A Simple-Minded Network Model with Context-like Objects

Bruce Edmonds,
Centre for Policy Modelling,
Manchester Metropolitan University.
http://www.fmb.mmu.ac.uk/~bruce
Thursday, February 13, 1997

Abstract

The aim of this paper is to describe a simple extension of semantic nets. In this formulation we have labelled nodes with directed arcs, but the directed arcs can lead to other arcs as well as nodes. In this model contexts are not differentiated as special objects, but rather that some nodes to a greater or lesser extent have roles as encoders of contextual information.

This formulation is shown to be expressive enough to capture several aspects of context, namely: context-dependent inference, context specific learning, the selection of a relevant context and the generalisation of knowledge. Its strengths are its simplicity, the fact that it can relate and integrate several aspects of context and its connections with formal logic. It is not claimed that this is a model of any type of context found in human activity.

1 Introduction

The purpose of this paper is *not* to formalise the many human uses of context or to determine the properties of context in an *a priori* way, but to put forward a formal model that could be used to relate different conceptions of context in a simple way.

As has been frequently pointed out (e.g. by Pat Hayes [2]), there are many different meanings of “context”. For example there is the context one may inhabit, the shared linguistic context which is used to enable effective communication and the specific mental constructs that acts as frameworks for inference and learning. Since my aim in this paper is to relate different kinds of context I want to abstract from these to possible shared properties, in order to motivate what follows.

To do this I concentrate on a mental-type construct since the other two are related to it. The context-one-may-inhabit only impinges upon me if I have perceived and represented it internally (though not necessarily symbolically), it is this internal reflection of the external context that is active from my point of view. The shared context-as-a-resource can be seen as comprising a set of internal contexts and some mechanism of reasoning about which is appropriate at any instance, for at any particular time the participants will need to have chosen an appropriate internal context from the context-as-a-resource in order to correctly interpret the communication. In these cases there is added complexity from the problem of how the participants select an internal context from either their perceptions of the environment or the communications of the other participants, but this does not stop us discussing the core of the composition of this common internal structure.

Although the model is based upon semantic-web models of (presumed) mental constructs, it is abstracted from it so as to be as general as possible. I call this a *generalised context* to indicate its distance from more grounded concerns about context.

In section 2 I list some properties of such a generalised conception of context; section 4 describes the basic model; section 5 discusses some simple extensions of the model and, finally, section 6 traces the relation of this model to other formalisations.

2 Some context-related properties a bridging model might need to include

In general, a model that seeks to relate different kinds of context needs to be able to include the following features.

2.1 (more specific) contexts increase inferential power

The power of contexts is that they greatly restrict the possible inferences so that it is easier to deduce the relevant facts about any particular situation. In addition the context may provide extra facts to be used in any such inference.

For example, when told that *a certain car stops if the light is red* and that *the car stops* people often conclude that *the light is red*. This seems to be because the initial facts also set the context (cars stopping at lights) which, if presumed, allows this conclusion. This is in contrast to trained logicians (when thinking in that mode) who have learned to restrict themselves (as far as possible) to inferences *without* presumptions and, in particular, without presumed contexts so that they will not be able to make any conclusions in this case.

2.2 learning (new information) occurs in specific contexts

Practical learning frequently seems to be context-specific. The difficulty with learning in the presence of a complex of contextual facts is that is a non-trivial task to distinguish between those that are relevant and those that are not (unless you are dealing with established contexts), so as to record only the useful contextual information with the learnt fact. It is only when such sets of contextual information form a coherent and frequently used group can one talk about *a context*, rather than *a set of contextual information*. In this case the set acquires an identity (and possibly an identifying archetype as a label). In other words, contexts themselves have to be learnt in parallel with other facts.

For example, children will often learn complete phrases for use in particular contexts and initially maintain a close correspondence between such phrases and their original contexts.

2.3 knowledge can be generalised from specific contexts (to more general ones)

One can separate learning facts in specific contexts from the process of generalising these to more general contexts. For this to happen there must presumably be some commonality between the facts which may also reflect some commonality between the different contexts. In any case it is clear that there is some mechanism for generalising knowledge from specific contexts to wider ones.

For example, if you know that you are nervous about public speaking and you find (to your surprise) that your friend is also nervous, you might well conclude that many people are similarly nervous.

2.4 contexts themselves can be objects of inference

Contexts are always just frameworks for inference and learning, but can sometimes themselves be objects of thought. This may be the result of obvious deliberation or not, but there must be some mechanism for selecting the correct context based on the available information and previous contexts.

For example, if one is talking at cross-purposes with someone, this can be realised and the correct context (i.e. the other person’s) inferred to re-establish effective communication.

2.5 different contexts can be selected depending on previous contexts (as well as other facts)

The order of previous contexts could affect the context assumed as the interpretation of any new facts is affected by the present context, including judgements of relevance. Such judgements of relevance may well affect the selection of a subsequent context.

For example, given a context of *the 30’s* on might well think of the economic depression, if one was then given a further context of *the U.S.* one might think of the effects of the American depression, but if one was given the context of *the U.S.* first, one might fix on something different (say gangster movies) so that the subsequent information indicating *the 30’s* might bring to mind prohibition.

2.6 whether something acts as a context or not could itself be context-dependent

Given that contexts can be objects of reasoning and learning as well as acting as a context for other reasoning, it would be surprising if *this* fact could not itself be context-dependent. In any case this is a possibility that needs to be catered for.

For example, *motor racing* can be a context for the discussion of the tactics involved in cornering but may be without any contextual properties in a comparative discussion of danger in sport.

3 Modelling context as an emergent property of a more basic structure

The considerations above suggest the potential dual nature of contexts (as context and object of reasoning). This means that a model which formalises contexts as a different class of object will need some method of strongly associating contexts with other objects (as in McCarthy's conception of first order contextual logic where reasoning can be about contexts).

Perhaps a more natural way of formalising context in order to compare different conceptions, is to not distinguish contexts from other object of reasoning, learning etc. but to be chosen such that certain object act *like* contexts in certain circumstances. Thus one would like a reasonably simple formalism such that particular kinds of model of context and context-related structures can be seen as special cases of the more general structure. I put forward one candidate below - I have suspicions that it is too simple for the job, but it is better to start simply and elaborate rather than start with an over-complex formalism that may obscure the essential details and hence the debate.

4 An enhanced network model

The basic model comprises of a set of labelled nodes and directed arcs. Unlike usual networks, directed arcs can go from node to another arc as well as from node to node. I will call this an *extended net*. An example is shown below in figure 1. Arcs such as the one in figure 1 between the tweety node and the bird node which are not pointed to by a directed arcs, I will call an *unconditional arc*. Other arcs (such as the one from the bird to the flyer node) will be called *conditional arcs*.

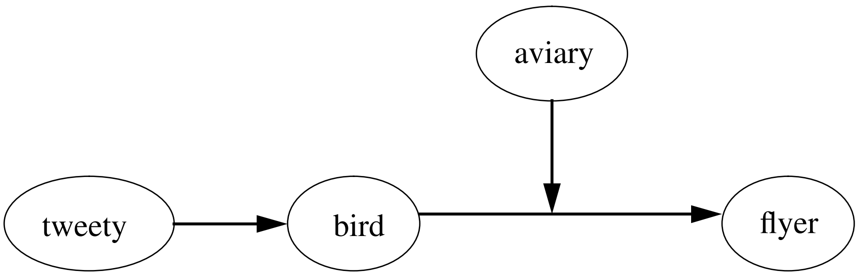

Figure 1: Example of an extended net

The simple idea is that nodes and arcs are either activated or not. The activation of one node can cause unconditional arcs leading from it to be activated. Conditional arcs may be activated if the node they lead from is activated and any arc leading to it is active. Nodes are activated by activated arcs leading to it. The spreading of activation from one node to another represents an inference. Chains of activated arcs will be called a *path*. There are various possible elaborations of this scheme, but I will leave these to later (section 5).

Nodes that have many paths leading from it to other arcs then act like a context. If they are activated they enable many inferences to occur that would not be possible otherwise. I will call such nodes *contextual nodes*. ***Note:*** I am not claiming these *are* contexts or that they model *real* properties of natural contexts but just that they have context-like features.

Context-dependent learning could take place when there are a number of “background” facts associated with the induced/observed connection. For example, when walking into an aviary one may be aware that it is a specialised enclosure designed for keeping birds, in that aviary one may observe that all the birds fly. Here an association between birds and flight is made in the context of the group of properties that distinguish the situation one finds oneself in (see figure 2). Later one may separately learn to distinguish the relevant features of an aviary.

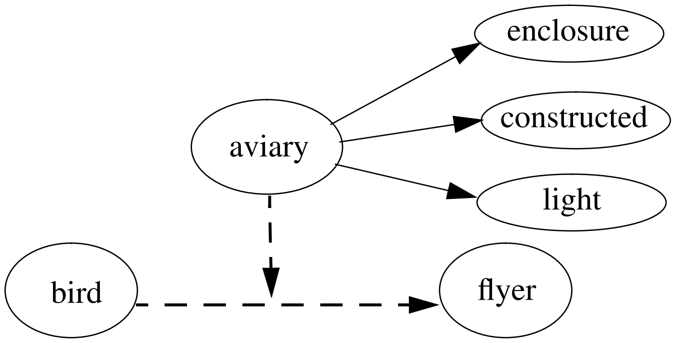

Figure 2: Learning that birds fly in the context of an aviary

Since nodes representing contexts are just as other nodes, they can themselves easily be part of an inference network. This allows the next context to be selected base on the previous context and other information. Contexts need not be in a strict hierarchy. Contexts may be active one at a time or multipley. Contexts can have a specific identity as in the aviary example shown in figure 2 above, where the collection of context related facts have been abstracted to that of an aviary.

Generalisation of facts from a set of specific contexts to a more general one can be done by learning common associations in several related contexts. There are many possible schema for generalisation expressible using this model. One such is shown in figure 3 below; here there is an arc (A to B) shared between two contexts, C1 and C2, (alternatively they may be duplicated - that is a matter of the semantics of the formalisation), the condition of this arc may then be generalised to a wider context (GC).

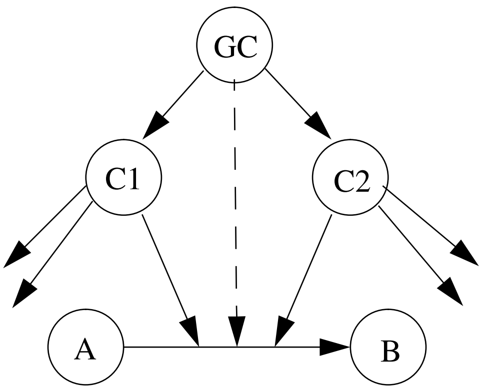

Figure 3: a possible schema for generalisation from specific contexts

Due to limitations of space I have not described the possible processing of these networks in any detail, of which there are several possibilities.

5 Some extensions of the model

I briefly describe two extensions to the model above, there are many other obvious ones.

5.1 Negation

Negation can be captured in two ways. There is the weak negation of simply not being activated ($\bar{p}$) and the strong negation of being forced into an inactive state by the structure of the network ($\neg p$). Weak negation is already included by just taking those parts of the network that are not activated.

Strong negation can be introduced with the introduction of a special node - the *absurdum* or *bottom* ($\perp$). This is a simply a node that must not be activated, acting rather like the ground terminal in electrical circuits. It can either be seen as forcing all nodes that lead to it as inactive or that is activation implies a negative feedback to the present structure (thus weakening or eliminating relations). Negation $\neg p$ can now be defined as $p \to \perp$. This sort of negation is a sort of intuitionistic negation so that, for example, the law of the excluded middle, $p \vee \neg p$, does not hold in general.

Bottom here could be used to symbolise different things: falsehood or an action representing negative feedback. In the first instance it could be used to merely indicate strong negation, in the second if bottom was ever activated this could indicate that the arcs leading to it could be removed. This could be done in a variety of ways, consistency resolution (like the "fixes in [6]), negative endorsement [1] or in a manner similar to some neural networks (as surveyed in [5]).

5.2 Necessity

In a similar way to bottom we can also define a special top node (T). This acts like the positive terminal in a circuit - it forces certain nodes to be activated. Thus nodes that it connects to with an unconditional arc must be necessarily active. This can be seen to formalise a type of necessity, $\Box p$, defined as $T \to p$. For completeness one could view all unconditional arcs $p \to q$ as arcs that are conditional upon T.

The presence of both top and bottom allows for the possibility of strong inconsistency - a sort of "short circuit" from T to $\perp$. This might indeed occur when mislearning has occurred or when there is a basic conflict in the agents goals - it might thus be the trigger for a radical rewiring.

6 Relation of the model to some other formalisations

6.1 Semantic networks

The enhanced network model described is strictly as powerful as a normal semantic network. In fact a semantic network can be seen as a special case where all contextual nodes are not involved in inference relations but act as a typing mechanism on the relations between labels (as illustrated in figure 4). The is-a relation can be seen as the untyped relations. Here each relation is typed by at most one contextual node. Here tweety *is-a* bird, a bird *is-a* animal, bird *has-property* fly and bird *has-property* walk.

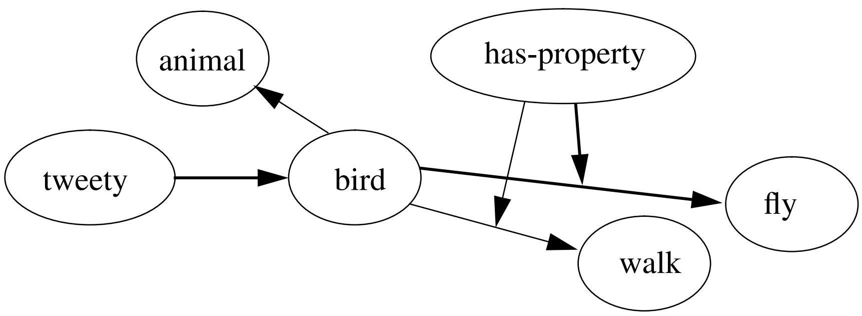

Figure 4: Implementing a semantic network with typed relations

6.2 Mixed defeasible inheritance networks

Mixed defeasible inheritance networks have strict and defeasible links between nodes. Basically in the lack of information from strict inheritance relations (corresponding to is-a relations), one can make deductions via the defeasible links under certain conditions (chiefly that the inference is not undermined by an inference with priority over it). Thus defeasible inference networks can be seen as an approximation of the enhanced network models where unconditional relations are mapped onto strict inheritance links and conditional ones to defeasible relations. In this case the logic of such inheritance networks can be seen as the logic of what one might infer from context-dependent knowledge when you do not know the present context.

6.3 McCarthy Style Contextual logic

One can Map the *ist(c, p)* relation to the enhanced network model with negation. $ist(c, a \to b)$ is mapped to a relation from *a* to *b* which is conditional upon *c*, and *ist(c, p)* is mapped to a relation from *c* to *p*. Here one is necessarily talking about the first-order logic as the reified contexts, *c*, can be used in the reasoning itself.

7 Conclusion

Regardless of whether the simple-minded model above turns out to be useful for its designated task, the utility of a similar base model to relate different conceptions of contexts as special cases is needed to provide some reference points for the debate. If the model only performs the function of allowing everybody to state exactly where and how they disagree, then it has succeeded.

References

[1] Cohen, PR; (1985). Heuristic Reasoning about Uncertainty: An ARtificial Intelligence Approach, Pitman, Boston.

[2] Hayes, P; (1995). Verbal summary of AAAI fall symposium on formalising context, MIT, Cambridge, MA.

[3] Horty, JF; 1994. Some Direct Theories of Nonmonotonic Inheritance. In Gabbay, DM; Hogger, CJ and Robinson, JA (eds.); *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Clarendon Press, Oxford, 111-187.

[4] McCarthy, J, 1993. Notes on Formalizing Context. *IJCAI '93 - proceedings of the thirteenth international joint conference on artificial intelligence*. Morgan Kaufman, Palo Alto, 555-560.

[5] Michie, D; Spiegelhalter, DJ; & Taylor, CC; 1994. *Machine learning, neural and statistical classification*. Ellis Horwod, New York, chap. 6, 84-106.

[6] Perlis, D; 1986. On the consistency of commonsense reasoning. *Computational Intelligence*, 2, 180-190.