MT SUMMIT PANEL :

"Have We Found the Holy Grail?"

grail, n. (definition 2): "the object of an extended or difficult quest"

At the TMI-92 Conference in Montreal, Robert Mercer, one of the leading researchers on the IBM Candide project, provocatively asserted that "rationalist methods in MT will be on the scrapheap five years from now." It turns out that Mercer was wrong, at least in his timeframe – but perhaps not quite as wrong as some people might have wished. Although Candide performed surprisingly well in the DARPA competition organized in the mid 1990's, it did not actually surpass SYSTRAN. At this year's NIST competition, on the other hand, statistical MT (SMT) systems similar to Candide did outperform all the participating commercial off-the-shelf systems, according to the NIST score. While the pertinence of automated scoring methods such as NIST's remains controversial, the maturation of SMT is undeniable. This maturation has not gone unnoticed in the mainstream media. A story that appeared in the New York Times on August 7 quoted a well-known expert in SMT (Kevin Knight) as claiming that "the progress and accuracy of statistical machine translation had recently surpassed that of the traditional machine translation programs used by Web sites like Yahoo and BabelFish." The Times entitled the article "From Uzbek to Klingon, the Machine Cracks the Code."

Many lay people would probably be wont to interpret this headline to mean that the new data-driven methods in MT have finally succeeded in cracking, not just the code, but the entire problem of automated translation. Of course, we as experts know better…

This was the theme of the panel discussion that closed the MT Summit conference in New Orleans: What exactly is the significance of the results of the recent NIST competition? The following are some of the questions that the invited panelists were asked to comment on before the discussion was opened up to the floor:

1. Have we found the ultimate solution to MT's long quest? If not, is the Holy Grail just around the corner?

2. Will progress in data-driven MT continue unabated? Or is there an inherent ceiling on MT quality that will resist even the most sophisticated data-driven methods?

3. Has the data-driven paradigm been able to model information that was not present in rule-based systems? Or has it 'simply' been able to model the same kind of information more thoroughly and efficiently?

4. Was the metric used to rank participating systems in the NIST competition fair, or was it somehow biased in favor of data-driven systems?

5. Even if the evaluation metric used at NIST was somewhat biased, can we still assume that SMT has indeed surpassed traditional rule-based systems? And if so, at what exactly?

6. Are there niche applications for which the new data-driven techniques are particularly well suited?

7. Is there a danger that SMT's recent success may lead the public – and worse yet, the funding agencies – to believe that the MT problem has finally been solved, and so to reduce the level of R&D grants to our field? If so, what we do to combat this misperception?

8. Would the results of the NIST competition have been different if the languages involved had been English and French? If so, why?

9. In previous debates on this question (e.g. at TMI-92), many people concluded that hybrid systems were the way of the future. What role do rule-based components play in today's leading data-driven systems, and what are the prospects for their future contribution?

Elliott Macklovitch

MT Summit Program Chair