

CRANFIELD CONFERENCE AFTER-DINNER SPEECH

I. D. K. Kelly

In machine translation we are looking for "Meaning In, Meaning Out". If poetry is what is lost in translation by ordinary means, and prose is that which is lost in adaption for television, what is it that is lost in Machine Translation? All too often the answer is "sense", but we're getting better at that. I would like to see MT systems measured against more exciting criteria than merely preserving the meaning of sentences about plugged drains - drain plugs. I would like to see MT preserving style and humour. I would like - and I am not entirely flippant about this - to see the Machine Translation of jokes.

New theories and processes in Artificial Intelligence are often tried out on the game of chess. This is not because AI workers see the game of chess as the ultimate aim of what they are doing, but because chess constitutes, for many purposes, an isolatable and controlled universe with objective measures of success.

Why jokes for MT?

Well, firstly because of their linguistic complexity. (Though remember, with Aquinas, that not everything that is more difficult is more meritorious).

The smallest unit of translation would have to be the whole joke itself, of whatever length, which may include dialogue, incomplete sentence fragments, linguistic caricatures in deviant orthography of racial, social or political stereotypes, and so on. (That means, in jokes the "other" guy talks odd). There are even puns, and very long word-plays which the best human translators may spend hours chewing a pencil over.

There are simple positions, and chess problems.

Secondly, because jokes cover a very wide universe of discourse. There are the wonderful tales of the Mulla Nasrudin which encapsulate philosophical thought and moral tuition in the form of jokes; there are children's surrealist jokes of Zen incongruity ("What's yellow and dangerous? Shark-infested custard!"); there are

scatological jokes which no-one claims to tell or hear in polite society, but which everybody none-the-less knows. There are jokes which bring in topicality and locality for their interpretation, and so on. But you are unlikely to need vast term-banks: jokes about "ternary incremental representation" or "loop invariance within irrotational fields" are unknown. Well, to me, at any rate. There may be as many possible games of chess as there are atoms in the universe, but there are only six kinds of chess piece.

Thirdly, I think it would be even more fun to work on. Much of what we translate is - let's face it - dull. Stuff you find it hard to put your heart into: indeed, the kind of stuff you suspect no-one's heart has ever been in. As well as fascinating challenges in informatics and linguistics, why should we not, as we research, have interesting inputs and outputs as well? AI researchers sometimes enjoy a game of chess, though that might not have been the primary aim of the research.

And finally, because humour touches something so fundamental and deep-rooted in our own, human, psychological make-up that the study of it can only serve to humanize the machines, and open up for them an understanding of what we are - which we, of course, will first have to discover. An objective measure of the successful translation of jokes would have to include laughter. Dr. Howard Pollio, a psychologist at the University of Tennessee, has been quoted as saying:

The stimulus of a laugh is an intellectual event, yet it quickly goes on to block all else. There are only two other phenomena that so completely take over your awareness - the orgasm, and the sneeze.

Besides that, the simple elation of winning a game of chess is insignificant indeed.

Hence I would like to institute a new award. The award itself will have no intrinsic value - merely kudos. It will be awarded in several categories:

- * for the best new computer joke,
- * for the most elegant machine translation of a joke,
- * for the most convincing evidence of an

automated appreciation of a joke, and -
of course -
* for the funniest example of a machine
mis-translation.

And there will be special awards for the
introduction of the funniest corpus of test
material.

You have heard of the Oscar; you have heard
of the Emmy; I offer you the Roderick.

The Roderick is named after the small
robotic hero of a pair of books by that
brilliant author John Sladek. In the books
- "Roderick" and "Roderick at Random" - we
follow a baby robot from its first
awakenings, through its growing pains and
education up to its adulthood - or should I
say "machinehood"? John Sladek has
obviously done his homework, and amongst
the very funny and percipient prose there
are lots of "in" jokes on AI and
computational linguistics. Sladek suggests
that - maybe - robotics is the next step in
our own, human, evolution; and that
research into robotics is fuelled by our
own impulse to procreate images of
ourselves. Those of you who do not receive
any of the Rodericks on offer tonight would
be well advised to borrow and read a copy
soon.

On this occasion the Rodericks will be
awarded without ceremony, and purely at
whim as we come across sufficiently
entertaining examples. But next year - or
whenever we decide to hold the next one of
these conferences - we will be able to
award them in their correct categories.
Contributions, please, to the Editor of the
Newsletter, Ulla Magnusson Murray.

I hope we shall be able to hold a series of
these conferences. When "Translating and
the Computer" started, the growth in
interest could hardly have been
anticipated. I would hope that following
this conference - which I like to refer to
as "MT84" - there will be an MT85, MT86,
and so on, to complement "Translating and
the Computer" from the computing side. And
who better to host such conferences than
the British Computer Society through its
specialist groups, specifically the Natural
Language Translation Specialist Group.

We know that Machine Translation is going
to grow in importance. As we are all aware,
MT is specifically one of the ingredients

of the Japanese 5th Generation. Public awareness of the reality of MT needs to be fostered. A lot of wrong impressions still abound: it's all very well for me to inaugurate a prize for a humorous example of a machine mis-translation, but it is quite another thing for an otherwise excellently informed publication like "New Scientist" to quote the old joke about "out of sight: out of mind" being mis-translated as "invisible idiot" as if that were the best that MT can do. The message that must be got across is: there is a language crisis now, caused both by the information explosion and the increasing insistence by nations and peoples to be communicated with in their own languages, the increasing awareness that just as you should respect other people's religion and should not - if you are civilized - expect them arbitrarily to conform with yours - so you should respect the whole of their culture, of which language is a vital part.