

PANEL: IS MT RESEARCH DOING ANY GOOD?

Research can be, and often is, conducted purely for its own sake. In this case, lacking external judgment, the criteria for what constitutes good research can be rather shaky, and the research field may go the way of alchemy or astrology. On the other hand, research usually requires time for pure experimentation before it can deliver results of practical import. What is the position with current MT research? Is it useful to pursue pure research instead of using the time and money to develop better MT tools and resources for the public domain, which companies could then pick up and deploy? What kind of research is worth pursuing? How does one know?

Moderator: Eduard Hovy, USC Information Sciences Institute

Panelists: Kenneth Church, AT&T Bell Laboratories
Bonnie Dorr, University of Maryland
Sergei Nirenburg, CRL, New Mexico State University
Bernard Scott, Logos Corp.
Virginia Teller, CUNY

How to Decide if Research Delivers

Eduard Hovy, USC Information Sciences Institute

Fully automatic, high-quality, unrestricted-domain translation will remain as elusive for machines as it is for humans — creating a superhuman machine when it comes to language in general is tantamount to creating a machine of superhuman general intelligence, and neither we nor our great-great-grandchildren are going to see this. No amount of research will do it! But careful research in well-motivated directions cannot but improve the current MT situation; surely no-one is going to claim that we know everything already. Thanks to research on representation techniques, statistical NLP and the like, systems coverage and results comparable to the Grand Old Ones can be built in a fraction of the time and under modern software principles.

The question is: how can we identify which new directions are ripe for being addressed, and how can we ensure that the research stays on track to identify productive new methods (or blind alleys) with some accuracy?

Although research is by its nature unpredictable, we can lay out some general considerations:

- the tradeoff curve between the ambitiousness of the research and the size of the project
- the nature of mid-course evaluations and the amount of course alteration they may cause
- the amount of additional non-research material required before the research is judged (that is, for example, is it enough to develop a wonderful new grammar or should the research also integrate it with a realistically sized lexicon, date and name recognizers, etc.?)
- the value of pilot projects and small-scale demonstrations

Inevitably, much of the discussion will focus on evaluation. But there's more to this question. I will argue that we can develop a "science of MT science", or at least a set of guidelines, to help the MT researcher, production system manager, and funder decide when and how to embark on a new research path.

Four Questions and Answers

Bonnie J. Dorr, University of Maryland

Why is it important to do research?

Because we seek a theory of MT.

Why do we care about a theory of MT?

Because we seek a common ground: a set of principles of universal validity. For example, at Maryland, we seek to define a theory of the interlingua. While many interlingual representations have been proposed, no one has proposed a fundamental theory underlying these representations, i.e., a set of basic elements and operations that can be used to specify an interlingua, prove properties about it, and test it in a working system.

How do we get to a common ground?

We need to determine the "constants" in MT, i.e., those aspects that must necessarily be addressed in order for successful translation to take place. These "constants" can then be the focus of our future debates and panels, i.e., these are the hypotheses to be tested and confirmed or refuted. Once the "constants" are factored out, we are left with the residual aspects (i.e., those components that must be transferred during translation, but which aren't inherent in language in general).

What will a theory of MT buy us?

1. It's hard to evaluate MT when we don't really know what it's about. A theory of MT would tell us what it's about, or at least head in that direction.
2. It would give us a common starting point from which to work from on an international basis.
3. It would provide a means of judging whether we are making progress in our field.

The Logos View

Budd Scott, Logos Corporation, Mt. Arlington, NJ

Thirty years of rationalist-based research and development has not yielded systems more powerful than those empirically based efforts rejected 30 years ago. The standard assumption that models demonstrating the effectiveness of formal grammars and their grammatical formalisms with respect to lists of sentences only need to be scaled up to achieve working versions has never been borne out. Eurotra is our latest and most dramatic (and most costly) example of that failure. Empiricists for their part do not regard language as a formal problem awaiting formal solutions. We at Logos discovered very early in our 25-year-long effort in learning how to do MT that you succeed in MT in the measure that you learn to cope with ambiguity and complexity. Efforts at handling ambiguity (typically via more information) lead to complexity and complexity is the ultimate barrier that formalist models cannot scale over. This tells us that MT is fundamentally a decoding problem (semantico-syntactic analysis) which in turn explains why generative grammar and its derivatives are not likely ever to result in industrial-strength MT. The pendulum is now swinging to the empiricists' side but problems remain here as well. If formal grammar sought to find a set of rules that could account for all sentences of a language, statistically-based MT, for its

part, seems able only to build systems potentially adequate for a fixed set sentences (corpus). The inability to generalize tells us that we still need linguistics, and a computational paradigm that will allow for handling generality as well as massive amounts of detail. We at Logos believe the place to look for this is in hybrid symbolic/connectionist paradigms. We agree with Prof. Nagao's observation that there appears to be a congenital problem with US research. US researchers like to build small models demonstrating some theoretical point or other and then give a paper on it. The Japanese researcher, in contrast, considers it an honor to work on a long-term large-scale system project. Our experience at Logos reinforces this observation that it is only by working inductively on large scale systems, over time, with an open mind, that one learns to do MT. Sponsors of research would do well to consider the sociological factors underlying the kind of research that appears to be needed in this field to get anywhere.

Big Bucks or Boutique Research?

Virginia Teller, Hunter College CUNY, New York

Broadly speaking, MT research can be divided into two categories: big bucks and boutique. Big bucks research is generally driven by specific requests for proposals from government agencies or the private sector and tends to be fueled by relatively generous funding from the Department of Defense or corporate contracts. Such projects may be staffed by teams of three or more scientists whose long term goal is to build a complete MT system but who may be faced with a series of short term subgoals for evaluation purposes en route.

Boutique MT research, differs sharply in all respects from the big bucks model. Driven primarily by intellectual curiosity and fueled by relatively modest grants from NSF and other similar sources, boutique projects are often carried out by one academic researcher, perhaps with the assistance of a graduate student, and typically focus on a specific aspect or a single component of an MT system. Since results are regularly reported in conference proceedings, the deadlines imposed by each year's program committees may appear to dictate research goals.

Which, if either, of these two approaches to MT research is more likely to produce lasting benefits to the MT community will be the focus of my panel presentation. On one hand, big bucks research clearly has a greater chance of producing a working system but the system may be prone to errors due to shortcuts taken and brute force methods used under pressure of deadlines. Boutique research, on the other hand, is unlikely to address the immediate needs of potential users, but the emphasis on basic research offers the promise of breakthroughs and the discovery of innovative solutions to the host of perennial problems that continue to plague the field.