Contextual Cognition in Social Simulation

Bruce Edmonds

Introduction

This chapter looks at the role of context when modelling human behaviour for social simulation. Models of human behaviour in social simulation vary from very simple cellular automata, where there is usually no cognition represented, through to simulations with relatively detailed cognitive models. In order to capture social dynamics, one needs enough agents to make a meaningful society, often including hundreds if not thousands of individuals. This means that it is infeasible to include all the features of a rich cognitive model. However, in order to capture this interplay (where social behaviour influences individual behaviour, and vice versa), it is important to incorporate at least some aspects of cognition into models of human behaviour if one wants to capture many kinds of social phenomena.

For example, the phenomenon of social norms involves both a cognitive dimension (for example what people believe is the norm) as well as the social dimension (for example what patterns of behaviour are most common). There is both an emergent process, from the "bottom up", whereby the beliefs of people interact and aggregate to result in a social regularity (e.g. which side of a corridor to walk down), as well as a process of "downward causation" where the dominant norms constrain the individual behaviour (Conte et al 2013). Here a social norm *is* a complex that arises from the interplay of both cognitive and social processes – if one only considered one of these one would miss something essential about the nature of norms (Xenitidou & Edmonds 2014). Here, I argue that contextual cognition and behaviour is a similar case.

Contextual cognition is simply the idea that much of the cognitive processes we might need to include within the agents in a simulation are, in fact, highly context-dependent. In other words, that the very context-dependency of these cognitive processes can be important because it results in a different set of social outcomes (compared to the case where the cognitive processes are context-free).

The next section will look at the nature of this context-dependency and the various concepts of context that this might involve. This will be followed by a brief review of how contextual cognition has been used in social simulation, followed by an abstract model of how such a cognitive model might be built with an example.

Context and Context-Dependency

Context pervades the human social and cognitive realms but due to its very nature it is often unnoticed or left implicit. However, when one is involved in trying to understand and model these realms it becomes an important factor. How one attempts to understand any kind of phenomena depends upon its properties. The implicit nature of context means that attempts to label it are often "over-loaded", with the result that the word "context" seems to have a variety of related, but distinct, meanings. Here I do not want to enter the debate concerning the "right" meaning of this word, since that has turned out to be a fairly fruitless enterprise, but I do want to make clear the way in which I am approaching the subject. The following discussion leads towards the approach that I will suggest later. It is also necessary to clear away some of the possible confusions that can arise from its use.

Unfortunately, "Context" is used in many different senses and has many different analyses. It is somewhat of a "dustbin" concept, in that if a theory or idea does not work the reason may be assigned to "the context". Thus to many (e.g. linguists) context is a subject that is to be avoided due to its difficulty. "Context" is closely related to (but not identical to) a number of other concepts, including: tacit knowledge (Polanyi 1966), the frame problem in AI (McCarthy and Hayes 1969), framing in psychology (Goffman 1974), and the "situation" (Barwise and Perry 1983).

Situational Context

The situation context is the particular situation where some events or other described phenomena takes place. This could include the time and location, but could include all that is the case about that situation, including: who was there, the knowledge of those people, the history of the place and all the objects present. In this sense the context is indefinitely extensive, it notionally includes all the circumstances in which an event or utterance occurs.

Such a context may be able to be specified adequately (if rather uninformatively) by giving the time and place of the events, but the relevant details might not be effectively retrievable from this. For example, the fact "I was reminiscing about our summer holiday" might well not be detectable from the time and place except by the person doing the reminiscing. Thus when talking about the situational context it is almost universal to abstract from this to what is relevant about that context, or what might be commonly understood (and hence safely not described but left implicit). Thus the phrase "the context" (as in the question "what was the context?") may *mean* "those factors that are relevant to understand this particular occurrence" even though it may *refer* to the situational context in general. Thus to understand what someone is saying to you, you might ask "what was the context?" and get a description of the relevant features of the circumstances, e.g. "I was on the train home".

Linguistic Context

Whilst the situational context could include anything, at least in theory, the linguistic context is composed of the words that surround an utterance or phrase. This typically indicates the words that precede or frame the target of understanding, but could also include the common knowledge that could be reasonably be expected to be known by the listener/reader, e.g. elements of the relevant culture. Sometimes this is taken to be the same as all that which is necessary to understand some example of natural language.

Historically this has been the last resort of the linguist in attempting to pin down the meaning of an utterance – what one appeals to if there seems to be no detectable foreground features to explain its meaning. However recently more positive attention has be focused on context in linguistics. For example, Peter Gardenförs (1997) has said (pragmatics being close to contextual considerations in linguistics):

> *Action is primary, pragmatics consists of the rules for linguistic actions, semantics is conventionalised pragmatics and syntax adds markers to help disambiguation (when context does not suffice).*

Clearly the linguistic context could refer to almost any of the language or culture that surrounds an utterance, and hence is not something that can be captured in its entirety. Often context is thought of as linguistic context because the social interactions that are being considered are composed of linguistic communications.

Cognitive Context and Framing

Clearly many aspects of human cognition are context-dependent, including: visual perception, choice making, memory, reasoning and emotion (Tomasello 1999, Kokinov and Grinberg 2001). A lot of recall, learning and inference is only done *with respect to* a recognised kind of situation. That is, some knowledge is acquired in a particular situation and then made available in similar situations. This abstraction of a situation in the brain – the recognised kind of a situation in which packages of knowledge etc. are relevant – is sometimes called the "cognitive context" (Hayes 1995). This is the cognitive correlate of the situational or linguistic context. Such cognitive contexts could be identified using a description of the kind of situational context that invokes them or else by the set of all the knowledge, norms, expectations, habits etc. that are immediately accessible once recognised. Humans seem to have an innate ability to recognise the cognitive contexts of others (Tomasello 1999).

This idea of cognitive context relates to the idea of framing in psychology as introduced by Goffman (1974), where frames are conceptualised as "schemata of interpretation" that allow individuals "to locate, perceive, identify, and label" experiences. In this way it can be said go make explicit thought possible by giving meaning, organizing these experiences, and guiding actions. This implies that framing is what happens to thought when a cognitive context is adopted. Indeed (Elliott & Hayward 1998) state the following.

> *"Frames are associated with, or even equivalent to, the numerous social and cultural contexts which define the appropriateness of particular norms of behaviour." (p.234)*

However, as (Entman 1993) points out there is no agreed definition of framing. Kahenman and Tversky (1984) use a clear but simple formulation of framing but this is restricted to the use of different reference points with respect to subjective utility of gains and losses. Shafir, Simonson and Tversky (1993) propose a more general "reason-based" model of choice in which

> *"Different frames, contexts and elicitation procedures highlight different aspects of the options and bring forth the different reasons and considerations that influence decision" (p.34)*

Thus the action of framing can be seen as one effect of considering something within a particular cognitive context – it is the result of cognitive context on opinion and choice. The two concepts are very close but the idea of Cognitive Context is more general – it encompasses other areas, such as how and when these patterns of salience are acquired and how it affects the acquisition of knowledge as well as its application.

It is essential that different kinds of contexts can be effectively and reliably recognised as cognitive context but this does not mean that they have to be consciously recognisable as distinct cognitive contexts and labelled. For example, they may be unconsciously recognised by all the members of a community but never named; or maybe they their features are distinctive and consciously recognisable but too complex and fuzzy to be completely specified. This difficulty is further discussed below.

Social Context

Some of the cognitive contexts we have learnt seem to correspond to recognisable kinds of social situation. Examples include: greeting, lecturing, and a political discussion. Once established these seem to be self-perpetuating, in that habits, conventions, norms, terms etc. can be developed by people who recognise the context, but in turn this might mean that the context is more recognisable as an important kind of situation which has its own characteristics. Thus social contexts can be co-constructed over time and passed-on (in terms of experience and social artefacts) to others.

When people are asked to describe the context of an event, they will often do it in social terms. Thus it is that the social context, although it is a special case of situational context is closely linked to the synchronised cognitive context that participants have learnt to associate with situation, because it is often the social aspects that are important in terms of communication and understanding. It is because of the context-dependency of human cognition that when the social context is recognised, experienced participants in that social context will know what set of norms, habits, terms, etc. are associated with it and automatically bring them to bear in their social behaviour and mutual organisation. Thus one of the consequences of the context-dependency of our cognitive capabilities is the prevalence and importance of social context in our understanding of the world.

Identifying and Talking about Context

One of the difficulties in discussing cognitive context is that they may well not (a) be accessible to us (b) identifiable even if they are accessible or even (c) definable in precise terms even if we can identify them. Thus although, in some way, the brain abstracts its incoming stream of information to some properties of its state that it can later use to recognise and retrieve knowledge that is relevant to the same kind of situational context, there is no reason to suppose that we can safely reify these properties that would correspond to the cognitive context. Rather we often have to try and deduce what the cognitive contexts are by introspection and other observation.

Despite this, we often talk about contexts as if they were discrete and identifiable "things", however it needs to be understood that for our conscious selves they may not be the case. Thus "the" context is an abstraction of the aspects of those background features that define it, whether or not this is a meaningful or reifiable entity for us. To simplify the discussion I will generally talk about contexts in the sections below as if they are well defined identifiable entities, but the caveats just mentioned need to be always taken into account. This difficulty means that the cognitive context for any situation is often not made explicit or represented – those involved may well not be aware of the cognitive context they are assuming.

The fact that the relevant cognitive context may not be directly accessible to our consciousness does not mean that it is totally immune to being partially identified or uncovered, just that this might be unnoticed, non-obvious, complex, fuzzy and only partially inferable. For example, although we may not be aware of what brought to mind a particular person in a situation, on introspection we might be able to work out that some music brought to mind a past event in which that person figured. Thus we may be able to work out something about what sort of cognitive context is relevant but still not be able to characterise it completely.

Approaches to Cognitive Contextuality in Social Simulation

Given that context-dependency seems to be fundamental to human cognition and human social behaviour, it is a notable fact that very few social or cognitive simulations represent any of the processes for dealing with such context-dependency. That is to say, the agents in social simulations tend to be endowed with cognitive processes which are not sensitive to, recognise or use context. In other words, agents in social simulations tend not to have anything that might act as a cognitive contexts. If the situation in which the agents are being represented can be considered as a single and fairly simple set of situational contexts, then this is reasonable since one only has to capture the behaviour and interactions within that.

However many simulations are intended not as a representation of something more general than those corresponding to a single cognitive context but aspire to be a more general theory of social interaction. In this case, one has to assume that either the simulation is to be taken only as an analogy or that the simulator does not think people's behaviour, norms etc. will be sufficiently similar between situational contexts so the context-free representation is adequate[1].

In the former case where the simulation is used only as an analogy, then this is valid because humans are experts at applying analogy in a context-dependent manner, adjusting its assumptions and form to be appropriate to its domain of application.

In the later case, where an essentially context-independent algorithm is used to represent a highly context-dependent process must, at least, be the legitimate target for doubt. Whilst the psychological realism that is necessary in a social simulation does depend upon the purpose of the simulation and the level of aggregation (Gilbert 2006), it is certainly not the case that the results of a

[1] These are the charitable assumptions, of course. More often one suspects that the simulator has simply not thought about the difficulties involved in such an enterprise.

simulation are robust against changes in the cognitive model being used (e.g. Edmonds & Moss 2001).

There are not many simulations which represent some aspects of context-dependency in their agents, but there are a few: (Edmonds 1998) used a cognitive learning model specifically because it included some aspects of context-dependency; (Schlosser et al 2005) argue that reputation is context-dependent, (Edmonds and Norling 2007) looks at the difference that context-dependent learning and reasoning can make in an artificial stock market, (Andrighetto et al. 2008) shows that learning context-dependent norms is different from a generic adaption mechanism, and (Tykhonov et al. 2008) argue that the definitions of trust mean that trust is also context-dependent. These show that, at least in some cases, that context-sensitive cognition can make a difference. The fact that it can make a difference is not very surprising given the apparently important role it plays in human cognition, means that there is a burden of justification on those who claim it is unnecessary – explaining why it *can* be safely ignored in their simulations.

There are approaches to including cognitive context within the learning and decision-making of agents. (Edmonds 2001b) which suggests a particular algorithm and approach to learning appropriate cognitive context (but did not achieve the co-development of cognitive context due to the anti-cooperative environment they were embedded within. (Andrighetto et al. 2008) use an approach based on social norms, whereby some of the habits and knowledge of agents are dependent upon the social context, in the sense of which group they are part of. (Alam et al. 2010) uses an endorsement mechanism to implement a kind of context-sensitive learning/decision-making mechanism in agents within a simulation of of some of the power structures within Afganistan. In particular they relate this to folk psychological accounts of how reasoning works and is of form that relates better to available observational and participant evidence. (Knoeri et al. 2011) look to Gidden's structuration theory and structural agent analysis. Within this framework they implement what they call a context-dependent Agent-based model using an analytical hierarchy process as the basis for the agent decision-making process in a model of mineral construction in Switzerland. (Dignum et al. 2004a, 2004b) describe a multi-layered system for specifying agents in simulations that explicitly includes the context-specific interpretation of social norms. (Antunes et al. 2000, Nunes et al. 2013) look at context specificity in terms of the context of different social networks, with switching between them in terms of different social influence operating in each.

A Model of Contextual Cognition

In this section I look at the outlines of a lightweight cognitive model that allows for context-dependent cognition to be implemented within social simulation models. This model integrates Machine-learning type of learning with an AI kind of reasoning via a context-structured memory.

Both learning and reasoning are far more feasible when their scope is restricted to a particular context because this means that only the relevant knowledge needs to be dealt with. However if any significant degree of generality is to be obtained in this manner (McCarthy 1987) then an intelligence must be able to appropriately change this focus as the external context, that is the context we inhabit in (Barwise and Perry 1983), changes. In other words there needs to be some internal correlate of the external context that allows an intelligence to identify which set of beliefs apply. We will call this internal correlate the cognitive context – this is the "internal" approach identified in (Hayes 1997). There are (at least) two tasks necessary for this:

- identifying the appropriate cognitive context from the perceptions of the environment;
- accessing the appropriate beliefs given the identified cognitive context.

The success of this strategy of assessing the relevance of knowledge via identifiable "contexts" depends upon whether the environment is usefully divided up in such a manner. This is a contingent matter – one can imagine (or devise) environments where this is so and others where it is not. The "pragmatic roots" of context, i.e. why context works, depends upon the underlying pattern of commonalities that occur in an environment or problem domain (Edmonds 1999). A cognitive context indicates the boundaries of what might be relevant in any situation.

Context serves not only to make it feasible to deal with our knowledge at any one time but also, at a more fundamental level, to make our modelling of the world at all feasible. The efficacy of our limited learning and inference in dealing with our complex world is dependent on the presumption that many of the possible causes or affects of events that are important remain relatively constant (Zadrozky 1997). Otherwise we would need to include all possible causes and affects in our models and decision making processes, which is clearly infeasible. It is the existence of relative constancy of many factors in particular situations that makes our limited modelling ability useful: we can learn a simple model in one circumstance and successfully use it in another circumstance that is sufficiently similar to the first (i.e. in the same "context").

It is the possibility of the transference of knowledge via fairly simple models from the circumstances where they are learnt to the circumstances in which they are applied that allows the emergence of context. The utility of "context" comes from the possibility of such transference. If this were not feasible then "context", as such, would not arise. For such a transference to be possible a number of conditions need to be met, namely that:

- some of the possible factors relevant to important events are separable in a practical way;
- a useful distinction can be made between those factors that can be categorized as foreground features and the others (the constant, background features);
- similar background factors are capable of being reliably recognized later on as the same "context";
- the world is regular enough for such models to be learnable;
- the world is regular enough for such learnt models to be useful where such a context can be recognized.

While this transference of learnt models to applicable situations is the basic process, observers and analysts of this process might identify some of these combinations of features that allow recognition and abstract them as “a context”. Note that it is not necessarily possible that such an observer will be able to do this as the underlying recognition mechanism may be obscure, too complex or difficult to analyze into definable cases.

Such a strategy answers those of the “frame problem” (McCarthy and Hayes 1969). Firstly, although the frame problem may be unsolvable in general it is learnable in particular contingent cases. Secondly, the identification of appropriate contexts are not completely accessible to reasoning or crisp definition – rather it is an unreliable, information-rich, and imprecise pro-cess. Thus knowing B in context A, is not translatable into statements like A→B, because the A is not a reified entity that can be reasoned about.

The power of context seems to come from this combination of “fuzzy” and fluid context identity and crisp, relatively simple context “contents”. Thus context straddles the fields of Machine Learning and Artificial Intelligence. Machine learning seems to have developed appropriate methods for complex and uncertain pattern recognition suitable for the identification of context. Artificial Intelligence has developed techniques for the manipulation of crisp formal expressions. Context (as conceived here) allows both to be used for different functions in an coherent way.

Context in Reasoning

In 1971, in his ACM Turing Award lecture, John McCarthy suggested that the explicit representation and manipulation of context might be a solution to the effective lack of generality in many AI systems; these ideas were later developed and written up in (McCarthy 1987). McCarthy's idea was to reify the context to a set of terms, *i*, and introduce an operator, *ist*, which basically asserts that a statement, *p*, holds in a context labeled by *i*. Thus:

$$c : ist(i, p)$$

reads "*p* is true in context *i*" which is itself asserted in an outer context c. ist is similar to a modal operator but the context labels are terms of the language. Reasoning within a single context operates in a familiar way, thus we have:

$$\forall i\,(ist(i, p) \wedge ist(i, p \rightarrow q) \rightarrow ist(i, q))$$

In addition one needs a series of ‘lifting’ axioms, which specify the relation between truth in the different contexts. For ex-ample if $i \geq j$ means that “*i*, is more general than context, *j*”, then we can lift a fact to a supercontext using:

$$\forall i \forall j (i \geq j) \wedge ist(p, i) \wedge \neg ab(i, j, p) \rightarrow ist(p, j)$$

where *ab* is an abnormality predicate for lifting to supercontexts. This framework is developed in (McCarthy and Buvac 1998). There are a whole series of formal systems which are closely related to the above structure, including, notably: the situations of (Barwise and Perry 1983); Gabbay's fibered semantics (Gabbay 1999); and the local semantics of the Mecha-nized Reasoning Group at Trento (Ghidini and Giunchiglia 2001).

One of the problems with this sort of approach is that it is likely that trying to apply generic reasoning methods to context-dependent propositions and models, will be either inefficient or inadequate (Greiner et al 2001). The generic approach forces a choice of the appropriate level of detail to be included, so that it is likely that either much information that is irrelevant to the appropriate context will be included (making the deduction less efficient) or much useful information that is specific to the relevant context may be omitted (and hence some deductions will not be possible).

Another problem is that, in practice, this type of approach requires a huge amount of information to be explicitly specified: contexts, contents of each context and bridging rules.

Context in Learning

The use of context in machine learning can be broadly categorized by goal, namely: to maintain learning when there is a hid-den/unexpected change in context; to apply learning gained in one context to different context; and to utilize already known information about contexts to improve learning. There are only a few papers that touch on the problem of learning the appropriate contexts themselves. Included in those that do, Widmer (1997) applies a meta-learning process to a basic incremental learning neural net; the meta-algorithm adjusts the window over which the basic learning process works. Here it is an assumption that contexts are contiguous in time and so a time-window is a sufficient representation of context. Harries et al. (1998) employ a batch learner as a meta-algorithm to identify stable contexts and their concepts; this makes the assumption that the contexts are contiguous in the "environmental variables" and the technique can only be done off-line. Aha de-scribes an incremental instance-based learning technique which uses a clustering algorithm to determine the weight of features and hence implicitly adjust context (Aha 1989).

Contextual knowledge has been used to augment existing machine learning techniques in a number of instances. Turney (1993) used explicit identification of what the contextual factors would be, but others have used implicit features (e.g. Aha (1989)). Turney (1996a) discusses the problem of the effects of context on machine learning and surveys some heuristics used to mitigate these effects (Turney 1996b).

Combining Context-Dependent Learning and Reasoning

Restricting both reasoning and learning to an appropriate context makes both more feasible. However, as with any other technique, there are a number of difficulties with applying a context-dependent approach to reasoning. *Firstly*:

- Explicitly specifying a set of knowledge appropriate for a whole set of potential contexts is both time-consuming and labor-intensive.

Thus with a few honorable exceptions (e.g. CYC (Lenat 1995)), most systems of context-dependent learning or reasoning are only tried out with a few contexts. A possible answer to this (and the one employed here) is to learn the contexts and the context-dependent knowledge. The second is easier than the first; for, as indicated above, there are a number of techniques to learn the knowledge associated with contexts.

The learning of the contexts themselves (i.e. how to recognize when a set of beliefs learnt in a previous situation are again applicable) requires a sort of meta-learning. As documented above, there are such techniques in existence. However most of these either require reasonably strong assumptions about the particular nature of the contexts concerned. An exception is (Edmonds 2001) which describes how contexts can be co-learnt along with the knowledge associated with those contexts. This applies an evolutionary learning algorithm where the knowledge is distributed across a space, where different positions in that space are associated with different set of perceptions or different parts of a problem. This can be clearly understood via the following ecological analogy. If the space can be thought of as a landscape where different parts of the landscape have different properties, and different plants require different properties (some might thrive in marshy land, others sunny and dry etc.). The set of solutions can be seen as varieties of a plant. The different varieties propagate and cross with others in each locality so that, eventually, each variety adapts and, at the same time, spreads across the areas that it is best adapted for. The patches where different sets of varieties thrive define the different ecological niches – corresponding to the different contexts via this analogy.

The ability to learn context allows us to progress beyond the 'loose' loop of:

```
repeat
   learn/update beliefs
   deduce intentions, plans and actions
until finished
```

to a more integrated loop of:

```
repeat
   repeat
      recognise/learn/choose context
      induce/adapt/update beliefs in that context
      deduce predictions/conclusions in that context
   until predictions are consistent
         and actions/plans can be determined
   plan & act
until finished.
```

Such a co-development of cognitive contexts along side their "contents" gives rise to a new problem when the knowledge in these contexts is used to infer predictions and decisions. Thus a *second* problem is this:

- When some of the contents turn out to be wrong, how can one tell when it is the context that is wrong and when it is the contents that are wrong?

There is no universal answer to such a question – it will, in general, depend upon the nature of the domain and hence the appropriate contexts in that domain. However there is a heuristic, as follows: if only a few of the elements of knowledge associated with a context are disconfirmed, it is likely that these are wrong (update the set); if many of the elements are dis-confirmed then it is likely that the context is wrong (change it and learn from this).

Thus in the proposed architecture there are four modules: (1) the context identification system; (2) the context-dependent memory; (3) the local learning/induction algorithm; and (4) the inference system, as shown in Figure 1.

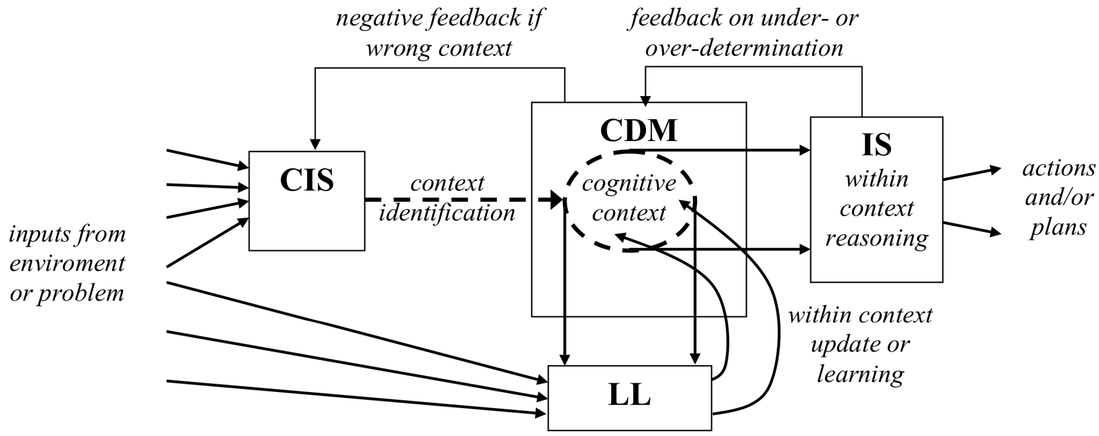

Figure 1. How the context-identification system (CIS), the context-dependent memory (CDM), the local learning algorithm (LL), and Inference system (IS) work together.

The context identification system (CIS) takes a rich range of inputs and learns in a flexible and imprecise way an indication of the context (which it outputs to the memory). The CIS learns as the result of negative feedback when too much of the knowledge in the cognitive context is disconfirmed.
The context-dependent memory (CDM) takes the indication given by the CIS and identifies all those memory items stored within that context. It evaluates the (current) truth of these and if too many are false it returns negative feedback to the CIS, which will identify another context. If a sufficient number of indicated contents are true, then the local learning updates the items within that context. Those items that are (currently) true are passed to the inference system.

The local learning algorithm (LL) performs a local update of the knowledge in the memory. It may include the propagation of successful items towards the focus, but may also include the deletion/correction of items that were false and the possible insertion of new induced/learned.

Finally the planning/inference system (IS) tries to deduce some decisions as to the actions or plans to execute. It could do this in a number of ways, but this could include trying to predict the future states of the world given possible actions and comparing the predictions using its goals.

Two common problems with inference systems that attempt to deduce predictions or decisions from an arbitrary collection of knowledge are under- and over-determination. Under-determination is when there is not enough information to come to a conclusion or decision that needs to be reached. In other words there may be a key proposition, α, such that neither α nor $\neg\alpha$ can be inferred. Over-determination is when there is contradictory information in the knowledge, i.e. when there is an α such that both α and $\neg\alpha$ can be deduced.

This architecture allows a useful response in these two situations. In the case of under-determination the context can be expanded so that more knowledge can be made available to the IS so that it may make more inferences. In the case of over-determination the context can be reduced so that some of the knowledge can be excluded, the knowledge that is peripheral to the context.

Many non-monotonic logics can be seen as attempts to solve the above problems in a generic way, i.e. without reference to any contingent properties obtained from the particular contexts they are applied in. So, for example, some use 'entrench-ment' to determine which extra information can be employed (e.g. oldest information is more reliable (Gärdenfors 1984)), and others allow a variety of default information to be used (e.g. using extra negative knowledge as long as it is consistent (Reiter 1980)). These may work well on occasion, but they can not exploit any of the relevance relations specific to the particular knowledge and context.

Learning Context

In order for context-dependent reasoning to occur, the context-dependent information (or beliefs) need to be captured. If the relevant contexts are already known by the designer (and there is some effective way of recognizing when they apply), then either the relevant information can be entered or a context-enhanced learning algorithm can be employed to learn the information with respect to each context. The former case can be onerous because one not only has to enter the relevant facts as well as specifying each fact's domain of application, but one also has to define all the 'lifting-rules' to allow the integration of the context-dependent information. In the later case the context-dependency of the learning means that one needs correspondingly more information within each context for the learning to be complete.

Thus in order for the desired efficiency in terms of context-constrained reasoning to occur (without a laborious entry of information) for each appropriate context, this information (that is both the contexts and the content in the contexts) should be learned by the agent, at least to some extent.

The basic idea is to simultaneously learn the models and the circumstances in which they work best. If there is sufficient regularity in the environment to allow it this will allow some clusters of similar circumstances to be identified and the corresponding models to be induced. However the clustering and induction parts of the algorithm can not work independently; i.e. clusters of like circumstances being identified and then models induced for these clusters. The reason for this is the contexts are identified by those circumstances where particular models work best. These may correspond to a neat (i.e. humanly identifiable) cluster but this is not inevitable – they may be (to the human eye) inextricably intertwined or overlapping.

There is a population of candidate beliefs, each of which is composed of two parts: a crisp model in a formal language (the content) and some information that specifies the model's domain of application (the domain). In the examples given here the designer specifies what inputs will be used for context recognition and which can be referred to in the model content (some may be in both). Repeatedly a particular circumstance is chosen (for example, these are the ones that simply occur to the agent), and those beliefs who are recognized as most probably relevant (or 'closer') are selected. Out of these the ones that work best are preferentially selected and crossed into future generations of the population. Beliefs that are never anywhere near occurring circumstances are, over time, forgotten.

The basic learning algorithm is as follows:

```
Randomly generate candidate models and place them randomly about
the domain, D
for each generation
  repeat
    randomly pick a point in D, P
    pick n models, C, biased towards those near P
    evaluate all in C over a neighbourhood of P
    pick random number x from (0,1)
    if x < propagation probability
      then propagate the fittest in C to new generation
      else cross two fittest in C, put result into new
           generation
  until new population is complete
next generation
```

A biological analogy makes this clear. Imagine that each belief is an plant. These plants exist in a space defined by the factors that allow context recognition. They compete locally, and those that are better replicate themselves into a neighbourhood (by propagation and sexual reproduction). Thus slowly the successful plants adapt and spread to fill all of the space in which they are relatively successful. Different plants will occupy different areas in the space. The contexts correspond to the ecological niches.

This is an example of the some more general heuristics for learning context.

- *Formation*: A cluster of models with similar or closely related domains suggests these domains can be meaningfully abstracted to a context.
- *Abstraction*: If two (or more) contexts share a lot of models with the same domain, they may be abstracted (with those shared models) to another context. In other words, by dropping a few models from each allows the creation of a super-context with a wider domain of application.
- *Specialisation*: If making the domain of a context much more specific allows the inclusion of many more models (and hence useful inferences) create a sub-context.
- *Content Correction*: If one (or only a few) models in the same context are in error whilst the others are still correct, then these models should either be removed from this context or their contents altered so that they give correct outputs (dependent on the extent of modifications needed to "correct" them)
- *Content Addition*: If a model has the same domain as an existing context, then add it to that context.
- *Context Restriction*: If all (or most) the models in a context seem to be simultaneously in error, then the context needs to be restricted to exclude the conditions under which the errors occurred.
- *Context Expansion*: If all (or most) of the models in a context seem to work under some new conditions, then expands the context to include these conditions.
- *Context Removal*: If a context has only a few models left (due to principle 2) or its domain is null (i.e. it is not applicable) forget that context.

These, the above algorithm and its properties is discussed in much greater detail in (Edmonds 2001).

Example: An Artificial Stock Market Model

In order to demonstrate this approach to learning, I needed an environment that was sufficiently complex yet having emergent contexts (i.e. ones difficult to predict in advance). I have chosen a stock market model, composed of many trading agents and one market maker (roughly following the form and structure of (Palmer et al)). The traders can choose to buy or sell one of a number of shares (if this is possible for them) from or to the market maker. The only fundamental in the market is a dividend rate for each of the shares which slowly change in a random walk. There are only a limited amount of each stock available to the market as a whole. The market maker sets prices as a result of the demand – if there is net demand for a stock it raises the price and if there is a net negative demand it lowers the price. There is a small transaction cost to the traders for every trader, so rapid random trading is unlikely to benefit it.

The goal of the traders is to maximise the total value of their assets (cash plus shares at current value). Thus the traders are in competition with each other – one trader tends to gain at another's expense. However this is not a zero-sum game due to the dividends paid on stocks and the possibility of making money at the market maker's expense.

Each time period the traders simultaneously buy or sell each of the stocks, assuming they have enough cash to fund the net price, the stocks to sell, and the market maker has the stocks to sell. Traders do not have to trade in any stock. Thus the decision that each of the traders has to make is how much to attempt to buy or sell of each stock each time period.

Traders can observe the following:

- the current and past prices of all stocks;
- the past actions of all traders;
- the current and past dividend rates.

In addition the traders are provided with primitives for:

- the current and past market index (average of all prices);
- recent trend of the index;
- recent total volume of trading;
- recent market volatility;
- the maximum historical price of any stock.

The operators available to the agents to build models with are:

- basic arithmetic (+, −, ×, ÷);
- the ability to refer back in time (last and lag operators).

They also have some constants, namely:

- the names of the other traders,
- the names of the stocks
- and a selection of random constants.

Basically the traders try to learn to predict what each of the stocks will be in the next time period and then buy or sell if they predict it will rise or fall sufficiently for this to be worthwhile.

This sort of set-up produces a rich series of dynamics as the traders participate in sequences of modelling 'arms-races' and imitation 'games'. Any successful prediction schema will not last forever as the other traders will soon spot your trading pattern and exploit it to your disadvantage. However, as with real stock markets, there are definitely patterns and market 'moods' (if there are enough traders and stocks), for example bull markets and speculative bubbles. There will be periods of relative quiet as traders sit on stock and so effectively prevent trading and periods of high volatility as subgroups of traders engage in bouts of activity trying to exploit each other. The dynamics are related to those of the "minority game" (Arthur 1994), and similar (Akiyama & Kaneko 1995) but are more varied and complex. Thus, although this is an artificial setting, it goes way beyond a "toy" problem in scope and complexity.

There are two types of traders: which I will call *generic* and *context* traders. Both types maintain a population of 20 models, each of which is composed of a separate expression to predict the future price of each stock. All models are initially randomly generated to a depth of 5 using the inputs, primitives, operators and constants already listed. Both agents use an evolutionary learning algorithm which evaluates fitness by the profit the agent would have made over the past 3 time periods had it used these models to predict prices.

The generic traders use a genetic programming learning algorithm to evolve their predictive models and the context traders have an adapted version of this algorithm to allow the simultaneous learning of context for its models. The types are otherwise identical.

The learning algorithm for the generic trading agent is as follows:

```
Randomly generate initial population of candidate models
for each generation
  for each model
    evaluate what the total wealth of the agent would be if
     it had used this model in trading over the past few
     time periods,  this is the model's fitness
  next model
  repeat
    randomly pick two models with a probability proportional
      to their current fitnesses
    pick random number x from (0,1)
    if x < propagation probability
      then propagate them to new generation
      else cross them and put results into new generation
  until new population is complete
next generation
```

The context trader's algorithm differs a little from the basic version outlined in the last section. This is because from an agent's point of view the only relevant circumstances (in terms of the space of possible ones) are those that actually occur. Therefore instead of randomly picking a sequence of circumstances until the new population is generated, we use only the present circumstance repeatedly and we propagate the rest into the next population with a bias against those that are furthest from any circumstance that has occurred. Also in this model we have associated with each model content a set of positions, so that its domain of application is indicated by a small cloud of points, not a sharply defined region.

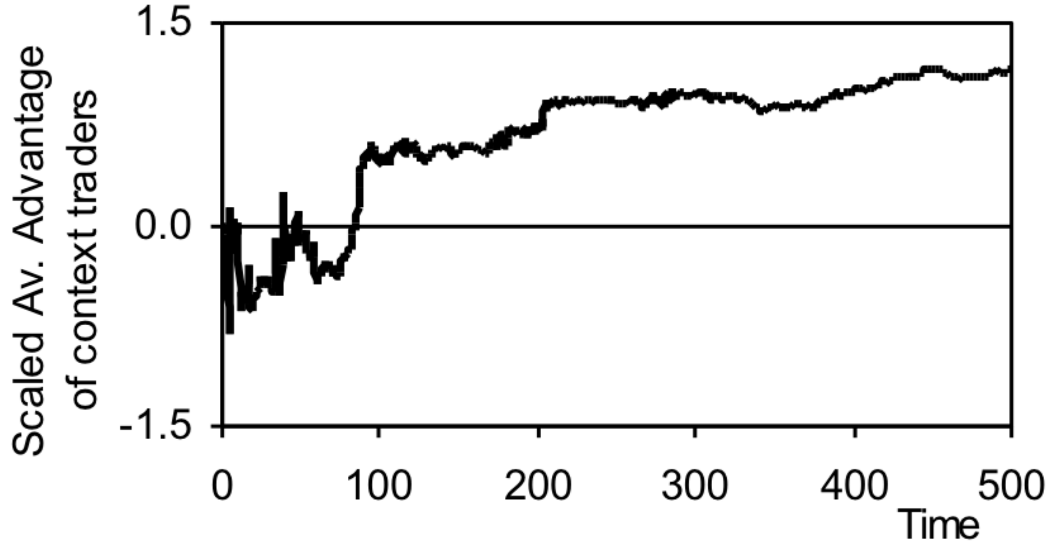

Figure 2. Difference of average asset values of context and generic traders, scaled by current asset spread

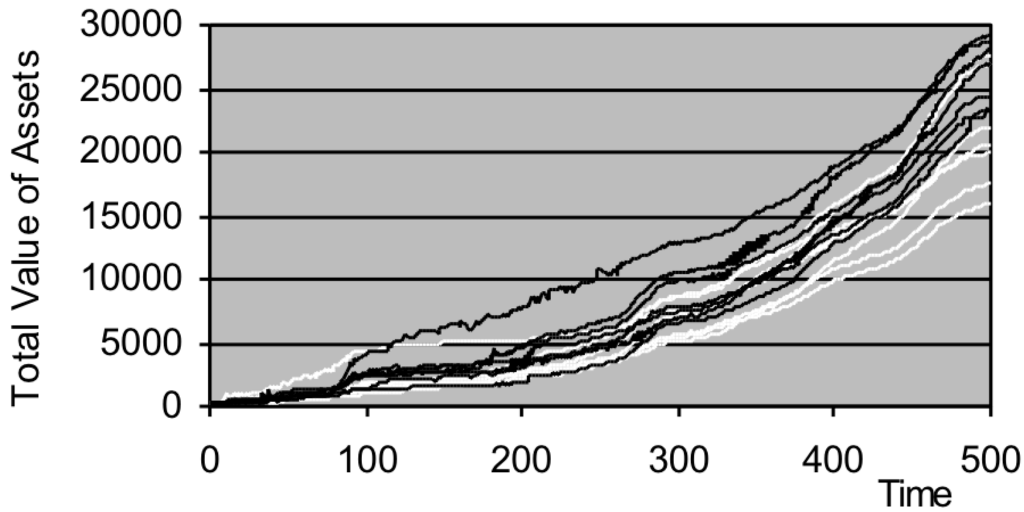

Figure 3. Growth in Agents' Assets over time (context traders in black, generic in white)

It is not obvious that the context trader is a better learner than the generic trader. The context algorithm restricts which models can be crossed to produce new variants to those that are in the same neighbourhood of an occurring circumstance, whilst the generic algorithm allows a more global search for solutions. Thus one might expect that the context traders do better only if there is a context-dependency in the environment to exploit. As we shall see this appears to be the case in this model.

The model was run with 7 of each type of agent (thus 15 including the market maker) trading 5 different stocks over 500 time periods. The model was implemented in SDML (Moss et al 1998).

For the first 80 periods one of the generic traders was doing substantially better than the others, but after this the context traders clearly did better, on the whole (see Figure 3). To make clear the significance of the difference between context and generic traders I have plotted the difference between the average value of context traders' assets minus the average value of the generic trader's assets, scaled by the current standard deviation of the spread of total asset values (Figure 2).

It is notable that the generic traders did better if there were only 2 or 3 of each type of trader – the context traders only reliably out-perform the generic traders (on the whole) with larger populations of traders. The context traders do particularly well if they are in a minority among many generic traders. It is postulated that it is only with larger numbers of the same type of trader that learnable contexts appear in the trading patterns for the context traders to learn and exploit.

To show that the context traders are, in fact, identifying meaningful contexts (at least sometimes), I have taken a snapshot of the positions indicating the domain of the 6 of the models in one agent for one stock at one time (the best performing agent halfway through the run). These clusters are shown in Figure 4. The contents of these six model are shown in Table 1.

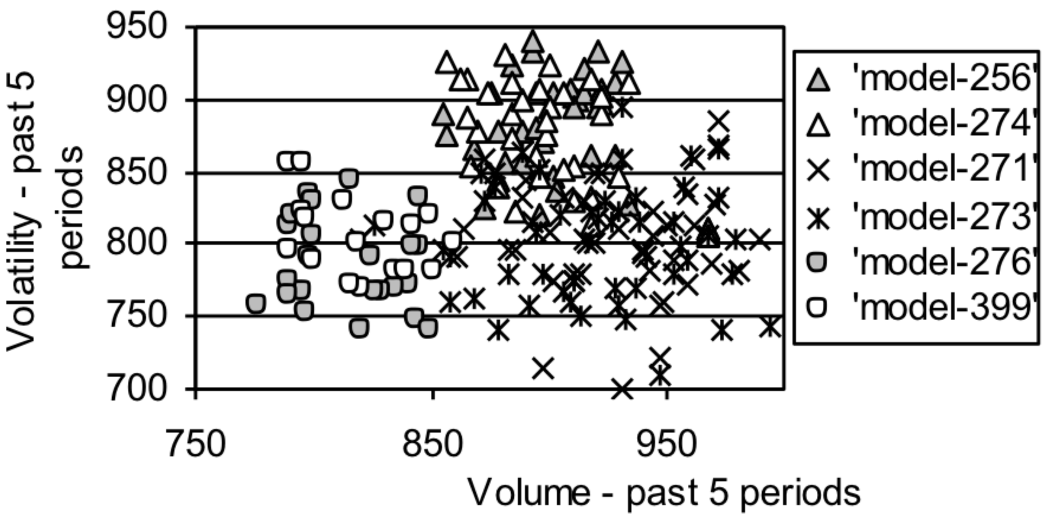

Figure 4. Snapshot of clusters of positions of 6 action models for a context trader indicating three distinct contexts.

Table 1. The action models (for stock 3) in Figure 4.

model-256	priceLastWeek (stock-4)
model-274	priceLastWeek (stock-5)
model-271	doneByLast (normTrader-5) (stock-4)
model-273	IDidLastTime (stock-2)
model-276	IDidLastTime (stock-5)
model-399	minus (divide (priceLastWeek (stock-2)) (priceLastWeek (stock-5))) (times (priceLastWeek (stock-4)) (priceNow (stock-5)))

For this agent at this time there seem to be three regions that might correspond to different contexts: one for lower volatility and higher volume, one for lower volatility and lower volume and one for higher volatility and middle volume. It is notable that, even within each of these there are a mixture of two models that are appropriate. Thus, even given the circumstances, the model selected for will be determined by recent predictive performance: for example, in the case of stock 3 in the above snapshot its price may be modelled best by either the price of stock 4 or stock 5 last time period.

The point of this example is two-fold:

- That implementing a context-dependent cognitive model within a social simulation is feasible, and
- That it makes a difference – there is an observable difference in behaviour between the context and generic traders in this simulation.

The point is *not* that context-dependent cognition will be perform better in all circumstances, since as with all cognitive processes (Edmonds 2008), context-sensitivity will be helpful in some circumstances and not in others.

Conclusions

The lack of agents endowed with the cognitive ability to recognise social context must limit or change the social complexity that results when they interact. In particular, the co-development of social contexts will be lacking, where the recognisability of a distinct social context will allow new and specific habits, norms etc. to be developed for that situation, enabling that social context to become more recognisable etc. This will limit the ability of such simulations to capture some classes of social phenomena where the co-development of social context is a key part. Thus it may be, for example, that such things as a "jittery market" might correspond to a co-developed cognitive context, recognised and reinforced by the market traders in that market[2].

Thus this suggests that:

- That a simulation composed of agents with essentially non-context cognitive models might be giving deceptive results, especially in cases where the agents are learning and/or making decisions in a wide variety of situations.
- Sometimes less "smooth" learning and inference algorithms in the agents in a simulation, that mimic some aspects of context-dependency, as observed in the humans that are being modelled, might well produce a simulation that matches the observed outcomes better.

In other words, the cognitive model encoded in the agent can matter. One can not hope that an "off-the-shelf" model based on something from another context, like AI or machine learning, will be good enough.

Context-dependency pervades the subject matter of social phenomena, with feasible modelling possible only within specific varieties of context. At the very beginnings of sociology Max Weber did point out the inherent context-dependency of social phenomena, also pointing out that this does not stop a scientific study of it (Coser 1977). These roots seem to have been somewhat forgotten.

Acknowledgements

khjgkjghjk

References

Akiyama, E. and Kaneko, K. (1995) Evolution of Cooperation, Differentiation, Complexity, and Diversity in an Iterated Three-person Game. Artificial Life, 2. 293-304.

Alam, Shah Jamal, Geller, Armando, Meyer, Ruth and Werth, Bogdan (2010) 'Modelling Contextualized Reasoning in Complex Societies with "Endorsements"' Journal of Artificial Societies and Social Simulation 13(4):6 <http://jasss.soc.surrey.ac.uk/13/4/6.html>.

Akman, V. and Surav, M. (1996) Steps Toward Formalizing Context. AI Magazine, 17. 55-72.

Akman, V., Bouquet, P., Thomason, R. and Young, R.A. (eds.). (2001) Modeling and Using Context: Proceedings of the Third International and Interdisciplinary Conference, CONTEXT'2001, Dundee, Scotland, 2001. Springer-Verlag, Berlin, 2001.

Andrighetto, G., Campennì, M., Conte, R., and Cecconi, F. (2008). Conformity in Multiple Contexts: Imitation Vs Norm Recognition,. In World Congress on Social Simulation 2008 (WCSS-08) George Mason University, Fairfax, USA.

Antunes, L. Nunes, D., Coelho, H., Balsa, J. and Urbano, P. (2000) Context Switching versus Context Permeability in Multiple Social Networks. EPIA 2009:547-559.

Nunes, D., Antunes, L. and Amblard, F. (2013) Dynamics of Relative Agreement in Multiple Social Contexts. EPIA 2013:456-467.

---

[2] As well as many other factors, of course.

Arthur, B. (1994) Inductive Reasoning and Bounded Rationality. American Economic Association Papers, 84. 406-411.

Barwise, J. and Perry, J. (1983). Situations and Attitudes. Cambridge: MIT Press.

Bouquet, P., Serafini, L., Brézillon, P., M. Benerecetti and Castellani, F. (eds.). (1991) Modeling and Using Context: Proceedings of the Second International and Interdisciplinary Conference, CONTEXT'99, Trento, Italy, September 1999. Springer-Verlag, Berlin, 1999.

Chattoe, E. (1998) Just How (Un)realistic are Evolutionary Algorithms as Representations of Social Processes? Journal of Artificial Societies and Social Simulation 1(3):2 http://jasss.surrey.ac.uk/1/3/2.html

Conte, R., Andrighetto. G. & Campennì, M. (eds.) (2013) Minding Norms - Mechanisms and dynamics of social order in agent societies. Oxford University Press

Deacon, T.W. (1998) Symbolic Species: The Co-Evolution of Language and the Brain. W. W. Norton & Co.

Dignum, V. Vazquez-Salceda, J. and Dignum, F. (2004a). A Model of Almost Everything: Norms, Structure and Ontologies in Agent Organizations. In Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 3 (AAMAS '04), Vol. 3. IEEE Computer Society, Washington, DC, USA, 1498-1499. DOI=10.1109/AAMAS.2004.20 http://dx.doi.org/10.1109/AAMAS.2004.20

Dignum, V. Vazquez-Salceda, J. and Dignum, F. (2004b) OMNI: Introducing Social Structure, Norms and Ontologies into Agent Organizations. PROMAS 2004:181-198.

Edmonds, B. (1998). Modelling Socially Intelligent Agents. Applied Artificial Intelligence, 12, 677-699.

Edmonds, B. (1999a) The Pragmatic Roots of Context. CONTEXT'99, Trento, Italy, September 1999. Lecture Notes in Artificial Intelligence, 1688:119-132.

Edmonds, B. (1999b). Capturing Social Embeddedness: a Constructivist Approach. Adaptive Behavior, 7:323-348.

Edmonds, B. (2001a) The Use of Models - making MABS actually work. In. Moss, S. and Davidsson, P. (eds.), Multi Agent Based Simulation, Lecture Notes in Artificial Intelligence, 1979:15-32.

Edmonds, B. (2001b) Learning Appropriate Contexts. In: Akman, V. et. al (eds.) Modelling and Using Context - CONTEXT 2001, Dundee, July, 2001. Lecture Notes in Artificial Intelligence, 2116:143-155.

Edmonds, B. (2002) Learning and Exploiting Context in Agents. Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), Bologna, Italy, July 2002. ACM Press, 1231-1238.

Edmonds, B. (2007) The Practical Modelling of Context-Dependent Causal Processes – A Recasting of Robert Rosen's Thought. Chemistry and Biodiversity, 4(1):2386-2395.

Edmonds, B. (2008) *The Social Embedding of Intelligence: How to Build a Machine that Could Pass the Turing Test*. In Epstein, R., Roberts, G. and Beber, G. (Eds.) Parsing the Turing Test. Springer, 211-235.

Edmonds, B. (2009) The Nature of Noise. In Squazzoni, F. (Ed.) Epistemological Aspects of Computer Simulation in the Social Sciences. LNAI 5466:169-182.

Edmonds, B. and Hales, D. (2005) Computational Simulation as Theoretical Experiment, Journal of Mathematical Sociology 29(3):209-232.

Edmonds, B. and Moss, S. (2001) The Importance of Representing Cognitive Processes in Multi-Agent Models, Invited paper at Artificial Neural Networks - ICANN'2001, Aug 21-25 2001, Vienna, Austria. Published in: Dorffner, G., Bischof, H. and Hornik, K. (eds.), Lecture Notes in Computer Science, 2130:759-766.

Edmonds, B. and Norling, E. (2007) Integrating Learning and Inference in Multi-Agent Systems Using Cognitive Context. In Antunes, L. and Takadama, K. (Eds.) Multi-Agent-Based Simulation VII, 4442:142-155.

Gabbay, D.M. Fibring logics. (1999) Clarendon, Oxford.

Gärdenfors, P. (?1984) Epistemic Importance and Minimal Changes of Belief. Australasian Journal of Philosophy, 62 (2). 136--157.

Gärdenfors, P., (1997) The pragmatic role of modality in natural language. in 20th Wittgenstein Symposium, (Kirchberg am Weshel, Lower Austria, 1997), Wittgenstein Society.

Ghidini, C. and Giunchiglia, F. (2001) Local Models Semantics, or Contextual Reasoning = Locality + Compatibility. Artificial Intelligence, 127 (3). 221-259.

Gilbert, N. (2006) When Does Social Simulation Need Cognitive Models? In Sun, R. (ed.) Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation. Cambridge University Press, 428-432.

Greiner, R., Darken, C. and Santoso, N.I. (2001) Efficient reasoning. ACM Computing Surveys, 33 (1). 1-30.

Harries, M.B., Sammut, C. and Horn, K. (1998) Extracting Hidden Contexts. Machine Learning, 32. 101-112.

Hayes, P. (1995). Contexts in Context. Context in Knowledge Representation and Natural Language, AAAI Fall Symposium, November 1997, MIT, Cambridge.

Kaneko, K. (1990) Globally coupled chaos violates the law of large numbers but not the central-limit theorem. Physics Review Letters 65:1391–1394

Knoeri, Christof, Binder, Claudia R. and Althaus, Hans-Joerg (2011) 'An Agent Operationalization Approach for Context Specific Agent-Based Modeling' Journal of Artificial Societies and Social Simulation 14 (2) 4 <http://jasss.soc.surrey.ac.uk/14/2/4.html>.

Kokinov, B. and Grinberg, M. (2001) Simulating Context Effects in Problem Solving with AMBR. in Akman, V., Bouquet, P., Thomason, R. and Young, R.A. eds. Modelling and Using Context, Springer-Verlag, 2116:221-234.

McCarthy, J. (1971) Generality in Artificial-Intelligence - Turing Award Lecture. Communications of the ACM, 30 (12). 1030-1035.

McCarthy, J. and Buvac, S. Formalizing Context (Expanded Notes) (1997) . in Westerstaahl, A.A.a.R.v.G.a.D. ed. Computing Natural Language, CSLI Publications, Stanford, California, 1998, 13--50.

McCarthy, J. and Hayes, P.J. (1969). Some philosophical problems from the standpoint of artificial intelligence. Machine Intelligence, 4:463-502.

Moss, S., Gaylard, H., Wallis, S. and Edmonds, B. (1998) SDML: A Multi-Agent Language for Organizational Modelling. Computational and Mathematical Organization Theory, 4 (1). 43-69.

Palmer, R.G. e.a. (1994) Artificial Economic Life - A simple model of a stockmarket. Physica D, 75. 264-274.

Pearl, J. (2000) Causality. Cambridge University Press.

Reiter, R. (1980) A Logic for Default Reasoning. Artif Intell, 13. 81-132.

Schlosser, A., Voss, M. and Brückner, L. (2005). On the Simulation of Global Reputation Systems. Journal of Artificial Societies and Social Simulation 9(1)4 <http://jasss.soc.surrey.ac.uk/9/1/4.html>.

Terán, O. (2004) Understanding MABS and Social Simulation: Switching Between Languages in a Hierarchy of Levels. Journal of Artificial Societies and Social Simulation 7(4):5 http://jasss.soc.surrey.ac.uk/7/4/5.html

Turney, P. and Halasz, M. (1993) Contextual Normalization Applied to Aircraft Gas-Turbine Engine Diagnosis. Applied Intelligence, 3 (2). 109-129.

Turney, P.D., (1993) Robust classification with context-sensitive features. in Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE-93, (Edinburgh, 1993), Gordon and Breach, 268-276.

Turney, P.D., (1996) The identification of context-sensitive features: A formal definition of context for concept learning. in ICML-96 Workshop on Learning in Context-Sensitive Domains, (Bari, Italy, 1996), 53-59.

Turney, P.D., (1996) The management of context-sensitive features: A review of strategies. in ICML-96 Workshop on Learning in Context-Sensitive Domains, (Bari, Italy, 1996), 60-66.

Tykhonov, D, Jonker, C, Meijer, S and Verwaart, T (2008), Agent-Based Simulation of the Trust and Tracing Game for Supply Chains and Networks. Journal of Artificial Societies and Social Simulation 11(3)1 <1 http://jasss.soc.surrey.ac.uk/11/3/1.html>.

Wheeler, M. and Clark, A. (1999) Genic representation: reconciling content and causal complexity. British Journal for the Philosophy of Science, 50, 1, pp. 103-135.

Widmer, G. (1997) Tracking Context Changes through Meta-Learning. Machine Learning, 27. 259-286.

Xenitidou, M. & Edmonds, B. (in press, 2014) The Complexity of Social Norms. Springer.

Zadrozny, W. (1997). A Pragmatic Approach to Context. Context in Knowledge Representation and Natural Language, AAAI Fall Symposium, November 1997, MIT, Cambridge.