

Re: Machine translation

Source: <http://www.derkeiler.com/Newsgroups/sci.crypt/2003-07/2105.html>

From: Mok-Kong Shen (mok-kong.shen_at_t-online.de)

Date: 07/30/03

Date: Wed, 30 Jul 2003 09:48:21 +0200

"Douglas A. Gwyn" schrieb:

- >
- > *Mok-Kong Shen wrote:*
- > > *"Give me enough parallel data, and you can have a*
- > > *translation system for any two languages in a matter*
- > > *of hours," said Dr. Och, ...*
- >
- > *Of course this doesn't result in high-quality machine*
- > *translation of natural languages. MT has been an*
- > *active field of research since the 1950s, and still*
- > *has a long way to go, although there has been real*
- > *progress. For that entire time, there have been*
- > *people making statements to the effect that they knew*
- > *how to do it and that high quality MT would be attained*
- > *any day now..*
- >
- > *One significant problem is that high-quality translation*
- > *often requires *conceptual* understanding and rephrasing,*
- > *something we are a long way from accomplishing very well*
- > *with computers.*

Here is another quotation from the news:

Och's translations proved best in the 2003 head-to-head tests against 7 Arabic systems (5 research and 2 commercial-off-the-shelf products) and 14 Chinese systems (9 research and 5 off-the-shelf). In the previous, 2002 evaluations they had proved similarly superior.

I think what is remarkable is that (if I correctly understand the news) there are no linguistic aids in the sense of grammars of the source and target language etc. programmed into it and that everything depends on statistics. The machine has to learn all by itself (possibly quite similar to the learning by a neural network). BTW, this would mean, if one interprets (and certainly highly exaggerates) in a crypto context, that a machine could all

sci.crypt: Re: Machine translation

by itself crack a cipher, if only given sufficient amount
of plaintext ciphertext pairs.

M. K. Shen