A Formal Specification Language for Agent-Oriented Software Engineering

Hong Zhu
Department of Computing, School of Technology, Oxford Brookes University
Wheatley Campus, Oxford OX33 1HX, UK. Tel.: ++44 1865 484580, Email: hzhu@brookes.ac.uk

ABSTRACT
One of the most appealing features of agent technology is its natural way to modularise complex systems in terms of multiple interacting autonomous components. This feature is supported by the language facility castes in the formal specification language SLABS, which is designed for modular and composable specification of multi-agent systems. The paper reports the syntax and semantics of the language, and illustrates its style of formal specification by a distributed synchronisation algorithm.

Keywords
Formal specification language, Multiagent Systems, Scenario, Software Agents, Castes

1. INTRODUCTION
In recent years, agent technology has been applied to more and more critical application areas such as telecommunications [1], space adventure [2], power grid control, military [3], etc. to provide viable solutions to the problems that was unable to be solved satisfactorily by other existing techniques. It is also clearly demonstrated that agent technology is particularly suitable to solve problems related to web-based applications such as e-commerce and web search engines [4]. However, developing agent-based systems is extremely difficult because the dynamic behaviours of agent-based systems are difficult to specify, analyse, verify and validate. Being autonomous, proactive and adaptive, an agent-based system may demonstrate emergent behaviours that are hard to predict, difficult to design, and expensive to test. An early incident of software failure attributed to autonomous agents in particular is the crash of Air France's Airbus 320 at an air show in June 1988 [5]. Airbus 320 was the first fly-by-wire passenger aircraft in the world. In other words, an autonomous agent controls the aircraft. The incident was caused by a conflict between the human pilot's instruction and the autonomous control by the software. While the pilot intended to fly over the airport in the air show, the fly-by-wire control software seems to have instructed the aircraft to land, which was believed to be the cause of the accident [6].

The new features of agent-based systems demand new methods for the specification of agent behaviours and for the verification and validation of their properties to enable software engineers to develop reliable and trustworthy agent-based systems. It has been recognised that the lack of rigour is one of the major factors hampering the wide-scale adoption of agent technology [7]. On the other hand, the modularity inherent in multi-agent systems can offer a new approach to decomposing complicated formal specifications into composable modular components.

The past few years have seen increasing research interests in agent-oriented software development methodology. Existing work falls into three main classes. The first is towards the theoretical foundations for understanding agent-based systems. Much work has been focused on modelling and reasoning agents' rational behaviour by introducing modalities for belief, desire and intention, e.g. [8, 9, 10, 11]. Game theory has also found its position in the formalisation of agent models, e.g. [12]. A great number of formal models of agents have been proposed and investigated in the literature; see e.g. [13, 14]. Most of them are based on an internal mental state model of agents, yet some are based on a model of the external social behaviours of collaborative agents, e.g. [15]. Although these formal models of agents significantly improved our understanding of agent-based systems, they do not immediately facilitate the development of agent-based systems. As pointed out in [16], a specification method based on a specific model of agents may result in the existence of certain agent theory and systems that do not match the concept in the specification formalism. Moreover, temporal logics, particularly when combined with modalities for belief, desire, etc., can be very complex. The second group of researches is on the development process and development methods for engineering agent-based systems, see e.g. [17, 18, 19, 20, 21]. These works mostly focused on diagrammatic notations that support the analysis and design of multi-agent systems. Some of the notations extend object-oriented methods and notations such as UML. Some introduce new models and new diagrammatic notations. How such diagrammatic notations are related to the logic and formal models of agents remains an open problem. The third group consists of the researches on the language facilities and features that support the formal specification and verification of agent-based systems in a software engineering context, although there is little such work reported in the literature [22]. The use of existing formal specification languages, such as Z, has also been explored to specify agent architecture [23] and concepts [24, 25]. Despite the large number of publications on agents in the literature, we are lack of researches on language facilities that support the development of large-scale complicated multi-agent systems. In particular, we are lack of language facilities to explicitly specify the environment of agents and agent-based systems although it is widely recognised that an important characteristic of agents is that they are entities situated (embedded) in a particular environment [26]. We are lack of facilities that can clearly state how agents' behaviours are related to the environment. We are lack of language facilities that enable us to maximise the power of the way that multi-agent systems modularise complex systems into cooperative autonomous components. Addressing these problems, in the past 3 years we have been searching for language facilities to support the analysis, specification, design and implementation of agent-based systems in the context of software engineering. Some of the results have been incorporated in the design of a formal specification language called SLABS [27, 28].

In this paper, we report the main features of the language SLABS and demonstrate its uses in the development of multi-agent systems. The remainder of the paper is organised as follows. Section 2 defines the syntax and semantics of the language. Section 3 illustrates SLABS' style of specification by an example. Section 4 concludes that paper with discussions on related work.
2. THE LANGUAGE SLABS

The meta-language used to define the syntax is EBNF; see Table 1. In a syntax definition, meta-symbols are in bold font such as ::=; Terminals are in italic font such as Var. Non-terminals are in normal font such as Agent-Description.

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>::=</td>
<td>A ::= B means that A is defined as B.</td>
</tr>
<tr>
<td>Concatenation</td>
<td>AB</td>
<td>A is followed by B.</td>
</tr>
<tr>
<td>Optional</td>
<td>[ ]</td>
<td>[A ] means that A is optional.</td>
</tr>
<tr>
<td>Choice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetition</td>
<td>{ }</td>
<td>{A } means that A may appear any times including zero times or more times.</td>
</tr>
<tr>
<td>Repetition with separator</td>
<td>/</td>
<td>{A</td>
</tr>
<tr>
<td>Positive repetition</td>
<td>+</td>
<td>{A } means that A may appear at least once.</td>
</tr>
<tr>
<td>Parenthesis</td>
<td>( )</td>
<td>They are used to change preference.</td>
</tr>
</tbody>
</table>

2.1 Agents and Castes

The specification of a multi-agent system in SLABS consists of a set of specifications of agents and castes.

System ::= \{Agent-description | caste-description\}*

Caste is one of the novel concepts in SLABS, which is a natural evolution of the concepts of classes in object-orientation. Castes can play a significant role in the requirements analysis and specification as well as design and implementation of multi-agent systems [29]. There is a most general caste, called AGENT, such that all castes are its sub-castes. The main body of a caste description contains a description of the structure of its states and actions, a description of its behaviour, and a description of its environment. The following gives the syntax of caste description in EBNF.

caste-description ::= 
Caste name \{ (Agent-description | caste-description) \} \{ instantiation ; \} 
[ environment-description ; ] 
[ structure-description ; ] \{ behavior-description ; \} 
end name

It can also be equivalently represented as follows in a graphic form similar to schemas in Z [30].

The clause 'Caste C ::= C₁, C₂, ..., Cₙ' specifies that caste C inherits the structure, behaviour and environment descriptions of existing castes C₁, C₂, ..., Cₙ. When no inherited caste is given in a caste specification, it is by default a sub-caste of the predefined caste AGENT. Thus, a binary inheritance relation \(<\) is defined on the castes C₁ and C₂, if C₁ is specified as a sub-caste of C₂. The inheritance relation is required to be a partial ordering on castes. Notice that, the inheritance relation is static in the sense that it does not change at run-time.

The relationship between agents and castes is similar to what is between objects and classes. What is different is that an agent can join into a caste or quit from a caste at run-time. If an agent is an instance of a caste, it has all the structural, behaviour and environment descriptions given in the caste's specification. The following gives the syntax of agent descriptions in SLABS and its equivalent graphic form.

agent-description ::= 
agent-name \{ (caste-name | ,\} \} \{ instantiation ; \} 
[ environment-description ; ] 
[ structure-description ; ] \{ behavior-description \} 
end name

When caste names are given in an agent description, the agent is an instance of the castes when it is created. If no caste name is given in an agent specification, the caste of the agent is by default AGENT. All the parameters in the specification of the caste must be instantiated. Moreover, it may have additional structural, behaviour and environment descriptions to extend its state space, to enhance its ability to take actions and to widen its view of the environment.

We define that a multi-agent system consists of a finite set of agents \{A₁, A₂, ..., Aₙ\}. These agents belong to a partially ordered set of castes C₁, C₂, ..., Cₙ. Let \(A \in C\) denote that agent A belongs to caste C at time t. We require that for all agents A and castes C and \(C\), for all times t, \(A \in C \land C \prec C' \Rightarrow A \in C\).

2.2 Environment

The SLABS language enables software engineers to explicitly specify the environment of an agent as a subset of the agents in the system that may affect its behaviour. This is another fundamental difference between agents (castes) and objects (classes). The syntax for the description of environments is given below.

Environment-description ::= 
ENVIRONMENT \{ (agent-name | All ; caste-name \{ variable : caste-name \} | )\}*

where an agent name indicates a specific agent in the system. 'All' means that all the agents of the caste have influence on its behaviour. As a template of agents, a caste may have parameters. The variables specified in the form of 'identifier: class-name' in the environment description are parameters. Such an identifier can be used as an agent name in the behaviour description of the caste. It indicates an agent in the caste when instantiated. The instantiation clause gives the details about how the parameters are instantiated.

Instantiation ::= \{ variable : agent-name \}*

The environment of an agent is a subset of the agents in the system \(Env_A \subseteq \{A₁, A₂, ..., Aₙ\}\), which may change at run-time. Let EC be an environment description expression. Then, the meaning
2.3 State and Action Spaces

In SLABS, the state space of an agent is described by a set of variables with keyword VAR. The set of actions is described by a set of identifiers with keyword ACTION. An action can have a number of parameters. An asterisk before the identifier indicates invisible variables and actions.

\[
\text{structure-description ::= } \begin{cases} \text{Var \{ \{ \* \} identifier: type; \} } \\
\text{Action \{ \{ \* \} action-declaration \{ f \}; \} } \\
\text{action-declaration ::= identifier \{ identifier \{ \{ parameter \} type \}; \} } 
\end{cases}
\]

2.4 Inheritance and Dynamic Caste

As informally stated above, a caste can inherit from multiple super-castes. We now formally define the semantics of inheritance and agent’s dynamic joining a caste or retreat from a caste. Let \( C \) be a caste specified as follows.

\[
\text{Caste } C \iff C(x_1: \text{expr}_1, \ldots, x_n: \text{expr}_n), \\
\ldots C(x_1: \text{expr}_1, \ldots, x_n: \text{expr}_n)
\]

ENVIRONMENT \( E_1, \ldots, E_n \):

VAR \( *v_1:T_1, \ldots, *v_n:T_m, u_1:S_1, \ldots, u_i:S_i \):

ACTION \( *A_1(p_1), \ldots, *A_i(p_i), B_1(q_1), \ldots, B_i(q_i) \):

RULES \( R_1, R_2, \ldots, R_k \):

Let \( \Sigma^{I}(C) \) and \( \Sigma^{E}(C) \) denote the visible and invisible state spaces of a caste \( C \), respectively. We define \( S^{I}(C) \) and \( S^{E}(C) \) as follows.

\[
S^{I}(C) = \{ C.v:T_1, \ldots, C.v:T_m, U_1:S_1, \ldots, U_i:S_i \}
\]

Let \( \Sigma^{V}(C) \) and \( \Sigma^{E}(C) \) denote the visible and invisible action spaces of caste \( C \), respectively. Similarly to state spaces, we define \( \Sigma^{V}(C) \) and \( \Sigma^{E}(C) \) as follows.

\[
\Sigma^{V}(C) = \{ C.A_1(p_1), \ldots, C.A_i(p_i) \}
\]

\[
\Sigma^{E}(C) = \{ C.B_1(q_1), \ldots, C.B_i(q_i) \}
\]

Let \( RULE(C) \) denote the set of behaviour rules of caste \( C \). We have that

\[
RULE(C) = \{ R_1, \ldots, R_k \} \cup \bigcup_{x=1}^{k} RULE(C_x).
\]

Let \( ENV(C) \) denote the environment for caste \( C \), we have that

\[
ENV(C) = \{ E_1, \ldots, E_n \} \cup \bigcup_{x=1}^{n} ENV(C_x).
\]

Let agent \( A \) be a member of castes \( C_1, C_2, \ldots, C_n \) at run-time \( t \). Let \( S_{A,t} \) denote the state space of agent \( A \) at time \( t \). Each state consists of two disjoint parts, the externally visible part and the internal part. The external part is visible for all agents in the system, while the internal part is not visible for any other agents in the system. Therefore, \( S_{A,t} = S_{A,t}^{V} \times S_{A,t}^{E} \), where \( S_{A,t}^{V} \) and \( S_{A,t}^{E} \) are the externally visible part and internal part of the state space, and defined as follows, respectively.

\[
S_{A,t}^{V} = \bigcup_{i=1}^{n} S_{i,t}^{V}(C_i), \quad \text{and} \quad S_{A,t}^{E} = \bigcup_{i=1}^{n} S_{i,t}^{E}(C_i).
\]

Let \( \Sigma_{A,t} \) denote the set of actions that an agent can take at time \( t \).

An action can also be either externally visible or internal (hence externally invisible). Assume that an agent cannot take two actions at the same time, thus \( \Sigma_{A,t} = \Sigma_{A,t}^{V} \cup \Sigma_{A,t}^{E} \), where \( \Sigma_{A,t}^{V} \cap \Sigma_{A,t}^{E} = \emptyset \).

\[
\Sigma_{A,t}^{V} = \bigcup_{i=1}^{n} \Sigma_{i,t}^{V}(C_i), \quad \text{and} \quad \Sigma_{A,t}^{E} = \bigcup_{i=1}^{n} \Sigma_{i,t}^{E}(C_i).
\]

The environment of agent \( A \) at run-time \( t \), denoted by \( ENV_{A,t} \), is defined as follows.

\[
ENV_{A,t} = \bigcup_{i=1}^{n} \{ EC | EC \in ENV(C_i) \}.
\]

The set of rules that agent \( A \) must satisfy at time \( t \) is the set \( RULE_{A,t} \), which is defined as follows.

\[
RULE_{A,t} = \bigcup_{i=1}^{n} RULE(C_i).
\]

Notice that, the value of the set \( ENV_{A,t} \) depend on time \( t \) for two reasons. First, the set of agents in a caste \( C \) may change when the agent joins into or retreats from a caste. Second, the agent \( A \) may join a caste or retreat from a caste so that the set \( \{ C_1, \ldots, C_n \} \) changes from time to time. For the sake of simplicity, when there is no risk of confusion, we will omit the subscript \( t \) in the sequel.

For example, consider the following specification.

Caste Person;

VAR *Surname: String, *Name: String, *Birthday: DATE, Nationality: String;

ACTION *Speak(String);

RULES

RULE1: [] \rightarrow Speak("Hello World"),

RULE2: [] \rightarrow Speak("My name is ", Name, Surname)

End Person;

Caste OBU-Students \iff Persons;

ENV Tutor: OBU-Staff;

VAR *Number: Integer;

Field: [Computing, Business]

ACTION *ChangeField(Field);

RULES

RULE1: [] \rightarrow ChangeField(NewField),

RULE2: [ChangeField(NewField)] \rightarrow !Field'=NewField, if Tutor: [ApproveChangeField(Self, NewField)]

End OBU-Students;

Caste OBU-Staffs \iff Persons;

ENV Tutee1, Tutee2: OBU-Students;

VAR *Number: Integer;

*Dept: [Computing, Math, Business]
ACTION *ApproveChangeField(OBU-Students, Field);
RULES <st1> : [] \rightarrow ApproveChangeField(St, NewField),
       If st\in OBU-Student: [ChangeField(NewField)],
       Where St=Tutee1 or St=Tutee2.

End OBU-Staffs;

The visible state variables of caste OBU-Students contains the following elements: {Person.Speak(String), Person.Name: String, Person.Nationality: String, OBU-Students.Number: Integer}. The two additional state variables OBU-Staffs.Number and Dept, and two new state variables OBU-Students.Number and OBU-Students.Field. This does not affect his other parts of state space and action space. Similarly, John can also quit from the OBU-Staff caste.

2.5 Behaviour

Agents behave in real-time concurrently and autonomously. To capture the real-time features, an agent’s behaviour is modelled by a set of sequences of events indexed by the time when the events happen.

2.5.1 Runs and Time

A run of a multi-agent system is a mapping from time to the set \( \prod_{i=1}^{n} \Sigma_{A_i} \times \Sigma_{\tau_i} \). The behaviour of a multi-agent system is defined to be a set \( R \) of possible runs. Instead of defining a fixed set of time moments, the set of time moments are characterised by a collection of properties.

Definition 1.

Let \( T \) be a non-empty subset of real numbers. \( T \) is said to be a time index set, or simply, the time, if

1) Bounded in the past, i.e. \( \exists t_0 \in T. \forall t \in T. (t_0 \leq t) \); \hspace{1cm} (14)
2) Unbounded in the future, i.e. \( \forall r \in \mathbb{R}. \exists t \in T. (t > r) \); \hspace{1cm} (15)
3) Uniformity, i.e. \( \forall t_1, t_2, t_1 \in T. (t_2 > t_1 \Rightarrow t_1 + t_2 - t_1 \in T) \). \hspace{1cm} (16)

The following lemma states that a time index set \( T \) can be characterised by two real numbers: the start time \( t_0 \) and the time resolution \( \rho \), where \( \rho > 0 \). Readers are referred to [28] for the proof of the lemma.

Lemma 1.

For all subsets \( T \) of real numbers that satisfy properties (14), (15) and (16), we have that either \( T = \{ t_0 | t_0 = t_0 + \rho n, n = 0, 1, 2, \ldots \} \) for some positive real number \( \rho \), or \( T = \{ r | r \in R \text{ and } r \geq t_0 \} \). In the former case, we say that the time index set \( T \) is discrete, and in the latter case, we say that \( T \) is continuous and \( \rho \) is called the resolution of the time index set \( T \).

For any given run \( r \) of the system, we say that a mapping \( h \) from \( T \) to \( S_{\tau_1} \times \Sigma_{\tau_2} \) is the run of agent \( A \) in the context of \( r \), if \( \forall t \in T. h(t) = r_A(t) \), where \( r_A(t) \) is the restriction of \( r(t) \) on \( S_{\tau_1} \times \Sigma_{\tau_2} \). Let \( r_A \) denote the run of agent \( A \) in the context of \( r \), and \( R_A = \{ r_A | r \in R \} \) denote the behaviour of agent \( A \) in the system.

2.5.2 Assumptions

We assume that a multi-agent system has the following properties. Instantaneous actions. We assume that actions are instantaneous, i.e. they take no time to complete.

Silent moments. We assume that an agent may take no action at a time moment \( t \). In such a case, we say that the agent is silent at time \( t \). For the sake of convenience, we treat silence as a special action and use the symbol \( t \) to denote silence. Therefore, we assume that for all agents \( A, \tau \in \Sigma_{A_i} \).

Separability. We assume that the actions taken by an agent are separable, i.e. for all runs \( r \), and all agents \( A \), there exists a real number \( c_A > 0 \) such that \( r_A(t) \neq \tau \) implies that for all \( x \in T \), \( t < x \leq t + e \Rightarrow r_A(t) = \tau \), where \( r_A(t) \) denotes the action taken by agent \( A \) at time moment \( t \) in the run \( r \). Consequently, an agent can take at most a countable number of non-silent actions in its lifetime.

Initial time and sleeping state. An agent can join the system at a time, say \( t_{init,A} \), later than the system's start time. We say that the agent is sleeping before time moment \( t_{init,A} \). We use a special symbol \( \perp \notin \Sigma_{A_i} \) to indicate such a state of an agent. Of course, we require that if an agent is sleeping, it will take no action but silence, i.e. \( \forall t \in T. (r_A(t) = \perp + r_A(t) = \tau) \), where \( r_A(t) \) denotes agent \( A \)’s time at moment \( t \) in the run \( r \). The initial time \( t_{init,A} \) of an agent \( A \) in a run \( r \) can be formally defined as the time moment \( t \in T \) that \( r_A(t) = \perp \land \forall t' \in T. (t' < t \Rightarrow r_A(t') = \perp) \).

2.5.3 Agent’s View of the Environment

The global state \( S_g \) of the system at any particular time moment \( t \) belongs to the set \( \prod_{i=1}^{n} S_{A_i} \times \Sigma_{\tau_i} \). However, each agent \( A \) can view the externally visible states and actions of the agents in \( Env_{A_i} \). In other words, an agent \( A \) can only view the part of the space \( S^v_{A_i} \times \Sigma^v_{\tau_i} \). Agent \( A \)’s view of the system state at a time moment \( t \) is defined as a mapping \( View_{A_i} \) from global state \( S_g \) to a value in \( S^v_{A_i} \times \Sigma^v_{\tau_i} \) as follows.

\[
View_{A_i} \left( \{ s_1, s_2, c_1, \ldots, s_v, c_v \} \right) = \{ s_i, c'_i \} \left( s_i, c_i \right) \right) \quad (17)
\]

where \( Env_{A_i} = \{ A_1, A_2, \ldots, A_n \} \). Let \( i = v \) implies that \( s_i = s_v \), \( c'_v = c_v \), if \( c_v \in \Sigma^v \), and \( c'_v = t \) if \( c_v \in \Sigma_{A_i} \). Because an agent’s view is only a part of the system’s global state, two different global states become equivalent from its view. The following formally
defines the relation.

\[ \forall x, y \in \prod_{i=1}^{k} S_i \times \Sigma_i . (x \approx_A y \iff \text{View}_A(x) = \text{View}_A(y)). \quad (18) \]

It is easy to see that the binary relation \( \approx_A \) is an equivalence relation.

### 2.5.4 Execution History

Although an agent may not be able to distinguish two global states, the histories of the runs leading to states may be different. An intelligent agent may decide to take different actions according to the history rather than only depending on the visible global state.

Let \( t \) be any given time moment. The **history of a run** up to \( t \), written as \( r^t \), is a mapping that is the restriction of \( r \) to the subset \( \{ x \mid t(x) \leq t \} \) of \( T \). The history of a run up to \( t \) in the view of an agent \( A \), denoted by \( \text{View}_A(r^t) \), is the mapping from the subset \( \{ x \mid t(x) \leq t \} \) of \( T \) of time moments to its views of the system's states in the run \( r \). It can be defined as follows.

\[
\text{View}_A(r^t) = \{ u \mid \text{View}_A(r(u)) \}, \quad \text{for all } u \in T \text{ and } t \leq t.
\]

Similarly, we define \( \text{View}_A(r) \) to be an agent \( A \)'s view of a run \( r \), and \( \text{View}_A(r^t) \) to be agent \( A \)'s view of agent \( B \)'s behaviour in a run \( r \). The equivalence relation defined on the state space can be extended to histories and runs as follows.

\[
\begin{align*}
\eta \approx_A \eta' & \iff \text{View}_A(\eta) = \text{View}_A(\eta') \\
(r_1 \downarrow t_1) \approx_A (r_1 \downarrow t_1) & \iff \text{View}_A(r_1 \downarrow t_1) = \text{View}_A(r_1 \downarrow t_1)
\end{align*}
\]

Before we finish this section, we introduce some further notation. Let \( A \) be any given agent in a multi-agent system. Let \( c_1, \ldots, c_m, \ldots \in \Sigma_i = \{ \} \) be the sequence of non-silent actions taken by agent \( A \) in a run \( r \) and \( t_1, t_2, \ldots, t_m \in T \) are the times of the actions, i.e. \( r^t = (t_1, t_2, \ldots, t_m) \). At a time moment \( t \in T \), we say that \( c_i \) is agent \( A \)'s current action, and \( c_{i+1} \) the next action, if \( t_{i+1} \leq t_i \), \( i \in \mathbb{N} \). We write

\[
\begin{align*}
\text{Current}(r) & = \langle t, s, c \rangle, \\
\text{Next}(r) & = \langle t, s, c \rangle, \text{ and} \\
\text{Events}(r) & = \langle t, s, c \rangle.
\end{align*}
\]

### 2.6 Specification of Behaviour

#### 2.6.1 Patterns of Behaviours

A pattern describes the behaviour of an agent by a sequence of observable state changes and observable actions. A pattern is written in the form of \( \{ p_1, p_2, \ldots, p_n \} \), where \( n \geq 0 \).

**Table 2. Meanings of the patterns**

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ $</td>
<td>The wild card, which matches with all actions</td>
</tr>
<tr>
<td>~</td>
<td>The silence event</td>
</tr>
</tbody>
</table>

Formally, let \( p \) be a pattern. We write \( B : r_f \downarrow t \models p \) to denote that from agent \( B \)'s viewpoint the behaviour of an agent \( A \) in a run \( r \) matches the pattern \( p \) at time moment \( t \). The relationship \( \models \) can be defined inductively as follows.

**Definition 2.**

We write \( B : r_f \downarrow t \models p \), if there is an assignment \( \alpha \) such that \( \text{B} : r_f \downarrow t \models_{\alpha} p \), which is inductively defined as follows.

- **\( B : r_f \downarrow t \models_{\alpha} [S] \)**, for all agents \( A, B \), runs \( r \) and time moments \( t \);
- **\( B : r_f \downarrow t \models_{\alpha} [t] \)**, if \( \text{View}_B(r_f \downarrow t) = t \);
- **\( B : r_f \downarrow t \models_{\alpha} [x] \)**, if \( \text{Current}(\text{View}_B(r_f \downarrow t)) = x(\alpha) \);
- **\( B : r_f \downarrow t \models_{\alpha} [\{ x \} \] \)**, if \( \text{Current}(\text{View}_B(r_f \downarrow t)) = (x, S, (\{ x \}, \{ x \})) \), \( S \) satisfies the predicate \( \alpha(p(\{ x \})) \), \( \alpha(t) = t \);
- **\( B : r_f \downarrow t \models_{\alpha} [p^k] \)**, if \( \text{Events}(\text{View}_B(r_f \downarrow t)) = \langle t_1, s_1, c_1 \rangle, \ldots, \langle t_2, s_2, c_2 \rangle, \ldots, \langle t_\alpha, s_\alpha, c_\alpha \rangle \), where \( v = \alpha(k) \), and for all \( i = 1, 2, \ldots, \alpha \);
- **\( B : r_f \downarrow t \models_{\alpha} [p] \)**;
- **\( B : r_f \downarrow t \models_{\alpha} [p \parallel \text{Constraint}] \)**, if \( B : r_f \downarrow t \models_{\alpha} p \) and \( \alpha(\text{Constraint}) \) is true.

Informally, \( B : r_f \downarrow t \models_{\alpha} p \) means that agent \( A \)'s behaviour in a run \( r \) matches a pattern \( p \) at time moment \( t \) from an agent \( B \)'s point of view under assignment \( \alpha \). An assignment \( \alpha \) for a set \( X \) of variables is a mapping that assigns values to variables in \( X \).

#### 2.6.2 Scenarios of Environment

The use of scenarios in agent oriented analysis and design has been proposed by a number of researchers, for example [21, 31, 19]. We define scenario as a set of typical combinations of the behaviours of related agents in the system. In addition to the pattern of individual agents' behaviour, SLABS also provides facilities to describe global situations of the whole system. The syntax of scenarios is given below.

**Scenario**: Agent-Name : pattern | atomic-predicate |

\[
\exists [\text{atomic-predicate}] \quad \text{Agent-Var} \in \text{Caste-Name} : \text{Pattern} \\
\forall \text{Agent-Var} \in \text{Caste-Name} : \text{Pattern} \\
\text{Scenario} \land \text{Scenario} \lor \text{Scenario} \land \text{Scenario} \land \text{Scenario}
\]

An atomic predicate in a scenario can be an expression in one of the following forms:

- **Agent :: Caste**, the agent is in the caste at that time;
- **AgentA :: AgentB** (or **AgentB :: AgentA**), the identifiers indicates the same (or different) agent;
- **A set relation expression**, which may contain expressions in the form of \( \{ x \in \text{Caste} \mid x : \text{Pattern} \} \), which is the set of agents
whose behaviour matches the pattern; or

- An arithmetic relation, which may contain an expression in the form of $\mu X \in In\{\text{Caste, Pattern}\}$, which is the number of agents in the caste whose behaviour matches the pattern.

The semantics of scenario descriptions are given in Table 3.

### Table 3. Semantics of scenario descriptions

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A \vdash P$</td>
<td>The situation when agent $A$'s behaviour matches pattern $P$</td>
</tr>
<tr>
<td>$\forall X \in C: P$</td>
<td>The situation when the behaviours of all agents in caste $C$ match pattern $P$</td>
</tr>
<tr>
<td>$\exists m \ni X \in C: P$</td>
<td>The situation when there exists at least $m$ agents in caste $C$ whose behaviour matches pattern $P$ where the default value of the optional expression $m$ is 1</td>
</tr>
<tr>
<td>$S_1 &amp; S_2$</td>
<td>The situation when both scenario $S_1$ and scenario $S_2$ are true</td>
</tr>
<tr>
<td>$S_1 \lor S_2$</td>
<td>The situation when either scenario $S_1$ or scenario $S_2$ or both are true</td>
</tr>
<tr>
<td>$\lnot S$</td>
<td>The situation when scenario $S$ is not true</td>
</tr>
</tbody>
</table>

The following are two examples of scenarios.

(1) $\exists p \in \text{Parties}: p \in \text{bacco} \land [\text{nominate-president(Bush)}] \land \text{bacco}(\text{March/2000})$.

It describes the situation that at least one agent in the caste called Parties took the action nominate-president(Bush) at the time of March 2000.

(2) $\left(\mu x \in \text{Voter}: [\text{vote(Bush)}] \lor \mu x \in \text{Voter}: [\text{vote(Gore)}]\right)$

It describes the situation that there are more agents in the caste Voter who took the action of vote(Brush) than those in the caste who took the action of vote(Gore).

Let $Sc$ be a scenario. We write $A \vdash r \downarrow t \models Sc$ to denote that from agent $A$'s point of view, the scenario $Sc$ occurs at time moment $t$ in a run $r$.

**Definition 3.**

From an agent $A$'s point of view, a scenario $Sc$ occurs at time moment $t$ in a run $r$, iff $A \vdash r \downarrow t \models Sc$, which is inductively defined as follows.

- $A \vdash r \downarrow t \models B: p \iff A \vdash r \downarrow t \models p$; (22)
- $A \vdash r \downarrow t \models Sc_1 \land Sc_2 \iff A \vdash r \downarrow t \models Sc_1$ and $A \vdash r \downarrow t \models Sc_2$; (23)
- $A \vdash r \downarrow t \models \lnot Sc \iff A \vdash r \downarrow t \models Sc$ is not true; (24)
- $A \vdash r \downarrow t \models \forall x \in G(x: Sc) \iff A \vdash r \downarrow t \models Sc$, for all $x \in G$; (25)
- $A \vdash r \downarrow t \models \exists x \in G(x: Sc) \iff A \vdash r \downarrow t \models Sc$, for some $x \in G$ (26)

### 2.6.3 Rules

In SLABS, an agent's behaviour is defined by a set of transition rules to describe its responses to environment scenarios.

**Behaviour-rule**: $\triangleright=

\langle rule-name \rangle \[ pattern \rightarrow \langle prob \rangle \rightarrow \langle event \rangle \[ Scenario \] [ where \ pre-cond ]$

In a behaviour rule, the pattern on the left-hand-side of the $\rightarrow$ symbol describes the pattern of the agent's previous behaviour. The scenario describes the situation in the environment, which specifies the behaviours of the agents in its environment. The where-clause is the pre-condition of the action to be taken by the agent. The event on the right-hand-side of $\rightarrow$ symbol is the action to be taken when the scenario happens and if the pre-condition is satisfied. The agent may have a non-deterministic behaviour. The expression $\text{prob}$ in a behaviour rule is an expression that defines the probability for the agent to take the specified action on the scenario. SLABS also allows specifications of non-deterministic behaviours without giving the probability distribution. In such cases, the probability expression is omitted. It means that the probability is greater than 0 and less than or equal to 1.

Let $R$ be the set of runs in a formal model of agent-based system. To define the semantics of rules, we first define a probabilistic space as follows.

For each scenario $Sc$, agent $A$, and constraint $Cn(r, t) \rightarrow \{tt, ff\}$, we define $R_{\alpha}(Sc, A, Cn)$ as a subset of histories $H = \{r \in \text{Run} \mid r \in T\}$ such that

$R_{\alpha}(Sc, A, Cn) = \{ r \in \text{Run} \mid r \in T, A \vdash r \downarrow t \models Sc, Cn(r, t) \}$ (27)

Let $H^*$ be the set that contains $H$ and all the subsets in the form of $R_{\alpha}(Sc, A, Cn)$ and closed under set complement, finite intersections and countable unions. Therefore, $H^*$ constitutes a $\sigma$-field. A probabilistic space can then be constructed over the $\sigma$-field $H^*$ by associating a probabilistic distribution over $H^*$. Let $\Pr(R_{\alpha}(Sc, A, Cn))$ be the probability that the scenario $Sc$ with constraint $Cn$ occurs from agent $A$'s point of view. Notice that agent $A$'s behaviour matches a pattern $p$ can be expressed equivalently as a scenario $(A; p)$, the order pair $<R, Pr>$ is called the probabilistic model of the agent-based system.

**Definition 4.**

Let $R_{\alpha} = \langle p \mid (exp) \rightarrow e \rangle$ if $Sc$ where $Cn$ be a rule for agent $A$, where $Sc$ is a scenario and $Cn$ is a constraint. We say that in a probabilistic agent-based system $<R, Pr>$ the agent $A$'s behaviours satisfy the rule $R_{\alpha}$ and write $<R, Pr>: A \vdash R_{\alpha}$, if

$\Pr(R_{\alpha}(A; p\#e, A, True)) \mid R_{\alpha}(Sc \# (A; p), A, Cn) = \exp$, (28)

where $p\#e = [p_1, p_2, ..., p_n, e]$, if $p = [p_1, p_2, ..., p_n]$. □

### 3. EXAMPLE

A number of examples of intelligent agents have been specified in SLABS in our previous papers, which include the Mae's personal assistant Maxim [27], Ants [32], a simple communication protocol, speech-act [28] and the evolutionary multi-agent ecosystem Amalthea [35], etc. In this paper, we use an example of a distributed algorithm to demonstrate how SLABS supports the modularity and composability of formal specifications.

The algorithm is for synchronisation of the accesses to critical regions in distributed systems. The following is the original informal specification of the algorithm given in the textbook [33] (Page 267).

When a process wants to enter a critical region, it builds a message containing the name of the critical region it wants to enter, its process number, and the current time. It then sends the message to all other processes. The sending of messages is assumed to be reliable; that is, every message is acknowledged. Reliable group communication if available, can be used instead of individual messages.

When a process receives a request message from another process, the action it takes depends on its state with respect to
the critical region named in the message. Three cases have to be distinguished:

1. If the receiver is not in the critical region and does not want to enter it, it sends back an OK message to the sender;
2. If the receiver is already in the critical region, it does not reply. Instead, it queues the request.
3. If the receiver wants to enter the critical region but has not yet done so, it compares the timestamp in the incoming message with the one contained in the message that it has sent everyone. The lowest one wins. If the incoming message is lower, the receiver sends back an OK message. If its own message has a lower timestamp, the receiver queues the incoming request and sends nothing.

After sending out requests asking permission to enter a critical region, a process sits back and waits until everyone else has given permission. As soon as all the permissions are in, it may enter the critical region. When it exits the critical region, it sends OK message to all processes on its queue and deletes them all from the queue.

It is assumed that processes are executed in a distributed system concurrently. The communications between the processes are reliable. They all use the synchronisation algorithm to access shared resources as critical regions. Regarding each process as an instantiations and used through instances, i.e. agents. This example shows that specifications can be modularised by using caste with parameters and composed through sub-castes and instantiations and used through instances, i.e. agents.

4. CONCLUSION

In this paper, we presented the formal specification language SLABS for multi-agent systems. The SLABS language integrates a number of novel language facilities that support the development of agent-based systems, especially the specification of such systems. Among these facilities, the notion of caste plays a crucial role. A caste represents a set of agents in a multi-agent system that have same capability of performing certain tasks and have same behaviour characteristics. Such common capability and behaviour can be the capability of speaking a language, using an ontology, following a communication and/or collaboration protocol, and so on. It is a notion that generalises the notion of types in data type and the notion of classes in object-oriented paradigm. This facility can be effectively used to specify or implement a number of notions proposed in agent-oriented methodologies, such as the notions of role, team, agent society, organisation, and so on. For example, a caste can be the set of agents playing the same role in the system. However, agents of the same caste can also play different roles especially when agents form teams dynamically and determine their roles at run time. Based on the caste facility, a number of other facilities in SLABS are defined. For example, the environment of an agent can be described as the agents of certain casts. A global scenario of a multi-agent system can be described as the patterns of the behaviours of the agents of a certain caste. The example systems and features of agent-based systems specified in SLABS have shown that these facilities are powerful and useful for the formal specification of agents in various models and theories in a modular and composable way.

4.1 Related Work

The model of software agents used in this paper is closely related to the work by Lesperance et al [34], which also focused on the
actions of agents. However, there are two significant differences. Firstly, they consider objects and agents are different types of entities, while we consider them as the same type of encapsulated computational entities. As argued in [28], we consider objects as a degenerate form of agents that obey simple behaviour rules and open to its environment so that everything in the environment can affect its behaviour. As a consequence of regarding objects as deferent entities from agents, they allow an agent to change the state of objects in the environment while we only allow an agent to modify its own state. Secondly, the most important difference is, of course, there is no notion of caste or any similar facility in their system.

The notion of groups of agents has been used in a number of researches on the multimodality logic of rationale agents, such as in Wooldridge's work [11], etc. However, such notion of groups of agents is significantly different from the notion of caste, because there is neither inheritance relationship between the groups, nor dynamic instance relationship between an agent and a group. Their only relationship is the membership relationship.

Many agent development systems are based on object-oriented programming. Hence, there is a native and primitive form of castes as classes in OO paradigm. However, although agents can be regarded as evolved from object and caste as evolved from class, there are significant differences between agents and objects, and thus between caste and class. Therefore, the new notion deserves a new name.

In our previous papers [27~29], an agent's membership to a caste is statically determined by agent description. Static membership has a number of advantages, especially its simplicity and easiness to prove the properties of agents. Examples have shown that introducing some state variables to represent the role that an agent is playing can specify dynamic team formation [29]. In this paper, we revised the notion of caste to allow a dynamic membership facility in order to specify and implement dynamic team formation. The formal and informal definitions of the language SLABS given in this paper supersedes our previous ones. An advantage of this approach is that the dynamic formation of a team can be explicitly specified. It is natural to specify the dynamic process of evolution in a multi-agent system [35].

Another design decision that we faced in the design of SLABS was whether we should allow redefinition of behaviour rules in the specification of sub-castes. An advantage of disabling redefinition is that provable properties of a supper caste are inherited by all sub-castes. The example given in this paper shows that it enables composable modular specifications.

4.2 Further Work

There are a number of open problems that need further investigation. Although the language facilities in SLABS, such as caste and scenario, were first introduced as a specification facility, we believe that they can be easily adopted in an agent-oriented programming language for the implementation of multi-agent systems. How to implement these facilities is an important issue in the design and implementation of agent-oriented programming languages.

The design of SLABS is aimed to support as many agent-oriented methodologies as possible. How to link from such methodologies to formal specifications in SLABS deserves further investigation. We are currently working on tools and graphic notations to support the development of formal specifications in SLABS. In [35], a process and a diagrammatic notation for modelling multi-agent systems are proposed so that formal specifications in SLABS can be derived from models of multi-agent systems represented in diagrams.

ACKNOWLEDGEMENT

The author is most grateful to his colleagues at Oxford Brookes University, especially Mr. Ken Brownsey, Prof. David Duce, Ms. Sue Greenwood, Mr. John Nealon, Dr. Nick Wilson, et al., for discussions on agent technology and many related subjects. The author would also like to thank Prof. Huaglory Tianfield for many invaluable discussions.

REFERENCES

[17] Kimny, D., Geogreff, M., and Rao, A. A methodology and
modelling technology for systems of BDI agents, in Agents
Breaking Away: Proc. of MAAMAW ’96, LNAI 1038,
Springer-Verlag, 1996.

[18] Moulin, B. and Cloutier, L. Collaborative work based on
multiagent architectures: a methodological perspective, in
Soft Computing: Fuzzy Logic, Neural Networks and
Distributed Artificial Intelligence, Aminzadeh, F. and

[19] Moulin, B., and Brassard, M. A scenario-based design
method and an environment for the development of
multiagent systems, in First Australian Workshop on
Distributed Artificial Intelligence, Lüksuse, D. and Zhang C.

for agent-oriented analysis and design, in Proc. of ACM
Third International Conference on Autonomous Agents,
Seattle (WA, USA, 1999) 69-76.

[21] Iglesias, C. A., Garijo, M. Gonzalez, J. C. A Survey of
Agent-Oriented Methodologies, in Intelligent Agents V,

[22] Conrad, S., Saake, G., Turker, C. Towards an Agent-
Oriented Framework for Specification of Information
Systems, in [13], 1999, 57-73.

[23] D'Inverno, M., Kinny, D., Luck, M., and Wooldridge, M. A
formal specification of dMARS, in Intelligent Agents IV,
Singh, M. P., Rao, A. Wooldridge, M. (eds.) LNAI 1365,

[24] Luck, M. and d'Inverno, M. A formal framework for agency
and autonomy, in Proc. of First Int Conf on Multi-Agent


[26] Jennings, N. R. Agent-Oriented Software Engineering, in
Multi-Agent System Engineering, Proc of 9th European
Workshop on Modelling Autonomous Agents in a Multi-
Agent World (Valencia, Spain, June/July 1999) Garijo, F. J.,

[27] Zhu, H. Formal Specification of Agent Behaviour through
Environment Scenarios, Formal Aspects of Agent-Based
Systems, Rash, J. et al. (eds.), LNCS 1871, Springer, 263-
277.

Agent-Based Systems, Int. J. of Software Engineering and

[29] Zhu, H. The role of caste in formal specification of MAS, in


[31] Iglesias, C. A., Garijo, M., Gonzalez, J. C., Velasco, J. R.
Analysis And Design of Multiagent Systems Using MAS-
Common KADS, in Intelligent Agents IV, Singh, M. P., Rao,
313-327.

[32] Zhu, H. A formal specification language for MAS
engineering, in Proc. of 2nd International Workshop on
Agent-Oriented Software Engineering (May 29, 2001).

[33] Tanenbaum, A. S. and van Steen, M., Distributed Systems:

[34] Lesperance, Y., levesque, H. J., Lin, F., Marcu, D., Reiter, R.
and Scherl, R. Foundations of logical approach to agent
programming, in Intelligent Agents II, Wooldridge, M.,