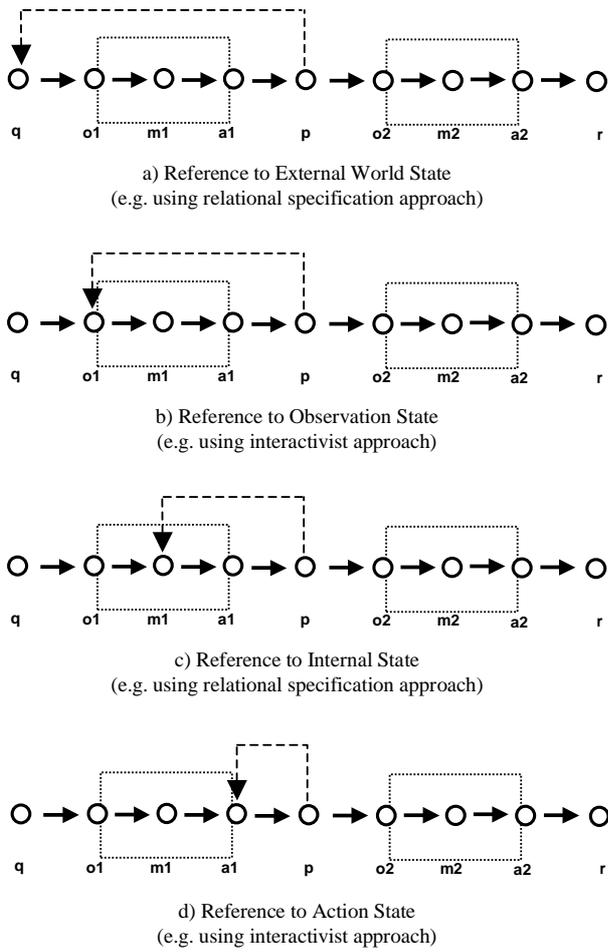
Figure 2: Backward Representation Relations.
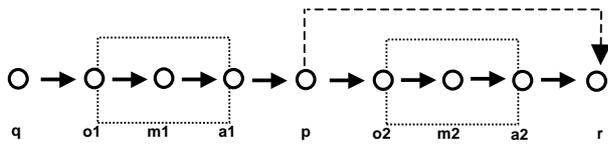
a) Reference to External World State
(e.g. using relational specification approach)
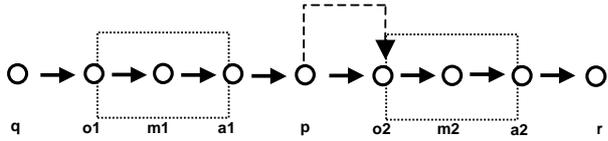
b) Reference to Observation State
(e.g. using interactivist approach)
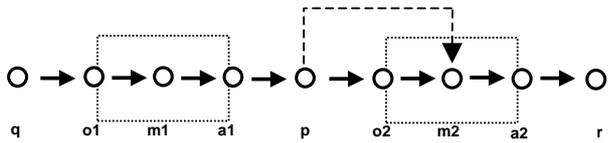
c) Reference to Internal State
(e.g. using relational specification approach)
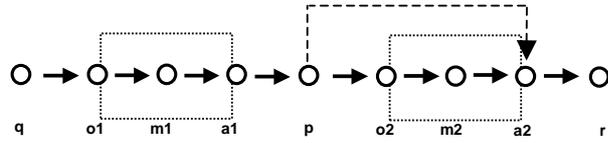
d) Reference to Action State
(e.g. using interactivist approach)

Figure 3: Forward Representation Relations.
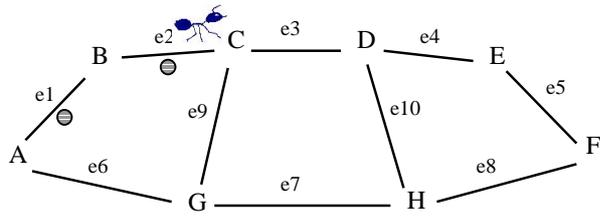
Figure 4: An Ants World.
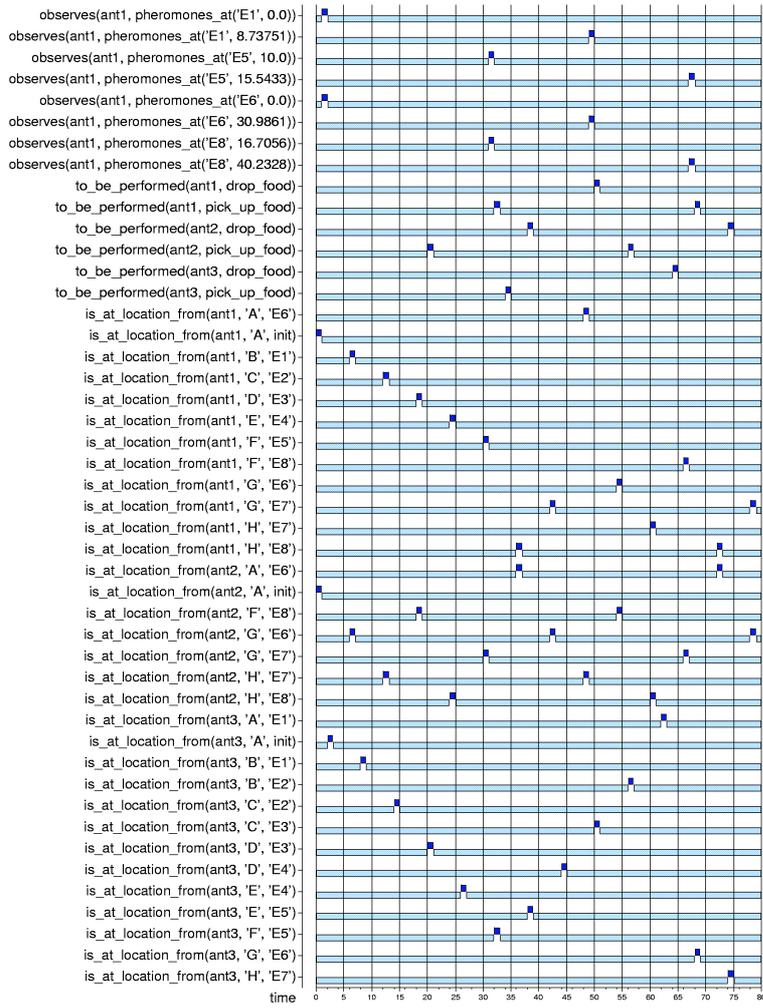
Figure 5: Simulation Trace - Ants Example.
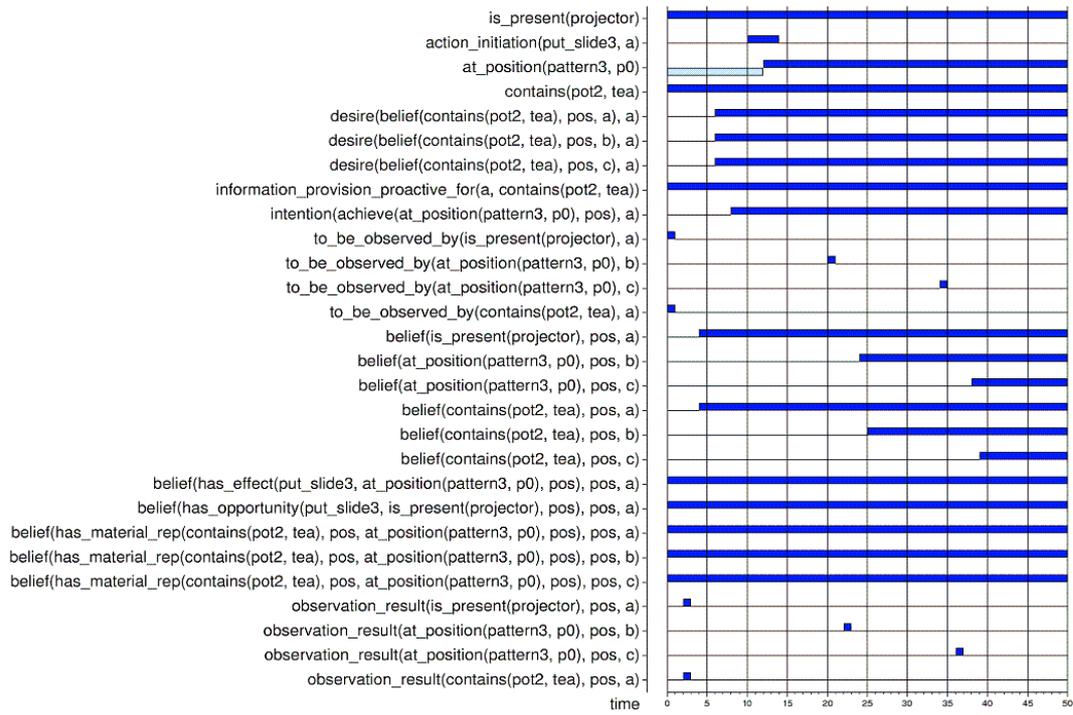
Figure 6: Simulation Trace - Slide Example.

Appendix A - Ants Simulation Model

LP1 (Initialisation of Pheromones)
This property expresses that at the start of the simulation, at all locations there are 0 pheromones. Formalisation:
start ●→» pheromones_at(E1, 0.0) and pheromones_at(E2, 0.0) and pheromones_at(E3, 0.0) and pheromones_at(E4, 0.0) and pheromones_at(E5, 0.0) and pheromones_at(E6, 0.0) and pheromones_at(E7, 0.0) and pheromones_at(E8, 0.0) and pheromones_at(E9, 0.0) and pheromones_at(E10, 0.0)

LP2 (Initialisation of Ants)
This property expresses that at the start of the simulation, all ants are at location A. Formalisation:
start ●→» is_at_location_from(ant1, A, init) and is_at_location_from(ant2, A, init) and is_at_location_from(ant3, A, init)

LP3 (Initialisation of World)
These two properties model the ants world. The first property expresses which locations are connected to each other, and via which edges they are connected. The second property expresses for each location how many neighbours it has. Formalisation:
start ●→» connected_to_via(A, B, l1) and … and connected_to_via(D, H, l10)
start ●→» neighbours(A, 2) and … and neighbours(H, 3)

LP4 (Initialisation of Attractive Directions)
This property expresses for each ant and each location, which edge is most attractive for the ant at if it arrives at that location. This criterion can be used in case an ant arrives at a location where there are two edges with an equal amount of pheromones. Formalisation:
start ●→» attractive_direction_at(ant1, A, E1) and … and attractive_direction_at(ant3, E, E5)

LP5 (Selection of Edge)
These properties model the edge selection mechanism of the ants. For example, the first property expresses that, when an ant observes that it is at location A, and both edges connected to location A have the same number of pheromones, then the ant goes to its attractive direction. Formalisation:
observes(a, is_at_location_from(A, e0)) and attractive_direction_at(a, A, e1) and connected_to_via(A, l1, e1) and observes(a, pheromones_at(e1, i1)) and connected_to_via(A, l2, e2) and observes(a, pheromones_at(e2, i2)) and e1 \= e2 and i1 = i2 ●→» to_be_performed(a, go_to_edge_from_to(e1, A, l1))

observes(a, is_at_location_from(A, e0)) and connected_to_via(A, l1, e1) and observes(a, pheromones_at(e1, i1)) and connected_to_via(A, l2, e2) and observes(a, pheromones_at(e2, i2)) and i1 > i2 •↠ to_be_performed(a, go_to_edge_from_to(e1, A, l1))

observes(a, is_at_location_from(F, e0)) and connected_to_via(F, l1, e1) and observes(a, pheromones_at(e1, i1)) and connected_to_via(F, l2, e2) and observes(a, pheromones_at(e2, i2)) and i1 > i2 •↠ to_be_performed(a, go_to_edge_from_to(e1, F, l1))

observes(a, is_at_location_from(l, e0)) and neighbours(l, 2) and connected_to_via(l, l1, e1) and e0 ≠ e1 and l ≠ A and l ≠ F •↠ to_be_performed(a, go_to_edge_from_to(e1, l, l1))

observes(a, is_at_location_from(l, e0)) and attractive_direction_at(a, l, e1) and neighbours(l, 3) and connected_to_via(l, l1, e1) and observes(a, pheromones_at(e1, 0.0)) and connected_to_via(l, l2, e2) and observes(a, pheromones_at(e2, 0.0)) and e0 ≠ e1 and e0 ≠ e2 and e1 ≠ e2 •↠ to_be_performed(a, go_to_edge_from_to(e1, l, l1))

observes(a, is_at_location_from(l, e0)) and neighbours(l, 3) and connected_to_via(l, l1, e1) and observes(a, pheromones_at(e1, i1)) and connected_to_via(l, l2, e2) and observes(a, pheromones_at(e2, i2)) and e0 ≠ e1 and e0 ≠ e2 and e1 ≠ e2 and i1 > i2 •↠ to_be_performed(a, go_to_edge_from_to(e1, l1))

LP6 (Arrival at Edge)
This property expresses that, if an ant goes to an edge e from a location l to a location l1, then later the ant will be at this edge e. Formalisation:
to_be_performed(a, go_to_edge_from_to(e, l, l1)) •↠ is_at_edge_from_to(a, e, l, l1)

LP7 (Observation of Edge)
This property expresses that, if an ant is at a certain edge e, going from a location l to a location l1, then it will observe this. Formalisation:
is_at_edge_from_to(a, e, l, l1) •↠ observes(a, is_at_edge_from_to(e, l, l1))

LP8 (Movement to Location)
This property expresses that, if an ant observes that it is at an edge e from a location l to a location l1, then it will go to location l1. Formalisation:
observes(a, is_at_edge_from_to(e, l, l1)) •↠ to_be_performed(a, go_to_location_from(l1, e))

LP9 (Dropping of Pheromones)

This property expresses that, if an ant observes that it is at an edge e from a location l to a location l1, then it will drop pheromones at this edge e. Formalisation:

observes(a, is_at_edge_from_to(e, l, l1)) •→» to_be_performed(a, drop_pheromones_at_edge_from(e, l))

LP10 (Arrival at Location)
This property expresses that, if an ant goes to a location l from an edge e, then later it will be at this location l. Formalisation:

to_be_performed(a, go_to_location_from(l, e)) •→» is_at_location_from(a, l, e)

LP11 (Observation of Location)
This property expresses that, if an ant is at a certain location l, then it will observe this. Formalisation:

is_at_location_from(a, l, e) •→» observes(a, is_at_location_from(l, e))

LP12 (Observation of Pheromones)
This property expresses that, if an ant is at a certain location l, then it will observe the number of pheromones present at all edges that are connected to location l. Formalisation:

is_at_location_from(a, l, e0) and connected_to_via(l, l1, e1) and pheromones_at(e1, i) •→» observes(a, pheromones_at(e1, i))

LP13 (Increment of Pheromones)
These properties model the increment of the number of pheromones at an edge as a result of ants dropping pheromones. For example, the first property expresses that, if an ant drops pheromones at edge e, and no other ants drop pheromones at this edge, then the new number of pheromones at e becomes i*decay+incr. Here, i is the old number of pheromones, decay is the decay factor, and incr is the amount of pheromones dropped. Formalisation:

to_be_performed(a1, drop_pheromones_at_edge_from(e, l1)) and $\forall$l2 not to_be_performed(a2, drop_pheromones_at_edge_from(e, l2)) and $\forall$l3 not to_be_performed(a3, drop_pheromones_at_edge_from(e, l3)) and a1 ≠ a2 and a1 ≠ a3 and a2 ≠ a3 and pheromones_at(e, i) •→» pheromones_at(e, i*decay+incr)

to_be_performed(a1, drop_pheromones_at_edge_from(e, l1)) and to_be_performed(a2, drop_pheromones_at_edge_from(e, l2)) and $\forall$l3 not to_be_performed(a3, drop_pheromones_at_edge_from(e, l3)) and a1 ≠ a2 and a1 ≠ a3 and a2 ≠ a3 and pheromones_at(e, i) •→» pheromones_at(e, i*decay+incr+incr)

to_be_performed(a1, drop_pheromones_at_edge_from(e, l1)) and to_be_performed(a2,

drop_pheromones_at_edge_from(e, l2)) and to_be_performed(a3,

drop_pheromones_at_edge_from(e, l3)) and a1 ≠ a2 and a1 ≠ a3 and a2 ≠ a3 and pheromones_at(e,

i) ●→» pheromones_at(e, i*decay+incr+incr+incr)

LP14 (Collecting of Food)

This property expresses that, if an ant observes that it is at location F (the food source), then it will pick up some food. Formalisation:

observes(a, is_at_location_from(l, e)) and food_location(l) ●→» to_be_performed(a, pick_up_food)

LP15 (Carrying of Food)

This property expresses that, if an ant picks up food, then as a result it will be carrying food. Formalisation:

to_be_performed(a, pick_up_food) ●→» is_carrying_food(a)

LP16 (Dropping of Food)

This property expresses that, if an ant is carrying food, and observes that it is at location A (the nest), then the ant will drop the food. Formalisation:

observes(a, is_at_location_from(l, e)) and nest_location(l) and is_carrying_food(a) ●→»

to_be_performed(a, drop_food)

LP17 (Persistence of Food)

This property expresses that, as long as an ant that is carrying food does not drop the food, it will keep on carrying it. Formalisation:

is_carrying_food(a) and not to_be_performed(a, drop_food) ●→» is_carrying_food(a)

LP18 (Decay of Pheromones)

This property expresses that, if the old amount of pheromones at an edge is i, and there is no ant dropping any pheromones at this edge, then the new amount of pheromones at e will be i*decay. Formalisation:

pheromones_at(e, i) and ∀a,l not to_be_performed(a, drop_pheromones_at_edge_from(e, l)) ●→»

pheromones_at(e, i*decay)

Appendix B - Slide Simulation Model

LP1 (Initialisation of World)

These properties express that at the start of the simulation, pot 2 contains tea, and an overhead projector is present. Formalisation:

start •↠ contains(pot2, tea)

start •↠ is_present(projector)

LP2 (Initialisation of Agent Beliefs)

These properties express that at the start of the simulation, all agents belief that the information that pot 2 contains tea may be materially represented by pattern 3 at position p0. Moreover, agent A beliefs that an action 'put slide 3' exists which has as an effect that pattern 3 is at position p0 and as opportunity that an overhead projector is present. Formalisation:

start •↠ belief(has_material_rep(contains(pot2, tea), pos, at_position(pattern3, p0), pos), pos, a)

start •↠ belief(has_material_rep(contains(pot2, tea), pos, at_position(pattern3, p0), pos), pos, b)

start •↠ belief(has_material_rep(contains(pot2, tea), pos, at_position(pattern3, p0), pos), pos, c)

start •↠ belief(has_effect(put_slide3, at_position(pattern3, p0), pos), pos, a)

start •↠ belief(has_opportunity(put_slide3, is_present(projector), pos), pos, a)

LP3 (Initialisation of Agent Characteristics)

This property expresses that at the start of the simulation, agent A is willing to provide information about the fact that pot 2 contains tea. Formalisation:

start •↠ information_provision_proactive_for(a, contains(pot2, tea))

LP4 (Initialisation of Agent Observations)

These properties express that at the start of the simulation, agent A initiates the observation whether pot 2 contains tea, and whether an overhead projector is present. Moreover, agent B and C initiate (after a while) the observation whether there is a pattern at position p0. Formalisation:

start •↠ to_be_observed_by(contains(pot2, tea), a)

start •↠ to_be_observed_by(is_present(projector), a)

start •↠$_{20,20,1,1}$ to_be_observed_by(at_position(pattern3, p0), b)

start •↠$_{34,34,1,1}$ to_be_observed_by(at_position(pattern3, p0), c)

LP5 (Observation Effectiveness)

This property expresses that, if an agent initiates an observation about something that is true in the world, it receives a positive observation result. Formalisation:

i ∧ to_be_observed_by(i, a) •↠ observation_result(i, pos, a)

LP6 (Belief Generation)

This property expresses that, if an agent receives an observation result about certain information, it will believe this information. Formalisation:

observation_result(i, s, a) •↠ belief(i, s, a)

LP7 (Desire Generation)

This property expresses that, if an agent beliefs something, and it is willing to share this type of information with others, it will have the desire that all other agents have the same belief. Formalisation:

belief(i, s, a) ∧ information_provision_proactive_for(a, i) •↠ ∀b [desire(belief(i, s, b), a)]

LP8 (Intention Generation)

This property expresses that, if an agent desires that other agents belief something, and it beliefs that this information can be materially represented by some pattern, then it will have the intention to create this pattern. Formalisation:

desire(belief(i1, s1, b), a) ∧ belief(has_material_rep(i1, s1, i2, s2), pos, a) •↠ intention(achieve(i2, s2), a)

LP9 (Action Initiation)

This property expresses that, if an agent has the intention to create a pattern, and it beliefs that an action ac exists which results in that pattern and for which there is an opportunity, and the pattern is not present yet, then the agent will initiate that action ac. Formalisation:

intention(achieve(i1, s1), a) ∧ belief(has_effect(ac, i1, s1), pos, a) ∧ belief(has_opportunity(ac, i2, s2), pos, a) ∧ belief(i2, s2, a) ∧ not i1 •↠ action_initiation(ac, a)

LP10 (Action Effectiveness)

This property expresses that, if an agent performs the action 'put slide 3', this will lead to pattern 3 being at position p0. Formalisation:

action_initiation(put_slide3, a) •↠ at_position(pattern3, p0)

LP11 (Belief Derivation)

This property expresses that, if an agent beliefs that some pattern exists, and that this patterns is the material representation of some information, then it will also believe this information. Formalisation:

belief(i1, s1, a) ∧ belief(has_material_rep(i2, s2, i1, s1), pos, a) •↠ belief(i2, s2, a)

LP12 (World Persistence)

This property expresses that, if some information exists in the world, then this information will persist forever (assuming for simplicity that no event will occur that destroys the information). Formalisation:

i ●⟫ i

LP13 (Belief Persistence)

This property expresses that, if an agent has a certain belief, then it will have this belief forever (assuming for simplicity that it does not forget anything). Formalisation:

belief(i, s, a) ●⟫ belief(i, s, a)

Footnotes

[1] Note that this picture can also be used to describe the 'traditional' situation of a (non-shared) extended mind for a single agent. In that case, both rectangles would correspond to the same agent.

[2] To obtain interesting simulation traces, different attractive directions were assigned to different ants. However, another possibility (that is supported by the software) is to assign attractive directions at random.