ware vendor, want the term "dialog box" to be translated in their manuals? Technical terms such as "dialog box" are difficult for translators because they are generally not as familiar with the subject domain as either the author of the source text or the reader of the target text. In the past, translators had to read a lot of background material in both the source and target languages until they mastered the terminology in both languages, an extremely labor-intensive process.

Parallel texts could be used to help translators overcome their lack of domain expertise by providing them with the ability to search previously translated documents for examples of potentially difficult terminology and see how they were translated in the past.

"Existing translations contain more solutions to more translation problems than any other existing resource."  

In this way, the statistical approach is producing a set of useful terminology and reuse tools. Unlike traditional MT, these tools do not attempt to compete with the human at what the human does best (translating the easy vocabulary and the easy grammar), but complement the human in areas where they know they need help (difficult vocabulary and reuse). The tools approach was proposed by Martin Kay 15 years ago in "The Proper Place of Men and Machines in Language Translation." In contrast with fully-automatic MT and largely automatic approaches such as machine-assisted translation followed by post-editing, Kay advocated the more modest goal of building tools that human translators would want to use.

It would be ironic if statistical MT ended up producing a toolbench that isn't statistical and isn't MT. But at least it isn't stone soup . . .


---

The inflexible fickleness of fashion
Sergei Nirenburg, New Mexico State Univ.

Machine translation has been a fashionable field for at least 40 years of its 50-year history. The reasons for this vary from R&D glory to commercial payoff. Over the years, researchers have used an impressive variety of methods as the basis for translation programs. The problem, however, has proved so complex that the quality of the result has not correlated significantly with the method chosen. Rather, it typically correlated with the amount of descriptive work on language that was carried out.

Of course, MT research has brought about significant side benefits. Entire scientific fields have emerged largely because of MT efforts, witness the nascent field of computational linguistics. Often, MT served as an application of choice for various workers to test and attempt to corroborate their theories of language and human thinking capacity. Characteristically, the Eurotra project's final report listed as its major success the creation of computational-linguistic infrastructure in the countries of the European Community, emphasizing the fact that no realistic MT system was built under its auspices. Many factors contributed to the lack of engineering achievement in this project, among them the relative lack of emphasis in Eurotra on actual description and system building, with preference given to designing detailed formal specifications of (largely syntactic) levels of analysis and their corresponding formalisms.

Is the Eurotra case prototypical for the entire field of MT? One problem with the field has been that the descriptive work is, frankly, rather monotonous and boring. This is why attempts were made either to make it less boring (by adding an independently motivated theoretical angle to the descriptive work), or to try to avoid it altogether.

The latter objective manifested itself in

- attempts to use AI learning techniques or more practical semiautomatic procedures for knowledge acquisition, and
- the application of statistical methods for establishing cross-linguistic correspondences in lieu of language-description work.

The former solution emerged in viewing MT as a testbed for one's favorite linguistic or computational-linguistic theories, such as the currently fashionable "principle-based" approach to syntax. MT is indeed a tempting avenue of computational inquiry into modeling human mental and language processes, and a number of approaches to NLP in AI dabbled in MT as a potential application. Knowledge-based MT is a direct offshoot of the AI tradition.

The most remarkable feature of the statistical methods in MT is that they are not at all specific to their subject matter: the same techniques applied to processing language could be and are used, for example, in the studies of the human genome.

The current R&D-oriented MT approaches, whether rule-based or statistical

---

References


or hybrid, are based on "imported" ideas. Also, the best systems on the market cannot boast much of technological or scientific advances. Instead, they rely on brawn: huge, hand-crafted dictionaries and grammars, and a plethora of specialized translation routines. All of us are curious to see how well the R&D approaches will work once sufficient resources are allocated for one or more of them to reach the status of a product. The question is: What kind of imported techniques show the most promise? The answer is not obvious and is determined by sociological (read: the vagaries of funding) as well as scientific and technological trends.

The major scientific (or methodological) trend in the field is experimenting with how well the statistics-oriented methods will advance the state of the art in MT without the need for massive, manual knowledge acquisition. The major technological trend in the field is looking for the best ways to mix the statistical and the rule-based methods. I have been an early advocate of mixing such methods at the level of their final results, an approach called multiengine MT. Other approaches seek a more involved interaction, with statistics used not only during the process of MT but also to support development of background resources—dictionaries and grammars.

The major sociological trend, at least in the US, is the emphasis on a regimen of evaluations and competitions among MT (and, more broadly, NLP) systems. This promotes rigor and discipline as well as conformity and search for local solutions, which are not necessarily the most promising ones in the long run. Approaches that show a steady improvement are rewarded. Approaches with long gestation periods are punished.

Emphasis on mixed approaches is, for nonstatisticians, a rearguard regrouping action, while for statisticians—witness the evolution of the claims and practices of the Candide IBM MT group—it is a search for any avenue for improving the rather modest results.

The knowledge- and linguistics-based methods will do well to regroup and concentrate on those tasks and situations in which statistical approaches fail to deliver. One must remember, however, the lesson of computer chess: at present, the best chess-playing systems are not terribly knowledgeable about chess strategy and tactics, but they consistently beat AI-based programs and compete on equal terms with grandmasters. The $64,000 question is: How much more complex is human translation ability compared with the human chess-playing ability? That is, for how long will there be an opportunity to study language use through MT? If statistical methods succeed, rule-based MT may go the way of the AI-based chess programs.

My opinion is that MT is too complex for the current statistical processing methods to handle, even though these methods do not aspire to building representational models of human language capacity and rely only on the input-output behavior of such models (in MT, a text and its translation). In the final analysis, the open-endedness of language will become the stumbling block for these methods conceptually, just as, logistically, the chronic shortage of resources (linguistic corpora) may precipitate the swing of the pendulum of MT R&D fashion back to the mentalist camp from its current behaviorist direction.

How long will this take? If history is any guide, such swings come roughly every 30 years: Mentalism was in scientific ascendency between 1960 and 1990, while behaviorism reigned, at least in the US, for about 30 years prior to that. Of course, we cannot be certain that we are witnessing this pendulum swing and not some other, unconnected development. Time will show.

A more intriguing thought is that, just possibly, the rule-based/corpus-based dichotomy is not as important as we think. Maybe the real problem of MT as technology is that AI researchers do not generally understand how difficult the problem actually is. The confident claims, made by newcomers to MT (including me some 15 years ago), help stoke the high expectations of getting the desired result with a modest expenditure. At the current level of MT R&D, either we should lower the expectations or significantly extend the time scale for getting results. For best results, we might need to fund a language description effort of truly Tower of Babel proportions.

Deepening wisdom or compromised principles?—the hybridization of statistical and symbolic MT systems

Eduard H. Hovy, USC/ISI

In late 1991, at the outset of ARPA’s MT program in the US, the Statistics Wars were getting into full swing. On the one side stood the Candide system (built at IBM, New York), untouched by any taint of symbolic or linguistic methodology or knowledge; on the other, in statisticless purity, stood the Interlingual systems out of which Pangloss would be built (by a collaboration of the Center for Machine Translation at Carnegie Mellon University, the Computing Research Laboratory at NMSU, and the Information Sciences Institute of the University of Southern California). A third system, Lingstar, refused to enter the Wars, opting instead to mix and match whatever techniques would help in the next evaluation.

Five years and four evaluations later, the picture looks completely different. All three systems, as well as a new system called Japangloss, built at USC/ISI, ended up employing methods from both statistics and linguistics. Although fundamental differences remain, it is informative for all future MT systems (and, in fact, for other NLP systems as well) to identify what parts of the systems tend toward linguistics, what parts toward statistics, and why this should be so.

Hybridization: natural and inevitable

Departing from your principles is hard. The frequency and volume with which the principles of statistical and symbolic MT were repeated during the Statistics Wars