

Hypothesis

Shiro Takagi

2021/4/12

1 question

Q: How to make an intelligent system that learns to manipulate symbols?

2 survey

2.1 cognitive science

Chomsky says that the power to taming syntax is innate property of human brain (universal grammar) [1]. Ibbotson says that “the complexity of language emerges not as a result of a language-specific instinct but through the interaction of cognition and use” (usage-based theory) [6]. According to the usage-based theory, linguistic structure develops by 1. categorization, 2. chunking, 3. rich memory, 4. analogy, and 5. cross-modal association [2, 6]. Children use a limited number of reliable short frames.

“ Overall it seems there is good evidence to support the usage-based prediction that language structure emerges in ontogeny out of experience (viz. use) and when a child uses core usage-based cognitive processes – categorization, analogy, form-meaning mapping, chunking, exemplar/item-based representations – to find and use communicatively meaningful units. [6] ”

The meaning of symbols is established by convention [10, 11, 7].

2.2 transformers

Transformer learns syntactic information [9, 4, 3, 12].

2.3 comparative study

Watson et al. claims that “*nonadjacent dependency processing, a crucial cognitive facilitator of language, is an ancestral trait that evolved at least 40 million years before language itself*” [13]. Wilson et al. explains that sufficient cues play crucial role for human nonadjacent dependency learning [14]. Okanoya and Merker propose the hypothesis that human language is established through string-context mutual segmentation: “*song strings and behavioral contexts are mutually segmented during social interactions*” [8].

3 Discussion

3.1 transformers

If transformers really capture syntax, **how it develops the syntactic representation during the pre-training?** Previous studies seem to find that pre-trained transformer have syntactic representation but how to do that remains to be answered. If we can single out the cause of the syntax emergence, we may be able to model a guiding principle for an intelligent agent to learn syntax.

3.2 syntax

It might be plausible that infants first identify phrases in sentences. To identify phrases, it might be necessary that the phrase is used in multiple sentences. By observing the subset of sentences in multiple sentences repeatedly, infants could identify which subset is the the phrase. However, a phrase less likely to appear many times in multiple contexts. Thus, we could hypothesize that infants first understand too common and too often appearing phrase. Then, they understand that theres is the concept “phrase” in their society. Finally, they could start to generalize their knowledge and to do try and error to manipulate phrase order. In sum, **we hypothesize that artificial intelligence should follow the following path for language acquisition: phrase identification - phrase order arrangement - phrase manipulation.** If this is the case, we should create an environment where identifying key phrase will give reward to the agent.

3.3 keyword detection

Dual-coding memory may be a key because visual information is pseudo label there [5]. Even if no instructor exists, the agent can learn the concept him/herself.

4 Note

References

- [1] Noam Chomsky. *Syntactic structures*. Walter de Gruyter, 2002.
- [2] Maria Angélica Furtado da Cunha. Language, usage and cognition. cambridge: Cambridge university press. 252 págs. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, 29(1), 2010.
- [3] Yoav Goldberg. Assessing bert’s syntactic abilities. *arXiv preprint arXiv:1901.05287*, 2019.
- [4] John Hewitt and Christopher D Manning. A structural probe for finding syntax in word representations. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4129–4138, 2019.
- [5] Felix Hill, Olivier Tieleman, Tamara von Glehn, Nathaniel Wong, Hamza Merzic, and Stephen Clark. Grounded language learning fast and slow. In *International Conference on Learning Representations*, 2021.
- [6] Paul Ibbotson. The scope of usage-based theory. *Frontiers in psychology*, 4:255, 2013.
- [7] James L McClelland, Felix Hill, Maja Rudolph, Jason Baldridge, and Hinrich Schütze. Placing language in an integrated understanding system: Next steps toward human-level performance in neural language models. *Proceedings of the National Academy of Sciences*, 117(42):25966–25974, 2020.

- [8] Kazuo Okanoya and Bjorn Merker. Neural substrates for string-context mutual segmentation: A path to human language. In *Emergence of communication and language*, pages 421–434. Springer, 2007.
- [9] Emily Reif, Ann Yuan, Martin Wattenberg, Fernanda B Viegas, Andy Coenen, Adam Pearce, and Been Kim. Visualizing and measuring the geometry of bert. *Advances in Neural Information Processing Systems*, 32:8594–8603, 2019.
- [10] Adam Santoro, Andrew Lampinen, Kory Mathewson, Timothy Lillicrap, and David Raposo. Symbolic Behaviour in Artificial Intelligence. *arXiv preprint arXiv:2102.03406*, 2021.
- [11] Tadahiro Taniguchi, Emre Ugur, Matej Hoffmann, Lorenzo Jamone, Takayuki Nagai, Benjamin Rosman, Toshihiko Matsuka, Naoto Iwahashi, Erhan Oztop, Justus Piater, and Florentin Wörgötter 1. Symbol Emergence in Cognitive Developmental Systems: a Survey. *IEEE transactions on Cognitive and Developmental Systems*, 11(4):494 – 516, 2018.
- [12] Ian Tenney, Patrick Xia, Berlin Chen, Alex Wang, Adam Poliak, R Thomas McCoy, Najoung Kim, Benjamin Van Durme, Samuel R Bowman, Dipanjan Das, et al. What do you learn from context? probing for sentence structure in contextualized word representations. *arXiv preprint arXiv:1905.06316*, 2019.
- [13] Stuart K Watson, Judith M Burkart, Steven J Schapiro, Susan P Lambeth, Jutta L Mueller, and Simon W Townsend. Nonadjacent dependency processing in monkeys, apes, and humans. *Science advances*, 6(43):eabb0725, 2020.
- [14] Benjamin Wilson, Michelle Spierings, Andrea Ravignani, Jutta L Mueller, Toben H Mintz, Frank Wijnen, Anne Van der Kant, Kenny Smith, and Arnaud Rey. Non-adjacent dependency learning in humans and other animals. *Topics in Cognitive Science*, 12(3):843–858, 2020.