


In MultiContext systems, these compatibility relations are formalized as a special kind of inference rules, called bridge rules\textsuperscript{11}. The form of a bridge rule is the following:

\begin{equation}
\phi_{n+1} \vdash c_j : \phi_1, \ldots, \phi_n
\end{equation}

where $c_i$ and $c_j$ are contexts and $\phi_1, \ldots, \phi_n$ are formulae belonging to the language of $c_i$ and $\phi_{n+1}$ is a formula belonging to the language of $c_j$. Semantically, the effect of a rule like this is to put a constraint on what counts as a model of a context: a local model of $c_j$ that does not satisfy $\phi_{n+1}$ is not compatible with a local model of $c_i$ that satisfy $\phi_1, \ldots, \phi_n$. It's quite simple to generalize bridge rules in such a way that the premises belong to different contexts.

Acknowledgments

This work is a first attempt of clarifying some intuitions underlying most of the work which has been done by the Mechanized Reasoning Group in the last ten years. The authors are particularly indebted with Massimo Benerecetti, who read earlier versions of this paper and made lots of valuable comments, and with Chiara Ghidini, who helped clarifying some intuitions behind the semantics of contextual reasoning. A very important contribution came from discussions with Alessandra Ciceri, Alessandro Cimatti, Luciano Serafini, and with all other members of the group.

References


\textsuperscript{11}Bridge rules were introduced and intuitively discussed in [11]. A more technical presentation can be found in [13], where it is proved a syntactical equivalence between the most important classes of modal systems and a multi-context framework.

\textsuperscript{12}
everyday life, we all often compare beliefs, discuss beliefs, learn from other people’s beliefs, and so on. Does locality entails that we must give up all of this? Of course not. However, a different model is needed. Consider the following situation. There are two agents (say Paolo and Fausto), and Fausto is ready to believe whatever Paolo says. Now Paolo tells Fausto that he believes that \( \phi \) (e.g. that Sofia is a nice city). The traditional formalization of this situation would involve an axioms whose content can be approximately described as follows: if Paolo believes \( \phi \) and Paolo tells Fausto that he believes \( \phi \), then Fausto believes \( \phi \). This axiom is in contrast with locality because the meaning of the sentence \( \phi \) is given independently of Paolo’s and Fausto’s belief contexts. Our explanation of the matter is therefore the following:

- Paolo believes a fact \( F \) that he represents as \( \phi \);
- Paolo wants to communicate this fact to Fausto and therefore utters the sentence \( \phi \);
- Fausto hears \( \phi \) that, in his own representation language, represent some fact \( F' \);
- Fausto believes the fact \( F' \).

This description seems the description of a communication failure, but it is not. On one side, we take seriously the idea that different contexts (in this case belief contexts) have different languages, which are interpreted locally; hence we do not “hardwire” in our explanation the fact that the same string (or sound) must always have the same meaning for Paolo and Fausto. On the other side, however, we are not imposing that \( F \) and \( F' \) must be different facts. It can certainly be the case that they are the same fact (and in 99% of cases it is really convenient to assume as a default that they are the same fact!), but our explanation shows that the basic case is the one where \( F \) and \( F' \) are not the same fact.

The simple example above is an instance of a more general phenomenon of contextual reasoning: the truth of a sentence in a context may be connected with the truth of a sentence (or set of sentences) in other contexts. We generalize this fact in the following principle of compatibility: two contexts may be related in such a way that the truth of a sentence (or set of sentences) in one of them entails the truth of some other sentences in the second. Such a relation between contexts is called a compatibility relation. Some examples:

- the truth of the sentence, “Yesterday it was very hot” in the context of Friday July 19, 1996 is entails the truth of the sentence, “Today is very hot” in the context of Thursday July 18, 1996;

- the truth of the sentence, “Fausto believes that winter in Trento is cold” in the context of Paolo’s beliefs entails the truth of the sentence, “Winter in Trento is cold” in the context of Paolo’s beliefs about Fausto’s beliefs;

- the truth of the sentence, “Holmes is a detective” in the context of Sherlock Holmes stories entails the truth of the sentence, “In the context of Sherlock Holmes stories, Holmes is a detective” in an outer context.
systems. Intuitively, in a propositional modal system, what can be said is represented as a set of atomic sentences, each of which is true or false (but not both) at every world. But then it is easy to see that what can be said is the same at any world. This is equivalent to saying that in modal systems expressibility is not a local, but a global property.

The language of each context must be given a local interpretation. Even if two terms (sentences) of two different contexts “look the same” (i.e. the symbols we use to build them are the same), they are not the same term (sentence). Suppose that both Paolo and Fausto have in their language the term ‘the US President’: the interpretation of the two terms (we stress that Paolo and Fausto have two terms, and it is not the case that they use the same term) is in principle independent. The adjective ‘slow’ in the context of geology and in the context of computer science has a very different interpretation. Likewise for the interpretation of sentences. The sentence “It’s cold” may be used by Paolo to describe a situation where the temperature is 10 degrees below zero, and by Fausto to describe a situation where the temperature is 5 degrees above zero. As we will see, it is possible (and usually convenient) to force some relation between terms (sentences) of different contexts. But this does not mean that we can assign to them an interpretation independently of the context.

The example of the train schedule and the phone book shows that, depending on the context, we may adopt different reasoning strategies. From a logical point of view, this means that we have different relations of logical consequences in different contexts. In the example, in the context of railway system, we are allowed to apply a closed world assumption (i.e. we derive negative information from the absence of positive information); in the context of a phone book, this cannot be done. Another case is that of cognitive systems with different computational abilities and resources.

[11] proposes to formalize the principle of locality by defining each context as a logical theory presented as an axiomatic system. This means that a context is viewed as a triple \( \langle L, \Omega, \Delta \rangle \), where \( L \) is said the language of the context, \( \Omega \) is a set of axioms, and \( \Delta \) is a set of inference rules defined over \( L \). Two contexts have distinct languages, so expressibility is local; have different axioms; have their own set of inference rules. This means that each context may have its own logic.

### 5.2 Compatibility

The emphasis on locality does not mean that there is no relation between the contents of different contexts. On one side, the principle of locality is not compatible with the idea that the “same” sentence can be evaluated in different contexts; on the other, it does not exclude that sentences in different contexts may be somewhat related. In order to give some intuitions, let us start from a simple example involving beliefs.

In many traditional approaches to formalizing beliefs, on one hand there is a sentence (say \( \phi \)), and on the other hand there are agents that believe (not believe, disbelieve) this sentence (incidentally, notice that every agent must either believe, disbelieve or not believe every sentence. This overload of beliefs is a consequence of having a unique – i.e. global – language in which all is expressed). This is in contrast with the principle of locality, as it implicitly assumes that \( \phi \) must have the same meaning for every agent. However, in
5.1 Localy

The name ‘Holmes’ refers to the detective Holmes in the context of Sherlock Holmes stories; it refers to the judge Holmes in the context of the US legal history; and perhaps does not refer to anybody in the context of Ming sculpture history. The sentence “France is exagonal” is true in the context of an informal discussion among friends about the shape of European countries; is patently false in a geometrical setting; is nonsense in the context of Greek geography of IV century B.C. (intuitively, because France doesn’t even exist in the ontology of that context). As to reasoning, if a train schedule booklet says that there is a connection from Trento to Verona at 9am and another at 11am, we are allowed to derive that there is not a train at 10am; however, if Mr. Clinton’s phone number is not on the Washington D.C. phone book, we are not entitled to infer that he has no phone number.

These examples, and many others that could be done, show that what can be said and what is true (or false) are always relative to a context. We state this as a general principle, the principle of locality: expressibility and the basic semantical relations (denotation, truth, logical consequence) are local to a context. Let us analyze some consequences of this principle in a formalization of context.

What can be said depends on the language that one has at his/her disposal. We mentioned this fact when discussing figure 4. In a formal setting, this is tantamount to say that there are multiple distinct languages, each of which is local to a context. One single language would not do, since it would make expressibility uniform. Notationally, we stress the fact that language is local to a context with the notation $c : \phi$. Its intuitive meaning is: $\phi$ is a formula belonging to the language of context $c$.

Each expression – in particular any sentence – belongs to the language of a context: there is not something like a universal language. The fact that an agent chooses a language to describe what he knows about something presupposes the choice of an ontology, and manifests some (implicit) assumptions concerning what is relevant and what is not. If Fausto were required to describe his office, he could start speaking of chairs, table, windows, and so on. As we said, this presupposes the choice an ontology (which includes those objects and not – for instance – powder or wires), and presupposes that his description matches the inquirer’s intentions. Were Fausto speaking with someone who is supposed to fix some electrical problem, he would probably mention plugs, lights, and so on. Both the choice of an ontology and the choice of a level of description are mostly implicit, and depend on the context. In a formalization of context, the choice of the language of a context is therefore a way of modelling part of the effects of the cognitive function which generated the corresponding view.

Philosophically, there is an essential difference between saying that any linguistic expression belongs to the language of a context and saying that there is a unique language which can be used to state facts in different contexts. This is one major difference with modal

---

10 This is our way of intending what Bar-Hillel hints in the definition of the relation $RP(a,b,c)$, namely that context includes also a reference to a language. This idea may also be connected with Lewis’ idea that a grammar for English “culminates with a specification of the conditions under which someone tells the truth-in-English”. In other words, a grammar is always a grammar for some language, and gives the conditions under which a speaker tells the truth in that language.
In the latter, the cognitive function that constructs a view is taken as a black box, and only its results are taken into account. So, if from a cognitive point of view, we are interested in knowing how is it that Fausto came represent his belief that winter in Trento is cold with the sentence “In Trento winter is cold”, from an epistemological perspective we are interested in the fact that such a sentence happens to belong to the set of Fausto’s beliefs. The set of sentences that describe the contents of a view is the object that intuitively corresponds to what we call context (see figure 5).

Figure 5: Context

In our opinion, this notion of context captures all the intuitions that are behind what we have been calling cognitive context so far. First of all, it is easily mapped onto the intuition that cognitive contexts belong to an individual’s cognitive state rather than to the world. Indeed, a context is thought of as the set of sentences that describe a view, and a view is part of an individual’s cognitive state. Any sentence is not just true or false, but it is true or false with respect to a context: true if it describes something that belong to a view, false otherwise. It should also be clear that a context can be considered a space in the sense of [6], namely some logically coherent reality, in which various propositions are treated as true, objects are assumed to exist, and relations between objects are supposed to hold. ‘Holmes’ is part of any view of Sherlock Holmes stories (but an agent might not have such a view), and also of the view of US legal history that can be found in a book on that topic. Also, only some objects belong to a context (namely those that belong to the corresponding view), and they are all the objects that have some causal rôle in reasoning (see Kokinov’s definition of context in section 2.2); any object that does not belong to a context cannot affect reasoning inside the context itself. As a view depends on the parameters of the cognitive function that builds it, so a context implicitly depends on a set of assumptions that partially represent the parameters of the cognitive function. In this sense, context as defined here is similar to context as defined by Sperber and Wilson.

5 Contextuality = Locality + Compatibility

The idea of contexts as views provides us with an intuitive model of a logic of contextual reasoning. In this section, we present two principles that, in our opinion, must be put are at the basis of any such a logic: the principle of locality and the principle of compatibility. Here these two principles are discussed rather informally; they are formally embodied in the definition of MultiContext systems as given in [11].
views partly depend on the way the world is, but they are not just “pictures” of the world. A view can rather be thought of as the result of a cognitive function from a domain to a representation of that domain. This function may have several parameters, like: physical capabilities, observation opportunities, individual biases, background knowledge and beliefs, mood, goals, desires, and so on. The value of these parameters depends (among other things) on the cognitive state of the agent that builds a view.

Views can be studied from two very different perspectives:

- the first is cognitive, and concerns the process of view-construction, that is the process by which – in a given circumstance – a given cognitive system constructs a view. This kind of investigation is empirical, and has to do with the cognitive faculties of an agent;

- the second is epistemological\(^2\), and concerns the contents of a view independently of any cognitive (empirical) mechanisms of view-construction and even of any particular cognitive system (whether human or artificial).

\(\text{could even see Kant's critic philosophy in this light), and with the opposite idea that there are idiosyncratic ways of building views.}\)

\(^2\text{We use the word ‘epistemology’ in the sense of [25].}\)
belongs to the structure of the domain we want to formalize (the world): a domain is not just a flat collection of objects, but a more complex structure which includes some ‘points’ (i.e. contexts), each of which gets associated an agent, a time, a place and a possible world. This is reflected in the notion of a structure for L.D. This is the technical conterpart of the intuitin that pragmatic contexts are part of the world.

However, the approach of figure 3 does not seem to account for many aspects of cognitive context that we have been presenting throughout this paper. A cognitive context is part of the cognitive state of an agent; it is part of the representation of the world rather that of the world itself. This fact has many consequences. For instance, a context can describe only a portion of a domain; the same domain can be described at many different levels of detail; different agents may have different descriptions of the same domain (even contradictory). How are these phenomena to be accomodated in a formalization?

Our solution is to complicate a little the traditional model. The idea is that we do not formalize (a portion of) the world, but rather a view of some portion of the world, as depicted in figure 4. Each view is thought of as an agent-centered representation of some domain\(^8\) and not just a description of the world “as it really is”. In other words,
to be used – in the speaker’s intentions – in interpreting what is said. This impossibility is much more common than one might think at first. Every communicative act is potentially subject to this kind of partial or complete misunderstanding, since in practice we have no access to other agents’ cognitive state. Therefore, any theory of communication assuming that speaker and hearer share the same communication context suffers from the same cognitive implausibility that we ascribed to Kaplan’s logic.

As to the second kind of impossibility, many cases could be mentioned. Recall, for instance, the discussion on the problem of generality and on the qualification problem: every common sense general axiom depends on a huge number of qualifications, most of which we are not aware of. The more qualifications we include, the more a representation is general. However, we are never guaranteed that we have reached the most general representation! Even the most sophisticated representations of the world, i.e. scientific theories, always depend on some hidden qualification, and sometimes the discovery of one such qualification results in a confutation of the theory itself.

In conclusion, the notion of context is essential in a theory of knowledge representation because on one side we may lack information on the assumptions underlying a given representation, and on the other side we may even lack information on the fact that a representation depends on something. This radical incompleteness must therefore be a central property of any formalization of context.

4 Toward a formalization of context and contextual reasoning

Formalizing knowledge can in general be viewed as a two step process. First, we select a universe of discourse (or domain) and construct a formal model (an abstract representation) of it; then we formally specify the language to be used for describing the selected domain and define a mapping from the language to the formal model of the domain. This process is completely general, namely it does not depend on what we take as our universe of discourse. Traditionally, the universe of discourse is thought of as a portion of the world. The situation is depicted in figure 3. At the bottom there is the word with some possible domains (e.g. cooking and weather). In a domain there are objects (e.g. salt, knives, tables, ovens, stoves, clouds, tornadoes, thunderstorms, temperature, pressure, satellites, and so on); objects have some properties; among objects some relations hold. The second level is the level of formal models, where a domain is usually represented in a very abstract form using the mathematical concept of set; so a property, e.g. being red, is representend as the set of all objects that are red; a binary relation, e.g. an object being on another, is represented as the set of all ordered pairs of objects that are one on the top of the other; and so on. Finally, the third layer is the level of the language that is used to describe the domain. The relation between a formal model and a formal language\textsuperscript{7} is the object of formal semantics [30].

The question is: where are contexts in this picture? Kaplan would say that contexts

\textsuperscript{7}In general, a formal language has more than one model. However, this fact is not relevant for our argumentation, and so we ignore it.
and fully competent. However, we think it incorrect to embody this assumption in the logic itself.

Our conclusion is therefore the following: in reasoning, information which depends on pragmatic context is taken into account and dealt with only as far as it is represented as part of the state of a cognitive system. This is the precondition for its use in any reasoning task. In Dinsmore's framework, we could say that pragmatic information must be part of the primary context of any indexical sentence. In Koko's terms, pragmatic features of context are reasoned about only if they are included in the set of the entities that influence the interpretation of a sentence in a particular circumstance. In McCarthy's approach, pragmatic information are part of the facts that one knows about context (and are therefore stated in the outer context). According to Giunchiglia, pragmatic information must be part of the portion of the state of the individual when he/she evaluates an indexical sentence. All these frameworks share the idea that nothing can be reasoned about which is not represented (and in some sense "activated") in the current state of a cognitive system. And this is what we mean when we say that, in modelling reasoning, dependence on pragmatic context either is viewed as a form of dependence upon cognitive context or does not play any rôle at all.

3.2 Why we do need cognitive context

We haven't still proved that context is an essential building block in a theory of knowledge representation. To do that, we need a refutation of the following more radical version of the transformability statement: any context-dependent sentence can be transformed into a sentence whose semantic value independent of context. If we can argue that this generalized transformability statement (GTS) is false, then we are almost done with the objection on the necessity of contexts.

Our argument against GTS has been partially anticipated in the previous sections. Very concisely, it can be stated as follows: there are sentences (or sets of sentences) that cannot be fully decontextualized because in practice a finite agent cannot reach a complete knowledge of the context. Notice that, in this formulation, the emphasis is on the phrase "in practice". This phrase must be read in a very strong sense. It is not just the case that sometimes an agent, for some reason, lacks knowledge about the value of some contextual parameter a sentence (or set of sentences) depends upon (as it was the case of Tom Brown, who did not know the time because his watch was under repair). The situation is much worse:

• on one side, there are dependencies that cannot be accessed by an agent;
• on the other side, context-dependence can be so complex and deep that no finite agent can in general have a full knowledge of it.

A typical example of the first kind of practical impossibility is communication. Recall the example of the cup of coffee: a hearer cannot know what are the assumptions that ought
Brown wants to have breakfast in his bed”, his wife must know that the current date is the first of January 1951, at 9 o'clock in the morning, that the person who is speaking is Tom Brown, and that the bed where he is lying is his bed. Since it is possible for an agent to have one kind of knowledge without having the other, it follows that – in actual communicative situations – such an agent might lack the ability of transforming an indexical into a non-indexical sentence:

Since our knowledge is limited, the use of indexical expressions seems therefore to be not only most convenient in very many situations – nobody would doubt this fact – but also indispensable for effective communication [2] [p. 368-369].

Bar-Hillel’s argument is very different from Kaplan’s, since it shifts the focus of the question. Kaplan assumes the very idealized perspective of an external (omniscient) observer, who knows the value of the pragmatic parameters of every context, actual or merely possible. Bar-Hillel assumes the perspective of a limited agent who uses sentences to represent cognitive contents and may lack information on the value of the pragmatic parameters of some context. This difference has dramatic consequences: if we were ideal (omniscient) agents, we could get by without demonstratives, since we would know how to express every cognitive content in a non-indexical form. So a logic like I.D, for an ideal agent, is pointless. But since we are limited agents (with limited cognitive abilities), we need a logic of indexical languages different than LD, since it presupposes knowledge that in general we cannot assume to have. A real agent may lack information about the values of contextual parameters and therefore the ability of seeing the equivalence between the indexical and the non-indexical representation of the “same” fact. The moral is the following: it is certainly true that the content of a sentence like “I'm here now” depends (among other things) on the speaker, the place and the time of production. However, for an individual to be able to use this fact in reasoning, two more conditions must be fulfilled:

(a) such an individual must know that the meaning of the sentence “I’m here now” pragmatically depends on the speaker, the time, and the place;

(b) the value of these pragmatic parameters (speaker, time, and place) must be represented as a part of the cognitive state of such an individual.

Unless we take into account these two conditions, we will not be able to make a plausible logic of contextual reasoning, since it is possible that some agent lacks knowledge about (a) or (b) or even both6. Sure, in some cases it is useful to assume that agents are omniscient

---

6A nice classical example of this lack of information is proposed by Perry in his paper on The problem of essential indexical [26]:

I once followed a trail of sugar on a supermarker floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me.
I was the shopper I was trying to catch [26] [p. 3].

On the same line of thought are the arguments of Lewis in his paper “Attitudes De dicto and De se” [18]: Lewis argues that we cannot get by without indexical languages because indexicals are used by agents to
A different argument is proposed by Bar-Hillel. He notices that behind the explicit
decision of some philosophers (e.g. Carnap in his *The logical construction of the world* [5])
not to undertake a logic of indexical language there is the assumption that “non-indexical
languages are sufficient for the formulation of any given body of knowledge”. However, Bar-
Hillel notices that language is not used only for the formulation of bodies of knowledge,
but also for communication. So, the question is:

Could we assert [...] that non-indexical languages are sufficient for every communica-
tive purpose? If this were true, if one could always express every cognitive content
in a non-indexical language, the urgency of an investigation of the logic of indexical
languages would be somewhat reduced. [2] [p. 367]

This shift from the problem of representing “bodies of knowledge” to the problem of
representing “cognitive contents” is crucial in Bar-Hillel’s strategy. Indeed, if on one
side Bar-Hillel takes for granted that any body of knowledge can be represented in non-
indexical languages, on the other side he shows that in general it is not true that in every
communicative situation an indexical sentence expressing some cognitive content can be
transformed into a non-indexical sentence (statement) expressing the same cognitive con-
tent. He makes this point with the following thought experiment (*Gedankenexperiment*):

Assume that Tom Brown is a logician interested in our problem who has decided to
find out whether he could get along, for just one day, the first of January 1951, using
the non-indexical part of ordinary English only. He told, of course, his wife about
this experiment. At the morning of the mentioned day Tom awakes and since it is a
holiday, he decides to have breakfast in his bed. His watch is under repair and he,
therefore, does not know the time. How shall he inform his wife about his wish? [2]
[p. 367].

The description of the experiment goes on with the Tom’s attempts of communicating
his wish to his wife, each of which fails. The reason is that

[...] effective communication by means of indexical sentences requires that the recip-
ient should know the pragmatic context of the production of the indexical sentence-
tokens. [...] To communicate the same amount of information by using non-indexical
sentences only, knowledge of the context by the recipient is not required, but in its
stead additional knowledge of some other kind may be necessary. Not in every actual
communicative situation could every indexical sentence be replaced, without loss of
information, by a non-indexical sentence [...] [2] [p. 368-369].

The experiment shows that understanding an indexical and a non-indexical sentence
which are supposed to be equivalent requires a different kind of knowledge. If Tom tells
his wife “I want to have my breakfast here”, the wife needs only to know that ‘I’ is the
person who is speaking, that ‘here’ is the position that this person occupies, and that
the tense indicates that he wants to have breakfast at the time of the utterance; instead,
if Tom tells his wife “On the first of January 1951, at 9 o’clock in the morning, Tom
a theory, or it is just a smart device that allows us to give more compact representations and to make reasoning more efficient? Our answer to this difficult question is divided into two parts. On one side, we argue that in a theory of knowledge representation we do not need the notion of pragmatic context (section 3.1). On the other side, we give some partial arguments in favour of the thesis that we cannot get by without the concept of cognitive contexts (section 3.2).

3.1 Why we don’t need pragmatic context

We’ll tackle the problem of whether we need pragmatic context in a more definite form: can we represent the same information that is represented in indexical languages using non-indexical languages without any loss of information? Can we represent the reasoning processes that involve knowledge expressed in some indexical language as reasoning processes that involve the same knowledge expressed in a (supposedly) equivalent non-indexical language? From our perspective, a positive answer to these two questions entails the conclusion that in principle we do not need pragmatic context. Indeed, if it is always possible to “transform” an indexical representation into a non-indexical one, we do not see any reason for introducing the further complication of the context mechanism into a theory of knowledge representation.

The thesis that any indexical sentence can be transformed into a context-independent sentence (a statement) without any loss of information is particularly tempting. According to Bar-Hillel, its (explicit or implicit) acceptance can explain the “strange neglect” of a logic of indexical languages in the past:

What can be the explanation of this strange neglect of such very obvious trait of ordinary languages? I venture the following hypothesis: Since a judgment with an indexical sentence as first component can always, without loss of information, be transformed into a judgment with a statement as a first component, keeping the second component intact, we might easily be tempted to drop the common phrase ‘a judgment with . . . as first component’ from both sides of this transformability statement and arrive at the result that any indexical sentence can be transformed into a statement, a patent falsity [. . .] [2] [p. 366]

The problem is: why is such a transformability thesis a patent falsity? A first argument is provided by Kaplan: an indexical sentence and its corresponding transformation into a non-indexical sentence have different logical properties, and the logical properties of the former cannot be formalized in a traditional intensional framework unless we include the notion of context. As a paradigmatic example, recall Kaplan’s discussion of the sentence “I’m here now” (section 2.1.2). But this argument cannot be used as an argument of the necessity of the notion of context in a theory of knowledge representation. Indeed, the problem is not whether we can give a logic of demonstratives without using the notion of context, but rather if we need to introduce demonstratives in our language. Kaplan does not motivates the usage of indexical languages, and takes it for granted. However, if we do not prove first that we need indexical languages, Kaplan’s argument is founded on a sort of petitio principii.
leaving a context as defined by McCarthy, even though technically different. Every space (except a distinguished space called base) is assumed to have a single primary context.

Dinsmore introduces also a notion of secondary context. It is meant to provides a kind of mapping from the content of one space to the contents of another; this mapping ought to be a consequence of the semantics of the primary contexts involved. Intuitively, a secondary context opens a channel of communication between two spaces. Syntactically, a secondary context term is written as [S], where S is a space tag. An example of secondary context is an inheritance context, namely a context that inherits all the facts of another context. For instance, let us imagine that [sp.3] models Warren’s beliefs about Alma and [sp.9] is the space containing all of Warren’s beliefs (no matter about what). Clearly, every fact contained in [sp.3] is inherited by [sp.9]

\[
\text{sp.9} \mid \text{[sp.3]}
\]

Thus if we have:

\[
\text{sp.3} \mid \text{Alma likes frogs}
\]

then this fact is inherited by \text{sp.9}:

\[
\text{sp.9} \mid \text{Alma likes frog}
\]

In the framework of mental spaces (or in that of partitioned representations), context is thought of on one side as an assumption that contributes to assign a meaning to the contents of spaces (primary context), on the other as the way of putting a relationship between the contents of different spaces (secondary context). It is clear that these two notions of context partially overlap with other notions we have discussed. For instance, a space can be built during a conversation, and so it can work as a communication context; or it can be built during a problem solving task, and so it functions as a problem solving context. Moreover, the idea of primary context recalls McCarthy’s idea of context as a set of assumptions, whereas the idea of secondary context is a special case of a lifting rule. Be as it may, we do not need to stress again the cognitive significance of the notion of context in Fauconnier’s or in Dinsmore’s framework.

3 Do we really need context?

At the beginning of this paper, we said that our interest was in the rôle of context in modelling reasoning. We see the task of modelling reasoning as a part of a more general theory of knowledge representation, whose object is a systematic explanation of the way a linguistic expression can represent a “piece” of an individual’s knowledge about the world. In other words, such a theory should allow us to determine the cognitive content of a linguistic expression when such an expression is used to ascribe some knowledge to an individual. From this perspective, a crucial question is: do we really need the notion of context in a theory of knowledge representation? Is context an essential concept in such
One of the main functional motivations for having multiple spaces is that they allow us to model parochial reasoning, namely reasoning happening within a single space. For instance, Dinsmore shows how to translate the standard rules of predicate calculus into rules for parochial reasoning, namely applicable only to facts belonging to the space in which reasoning happens.

Dinsmore introduces the notion of context when he discusses forms of reasoning involving different spaces. Let us consider figure 2, and suppose that space sp_3 is meant to represent (some of) Warren’s beliefs, e.g. that sp_1 is in Alma’s pocket. The expression sp_3 | frog_1 is in Alma’s pocket alone does not say that this is one of Warren’s beliefs. We must know the context of sp_3. In Dinsmore’s book, a primary context is defined as a function that maps the satisfaction of a proposition in one space onto the satisfaction of a (more complex) proposition into another space. In our example, the expression:

$$sp_1 | Warren believes that [[sp_3]]$$

is said to be the primary context of sp_3 ([[sp_3]] is said a primary space term). As a consequence, the proposition:

$$sp_3 | frog_1 is in Alma’s pocket$$

maps implicitly into:

$$sp_1 | Warren believes that (frog_1 is in alma pocket)$$

The reasoning step from sp_3 | frog_1 is in Alma’s pocket to sp_1 | Warren believes that (frog_1 is in alma pocket) is called context climbing. This is one of the most important rules of parochial reasoning (intuitively, it is quite similar to the operation of

---

Figure 1: Spaces

<table>
<thead>
<tr>
<th>sp_1</th>
<th>George Bush is President</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The Democrats control the Senate</td>
</tr>
<tr>
<td></td>
<td>Beer is good for you</td>
</tr>
</tbody>
</table>

Figure 2: Primary context
to another. It is clear that different qualifications can be reasoned about in different circumstances. For instance, the existence of the flight and having the ticket are very likely to be included, whereas the possible failure of part ATZ245124 is not. However, the situation could be reversed if the traveller were an engineer who is testing the security of a new aircraft. This simple example shows that not only we may adopt a different description language in different contexts, but also that the set of facts that we use in reasoning may change\(^5\).

The problem of generality (and to the qualification problem as well) suggests that context can be conceived of as the reification of all the implicit assumptions a representation of some domain depends upon. Since in general we are not aware of all these assumptions, McCarthy says that contexts are rich objects [24]. The idea of contexts as rich objects is presented in more details in Guha’s PhD. dissertation, written under McCarthy’s supervision:

> Contexts are objects in the domain, i.e. we can make statements “about” contexts. They are rich objects […] in that a context cannot be completely described. The contextual effects on an expression are often so rich that they cannot be captured completely in the logic. […] In other words, the context object can be thought of as the reification of the context dependencies of the sentences associated with the context. [14] [p. 6]

### 2.2.4 Contexts and (mental) spaces

Fauconnier’s book on Mental Spaces [7] and Dinsmore’s book on Partitioned representations [6] provide a differently motivated approach to context. Fauconnier’s work is an investigation on human knowledge representation and linguistic processing; it is meant to explain problems like referential opacity and transparency, specificity of reference, definite reference in discourse, the projection problem for presuppositions, the semantic processing of counterfactuals. Dinsmore’s work takes a more functional perspective, and motivates partitioned representations as a functional system of mental representation and as a principled basis for language understanding. We will present some of Dinsmore’s basic ideas.

A space represents some logically coherent situation or potential reality, in which various propositions are treated as true, objects are assumed to exist, and relations between objects are supposed to hold. Examples are: belief spaces; hope and wish spaces; fictional, dream, and pretense spaces; spaces representing specific places, times and situations; spaces representing the scope of certain existential assumptions; spaces expressing generalizations; spaces representing the implications of certain propositional assumptions, either conditional or counterfactual. The fact that a sentence is asserted in a space is graphically represented as in figure 1 and linearly as follows:

\[
\text{sp}_x \mid P
\]

where \(\text{sp}_x\) is a space tag and \(P\) is a sentence.

\(^5\)The analogy with the idea of a problem solving context should be evident in this formulation of the qualification problem.
book is on the table', omitting reference to time and precise identification of what book and what table. [...] A possible way out involves formalizing the notion of context [...] [23]

Here the problem is the tradeoff between needed generality and an excess of generality. Depending on the context, the “same” fact can be given several different representations with a different degree of generality. If, on one side, a more general representation can be applied to a larger class of circumstances, on the other side too much generality is a problem from the standpoint of implementing a reasoning system. In many contexts, some information can be left implicit that in other contexts is necessary to include. McCarthy’s paradigmatic example is the so-called above–theory. This theory contains very simple axioms on the blocks world. For instance, the axioms:

\[
on(x, z) \supset above(x, z) \tag{8}
\]

\[
above(x, y) \land above(y, z) \supset above(x, z) \tag{9}
\]

say that an object \(x\) is above an object \(y\) if either \(x\) is on \(y\) or \(x\) is above an abject \(y\) which in turn is above \(z\). Most times, these two axioms are sufficient for reasoning about the property of being above in the blocks world. However, there are cases in which they are not general enough; for instance, (9) is true only if \(above(x, y)\) and \(above(y, z)\) are true at the same time. One way of making explicit this qualification is to add a parameter for the time to the predicates \(on\) and \(above\):

\[
on(x, z, t) \supset above(x, z, t) \tag{10}
\]

\[
above(x, y, t) \land above(y, z, t) \supset above(x, z, t) \tag{11}
\]

A formalization of context should allow us to use the “right” axioms in the “right” context, i.e. the less general axioms (8)–(9) if the context allows us to disregard the time, and the more general axioms (10)–(11) in a context where time is relevant.4

Notice that the same difficulty arises not only with the choice of primitives, but also with the set of facts that are explicitly reasoned about in a given circumstance. This can be easily seen if we consider the qualification problem as defined in [12], namely as the problem of qualifying the truth of any common sense general axiom. Consider a set of common sense axioms stating the preconditions and the effects of flying from a city

---

4McCarthy and his group proposed a formalization of context that is meant to deal with all these problems. The reader interested in this approach may refer to [24, 14] for the general intuitions and many motivating examples, and to [4, 3] for a semantics of a propositional and quantificational logic of context based on an extension of a Kripke semantics. The basic intuitions are: context must be reified as reasoning objects (so that, for instance, we can write formulae like \(ist(c, w)\), whose intended meaning is that the formula (named) \(w\) is true in the context (named) \(c\); second, that there exist some specific patterns of contextual reasoning, e.g. \(entering\) and \(leaving\) a context; third, that axioms can be \(lifted\) from a context to another modulo suitable modifications, which are stated in special lifting rules. Buvac and Mason’s semantics of McCarthy’s logic of context views \(ist\) as a new modality. As far as we know, McCarthy is still uncommitted on this approach to the semantics of context.
A slightly different version of the idea of problem solving context can be found also in the literature of AI. A good example is given by McCarthy’s discussion of the cannibals and missionaries puzzle [22]: three missionaries and three cannibals must cross a river; a boat holding two people is available; if cannibals outnumber missionaries on either side of the river, they eat the missionaries; how is the river to be crossed in such a way that missionaries are not eaten? McCarthy makes two interesting observations:

• on one side, most knowledge that is needed to understand and solve the problem is not explicit in the statement of the puzzle. For instance, nothing is said about how boats are used to cross rivers, and so this information must in some sense be “loaded” from memory (if the solver is a computer program, from its common sense knowledge base);

• on the other side, a lot of potentially relevant information is not considered by the average solver. For instance, it might be the case that there is a bridge one mile upstream. However, this possibility is systematically overlooked, since people tend to focus only on the objects that are explicitly mentioned in the statement of a task.

Again, we do not want to propose solutions to any of the issues related to the notion of problem solving context. We only stress the cognitive flavour of this notion. A problem solving context is clearly relative to a particular problem solving task, which is performed at some time and place by a given agent. However, the contents of such a context depend only on the representation of the problem and not on the fact that the problem is reasoned about at some time and place (not in the pragmatic sense).

2.2.3 The problem of generality

Another point of contact between context and AI is the so-called problem of generality (identified by McCarthy in his paper on Generality in Artificial Intelligence [23]) and its dual problem, the qualification problem\(^3\). McCarthy characterizes the problem of generality and its relationship with the problem of context as follows:

Whenever we write an axiom, a critic can say that the axiom is true only in a certain context. With a little ingenuity the critic can usually devise a more general context in which the precise form of the axiom doesn’t hold. Looking at human reasoning as reflected in language emphasizes this point. Consider axiomatizing on so as to draw appropriate consequences from the information expressed in the sentence, ‘The book is on the table’. The critic may propose to haggle about the precise meaning of on, inventing difficulties about what can be between the book and the table, or about how much gravity there has to be in a spacecraft in order to use the word on and whether centrifugal force counts. Thus we encounter Socratic puzzles over what the concept mean in complete generality and encounter examples that never arise in life. There simply isn’t a most general context.

Conversely, if we axiomatize at a fairly high level of generality, the axioms are often longer than is convenient in special situations. Thus humans find it useful to say, ‘The

\(^3\)Also the qualification problem was originally identified by McCarthy in [22]. However, its duality with the problem of generality was highlighted in [12].
then added new assumptions as the conversation goes on? Or is there something as an initial context which is then suitably modified? What is to be included in this initial context? One might think that perhaps the initial context must include our beliefs about the speaker's beliefs, cultural background, linguistic habits, intentions, and so on. On one side, this idea undermines the concept of a shared communication context; on the other hand, it seems to require the construction of a very large and complex context, which is psychologically quite implausible.

We do not want to suggest any solution to these problems. We only stress the cognitive aspects of the notion of context in a theory of communication. It is clear that such a context comes into existence at some point in time and place and is relative to some agent (the speaker or the hearer, depending on the perspective we adopt), but most assumptions have nothing to do with such a location (in Lewis' sense). Of course, some of the premises will concern the time and place where the communication process happens (e.g. if the speaker says "it's raining", some of the hearer's assumptions must concern the where and the when); however the interpretation will not depend on the location itself, but rather on the fact that some features of this location are part of the assumptions that speaker and hearer use in interpreting indexical sentences. We'll come back to this point in section 3.1.

2.2.2 Context in problem solving

Another typical area in which the notion of context plays an important rôle is problem solving. Indeed, there is a lot of evidence that people tend to solve problems using only a portion of what they globally know. The set of facts that are "active" in a problem solving task is very often referred to as a problem solving context.

What is included in a problem solving context depends on many factors, for example the way the task is stated (in particular for verbally posed puzzles), the objects that the solver has at his disposal in the environment, his/her background knowledge, the solver's familiarity with the kind of task, the available time, and so on. All this is very concisely expressed in the following definition of context by Kokinov [17]:

"Context is the set of all entities that influence human (or system's) cognitive behaviour on a particular occasion. [...] There are many things in the universe that do not influence human behaviour in a particular moment and only very few that do influence it. Moreover, different people will be influenced by different elements of the same environment. So, all the entities in the environment which do influence human behaviour are internally represented and it is the representation which actually influence the behaviour [...] That is why context is considered as a 'state of the mind' of the cognitive system. [17] [p. 200]

A very interesting aspect of this definition is the emphasis on the internal representation of the entities that belong to a problem solving context (and, more in general, to the context of any cognitive behaviour). It is easy to experiment that all often people have the "right" object to solve a problem at their disposal, but they cannot see how that object can be used in that context. Or that the presence of some useless objects can sidetrack people and lead them to search for wrong solutions."
ture, in which points of productions (contexts) are taken into account. We said that, in this conceptual framework, contexts are thought of as part of the state of the world with respect to which a sentence is given a content and a truth value.

However, in other research fields, context is not conceived of as part of the state of the world, but rather as part of the cognitive state of an agent. Context dependence is not dependence on the point of production, but rather dependence on a set of implicit assumptions that affect cognitive processes, like communication, problem solving, common sense reasoning, and so on. The work on this second notion of context, which we name cognitive context, is much less homogeneous that the work on pragmatic context, and an exhaustive presentation of it is out of the scope of this paper. In the remaining of this section, we try to give an idea of the cognitive phenomena that context is supposed to account for.

2.2.1 Context in communication

Imagine that S invites H for dinner. H had planned to work all night, but it’s been a long time since he saw S and so he accepts the invitation. After dinner, S asks H: “Would you like a cup of coffee?” H replies: “Coffee would keep me awake”. H’s intention was to accept his host’s offer, because he actually wants to stay awake. However, S assumes that his guests does not want to stay awake, and therefore interprets the answer as a refusal. This simple example of communication failure shows that the same sentence can be given very different interpretations depending on the set of premises used in its interpretation. In many theories of communication, this set of premises is called a context. This idea is very clearly stated by Sperber and Wilson give in their book on Relevance. Communication and Cognition [29]:

The set of premises used in interpreting an utterance […] constitutes what is generally known as the context. A context is a psychological construct, a subset of the hearer’s assumptions about the world. It is these assumptions, of course, rather than the actual state of the world, that affect the interpretation of an utterance. A context in this sense is not limited to the information about the immediate physical environment or the immediately preceding utterances: expectations about the future, scientific hypotheses or religious beliefs, anecdotal memories, general cultural assumptions, beliefs about the mental state of the speaker, may all play a rôle in interpretation. [29] [p. 15-16].

Context is viewed as the set of assumptions that a hearer uses in interpreting a speaker’s sentence. This notion of context raises very interesting problems. A typical one is: is the context shared between speaker and hearer? On one side, assuming that it is shared would allow us to give a simple explanation of how people communicate; however, it would be very hard to explain cases of communication failure as the one reported above. On the other side, holding that it is not shared leads to the opposite problem: how is it possible to guarantee that people really understand to each other? Indeed, unless we can prove that every assumption is shared, there is the possibility that people believe they understand to each other, but actually they don’t. Another problem is: how is the communication context generated? Does a conversation start in a sort of empty context at which are
2.1.3 Context and index

Kaplan shows that, in order to give a logic of demonstratives, we need both the notions of context (thought of as a quadruple of parameters) and index (thought of as a pair time-world). The relation between these two technical notion is discussed in more detail in Lewis’ paper on “Index, context, and content” [19].

Similarly to Kaplan, Lewis defines a context as a location (i.e. a time, a place, and a possible world). However, he stresses the fact that a context has countless features, determined by the character of the location. Examples of features are: the speaker (the one who is speaking at that time, at that place, at that world); the audience, the standard of precision, the salience relations, the presuppositions, . . . , which are determined by such things as the previous course of the conversation that is still going on at the context, the states of mind of the participants, and the conspicuous aspects of their surroundings. The truth of a sentence in a context depends on all these features. It is sufficient to think at the huge number of variables that are introduced by the states of mind of the participants. Any “package” of features is called an index.

The problem is: are we able to build indices rich enough to include all relevant features of a context? Lewis thinks that we are not, and this is the reason why we need both context and index: “Indices are no substitute for contexts because contexts are rich in features and indices are poor” [p. 88]. As a consequence, an index is a package of features that one chooses to consider according to some criterium; but in general, this package does not include all relevant features.

So the problem becomes: what are the features that ought to be packed into an index? Lewis proposes a possible criterium: we have to include only features that can be shifted. A feature is shiftable if the truth of a sentence in a context “depends on the truth of some related sentences when some feature of the original context is shifted”. Simple examples are the following. “Somewhere the sun is shining” is true here if and only if “The sun is shining” is true somewhere (i.e. if and only if the sentence “The sun is shining” is true in a context in which the place is changed with respect to the original context); “Aunts must be women” is true at our world if and only if “Aunts are women” is true at all worlds; and so on. Lewis’ proposal is summarized as follows:

Since we are unlikely to think of all the features of context on which truth sometimes depends, and hence unlikely to construct adequately rich indices, we cannot get by without context-dependence as well as index-dependence. Since indices but not contexts can be shifted one feature at a time, we cannot get by without index-dependence as well as context-dependence. An assignment of semantic values must give us the relation: sentence s is true at context c at index i, where i need not be the index that gives the features of context c. [19] [p. 79]

2.2 Cognitive context

The notion of context presented in section 2.1 is based on the intuition that the point at which a sentence is produced is essential in order to determine the content of indexical sentences. The idea is that a sentence is evaluated with respect to a more complex struc-
Kaplan’s notion of context accounts for the demonstratives ‘I’, ‘here’, ‘now’, ‘actually’, ‘yesterday’, and some tense operators (future, past, one day ago). To give an idea of the general approach, terms and sentences are given an interpretation in a context \( c \) under an assignment \( f \), with respect to a world \( w \) and a time \( t \) (in a structure \( U \)). The denotation of a term is written as \( | \alpha |_{cftw} \), whereas the truth of a sentence is written as \( |=_{cftw} \phi \). For instance:

\[
| I |_{cftw} = \alpha_A \quad \text{The demonstrative } I \text{ in a context } c \text{ is interpreted on the agent of } c \\
| \text{here} |_{cftw} = \alpha_p \quad \text{The demonstrative } \text{here} \text{ in a context } c \text{ is interpreted on the position of } c \\
|=_{cftw} N \phi \iff |=_{cftw} \phi \quad \text{The sentence ‘It is Now the case that } \phi \text{’ is true in a context } c \text{ iff it is true at the time of } c \\
|=_{cftw} A \phi \iff |=_{cftw} \phi \quad \text{The sentence ‘It is Actually the case that } \phi \text{’ is true in a context } c \text{ iff it is true at the world of } c
\]

A formula is valid in LD if and only if it is true in every context of every structure (written \( |= \phi \)). An interesting property of LD validity is that it does not entail logical truth. This means that necessitation (if \( |= \phi \) then \( |= \Box \phi \)) fails:

From the perspective of LD, validity is truth in every possible context. For traditional logic, validity is truth in every possible circumstance. Each possible context determines a possible circumstance, but is not the case that each possible circumstance is part of a possible context. In particular, the fact that each possible context has an agent implies that any possible circumstance in which no individuals exist will not form a part of any possible context. Within LD, a possible context is represented by \( \langle U, c \rangle \) and a possible circumstance by \( \langle U, t, w \rangle \). To any \( \langle U, c \rangle \), there corresponds \( \langle U, c_T, c_W \rangle \). But it is not the case that to every \( \langle U, t, w \rangle \) there exists a context \( c \) of \( U \) such that \( t = c_T \) and \( w = c_W \). [15] [p. 93]

A last remark on LD. In his paper, Kaplan hints to the possibility of a cognitive interpretation of LD. Indeed, he suggests that the difference between content and character is not only a technical, but also cognitive:

Although a lack of knowledge about the context (or perhaps about the structure) may cause one to mistake the Content of a given utterance, the Character of each well-formed expression is determined by rules of the language [... which are presumably known to all competent speakers. [15] [p.92–93]

Character is – so to say – linguistic knowledge, whereas the determination of the content partly depends upon what an individual knows (or does not know) about a context (or a structure). Relative to Bar-Hillel’s thought experiment, this means that, even though Tom and his wife might lack the information they would need in order to transform the indexical into a non-indexical communication, they are very likely not to lack information on the rules such a transformation would follow.
then (5) is true iff in the world $w$ it is true that $x$ is at place $p$ at time $t$. For instance, (5) might be equivalent to the statement:

**David Kaplan is in Los Angeles on April 21, 1973** (6)

But this “transformation” misses an essential point in our usage of demonstratives in ordinary language. Indeed, (5) is intrinsically true (whatever the context, it cannot be falsely uttered!), whereas the same cannot be said for (6), whose truth is contingent. This fact has a dramatic impact on the treatment of demonstratives we sketched above. Indeed, if (5) is always true (i.e. is true at every index), then it is a logical truth. But this seems quite counterintuitive, since it is not necessarily true that, for instance, David Kaplan is in Los Angeles on April 21, 1973. But the situation is even worse. Indeed, one might try to restrict the class of indices to the proper ones, namely to those such that in the world $w$, $x$ is located at $p$ at the time $t$. This would account for the difference between (5) and (6), but the side effect would be that

$$\square (I'm \ here \ now)$$

would become true as well, where it is clear that it is not necessary that $x$ is at $p$ in $t$.

Kaplan identifies the source of the difficulties in the confusion between two kinds of meaning. Ramifying Frege’s distinction between sense and denotation [9], Kaplan proposes two varieties of sense:

- **content**: what a sentence would say if it were uttered (evaluated) in a context. Roughly speaking, if David Kaplan on April 21, 1973 at Los Angeles utters (5), then what is said (the content) is (6);

- **character**: the aspect of the meaning of an expression which determines what its content will be in each context. For instance, the character of ‘$I$’ is “*the function which assigns to each context that content which is represented by the constant function from possible worlds to the agent of the context*” [15] [p. 84].

Notice that content is always taken with respect to a given context of use. Relative to a context, demonstratives have a stable content. For example, the content of ‘now’ in a context is the time associated to that context. As to the character, any expression which contains no essential occurrence of demonstratives has a stable character, namely it has the same content in every context. For example, (6) has a stable character.

Using Bar-Hillel’s terminology, a statement is a sentence whose content is the same in every possible context (*i.e.* it has a stable character); an indexical sentence is a sentence whose content depends on the context. (1) has a stable character, whereas (3) doesn’t, the reason being that the content of ‘$I$’ is different in two different contexts.

Technically, Kaplan defines a context $c$ as a quadruple $\langle c_A, c_T, c_P, c_W \rangle$, where $c_A$ is called the agent of $c$, $c_T$ is the time, $c_P$ is the position, and $c_W$ the world of $c$. When is a sentence like (3) true in a context $c$? The idea is that it is true if and only if at time $c_T$, in the world $c_W$, it is true that the individual $c_A$ is hungry. It should be clear why (6) is true in some contexts and false in others, whereas (5) is true in every context.
side the pragmatic reference (and thus the truth value) of indexical sentences like (2) and (3) may be different in different contexts, on the other side the pragmatic reference (and truth value) of a statement like (1) is constant in any context. This motivates Bar-Hillel’s introduction of a derived notion of reference, called semantical reference:

\[ RS(a,b) =_{def} \forall c \forall d \text{ (RP}(a,b,c) \text{ iff } \text{RP}(a,b,d)) \]

(4)

Semantical reference is obtained from pragmatic reference by abstracting from the context. In other words, a statement (e.g. (1)) has a semantical reference (e.g. the fact that ice floats on water) because its pragmatic reference is constant in any context, whereas the pragmatic reference of an indexical sentence may vary in different context (e.g. (2) pragmatically refers to the fact that at the time and place where of utterance it is raining and (3) to the fact that at the time of utterance the speaker is hungry).

It should be clear from the above observations that an indexical sentence lacks a reference if the pragmatic context of production is not taken into account. Therefore, since ordinary language is an indexical language, Bar-Hillel concludes that context is a key notion in a logic of ordinary language. The first (and maybe most well-known) attempt of giving such a logic is Kaplan’s Logic of Demonstratives (LD).

2.1.2 A logic of demonstratives

LD was presented in [15]. Kaplan largely agrees with Bar-Hillel’s idea that indexicals (Kaplan prefers the term demonstratives) are an essential component of ordinary language, and that an adequate treatment requires some notion of context. In addition, he shows that demonstratives have their own logic which is not reducible to other intensional notions (e.g. to the notion of index).

Kaplan’s goal is to show that not only the extension, but also the intension of demonstratives are determined by the context of their use. The intension of “eternal” terms (like “The Queen of England in 1973”) has generally been taken to be represented by a function which assigns to each possible world the Queen of England in 1973 of that world. If instead of possible worlds we consider more complex indices (e.g. including a time, a place, a speaker, a possible world, …), then one might think that by analogy the intension e.g. of ‘I’ can be given as a function from speakers to individuals (in fact, the identity function); the same for other demonstratives like ‘here’ and ‘now’. This approach has the advantage that the principle of modal generalization (or necessitation) is validated. Indeed, a sentence \( \phi \) is logically true iff it is true at every index (in every structure), and \( \Box \phi \) is true at a given index (in a given structure) just in case \( \phi \) was true at every index (in that structure). From this, necessitation follows: if \( \models \phi \), then \( \models \Box \phi \).

But this – says Kaplan – is technically wrong and, more importantly, conceptually misguided. Demonstratives depend on context in a different way. Consider the paradigmatic sentence

\[ I'm \text{ here now} \]

(5)

On one side, it is clear that – for many choices of index – it will be false. If the index is for instance a quadruple \( \langle w, x, p, t \rangle \) (where \( w \) is a world, \( x \) and \( p \) a place and \( t \) a time),
We must also know when and where the sentence is produced. If it were uttered on June 14, 1996 at 11.30 am at Trento, it would be false; but had it been uttered on June 22, 1996 at Trento, it would have been true. Typically, time and place of production are included among the parameters which identify a box, and their value is used to determine what is said when one utters “It is raining” in that box\(^1\). In more philosophical terms, such a box is called a **pragmatic context**.

The prototypical class of expressions whose semantic value depends on pragmatic context are the so-called **indexical expressions**, like “I”, “here”, “now”, and so on. However, it is clear that other elements of ordinary language depend on pragmatic context, in particular the tense (as proved by “It is raining”). In the following three sections, we make a short summary of the main ideas that have been proposed in the past about pragmatic context.

### 2.1.1 Sentences in their context

The first who put the problem of context as we presented it at the beginning of this section was Bar-Hillel in his paper on **Indexical expressions**\(^2\): “Even very superficial investigation into the linguistic habits of users of ordinary language will reveal that there are strong variations in the degree of dependence of the reference of linguistic expressions on the pragmatic context of their production” [2]. For example, if on one side the truth value of the sentence

\[
\text{Ice floats on water} \quad (1)
\]

has a very low degree of dependence on context (so low that in most practical situations it can be disregarded), on the other side the truth of the sentences

\[
\text{It’s raining} \quad (2)
\]

\[
\text{I’m hungry} \quad (3)
\]

effectively depends on the pragmatic context of their production. Bar-Hillel introduces the following terminology. He uses the word *sentence* – with respect to ordinary language – as grammarians use it. An ordered pair \(\langle \text{sentence-token, context} \rangle\) is called a *judgment*. The first component of a judgment is called a *declarative sentence*. A declarative sentence that, paired with any context whatsoever, forms judgments which refer always to the same proposition will be called a *statement*, otherwise an **indexical sentence**.

Bar-Hillel argues that it is not possible to assign a reference to indexical sentences, but only to a **sentence-token-in-a-certain-context**, namely sentence-tokens that occur as the first element of a judgment. For this reason, we have before us an essentially triadic relation \(RP(a,b,c)\): “(the sentence) \(a\) refers-pragmatically-to (the proposition) \(b\) in (the pragmatical context which includes also a reference to a language) \(c\)” [p. 364]. If on one

\(^1\)Strictly speaking, it is not necessary that a sentence is actually uttered. As David Kaplan points out, “it is important to distinguish an *utterance* from a *sentence-in-a-context*. The former notion is from the theory of speech acts, the latter from semantics” [15] [p. 91].

\(^2\)Actually, the same problem (but in a different form) was noticed also by Frege in some of his last papers on language and thought, for instance in [10]
notions of context *pragmatic context* and *cognitive context* respectively. In the literature on context, thought of in two different ways. In the second part of the paper (section 4) we show that, from the standpoint of a theory of knowledge representation, the notion of pragmatic context can be eliminated in favour of the notion of cognitive context, and that the latter is an essential building block in any theory of knowledge representation. This will be the starting point for some considerations on the place of context in a theory of knowledge representation (section 4). Finally in section 5 we introduce two basic principles of contextual reasoning, namely *locality* (section 5.1) and *compatibility* (section 5.2).

2 Context(s)

It is a quite common intuition that some sentences are true (polite, effective, appropriate,. . . ) in a context, and false (impolite, not effective, inappropriate) in others, that some conclusions hold only in some contexts, that a behaviour is good only in some contexts, and so on. For instance, “France is exagonal” (or “Italy is boot-shaped”) is true in contexts whose standard of precision is very low, false in the context of Euclidean geometry. In the context of the Computer Science Department at the University of Trento, the sentence “All professors are Italian” is true; in a larger context (e.g. in the context of all Italian Universities), it’s false. To shout is allowed in the context of a football game, but not in the context of a religious cerimony. All these examples seem to suggest that a context can metaphorically be thought of as a sort of “box”. Each box has its own laws and draws a sort of boundary between what is in and what is out.

A closer look to the literature on context will show that this metaphor can be given two very different interpretations. According to the first, a “box” is viewed as part of the structure of the world; according to the second, a “box” is viewed as part of the structure of an individual’s representation of the world. We call the two notions of context that result from these two different interpretations of the metaphor *pragmatic context* and *cognitive context* respectively. The next two subsections are devoted to present in some detail the intuitions underlying these different concepts of context.

2.1 Pragmatic context

The first way of looking at the metaphor of the box is to see a box as a piece of the world. The phrase “piece of the world” does not mean just a portion of the physical (actual) world. It can be viewed as a physical position at some instant of time at some (possible) world; usually, an agent is associated to every box. Each box is uniquely identified by some (more or less complex) collection of parameters. Depending on the parameters we choose, the box will have different properties. For instance, if we choose a point in time and space, the box “shrinks” to a point. If we choose a time interval and a larger spatial area, the box becomes bigger; this is needed if we want to evaluate sentences with respect to e.g. the Computer Science Department at the University of Trento.

The idea is that the value of the parameters that identify a box provides information that a linguistic expression may lack. For example, “It is raining” cannot be assigned a definite truth value if all we know is the meaning of the words “it”, “is” and “raining”.  

2
Introduction to Contextual Reasoning
An Artificial Intelligence Perspective

Fausto Giunchiglia\footnote{D.I.S.A. - University of Trento (Italy)} \footnote{Istituto per la Ricerca Scientifica e Tecnologica - Trento (Italy)} \hfill Paolo Bouquet\footnote{E-mail: fausto@irst.itc.it \hfill bouquet@cs.unitn.it \hfill URL: http://afrodite.itc.it:1024/~fausto \hfill http://www.cs.unitn.it/~bouquet}

1 Introduction

The notion of context is called to account for a multifarious variety of phenomena. Since Frege’s proposal of a principle of contextuality \cite{Frege1884}, context has had its place in philosophy of language (\emph{e.g.} \cite{Searle1989,McGinn1990,Searle1995}). In Artificial Intelligence (AI), McCarthy was the first to argue that formalizing context was a necessary step toward the designing of more general computer programs (\cite{McCarthy1960}, but see also \cite{Sowa1976,Salmon1983,Alchourron1985,Salmon1990}). Other work comes from cognitive science (\emph{e.g.} \cite{Dretske1981,Barrett1985,Fodor1983}), where context is viewed as a way of structuring knowledge and its usage in problem solving tasks. The motivations and the approaches to the problem of context are very different, and one might even wonder whether there is something as the problem of context, or rather a multiplicity of different problems very loosely related by the word “context”. We will argue that most of the proposed notions of context can be looked at from a unifying perspective. However, before we are ready to defend this claim, a lot of preliminary work is needed.

As a starting point, we will accept a very general notion of context as a collection of “things” (parameters, assumptions, presuppositions, …) a representation depends upon. The fact that a representation depends upon these “things” is called context dependence. Since our main interest is in building formal systems for modelling reasoning, we will consider context dependence only as far as it plays a rôle in reasoning. It has been argued that context has a rôle in modelling other cognitive processes (\emph{e.g.} perception, language understanding, pattern recognition; see for instance \cite{Hamblin1985,Brown1984}); we will not explicitly discuss these issues, even though we believe that they are somewhat related with our topic.

This paper mostly addresses the following (rather philosophical) questions: what is context? What does it mean for a representation to depend upon context? Is context an essential aspect of a theory of knowledge representation? Can we get by without context?

The structure of this paper is the following. In the first part of the paper (section 2), we argue that the notions of context proposed in the literature relies on two very different intuitions. According to the first, a context is thought of as part of the structure of the world (section 2.1); according to the second, a context is thought of as part of the structure of an individual’s representation of the world (section 2.2). We call the resulting
INTRODUCTION TO CONTEXTUAL REASONING:
AN ARTIFICIAL INTELLIGENCE PERSPECTIVE

Fausto Giunchiglia and Paolo Bouquet

May 1997

Technical Report # 9705-19

© Istituto Trentino di Cultura, 1997

LIMITED DISTRIBUTION NOTICE
This report has been submitted for publication outside of ITC and will probably be copyrighted if accepted for publication. It has been issued as a Technical Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of ITC prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article.