

ANALOGY-BASED ACQUISITION OF UTTERANCES RELATING TO TEMPORAL ASPECTS

V. Sembugamoorthy
Computer Centre
Indian Institute of Technology
Bombay 400 076 India

ABSTRACT

PLAS is an analogy-based language acquisition system that can be taught through examples. The primitives of its knowledge representation are language expressions and are learnt through interaction in the very same language that is being learnt. This paper proposes extensions to cope with utterances relating to temporal aspects.

I INTRODUCTION

Our approach to cope with the formidable complexity of dealing with Natural language as well as learning is to divide the effort into two stages. The first-stage system will begin with sensori-motor mechanisms and learn an interpretational machinery that enables comparatively primitive language processing. The second-stage system will start with such an interpretational machinery and learn formal logic and grammar of the language that support complex performance.

We have described elsewhere [1,2,3,4] the teaching, language acquisition and understanding machinery of PLAS, a language acquisition system belonging to the first stage. PLAS is analogy-based and learns through examples. The primitives of its knowledge representation are not grammatical constructs but natural language expressions themselves. These primitives are learnt through interaction in the very same language that is being acquired (arguments for the above features are available in [1 to 5]).

Though the language machinery of PLAS can cope, within its scope, with spatial as well as temporal aspects of the external world, its sensori-motor machinery cannot represent the temporal aspects. Moreover, though it can learn to understand new phrases and sentences, its capability to acquire words is comparatively limited. This paper proposes extensions to remove these limitations (call the extended PLAS as PLAS.T).

II INITIAL KNOWLEDGE, TEACHING AND ACQUISITION

The world of PLAS.T is a system of aspects (e.g.: the meaning of YESTERDAY, Male 10) and aspect-classes (e.g.: the class of past events, the class of males) of the external world, and utterances and utterance-classes i.e., language schemas or simply schemas (e.g.: SOMEBODY WENT SOMEWHERE SOMETIME i.e., '- WENT - -') of the linguistic world.

PLAS.T starts with a system of sensori-motor aspects (A sensori-motor aspect is an aspect of the external world that PLAS.T becomes aware of through its sensing and manipulating capabilities. For example, a specific action such as walking, a specific agent.) and aspect-classes.

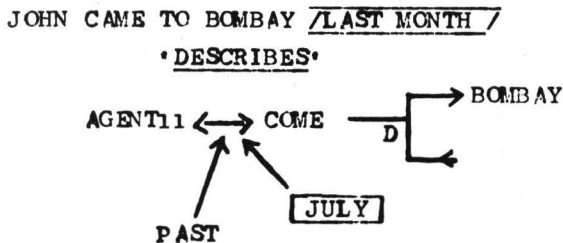
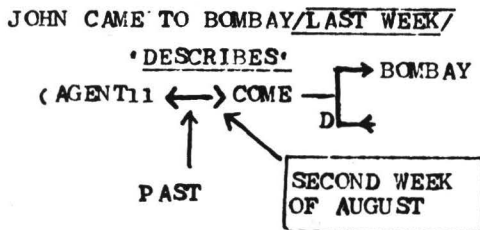
Since the case structures are closer to the structures learnt by PLAS [4] and Schank's CDN representation [6] is language-free, we propose to use CDN to represent the sensori-motor knowledge given above. However, CDN will not be used to represent the system of aspects and aspect-classes at the language level. We believe that a person's language plays a key role in determining his model of the world and thus structuring his thought [7].

Utterances, sensori-motor aspects and aspect-classes are primitively related by the teacher as in PLAS [1] by naming. For example,

CHAIR₁; CHAIR, A CHAIR; α

'α' stands for a CDN-representation of the sensori-motor aspect-class corresponding to the class of chairs. The example informs the system that CHAIR, A CHAIR are names of the aspect CHAIR₁ belonging to the aspect-class specified by α. All name-utterances need not be introduced by naming. PLAS.T, using the acquisition machinery of PLAS and the techniques discussed below, can learn them by analogy.

The teacher can also associate a linguistic and a sensori-motor representations of an event by the extra-linguistic marker 'DESCRIBES' which simulates the extra-linguistic behaviour viz., simultaneously pointing to, and describing, a situation. For Example,



(For lack of space, 'COME' has not been described fully, and structures of the linguistic representation were not shown). Applying the rule that any difference in the linguistic representation of two events relates to the difference in their sensori-motor representations, PLAS.T can associate LAST MONTH with JULY and LAST WEEK with SECOND WEEK OF AUGUST. Using another rule, PLAS.T can relate the only non-understandable unit in a linguistic representation with the only one in the sensori-motor representation that has not been related to any unit in the former. For example, the boxed units given below:



Next, linguistic representations of two events can be related by the teacher using the extra-linguistic markers, namely 'IS-A-RECOLLECTION-OF' and 'IS-AN-IMAGINATION-OF'. For example,

JILL DROPPED THE DOLL YESTERDAY
'IS-A-RECOLLECTION-OF' JILL DROPS THE DOLL

JILL WILL BE AT BOMBAY IS-AN-
IMAGINATION-OF JILL IS AT BOMBAY.

'IS-A-RECOLLECTION-OF' simulates the extra-linguistic means employed by a language community to enable the hearer to recollect a past event and associate with its representation the present one. The

other marker simulates the extra-linguistic means that enable the hearer to imagine a new situation and relate its representation with the present one. With the help of such associations between statements, and the analogy-based capability of PLAS.T to transform utterances (details in [1,2,4]), PLAS.T can convert the tense of a statement.

Teaching PLAS.T a text is as in PLAS [1,2,4] except that the extra-linguistic markers described above can also be used in addition to the ones already available in PLAS.

III UNDERSTANDING AND KNOWLEDGE REPRESENTATION

In PLAS.T a sensori-motor aspect is represented by a Picture Producer node, Action node, Picture Aider node, Action Aider node or a conceptualization [6]. An interpretation of an utterance is an aspect (at either the sensori-motor or the language level). The semantic category of an utterance is represented by a sensori-motor aspect-class (and not by a language level aspect-class to avoid complexity).

The surface-structure of a phrase or a sentence consists of the underlying schema and the surface-structures of its fillers. For example, that of JOHN BRUSHES HIS TEETH DAILY is: (- BRUSHES - -, JOHN, (HIS -, TEETH)). The aspect-class structure of an utterance (i.e., a phrase or a sentence) consists of the aspect-class of the utterance and the aspect-class structures of its fillers. The aspect-structure of an utterance consists of the aspect of the utterance and the aspect-structures of its fillers. The surface, aspect-class and aspect structure of a word is respectively the word itself, its aspect-class and aspect.

To understand an utterance, PLAS.T has to obtain its three structures (details of the understanding process are in [1,2,4]).

When an utterance is successfully understood, the utterance and its structures are integrated into the knowledge net of PLAS.T. The knowledge net called LBA-net is a Language Based (because the primitives are language expressions) Associative network [3]. It supports the following associative information retrieval.

1. Aspects represented by an utterance;
2. Utterances representing an aspect;

3. Aspect-class-structures associated with an utterance and vice versa;
4. Aspect-class-structures associated with an aspect and vice versa;
5. Fillers of a variable of a schema;
6. Utterances that can be segmented through a schema;
7. Language schemas segmenting an utterance.

These associative functions constitute the interface between the understanding process and the knowledge net (details in [3]).

IV COMMENTS

The acquired competence of PLAS.T can be tested as in PLAS [1,2,4]. If its response is incorrect, the teacher presently has to switch PLAS.T to the state of the latest correct response and modify his teaching sequence. Extensions to enable the system to debug its knowledge through question-answering have been discussed in [4]. Though some of the learning principles discussed above rely on the concept of 'near miss', the generalization techniques of PLAS which PLAS.T inherits do not require examples of 'near miss' type. Since neither PLAS nor PLAS.T is designed to learn a grammar, and the knowledge representation does not employ the usual grammatical and semantic primitives, the computational machinery of PLAS.T is bound to be unconventional.

ACKNOWLEDGEMENTS

The valuable guidance and encouragement of Prof. R. Narasimhan, Director, NCSDCT, TIFR, Bombay are gratefully acknowledged.

REFERENCES

- [1] Sembugamoorthy, V. "PLAS, a Paradigmatic Language Acquisition System; An Overview". In Proc. IJCAI -79. Tokyo, Japan, August 1979.
- [2] Sembugamoorthy, V. "An Analogy-Based Language Acquisition System that can be taught through examples". In Proc. 2nd Intern. Meet on AI. Repino, USSR, October 1980.
- [3] Sembugamoorthy, V. "A Language-Based Associative Knowledge Representation that Supports Analogy-Based Language Acquisition and Understanding". Technical Report, Computer Centre, Indian Institute of Tech., Bombay-76. 1980.
- [4] Sembugamoorthy, V. "A Paradigmatic Language Acquisition System" Ph.D. Thesis. Computer Centre, Indian Institute of Tech., Bombay-400 076. May 1978.
- [5] Narasimhan, R. Modelling Language Behaviour Vol. 10., Springer Series in Language and Communication (in print).
- [6] Schank, R.C. "Conceptual Dependency: A Theory of Natural Language Understanding". Cognitive Psychology, 1:4 (1974) 552-631.
- [7] Martin, W.A. "Knowledge Representation Panel Discussion". in Proc. IJCAI-77, Cambridge, MA, August 1977, 983-992.