VARIABLE - DEPTH NATURAL LANGUAGE UNDERSTANDING

Daniel KAYSER

E.R.A. 452 du CNRS
Laboratoire de Recherche
en Informatique
Bit.490 - Campus d'Orsay
91405 ORSAY - FRANCE

Daniel C0UL0N

L.A. 262 du CNRS
Centre de Recherche en Informatique
de Nancy
ENSM1M - Pare de Saurupt
54042 NANCY - FRANCE

ABSTRACT

Standard A I representations of knowledge operate at fixed depth (i.e. the objects manipulated are described by an amount of information which remains constant for every task). Contrary to this approach, Variable Depth Processing (VDP) uses a progressive description of objects, tries different strategies according to the quality of the result it needs, and continually controls this quality by means of an evaluation of the approximations it makes.

Contextual Production Rules are shown to be an effective way to implement some features of VDP.

We are currently developing a VDP question - answering system which works on texts concerning a non-technical subject, namely an excerpt of a general public - oriented encyclopaedia.

I INTRODUCTION

The slogan *"variable depth"* appears from time to time in the A I literature (e.g. f.3 1(7)) ; nevertheless, nearly all knowledge representation systems can be said to operate at fixed depth. We shall first attempt to define these terms (fixed *depth* vs. *variable depth)* ; we then propose a scheme to embody the latter concept.

In a *fixed - depth* representation, the objects manipulated (usually words) are described by an amount of information which remains constant for every task ; for instance : a formula in predicate calculus ; a diagram in conceptual dependency ; to some extent, a "perspective" in a KRL-like language.

Fixed - depth representation is usually a step towards an alleged canonical representation of meaning. A representation is canonical with respect to an equivalence relation on meanings : two expressions yield the same representation iff they have equivalent meanings.

We claim that no universal equivalence relation can be interestingly defined ; hence, the quest for canonical representations is fruitless, unless it is' acknowledged as a first - order approximation to a much more complex phenomenon.

This work has been done under contract n°80/270 granted by l'Agence de 1'Informatique

II VARIABLE - DEPTH PROCESSING (VDP)

We propose to bestow the label to any system having one or more of the throe following features : (a) progressive description ; (b) multiple strategies ; (c) evaluation of approximation.

A. Progressive Description

The sharpness of a computational analysis needs not be the same for every word to process [7]. However, we must be capable to deepen it on any request (triggered by internal factors : a coarse examination arouses our interest [6] or external factors : intonation or typography induces to think that something is important in the author's mind).

This implies that any word should be related to an ordered amount of knowledge, which is made progressively available as the depth of analysis increases.

Remark : this order is not rigid ; there ought to be strategies (see below) able to control order shifts (something analogous to the transformation of point of view in [4] section 1.8).

For instance, from a child's perspective, a Christmas - tree could first evoke the presents, the garlands, the evening spent with the family, then — if needed — be defined in terms of a tree, with trunk, branches, resin, and related to what the child happens to know about firs, then related to more general topics, as winter, religion, etc.

It is absolutely unnecesserary to make all these chunks of knowledge available at once : most of the utterances produced or understood by the child will need but the first levels. Nevertheless, any cut-off in the definition - which would correspond to some *">*a>wni('al"* level — would be arbitrary and exclude the comprehension of some situations which are well understood by the child.

B. Multiple Strategies

We get different — and hopefully compatible — interpretations when we skim through a text, or when we read it carefully. We make more or less text - driven inferences, take more or less time to cross - check what we understood with what we know, detect potential inconsistencies, point out spelling or syntactic mistakes, etc.

This means that we are able to process a same text under various strategies.

A VDP understander should be able to select a strategy according to its needs ; more importantly, it should switch from one strategy to the other as the analysis proceeds ; it shares its resources among different processes : syntax evaluation, tight logical processing, loose semantic evaluation,

Each of these processes may in turn send requests to the knowledge base, which recursively need VDP ; from time to time, the understander should look to what happens, and possibly modify the resource allocation or the strategy. Moreover, it should be able to perform a general or partial reordering of the features stored with the words, in order to reflect a change of perspective.

C. Evaluation of Approximation

The above strategies make use of incomplete descriptions, on the grounds of which they take decisions. This is all right, as long as we know the order of magnitude of the risk we take.

To stick a label "*probable*" on some knowledge, or *"approximate"* on some methods, is far from adequate. We want a measure which classifies the knowledge and the processes in such a way that, according to one's needs, one can select a method compatible with the required degree of accuracy.

We have argued elsewhere that such a measure could neither be probabilistic, nor "fuzzy"-like, and have proposed what we call "*plausibility calculus*" [2]. This calculus relies on the assumption that only two composition laws may affect a plausibility, namely weakening (symbolized by infix S) and reinforcement (infix //) .

III CONTEXTUAL PRODUCTION RULES (CPR)

Production rules are known to be an efficient tool for knowledge representation [5]. Surprizingly, slight changes confer them the ability to adapt to VDP. Here is the syntax we use :

$$\alpha \rightarrow \beta \quad \$p\ell \ (\gamma_1)\ \$p\ell_1 \ \&\ \ldots\ \&\ (\gamma_n)\ \$p\ell_n$$

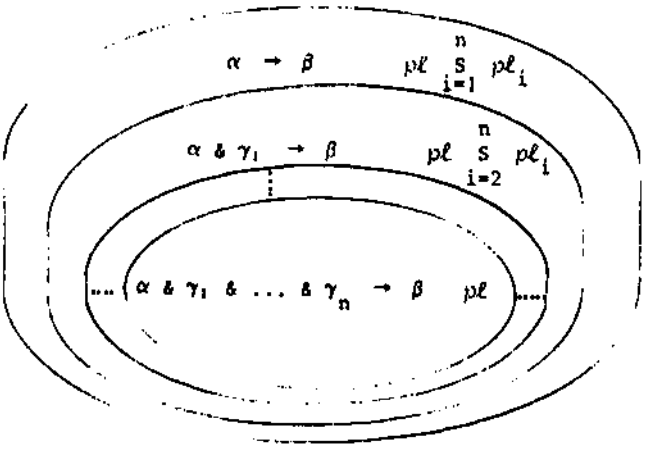
This rule means : an occurrence of α can be replaced by β provided that conditions γ_i hold. The plausibility of the sentence containing β , $p\ell(\beta)$, is computed from the plausibility of the sentence containing α , $p\ell(\alpha)$, and from the plausibility to which the γ_i have been satisfied, $p\ell(\gamma_i)$, by means of the following formula :

$$p\ell(\beta) = p\ell(\alpha)\ S\ p\ell_i \overset{n}{S}_1 (p\ell(\gamma_i)\ //\ p\ell_i)$$

Comments :

- **if the plausibility of a γ_i is unknown, the result of $p\ell(\gamma_i)$ // $p\ell_i$ is defined to be equal to $p\ell_i$; so $p\ell_i$ reflects how the final result is weakened when the condition has not been checked ; it is thus a measure of the importance of the condition ;**

- one interesting way to see the rule is to consider it as a set of n+1 rules arranged as an onion skin, ordered by decreased importance of the γ_i ; according to the depth of analysis, one uses a more or less internal rule :



- the plausibility pi is thus a residual uncertainty attached to the substitution *a* -* B , when all conditions are verified ; it copes for the factors one cannot or will not identify, but still influence the substitution ;

- the 7; can be either texts — which must tnen be recursively validated through CPRs — or function calls, to embody specific computations (generally at the morpho-syntactic levels) or local strategies.

IV CONCLUSION

A. On Canonical vs. Variable - Depth Representations

As already stated, canonical forme — aa every fixed - depth representation — are descriptions for which a degree of refinement has arbitrarily been selected.

Moreover, they have practical inconveniencies :

- they often force to resolve ambiguities, while the required level of text processing would not need it ;
- they land themselves poorly to approximation : if a procedure ia attached to one form, how to call it when the data are close, but not identical, to that form : what are then the parameters ? how to make sense of the results ? how to measure a distance between canonical forms t etc.

VDP has also its own problems (e.g. how to weight the importance of such or such factors ?) which canonical forms seem to avoid, but we believe that theae problems are intrinsically related to the nature of commonsense knowledge.

B. On Contextual Production Rules

CPEs are but one way to implement the idea of VDP in a computer. We are currently developing a question - answering system which operates on natural - language texts concerning a non - technical subject, extracted from a general public - oriented encyclopaedia.[1]

If this experiment turns out to be successful, we might consider other areas of application, because we feel that variable - depth is a feature common to many human activities, including — but not restricted to — natural language understanding

REFERENCES

[1] COULON D., KAYSER D. "Avoir réponse à tout ?" *in Proc. AFCET-RFIA Congress, Nancy (France), Sept. 1981 (forthcoming).*

[2] KAYSER D. "Vers une modélisation du raisonnement 'approximatif'" *in Proc. Colloque sur la représentation des connaissances et du raisonnement, INRIA, 1980, 440-457.*

[3] LEHNERT W., WILKS Y. "A Critical Perspective on KRL" *Cognitive Science 3 : 1 (1979) 1-28.*

[4] MINSKY M. "A Framework for Representing Knowledge" *in The Psychology of Computer Vision (P.H.Winston, ed.) , Mc Graw Hill, New York (1975) 211-277.*

[5] NEWELL A. "Production Systems : models of control structures" *in Visual Information Processing (W.C.Chase, ed.), Academic Press, New York (1973).*

[6] SCHANK R.C. "Interestingness : controlling inferences" *Artificial Intelligence 12 : 3 , (1979) 273-297.*

[7] SCHANK R.C., LEBOWITZ M., BIRNBAUM L. "Integrated Partial Parsing" *Technical Report 143, Computer Science Department, Yale University, New Haven, CT., 1978.*