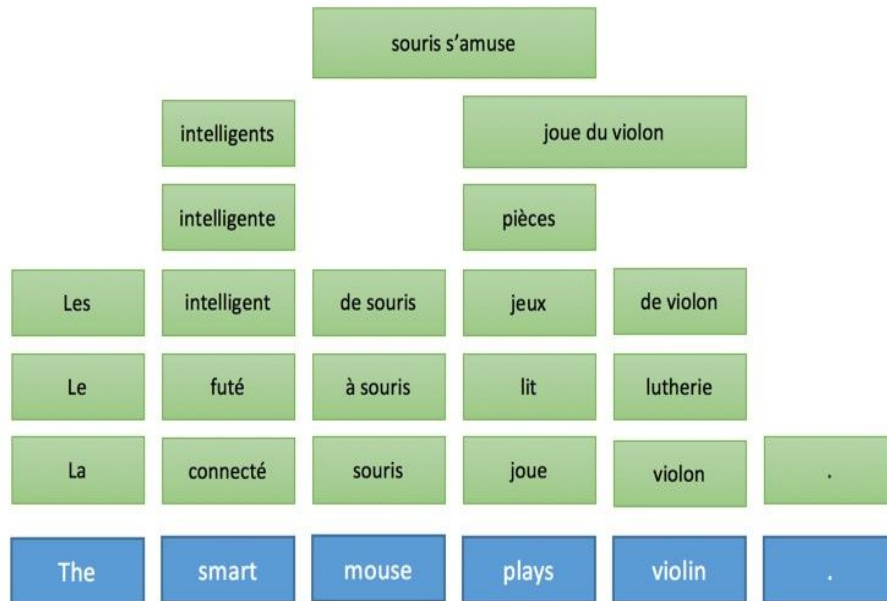# Deep Learning in Machine Translation

Chaoran Lin

The old approach: Phrase-based machine translation (PBMT)

- Work with a huge data set of existing translations
- Analyzed text word by word (or phrase)
- Associates corresponding words and phrases in different languages with each other
- Problems?

Deep learning

- Connected layers of computational units in artificial neural network
- Combined with enormous data sets
- Has beaten other machine approaches in:
  - Image recognition
  - Game-playing
- Now applied towards language translation
 - Google Neural Machine Translation system (NMT)
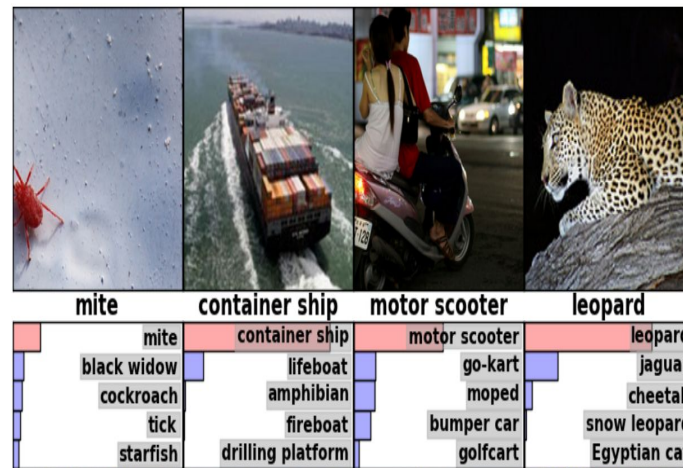
Google Neural Machine Translation (GNMT)

- Takes entire input sentences as a unit
- Looks at word segments in relation to one another
- Inside neural network, segments combine to form entire semantic context
- Generates most relevant translation

| Input sentence:	Translation (PBMT):	Translation (GNMT):	Translation (human):
李克強此行將啟動中加總理年度對話機制，與加拿大總理杜魯多舉行兩國總理首次年度對話。	Li Keqiang premier added this line to start the annual dialogue mechanism with the Canadian Prime Minister Trudeau two prime ministers held its first annual session.	Li Keqiang will start the annual dialogue mechanism with Prime Minister Trudeau of Canada and hold the first annual dialogue between the two premiers.	Li Keqiang will initiate the annual dialogue mechanism between premiers of China and Canada during this visit, and hold the first annual dialogue with Premier Trudeau of Canada.

Example
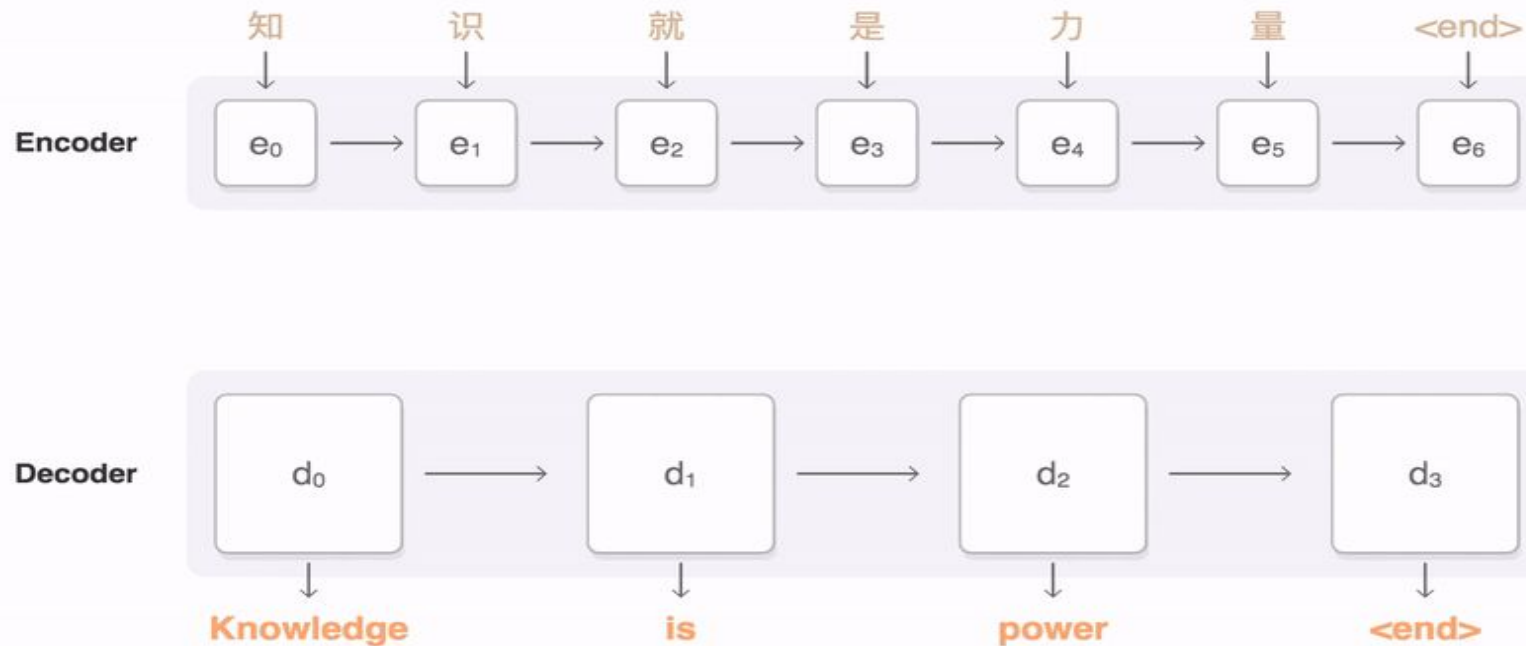
Results & Limitations

- Results
  - Reduces translation errors by more than **55%-85%**
 - Translation quality close to that of human translations
- Limitations
  - Test data is too simple/well-crafted sentences
  - Mistranslation of proper names/rare terms
  - Some difficult translation pairs still lag relatively (e.g. Chinese to English)
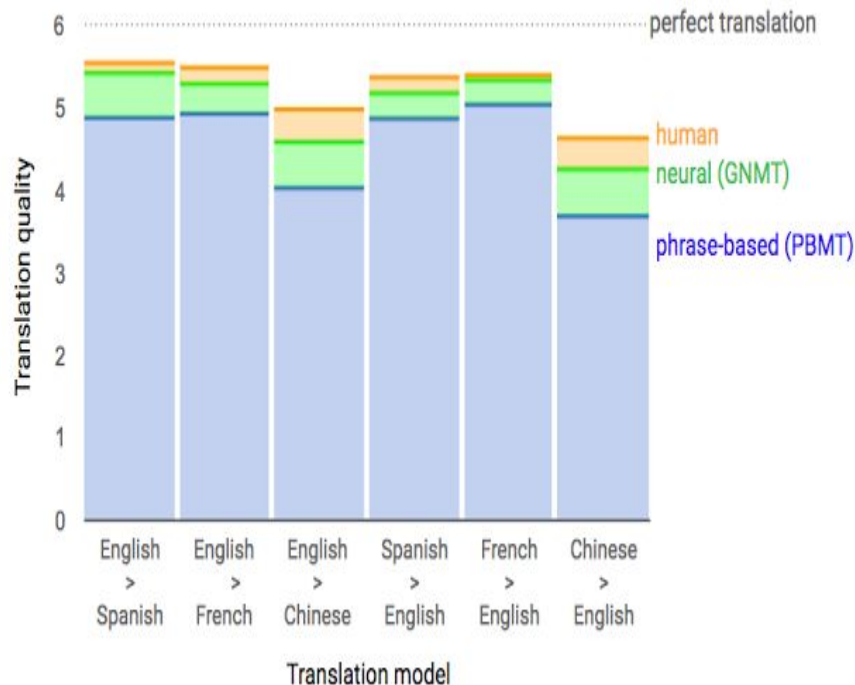
What's next

- Google Translation and other translation services
  - Deep learning: more input, more accuracy
- Integration with intelligent personal assistant tools
  - Google Assistant
- Integration with computer-assisted learning software
  - Rosetta Stone, Duolingo...
- Connecting the world's languages and people
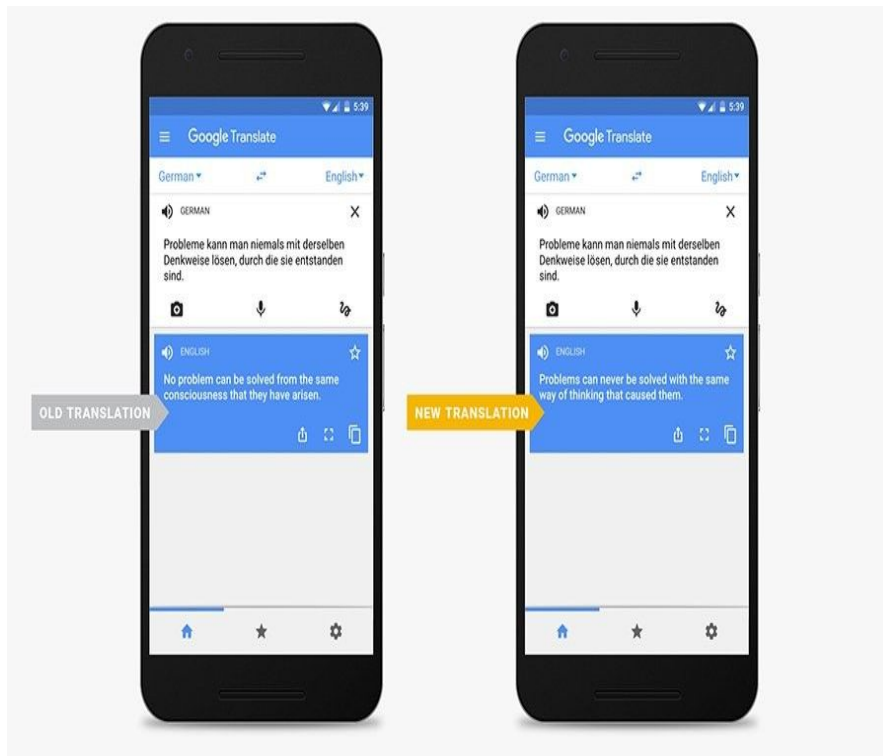
References

- Castelvecchi, Davide. "Deep learning boosts Google Translate tool." Nature.com. Macmillan Publishers, 27 Sep. 2016. <http://www.nature.com/news/deep-learning-boosts-google-translate-tool-1.20696>
- Jones, Nicola. "Computer Science: The Learning Machines." Nature.com. Macmillan Publishers, 08 Jan. 2014. Web. 01 Dec. 2016. <http://www.nature.com/news/computer-science-the-learning-machines-1.14481>
- Chouard, Tanguy. "The Go Files: AI Computer Wraps up 4-1 Victory against Human Champion." Nature.com. Macmillan Publishers, 15 Mar. 2016. Web. 30 Nov. 2016. <http://www.nature.com/news/the-go-files-ai-computer-wraps-up-4-1-victory-against-human-champion-1.19575>
- Le, Quoc V., and Mike Schuster. "A Neural Network for Machine Translation, at Production Scale." Google Research Blog. Google, Inc., 27 Sept. 2016. Web. 30 Nov. 2016. <https://research.googleblog.com/2016/09/a-neural-network-for-machine.html>