What Is the "Context"

for Contextual Vocabulary Acquisition?

William J. Rapaport

Department of Computer Science and Engineering,

Department of Philosophy, and Center for Cognitive Science

State University of New York at Buffalo, Buffalo, NY 14260-2000

rapaport@cse.buffalo.edu

http://www.cse.buffalo.edu/∼rapaport/cva.html

August 13, 2004

Abstract

"Contextual" vocabulary acquisition is the active, deliberate acquisition of a meaning for a word in a text by reasoning from context, without external sources of help such as dictionaries or people. But what is "context"? Is it just the surrounding text? Does it include the reader's background knowledge? I argue that the appropriate context for contextual vocabulary acquisition is the reader's "internalization" of the text "integrated" into the reader's "prior" knowledge via belief revision.

> … *every linguistic representation of some circumstance is in principle incomplete and must be supplemented from our knowledge about the circumstance.*
>
> — Karl Bühler 1934: 255; cited by Kintsch 1980: 595.

1 The Nature of Contextual Vocabulary Acquisition

"Contextual vocabulary acquisition" (CVA) is the acquisition of a meaning for a word in a text by reasoning from textual clues and prior knowledge, including language knowledge and hypotheses developed from prior encounters with the word, but without external sources of help such as dictionaries or people. It is the task faced by anyone coming upon an unfamiliar word while reading, who has no outside source of help, but who needs to figure out a meaning for the word in order to understand the text being read.

I purposely use the indefinite noun-phrase 'a meaning *for* a word', rather than the more common definite noun-phrase 'the meaning *of* a word'. I do this in order to emphasize that the meaning that is produced by CVA is a hypothesis that is constructed and assigned to the word, rather than "the correct" (dictionary-style) definition that, in some sense, "belongs" to the word. CVA is not a process of discovery of a word's "correct" meaning (whatever that might be). Rather, it is a process of developing *a* theory about *the* meaning that a particular use of a word (the one appearing in this particular textual passage) *might* have, *temporarily* assigning that meaning to the word, and testing the hypothesis when future occurrences of the word are encountered.

Wittgenstein appears to agree with our de-emphasis on "correct" meanings, having asserted that

> a word hasn't got a meaning given to it, as it were, by a power independent of us, so that there could be a kind of scientific investigation into what the word *really* means. A word has the meaning someone has given to it. (Wittgenstein 1958: 28, as cited in Glock 2004: 426.)

But one must be careful not to fall into Humpty-Dumpty's claim that a word "means just what I choose it

to mean—neither more nor less" (Carroll 1896, Ch. 6), as Wittgenstein's last sentence suggests. What is the reader's task in CVA? I have urged that it is not to determine the "correct" meaning of the word (as it appears in the text). Neither is it (necessarily) to determine the meaning of the word *that the author intended*. Rather, all that is necessary is for the reader to determine *a* meaning for the word (as it appears in the text) *that enables the reader to understand the text sufficiently to continue reading*.

My colleagues and I are engaged in a project with a dual goal: (1) To develop a computational theory of active, deliberate CVA (Ehrlich 1995, Ehrlich & Rapaport 1997, Rapaport & Ehrlich 2000) and (2) to adapt our computational strategies for doing CVA to an educational curriculum for teaching CVA strategies to students in classroom settings (Rapaport & Kibby 2002). CVA has been investigated (though not hitherto in an integrated fashion) in artificial intelligence (AI), psychology, first- and second-language (L1, L2) acquisition, and reading education. (See Rapaport 2004 for a (partial) bibliography of CVA research in these fields.)

CVA is not restricted to fluent readers faced with a new word. Most of our vocabulary is acquired this way: People know the meanings of more words than they are explicitly taught, so they must have learned most of them as a by-product of reading or listening (Nagy & Anderson 1984, Nagy & Herman 1987). The *average* number of word families (e.g., 'help', 'helps', 'helped', 'helping', 'helper', 'helpless', 'helpful' are one word family) known by high school graduates is estimated at between 45,000 (Nagy & Anderson 1984) and 60,000 (Miller 1991). *Excellent* students who read a great deal may know 120,000 word families (Miller 1991). Learning even 45,000 words by age 18 means learning an average of some 2500 words each year; yet no more than 400 words per year are directly taught by teachers (Nagy & Anderson 1984)—4800 words in 12 years of school. Therefore, around 90% of the words we know and understand are learned from oral or written context. Learning words from context is not a once-in-a-while thing; it averages almost 8 words learned per day (Nagy & Anderson 1984).

Some of this "incidental" acquisition is the result of conscious, active processes of hypothesizing a

meaning for unfamiliar words from context. How do readers do this? The psychology, L1, and L2 literatures suggest various strategies (e.g., Ames 1966, Clarke & Nation 1980, Van Daalen-Kapteijns & Elshout-Mohr 1981, Sternberg et al. 1983, Sternberg 1987, Kibby 1995, Blachowicz & Fisher 1996, Wesche & Paribakht 1999).

But most of these strategies are quite vague. For example, Clarke & Nation 1980 (cf. Nation & Coady 1988) gives these directions (my emphasis):

step 1: "look at the word itself and its surroundings to decide on the part of speech" (this is fairly straightforward, depending only on the reader's knowledge of grammar);[1]

step 2: "look at the immediate grammar context of the word, usually within a clause or sentence" (this presumably gives such information as who does what to whom, etc.);

step 3: "look at the wider context of the word usually beyond the level of the clause and often over several sentences" (looking for causal, temporal, class-membership information, etc., of the sort recommended by Sternberg et al. 1983, Sternberg 1987, and the others cited in the previous paragraph);

step 4: "*guess* . . . *the word* and check . . . that the guess is correct".

This is hardly a detailed algorithm that could easily be followed by a student: In particular, step 4 is reminiscent of a famous Sidney Harris cartoon[2] showing a complicated mathematical formula, in the middle of which occurs the phrase, "then a miracle occurs"! (The cartoon has one mathematician looking at this and commenting to the one who wrote it: "I think you should be more explicit here . . . ".)

For another example, Sternberg et al. 1983: 139–140 offers the following "general strategy for context use" (my emphases):

[1] Note that the reader's knowledge of grammar is not in the text, but "in" the reader's mind.

[2] [http://www.sciencecartoonsplus.com/gallery.htm]

step 1: "Attempt to infer the meaning of the unknown word from the general context *preceding* the word....";

step 2: "Attempt to infer the meaning of the unfamiliar word from the general context that *follows* the word....";

step 3: "Attempt to infer the meaning of the unknown word by looking at the word parts...." (i.e., by "looking at" its morphology);

step 4: "If it is necessary ["to understand the word's meaning in order to understand the passage . . . in which it is used"], estimate how definite a definition is required; if it is not necessary, further attempts to define the word are optional....";

step 5: "Attempt to infer the meaning of the unknown word by looking for specific cues in the surrounding context ....";

step 6: "Attempt to construct a coherent definition, using internal and external cues, as well as the general ideas expressed by the passage and general world knowledge...."

step 7: "Check definition to see if meaning is appropriate for each appearance of the word in the context...."

This may appear to be a bit more detailed than Clarke & Nation's strategy. However, note first that steps 1 and 2 do not specify *how* to make the inference from context, nor do they specify how to relate these *two* inferences. Second, note that step 5 appears at best to be part of the needed specifications for steps 1 and 2, and at worst appears to merely repeat those two steps. In any case, step 5 is no more detailed than Clarke & Nation's step 3. Many questions remain: *What* should the reader do with the information found in steps 1–5? *How* should the reader make the required inference? For that matter, is it as simple as a deductive inference? And *how* should the reader "construct" a definition (step 6)?

Part of the problem is that, while many authors suggest what contextual clues to look for, few (if any) provide specific advice on what to *do* with these clues once they have been found, i.e., what reasoning

or other cognitive processes and what prior knowledge should be applied to them. One other suggested strategy is *slightly* more explicit (perhaps): Buikema & Graves 1993 have the students *brainstorm* about what the word might mean, based on textual cues and past experience. This is fair, but not very precise; nor is it easily done or easily taught (or easily computed).

Moreover, unfortunately, little (if any) of the *computational* research on the *formal* notion of reasoning within a context is directly relevant to CVA (e.g., Guha 1991, McCarthy 1993, Iwańska & Zadrozny 1997, Lenat 1998, Stalnaker 1999; Hirst 2000 suggests why this might be the case—see below). Knowing more about the nature of context, having a more precise theory of CVA, and knowing how to teach it will allow us to more effectively help students identify context clues and know better how to use them, leading to larger vocabularies and better reading comprehension.

Learning new concepts and their words is not simply "more facts" or memorizing a definition. Concept learning requires making ever more refined discriminations of ideas, actions, feelings, and objects; it necessitates "assimilating" (Piaget 1952) the newly learned concept with prior knowledge, via inference, belief revision, and reorganization of existing cognitive schemata. We are investigating ways to facilitate readers' natural CVA by developing a rigorous computational theory of how context is used and creating a systematic, viable curriculum for teaching CVA strategies, based on our AI algorithms and on analysis of CVA processes used by good readers.

Learning both concepts and their words—especially when the concept is new (i.e., not part of the reader 's background knowledge) and more especially when it is "accessible" (i.e., such that the reader has the background knowledge needed to learn it quickly)[3]—increases the learner's perception and conception

[3]E.g., 'pentimento' describes that portion of an old oil painting painted over by a new one that can be seen when the top layer chips or fades. Most readers would not know this word, nor are they likely to have ever seen pentimento in a painting, but even an unsophisticated reader has the prior knowledge necessary to learn this "pentimento" concept. By contrast, 'kurtosis' refers to the relative flatness or peakedness of a frequency distribution as contrasted with a normal distribution. Not only would readers not know this word either, but they would have to be relatively knowledgeable about statistics to learn it. See Kibby 1995, from whom

of the world and helps "students expand their ability to perceive similarities, differences, and order within the world" (Kibby 1995: 210). Learning new concepts and their words is not simply "additional knowledge" or learning a definition. Concept learning requires making ever more refined discriminations of ideas, actions, feelings, and objects; it necessitates "consolidating" (Hansson 1999) or "assimilating" (Piaget 1952) the newly learned concept with prior knowledge, which might include inference, belief revision, or reorganizing existing cognitive schemata.

2 The Problem of "Context"

Almost everyone working on this topic believes that it is possible to "figure out" a meaning for a word "from context". Other terms for this process that can be found in the literature include "construct", "deduce", "derive", "educe", "guess", "infer", and "predict". I prefer to say that the reader "computes" a meaning for an unknown word; that is what our software does, and what our algorithm-based curriculum teaches. It is also what—on the computational theory of mind—human readers do. (On the computational theory of mind, see: Putnam 1960, Fodor 1975, Rapaport 1998, Horst 2003.)

But what is "context"? Most CVA researchers in all disciplines have in mind *written* contexts, as opposed to *spoken* contexts and as opposed to a broader notion of "context" that might include visual or "situative" contexts (speaker, location, time, etc.). Still, there is ambiguity (see Engelbart & Theuerkauf 1999 for a survey): Informally, many researchers say something like this: "The reader can infer/guess/figure out, etc., the meaning of a word *from context* . . . " (e.g., Werner & Kaplan 1952, McKeown 1985, Schatz & Baldwin 1986). Sometimes they *say* that, but *mean* something like this: ". . . from context *and* the reader's background knowledge" (e.g., Granger 1977, possibly Sternberg et al. 1983, Sternberg 1987, Hastings & Lytinen 1994ab). Sometimes, instead of talking about "context *and* background knowledge", they talk about "context *including* background knowledge" (e.g., Nation & Coady 1988; see also Graesser & Bower

this footnote is paraphrased with permission.

1990). But whereas 'context' as used in these studies has the connotation of being in the *external world* (in particular, in the text containing the word), 'background knowledge' has a connotation of being in the reader's *mind*.[4] What exactly is, or should be, meant by the 'context' for contextual vocabulary acquisition?

Interdisciplinary cognitive scientists, especially, face the problem that many terms are used differently by different researchers, without any notice of the differences, often resulting in confusion. One should always try to figure out (from context, if by no other way!) how an author uses such terms. On the other hand, one should never *use* any of these terms without clearly explaining how one is using it. 'Context' is one such term. Here, I propose some definitions related to this term. I think the concepts are more important than the words we use for them, but we need to use words that we clearly define (so that they *don't* have to be figured out from context!) and that are not too unusual (so that readers don't have to figure out why we're using such a strange word).[5]

Graeme Hirst (2000) is justifiably skeptical of such attempts to pin down 'context'. But his skepticism seems to be based on the widely different uses that the term has in different disciplines (such as knowledge representation and natural-language understanding), exacerbated by formal investigations (such as McCarthy 1993) that take the term as primitive. He points out that, since anaphora is "interpreted with respect to the preceding text, … so any preceding text is necessarily an element of the context." And then he observes that the sky's the limit: Context can "include just about anything in the circumstances of the utterance, and just about anything in the participants' knowledge or prior or current experience" (Hirst 2000,

[4]Cf. footnote 1's comment on Clarke & Nation's step 1.

[5]Indeed, not only can it be dangerous to use a familiar word in an unfamiliar way, it can also be dangerous to use unusual words, as in the sad case of Baron Stueckelberg, who "anticipat[ed] some of the major discoveries" in nuclear physics but whose "curious indifference to the accepted forms and forums of communication with colleagues combined to keep [his] work peripheral to" mainstream physics. He wrote "in a convoluted style that was further complicated by his habit … of inventing a special notation to replace … the universal language of physics … [and] switch[ing] the terms for variables and parameters, revers[ing] the order of symbols and subscripts", and so on (Crease & Mann 1985: 18, 21).

§4). My point will be that, when it comes to CVA, the sky *must* be the limit.

A clue to the nature of context as needed for our purposes can be found in our CVA software: We use a knowledge-representation-and-reasoning system (SNePS; Shapiro & Rapaport 1987, 1992, 1995) to represent, in a *single* semantic-network knowledge base, *both* the information in the text *and* the reader's background knowledge. This strongly suggests that the relevant "context" for CVA of the unknown word is the entire surrounding *network* (or at least a subpart of it; for a defense of such holistic semantic theories, see Rapaport 2002, 2003).

What follows is a sequence of proposed terms and their definitions, leading to a proposal for the proper definition of 'context' as, arguably, it should be used in CVA and that is consistent with our computational cognitive approach.

3 Preliminary Definitions

3.1

An **unfamiliar word** [or, sometimes, an **unknown word**] for a reader is by definition (is_{def}) a word (or phrase) that the reader has either never seen before, or is such that he or she has only the vaguest idea about its meaning. (For a discussion of levels of familiarity with the meaning of a word, see Kibby 1995.) For convenience, let's symbolize this by '*X*'.

3.2

A **text** is_{def} a (written) passage. It could be as short as a sentence or as long as a paragraph or even several paragraphs, and it will usually contain *X*. It is not essential that the text be written: Presumably the same techniques could be applied to oral CVA (though there would be attentional and memory limitations); in any case, most CVA research concerns texts that are read, rather than heard. There have been some studies of

oral CVA, e.g., Gildea et al. 1990, Beals 1997, Aist 2000; and we have considered experimenting with the MICASE (2002) on-line corpus of spoken text (though, of course, it is transcribed, hence written).

3.3

The next definition uses a possibly awkward term of art, but it serves a useful role, and others have used it before (Brown & Yule 1983: 46–50, citing Halliday; Haastrup 1991): The **co-text** of *X* as it occurs in some text is$_{def}$ the entire text (be it one sentence or several paragraphs)[6] "minus" *X* (i.e., the entire text surrounding *X*). So, if *X* = 'brachet', and our text is:

(T1) There came a white hart running into the hall with a white **brachet** next to him, and thirty couples of black hounds came running after them. (Malory 1470: 66.)

then *X*'s co-text in (T1) is:

> There came a white hart running into the hall with a white _____ next to him, and thirty couples of black hounds came running after them.

The underscore marks the location of the missing *X*. Co-texts are often used in "cloze" tests, in which a passage with a missing word is presented to a subject, who must then "fill in the blank", e.g., determine what that word might have been. Note that in CVA, however, the reader is not usually trying to *find* a known but missing *word*; rather, the reader is hypothesizing a *meaning* for a visible but unknown word.

3.4

The reader's **prior knowledge** is$_{def}$ the knowledge that the reader has when s/he *begins* to read the text and is able to recall as needed while reading.

'Knowledge' is the common term in this phrase, though probably not the best one, since usually what is known is true (Plato, *Theaetetus* 201): But obviously a reader might have lots of mistaken beliefs,

[6] *Pace* Schatz & Baldwin 1986, however, the co-text should not be *limited* to a 3-sentence window around *X*.

so ‘prior *beliefs*’ is probably a better term. We can use ‘knowledge’ and ‘belief’ interchangeably as long as it’s clear that prior “knowledge” can be false.

Similar terms used by other researchers include: ‘background knowledge’, ‘world knowledge’, and ‘commonsense knowledge’. However, these all have slightly different connotations:

1. “*Prior* knowledge” usefully suggests that it’s what the reader has *before* reading, i.e., the beliefs that the reader brings to the text and has available for use in understanding it.

2. “*Background* knowledge” lacks that temporal connotation, but is otherwise synonymous for our purposes. It might, however, more usefully refer to the information that the text’s *author* assumes that the reader should have. (On assumptions that authors make about who their readers might be, see Ong 1975.) We could then distinguish the background knowledge *necessary* (or assumed) for understanding the text from the reader’s actual prior knowledge.

3. ‘*World* knowledge’ connotes general factual knowledge about things *other* than what the text is about.

4. Specialized, subject-specific knowledge about the text’s topic is often called ‘*domain* knowledge’.

5. ‘*Commonsense* knowledge’ connotes the sort of knowledge that “everyone” has (e.g., that water is wet, that dogs are animals, maybe that Columbus discovered America in 1492, etc., but no “domain” knowledge). I would include under this rubric both the sort of very basic commonsense information that the CYC knowledge-representation and reasoning system is concerned with (Lenat 1995) and the somewhat more domain-specific information that the “cultural literacy” movement is concerned with (Hirsch 1987, 2003).

There is much overlap among these different notions of knowledge. For instance, surely the reader’s prior knowledge includes much commonsense knowledge, and the author’s intended background knowledge might include much domain knowledge.

4 The Proper Definition of 'Context'

How, then, might we define the 'context' of *X*? 'Context' is the tricky word that should never be used without defining it. I begin with the following first approximation, with some caveats to be discussed in a moment (including what the plus-sign ('+') represents):

Definition 1 *The* **context** *of X for a reader is*$_{def}$ *the co-text of X + the reader's prior knowledge.*

I think it's quite clear that *both* co-text *and* prior knowledge are needed: To take a simple example, after reading:

(T2) Then the hart went running about the Round Table; as he went by the sideboard, the white brachet bit him in the buttock .... (Malory 1470: 66.)

most subjects infer that brachets are (probably) animals. But they do not make the inference solely from this textual premise (T2). They must also use an enthymematic premise from their prior-knowledge (Singer et al. 1990; cf. Suh & Trabasso 1993): If *x* bites *y*, then *x* is (probably) an animal. (Actually, it's a bit more complex: We don't want to infer merely that this particular white brachet is an animal, but that brachets in general are animals.)

Two claims were just made: that an enthymematic premise is needed and that it comes from prior knowledge. An enthymematic premise is a "missing premise", i.e., a premise that needs to be added to an argument to make it valid; Singer et al. 1990 call these "bridging inferences". They are indeed "bridges"—between the text and the reader's prior knowledge—though they are not themselves inferences. However, they do need to be inferred, though the inference involved is not (necessarily) deductive; rather, it is "abductive".[7]

[7] Abduction is inference to the best explanation. It is non-deductive, because it is based on circumstantial evidence. The general form of an abduction is the deductive fallacy of "affirming the consequent": If *P* implies *Q*, and *Q* is observed, then infer that *P* must (or might) have been the case; i.e., *P* can explain the observation *Q*, so perhaps *P* is the case. Cf. Hobbs et al. 1993.

That the missing premise comes from prior knowledge is not necessarily the case: It might be found among, or deductively inferred from, information in the surrounding text. But in every situation that I have come across, the missing premise does, indeed, come from the reader's prior knowledge.

To refine Def. 1, recall that "text" (and hence "co-text") is something "out there" in the world, while "prior knowledge" is something "inside" our heads, in our minds. But I believe (as do many other cognitive scientists and many, if not most, reading specialists) that, when you read, you "internalize" the text you are reading, i.e., you "bring it into" your mind (cf. Gärdenfors 1997, 1999ab; Jackendoff 2002, §10.4; Rapaport 2003). Moreover, *this "internalized" text is more important than the actual words on paper*. As a simple example, consider the following dialogue in a *Betty* comic strip (16 April 2004):

> **Betty** (to her husband, who is reading): "Is that one of those 'I Spy' books?"
>
> **Husband:** "Yes, I'm using it to sharpen up my powers of perception."
>
> **Betty:** "Is it working?"
>
> **Husband:** "I've just started. I'm looking for five red pens."
>
> **Betty** (bending over to look at her husband's book): "That's 'hens.' Five red hens."
>
> **Husband:** "I perceive I'm not off to a very good start."

Or suppose that the text says "I'm going to put the cat out", but you misread it as "I'm going to put the car out." Your subsequent interpretation or understanding of the rest of the text will be quite different from that of someone who didn't misread 'cat' as 'car'. So, what matters for your understanding of the text is not what the text actually *is*, but what you *think* the text is.

We need a name for this "internalized text". Some people might call it a "represented text", since that is what it is: a representation of the text. I hesitate about that, since 'representation' is one of those words that I warned you about at the start of this. So we might call it the reader's "mental model" of the text, but 'mental model'—though otherwise quite nice—is perilously close to being a brand name (Johnson-

Laird 1983) and best avoided. For now, I can't think of a better name than … 'internalized text'. I'll also use 'internalized co-text' for (you guessed it) the internalized co-text. Our second approximation to a better definition for the "context" of *X* would be this:

Definition 2 *The* **context** *of X for a reader is*$_{def}$ *the reader's internalized co-text of X + the reader's prior knowledge.*

But there's another problem: The internalized text "+" the prior knowledge might not be a simple "addition" or "conjunction" or "union" of the two things. An active reader will typically make some (possibly unconscious) inferences while reading. E.g., from this small bit of text:

John went to the store. He bought a book.

the reader will automatically infer that 'he' refers to John (some people would say that 'he' and 'John' co-refer; others would say that 'he' refers back to 'John'—the differences don't matter for our purposes) and may automatically infer that John bought the book in the store that he went to. Or, e.g., a reader of the phrase 'a white brachet' might infer (from prior, commonsense knowledge that only physical objects have color) that the brachet has a color or even that brachets are physical objects. Similarly, a reader might infer that, if person *A* is shorter than person *B*, who is shorter than person *C*, then *A* is shorter than *C*; or that if a knight picks up a brachet and carries it away, then the brachet (whatever 'brachet' might mean) must be small enough to be picked up and carried.

In these cases, the whole is greater than the sum of the parts: The integration of the prior knowledge with the internalized text might include some extra beliefs that are not in the text and that were not previously in the prior knowledge, i.e., that were not previously known; i.e., you can learn from reading!

But the whole might also be *less* than the sum of the parts: From reading, you can also learn that one of your prior beliefs was *mistaken.* (It's less likely, though possible—e.g., in the case of a typographical

error—that you'd conclude that a sentence *in the text* was in error; cf. Rapaport 1991; Rapaport & Shapiro 1995, 1999.) In that case, you'll be revising your beliefs by *eliminating* something.

So, that plus-sign in Defs. 1 and 2 should be taken with a grain of salt. There is a whole branch of AI, knowledge representation, and philosophy that studies this, called "belief revision". (See, e.g., Alchourrón et al. 1985, Martins & Shapiro 1988, Martins 1991, Gärdenfors 1992, Hansson 1999.) Here's a sample of some of their terminology applied to reading (but please also take some of this with a grain of salt, since the terminology isn't universally agreed on):

The plus-sign represents an operation that takes as input the reader's prior knowledge and internalized (co-)text, and that outputs an updated mental knowledge base that is a "belief-revised integration" of the inputs. As the reader reads the text, some passages from it will be *added* (i.e., unioned or conjoined) to the reader's prior knowledge, and perhaps new inferences will be drawn; this is called 'expansion' of the prior knowledge base. Other text passages will be added, followed by the *elimination* of prior–knowledge-base beliefs that are inconsistent with it (it is limited to prior beliefs, since a reader typically assumes that the text is correct, as just noted); this is called 'revision'. A few text passages (e.g., those involving typos) might be added, then rejected when seen to be inconsistent with the prior knowledge base; this is called 'semi-revision'.[8] Beliefs that are removed are said to be 'retracted'; such '*con*traction' of a knowledge base might also result in the *re*traction of other beliefs that inferentially depended upon the removed one. (This, too, often happens when reading; cf. Rapaport 1991, Rapaport & Shapiro 1995.) After the text has been fully read, the reader might consider all (relevant) beliefs in his or her newly expanded mental knowledge base, make new inferences, and eliminate further inconsistencies (such elimination is called 'consolidation'; cf. Hansson 1999 for definitions of these terms). Let's call the end result the '(belief-revised) integration' of the two inputs.

[8] A term that Baron Stueckelberg might have loved! See note 5.

Pictorially, it might look like Figure 1. That figure may be interpreted either as showing the computational system's knowledge base and the text or else as showing the reader's mind and the text. The left-hand rectangle represents the knowledge base or reader's mind; initially, it consists of (say) four propositions representing the reader's prior knowledge: PK1, PK2, PK3, and PK4. The right-hand rectangle represents the text being read; initially, it is empty (representing the time just before the reading begins). At the next time step, the reader (or computer) reads the first sentence of the text; call this sentence 'T1'. At the next time step, the reader "internalizes" T1, adding the (mental) proposition, I(T1), to the "integrated" knowledge base. Here, "I" is an internalization function, encoding most of the processes involved in reading the sentence. At the next time step (or possibly as part of the internalization process), the reader might draw an inference from I(T1) and PK1, concluding some new proposition, P5, which becomes part of the "belief-revised" integrated knowledge base. Next, T2 is read and internalized as I(T2), with perhaps a new inference to P6, and similarly for T3 and I(T3). I(T3), however, might be inconsistent with prior belief PK4, and the reader might decide to reject PK4 (i.e., to temporarily (while reading) or permanently stop believing that PK4) in favor of I(T3). Similarly, upon reading further sentences of the text, other prior beliefs (such as PK3) might be rejected and other inferences might be drawn (such as P7 from PK1 and PK2).

The important point to see is that any "contextual" reasoning is done in the "context" on the left-hand side, i.e., in the belief-revised, integrated knowledge base. The context for CVA does not consist solely of the text being read (better: the co-text of the unfamiliar word), but neither is it that (co-)text together with the reader's prior knowledge. Instead, it is the reader's internalization of the (co-)text integrated via belief revision with the reader's prior knowledge.

There's one final detail before I can present the formal definition: '*X*' was the unknown word *in the text*. But we need a mental counterpart for it, an "internalized *X*". We need this because everything else has been internalized. So, the final(?) definition of 'context' for CVA makes it a ternary relation among a reader, a word, and a text:

Definition 3 *Let T be a text.*

Let R be a reader of T.

Let X be a word in T unfamiliar to R.

Let $T - X$ be X's co-text in T.

Then **the context that** *R* **should use to hypothesize a meaning for** *R's* **internalization of** *X* **as it occurs in** $T =_{def}$ *the belief-revised integration of R's prior knowledge with R's internalization of* $T - X$.

I.e., the "context" that the reader should use in order to compute a meaning during CVA is the single, mental knowledge-base resulting from the belief-revised integration of the reader's prior knowledge with the reader's internalized (co-)text.

This view of what the full context is for CVA agrees with the experimental results of at least one reading researcher:

> Context has generally been assumed to refer to the immediate or local context that happens to surround a word. This conceptualization of context is limited in the sense that it does not take into account the mental representation that the reader is constructing on the basis of a variety of information contained in the text as well as prior knowledge. (Diakidoy 1993: 3.)

> The findings of this study point to the need to broaden our operationalization of context to include information that the reader has available in addition to information that is printed in close proximity to an unfamiliar word. In case the reader has been able to comprehend the text, then we must assume that *the amount of relevant information that the context provides is minimal when compared to the information contained in the mental representation*. (Diakidoy 1993: 84–85; my emphasis.)

It also meshes nicely with most cognitive-science and reading-theoretic views of text understanding as requiring schemata (e.g., scripts, frames, etc.; cf. e.g., Schank 1982, Rumelhart 1985), and also with most knowledge-representation and reasoning techniques in AI for processing text, including our own: The reader's mind is modeled by a knowledge base of "prior knowledge" (including commonsense knowledge, world knowledge, perhaps some domain knowledge, etc.) expressed in a semantic-network language (SNePS). As our computational cognitive agent ("Cassie"; cf. Shapiro & Rapaport 1987, 1995; Shapiro 1989) reads the text, she incorporates the information in the text into her knowledge base, making inferences and performing belief revision along the way (using SNeBR; Martins & Shapiro 1988). Finally, when asked to define one of the words she has read, she deductively searches this *single,* integrated knowledge base for information that can fill appropriate slots of a definition frame (for details, see Rapaport & Ehrlich 2000; the notion of a definition frame was adapted from Van Daalen-Kapteijns & Elshout-Mohr 1981, and the slots were inspired by Sternberg et al. 1983 and Sternberg 1987). Her definition is thus determined by relevant portions of the semantic-network knowledge base (this is a version of a conceptual-role semantics that avoids alleged evils of holism; cf. Rapaport 2002, 2003). Thus, from our computational point of view, the "context" that she uses to hypothesize a meaning for a word represented in her knowledge base consists of her prior knowledge together with that part of her knowledge base containing the information that she consolidated into it from the text. This matches our definition of 'context' for CVA.

5 Distinguishing Internalized Co-text and Prior Knowledge

Although all relevant information is in this single knowledge base, we may continue to need to distinguish between beliefs that came from the (co-)text, beliefs that were already in the reader's prior knowledge, and beliefs that arose from inferences from both of these.

First, when eliminating one of two inconsistent beliefs, we need to know their sources, so that we

would know whether to retract a prior belief or a belief originating from the text. We do this by marking each proposition with a "knowledge category": 'story', meaning that the proposition came from the text; 'life', meaning that it came from prior knowledge, etc. (for details and examples, see Ehrlich 1995, Rapaport & Ehrlich 2000).

Second, consider the following text containing the (presumably) unfamiliar word 'detritus':

(T3) The birds alert nearby anglers that a massive school of menhaden is under attack by bluefish. The razor-toothed blues tear at the menhaden like piranhas in a killing frenzy, gorging themselves, some killing even when they are too full to eat, some vomiting so they can eat again. Beneath the blues, weak fish begin to circle, snaring the **detritus** of the carnage. (Franklin 2001.)

What prior knowledge might be useful for computing a meaning for 'detritus' from this passage? One possibility is the following defeasible rule:[9]

(R) If fish *x* attacks fish *y*, and fish *z* is weaker than fish *x*, then fish *z* will only get leftovers.

From this rule and the following part of T3:

(T3.1) [W]eak fish begin to circle, snaring the detritus of the carnage.

we can infer that 'detritus' might be 'leftovers'. More precisely, one way to do this is as follows:

Consider the statements below, where R′—representing the version of R in the knowledge-representation language—and the items labeled 'PK*i*' are found in the reader's prior knowledge, 'WF1' is a Skolem constant referring to some item in the reader's prior knowledge that satisfies the conditions in PK3, and T3.1′ is the representation of T3.1):[10]

R′. $(\forall x,y,z)[(\text{Fish}(x) \wedge \text{Fish}(y) \wedge \text{Fish}(z) \wedge \text{Attacks}(x,y) \wedge \text{Weaker-than}(z,x)) \rightarrow \exists w[\text{Leftovers}(w) \wedge \text{Gets}(z,w) \wedge \forall v[\text{Gets}(z,v) \rightarrow v = w]]]$

PK1. Fish(bluefish) (i.e., bluefish are fish)

[9] A defeasible rule is, roughly, one that can be "defeated"—i.e., rejected—later, on the basis of new information.

[10] Here, we ignore difficulties in representing generics like 'bluefish', 'menhaden', etc.

PK2. Fish(menhaden) (i.e., menhaden are fish)

PK3. Fish(WF1) $\wedge$ Weaker-than(WF1, bluefish) (i.e., something—call it WF1—is a fish and is weaker than bluefish)

PK4. $(\forall x,y)[\text{Tears-at}(x,y) \rightarrow \text{Attacks}(x,y)]$ (i.e., if you tear at something, then you attack it)

PK5. $(\forall x,y)[\text{Snares}(x,y) \rightarrow \text{Gets}(x,y)]$ (i.e., if you snare something, then you get it)

T3.1′. Begin-to-Circle(WF1) $\wedge$ Snares(WF1, detritus) $\wedge$ Tears-at(bluefish, menhaden)

Using the substitution $\{x := \text{bluefish}, y := \text{menhaden}, z := \text{WF1}\}$, we can instantiate and apply *modus ponens* (MP) to R′, PK1–PK4, and T3.1′ to get:

$$\exists w[\text{Leftovers}(w) \wedge \text{Gets}(\text{WF1}, w) \wedge \forall v[\text{Gets}(\text{WF1}, v) \rightarrow v = w]]$$

(i.e., there are leftovers that WF1 gets, and the only thing that WF1 gets are those leftovers). Similarly instantiating and applying MP to PK5 and T3.1′ allows us to infer: Gets(WF1, detritus) (i.e., WF1 gets the detritus). Now, *if it were the case* that Leftovers(detritus) $\wedge$ $\forall v[\text{Gets}(\text{WF1}, v) \rightarrow v = \text{detritus}]$ (i.e., there is nothing else for the weak fish to get except the detritus of the carnage), then we would be able to *deductively* infer the consequent of R′. So, we can *abductively* infer Leftovers(detritus) (cf. Hobbs et al. 1993). This gives us a partial definition (or meaning hypothesis) for 'detritus'. This hypothesis, of course, is defeasible (i.e., it might be incorrect), yet it is plausible and can serve as a first approximation to a full definition. At the very least—but perhaps most importantly—*it enables the reader to understand this passage*.

However, we *don't* want to infer from T3.1, *which is from the text,* and (e.g.) "They (those weak fish) also snared worms.", which let's suppose is *also* in the text, that 'detritus' are worms. One way to block this is to only allow the previous inference to go through when we use prior knowledge together with internalized text information, rather than two pieces of information from the text. And one way to do *that* is to associate each proposition with its source: text or prior knowledge (or an inference from these). (As it happens, we already do this for independent reasons having to do with belief revision, as noted above.

To sum up: When we speak of figuring out a meaning for a word "from context", we should mean: from the belief-revised integration of the reader's prior knowledge with the reader's internalized co-text of the word, with each proposition in this single mental KB marked with its source.

Acknowledgments

The preparation of this paper was supported by grant REC-0106338 from the National Science Foundation. I am grateful to Marc Broklawski, Frances L. Johnson, Michael W. Kibby, Brian S. Morgan, Scott Napieralski, Stuart C. Shapiro, Karen Wieland, other members of the CVA and SNePS Research Groups, and Jan Chomicki for discussion and for comments on earlier drafts.

References

Aist, Gregory (2000), "Helping Children Learn Vocabulary during Computer Assisted Oral Reading", *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000; Austin, TX)* (Menlo Park, CA, & Cambridge, MA: AAAI Press/MIT Press): 1100–1101.

Alchourrón, Carlos E.; Gärdenfors, Peter; & Makinson, David (1985), "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", *Journal of Symbolic Logic* 50: 510–530.

Ames, Wilber S. (1966), "The Development of a Classification Scheme of Contextual Aids", *Reading Research Quarterly* 2: 57–82.

Beals, Diane E. (1997), "Sources of Support for Learning Words in Conversation: Evidence from Mealtimes", *Journal of Child Language* 24: 673–694.

Blachowicz, Camille, & Fisher, Peter (1996), *Teaching Vocabulary in All Classrooms* (Englewood Cliffs, NJ: Prentice-Hall/Merrill): 15–34.

Brown, Gillian, & Yule, George (1983), *Discourse Analysis* (Cambridge, UK: Cambridge University Press).

Bühler, Karl (1934), *Sprachtheorie* (Jena: Rischer Verlag).

Buikema, J.L., & Graves, M.F. (1993), "Teaching Students to Use Context Cues to Infer Word Meanings", *Journal of*

Reading 36(6): 450–457.

Carroll, Lewis (1896), *Through the Looking-Glass and What Alice Found There*, various editions; online at [http://www.cs.indiana.edu/metastuff/looking/lookingdir.html].

Clarke, D.F., & Nation, I.S.P. (1980), "Guessing the Meanings of Words from Context: Strategy and Techniques", *System* 8: 211–220.

Crease, Robert P., & Mann, Charles C. (1985), "The Physicist that Physics Forgot: Baron Stueckelberg's Brilliantly Obscure Career", *The Sciences* (July/August 1985): 18–22.

Diakidoy, Irene-Anna N. (1993), "The Role of Reading Comprehension and Local Context Characteristics in Word Meaning Acquisition during Reading" (unpublished Ph.D. dissertation, University of Illinois at Urbana-Champaign).

Ehrlich, Karen (1995), "Automatic Vocabulary Expansion through Narrative Context," *Technical Report 95-09* (Buffalo: SUNY Buffalo Department of Computer Science).

Ehrlich, Karen, & Rapaport, William J. (1997), "A Computational Theory of Vocabulary Expansion", *Proceedings of the 19th Annual Conference of the Cognitive Science Society (Stanford University)* (Mahwah, NJ: Lawrence Erlbaum Associates): 205–210.

Engelbart, Silke Maria, & Theuerkauf, Beate (1999), "Defining Context within Vocabulary Acquisition", *Language Teaching Research* 3(1): 57–69.

Fodor, Jerry A. (1975), *The Language of Thought* (New York: Thomas Y. Crowell Co.).

Franklin, H. Bruce (2001), "The Most Important Fish in the Sea", *Discover* 22(9) (September): 44–50; [http://newark.rutgers.edu/∼hbf/menhaden.htm].

Gärdenfors, Peter (ed.) (1992), *Belief Revision* (Cambridge, UK: Cambridge University Press).

Gärdenfors, Peter (1997), "Meanings as Conceptual Structures", in Martin Carrier & Peter Machamer (eds.), *Mindscapes: Philosophy, Science, and the Mind* (Pittsburgh: University of Pittsburgh Press).

Gärdenfors, Peter (1999a) "Does Semantics Need Reality?", in Alexander Riegler, Markus Peschl, & Astrid Von Stein (eds.), *Understanding Representation in the Cognitive Sciences: Does Representation Need Reality?* (Dordrecht: Kluwer).

Gärdenfors, Peter (1999b), "Some Tenets of Cognitive Semantics", in Jens S. Allwood & Peter Gärdenfors (eds.),

Cognitive Semantics: Meaning and Cognition (Amsterdam: John Benjamins).

Gildea, Patricia M.; Miller, George A.; & Wurtenberg, Cheryl L. (1990), "Contextual Enrichment by Videodisc", in Don Nix & Rand Spiro (eds.), *Cognition, Education, and Multimedia: Exploring Ideas in High Technology* (Hillsdale, NJ: Lawrence Erlbaum Associates): 1–29.

Glock, Hans-Johann (2004), "Was Wittgenstein an Analytic Philosopher?", *Metaphilosophy* 35(4): 419–444.

Graesser, Arthur C., & Bower, Gordon H. (eds.) (1990), *Inferences and Text Comprehension* (San Diego: Academic Press).

Granger, Richard H. (1977), "Foul-Up: a Program that Figures Out Meanings of Words from Context", *Proceedings of the 5th International Joint Conference on Artificial Intelligence (IJCAI-77, MIT)* (Los Altos, CA: William Kaufmann): 67–68.

Guha, R.V. (1991), "Contexts: A Formalization and Some Applications" (Stanford University Ph.D. Thesis).

Haastrup, Kirsten (1991), *Lexical Inferencing Procedures or Talking about Words: Receptive Procedures in Foreign Language Learning with Special Reference to English* (Tübingen: Gunter Narr Verlag Tübingen).

Hansson, Sven Ove (1999), *A Textbook of Belief Dynamics* (Dordrecht, The Netherlands: Kluwer Academic Publishers).

Hastings, Peter M., & Lytinen, Steven L. (1994a), "The Ups and Downs of Lexical Acquisition", *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94, Seattle)* (Menlo Park, CA: AAAI Press/MIT Press): 754–759.

Hastings, Peter M., & Lytinen, Steven L. (1994b), "Objects, Actions, Nouns, and Verbs", *Proceedings of 16th Annual Conference of the Cognitive Science Society* (Hillsdale, NJ: Lawrence Erlbaum Associates): 397–402.

Hirsch, Jr., E.D. (1987), *Cultural Literacy: What Every American Needs to Know* (Boston: Houghton Mifflin).

Hirsch, Jr., E.D. (2003), "Reading Comprehension Requires Knowledge—of Words and of the World", *American Educator* 27(1) (Spring): 10–13, 16–22,28–29, 48.

Hirst, Graeme (2000), "Context as a Spurious Concept", in *Proceedings, Conference on Intelligent Text Processing and Computational Linguistics* (Mexico City): 273–287.

Hobbs, Jerry; Stickel, Mark; Appelt, Douglas; & Martin, Paul (1993), "Interpretation as Abduction", *Artificial Intelligence* 63(1–2): 69–142.

Horst, Steven (2003), “The Computational Theory of Mind”, in Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, [http://plato.stanford.edu/entries/computational-mind/].

Iwańska, Łucja M. & Zadrozny, Wlodek (1997), Special Issue on Context in Natural Language Processing, *Computational Intelligence* 13(3).

Jackendoff, Ray (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution* (Oxford: Oxford University Press).

Johnson-Laird, Philip N. (1983), *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Cambridge, UK: Cambridge University Press).

Kibby, Michael W. (1995), “The Organization and Teaching of Things and the Words that Signify Them”, *Journal of Adolescent and Adult Literacy* 39(3): 208–223.

Kintsch, Walter (1980), “Semantic Memory: A Tutorial”, in Raymond S. Nickerson (ed.), *Attention and Performance VIII* (Hillsdale, NJ: Lawrence Erlbaum Associates): 595–620.

Lenat, Douglas B. (1995), “CYC: A Large-Scale Investment in Knowledge Infrastructure”, *Communications of the ACM* 38(11): 33–38.

Lenat, Douglas B. (1998), “The Dimensions of Context-Space” [http://www.cyc.com/].

Malory, Sir Thomas (1470), *Le Morte Darthur*, ed. by R.M. Lumiansky (New York: Collier Books, 1982).

Martins, João P. (1991), “The Truth, the Whole Truth, and Nothing but the Truth: An Indexed Bibliography to the Literature of Truth Maintenance Systems”, *AI Magazine* 11(5): 7–25.

Martins, João, & Shapiro, Stuart C. (1988), “A Model for Belief Revision,” *Artificial Intelligence* 35: 25–79.

McCarthy, John (1993), “Notes on Formalizing Context”, *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)* (San Mateo, CA: Morgan Kaufmann): 555–560.

McKeown, Margaret G. (1985), “The Acquisition of Word Meaning from Context by Children of High and Low Ability”, *Reading Research Quarterly* 20(4): 482–496.

MICASE: Michigan Corpus of Academic Spoken English (2002), [http://www.hti.umich.edu/m/micase/].

Miller, George A. (1991), *The Science of Words* (Scientific American Library).

Nagy, William E., & Anderson, Richard C. (1984), “How Many Words Are There in Printed School English?”, *Reading Research Quarterly* 19(3): 304–330.

Nagy, William E. & Herman, Patricia A. (1987), "Breadth and Depth of Vocabulary Knowledge: Implications for Acquisition and Instruction", in Margaret G. McKeown & Mary E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Hillsdale, NJ: Lawrence Erlbaum Associates): 19–35.

Nation, Paul, & Coady, James (1988), "Vocabulary and Reading", in Ronald Carter & Michael McCarthy (eds.), *Vocabulary and Language Teaching* (London: Longman): 97–110.

Ong, Walter J. (1975), "The Writer's Audience Is Always a Fiction", *PMLA* 90(1) (January): 9–21.

Piaget, Jean (1952), *The Origins of Intelligence in Children*, Margaret Cook, trans. (New York: International Universities Press).

Putnam, Hilary (1960), "Minds and Machines", in Sidney Hook (ed.), *Dimensions of Mind: A Symposium* (New York: New York University Press): 148–179.

Rapaport, William J. (1991), "Predication, Fiction, and Artificial Intelligence", *Topoi* 10: 79–111.

Rapaport, William J. (1998), "How Minds Can Be Computational Systems", *Journal of Experimental and Theoretical Artificial Intelligence* 10: 403–419.

Rapaport, William J. (2002), "Holism, Conceptual-Role Semantics, and Syntactic Semantics", *Minds and Machines* 12(1): 3–59.

Rapaport, William J. (2003), "What Did You Mean by That? Misunderstanding, Negotiation, and Syntactic Semantics", *Minds and Machines* 13(3): 397–427.

Rapaport, William J. (compiler) (2004), "A (Partial) Bibliography (in Chronological Order) of (Computational) Theories of Contextual Vocabulary Acquisition", [http://www.cse.buffalo.edu/∼rapaport/refs.vocab.html].

Rapaport, William J., & Ehrlich, Karen (2000), "A Computational Theory of Vocabulary Acquisition", in Łucja M. Iwańska & Stuart C. Shapiro (eds.), *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language* (Menlo Park, CA/Cambridge, MA: AAAI Press/MIT Press): 347–375; errata at [http://www.cse.buffalo.edu/∼rapaport/Papers/krnlp.errata.pdf].

Rapaport, William J., & Kibby, Michael W. (2002), "Contextual Vocabulary Acquisition: A Computational Theory and Educational Curriculum", in Nagib Callaos, Ana Breda, and Ma. Yolanda Fernandez J. (eds.), *Proceedings of the 6th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2002; Orlando, FL)* (Orlando: International Institute of Informatics and Systemics), Vol. II: Concepts and Applications of Systemics,

Cybernetics, and Informatics I, pp. 261–266.

Rapaport, William J., & Shapiro, Stuart C. (1995), "Cognition and Fiction", in Judith Felson Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 107–128; revised and expanded version of the second half of Rapaport 1991.

Rapaport, William J., & Shapiro, Stuart C. (1999), "Cognition and Fiction: An Introduction", in Ashwin Ram & Kenneth Moorman (eds.), *Understanding Language Understanding: Computational Models of Reading* (Cambridge, MA: MIT Press): 11–25; abridged and slightly edited version of Rapaport & Shapiro 1995.

Rumelhart, David E. (1985), "Toward an Interactive Model of Reading", in H. Singer & R.B. Ruddell (eds.), *Theoretical Models and Processes of Reading; 3rd edition* (Newark, DE: International Reading Association).

Schank, Roger C. (1982), *Reading and Understanding: Teaching from the Perspective of Artificial Intelligence* (Hillsdale, NJ: Lawrence Erlbaum Associates).

Schatz, Elinore Kress, & Baldwin, R. Scott (1986), "Context Clues Are Unreliable Predictors of Word Meanings", *Reading Research Quarterly* 21(4): 439–453.

Shapiro, Stuart C. (1989), "The CASSIE Projects: An Approach to Natural Language Competence", in João P. Martins & Ernesto M. Morgado (eds.), *EPIA 89: 4th Portuguese Conference on Artificial Intelligence Proceedings, Lecture Notes in Artificial Intelligence 390* (Berlin: Springer-Verlag): 362–380.

Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in Nick Cercone & Gordon McCalla (eds.), *The Knowledge Frontier: Essays in the Representation of Knowledge* (New York: Springer-Verlag): 262–315; shorter version appeared in *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86, Philadelphia)* (Los Altos, CA: Morgan Kaufmann): 278–283; revised version of the shorter version appears as "A Fully Intensional Propositional Semantic Network," in Leslie Burkholder (ed.), *Philosophy and the Computer* (Boulder, CO: Westview Press, 1992): 75–91.

Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family," *Computers and Mathematics with Applications* 23: 243–275; reprinted in Fritz Lehmann (ed.), *Semantic Networks in Artificial Intelligence* (Oxford: Pergamon Press, 1992): 243–275.

Shapiro, Stuart C., & Rapaport, William J. (1995), "An Introduction to a Computational Reader of Narrative," in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective*

(Hillsdale, NJ: Lawrence Erlbaum Associates): 79–105.

Singer, Murray; Revlin, Russell; & Halldorson, Michael (1990), “Bridging-Inferences and Enthymemes”, in Graesser & Bower 1990: 35–51.

Stalnaker, Robert C. (1999), *Context and Content* (New York: Oxford University Press).

Sternberg, Robert J. (1987), “Most Vocabulary is Learned from Context”, in Margaret G. McKeown & Mary E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Hillsdale, NJ: Lawrence Erlbaum Associates): 89–105.

Sternberg, Robert J.; Powell, Janet S.; & Kaye, Daniel B. (1983), “Teaching Vocabulary-Building Skills: A Contextual Approach”, in A.C. Wilkinson (ed.), *Classroom Computers and Cognitive Science* (New York: Academic Press): 121–143.

Suh, Soyoung, & Trabasso, Tom (1993), “Inferences during Reading: Converging Evidence from Discourse Analysis, Talk-Aloud Protocols, and Recognition Priming”, *Journal of Memory and Language* 32: 279–300.

Van Daalen-Kapteijns, M.M., & Elshout-Mohr, M. (1981), “The Acquisition of Word Meanings as a Cognitive Learning Process”, *Journal of Verbal Learning and Verbal Behavior* 20: 386–399.

Werner, Heinz, & Kaplan, Edith (1952), “The Acquisition of Word Meanings: A Developmental Study”, *Monographs of the Society for Research Child Development*, 15 (1950) 51(1).

Wesche, Marjorie, & Paribakht, T. Sima (eds.) (1999), “Incidental L2 Vocabulary Acquisition: Theory, Current Research, and Instructional Implications”, *Studies in Second Language Acquisition* 21(2).

Wittgenstein, Ludwig (1958), *Blue and Brown Books* (Oxford: Blackwell).

Note: All "contextual" reasoning is done in this "context":

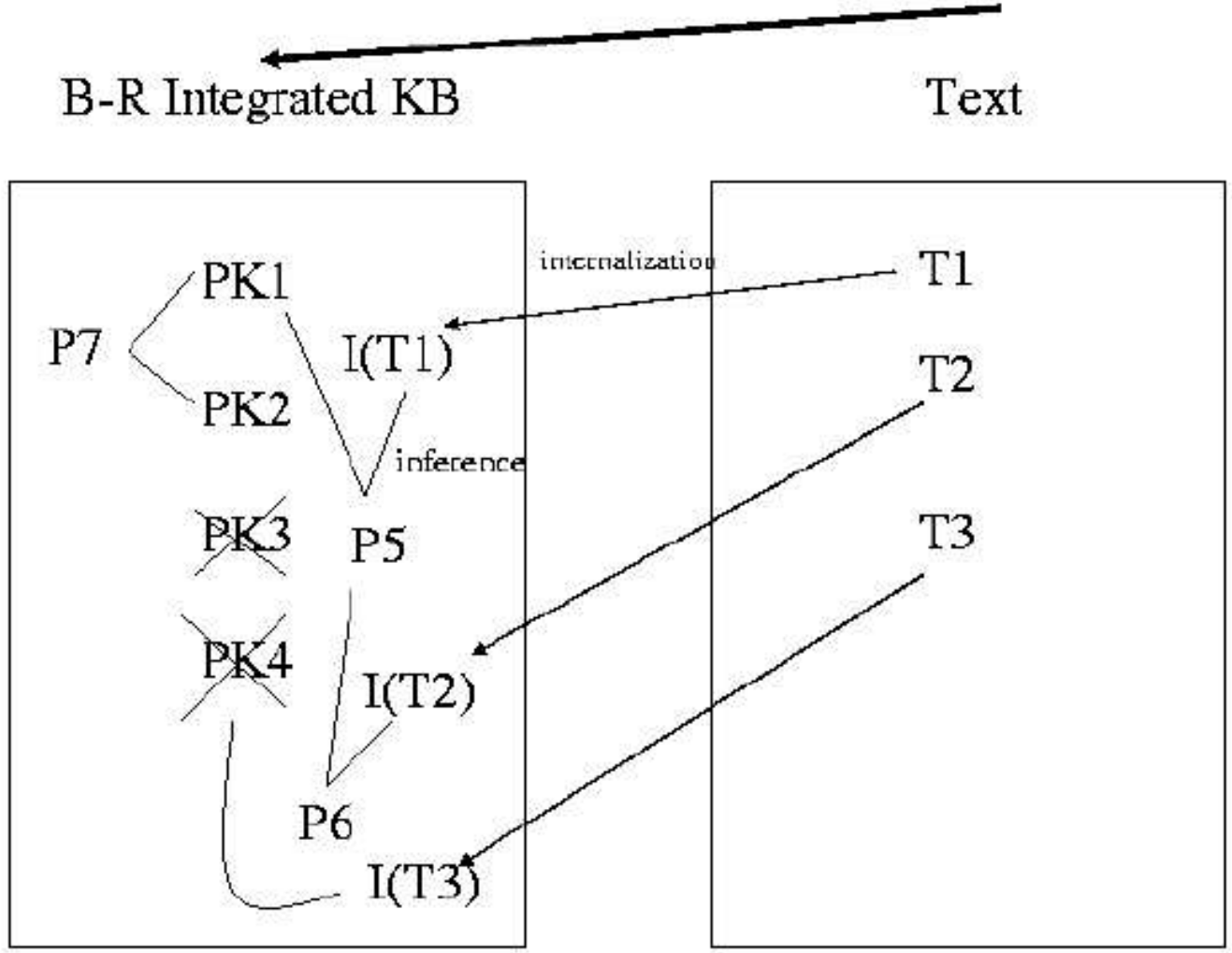

Figure 1: A belief-revised, integrated knowledge base and a text.