Contexts, Locality and Generality*

Paolo Bouquet[1,2], Enrico Giunchiglia[3] and Fausto Giunchiglia[1,4]
Mechanized Reasoning Group
[1] I.R.S.T., Trento (Italy).
[2] Dept. of Philosophy, Univ. of Genoa (Italy).
[3] D.I.S.T., Univ. of Genoa (Italy).
[4] D.I.S.A., Univ. of Trento (Italy).
e-mail: bouquet@irst.itc.it. enrico@dist.unige.it. fausto@irst.itc.it

Abstract

It has been recognized that AI programs suffer from a lack of *generality*, the first gross symptom being that a small variation to the problem being solved usually causes a major revision of the theory describing it. The lack of generality seems an unavoidable consequence of the process of approximating the world while building theories about it. In this paper we propose an approach where generality is achieved by formulating, for each problem at hand, an appropriate *local* theory, *i.e.* a theory containing the needed information. The process of theory formulation and reformulation is formalized using *contexts*.

1 Introduction to the Problem

Since the birth of Artificial Intelligence, many formalisms and programs have been devised to model human common sense. The range of applicability of such formalisms has been investigated and some industrial applications based on such formalisms have been made. However, these formalisms (in different modes and to a different extent) lack *generality* in the sense of [?]. In that paper, John McCarthy's writes:

> "It was obvious in 1971 and even in 1958 that AI programs suffered from a lack of generality. It is still obvious and now there are many more details. The first gross symptom is that a small addition to the idea of a program often involves a complete rewrite beginning with the data structures."

*This paper is a re-elaboration of the ideas already presented in [Giunchiglia, 1993, Giunchiglia *et al.*, 1993a, Bouquet and Giunchiglia, 1995].

Such a lack of generality is testified by our failure in building a "general" theory of the world by which we can (i) solve the problems as they stand, and (ii) reason about and solve variations of the same problems. This failure is not surprising: *all* our theories of the world are approximate, *i.e.* they do not describe the world completely but only from a certain perspective, at a certain level of abstraction. The approximations and abstractions performed in the formalization of a problem are necessarily dependent on the problem being considered. Any new previously unconsidered information will in general correspond to a new view of the world which may hide some details and highlight others previously not considered.

In our view, the lack of generality is intrinsic in any theory of the world, since any such theory will be (in different modes and to a different extent) approximate. Generality is achieved formulating, for each problem at hand, an appropriate *local* theory, *i.e.* a theory with the least possible amount of information necessary for solving the problem. Any of such theories will be very approximate, but general enough to allow for a meaningful solution to the problem being considered. New information about the problem will force a revision of the previously used approximate theory and the building of a new, reformulated theory, which will be still approximate but general enough for a meaningful solution to the new formulation of the problem.

In this paper, theories are called *contexts* in that each context is meant to embody the subjective perspective that an individual has about the world. As described in [**?**], contexts are *partial theories* (the individual may have different perspectives and/or models of the world, which is completely described —in the individual's mind— by the set of all its contexts) and also *approximate theories* (a context will never describe the world completely).

The paper is structured as follows. First, in Section **??**, we show how contexts can be formalized as logical theories. Furthermore we formalize the idea of reasoning inside a formal system allowing multiple contexts. Then, in Section **??**, we discuss the relations among contexts, locality and generality, and how we achieve locality and generality using contexts. As a case study, we focus on the qualification problem.

2 Contexts as Formal Objects

The basic intuition for introducing contexts as first class citizens in our formalism, is that reasoning is usually performed on a subset of the global knowledge base; we never consider *all we know* but only a very small subset of it [**?**]. The notion of context is used as a mean of formalizing this idea of localization of the reasoning. Roughly speaking, we take a context to be the set of facts used locally to prove a given goal plus the inference routines used to reason about them (which, in general are different for different sets of facts). The perspective is thus similar to that proposed in [**?**, **?**]. In the rest of this section we formalize these ideas by saying more precisely what we mean by "context" and by "reasoning with contexts". The goal is to model reasoning as deduction.

2.1 Reasoning Inside a Context

The starting point is that the set of facts which provide the context of reasoning is in general only a subset of the knowledge base. We therefore describe the knowledge base as structured into sets of facts, which we write $A_1, \ldots, A_n$. Taking a context to be any A_i would lead to a notion of context which is similar to the notion of a partition in partitioned data bases or to the similar notion of microtheories in CYC [?]. Even though these partitions do not need to be static and fixed once for all, this does not seem a satisfactory enough solution. In fact, in general, each A_i is written using different sets of constants, predicate, and function symbols. For instance, the set of facts about arithmetic will have, as part of the signature, 0, +, and $*$, while the theory of how to get on a plane will use constants like $existsflight$ and $plane$. The formalization of an agent's beliefs about the world or about its beliefs about its own beliefs requires building up theories with different signatures: the second set need only contain the belief predicate and the names of formulas, but not the formulas themselves (see [?], for example). We therefore require that each context come with its own signature. More interestingly, we also take the notion of wellformedness to be *localized* and distinct for each A_i (Notice that in partitioned data/knowledge bases (*e.g.* CYC) the notion of wellformedness is taken to be the same everywhere). This allows us, for instance, to have a context which is essentially a data base of atomic ground formulas and another whose facts express metalevel heuristics, all expressed in clausal form.

We formalize the requirement that each A_i come with its own signature and wellformedness rules by associating a language L_i to each A_i.

The next step is to model reasoning. The standard solution is to have a unique inference engine (possibly consisting of a set of inference modules) which can be applied to any set of facts or, even, to combinations of them. Our proposal is to associate a distinct inference engine to each distinct set of facts A_i. This allows us to localize the form of reasoning and, for instance, to define special purpose inference engines which exploit the local form of wffs. For example, we can use PROLOG on clausal first order languages and Davis Putnam decision procedure [?] on propositional languages.

If we call Δ_i the set of inference rules associated with a set of facts A_i, then we can define a context C_i to be the triple $C_i = \langle L_i, A_i, \Delta_i \rangle$. In other words, we take a context to be a logical theory, presented as an axiomatic formal system. This allows us to take the usual notion of deduction (see for instance [?]) as the formalization of reasoning *inside* a context.

2.2 Reasoning with Multiple Contexts

A knowledge base contains in general a set of interacting contexts $C_1, \ldots, C_n$. We need to capture the idea that reasoning in one context may influence reasoning in other contexts. We introduce therefore a new set of rules which allow us to derive a fact in a context because we have derived other facts in other contexts. Thus, if $c\!:\!\alpha$ denotes the formula α in the context c, such inference rules look like

$$[\rho] C\!:\!A C_1\!:\!A_1, \ldots, C_n\!:\!A_n$$

Inference rules like ρ (with premises and conclusions in different contexts) are called *bridge rules* [**?**] as they allow us to bridge deductions in different contexts. In particular, ρ allows us to derive A in C just because we have derived A_1 in C_1, ..., A_n in C_n. We say also that A in C is a *justified assumption* as it is an assumption we make in C which is justified by derivations in other contexts. Notice that a notion similar to that captured by McCarthy's lifting axiom [**?**] can be formalized by the following bridge rule:

$$\frac{c\!:\!A}{c'\!:\!ist(A,c)}$$

which intuitively says that, if we can prove A in context c, then we can prove (in a context c') that we can prove A in c.

Contexts and bridge rules are the components of *multicontext systems (MC systems)* defined as a pair $\langle \{C_i\}_{i\in I},\ BR \rangle$, where $\{C_i\}_{i\in I}$ is a family of *contexts* and BR is a set of *bridge rules*. Derivability in a MC system is defined in [**?**] and [**?**]; roughly speaking, it is a generalization of Prawitz' notion of deduction inside a natural deduction system.

MC systems are a powerful formal tool for the formalization of common sense. Even restricting the language of each context to be first order, it seems possible to build particular MC systems which are provably equivalent to richer formalisms than (first order) classical logic is (see for example [**?**, **?**, **?**]).

3 Contexts, Locality and Generality

A representation in a MC system is not necessarily local. For example, the representation could use a single context, and thus be reducible to a standard representation in an axiomatic formal system. However, MC systems provide the tools to achieve locality: the possibility of structuring knowledge in contexts, and of considering only a subset of the defined contexts.

Locality does not imply generality: a representation may be local in that it considers only the information necessary for the problem, but that representation is also not general since peculiar for solving that very problem. As we said, generality is achieved formulating and reformulating the local theory where the problem has to be solved. MC systems provide the formal tool to achieve generality: starting from a set of already defined contexts, each containing the knowledge necessary for a specific (sub-)problem, bridge rules allow to build up the appropriate context.

In the next subsections, we will show how to put these ideas into practice for solving the qualification problem.

3.1 The Qualification Problem

As pointed out in the introduction, the lack of generality of our representations is testified by our failure of building a unique theory of the world by which we can solve (variations of) the problems we have to deal with. A clear effect of this lack of generality is the so-called *qualification problem*, which arises in the attempt of formalizing reasoning about action and change [?]. Consider the following axiom about air traveling (free variables are universally quantified throughout the paper):

$$fly(x,y) \supset at(y) \tag{1}$$

Intuitively, axiom (**??**) says that flying from a place x to a place y causes being at y. This axiom can be useful in very generic circumstances. However, it is quite clear that it needs a lot of qualifications. For instance, there must be a flight connection between x and y; the traveler must be at x in order to get the flight; he must have the ticket. Let us abbreviate with P these three preconditions. Then, instead of (**??**), we can write:

$$P \supset (fly(x,y) \supset at(y)) \tag{2}$$

axiom (**??**) is more general than axiom (**??**), because it depends on a smaller set of implicit assumption. Still, it implicitly presupposes some temporal relation, namely that the preconditions are satisfied in the situation where the action takes place. This can be made explicit with an axiom like:

$$holds(P,s) \supset holds(at(y), result(fly(x,y),s)) \tag{3}$$

whose meaning is that for any situation s, if the preconditions P hold in the situation s, then in the situation resulting from flying from x to y, $at(y)$ holds. Are we done? Not quite. Many other facts qualify the truth of (**??**). A blatant example is a pilots strike, that would invalidate the conclusion. Also being clothed is usually a pre-requisite for getting on a plane. But then these two facts should be added to P. It is clear that this process can go on as long as we like. The problem is that axioms like (**??**), (**??**) and (**??**) are not eternal, in the sense of [?]. Transforming them into eternal sentences would require to explicit all contextual dependencies. Unfortunately, in general this is not possible. As McCarthy writes in [?],

> "...the axioms we devise to express common sense knowledge are too restricted in their applicability for a general common sense database. ...With a little ingenuity the critic can usually devise a more general context in which the precise form of the axiom doesn't hold [...]"

From this perspective, the qualification problem can be restated as the problem of *conjecturing* that a certain theory is detailed enough to solve a problem and of *revising* this conjecture once it turns out that this is not the case. The conjecturing activity can be divided in two steps:

1. first, we conjecture that a certain theory including only part of the knowledge of an agent is adequate for reasoning about a given problem;
2. second, we use the selected theory for reasoning *as if* such theory contained all we know about the problem to solve. This can be done by the "usual" nonmonotonic reasoning

However, one of these two conjectures can fail, causing a revision:

1. some new fact could force the reasoner to retract the conjecture that the theory is adequate for the problem (the set of considered qualifications is too small). An example is the pilot strike in the air-traveling example;
2. a new fact might cause a violation of an explicitly considered qualification. This case can be handled by the usual nonmonotonic revision inside the theory. For example, in the air-traveling example we could add the fact that the traveler does not have the ticket

The big picture of the solution we propose is as follows. We start with the formulation of a problem. In the background, we have a commonsense knowledge base (KB) which contains the knowledge of an agent. The KB is not a unique theory; it is structured as sets of facts, each about a particular topic. From now on, we informally call *KB contexts* these sets of facts. Each KB context is associated to one (or more) symbol(s) which can be used to formulate problems. This constitutes an *associative memory*. We use this associative memory to state a form of *relevance claim*: if a certain symbol is associated to a KB context, then this KB context is assumed to be relevant for any problem in which this symbol appears. For instance, we can imagine that mentioning tickets in the air-traveling example would imply that the KB context(s) containing knowledge about tickets will be inferred to be relevant. This exploits the implicit assumption that the statement of the problem fixes the description of the world which should be used to solve the problem; in other words, the amount of detail which should be considered. In our terminology this means that any specific problem formulation drives the choice of a theory at a given level of detail and, analogously, any specific problem reformulation allows us to confirm or revise this choice. This kind of relevance claim is quite simple. In principle, we can introduce arbitrarily complicated relevance claims. Their form may depend on many factors. In commonsense reasoning, for instance, it may depend on the confidence that people have in their ability of solving the problem, their current attitude (more or less oriented towards details), the time they have to solve the problem (taking into account more details requires more time) and so on. However, even such a simple form of relevance claim allows us to exploit the advantages of localizing the reasoning to subsets of a knowledge base (see for instance [**?**]).

The relevance claims and the associative memory define a space of theories. These theories are called *workspace contexts* (WS contexts from now on). A WS context is adequate for a problem if it contains *all* the assumptions of the problem

Figure 1: A contextual architecture for the GLM example

and *all* the relevant KB contexts. This is what we call a *theory adequacy* requirement. Clearly, we can only conjecture that this form of adequacy actually holds. Indeed, any new fact added to a problem formulation can make it fail.

3.2 A Simple Example: the GLM Problem

The example, originally proposed in [?], concerns an air traveling domain. The scenario is as follows. A traveler must go from Glasgow to Moscow and he knows that there exists a flight from Glasgow to London and another from London to Moscow. The idea is formalizing this reasoning: (a) after the two flights, the traveler will be at London; (b) if he loses the ticket at London, the plan fails; (c) a new plan including the purchase of a new ticket will succeed. The point is that the number of possible obstacles to the plan (losing the ticket is only an example, also a pilot strike or the theft of the traveler's clothes in the toilet of the airport are possible obstacles) is huge, and we cannot sensibly take all of them into account.

The proposed solution is based on the following idea: to keep the reasoning as simple as possible, reasoning must be carried on in a context where it is assumed that no obstacles occur unless explicitly stated. Therefore, given the part (a) of the problem, we want to explicitly use only the fact that the two flights exist. Since tickets are not even mentioned, we want our context to implicitly assume that knowledge about tickets can be disregarded in order to infer the success of the plan. When the loss of the ticket is mentioned, the above assumption must be revised and tickets must be reasoned about (in particular, the fact that having the ticket is a precondition for getting on a plane). Finally, using information about buying things, we can conclude (c).

The architecture of the solution is depicted in figure **??**. Each circle represents a context. The contexts on the left are the KB contexts; each of them contains portions of the information on the domain that the agent has at his disposal (*e.g.* on tickets, on flights, and so on); in particular, the context *Action* contains axioms for reasoning about action and change in general (actually the axioms of [?]). The contexts on the right are the WS contexts; they are built up by "merging" the information contained in some of the KB contexts. The context with label C is called the control context; C controls the flow of information from KB contexts to WS contexts. The idea is that in C are specified some conditions under which the system can infer that a KB context is relevant for a problem. For instance, if a problem does mention the loss of the ticket, the system assumes that the facts in the KB context *Ticket* must be loaded in the WS context of that problem. This is stated by axioms of this form:

$$COND \supset Lift(kb, wrkc(p)) \tag{4}$$

where *COND* states the condition for the KB context *kb* to be relevant for a problem. The axiom says: if we can prove that the KB context *kb* is relevant to the solution of a problem p, then lift *kb* in the workspace of the problem p. The notion of lifting a context into another is defined in the following axiom of C:

$$Lift(c1, c2) \equiv \forall w \ (Ist(c1, w) \supset Ist(c2, w)) \tag{5}$$

The consequence of this axiom is that $c1$ becomes a "subset" of $c2$, since it imposes that any formula that can be proved in $c1$ be provable also in $c2$. This intended meaning is given by the following bridge rules on the predicate *Ist*:

$$c\!:\!\alpha C : Ist(c, \alpha) \qquad\qquad C : Ist(c, \alpha) c\!:\!\alpha$$

where c is any context different from C (the control context). The rule on the left is called reflection down, that on the right reflection up. Reflection down says that we can derive α in c whenever we have derived $Ist(c, \alpha)$ in the control context C. Reflection up has the dual meaning.

Now let us imagine that the system is given the formulation of a problem $P1$ such that the only KB context provably relevant is *Flights* (it contains the information that the existence of the flight is a precondition for flying). Then, the axioms of C allow us to lift the context *Flights* into the WS context of $P1$, denoted by $wrkc(P1)$ (because of an axiom like (**??**) relative to *Flights*). In $wrkc(P1)$ we can prove that the only relevant precondition for the success of the plan is the existence of the flights from Glasgow to London and from London to Moscow. Since we know that the two flights exist, the success of the original plan can be proven.

This conclusion depends on the implicit assumption made in $wrkc(P1)$ that all the preconditions other than the existence of the flights can be disregarded. However, this assumption is defeated whenever the loss of the ticket is explicitly mentioned in the reformulation of the problem (say $P2$). In this case, an axiom like (**??**) is used to prove that also the context *Ticket* must be lifted in the WS context of $P2$. Indeed, in $wrkc(P2)$ having the ticket is now explicitly considered as a qualification, so that its loss is a sufficient reason for inferring the unsuccess of the original plan. However, since the context *Ticket* contains also the information that tickets can be bought, we can conclude that a new plan including the purchase of a new ticket will eventually succeed.

3.3 Related Work

The qualification problem has been first identified in [**?**]. Since then, many theories have been developed to deal with this problem (see for example [**?**, **?**, **?**]) despite some technical differences, all such theories share a common approach: they assume that all qualifications for the action under consideration are always explicited in the theory that is used to reason about the problem; then it is conjectured that all such qualifications are satisfied unless explicitly stated. (Notice that in our approach this would correspond to the case in which all the KB contexts are lifted in the WS context of the problem).

However, as Ginsberg and Smith point out [**?**], the overall qualification problem consists of three distinct difficulties:

1. the language or ontology may not be adequate for expressing all possible qualifications of an action;

2. it may infeasible to write down all the qualifications (even if the ontology is adequate);

3. it may be computationally intractable to check all the qualifications

Standard approaches do not address the first difficulty and may cause problems with the second and the third. In the same paper, Ginsberg and Smith propose two new approaches which (partially) solve the second and third difficulty, but not the first. On the other hand, our approach is aimed at solving the first difficulty. Each WS context has a language and a set of facts which may not be adequate for expressing all possible qualifications of an action. However, the WS context adequacy is only conjectured and this conjecture can be revised. This allows us to classify the qualifications for a problem into three groups: the set of all the facts the reasoner knows about and that potentially qualify the solution; the subset of these facts that the reasoner explicitly considers in a given formulation; all the unexpected or unknown qualifications. All the previous approaches don't distinguish between the first and the second kind of qualifications. Indeed, any qualification which is included in the theory has to be explicitly considered.

Acknowledgments

John McCarthy has provided many motivations and intuitions. Vladimir Lifschitz and Luciano Serafini have given important suggestions on previous versions of this paper.

References

[Bouquet and Giunchiglia1995] P. Bouquet and F. Giunchiglia. Reasoning about Theory Adequacy. A New Solution to the Qualification Problem. *To appear in Fundamenta Informaticae*, 1995. Also IRST-Technical Report 9406-13, IRST, Trento, Italy.

[Davis and Putnam1960] M. Davis and H. Putnam. A computing procedure for quantification theory. *Journal of the ACM*, 7:201–215, 1960.

[Ginsgerg and Smith1988] M.L. Ginsgerg and D.E. Smith. Reasoning about Action II: the Qualification Problem. *Artificial Intelligence*, 35:311–342, 1988.

[Giunchiglia and Giunchiglia1995] F. Giunchiglia and E. Giunchiglia. Ideal and Real Belief about Belief. Technical Report 9501-02, IRST, Trento, Italy, 1995. Also available as MRG-DIST Technical Report 95-0026, DIST, University of Genova, Italy.

[Giunchiglia and Serafini1994] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics (or: how we can do without modal logics). *Artificial Intelligence*, 65:29–70, 1994. Also IRST-Technical Report 9110-07, IRST, Trento, Italy.

[Giunchiglia and Weyhrauch1988] F. Giunchiglia and R.W. Weyhrauch. A multi-context monotonic axiomatization of inessential non-monotonicity. In D. Nardi and P. Maes, editors, *Meta-level architectures and. Reflection*, pages 271–285. North Holland, 1988. Also DIST Technical Report 9105-02, DIST, University of Genova, Italy.

[Giunchiglia *et. al.* 1993a] F. Giunchiglia, E. Giunchiglia, T. Costello, and P. Bouquet. Dealing with Expected and Unexpected Obstacles. In *Proceedings Second Symposium on Logical Formalizations of Commonsense Reasoning*, Austin, Texas, 1993. Also IRST-Technical Report 9211-06, IRST, Trento, Italy.

[Giunchiglia *et. al.* 1993b] F. Giunchiglia, L. Serafini, E. Giunchiglia, and M. Frixione. Non-Omniscient Belief as Context-Based Reasoning. In *Proc. of the 13th International Joint Conference on Artificial Intelligence*, pages 548–554, Chambery, France, 1993. Also IRST-Technical Report 9206-03, IRST, Trento, Italy.

[Giunchiglia1991] F. Giunchiglia. Multilanguage systems. In *Proceedings of AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning*, 1991. Also IRST-Technical Report 9011-17, IRST, Trento, Italy.

[Giunchiglia1993] F. Giunchiglia. Contextual reasoning. *Epistemologia, special issue on I Linguaggi e le Macchine,*. XVI:345–364, 1993. Short version in Proceedings IJCAI'93 Workshop on Using Knowledge in its Context, Chambery, France, 1993, pp. 39–49. Also IRST-Technical Report. 9211-20, IRST, Trento, Italy.

[Guha1990] R.V. Guha. Microtheories and contexts in cyc. Technical Report ACT-CYC-129-90, MCC, Austin, Texas, 1990.

[Lansky1991] A. Lansky. Localized search for multiagent planning. In *Proc. of the 12th International Joint Conference on Artificial Intelligence*, pages 252–258, 1991.

[Lifschitz1987] V. Lifschitz. Formal theories of action. In F. M. Brown, editor, *The frame problem in Artificial Intelligence*, pages 35–58. Morgan Kaufmann. Los Altos, CA, 1987.

[Lin and Reiter1994] F. Lin and R. Reiter. State constraints revisited. *Journal of Logic and Computation*, 4:655–678, 1994.

[McCarthy1977] J. McCarthy. Epistemological Problems of Artificial Intelligence. In *Proc. of the 5th International Joint Conference on Artificial. Intelligence*, pages 1038–1044, 1977. Also in V. Lifschitz (ed.), *Formalizing common sense: papers by John McCarthy*, Ablex Publ., 1990, pp. 77–92.

[McCarthy1987] J. McCarthy. Generality in Artificial Intelligence. *Communications of ACM*, 30(12):1030–1035, 1987. Also in V. Lifschitz (ed.), *Formalizing common sense: papers by. John McCarthy*, Ablex Publ., 1990, pp. 226–236.

[McCarthy1991a] J. McCarthy. Notes on Formalizing Context. Unpublished, 1991.

[McCarthy1991b] J. McCarthy. Overcoming an Unexpected Obstacle. Unpublished, 1991.

[Prawitz1965] D. Prawitz *Natural Deduction - A proof theoretical study.* Almquist and Wiksell, Stockholm, 1965.

[Quine1969] W. V. Quine. Propositions. In *Ontological Relativity and other Essays.* Columbia University Press, New York, 1969.