

THE CHINESE UNIVERSITY LANGUAGE TRANSLATOR (CULT) - A REPORT

P.H. Nancarrow (*University of Cambridge*)

\* \* \* \*

At a meeting of the British Computer Society Natural Language Specialist Group held at King's College, London, on 13 September 1977, Professor S.-C. Loh of the Chinese University of Hong Kong spoke about the Chinese University Language Translator (CULT) with which he has been concerned for some time.

Professor Loh outlined some of the problems of obtaining high quality translated output from machine translation techniques, and explained that, whereas other workers had submitted source language texts directly to the machine process and then relied upon post-editing to resolve syntactic problems in the resultant target language, the CULT approach was to pre-edit the source language prior to machine processing and thereby achieve a high quality output which did not require post-editing.

The input of the source language is currently achieved on 80-column punched cards using Standard Telegraphic Code to identify the characters. During the preparation of the cards, which has to be done by a Chinese-literate operator, grammatical markers are introduced in a step embraced by the term 'pre-editing', and Professor Loh considered that a fairly low level of skill was required to carry out this step. However, since the markers are being added to supplement the Chinese text at exactly those points where the language itself lacks features which will be necessary to fit to English language patterns, it seems probable that a bilingual operator would achieve a better standard of performance than a speaker of Chinese only.

One severe disadvantage in the scheme currently operated concerns the handling of symbols, formulae, and equations. We are here considering scientific texts, and the solution to the problem of non-standard symbols which has been adopted is to enter corresponding blank spaces in the punched card input. The consequence of this is that the target language printout contains gaps corresponding to, but not necessarily in the same order as, the gaps in the original text. This solution therefore further evokes the services of a rather skilled proofreader who checks that manuscript additions to the printout have in fact been correctly entered. In practice the proofreader is a professionally trained scientist with language skills in both Chinese and English. This has the advantage that any gross errors of the machine translation process can be detected by the proofreader, but if CULT is largely satisfactory in this respect, the use of other techniques for handling non-standard symbols would seem to be a priority in streamlining the system for day-to-day use. This and other developments from Hong Kong will no doubt be eagerly awaited by those interested in the problems and possibilities of Machine Translation.