Letters to the Editor

Pinpointing the Dates of Data Examined by the ALPAC

EDITOR:

Mr. Titus's article on "The Nebulous Future of Machine Translation," *Communications of the ACM*, Volume 10, Number 3, March 1967, deserves a comment or two, since I fear that some of the people he interviewed are not in full possession of the facts.

In the article it is stated that "The ALPAC group looked at data that was probably two years old. And that is the basic weakness of the ALPAC report." This is not true. Copies of the ALPAC report were sent to the government agencies which requested the ALPAC study on 1 September 1965. At least one—the one I can accurately date—of the samples of MT in the report was produced as late as April 1965. Therefore, the ALPAC looked at data that was, in one case at least, only four months old.

I am happy that I can shed some light on the enigma of "who was responsible for the distribution of the publicity [about the ALPAC report]." Preparation and distribution of the news release came from the NAS-NRC Office of Information, where publicity about NAS-NRC reports is usually issued.

As the ALPAC report states (p. 24), "No one can guarantee, of course, that we will not suddenly or at least quickly attain machine translation, but we feel that this is very unlikely." As Mr. Titus's article points out, ITEK and IBM have given up. The NSF washed its hands of MT with an alacrity unheard of in the annals of bureaucracy, and the article hints that even the CIA, the original supporter of MT, was glad to be rid of it.

It is barely possible that ALPAC's assessment of MT—although unpopular—might conceivably be correct.

A. HOOD ROBERTS
Former Executive Secretary, ALPAC

Further Clarification of Dates

EDITOR:

I would like to comment on the letter of Dr. A. Hood Roberts with reference to J. Titus's article "The Nebulous Future of Machine Translation," *Communications of the ACM*, Vol. 10, No. 3, March 1967.

Dr. Roberts is correct in stating that copies of the ALPAC report were sent to ALPAC sponsors on 1 September 1965. However, they were draft copies of the final report. As is generally known, the ALPAC report was published in November 1966. The NAS-NRC press release concerning this report is dated 24 November 1966. The second printed version of the ALPAC report (with one plate changed on p. 87) went on sale in February 1967.

Since no draft copy of the ALPAC report has been made available to Mr. Titus (or anybody else besides ALPAC sponsors on 1 September 1965), he is undoubtedly justified in asserting that "the ALPAC group looked at data that was probably two years old." This moderate statement requires some elaboration.

Appendix 10 of the ALPAC report ("An Experiment in Evaluating the Quality of Machine Translation," pp. 67–75) is based on the raw machine translation output produced in October-November 1962. Appendix 5 ("Machine Translation at the Foreign Technology Division, U.S. Air Force Systems Command," pp. 43–44) is based on data covering the period of June-September 1964. Appendix 15 ("Evaluation by Science Editors of JPRS and FTD Translations," pp. 102–106) is based on postedited FTD machine translations produced during the period of February-December 1964.

The Georgetown MT project was dismissed with one reminiscence about the Georgetown-IBM experiment in January 1954 (ALPAC report, p. 23) and with one short note on the attempt to "produce useful output in 1962" without postediting (*ibid.*, p. 19). The fact that the Georgetown MT system had provided routine translation service without postediting at EURATOM and ORNL since 1963 was not even mentioned in the ALPAC report.

I welcome this opportunity to shed some more light on the enigma of "who was responsible for the distribution of the publicity (about the ALPAC report)." Dr. Roberts is once again correct in stating that the publicity about NAS-NCR reports is handled by the NAS-NRC Office of Information. However, it is also undeniable that Dr. Roberts was directly and personally responsible for the accuracy of the NAS-NRC press release about the ALPAC report. The release contained at least one grave error consisting in the statement that postediting alone costs $36.18 per 1,000 words of Russian. It is regrettable that this error (and other discrepancies in the ALPAC report) have not yet been acknowledged and rectified.

I do not think that Dr. Roberts has carefully read Mr. Titus's statements about intentions of IBM and CIA. Titus has noted that the "(IBM) staff is studying more applications of language processing than just machine translation." As regards CIA, Mr. Titus has stated that "no one outside of CIA knows for sure."

It is not even "barely possible" that the ALPAC's assessment of machine translation is correct. If Bell's invention of the telephone were similarly assessed, this country would still be deprived of the telephone service.

Z. L. PANKOWICZ
105 Stanwix St.
Rome, N.Y. 13440

What Are the Most Promising Avenues of Approach to Attaining MT?

EDITOR:

I'd like to add a few remarks on the letter of Hood Robert's in relation to Titus's article.

It is clear that "acceptable machine translation" is not being used by Titus in a fashion which would seem obvious to everyone. It is of course gratifying that Oak Ridge is able to reach usable results with speed by the aid of machine; such sensible use of resources is to be applauded. It will be noted that the ALPAC report has specifically recommended (as the OSIS is now implementing) that machine-aided operations be fostered and improved by research. This is true even if "the grammatical quality leaves much to be desired," as Titus puts it. An uncharitable view could point out that 200,000 words a year of low quality output is not much to brag about, but even if the output were greatly increased in volume, speed, and quality, the fundamental problem is neither solved nor approached.

It has become clear to most linguists in the last decade that earlier linguistic theory had serious inadequacies, and views on a newer more adequate theory, while agreed on some fundamentals, are still in many ways unsettled. Most aspects of MT have been based on important portions of earlier theories. The entire status of MT itself is not at all clear in the framework of newer theory. Therefore the improvement of machine-aided operations (which could, naturally, be adapted to various and whatever stages of an ultimate acceptable MT) has little or nothing to do with attaining true MT.

In this fashion, acceptable MT as such in the senses in which it has been understood, cannot be a valid goal at present. Instead, we must foster, by whatever means possible, efforts to extend our basic understanding of linguistic theory and human language. Of course, we should exploit machines wherever we can; and that is one function of computational linguistics, which, as Titus points out, badly needs an accretion in numbers of good and well prepared scholars. If such studies ultimately point the way to an acceptable MT, among other things, well and good. But if they do, it is already clear on theoretical grounds that this will happen on a basis that is different from that used during past strivings for MT.

It is true, as both Titus and the ALPAC report mention, that we have learned substantive lessons from the failures in the search for MT. No sensible person or committee would advocate cutting off basic research on an interesting problem. But no amount of tinkering can rectify a basic theory that can be shown to be inadequate. The obvious course is to offer the broadest possible scope to research on basic theory without tying such research to one limited goal, e.g., MT.

Titus speaks of abandonment of support of MT "after only twelve brief years" as if it meant utter relinquishment of all approaches. As a member of ALPAC (though in this letter throughout I can pretend only to speak for myself), I conceived my task as one of inspecting evidence with a view to encouraging support for investigators to seek out the currently most promising avenues of approach. Whether or not they include MT would itself constitute a capital contribution.

Eric P. Hamp
University of Chicago
Chicago, Illinois 60637

On "Numerical Integration of a Function That Has a Pole"

Editor:

The paper by E. Eisner [*Comm. ACM 10*, 4 (April 1967), 239] describes a method for evaluating integrals when the integrand has a singularity outside the range of integration by determining weights which depend on the order and location of the singularity. An alternate approach described by Krylov [1] for dealing with such problems subtracts the singularity from the integrand and uses a conventional formula to evaluate the transformed integral. This approach appears to involve less work and to be directly applicable to multiple singularities or singularities which are not on the real axis. Even when experimental data is involved it should be possible to estimate the coefficient of the poles if its order and location are known.

For the example given by Eisner,

$$\int_0^X \sec^2 \pi X \, dX = \frac{4\pi^{-2}X}{1-2X} + \int_0^X \left(\sec^2 \pi X - \frac{4\pi^{-2}}{(1-2X)^2}\right) dX.$$

The midpoint rule was used to evaluate the integral on the right side. The resulting absolute error $\times$ 10^6 for several step sizes, h, were

x/h	1/30	1/60	1/120	1/240
.1	33	8	2	.5
.2	50	13	3	.8
.3	61	15	4	1.0
.4	69	17	4	1.1
$(.5 - h)$	73	19	7	9.8

This is comparable to Eisner's results.

Reference:

1. Krylov, V. I. *Approximate Calculations of Integrals.* Macmillan, New York, 1962.

William Squire
West Virginia University
Morgantown, West Virginia

Dr. Eisner's Reply

This is in reference to Squire's comments, [1] drawing attention to an approach by Kantorovich [2] that provides an interesting alternative to that described in my recent paper. [3] A proper comparison of the merits of the two methods can probably be made only after varied experience in using them both. I shall merely point out a source of inaccuracy in Kantorovich's method that has no counterpart in mine, without attempting to evaluate its importance. I shall restrict the discussion to the integration of a function $R(X)$ that has a single pole, of order n, at $X = T$.

Kantorovich takes the first $(k + 1)$ terms $(k \geq n)$ of the Taylor series about T of $[(X - T)^n R(X)]$. This enables him to split $R(X)$ into two parts, A and B, such that A is singular at $X = T$ but can be integrated analytically, while B must be integrated numerically, but has no singularity. B can therefore be integrated by conventional methods, but it must be evaluated with care, since it is a difference of two nearly-equal quantities.

As Squire points out, both my method and Kantorovich's require the order, n, and location, T, of the pole to be accurately known, but Kantorovich's requires the $(k + 1)$ coefficients in the Taylor series as well. Errors in these coefficients will appear direct in the integral of the analytical part, A, of the integrand. However, it may be more serious that such errors leave singularities in the nominally singularity-free part, B. The conventional formulae used to integrate B will therefore be inaccurate. If tabular values of $R(X)$ are to be used to estimate the Taylor coefficients by curve fitting and extrapolation, we have a procedure similar to that which underlies my method, but much less direct in use.

It would be interesting to know how important this source of error is. My guess is that Kantorovich's method may be simpler (though less automatic) when $R(X)$ is an analytical expression for which the Taylor coefficients can be found exactly, while my method is simpler and more accurate where $R(X)$ contains computed or experimental data (which was the case in the problem that gave rise to my work [4]). The potential user should not be deterred by the fairly complicated-looking formulae of my method: they are really very straightforward to program.

References:

1. Squire, W. Letter on "numerical integration of a function that has a pole." *Comm. ACM 10*, 10 (Oct. 1967) 608.
2. Kantorovich, L. V. On approximate calculation of certain types of definite integrals and other approximations of the method of removal of singularities. *Mat. Sbornik, Ser. 1. 41*, (1934), 235–245 (in Russian, with French summary); English translation deposited with the Special Libraries Translation Pool, John Crerar Library, Chicago. Described in part by V. I. Krylov, *Priblizhennoe Vychislenie Integralov* (Approximate Calculation of Integrals) (in Russian) Gos. Izd. Fiz.-Mat. Lit., Moscow, 1959, English translation by A. H. Stroud, Macmillan, New York, 1962; Chapter 11, "Increasing the Precision of Quadrature Formulas," especially Section 11.2, pp. 202–206 (English edition), "Weakening the Singularity of the Integrand." [Kantorovich's very interesting paper deals with many things I have not considered in [3], only some of which are described by Krylov. For instance, he considers in detail, multiple singularites, logarithmic singularities, and the application of his method to singular (ordinary and partial) differential equations, and integral equations.]
3. Eisner, E. Numerical integration of a function that has a pole. *Comm. ACM 10*, 4 (April 1967), 239–243.
4. Eisner, E. Complete solutions of the 'Webster' horn equation," *J. Acoust. Soc. Am. 41* (1967), 1126–1146.

E. Eisner
Bell Telephone Laboratories
Murray Hill, N. J. 07971