Context Mereology

Patrick Hayes
Florida IHMC, 40 South Alcaniz St.
Pensacola, Florida 32502, USA
phayes@ihmc.us

Abstract

We show that under a small number of assumptions, it is possible to interpret truth in a context as a quantification over truth in 'atomic' or pointlike contexts, which are transparent to all the connectives. We discuss the necessary assumptions and suggest conditions under which they are intuitively reasonable.

Introduction

This note is inspired by the context logic originally introduced by J. McCarthy (McCarthy 1993) and R. V. Guha, (Guha 1991) and subsequently developed by others. It is intended to address the issue discussed in (Makarios, Heuer & Fikes 2006), viz. the transparency of contextual assertions to the propositional connectives.

The central construction of context logic is ist(c P). Here c denotes a context, which is supposed to be a 'bearer of truth' in some broad sense, and P denotes a proposition, which is some entity that can be said to be true or false in a context. A proposition may fail to have a truth-value in a context, so that ist(c P) and ist(c ~P) might both be false.

Examples of contexts include *time-intervals*, where ist means that the proposition holds during the time-interval; *believers*, in which ist means that the proposition is believed; *information sources* such as databases or documents, in which ist means that the information source is a provenance for the proposition; and modal-alternative, imaginary or counterfactual *worlds* or *situations*, in which ist asserts that a proposition is true in the world or situation. Most of the intuitions underlying the development here arise from the first application, where contexts are thought of as time-intervals, but the formal results apply to any kind of context which satisfies the axioms.

Ist and its dual

It seems to be widely assumed that ist distributes over conjunctions in its second argument: that is, that ist(c P&Q) implies ist(c P) & ist(c Q). It is easy to see intuitively that this corresponds to the idea that the proposition P is true throughout the context c. However, not all attributions of truth to a situation distribute over conjunction. For example, a recent day of hard traveling might be summarized by saying "On April 30 I was in seven states" meaning, of course, that I was at various times of the day in one of seven states, not that I was in all seven states all of the day. In this case, the relevant notion of 'truth in' seems to *not* distribute over conjunction: for

(true-in 300405 (I am in Texas)) & (true-in 300405 (I am in Mississippi))

can be true, when clearly

(true-in 300405 ((I am in Texas) & (I am in Mississippi)))

must be false. It is intuitively clear what is meant, however: this is a different notion of 'true in' from the ist sense, and in fact it is precisely the classical dual, definable as

wist(c P) =df ~ist(c ~P)

Clearly, wist distributes over disjunction but not (in general) over conjunction; and true-in, as used above, is wist rather than ist: it means intuitively "at some time during" rather than "throughout".

The relationship between ist and wist is exactly analogous to the usual duality between the universal and existential quantifiers, and between the strong and weak modal operators. In fact, the two operators can be viewed as indexed modalities, with the particular context providing the index, and the standard transliteration of modal propositional logic into first-order logic then maps them into patterns of quantification. If we think of a context as a time-interval, and a time-interval as a set of points, then ist(c P) translates into *for all p in c, P(p)*, and wist(c P) translates into *there exists a p in c with P(p)*. This kind of translation provides for a useful and natural reduction of 'ist' language to a simpler subcase where the basic relationship between a proposition and a 'context-point' is transparent to all the connectives and hence merges ist and wist into a single relationship, which we could paraphrase as 'P is true *at* c'. The various cases of truth-in-a-context being opaque to the connectives, such as the example given above of wist(c P&Q) not being identical in meaning to wist(c P) & wist(c Q) , can then all be explained by the patterns of quantification; in this case, the fact that exists(x)(P(x) & Q(x) is not logically equivalent to (exists(x) P(x)) & (exists(x) Q(x)). Propositional context logic then reduces to classical quantifier logic plus a very simple, completely transparent, notion of true-at-a-context-point.

The purpose of this note is to identify some general conditions under which this reduction of contexts to sets of context-points can be done. We will show that a small number of axioms, only one of which is controversial, suffice.

Axioms

Following this analogy with modal logic, and to reduce notational clutter in what follows, we will adopt modality-style notation and write $[c]P$ for $ist(c\ P)$ and $<c>P$ for $wist(c\ P)$; the brackets are intended to suggest the box-diamond notation commonly used to indicate modalities.

Definition 1. $<c>P =df \sim [c](\sim P)$

The first axiom is the basic assumption about ist, that it distributes over conjunction.

Axiom 1. $[c](P \& Q)$ iff ($[c]P$ and $[c]Q$)

The second is a kind of internal coherence principle for contexts, that an overt contradiction cannot be true throughout a context:

Axiom 2. $\sim [c](P \& \sim P)$

Lemma 1. $[c]P$ implies $<c>P$

Proof. Assume $[c]P$, and suppose $\sim <c>P$. By def1, $[c]\sim P$. By axiom 1, $[c](P \& \sim P)$ contradicting axiom 2. **QED**

Parts of Contexts

Now we introduce a relationship of parthood on contexts. Intuitively, a part of a context is some piece or aspect of it which is also considered to be a context, and can be distinguished by there being a proposition which has a different truth-value in that part than in some other part. In the case of a time-interval, parthood seems to correspond naturally to being a subinterval. We will write $c<d$ for the parthood relationship, which we will assume is transitive, asymmetric and reflexive, i.e. a partial order:

Axiom 3. $<$ is a partial order

To be true throughout a context, then, is to be true throughout all of its parts, giving another axiom:

Axiom 4. ($[c]P$ and $d<c$) implies $[d]P$

This simply identifies $[c]$ as the 'strong' modality of the dual pair. It is easy to see that this has an alternative formulation:

Axiom 4a. $[c]P$ iff (for all $d<c$, $[d]P$)

since 'if' is trivial, because $c<c$.

We need a stronger way to relate truth in a context with parthood of a context. After some exploration of possibilities, the following seems to be the most reasonable assumption which is sufficient to establish the results. We call this the *truth locating axiom*. It is discussed at greater length later.

Axiom 5. (TLA). $<c>P$ iff there is a $d<c$ with $[d]P$

This could be stated as an implication, since "if" is a consequence of earlier axioms.

Finally, we assume that parthood and truth have an extensional relationship:

Axiom 6. (Separation) $c<d$ or there is a proposition P with $[d]P$ and $\sim [c]P$

That is, if c is not a part of d then there must be some proposition which characterizes the part of d 'outside' c. A consequence of this is that if, for every proposition P, $[c]P$ iff $[d]P$, then $c=d$. That is, the set of contexts satisfies an extensionality condition with respect to the set of propositions: two contexts differ only if they somewhere assign a different truth-value to some proposition, which we can call a *separating* proposition.

Axioms 1 through 6 are all that we require to show that any context can be viewed as a set of context-points. First we will establish some useful consequences.

The following equivalences show that axiom 4a can be re-cast, assuming the TLA axiom, into a stronger form:

$[c]P$ iff

$\sim <c>\sim P$ iff

$\sim$(exists $d<c$ with $[d]\sim P$) iff

for all $d<c$, $\sim [d]\sim P$ iff

for all $d<c$, $<d>P$

so:

Lemma 2: $[c]P$ iff (for all $d<c$, $[d]P$) iff (for all $d<c$, $<d>P$)

QED

The most important consequence of the TLA is that every context must somewhere make a *commitment* to any proposition:

Lemma 3: For any context c and proposition P there is a $d<c$ with $[d]P$ or $[d]\sim P$

Proof. By axiom 2, $\sim [c](P \& \sim P)$ so by axiom 1, $\sim [c]P$ or $\sim [c]\sim P$, i.e. $<c>\sim P$ or $<c>P$, so by axiom 5 the result follows. **QED**

The axioms imply that parthood of contexts satisfies a basic axiom of mereology, the supplementation axiom.

Definition 2. cOd =df there is an e with $e<c$ and $e<d$

Lemma 4: $c<d$ or there is an e with $e<c$ and $\sim(eOd)$

Proof. Suppose ~(c<d), then by axiom 6 there is a P with [d]P and ~[c]P, i.e. <c>~P; so by axiom 5, there is an e<c with [e]~P. If eOd then there is an f<e so [f]P by axiom 3, and f<d so [f]~P, violating axiom 2. **QED**

Points and Nests

Context-points are constructed from *nests*, i.e. descending chains of subcontexts. This construction is an application of the mathematical ultrafilter/ideal technique for constructing points from partial order structures, widely used in topology and underlying the Stone representation theorem for Boolean algebras.

The intuitive picture behind the following definitions is that such a nest either determines the truth-values of all propositions, or else can be extended by adding a subcontext which determines a new proposition: so by taking the limit, we can view any sufficiently deep nest as a 'point' which fully determines the truth-value of every proposition, and so is transparent to negation. We then say that a point is *inside* a context when the nest contains the context (since, intuitively, any nest determines a point which is inside all the contexts in the nest, as figure 1 illustrates), and identify a context with the set of all such points inside it.

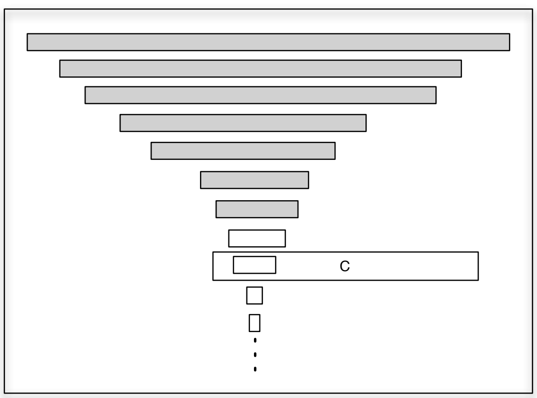

Figure 1. A nest defining a point inside the context C

Definition 3. *nest* =df a set of contexts totally ordered by <

Definition 4. c is the *bound* of n =df c in n and for all d in n, d<c

Definition 5. P is *true at* n =df (n)P =df for some c in n, [c]P

Definition 6. x is a *point* =df x is a nest and for every proposition P, either (x)P or (x)~P

Definition 7. x is *inside* c =df d<c for some d in x

We use the neutral notation (n)P rather than [n]P to emphasize the distinction between points and contexts, and also because, as shown below, the strong and weak forms of ist coincide for points, i.e. ~(n)P iff (n)~P. Notice that in definition 7, c is not required to be in the nest x (as figure 1 also illustrates).

The main result is a consequence of the axiom 5 and the axiom of choice:

Lemma 5. Every bounded nest is a subset of a point

Proof. Say that P is *determined by* c when [c]P or [c]~P, and let n be a bounded nest with bound b. By the axiom of choice, we can assume that the set of propositions is well-ordered. If n is not a point then let P be the first proposition in the well-ordering which is not determined by b, then by lemma 3 there is a d<b with [d]P or [d]~P, so (n union {d}) is a bounded nest which determines a superset of the propositions determined by n. By induction, the limit of this construction is a nest which determines every proposition in P. **QED**.

Since the singleton set of any context is a bounded nest, Lemma 5 can be re-stated as: **any context has a point inside it**. (The boundedness condition is required; the lemma does not hold for arbitrary nests, as shown by the well-known example of the interval-endpoint when a light switches off (Allen&Hayes 1989).)

Points – in fact, nests generally – are transparent to conjunction, like contexts:

Lemma 6. (n)(P & Q) iff (n)P and (n)Q

Proof. If (n)(P & Q) then for some c in n, [c](P & Q), so [c]P and [c]Q, so (n)P and (n)Q. If (n)P and (n)Q, then for c, d in n, [c]P and [d]Q, and either d<c or c<d. Assume c<d; then [c]P & [c]Q by axiom 4, so [c](P & Q) by axiom 1; so (n)(P & Q); and similarly if d<c. **QED**.

Since points determine the truth-values of all propositions, they are also transparent to negation, and hence to all the propositional connectives:

Lemma 7. If x is a point then (x)~P iff not (x)P

Proof. If (x)P and (x)~P then (x)(P & ~P) so for some c in x, [c](P & ~P), contradicting axiom 2. So not ((x)P and (x)~P). But either (x)P or (x)~P, since x is a point. So not (x)P iff (x)~P. **QED**

If a proposition holds anywhere in a context then it must hold at a point in the context:

Lemma 8. <c>P iff there is a point x inside c with (x)P

Proof. Suppose c in x and (x)P; then there is a d with [d]P and c<d or d<c, so either [c]P or there is a d<c with [d]P; so <c>P. Now suppose that <c>P; then by axiom 5, there is a d<c with [d]P. {c,d} is a bounded nest with bound d, so by lemma 4 there is a point x inside d, hence inside c. Since x contains d, (x)P. **QED**

It follows that truth throughout a context, i.e. the ist case, is exactly truth at all points inside the context:

[c]P iff
~<c>~P iff
~(exists x inside c with (x)~P) iff
for all x inside c, ~(x)~P iff
for all x inside c, (x)P

Finally, it remains to show that discriminating points is sufficient to map the subcontext relation into the subset relation on sets of points.

Lemma 9. c<d iff every point x inside c is also inside d

Proof. If c<d and x is inside c, then x is inside d by definition 7 and axiom 3. Suppose not c<d; then by lemma 4 there is an e with e<c and ~(e<d). {e} is a bounded nest, so there is a point x inside e. If x is inside d, then e and d would be in a nest, so d<e or e<d; so dOe; so, x is not inside d. **QED**

This establishes the main result we required: if we think of a context as *being* the set of points inside the context, then:

- being a part of a context means being a subset of the context;

- truth throughout a context, the original ist, means being true at all the points in the set; and dually, wist means being true at some of the points in the set.

- truth at a point is transparent to negation and conjunction, and hence to all the Boolean connectives.

It follows that any 'pattern' of truth in a context can be described by stating it in terms of truth at points in the corresponding set, and using quantification over these points.

Discussion of the Axioms

Are these axioms plausible? Certainly axiom 2 has given rise to controversy, and seems inappropriate for some notions of context; we discuss this in more detail below. I claim that axioms 1 and 3 are uncontroversial, and amount to little more than definitional choices. Similarly, it is easy to see that axiom 4 amounts to little more than identifying [c]P as the 'strong' modality of the pair.

Although axioms 5 and 6 may seem rather unusual, they amount to two halves of the plausible assumption that the sets of contexts and propositions are well-suited to one another, in the sense that neither set has structure which is irrelevant to the other.

Axiom 6 amounts to a claim that the set of contexts does not contain distinctions which are not reflected somehow in a proposition. This is in a sense trivial, since if the set of contexts does not satisfy this axiom, it can be replaced with an equivalent set which does, by taking the quotient under the obvious equivalence. Define c#d to mean: ([c]P iff [d]P) *for all propositions* P. Clearly this is an equivalence relation, and if we take the quotient of the set of contexts under # and re-define [x]P to mean: [c]P *for all* c *in the equivalence class* x, then the resulting context space is indistinguishable from the original as far as the truth of any proposition is concerned, and satisfies the separation axiom 6 by construction. Thus, axiom 6 can always be satisfied: it amounts to accepting the fact that differences between contexts which make no difference to the truth of propositions can be ignored, when we are dealing with questions concerning truth of propositions. In fact, some version of this axiom is often assumed without comment in discussions of truth in a context. (This argument would be less compelling, however, for a language which allowed other means that the use of sentences to express propositions; for example, if it allowed for quantification over propositions.)

If the propositions are expressed using a vocabulary which itself contains a symbol denoting the subcontext relation, and in which it is possible to refer to contexts and quantify over them, then axiom 6 is automatically satisfied, since one can express the necessary proposition directly in terms of the contexts themselves: it amounts to *part of d and not part of c*.

An example of the conditions under which separation holds for temporal contexts is the presence of a clock. If we think of a clock as a source of propositions of the form "*the time is ....* ", which are true just at the time-point when the proposition is asserted by the clock, then the resulting set of propositions is sufficient to separate any set of time-intervals down to the resolution provided by the clock (for example, to the nearest second, say). Another example might be contexts which are the episodes in a story or narrative, considered as subcontexts of the entire story, and separated by sentences or phrases in the text which describe some distinctive event or circumstance which is unique to that episode, and therefore can be used to refer to it, in a phrase like "the weekend when Joe came to the ranch to court Millie and the dog caught fire". In belief contexts, contexts which could not be separated would be those which were indistinguishable to the believer, so separation in this case amounts to the assumption that distinct states of belief can somehow be characterized as states in which distinct propositions are believed.

Axiom 6, then, is a requirement that the set of contexts is not too *large* for the propositional burden it is asked to bear; that it not contain distinct contexts which are propositionally indistinguishable. Axiom 5, in the same sense, is the requirement that the set of contexts is not too *small*. It insists that if a proposition holds somewhere in a context, then there must be a subcontext which captures that truth exactly.

To see this more clearly, consider a proposed counterexample to axiom 5, which might be called the 'irrational oscillator'. Let the contexts be subintervals of the unit interval on the reals, and suppose P is true at all irrational points and false at all rational points. Then <c>P is always true and [c]P always false, for any c. This fails to satisfy the axiom by virtue of the truth being too finely scattered, preventing a single 'piece' of it to be isolated by a context-point with enough precision. By assuming the axiom, therefore, we are excluding examples like this, where truth is more finely distributed than the set of context-parts is able to discriminate. Note however that if we took contexts instead to be the *subsets* of the unit interval, rather than merely the subintervals – or even if we took it to merely include closed single-point intervals [t,t] for every irrational t – then axiom 5 would be satisfied. Thus, the failure of this example to satisfy the axiom can be attributed to its failure to include enough context-parts in the space of contexts.

It might be thought that axiom 5 rules out examples such as "The light came on at some time during the afternoon", where a proposition is true at a single point in an interval; but that conclusion would be mistaken, since the d in the axiom might itself be a single point (or a set containing a single point). What the axiom requires, in cases such as this, is that the set of contexts – in this example, the set of of time-intervals – includes subcontexts which are small enough to be *the exact time referred to in the sentence* – in this case, the moment when the light came on. If there are propositions which are true only at isolated points, then the space of contexts must provide those point-like contexts. Another way to put it is that the space of contexts must provide for the possibility of understanding a logical implication P *implies* Q as the contextual assertion of Q in a context defined by P, hence requiring that such a context exist. This intuitively accords with the idea that the purpose of subcontexts is to be the repositories of patterns of truth: that a subcontext is in a sense identifiable as a 'place' throughout which some propositions are true but others may not be.

Axioms 5 and 6 are closely related to the Hausdorff property of topological spaces, which is often assumed to be a minimal criterion for a 'reasonable' topology. If we think of <c>P as saying that there is a point in the space c where P is true, and [c]P as saying that c is an open set throughout which P is true, then axiom 5 has the consequence that if <c>P and <c>~P, i.e. if there are two points in c which are propositionally distinguishable, then there are non-overlapping open sets d and e surrounding those points. Axiom 6 requires that *any* two distinct points will be propositionally separable, so this applies to any two distinct points: which is exactly the Hausdorff property.

Are contexts consistent?

This leaves axiom 2 as the only seriously controversial assumption among our axioms. It is often claimed that contexts representing beliefs or which represent the content of a document or information source, cannot be assumed to be consistent, since beliefs may be inconsistent and documents may contain inconsistencies; and of course these facts are indisputable. It would seem then that axiom 2 does not apply to cases like this. However, representing inconsistent beliefs or documents directly as *logical* inconsistencies in this way is a risky, and perhaps naive, strategy. For example, it is often claimed that contexts provide for efficient reasoning by defining 'microtheories' which allow for sets of relevant sentences to be selected from an large set. Allowing inconsistent microtheories would obviously be a destructive move, however, unless the logic used on the microtheories is itself able to handle inconsistencies in a non-classical way. The Cyc system, for example, which coined the term 'microtheory', requires them to be internally consistent, satisfying axiom 2.

It is in fact rather hard to see what classical logical principles could be applied to a context which violates axiom 2. Thus, while indeed a straightforward application of context-based representations to the description of beliefs and documents may require such violations, the resulting logic seems capable of little more than recording them, rather than reasoning usefully about them; unless the logic itself is paraconsistent. In any case, such 'psychological' contexts do not seem to support the basic intuitions of parthood on which the development above relies, so their failure to satisfy axiom 2 may be irrelevant.

What kinds of context have extents?

All of the above discussion is based on, and motivated by, an overarching intuition which thinks of contexts as entities which have an *extent* and which can therefore be thought of as having *parts*. This is an essentially spatial metaphor, well suited to notions of context which are essentially spatio-temporal; what Menzel [Menzel 99] calls 'objective' contexts. This topological/mereological framework seems less obviously well suited to more epistemic or psychological notions of context, however.

Take for example the idea of a document or information source as a context, where ist(c,P) is supposed to mean that the proposition P is asserted more or less directly by the source c. (Notice the required change from being *true in* a context to being *asserted by* a context.) Or, to take a related example, suppose a context is understood to be a work of fiction, and ist (c,P) means that P would hold in the "imaginary world" described by the fiction c. (We are here using "world" in the informal sense in which one might speak of "the world of Conan Doyle", rather than the modal-semantic sense of "possible world".) In cases like this, it is far from clear what the relation of context-parthood could be understood to mean; and certainly, the links we have noted to mereology and topology are far less relevant when applied to such cases. There seems to be no plausible reason why states of belief or fictional "worlds" should satisfy anything like the Hausdorff condition, for example.

Kinds of context

As this analysis shows, different 'kinds' of context seem to possess sharply different mathematical and logical properties. It is very difficult to come up with any assumptions which are both sufficiently nontrivial to support any useful level of mathematical or meta-theoretic analysis, and also seem to be plausible across *all* the proposed uses of the 'context' idea and the ist(c,P) notation. Our main conclusion, therefore, is to reiterate a thesis proposed some time ago (Hayes 1997): that *there is no useful single idea of "context"*, and in order to construct nontrivial theories of contextual truth, it is necessary to distinguish different conceptions of context and analyze their varying and different properties. In this spirit, then, the formal results are offered as a contribution towards a useful theory of spatiotemporal contextualization, with an accompanying suggestion that it might be useful to try to characterize the basic assumptions of other notions of context and contextualization of truth in the same formal, axiomatic style.

Acknowledgements

This work was supported in part by the DoD under the IKRIS project. I am grateful to Chris Menzel for suggesting the proof strategy used in lemma 4.

References

Allen, J. & Hayes, P. 1989 Moments and Points in an Interval-Based Temporal Logic. *Computational Intelligence*, 5: 225-238

Buvac, S., Buvac, V. & Mason, I. 1995 Metamathematics of Context. *Fundamentae Informaticae* 23(3): 263-301

Hayes, P. 1997 Contexts in context. In *Contexts in Knowledge Representation and Natural Language*, Proceedings of the 1997 AAAI Spring Symposium, Menlo Park, Calif.: AAAI Press.

Guha, R. V. 1991 *Contexts: A Formalization and some Applications*. PhD Thesis, Stanford University; Technical report STAN-CS-91-1399-Thesis

Makarios, S., Heuer, K & Fikes, R. 2006 *Computational Context Logic and Species of ist.* Unpublished memorandum KSL-06-07, Stanford Knowledge Systems Laboratory, Stanford University

McCarthy, J. 1993 Notes on Formalizing Context. In Proceedings of the 13th International Joint Conference on Artificial Intelligence, Menlo Park, calif.: International Joint Conferences on Artificial Intelligence, Inc.

McCarthy, J. & Buvac, S. 1998, Formalizing Context (Expanded Notes). In *Computing Natural Language* (ed. Aliseda, van Glabeek & Westerstahl), CSLI Lecture Notes vol. 81, CSLI Publications, Stanford University.

Menzel, C. (1999) The objective conception of context and its logic. *Minds and Machines*, 9(1), pp 29-56, 1999