Function-Behaviour-Structure: A Model for Social Situated Agents

John S. Gero and Udo Kannengiesser
University of Sydney
Sydney NSW 2006, Australia
{john, udo}@arch.usyd.edu.au

Abstract

This paper proposes a comprehensive schema to represent a social situated agent, based on the notions of function, behaviour and structure. Although this schema was originally developed to represent knowledge about design objects, it is sufficiently general to also describe knowledge about agents. This paper shows how such an FBS view is useful to support social interaction of situated agents.

1 Introduction

In the last few years, multi-agent systems (MASs) have been of growing interest for the Artificial Intelligence (AI) community [Wooldridge, 2002]. The reason for this can be seen in their modularity and their ability to deal with complex tasks in dynamic environments. A fundamental issue of MASs is how agents coordinate their local actions to achieve useful global behaviour. The more autonomous the agents, the more important is their ability to reason about each other's role in the achievement of a global goal, about each other's capabilities and actions and sometimes even about each other's reasoning (i.e. goals, beliefs) [Castelfranchi, 1998]. If an agent is to reason and act socially, it needs formal internal representations of other agents, which cover most of these aspects.

Agents that explicitly deal with internal representations are generally labelled cognitive. A large number of cognitive agents are based on the belief-desire-intention (BDI) model [Rao and Georgeff, 1991], differentiating their internal state into formal representations of beliefs, desires and intentions. While the BDI model exclusively deals with cognitive states, it does not account for an agent's actions or roles. Therefore its scope is limited to providing a framework for an agent architecture rather than for modelling social knowledge within an agent architecture.

In this paper we propose a comprehensive representation schema for social knowledge based on the notions of function (F), behaviour (B) and structure (S). This schema unites all the aspects of agents in one set of constructs, which is both uniform and consistent in its applicability. We outline how the FBS view supports situated social interaction within open MAS environments.

2 Situated Agents

2.1 Situatedness

In AI, the notion of situated agents has commonly been equated with "reactive" or "behavioural" agents, such as the ones used in Brooks' [1986] subsumption architecture. These "situated" agents base their decisions and actions on the current state of the environment without using any explicit representations of it. They use predefined stimulus-response mappings that have been shown to be useful in predictable (notably physical) environments. However, in more complex environments we need to broaden the scope for situatedness to include some form of internal representation. These representations are grounded in the agent's interactions with the environment [Bickhard and Campbell, 1996] rather than encoded by a third party. Situatedness, in the way we understand it, thus allows for autonomy and rationality of the agent [Smith and Gero, 2000].

A situated agent adapts its behaviour to changes in its environment based on its current goals, its knowledge and its interpretation of the environment [Clancey, 1997]. As a consequence, the agent can be exposed to different environments and produce appropriate responses. Situatedness has been recognised as an essential feature of designing (carried out by human or computational agents), which is an activity that inherently changes the world in which it operates [Gero, 1998]. This claim has been supported by empirical studies of human designers [Schön and Wiggins, 1992; Suwa *et al.*, 1999], which have characterised designing as an interaction of the designer with their environment: After performing actions to change the environment (e.g. by producing sketches of the design object), the designer observes and interprets the results of these actions and then decides on new actions to be executed on the environment. This means that the designer's concepts may change according to what they are "seeing", which itself is a function of what they have done.

Gero and Fujii [2000] have developed a framework for situated cognition in an agent, which describes the agent's interpretation of its environment as interconnected sensation, perception and conception processes. Each of them consists of two parallel processes that interact with each other: A *push process* (or data-driven process), where the production of an internal representation is driven ("pushed") by the environment, and a *pull process* (or expectation-driven process), where the interpretation is driven ("pulled") by some of the agent's current concepts, which has the effect that the interpreted environment is biased to match the current expectations.

The environment that is interpreted can be external or internal to the agent. The situated interpretation of the internal environment accounts for constructive memory, which is a notion whose foundations can be traced back to the work of Dewey [1896]. Quoting Dewey, constructive memory holds that "sequences of acts are composed such that subsequent experiences categorize and hence give meaning to what was experienced before". The implication of this is that memory is not laid down and fixed at the time of the original sensate experience but is a function of what comes later as well. Memories can therefore be viewed as being constructed in response to a specific demand, based on the original experience as well as the situation pertaining at the time of the demand for this memory. Each memory, after it has been constructed, is added to the agent's knowledge and is now available to be used later, when new demands require the construction of further memories. These new memories can be viewed as new interpretations of the agent's augmented knowledge.

2.2 A General Architecture of a Situated Agent

Gero's and Fujii's [2000] general architecture of a situated agent accounts for both situated cognition and constructive memory. Its major components and processes are depicted in Figure 1.

The agent's sensors monitor the environment to produce sense-data relevant for the agent. The sensors receive biases from the perceptors, which "pull" the sense-data to produce percepts. Percepts are grounded patterns of invariance over interactive experiences. Perception is driven both by sense-data and biases from the conceptor, which "pulls" the percepts to produce concepts. Concepts are grounded in the percepts as well as possible future interactions with the environment. The hypothesizor identifies mismatches between the current and desired situation and decides on actions that when executed are likely to reduce or eliminate that mismatch. Based on the hypothesized action, the action activator decides on a sequence of operations to be executed on the environment by the effectors.

The agent architecture depicted allows implementing different degrees of sophistication in an agent's reasoning and consequently different degrees of flexibility in its actions. A reflexive agent uses the raw sensory input data from the environment to generate a preprogrammed response. This does not involve any reasoning or intelligence. A reactive agent uses percepts to activate its actions, which can be viewed as a limited form of intelligence constrained by a fixed set of goals. The agent's perception will vary as a consequence of its experience. A reflective agent constructs concepts based on its current goals and beliefs and uses them to hypothesize possible desired external states and propose alternate actions that will achieve those desired states through its effectors. The agent's concepts may change as a consequence of its experience.

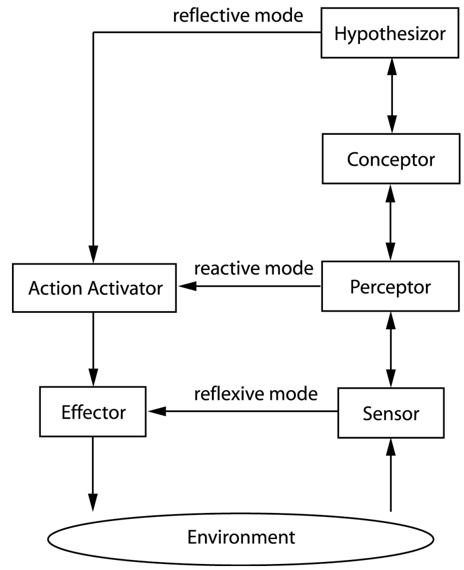

Figure 1. A general architecture for a situated agent allowing different modes of reasoning.

3 An FBS View of Objects

Gero [1990] has introduced a representation schema for design knowledge, i.e. knowledge about existing (physical) or to be designed (imaginary) objects, based on the notions of function (F), behaviour (B) and structure (S):

- *Function* (F) describes the teleology of an object.
- *Behaviour* (B) describes the attributes that are derived or expected to be derived from the structure (S) of an object.
- *Structure* (S) describes the components of an object and their relationships.

An agent that uses the FBS schema to represent a particular object constructs connections between function, behaviour and structure through experience with the object. Specifically, the agent ascribes function to behaviour and derives behaviour from structure. A direct connection between function and structure, however, is not established.

Take the example of a mobile phone as the design object. Figure 2 shows some of the structural properties of the Nokia 6100 model, which have been designed to fulfil various functions via appropriate behaviours. These functions include, for example, ease of navigation. A behaviour that has been chosen to achieve this function is the number of degrees of freedom (here two) when scrolling through the menu options. The structure that allows this behaviour includes a 4-way scroll key besides other structural components not shown in the figure (e.g. electronic and software elements). Another potential function of this design may be the ease with which the phone can be carried in the owner's pocket. Here a relevant behaviour is a reduced volume of the phone case, which is established by sufficiently small values for the structure variables length, width and thickness.

Figure 2. Some structural properties of a Nokia 6100 mobile phone.

The FBS schema is sufficiently general to cover all interpretations of objects. This is particularly important for situated agents, as their views of the same object are different when dissimilar goals and dissimilar prior knowledge are used as interpretation biases. This is most obvious for interpretations of the teleology (F) of an object, since it is closely connected to the agent's current goals that are likely to change for one agent as well as differ for different agents. However, the interpretation of behaviour (B) and structure (S) of an object is also situated. Take our mobile phone example, we only highlighted those properties of the structure that are visible for the user. However, the structure relevant for the electrical engineer (or for an interested user) may include the electronic circuits in the mobile phone. This more technical knowledge about the structure can then be used to account for an additional range of behaviours, such as the number of ring tones, the Specific Absorption Rate (SAR) or the ability to browse Wireless Application Protocol (WAP). Another, less mechatronic example is a simple house: For an agent specialised in architectural design, the structure (S) of a house may be composed of a configuration of spaces, whereas an agent specialised in structural engineering may view the structure (S) of the same house as a configuration of walls and floors (spaces, as they are derived from this structure, are interpreted as behaviour (B)). However an agent interprets an object, the resulting internal representation can always be modelled as an FBS view.

Once a number of experiences with objects has been gained and represented in the FBS form, the agent is able to generalise by clustering sets of like experiences. When the agent needs to access its knowledge about a particular object, it can derive a large part of this knowledge without much computational effort from its generalized experiences. This derived knowledge may even add (default) assumptions about an object where information gained from directly interacting with that object is missing.

The FBS schema provides a uniform set of constructs to model objects at all levels of generality and thus significantly supports this generalisation. Gero [1990] has introduced so-called design prototypes that represent generalised design knowledge, from which specific design objects can be instantiated. Design prototypes are useful for a design agent to start designing even if only incomplete information is available about the function, behaviour and structure of the object to be designed.

4 An FBS View of Agents

The FBS view is sufficiently general to comprehensively represent all kinds of objects. Even processes can be easily thought of as objects (this is what object-oriented software engineering does) and can similarly be represented in terms of F, B and S. One particular class of "objects" comprises agents. Figure 3 shows how an agent can be represented using the FBS view.

The function (F) of an agent is the purpose that an observer ascribes to its behaviour and usually refers to the agent's role in some environment.

An obvious interpretation of an agent's behaviour (B) is how the agent acts given a set of conditions (which are shown in Figure 3 as input from the environment). This corresponds to the notion of a "black-box" or "input-output" view of the agent.

As illustrated, we distinguish two kinds of structure (S) of an agent. One refers to the "fixed" parts of the agent (S^f), i.e. those components or processes that are given by design and that are not subject to significant change. This type of structure is the same as for objects that have no agency, and typically includes "visible" components such as the sensors and effectors of the agent. The other kind of structure refers to the "situated" parts of the agent (S^s), i.e. those internal representations or processes that are constructed by the agent's interaction with its environment. The situated structure of an agent may be

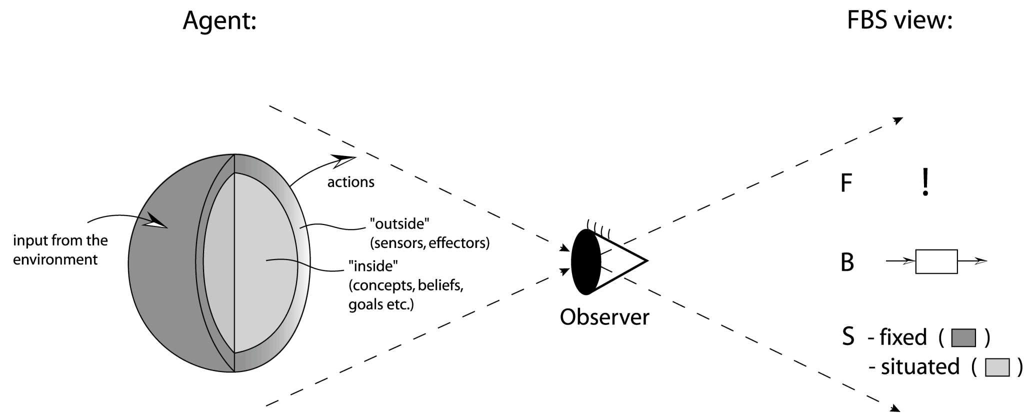

Figure 3. An FBS view of an agent.

interpreted as its grounded concepts, beliefs, goals etc., which are often hidden from the observer (and thus depicted "inside" the fixed structure in Figure 3).

It is apparent that the notion of structure (S) is the most complex class of properties for an agent. The agent's structure can be instantiated in several ways depending on the particular observer or the particular point of view. There are three important types of observers, whose knowledge about an agent's structure differs considerably from each other both quantitatively and qualitatively: The designer (developer) of the agent, other agents that interact with it and the agent itself.

If the observer is the designer of the agent, obviously a high amount of information about the agent's structure is available. This information includes all aspects of the fixed structure (S^f) of the agent architecture, such as the one shown in Figure 1, as well as its implementation. In addition, the designer might also have some information about the situated structure (S^s) of the agent, either by directly accessing and tracking its internal states and processes, or via inferring them from the agent's behaviour (B). For example, a particular behavioural change can make the designer suppose that a particular set of internal states (such as percepts, concepts, actions etc.) or processes has been activated.

If the observer is another agent interacting with the agent, its amount of knowledge about the agent's structure is rather small compared to the one available to the designer of that agent. In the most extreme case, this knowledge only consists of the most crucial parts of the agent's (fixed) interfaces with its environment, which are typically the agent's sensors and effectors. For simple forms of behavioural interaction (e.g. in subsumption architectures), this small amount of S^f knowledge can be sufficient. When there are higher demands on multiple situated agents coordinating their actions, it is desirable for the agents to shed some more light into the "black box" structure of their models of each other. This additional knowledge usually does not include the hitherto unknown parts of the agents' architectures; rather it refers to mentalistic (intentional) notions, as they facilitate explaining and predicting agent behaviour without having to know its implementation [Dennett, 1987]. Mentalistic notions can be viewed as referring to the fixed (S^f) or to the situated (S^s) structure of the agent. Examples of mentalistic properties describing the fixed structure include identifiers such as "reflexive", "reactive" and "reflective" (Figure 1). Examples for the situated structure of an agent include interpretations such as beliefs, desires and intentions (BDI) or function, behaviour and structure (FBS) of other objects or agents.

The observer can also be the agent itself. The structure that the agent "sees" of itself is obviously very rich and detailed. Constructing an FBS view of itself generally allows the agent to reflect on its own function, behaviour and structure in order to better adapt future interactions with its environment. Similar to its view of the structure of other agents that it interacts with, its view of its own structure relates to its "mental" states rather than its architecture (unless the agent wants to re-design or clone itself).

It is not hard to see that the situated structure (S^s) of an agent serves as a filter to access other agents' FBS views. This allows modelling nested social beliefs to an arbitrary order. For instance, S^s_0 (S^s_1 (F_2)) denotes agent 0's (first-order) belief about agent 1's view of agent 2's function, S^s_0 (S^s_1 (S^s_0 (B_1))) denotes agent 0's (second-order) belief about agent 1's belief about agent 0's view of agent 1's behaviour, etc.

5 FBS for Situated Social Interaction

We have outlined how the scope of the FBS schema as a knowledge representation for (design) objects can be extended to represent agents. We have also given a first glimpse of how the FBS schema accounts for situatedness, by illustrating its instantiation depending on the perspective of the observer. Different contexts and purposes shape this perspective.

The context for which the FBS schema was originally developed is designing. It has been used as the basis for modelling designing as a set of processes that ultimately transform a set of functional requirements (F) into structure (S) via behaviour (B) [Gero, 1990; Gero and Kannengiesser, 2002]. It has also been used in the development of design agents [Kulinski and Gero, 2001] and design agent memories [Liew and Gero, 2002]. Viewing agents as design objects, the FBS schema can similarly be employed to understand and support the process of designing agents.

Designing is just one instance of an agent's interaction with its environment, where the environment that is acted upon consists of passive objects that have no agency at the time of the interaction. If the environment of an agent contains other agents whose situated agency (i.e. their autonomy and rationality) is taken into account by the agent for devising its actions, then we can speak of social (inter-) action [Castelfranchi, 1998]. While designing is concerned with generating the fixed structure (S^f) of an object or an agent, social interaction commonly deals with generating (or changing) an agent's situated structure (S^s). Both kinds of interaction, however, involve some view on structure (S) and aim at producing a behaviour (B) that is to bring about a function (F). Thus the FBS schema can account for all interactive contexts. As the FBS schema has been well explored in the context of designing, we now focus on examining its use in the context of socially interacting situated agents.

As an interaction becomes social only if the interacting agents account for each other's situatedness, a purely behavioural interaction, as a reduced form of situatedness, does not fulfil this criterion. As mentioned in the previous section, the FBS view of a socially interacting agent should contain a minimum amount of S^s knowledge. The social action occurs in a way that ensures a shared understanding of the meaning of the interaction. In other words, there must be shared knowledge among socially interacting agents.

Most MASs have implemented shared knowledge as common ontologies [Gruber, 1993]. They pre-define all the knowledge that is relevant for meaningful interaction (communication) and encode them into the agents. However, these agents are then no longer completely autonomous and thus not completely situated. In contrast, a situated approach is based on the ability of each individual agent to construct appropriate models of the agent(s) it interacts with. These models, represented as FBS views, are then used to adapt the interaction to the respective agent. Common ground [Clark, 1992] between interacting situated agents is the notion that affirms that these agents have constructed appropriate models of each other to an extent sufficient for the purpose of the current interaction. Common ground is thus an emergent property of the current interaction rather than static knowledge that is not affected by its use in the current situation. Figure 4 depicts the pairs of FBS models that have to adequately match to establish the common ground between two agents.

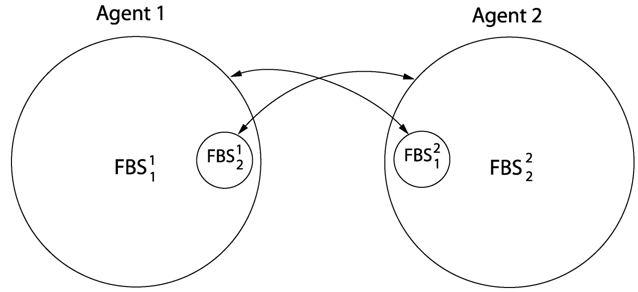

Figure 4. Pairs of matching FBS models that establish the common ground of two agents.

It has often been suggested that for an agent to evaluate the entire common ground it shares with another agent it would have to go through a process of infinite recursion [Clark and Marshall, 1981]. It would have to construct a belief about the other agent's belief about the common ground, which itself entails constructing a belief about the agent's belief about the common ground, and so on. However, situated agents avoid this recursion by "satisficing" [Simon, 1969] the construction of their models of each other according to the needs of the current interaction. Even non-social interactions based on modelling the agents' behaviour (B) without considering their structure (S) (e.g. habitual interaction developed to simplify what was formerly social interaction) could then be interpreted as conforming to common ground. The FBS schema supports this view of satisficed common ground by clearly distinguishing between the teleology, the behaviour and the (mentalistic) structure of an entity or agent.

In most cases, however, interaction between situated agents requires the construction of FBS views that include a sufficient amount of structure (S), notably the "hidden" situated structure (S^s). The agent can generally use two sources of information to access S^s. The first one includes those parts of S^s that the other agent makes directly available by communicating them. The second one includes generalisations over a set of experiences with other agents. Cues for constructing these generalisations are often provided by observations of the other agent's behaviour (B). Usually both sources of information are employed, with generalisations typically providing default assumptions when only incomplete information is available from direct communication.

Figure 5 shows an agent (0) having constructed FBS models of other agents (1, 2, 3 and 4). As the differently sized FBS models in the figure suggest, some agents (1 and 2) are better known (grounded) than others (3 and 4), and the best-known agent is certainly agent 0 itself. When the agent wants to interact with one of the other agents but has too little knowledge about that agent (here agent 4) to establish sufficient common ground for this interaction, it complements the existing FBS model with assumptions reflecting its generalised knowledge about similar agents. This generalised knowledge is derived mainly from those instances the agent (0) is most familiar with, as indicated by the different weights of the arrows in Figure 5, which principally includes the agent (0) itself. When a new, previously unknown agent (5) enters agent 0's environment, the generalised knowledge may still suffice to construct an adequate FBS model of that agent using the generalised knowledge about F, B and S individually and their relationships. If there is a conflict between the generalised knowledge and the interactions with a specific agent then a specialised FBS view of that agent needs to be constructed.

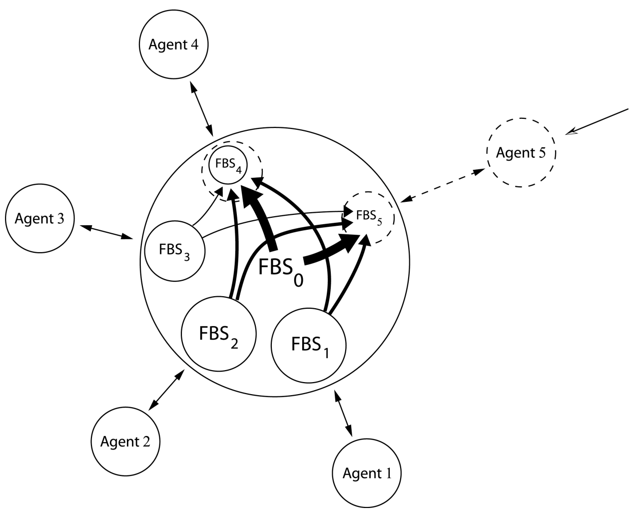

Figure 5. New FBS models are constructed using generalisations of previously constructed FBS models.

6 Discussion

We have presented a formal representation schema for agents, which is sufficiently comprehensive to capture all aspects of an agent. In contrast to other representation schemas (such as BDI) that focus on the architecture (structure) of agents, the FBS model can explicitly deal with the actions (behaviour) and roles (function) of agents. It can therefore serve two purposes: to support the process of designing an agent and, most importantly, to provide an agent with a framework to represent the world of objects as well as of agents (comprising itself and other agents).

A uniform and consistent framework such as the FBS schema helps an agent to satisfice and to generalise and thus to cope with increasing amounts of grounded knowledge. The FBS schema can be considered as an ontology. This ability is an important condition for an agent to interact with other agents in a situated way. Situated social interaction has the advantage that the agents can operate in open environments, with new agents regularly or irregularly entering the MAS. New interactions can be commenced using the FBS ontology of similar, previously experienced social situations. If the new social situation turns out not to be close enough (and thus the ontology inadequate), the agents modify this ontology until it matches the needs of the new interaction. We see situated social interaction based on FBS ontologies that are constructed on the fly as a potential alternative to predefined approaches to agent interaction, thus making coordination in MASs more flexible.

Acknowledgments

This research is supported by a Sesqui Research and Development grant from the University of Sydney and by an International Postgraduate Research Scholarship.

References

[Brooks, 1986] Rodney A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation*, 2(1): 14-23, 1986.

[Bickhard and Campbell, 1996] Mark H. Bickhard and Robert L. Campbell. Topologies of learning. *New Ideas in Psychology*, 14(2): 111-156, 1996.

[Castelfranchi, 1998] Cristiano Castelfranchi. Modelling social action for AI agents. *Artificial Intelligence*, 103(1-2): 157-182, 1998.

[Clancey, 1997] William J. Clancey. *Situated Cognition*. Cambridge University Press, Cambridge, 1997.

[Clark, 1992] Herbert H. Clark. *Arenas of Language Use*. University of Chicago Press, Chicago, IL, 1992.

[Clark and Marshall, 1981] Herbert H. Clark and Catherine R. Marshall. Definite reference and mutual knowledge. *In* A. K. Joshi and B. Webber (eds), *Linguistics Structure and Discourse Setting*. Cambridge University Press, Cambridge, pages 10-63, 1981.

[Dennett, 1987] Daniel C. Dennett. *The Intentional Stance*. MIT Press, Cambridge, MA, 1987.

[Dewey, 1896] John Dewey. The reflex arc concept in psychology. *Psychological Review*, 3: 357-370, 1896 reprinted in 1981.

[Gero, 1990] John S. Gero. Design prototypes: A knowledge representation scheme for design. *AI Magazine*, 11(4): 26-36, 1990.

[Gero, 1998] John S. Gero. Conceptual designing as a sequence of situated acts. *In* I. Smith (ed.), *Artificial Intelligence in Structural Engineering*. Springer-Verlag, Berlin, pages 165-177, 1998.

[Gero and Fujii, 2000] John S. Gero and Haruyuki Fujii. A computational framework for concept formation for a situated design agent. *Knowledge-Based Systems*, 13(6): 361-368, 2000.

[Gero and Kannengiesser, 2002] John S. Gero and Udo Kannengiesser. The situated Function-Behaviour-Structure framework. *In* J. S. Gero (ed.), *Artificial Intelligence in Design'02*. Kluwer, Dordrecht, pages 89-104, 2002.

[Gruber, 1993] Thomas R. Gruber. A translation approach to portable ontologies. *Knowledge Acquisition*, 5(2): 199-220, 1993.

[Kulinski and Gero, 2001] Jaroslaw M. Kulinski and John S. Gero. Constructive representation in situated analogy in design. *In* B. de Vries, J. van Leeuwen and H. Achten (eds), *CAADFutures 2001*. Kluwer, Dordrecht, pages 507-520, 2001.

[Liew and Gero, 2002] PakSan Liew and John S. Gero. An implementation model of constructive memory for a situated design agent. *In* J. S. Gero and F. Brazier (eds), *Agents in Design 2002*, Key Centre of Design Computing and Cognition, University of Sydney, Australia, pages 257-276, 2002.

[Rao and Georgeff, 1991] Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a BDI-architecture. *In* R. Fikes and E. Sandewall (eds), *Proceedings of Knowledge Representation and Reasoning (KR&R-91)*. Morgan Kaufmann Publishers, San Mateo, CA, pages 473-484, 1991.

[Schön and Wiggins, 1992] Donald A. Schön and Glenn Wiggins. Kinds of seeing and their functions in designing. *Design Studies*, 13(2): 135-156, 1992.

[Simon, 1969] Herbert A. Simon. *The Sciences of the Artificial*. MIT Press, Cambridge, MA, 1969.

[Smith and Gero, 2000] Gregory J. Smith and John S. Gero. The autonomous, rational design agent. *In* H. Fujii (ed.), *Workshop on Situatedness in Design*, Artificial Intelligence in Design'00. Worcester, MA, pages 19-23, 2000.

[Suwa *et al.*, 1999] Masaki Suwa, John S. Gero and Terry Purcell. Unexpected discoveries and s-inventions of design requirements: A key to creative designs. *In* J. S. Gero and M. L. Maher (eds), *Computational Models of Creative Design IV*. Key Centre of Design Computing and Cognition, University of Sydney, Sydney, Australia, pages 297-320, 1999.

[Wooldridge, 2002] Michael Wooldridge. *An Introduction to Multiagent Systems*. John Wiley & Sons Ltd., Chichester, 2002.

This is a copy of the paper: Gero, JS and Kannengiesser, U (2003) Function-behaviour-structure; A model of social situated agents, *IJCAI Agents Workshop* (to appear)