On 'Extracting knowledge from text':
Modelling the architecture of language users

Stephen Regoczei
Computer Studies
Trent University
Peterborough, Ontario
CANADA K9J 7B8

Graeme Hirst
Department of Computer Science
University of Toronto
Toronto, Ontario
CANADA M5S 1A4

Abstract

We propose a unified view of natural language understanding and knowledge acquisition. Knowledge is not 'extracted' from a text, but rather is *added* to the text by a 'cogniting' agent. The text, and whatever is contained in it, serves only as a triggering mechanism. This process of addition is **concept cluster attachment**. This can be generalized to artificial notations such as mathematical formulas and diagrams (even Rorschach tests!) and general signing such as facial expressions and gestures. We develop a minimal three-level architecture for a cogniting agent, consisting of verbal, conceptual, and sub-conceptual levels.

We further propose that natural language understanding and knowledge acquisition from text requires expertise. We discuss how this expertise may be acquired and incorporated into an expert system, and incrementally build up the architecture of the theoretical version of such an expert system, which we call LUKES. We discuss its implementation as LOGOS using the sortal analysis tool SORTAL.

1 Introduction

> *What is controversial is the contention that a concept is something mental and proprietary to a particular mind: for if I am to understand what you say I must attach the same meaning to your words as you do.*
> —*Anthony Quinton explaining conceptualism [Bullock et al 1988, p. 160]*

Intuitively, one feels that there should be a close connection between natural language understanding (NLU) as studied in computational linguistics, and knowledge acquisition (KA) as studied in the context of knowledge-based systems. Yet the connection between NLU and KA is seldom described explicitly. Even natural language–mediated KA [Regoczei and Plantinga 1987, Regoczei and Hirst 1989] dodges the issue.

In this paper, we would like to tackle this problem by developing requirements specifications for a system that performs both natural language understanding and knowledge acquisition from text. Our strategy is to incrementally build an architecture for a theoretical (*gedanken*) system that we will call LUKES (for Language Understanding and Knowledge Extraction System). The enlarging of the specification for LUKES is guided by the question of what features and components such a system would have to have to do a good job.

Thus we will examine the difficulties associated with NLU and KA. As we come to an understanding of what makes a particular subtask difficult, we posit a module or an operationally or computationally effective strategy to cope with the difficulty. This gives the final version of LUKES a very comprehensive architecture. While some components of this architecture may not be implementable at the present date for reasons of theoretical or technological limitations, or resource limitations, LUKES nevertheless serves as a benchmark against which the NLU component of KA tools and expert system tools can be compared. Also, if at some future date a toolmaker wants to include an NLU component in their KA tool, LUKES provides a comprehensive shopping list of 'must-have', 'should-have', and 'nice-to-have' features.

We take the implementation issue seriously, including the problem of building up large knowledge bases in order to acquire the knowledge for other, even larger knowledge bases. This bootstrapping operation is addressed through the combination of LOGOS and SORTAL in section 7.2. The implementability of LOGOS, as a KA system with NLU capabilities, is discussed.

The building up of the architecture for LUKES through a series of versions of incrementally
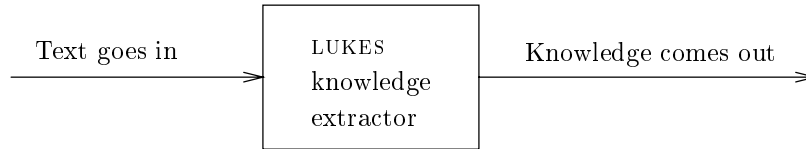
Figure 1: Knowledge acquisition made easy.

greater sophistication and power is accomplished through a series of steps. Each increment is motivated by a theoretical discussion of some severe difficulty in NLU, KA, cognitive modelling, or conceptual analysis. A solution to the difficulty is postulated, and the computationally or constructively operationalized solution is incorporated into the architecture of LUKES as a feature or a separate module. In this way we hope to build up a well-motivated, realistic, rich architecture that addresses the real problems, not merely the toy, laboratory issues that ignore the difficulties of practice.

To help the reader focus on the essential point of the paper, we should mention at the outset that the central mechanism that we will put forward both for NLU and KA is **concept cluster attachment** (CCA) [Regoczei and Hirst 1988], which we see as unifying NLU and KA. In addition, CCA generalizes from text understanding to the understanding of a broad range of symbols, signs, and other means of communication. We hope that this helps to throw some light on a severe difficulty in KA, namely the oft-repeated objection that the KA analyst does not only work with what the informant says, but with extra-linguistic clues such as gestures and facial expressions.

Throughout the paper we emphasize the role of **the agent**, and in particular, we incorporate into LUKES a model of the cognitive architecture of the agent. The agents that are being modelled are not only the one that is doing the NLU and the KA but also those that are mentioned during the NLU or KA process.

Because of space limitations, the present paper can do no more than outline certain key components of LUKES and refer briefly to the rest, in the hope that the overall argument will still be cogent. This paper is one of a series on **domain of discourse creation**, **conceptual analysis**, the **meaning triangle**, **multiple domains of discourse**, and the uses of **sortal analysis** and SORTAL [Regoczei and Plantinga 1987, Regoczei and Hirst 1988, Regoczei and Hirst 1989, Regoczei and Hirst in prep].

2 Knowledge acquisition as mining

2.1 Taking a metaphor literally

Phrases such as *extracting knowledge from text* are metaphors. Such metaphors are very useful, but may be misleading in certain crucial circumstances. The longer phrase, *extracting and modelling knowledge from text*, taken from the announcement of this conference, raises an additional issue: is it possible to 'model from text'? Can modelling be 'from' something? Is there an **operationalizable** interpretation of these phrases? That is, can we write an algorithm for the kind of KA activity that these phrases refer to that would run on an algorithmic machine?

The metaphor *knowledge extraction from text* suggests a sort of mining operation. We dig into the text, cut out the little bits of knowledge, and throw away the debris. Perhaps the 'ore' that results needs further refining before being cast into ingots of pure knowledge.

Now, some people are more expert at mining for knowledge—*i.e.*, text understanding—than others, and this varies with context, domain knowledge, and medium of communication. Therefore, according to the basic tenets of knowledge-based systems research, these people's expertise should be acquirable, and should be able to be built into an expert system. Let us take the metaphor literally. Figure 1 shows LUKES, our *gedanken* system for the task: text goes in and knowledge comes out.

There is a parallel between natural language understanding and knowledge acquisition from text. In the case of KA, we start off with text (be it printed text or an interview with an expert) and 'extract knowledge from it'. In NLU, we also start with text and 'find meaning in it' (figure 2). There seems to be a close similarity, so let us propose a strategy: let us identify 'mean-
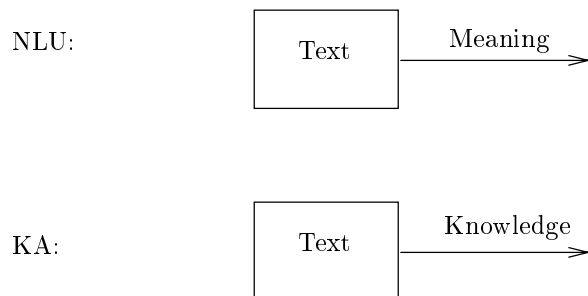
NLU:

KA:

Figure 2: NLU parallels KA.

verbal level

- - - - - - - - - - - - - - - - - - - -

conceptual level

- - - - - - - - - - - - - - - - - - - -

sub-conceptual level

Figure 3: Three levels of cogniting.

ing' with 'knowledge', thereby creating a parallel between the two activities. The support for this strategy is to be found in the ease with which we will use insights from one of the fields to help in the analysis of the other.

2.2 Operationalizing metaphors

In the computational context, **operationalizability** of a task refers to the feasibility of writing algorithms that will run on an algorithmic machine to perform the task. Thus operationalizability entails being able to find effective procedures. Operationalizability is closely linked to constructivism in mathematics. For example, in Euclidean geometry, trisection of an angle is a perfectly well understood notion, but it is not operationalizable because it cannot be performed with a compass and straightedge. In arithmetic, construction of *all* the natural numbers is not operationalizable, though constructing the first few, or even the first few million, is.

In physics, Bridgman [1950] based the ontology of physical entities on experiment and measurement. A physical quantity exists if effective measures can be taken and physical operations can be performed that monitor the effect of the hypothesized physical entities. In general, as described by Rapaport [1953], operationalism links thought with action and insists on specification of actions to be carried out as giving meaning to both linguistic utterances and theoretical entities.

It is in this spirit that we are seeking, but at present only partially finding, an operational interpretation of NLU and KA from text.
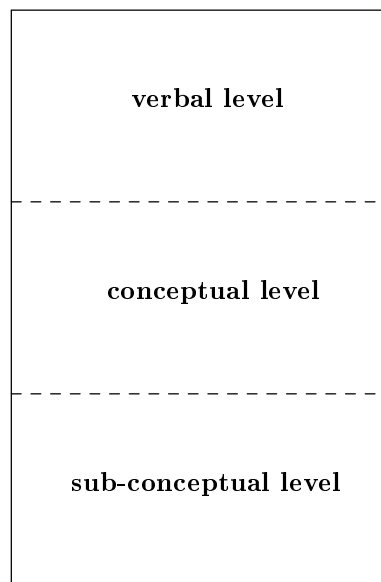
3 A three-level architecture

3.1 Three levels of cogniting

To establish our terminology, we start with a simplified model of 'cogniting' agents. Our neologism *cogniting* is a verbal form of *cognition* [Regoczei and Plantinga 1987]. Both computers and people are agents, but (so far) only people are cogniting agents. We would like to get computers to do more cogniting. (In contrast, bricks are not agents, nor do they cognit.)

A cogniting agent can be seen as operating at three levels: verbal, conceptual, and sub-conceptual (see figure 3). The **verbal** level includes the syntactic component of language use, essentially handling language as character strings without any regard for meaning, significance, or the concepts 'behind' the words. The **conceptual** level involves the construction of conceptual mental models, and the manipulation of these models to understand the external world. The **sub-conceptual** level of the agent includes its physiological embodiment and functioning, as well as all the emotions, gut feelings, and vague hunches that are very much part of human cognitive activity.

We can picture verbal activity as being at the surface level, with the deeper level of meaning provided by either the conceptual or the sub-conceptual level. At the conceptual level, we can build conceptual models of sub-conceptual

affects,[1] external-world entities, and the meaning 'behind' the words.

The main purpose of the three-level architecture is to separate the conceptual from the verbal, and to distinguish both of these from the rest, *i.e.,* from the sub-conceptual. If LUKES is to be a cogniting agent, it too must have such an architecture.

3.2 A minimal architecture for a cogniting agent

On the basis of the discussion above, we now propose a model for a language-using cogniting agent that in a sense is a minimal architecture: it contains components (though not necessarily all of them) that must be present to give an account of what happens in language use, language understanding, knowledge acquisition, and cognition in general.

Our model is shown in figure 4. The model receives both verbal and non-verbal input from the world, and can respond verbally or non-verbally. The non-verbal input could be visual or tactile, for example, and the non-verbal output could be physical actions. The numbered arrows show operations within and between levels, as follows:

(1) A verbal feedback loop, as in talking to oneself.

(2) Verbal input producing a sub-conceptual affect.

(3) Attaching concepts to words.

(4) Using conceptual structures and models to generate text.

(5) Verbalizing a sub-conceptual affect (*e.g.,* a cry of pain).

(6) A conceptual feedback loop: thinking in concepts, not words.

(7) Building a conceptual model of sub-conceptual affects.

(8) A sub-conceptual feedback loop: sub-conceptual thought (*e.g.,* making aesthetic judgements).

This is a simple, but adequate model. It may seem complicated, but it is the smallest model that suffices: three levels, one external entity, and eight operations. The only extension that might be convenient to have is a pre-processor that distinguishes verbal from non-verbal input (figure 5). This helps us to talk about the sorting, *e.g.,* how do we know that a piece of visual stimulus is actually text.[2]

To motivate our model, we mention a simple but important example from the field of communications. Much advertising, whether verbal or pictorial in nature, is directed to the sub-conceptual level of the reader. This sub-conceptual affect is often very hard to describe—it takes a good deal of analysis to be able to make the emotional reaction clear. In fact, what happens is that a conceptual model is built that reflects the sub-conceptual affect. In articulation at the verbal level, it is this conceptual model that gets described. Likewise, verbal input can act as a stimulus for the formation of conceptual structures and mental models of a conceptual kind. These in turn can be described verbally. Thus the conceptual level is central both in mediating between the sub-conceptual and the external world and supplying the deep-structure-level meaning for text. Concept cluster attachment, to be discussed in section 4.6, provides the mechanism of interaction between the surface, verbal level and the deep-structure, conceptual level.

3.3 Implications for LUKES

Following this architecture, LUKES would have to have three levels (figure 6): a verbal level that is essentially free of non-linguistic knowledge, a conceptual level using, say, conceptual graphs [Sowa 1984], and a sub-conceptual level that uses either a connectionist model or spreading activation in a network. Because LUKES is a computer program, all input is text—strings of characters—going to the verbal level. However, some strings are treated as natural language prose and some as (names and descriptions of) concepts. As output,

---

[1]Do not misread *affects* as *effects.*

[2]We are not claiming any psychological reality for such a pre-processor. On the contrary, humans almost certainly process all input in parallel as both verbal and non-verbal. For example, in the well-known Stroop colour-naming task [Stroop 1935], subjects are unable to fail to read the words they are shown, even though their task is only to name the colour of the ink in which they are written.
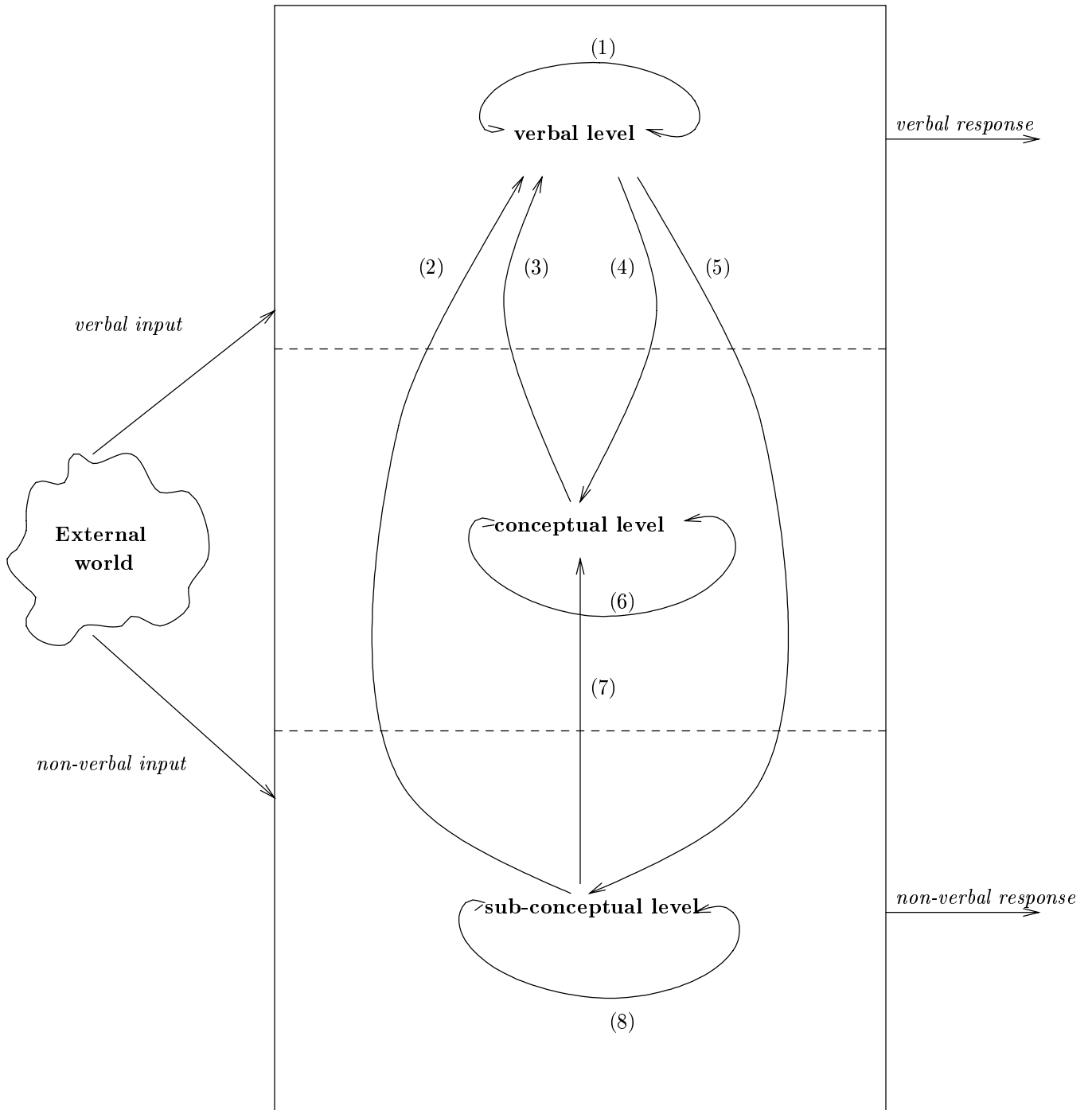
Figure 4: Minimal architecture for a cogniting agent. The numbers are explained in section 3.2.
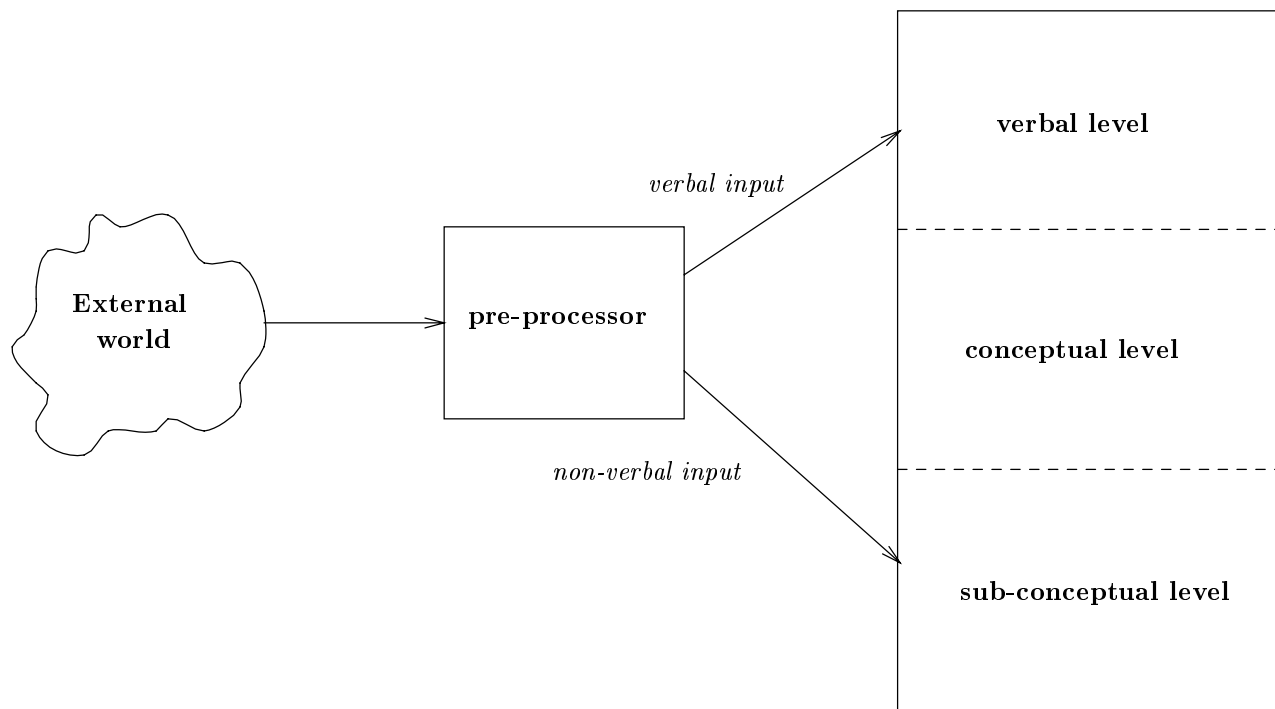
Figure 5: Adding a pre-processor that distinguishes verbal from non-verbal input.

LUKES would produce knowledge bases, with the knowledge encoded as (symbolic representations of) conceptual structures.

4 Comparing different theories of knowledge

In this section, we compare some different theories of knowledge/meaning and text understanding. We elaborate on these to motivate our own theory of understanding as **concept cluster attachment** (CCA). CCA is not a rival theory to any of the others; it can coexist, with greater or lesser harmony, with each of them. Rather, it is a synthesis, for all the other theories point to CCA as an essential component of their operationalizability. We will try to show that concept cluster attachment is the main mechanism in, and the common mechanism between, natural language understanding, artificial notation understanding, general sign understanding, and knowledge acquisition.

4.1 The content theory of knowledge

The **content theory of knowledge** is 'intuitively obvious'. The theory holds that some texts are more meaningful, more informative, more capable of conveying knowledge, because they have more **content**. This content, often referred to as the **meaning**, is what is 'extracted' from the text by the knowledgeable reader. The knowledge is there in the text, ready to be seen, extracted, and acquired. This is a coherent and perfectly satisfactory position to take to explain how meaning is conveyed from author to reader. It only breaks down when we try to implement it on a computer, or we are confronted by an individual who sees things differently, and does not find quite the same 'obvious' content in the text that we find and surmise that the author intended. In other words, the content theory works just fine until it breaks down. *Extracting knowledge from text* is a phrase firmly founded in the content theory of meaning. It is one of the 'metaphors we live by'.

In what follows, we accept the appropriateness of the content theory of knowledge or meaning, but want to extend it or reinterpret it for cases where it is not operationalizable or it breaks down. What exactly is wrong with the content theory? Mainly, it cannot be operationalized in a computationally effective way. Whatever the 'content' of the text is, it is not pure, refined knowledge or meaning. Rather, the content is best construed as a set of clues as to what kind of association can be made with the text. We
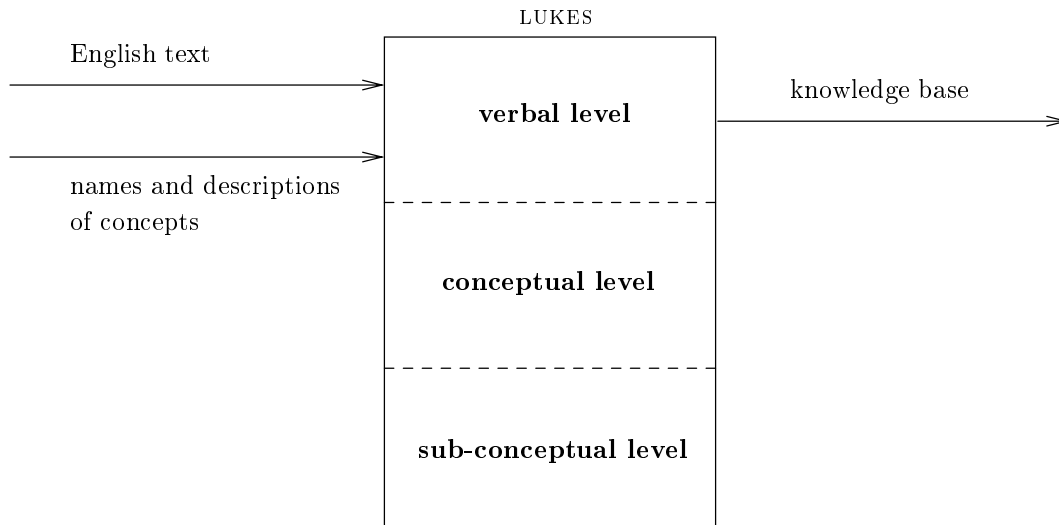
English text

names and descriptions
of concepts

verbal level

knowledge base

- - - - - - - - - -

conceptual level

- - - - - - - - - -

sub-conceptual level

Figure 6: A three-level architecture for LUKES.

NLU:

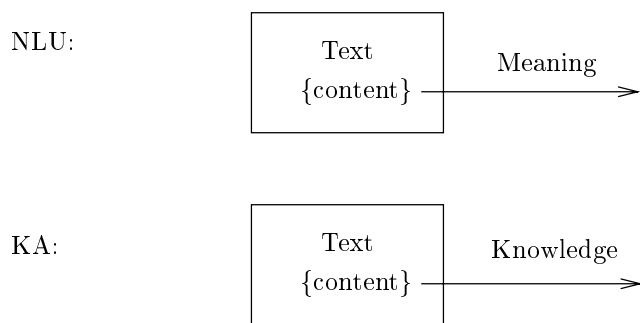Text
{content}

Meaning

KA:

Text
{content}

Knowledge

Figure 7: The content theory of knowledge.

are not enquiring about one single individual on one particular occasion; we are asking about the normally expectable associations. This requires knowledge. We have to know what knowledge can be attached—where the word *can* is to be construed as *allowable within a domain of discourse*.

To summarize: Text is seen as possessing a content, and the knowledge or meaning derived from text (see figure 7) is nothing more than the extraction of this content.

4.2 The representation theory of knowledge

The **representation theory of knowledge** holds that natural language text does not so much contain knowledge as **represent** it—almost like a pointer, it points to it. In other words, the significance of texts, notations, and signs comes from what they **stand for**. For example, a fine car represents the wealth of its owner. The car,

as a general sign, is something that semioticians and cultural and communication specialists would study. The understanding of such a sign, or the knowledge to be acquired from the sign, or the 'content' of the sign, stems from the representational role that it plays.

Like the content theory of knowledge, this is a coherent position. To make it operationalizable, we must first realize that the car does not 'represent' all by itself; there must be an interpreter, a cogniting agent. The car acts as a **clue** or **trigger** that activates certain processes within the agent. We can hypothesize some simple models for these processes. For example we might say that the interpreter has a rule in her head:

> **if** car is large and expensive-looking
> **then** owner is rich.

This is part of the knowledge base of the interpreter, a result of socio-cultural training and the background of the agent. In the absence of an interpreter, the car is not representing anything.

4.3 The hermeneutical theory of knowledge

In the **hermeneutical** tradition, knowledge of the true meaning of text is the 'holy grail' of the **hermeneutical process**. This is just a technical term for a way of thinking that has a strong grip on the popular imagination. The core of the theory is that to understand text, one performs a search for its 'true meaning'. As with the content

theory, the meaning is 'in' there—but getting it out is not easy. The popular analogy is to a sculpture that is 'in' a block of granite, and all that the sculptor has to do is get it out of there. This is a powerful image, and a very potent metaphor, easily applicable to text understanding. But it is a misleading image. There is another kind of sculpture: an assemblage of metal, produced with a blowtorch, both adding and subtracting pieces to or from the original material. If understanding is like sculpting, it is like this second kind of sculpting.

We mention these issues because of the powerful hold that the search for the 'true meaning of the text' had for a long time on computational linguistics. This manifested itself by the desire to be able to *deduce* the meaning of a text from the character string alone. This paradigm is fading, to be replaced by various knowledge-based processes.

The hermeneutical theory is the strongest challenge to concept cluster attachment. Its proponents would say that if (as we concede in sections 4.6–4.7) you are allowed to attach to a text any old concept cluster you like, you can go right ahead and do so—but that won't make it the one correct true meaning. The contest thus is concept cluster attachment versus 'the true meaning of the text'.

4.4 The information-theoretic theory of knowledge

Just as we identified knowledge with meaning, we can identify both knowledge and meaning with **information** in the Shannon–Weaver [1949] sense. On this view, information is a substance, like energy, that can be 'sent'. In fact, it can be measured, which may be the strongest evidence of its 'true existence'. The **information-theoretic** approach supports, and is supported by, powerful images and metaphors centred on communication as transmitting a message from a sender to a receiver. What these metaphors conceal is the detailed operationalized mechanism that Shannon offered [1948]: selecting from a pre-defined set of alternatives. That is, the receiver is 'pre-wired' to act or respond in accordance with the pre-defined messages that can be sent. The messages could include a request that the agent alter or 'rewire' itself—as, for example, in an ASCII transmission in which the ESC character signals the receiver to

shift to a different set of actions. But there are many different kinds of alterations one might want the agent to make—to rewire itself, say, to store uninterpreted strings, or to accept a description of a new set of responses, or even to change the way in which it accepts such descriptions. The theory is not amenable to this. It would be like trying to reprogram a pocket calculator by entering numbers.

Thus we see the difference between a signalling theory such as Shannon's and a semantic/epistemological theory that requires not only the ability of the agent to rewire itself but also a knowledge base in the agent to make possible a higher-level interpretation of the signal. Shannon is quite explicit [1948, p. 3] that he is not concerned with meaning *per se*, but rather is focusing on the physical channel that conveys information that may, in turn 'package' a kind of meaning. His analysis is thus at a different level from the one at which concept cluster attachment takes place. For example, one can correctly perceive a text in, say, German, character by character, but that's not the same as understanding it.

4.5 Miscellaneous theories

There are many other theories of knowledge that we cannot give space to here. These include the **mental-models theory** (which we used in [Regoczei and Hirst 1989]), the **deduceability theory** alluded to in section 4.3, and a **surrogate theory** that may also be called 'representational'. In the surrogate theory, the text is a surrogate for the knowledge in it. When we look at a text, the theory says, we aren't looking at knowledge, but we are as good as doing so.

We mention these theories for the sake of completeness; the objections that they raise against our position can be countered much as we did with the theories above.

4.6 The attachment theory of knowledge

We turn now to the approach that we will use in LUKES, a synthesis of the theories we have reviewed.

The question *Is there knowledge in the text?* can be answered in three ways:

1. Yes, there is knowledge in the text. This is the content theory or the hermeneutical theory.

2. No, there is no knowledge in the text. Text is lexical, knowledge is conceptual. To attribute conceptual entities as somehow 'contained' in lexical entities is a category mistake. This position cannot explain why some texts seem more 'meaningful' and better suited for the purposes of communicating than others.

3. A compromise position would claim that there is *something* in the text, even if that something cannot be characterized as knowledge. Perhaps this something is information, or clues that the astute reader, who is an expert in language understanding, picks up and uses in the process of understanding.

One does not have to take a position on these three views, because a synthesis is possible. We note that both 1 and 3 require special expertise: the first, the expertise of being able to see the content in the text, and the third, the expertise of being able to **attach** concepts to text, given the clues contained in it. Thus there is no knowledge *per se* 'in' text; what there is is a set of perceivable clues that stimulate the reader to synthesize knowledge, attach it to the text, and even attribute it to the author. This is what we are calling **concept cluster attachment**.

Variant positions can be developed. For example, one could argue that a certain amount of knowledge *is* perceived as being 'in' the text, and to this meagre stock of initial knowledge a greater amount of knowledge is added. This would be another form of the concept cluster attachment process.

What exactly are these clues in the text? They need be nothing more than substrings. It is quite remarkable that when presented with text as a character string, such as *The cat sat on the mat*, the substring *cat* will be associated with a concept such as [FELINE], whereas most readers of this paper will look at the character string *Felugrott a macska az asztalra* and associate nothing more with the substring *macska* than [NONSENSE], [GIBBERISH], or [HUNGARIAN]. The clues may also be construed as concepts themselves, *i.e.*, the sparse initial knowledge that might be seen in the text.

CCA enables us to enlarge our set of techniques for linguistic ambiguity resolution [Hirst 1987, Hirst 1988]. For example, consider the sentence

[Hirst 1987, p. 112] *Ross was escorted from the* <u>*bar*</u> *to the* <u>*dock*</u>. Each of the underlined words is ambiguous and has more than one reading that is consistent with a reading of the other. In the older paradigm of the deduction theory of knowledge, there was no way of deducing a unique interpretation from the text alone. Now that we can attach concepts to text, we know that the ordered pairs

⟨*Ross was escorted from the bar to the dock*, [COURTROOM]⟩

and

⟨*Ross was escorted from the bar to the dock*, [HARBOURSIDE]⟩

are of considerable help. Assisting an NLU program by feeding it concepts would have been considered cheating in the earlier paradigm, which restricted processing strictly to the verbal level. Given a central concept, such as [COURTROOM] or [HARBOURSIDE], an entire concept cluster **centred on** the concept, can be attached to the sentence. It doesn't matter whether the concept clues are supplied externally or perceived as part of the initial knowledge in the text; in either case CCA starts with a small amount of knowledge and generates a larger amount.

The arbitrary nature of CCA is immediately evident, and has to be addressed. Could wildly different things be added to the sentence? Yes, and they often are. These spurious additions are not weeded out until the interlocutors have harmonized their mental models. Often they are not weeded out at all, but are killed and thrown away along with the rest of the **microdomain** that was being formed as part of the harmonization process.

We need a theory like CCA to explain how an otherwise meaningful text can be given to somebody who fails to understand it or fails to find the meaning 'contained' in the text. The phenomenon of 'not-understanding' is of great significance. In CCA theory, not-understanding is equivalent to not being able to attach any concepts to the text. For example, when the reader is confronted by the words *kutya, farkas, róka,* she may baffled—no concepts 'come to mind'. This contrasts sharply with what does happen when the reader is confronted with *dog, wolf, fox*. Any coherent theory

Notation	Meaning
cat	word
[CAT]	concept
⟨⟨⟨[CAT], [DOG], ... ⟩⟩⟩	concept cluster
⟨⟨⟨[CAT] | [DOG], ... ⟩⟩⟩	concept cluster centred on the concept [CAT]

Figure 8: Notation used in concept cluster attachment.

of NLU or KA has to come to grips with the question of what does or doesn't happen when we *don't* understand something.

In figure 8, we summarize our notation for words, concepts, and concept clusters.

In what follows, we focus on the position that the knowledge, or at least the right level of knowledge, is not in the text, but has to be added by an agent. We then explore further the mechanism of concept cluster attachment, and the expertise required to be good at it.

4.7 Examples of concept cluster attachment

The terms *concept, concept cluster, concept cluster centred on a concept*, and *attachment of concept clusters to text* require some clarification. A **concept** is a cognitive construct [Regoczei and Hirst 1989], a component of a mental model used to understand the world. These concepts are specific to an agent. They are typically personal versions of public concepts [Regoczei and Hirst 1988]. A **concept cluster** is a loose collection or set of concepts. We use this as an informal precursor to the more formal and structured mechanisms such as conceptual graphs [Sowa 1984]. As for the other terms, the best way to indicate what they mean is through examples.

A good way to illustrate concept cluster attachment and its fundamentally arbitrary nature is through the old Rorschach-test joke: The new recruit goes to the army psychiatrist who shows him a Rorschach inkblot and asks, "What does this remind you of?" "Sex," answers the recruit. All of the other inkblots get the same response. The punchline of the joke varies in different versions;[3] here, we only need note that the recruit effected a concept cluster attachment, expressed in our notation as ⟨*inkblot_i*, [SEX]⟩. It is futile to ask whether sex is in the inkblot (content theory),

whether it symbolizes sex (representation theory), whether sex is the true meaning of the inkblot (hermeneutical theory), or whether there is an information component in the inkblot message that can be decoded as sex (information-theoretic theory). No, the attachment has more to do with the pre-occupation of the young recruit than anything else. In spite of this, however, we may accept the claim that in some sense at least, the recruit's concept cluster attachment was completely arbitrary.

Cutting down on the arbitrariness of the attachment is a social process—in particular, a **harmonization of mental models** [Regoczei and Hirst 1989]. The uninitiated would attach no concept to the word *kutya*, but those expert at handling such words would attach a concept cluster such as

⟨⟨⟨[DOG] \| [PET], ... ⟩⟩⟩

(if Hungarian) or

⟨⟨⟨[PORRIDGE] \| [CHRISTMAS], ... ⟩⟩⟩

(if Ukrainian).

There is a social process of training that draws it into an individual's head that the word *cat*, for example, is to be associated with a cluster such as

⟨⟨⟨[TABBY], [TOMCAT], [PERSIAN], ... ⟩⟩⟩

but not

⟨⟨⟨[PETAL], [SEPAL], [PISTIL], ... ⟩⟩⟩

The knowledge base that enables one to effect the concept cluster attachments that are 'normal', 'accepted', 'common', or 'according to the dictionary' may look like the conceptual lexicon given by Sowa [1984]. Acquiring such a knowledge base is a long learning experience. Endowing an NLU system with such a knowledge base is a basic necessity.
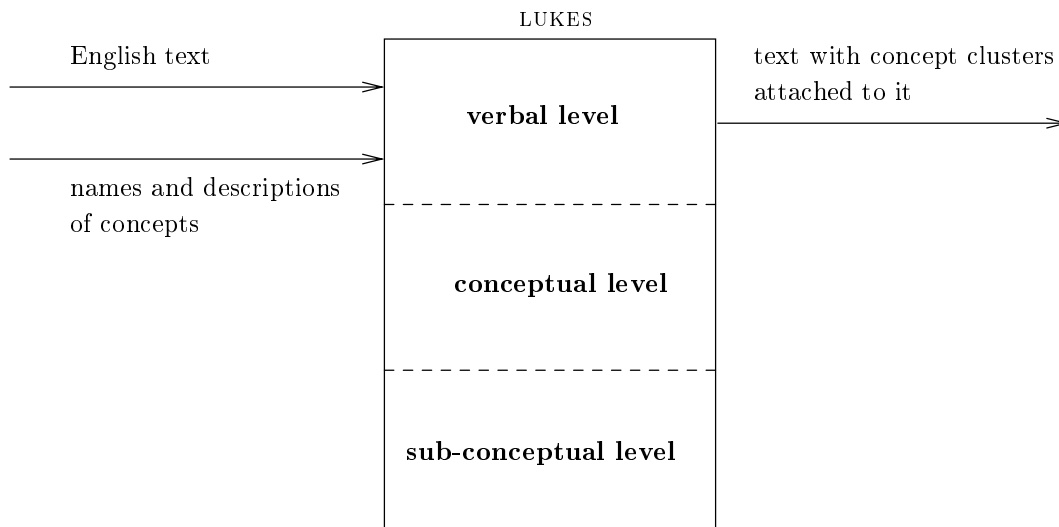
---

[3]For example: **Psychiatrist:** "I think you have a fixation on sex." **Recruit:** "I have a fixation on sex?! *You're* the one who's been showing me all these dirty pictures!"

LUKES

```
                     English text                    text with concept clusters
          ┌──────────────────────────┐               attached to it
          │                          │
          │       verbal level       │─────────────────────────────────────▶
          │                          │
          │ - - - - - - - - - - - -  │
  names and descriptions
  of concepts
          │     conceptual level     │
          │                          │
          │ - - - - - - - - - - - -  │
          │                          │
          │   sub-conceptual level   │
          │                          │
          └──────────────────────────┘
```

Figure 9: LUKES with concept cluster attachment. The output consists of ordered pairs, $\langle T, C \rangle$.

4.8 Implications for LUKES

Concept cluster attachment suggests a new approach to knowledge bases. A knowledge base is now a collection of concept clusters attached to text. So knowledge is captured both at the verbal level and the conceptual level. The architecture of LUKES is now modified as follows: The output is sortally composite entities consisting of ordered pairs $\langle T, C \rangle$ where the first component, T, is a lexical entity and the second, C, is a set of concept clusters (figure 9).

4.9 Unifying natural language, artificial notation, and signs of a general nature with concept cluster attachment

By **natural language** we mean not just ordinary text, as it is usually found in 'natural' contexts, but also technical sublanguages such as those of medicine, law, politics, and mathematics. **Artificial notations** are also often used within the context of a sublanguage, but may transcend sublanguages. They are often 'language-like', but need not be; they can be like two-dimensional diagrams, relying upon spatial arrangement on a page for effect. Examples include calculus notation, first-order logic, dataflow diagrams, blueprints, and conceptual graphs. Such notations are artificial, *i.e.,* deliberately invented, unlike natural language.

By **signs of a general nature** we mean any

non-verbal communication. This includes human actions such as gestures, facial expression, or other 'body language' (sic!). But it also includes things like the communicative aspects of clothing, three-dimensional models of airplanes, and punching someone in the mouth.

All of these we want to treat in a unified way. While they are in some ways different, they are amenable to a uniform treatment as they all have a suggestive power that makes them usable for communication. However, our comments about artificial notation and diagrams or general signs are subsidiary points to illustrate the universality of our approach. Mostly, we restrict our attention to natural language text, and the knowledge or meaning that may be 'extracted'.

We explicitly posit concept cluster attachment as an explanatory mechanism for the understanding all these kinds of communication in response to the problem that the KA analyst does not only work with what the informant says, but also picks up extra-linguistic clues from the behaviour, intonation, and gestures of the informant. It has been an irritant of long standing in KA that while the importance of these extra-linguistic clues was recognized, and they can be even captured by tape recorder, or, better still, by video recorder, how they were to be 'translated' into knowledge to be represented in a knowledge base was totally obscure.

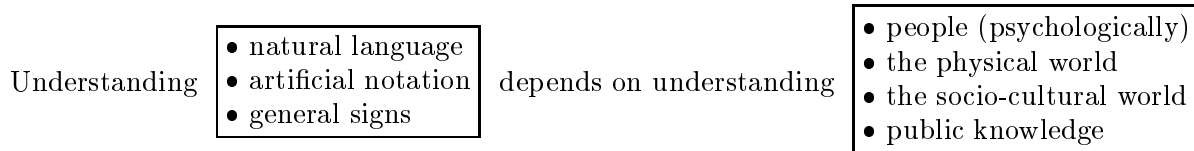| Understanding	• natural language • artificial notation • general signs	depends on understanding	• people (psychologically) • the physical world • the socio-cultural world • public knowledge

Figure 10: Two kinds of understanding.

5 The pessimistic part

NLU or KA from text is an inherently difficult process. The various theories above are not final solutions to the problem but little more than indications of the multi-faceted approach that has to be taken. Our suggestion of emphasis on the attachment theory of knowledge, in particular, concept cluster attachment and attribution, is a synthesis—a promising research direction and a way of developing software architectures. In putting it forward, we are not blind to the severe difficulties of NLU and KA from text. In this section, we discuss these difficulties, and in section 6 suggest a way out that is based on CCA.

5.1 Understanding understanding: A crucial distinction

There is a crucial difference between understanding sentences and understanding people. Most often, we understand each other without the necessity of using language. In fact, the use of language indicates that understanding broke down [Regoczei 1987]. Often the difficulty of understanding a sentence boils down to knowing what is 'behind' the sentence, who the agent that uttered it is, what fragment of the physical or socio-cultural world it is about, and what bodies of public knowledge are being referenced (see figure 10).

These problems can be placed in a general context that illuminates both the problem of 'finding meaning in the text' and the 'extraction of knowledge from text'. Understanding the items in the left column of figure 10 often depends on understanding items in the right column. The physical world or the socio-cultural world can be understood through direct observation, or being informed through private versions of the public knowledge. This distinction captures the 'aboutness' of signs. They refer to something outside themselves. Understanding the signs (left column) has to be preceded by understanding what the signs are about (right column). The difficulty of understanding text entails the even greater difficulty of understanding the world.

5.2 Understanding text

Quite apart from the difficulties discussed in the previous section, there are severe difficulties just at the text level:

Syntax: Natural language is syntactically complex, and syntactic structure and semantic interpretation each depend upon the other. In addition, much natural text is syntactically ill-formed.

Ambiguity: Words, syntactic structures, and modes of reference can all be ambiguous. The listener has to decide among the possibilities.

Semantic problems: Utterances can be vague, *i.e.,* with no fully determined meaning, or involve non-literal devices such as metaphor and metonymy.

Reference: Pronouns and other definite references must be resolved.

Discourse connections: The meaning of a paragraph is more than just that of its sentences, but also involves the relationship between them. The relationship is usually not explicit, but must be determined by the listener. Likewise, a longer text is more than just the sum of its paragraphs.

Research in computational linguistics is addressing all these issues, and there are already many interesting (partial) solutions.

6 The optimistic part

The expert-system approach assumes that a difficult activity can be handled by a system with a knowledge base that incorporates expertise. Then the difficulty of performing the task reduces to the difficulty of building the knowledge base. We will follow this paradigm by investigating how to turn LUKES into an expert system.

Given LUKES's three-level architecture, it needs a knowledge base at each level. Furthermore, it needs a knowledge base for each of the major activities that require special expertise, such as the operations of CCA, formulation of mental models for sub-conceptual affects, and expression of conceptual models in verbal form. Following the expert-system paradigm, the best way to build up the knowledge base is by asking an expert to tell us how she does it. This we do in the next section.

6.1 Interviewing the expert

Most people are experts at understanding their native tongue, yet when asked to describe how they understand text, they become surprisingly tongue-tied.

For example, we may go to an expert/informant with the well-known example *The cat sat on the mat* and ask her how she goes about understanding it or determining what it means.

Expert: I don't know. It doesn't mean anything. I can't tell you what it means unless you tell me what you are talking about. What are you talking about?

Analyst: Well, actually I didn't mean anything by it. It was just an example that I used.

Expert: Well, if it's empty words, and you don't mean anything by it, then how should I be able to determine the meaning?

Analyst: Well, what would people usually mean by it? How would you expect other people to understand this phrase?

Expert: I wouldn't. I recognize *The cat sat on the mat* as a silly example that philosophers use when they don't want to face issues about how people really use language in their everyday life.

In other words, the expert applied a great deal of knowledge to contextualize and determine that the string is a fake, just empty words. There is no meaning in it at all.

With a different expert, we might get a very different response.

Expert: *The cat sat on the mat?* There's not much here to understand, but I recognize the word *cat*. It looks like English. I do know

quite a bit about cats. I've owned a few cats, and I know what a cat is but I don't know if they're talking about a specific cat in this case. My cats didn't sit on the mat. They usually preferred an armchair or the newspaper that I wanted to read. I remember the time that . . .

The analysis of this seemingly simple dialogue is quite revealing of the strategies people use to make sense of text. For example, the text string is repeated to make sure the listener heard it correctly. A remark is made at the lexical level about the word *cat* being English. Next, the referent level is broached in claiming possession of world knowledge about cats. Specific instances are cited (personal ownership) and the phrase claiming to "know what a cat is" is a claim about having possession of the concept [CAT]. Finally, the expert draws a distinction between understanding the sentence as opposed to understanding the speaker, *i.e.*, she does not know what the speaker means but indicates that she would like to continue talking by mentioning some particular feline behaviour concerning armchairs and newspapers.

Thus, careful analysis reveals a surprising richness in all this. The expert/informant does concept cluster attachment, sortal analysis, conceptual analysis, conceptual modelling, contextualization, categorization, and modelling the agents involved. She is hypothesising agents as needed (see section 6.2). The processes performed by the expert/informant can be modelled, interpreted, or pictured by drawing diagrams (*e.g.*, agent-centred meaning triangles—see [Regoczei and Hirst 1988]). We model the agent's modelling of agents and their mental models.

6.2 Positing other agents

There is a special form of expertise that helps us, experts at NLU, to cope with the difficulty of understanding not only sentences but each other—the ability to imagine what the other person is thinking and, so to speak, direct our communication to this hypothetical agent. In particular, and this is perhaps the most significant activity, one hypothesizes into being prototypical public agents, attributing to them 'normal expectations'. In other words, using the terminology of mental models, not only do we form mental models to understand the world, but we also form mental mod-

els of other agents' mental models. It is this ability that enables us to apply Gricean rules [Grice 1975], use conversational implicature, or change style to conform to the (posited) domain of discourse that others are using.

LUKES should have the facility to hypothesize cogniting agents, with the three-level-architecture specifying and distinguishing:

- what the posited agent *says*;

- what conceptual structures the posited agent may have associated with the words he or she uses; and

- what sub-conceptual states the agent may be in.

This ability of LUKES would carry out knowledge acquisition as attribution, and perform version control of the changing state of the posited agent as the text is read or the dialogue progresses.

The arbitrariness of CCA, discussed above, is overcome by the facility of being able to **harmonize** mental models; more precisely, the mental models as posited by LUKES are attributed to the agent and attached to the text received from the agent.

6.3 LUKES **as an expert system**

The concept cluster attachment module of LUKES is the most important component. To generate knowledge bases, LUKES needs a wide range of knowledge on how to attach concept clusters. Looking upon it as an expert system, CCA is its main expertise—that is, its expertise is in knowledge base formation from textual input.

Given the three-level architecture for a cogniting agent that we described in section 3, we can now list the main components of LUKES:

1. A verbal module.

2. A conceptual module.

3. A sub-conceptual module.

4. A concept cluster attachment module, consisting of:

    (a) an episodic microdomain creation submodule; and
    (b) a domain of discourse creation submodule.

5. A module to form conceptual models of sub-conceptual states.

6. A text generation module to output a symbolic description of a knowledge base.

7. A social interaction module, consisting of:

    (a) an agent-positing submodule; and
    (b) an inventory of posited agent surrogates.

In the next section, we address the implementability of such a system.

7 Implementation

7.1 **What constitutes implementability?**

The implementability of a piece of software that exists only in the form of a wish-list, *i.e.,* a requirements specification and perhaps an architecture, has to be investigated using certain stringent, practical criteria. We can consider such software as implementable (but not yet implemented) if:

1. The architecture is modularized;

2. Each module can be prototyped;

3. There is a mechanism for scaling up from the prototype to a production version; and

4. A reasonable estimate of time and money requirements for the scaling up is obtainable.

While we tried to provide a rich enough architecture for LUKES, and even looked at some prototyping possibilities inspired by the Absity system [Hirst 1987, Hirst 1988], there are serious obstacles in the way of achieving task 3 above, and hence task 4 is not possible. In the rest of this section, we investigate a direction towards implementability and some of the difficulties.

7.2 **Implementing** LOGOS **supported by** SORTAL

We will now briefly outline a possibly-implementable version of LUKES, which we will call LOGOS ("In the beginning, was the word ... "). LOGOS is a system whose architecture is derived from that of LUKES by cutting LUKES down to an implementable configuration. This necessarily entails the loss of certain capabilities and a departure from the realistic functioning of cognitive agents.
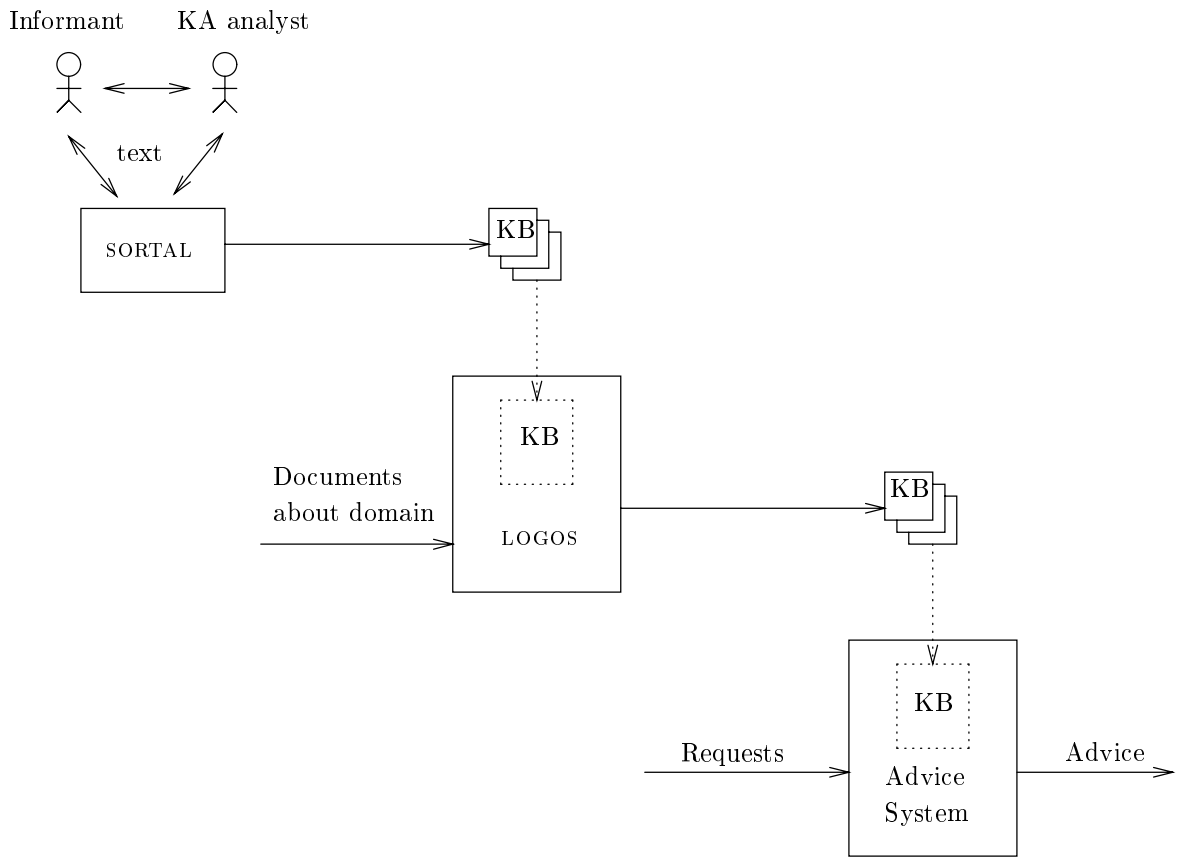
Figure 11: Bootstrapping LOGOS.

For implementation, we will drop the sub-conceptual module, because at present there is no clear way of interfacing a connectionist model of computation with an algorithmic one. LOGOS emphasizes the CCA module, with rich capabilities to create domains of discourse, including multiple domains and episodic microdomains, possibly using a concept-cluster version of Hirst's Polaroid Words mechanism [Hirst 1987, Hirst 1988]. The generation of knowledge bases, including knowledge at both the textual and conceptual levels (but expressed as symbols, *i.e.,* character strings) requires a text/symbol–generating module. To read documents, extract knowledge from text, and monitor the informant-analyst–based, natural language–mediated knowledge acquisition process, with version control of the emerging domain of discourse, requires a 'social interaction' module that posits models of language-using agents.

Because these modules are all knowledge-based, building up the knowledge bases for LOGOS becomes an important issue. We propose to use SORTAL [Regoczei and Hirst in prep], a knowledge acquisition tool that assists the analyst in sortal analysis [Regoczei and Hirst 1988], to get LOGOS started on its path of acquiring knowledge (see figure 11). SORTAL is a program to be used by a KA analyst to help set up a knowledge base with a small sample of representative text—priming the pump. The basic concepts, established using SORTAL, are passed to LOGOS. LOGOS reads the domain documents to set up the knowledge base for, say, an advisor system.

8   Conclusion

In this paper, we outlined a three-level architecture for LUKES, a theoretical natural language understanding, knowledge extraction system based upon concept cluster attachment. We discussed how LUKES could be made implementable as LOGOS by cutting its architecture down.

Experts may have hundreds, even thousands, of different kinds of techniques that they use for CCA. Our intention was not to catalogue these, but to point to a new research direction that:

1. Unites NLU and KA;

2. Places concept cluster attachment in the centre of investigations;

3. Makes the expert-system paradigm central, in that it claims that both NLU and KA can be modelled as expert systems;

4. Asks what kind of knowledge has to be in the knowledge bases of the expert system for NLU and KA; and

5. Explores these issues in the context of developing an architecture for LUKES.

In addition, with CCA being at the centre of the understanding process, there is no reason to treat natural language as a special case. Whatever we state about attaching concept clusters to natural language texts, we can say about attaching them to artificial notations such as mathematical symbols, logical formulas, and diagrams, and general signs, such as facial expressions and gestures.

Some concept clusters have privileged status. There is a long tradition of distinguishing meaning from allusion, or denotation from connotation. From the point of view of CCA, there is no difference. From the point of view of evaluation of the output, *i.e.,* the composite entity $\langle T, C \rangle$, there could be a wide range of 'correctness' ratings according to different, well-established valuations.

The theory presented here is intended to be a descriptive model, depicting what people actually do. However, we have to note that it is a formalized model, and hence deviates from the realistic messiness of the world. People are seldom aware that they are attaching concept clusters, and even less frequently do they manage to keep the separate clusters apart. In this respect, our model is too tidy. In actual, experiential terms, people just 'catch on', 'manage to grasp things', and in a bumbling sort of way, figure out 'what stuff means'. Computers, however, have to take a neater route.

Acknowledgements

References

[Bridgman 1950] Bridgman, Percy Williams (1950). *Reflections of a physicist.* NY: Philosophical Library.

[Bullock *et al* 1988]
Bullock, Alan; Stallybrass, Oliver; and Trombley, Stephen (editors) (1988). *The Fontana dictionary of modern thought* (second edition). London: Fontana Press.

[Grice 1975] Grice, H. Paul (1975). "Logic and conversation." In: Cole, Peter and Morgan, Jerry L. (editors). *Syntax and semantics 3: Speech acts.* NY: Academic Press. 41–58.

[Hirst 1987] Hirst, Graeme (1987). *Semantic interpretation and the resolution of ambiguity* (Studies in natural language understanding). Cambridge University Press.

[Hirst 1988] Hirst, Graeme (1988). "Semantic interpretation and ambiguity." *Artificial intelligence*, **34**(2), March 1988, 131–177.

[Rapaport 1953] Rapaport, Anatol (1953). *Operational philosophy: Integrating knowledge and action.* NY: Harper & Brothers.

[Regoczei 1987] Regoczei, Stephen (1987). Review of: Winograd, Terry and Flores, Fernando. *Computers and cognition: A new foundation for design*, Norwood, NJ: Ablex Publishing Company, 1985. In: *Computational linguistics*, **13**(3–4), July–December 1987, 340–344.

[Regoczei and Hirst 1988]
Regoczei, Stephen and Hirst, Graeme (1988). "The meaning triangle as a tool for the acquisition of abstract, conceptual knowledge". *Proceedings, Third Knowledge Acquisition for Knowledge-Based Systems Workshop*, Banff, November 1988. To appear in *International journal of man–machine studies*, 1989.

[Regoczei and Hirst 1989] Regoczei, Stephen and Hirst, Graeme (1989). "Knowledge acquisition as knowledge explication by conceptual analysis." In: Sowa, John F; Foo, Norman; and Rao, Anand (editors). *Conceptual graphs for knowledge systems* [to appear]. Also published as technical report CSRI-205, Computer Systems Research Institute, University of Toronto, January 1988.

[Regoczei and Hirst in prep] Regoczei, Stephen and Hirst, Graeme (1989). "Sortal analysis with SOR-TAL, a software assistant for knowledge acquisition." Technical report, Department of Computer Science, University of Toronto, in preparation.

[Regoczei and Plantinga 1987]
Regoczei, Stephen and Plantinga, Edwin Peter Owen (1987). "Creating the domain of discourse: Ontology and inventory." *International journal of man–machine studies*, **27**(3), September 1987, 235–250.

[Shannon 1948] Shannon, Claude Elwood (1948). "The mathematical theory of communication." *Bell System technical journal*, **?**(?), July and October 1948. Reprinted in [Shannon and Weaver 1949], 3–91.

[Shannon and Weaver 1949] Shannon, Claude Elwood and Weaver, Warren (1949). *The mathematical theory of communication.* Urbana: The University of Illinois Press.

[Sowa 1984] Sowa, John F (1984). *Conceptual structures: Information processing in mind and machine.* Reading, MA: Addison-Wesley (The systems programming series).

[Stroop 1935] Stroop, J.L. (1935). "Studies of interference in serial verbal reactions." *Journal of experimental psychology*, **18**, 1935, 643–662.