A CASE STUDY IN SOFTWARE EVOLUTION: FROM ARIANE-78.4 TO ARIANE-85

presented for the seminar on MT
(August 14-16, 1985),
Hamilton, Colgate University
by
Ch.BOITET, P.GUILLAUME & M.QUEZEL-AMBRUNAZ

GETA, BP 68
Université de Grenoble
38402 Saint-Martin-d'Herès, FRANCE

ABSTRACT

ARIANE-78 has been used for years at GETA as the underlying programming environment for writing many MT systems or subsystems, 1n a set of Specialized (rule based) Languages for Linguistic Programming (SLLP). We present briefly its recent evolution, which has been prompted by the feedback from the users, and has led the implementors to a deep reshaping. In particular, the control structure of the entire environment has been parametrized to a large extent, due to the Introduction of a specialized (finite state based) language used for describing sets of possible sequences of linguistic processes ("phases"), such as structural analysis or lexical expansion.

Keywords : Specialized Programming Environment, Machine Translation, Specialized Languages for Linguistic Programming. ARIANE-78, ARIANE-85.

INTRODUCTION

Various aspects of ARIANE-78 and of applications Implemented 1n ARIANE-78 have been presented elsewhere (5.6,8). In the sixties, the CETA (former name of GETA) had already developed a large scale MT system, largely based on augmented context-free techniques, and on a specialized transformational component.

In the seventies, the researchers at GETA turned towards the use of computational models based on transducers rather than on analyzers. This led to the definition and implementation of ATEF, ROBRA (its first version by J.Chauché was ambiguously called CETA), TRANSF and SYGMOR.

In a parallel way, linguistic techniques were refined, and B.Vauquois proposed his now famous "multilevel decorated tree structures" to represent units of translation at various levels of linguistic interpretation, ranging from lexical properties to semantic and logical relations.

Already 1n the sixties, the MT system had been designed (by the third author) as a software product, complete with *a* kind of command language, to be usable (in batch mode on an IBM 7044) by an operator having no special background.

This trend continued in the seventies. Then, the interactive character of modern operating systems such as CP/CMS were put to profit, and ARIANE-78.1 was release at the beginning of 1978 as the first completely integrated programming environment for MT. This means that, on any user space, ARIANE-78 maintains two specialized data-bases, one for the texts, and one for the "lingware" (essentially grammars and dictionaries written in the SLLPs). and ensures their coherency.

As the years passed by new possibilities were included, the most noticeable being the integration of a subenvironment for human revision and the possibility to use a priority scheme for the use of lexical transfer dictionaries.

Version 4, delivered 1n 1981, was chosen to be the basis for the work of the French national CAT (Computer Aided Translation) project, and is still used for the associated linguistic developments.

In the first part of this article, we show how this "freeze" has allowed the implementors to envisage more fundamental revisions, and to specify a largely new release, called ARIANE-85. We then present the new characteristics of ARIANE-85. Part II is devoted to the user's point of view, and part III to the implementor's point of view.