CAN COMPUTERS LEARN LANGUAGES?

The Man
I'm a theoretical particle physicist by training (think "Sheldon Cooper"), an oil executive by experience, and a professor of business IT by profession. I was one of the original developers of the World-Wide Web. As a postgraduate, I wrote papers on neural nets (applied to particle physics theory) - incidentally, neural nets are the dominant method to teach computers language today. I am involved in a Master program on Business Intelligence, which really stands for machine learning and various data and process mining techniques. So I have a fairly good overview of what's cooking in this area right now - though from an (applied) IT perspective. Though I am not a linguist, I speak several languages. I've also dabbled in fiction writing (English only). At home (in Berlin, Germany), I am immersed in an English-speaking environment since my wife, just like my dear sister, is a US citizen (in fact, since last month, my daughter, too - leaving me as the only one without the privilege and benefits of dual citizenship *sigh*).

The Mess
This field is very much in flux - in the area of "Natural Language Processing" (NLP), which is the dominant discipline in which computers are being instructed to "learn languages", lots is happening right now (Source e.g.: Deep or Shallow, NLP is Breaking out, 2016, or for an overview, see this blog, Machine Learning for Natural Language Processing).

PEOPLE! Our time is too short, our canvas too small, and our knowledge is too heterogeneous to discuss the details of NLP theory or practice with any gain. Instead, possibly more interesting for a discussion is the question how computers learn anything today, and how. If this is understandable, we can then compare the facts with the ways in which humans acquire language. The field of inquiry we're moving in are the sciences of complexity, or the solving of complex problems. I happen to know something about this, so you're in luck today.

Structurally, I'm generally not looking for a collection of facts (remember that I am a teacher, too) but instead for a **story** that I can tell and that you will then also be able to tell.

The Map
I started exactly 1 week before this event to prepare myself and these are my notes, which I scribble continuously as I traverse articles and books & think about the topic at hand. (This is unusual for me: normally, I'd like at least another 6 months or so to really get my head around something.) I also created this **mind-map** (PDF copy without links) for our session as something to look at while I deliver my initial remarks (for this interactive mind-map with links, you need to download xmind.net)

I wish to address the following three aspects to the question:

- Players (Humans vs Machines)
- Problems (Simple vs Complex)
- Purpose (Applications today and tomorrow)

1) THE PLAYERS - HUMANS VS MACHINES

Learning
One problem here is already the use of the word "learning" for two very different entities: humans are part of the **natural** world, and (leaving out issues of "intelligent design"), not artificial. Computers on the other hand, are **artefacts** designed by humans. The difference can also be expressed as the difference between the "**necessary**" (natural, factual, without human interference) and the "**contingent**" (subject to chance, open to design, controllable).

Human Nature
This distinction is a philosophical powder-keg! It goes back to the question of the nature of humans as machines who can speak (materialist view) or special beings, who e.g. have an immortal soul, desire knowledge and become political (Plato's view), or who were made in the image of God with all that entails, like for example being "fallen", as in having to struggle with temptation and evil (Christian view).

Artificiality
Herbert Simon: "Artificiality is interesting principally when it concerns complex systems that live in complex environments." (Source: THE SCIENCES OF THE ARTIFICIAL - an important classic: such sciences include psychology, economics, computer science etc.)- Both language acquisition and use quali

2) THE PROBLEMS - SIMPLE VS COMPLEX

Infinite Monkey Theorem
"Given enough time, a monkey randomly striking keys on a typewriter will end up banging out a copy of *Hamlet*.

Crazy as it seems, the infinite monkey theorem can be proved using basic probability (the trick is having either an infinite number of monkeys or an infinite amount of time, or both). What you could not do, of course, was experimentally verify the monkey theorem.

But that was before cheap supercomputers."

Mystery of language evolution
What about "Kingdom of Speech" by Tom Wolfe? Mentions admission by leading linguists (including Noam Chomsky) that they (still)

don't understand language at all:

"In the last 40 years, there has been an explosion of research on [the] problem [of language evolution] as well as a sense that considerable progress has been made. We argue instead that the richness of ideas is accompanied by a poverty of evidence, with essentially **no explanation of how and why our linguistic computations and representations evolved.**" (My emphasis)
Source: The Mystery of Language Evolution (Front. Psychol. 7 May 2014)

Reviews: negative review in the Guardian, positive review in the National Review

EXCURSION: MATHEMATICAL MODELLING
If you cannot express it mathematically, you cannot model it. This is called designing an algorithm. It means explaining step by step how to execute a certain procedure. If you cannot model it, you cannot compute it (with or without a computer, irrespective of the computer's processing power).

The art of devising algorithms is ancient: I just read in a book (Stepanov and Rose, 2015) that the earliest known algorithm goes back to Egypt almost 4000 years ago. It says how to multiplication of natural numbers (in a culture that did not know the number 0). It is mastered by young school children today but in Ancient Egypt, there was a class of priests whose secret it was.

In words, it looks like this:
"Add something to itself a number of times."

In abstract algebra, the algorithm looks like this:
alg0r1thN_
1a = a
(n+1)a = na + a

In the high level programming language C++, it looks like this:

```
int multiply0(int n, int a) {
    if (n == 1) return a;
    return multiply0(n-1,a) + a
}
```

Though we've had thousands of years to improve on algorithms, the design of algorithms was only formalised in the last century - and most algorithms in use were discovered since then. Partly because of the relationship between algorithms as procedures and computers as artefacts that can execute the algorithms faster and faster, using more and more data.

Modern example (Sedgewick): network connectivity - what is the shortest path between two phones in a network.

Niklaus Wirth: Algorithms + Data structures = Programs (data + models = information, Bakopoulos, 1985).

Avi Widgerson: "Algorithms [are] a common language for nature, human and computer."

Simon (conclusion to ch.4 "The Psychology of Thinking" in "The Sciences of the Artificial"):

"Human beings viewed as behaving systems, are quite simple. The apparent complexity of our behaviour over time is largely a reflection of the complexity of the environment in which we find ourselves."

My summary:
There are many attempts, methodological, philosophical, scientific, artistic etc. to simulate the complexity of human language and human interaction. Any real breakthrough will have to use mathematical models that imply an understanding of the evolution of language (which, so far, is not forthcoming).

3) THE PURPOSE - APPLICATION TO THE PROBLEM OF LANGUAGE LEARNING

10 Applications (source: Davydova, 2017)

1) Text classification/categorisation(e.g. language="English")
2) Named Entity Recognition (e.g. "London"="City")
3) Part-of-Speech Tagging (e.g. "Language"="noun")
4) Semantic Parsing (e.g. answering natural language questions)
5) Paraphrase Detection (e.g. different ways to ask same question)
6) Language Generation (e.g. document summary generation)
7) Machine Translation (between different natural languages)
8) Speech Recognition (e.g. Alexa, Siri in the home)
9) Character Recognition (e.g. reMarkable)
10) Spell Checking (e.g. Grammarly)

If computers are so stupid, how come they excel at Chess, at Go and at Jeopardy?
The short explanation of all these successes in these (and other) games (source - enjoyable non-scientific read with many examples) is that the settings were not particularly complex. Though Chess and Go especially were considered for a long time to be hallmarks of human intelligence, it turns out that they were easy to master by a machine that otherwise is as dumb as a toaster. This is also sometimes expressed as the difference between Artificial General Intelligence (AGI, human-like), and Artificial Narrow

Intelligence (ANI, which is what we have right now).

NLP is breaking out
(Source: Deep or Shallow, NLP is Breaking out, 2016)

Natural Language Understanding is not NLP
A critical voice:
"None of the "big" [data sets] and none of the "deep" [learning models] can still understand a simple sentence that a 4-year-old can easily utter or effortlessly comprehend."

(Source: Communications of the ACM, 16-Sep-2018)

Chatbots
They are surprisingly bad at what they do (given that they only have to deal with digital text!). ELIZA is an example that you know already - but have you tried this program? It's available online here with some additional information & including a Javascript version of the program, too. It is based on a simple dictionary (the "semantic parsing" method of old).

Here are some examples (2018):
https://clevertap.com/blog/machine-learning-marketing/

Machine Translation
Basic Problems (that remain):
1) Data: they have to be "good" = not noisy, parallel corpora (like the Rosetta stone) & accurate. E.g. European parliament proceedings, movies with subtitles. Open parallel corpus (translated texts on the web). Sub-problems: domain-specificity (movie subtitles for scientific translation? Knowledge domain transfer is poor).

2) Evaluation: there have to be procedures. The measures to assess accuracy are of a statistical nature. However, for really complex problems, the principal ability of "neural nets" (which are not really "neural", just called that way) to be trained up to human abilities is not known (which does not mean that it is impossible, but it also is not certain).

History: MT discarded in the mid-60s. Next stop: 2006 (Google Translate & Moses training system); Neural Machine translation discussed since 2013 - in production since 2016 (.

Source: Coursera online course on NLP (video, 15')
[@Nat'l research univ higher school of economics, Moscow!]-
incidentally, this video shows the limitations and the beauty of the technical approach (critiqued in the 2018 article above).

SUMMARY
Natural Language Processing (NLP) has made enormous progress

through the application of neural nets in an extremely short time. With expected improvements in data processing and computing performance (e.g. through quantum computing and distributed methods like blockchain). But NLP is not the same as understanding natural language.

FINAL QUESTION:
So, does this mean that there is no point anymore to learn foreign languages, train as a translator etc.?

Earlier Notes on the basic texts:
(How do we acquire language, ch.3)

Second language learning vs first language learning

1) First and second language acquisition are different tasks – "the first language serves as the model, and errors often result from taking words from the target language and stringing them together by applying rules from the first language."

2) "Second language learning proceeds more quickly with people who have a high proficiency in their first language."

3) "Self-confidence, motivation, good self-image, and low anxiety are traits that improve facility in a second language, but none of these traits is important to first language acquisition."

4) "The complexity of the input affects second language learning but not first language acquisition."

5) "Practice is important for second language learning but not so much for first language acquisition […] but second language learning proceeds more quickly if the target language is used as the medium of instruction."

About Translation

- "Traduttori, traditori" (Translators, traitors)
- Intralingual translation: "explanation through rewording within a language"
- Intersemiotic translation: "interpretation of one medium in another"
- Interlingual translation: "interpreting from one human language into another human language."
    - Translating legal documents – strict adherence to the informational message
    - Poetry "surface" translation (sound rendering)
    - Word-by-word poetry translation (holistic)

- o Rhyme, tone & melody (musical aspects of the language), formality matter

Personal observation: reading this about translation brings back (bad) memories of various translators I've come across. What always impressed me about these people: how they value language over meaning. I remember a virtual exchange with one translator who didn't see what was so great about the Brothers Karamasov by Dostoyevksy. She completely missed out on the spiritual dimension of the story, which, in my view, constituted a large part of the novel's value. When I said so, she replied "Der Wert eines literarischen Werkes hängt immer an der Sprache." (The value of any work of literature always depends on the language.) There you have it. In fact, while I can appreciate language-conscious translations, even a bad translation of Dostoyevksy cannot destroy its impact and its value – because ultimately it does not depend on language. Except in a very generic way: of course, it does make a difference that D. was Russian and that the novel was written in Russian, but could it only have been in Russian? Can it only be truly appreciated and unfold its value in Russian? I don't think so – probably worth debating this point!

- • Dutch-to-English nursery rhyme translation (p. 42): for a Dutch speaker (or, for that matter, for a German speaker) the result sounds really poor and fails to transport the naivité of the original. Instead, the translation sounds stilted (especially the 4th line)
- • P. 45: "Ironically, the accuracy of a biblical translation is particularly thorny since the identity of the original manuscript is constantly under debate." *Is this true? What about Luther's first translation of the Bible into a modern (non-Latin) language, namely German – what about its undeniable impact on world history (namely, the reformation, and the secularization?*
- • "Before we can do a holistic translation of a text, we have to analyze it [and] break it down into all the components that make it precisely what it is." *Doesn't this presume that we KNOW what it is? How can we know this "precisely" if we are not native speakers of both languages?*
- • "Translation is not a mechanical act; it doesn't proceed by any sort of simple algorithm." *What about complicated algorithms and learning neural nets though?*
- • "So either we say it's impossible for us to truly understand each other even when we're speaking the same language, or we allow ourselves to enjoy the fluidity of our linguistic interactions – and thus to happily read all sorts of literature in translation. Furthermore, with this attitude we can study a

second language with optimism." *Love it!*

- *About my experiences learning, speaking and writing a second language (English) well: ...*

About computers learning languages

I don't know yet to which extent Alice and her class are interested in the computer/human interface as a linguistic boundary.

- *Chapter 7 is pretty out of date (references are from 2009, serious progress has been made since then – cp. e.g. this tracking site https://ruder.io/tracking-progress-nlp/ and here: https://nlpprogress.com/ for progress in different languages. The curator of this site writes a blog: https://ruder.io/. The latest research focuses on transfer learning.*
- *All of this is a little abstract. Here are some concrete examples worth mentioning: https://www.wonderflow.co/blog/natural-language-processing-examples*

More mixed resources:

- "Machines that learn language more like kids do - Computer model could improve human-machine interaction, provide insight into how children learn language: MIT researchers have developed a "semantic parser" that learns through observation to more closely mimic a child's language-acquisition process, which could greatly extend computing's capabilities." http://news.mit.edu/2018/machines-learn-language-human-interaction-1031
- Language Acquisition in Computers: Belzner et al (2012): "Using certain principles of language, we have designed a novel method by which a computer can gain an intuitive understanding of language rather than simply an artificial understanding. We have developed techniques by which a computer can learn and analyze the morphology of any given language, and hence understand differences between two languages. We have also developed a recursive learning system for understanding sentence patterns and constructs, which uses a minimum of initial information. At present, the program can interpret many basic sentences, and we have also provided possibilities and suggestions for extending the capabilities of the program. This approach is unique compared to common natural language processing systems because of this lack of need for significant initial input and its recursive design, and could have great potential in the field of natural language processing.

https://arxiv.org/abs/1206.0042

- • A Simple Introduction to Natural Language Processing: Michael J Garbade, 15 Oct 2018, "Natural Language Processing is the technology used to aid computers to understand the human's natural language." https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32
- • Learn Natural Language Processing: Siraj Javal, 30 June 2019, "I've designed a free natural language processing curriculum for anyone interested in improving their skills in order to start a startup, get consulting work, or find full-time work related to NLP. This curriculum is for beginners and starts with basic NLP terminology, then moves into basic language models and word embeddings. Then, it moves onto more advanced concepts like neural networks, sequence modeling and dialogue systems. At the end, I'll detail what the most experimental, modern-day techniques are in the field. I hope you find this curriculum useful!" https://youtu.be/GazFsfcijXQ
- On Programming Languages vs Human Languages: Ana V Harris (1 Nov 2018) - "Understanding the difference between programming languages and human languages is of great value for many fields that work on creating tools for important practical tasks such as machine translation, speech recognition, speech synthesis, information extraction from text, grammar checking, text mining and more. Think of Google Translate or Siri, and try to understand everything it takes to turn a simple sentence in English into a message that can be processed by a computer." https://medium.com/@anaharris/human-languages-vs-programming-languages-c89410f13252
- 20 Natural Language Processing Example for Business (15-Aug-2019)