

From the Archives...

Fifty years of the computer and translation

John Hutchins

In March this year we mark the fiftieth anniversary of the 'birth' of machine translation, or - more precisely, the beginnings of discussion about how electronic computers might be applied to the task of translating natural languages.

There had, of course, been earlier proposals for the mechanisation of some aspect of the translation process, primarily devices for a mechanical dictionary which would substitute words of one language by words of another. Patents for such devices were issued in 1933 in France to Georges Artsrouni and in Russia to Petr Troyanskii (Hutchins 1993). In conception, it is possible to trace back the idea of mechanical dictionaries to the seventeenth century, but the mechanisms for implementation did not become reality until the twentieth.

However, these earlier suggestions were unknown to Weaver, and indeed to many other MT pioneers until the late 1950s. In any case, Weaver's inspiration was the potential of the newly developed US computers, whose awesome power gave them the popular name at the time of 'electronic brains'.

It was on 4th March 1947 that Warren Weaver, Director of the Natural Sciences Division of the Rockefeller Foundation, wrote to his mathematician friend Norbert Wiener, shortly to become famous for his writings on cybernetics. Weaver had met Wiener during the War when both were involved in military research: Weaver on ballistics, Wiener on radar and prediction theory. Weaver (1947) wrote as follows:

One thing I wanted to ask you about is this. A most serious problem, for UNESCO and for the constructive and peaceful future of the planet, is the problem of translation, as it unavoidably affects the communication between peoples. Huxley has recently told me that they are appalled by the magnitude and the importance of the translation job.

Recognizing fully, even though necessarily vaguely, the semantic difficulties because of multiple meanings, etc., I have wondered if it were unthinkable to design a computer which would translate. Even if it would translate only scientific material (where the semantic difficulties are very notably less), and even if it did produce an inelegant (but intelligible) result, it would seem to me worth while.

Also knowing nothing official about, but having guessed and inferred considerable about, powerful new mechanized methods in cryptography - methods which I believe succeed even when one does not know what language has been coded - one naturally wonders if the problem of translation could conceivably be treated as a problem in cryptography. When I look at an article in Russian, I say "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."

Have you ever thought about this? As a linguist and expert on computers, do you think it is worth thinking about?

The letter was reproduced by Weaver in his memorandum of July 1949, which effectively launched MT research in the United States -- and the memorandum itself was included in the Locke and Booth collection (Weaver 1955).

Weaver's reference to the possible analogy of cryptography was fully explicable at the time. He had himself heard of an impressive feat of decipherment from Professor Prager at Brown University. It involved a message encoded in Turkish, which was taken to a mathematician ignorant of the language of the original. The text was decoded based on letter frequencies in English, which fortunately matched those of Turkish (after simplification of letters with diacritic marks not found in English). To Weaver, the feat seemed to demonstrate that there were statistical uniformities in all languages that could be used in some way in translation. Weaver was himself a prominent mathematician and was familiar with the work of Claude Shannon on cryptography; later, Weaver collaborated with Shannon on the statistical theory of communication (Shannon and Weaver 1949).

The idea of using computers for translation had evidently occurred to Weaver some time before writing to Wiener. In his autobiography (Weaver 1970) he states: "Early in 1947, having pondered the matter for nearly two years, I started to formulate some ideas about using computers to translate..." Quite possibly he spoke about it to others before writing to Wiener. Bar-Hillel (1952) says that "as early as 1945... Dr. Warren Weaver... started thinking and talking about the possibility..." One of those he may have mentioned it to was Professor J. Desmond Bernal of Birkbeck College (University of London), a physicist much interested in numerical computation, who arranged for Andrew D. Booth to visit the United States to investigate developments in computing.

Booth's first visit was in June 1946. He met Weaver at the Rockefeller Foundation to discuss the possibility of the Foundation's funding a computer at London University. He then went to the Moore School in Pennsylvania and Princeton University to see and learn about progress on the ENIAC and EDVAC computers. In the following year, he returned for a three month study period at Princeton, funded by a grant from the Rockefeller Foundation. He met Weaver again on 6th March 1947, just two days after Weaver's letter to Wiener. It was on this occasion that Weaver mentioned to Booth the possibility of using the London computer for "non-numerical" applications -- which were more likely to attract American funding support -- and suggested in particular mechanical translation. Booth claimed later in his well-known 'historical introduction' to the collection he edited with William Locke (Booth and Locke 1955) that MT had been discussed by Weaver and himself during the 1946 meeting. However, there is no documentary evidence to support the assertion. For example, in the report of his 1946 visit which Booth submitted to the Rockefeller Foundation, he makes no reference to such an application (Booth 1946), whereas, by contrast, after his 1947 visit, his report to the Foundation in February 1948 includes the following passage: A concluding example, of possible application of the electronic computer, is that of translating from one language into another. We have considered this problem in some detail and it transpires that a machine of the type envisaged could perform this function without any modification in its design. (Booth 1948)

In fact, in a number of articles on MT, Booth states explicitly that the discussion with Weaver took place in March 1947. It is unfortunate that Booth's historical account in the Locke/Booth collection of early MT articles has led so many later writers into error on this point.

There is, in short, no doubt. Although Weaver may have spoken to Bernal and others earlier about the possibility of using computers to translate, the first corroborated and definite mention was made by Warren Weaver in his letter to Norbert Wiener and in his discussion with Andrew Booth, in the early days of March 1947.

References

- Bar-Hillel, Y. (1952): Mechanical translation: needs and possibilities. 6 pp. (Presented at MIT conference, June 1952)
- Booth, A.D. (1946): Report on visit to American laboratories. 3rd July 1946 (Rockefeller Foundation Archives)
- Booth, A.D. (1948): Proposals for the establishment of a centre for the construction and use of an electronic computer, February 12, 1948. (Rockefeller Foundation Archives)
- Booth, A.D. and Locke, W.N. (1955): Historical introduction. In: Locke and Booth (1955), 1-14.
- Hutchins, J. (1993) The first MT patents, *MT News International* 5, 14-16.
- Locke, W.N. and Booth, A.D. (1955), eds.: *Machine translation of languages: fourteen essays*. Cambridge, Mass.: M.I.T.Press.
- Shannon, C. and Weaver, W. (1949): *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Weaver, W. (1947): Letter to Norbert Wiener, 4 March 1947.
- Weaver, W. (1955): Translation. In: Locke and Booth (1955), 15-23.
- Weaver, w. (1972): *Scene of change: a lifetime in American science*. New York: Scribner.