What’s logic got to do with it?

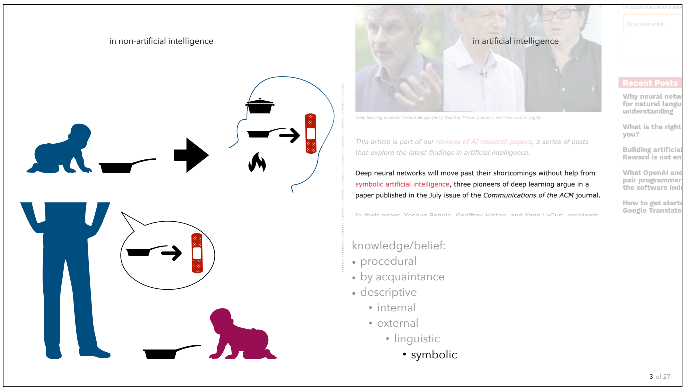
peter bloem, learning and reasoning group, vrije universiteit amsterdam

LEARNING
and REASONING

logic

- A formal language in which to express knowledge.
- A precise way to reason about that knowledge.

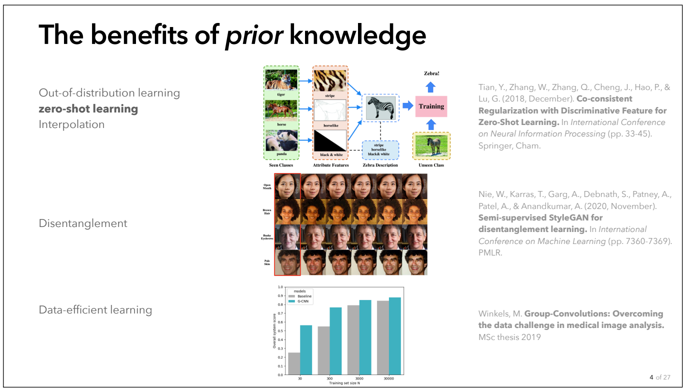
```
likes(john, jurassic_park)
```

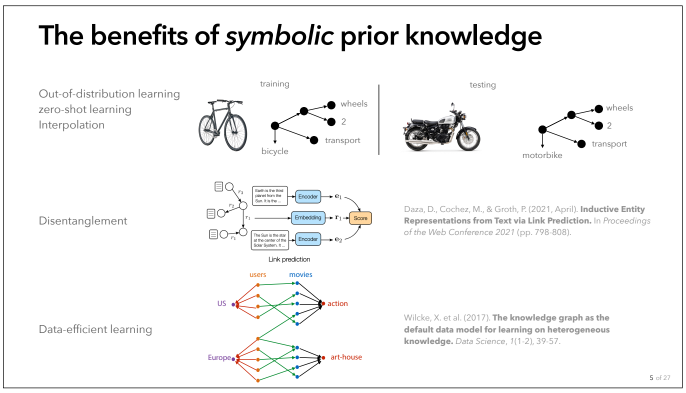
When we produce non-artificial intelligence (also known as children), combining knowledge and learning is the most natural thing in the world. A child may learn through experience that touching a hot pan hurts, but a responsible parent will try to limit such personal experience as much as possible. We do this by distilling our own **experiences** into **knowledge representations** (in this case the phrase "touching a hot pan will hurt") and hoping that the child heeds our warnings.

So why then, when it comes to artificial intelligence do large parts of the learning community seem to reject the help of such symbolic prior knowledge? Why do we insist on learning everything from scratch?

Note that I'm casting a slightly wider net with the definition of knowledge than the common definition of a "justified true belief", since the definition doesn't allow us to distinguish between the beliefs that are knowledge and those that aren't before we use them.

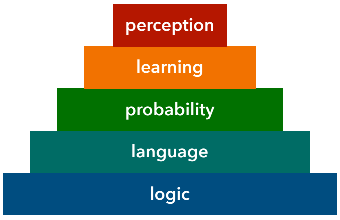


It's certainly not controversial to say that knowledge might help, in learning, or even be required. Here are three of the places where knowledge might help.



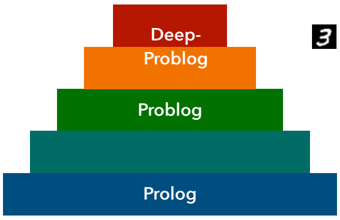
Downside: highly use case specific.

the pyramid of thought



---

neurosymbolic AI

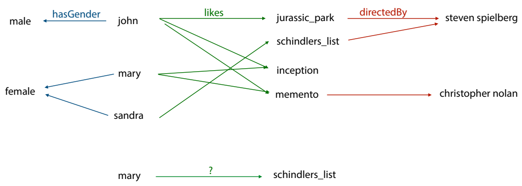


Manhaeve, R., Dumancic, S., …(2018). **Deepproblog: Neural probabilistic logic programming.** *NeurIPS*

van Krieken, E., Thanapalasingam, T. … (2024). **A-nesi: A scalable approximate method for probabilistic neurosymbolic inference.** *NeurIPS*

---

link prediction: `likes(mary, schindlers_list)?`



training: learn to separate true edges from randomly sampled *negatives*.

domain/range constraint:

marriedTo: subject is always a <human>, object is always <human>

<henry8, marriedTo, catherineOfAragon>

<henry8, marriedTo, janeFonda>

<henry8, marriedTo, towerOfLondon>

If we have the semantics, which negatives should weight more heavily?

Either we think of the semantically correct negatives as being “less wrong” so they should carry a lower loss, or as “hard negatives” which are more challenging to recognize as negatives, so they should carry a higher loss (or equivalently, be more likely to be sampled).

Treat Different Negatives Differently: Enriching Loss Functions with Domain and Range Constraints for Link Prediction

Nicolas Hubert[1,2][0000-0002-4682-422X], Pierre Monnin[3][0000-0002-2017-8426], Armelle Brun[2][0000-0002-9876-6906], and Davy Monticolo[1][0000-0002-4244-684X]

[1] Université de Lorraine, ERPI, Nancy, France
[2] Université de Lorraine, CNRS, LORIA, Nancy, France
[3] Université Côte d'Azur, Inria, CNRS, I3S, Sophia-Antipolis, France
{nicolas.hubert,armelle.brun,davy.monticolo}@univ-lorraine.fr
pierre.monnin@inria.fr

Abstract. Knowledge graph embedding models (KGEMs) are used for various tasks related to knowledge graphs (KGs), including link prediction. They are trained with loss functions that consider batches of true and false triples. However, different kinds of false triples exist and recent works suggest that they should not be valued equally, leading to specific negative sampling procedures. In line with this recent assumption, we posit that negative triples that are semantically valid w.r.t. signatures of relations (domain and range) are high-quality negatives. Hence,

In this recent paper, the authors get good results by giving the semantically correct triples a lower loss than the semantically incorrect ones. This causes the model to internalize the semantics.

Berlin » Mrs Eklöf-Berliner-Mauer: The woman who married the Berlin Wall

11.08.2023 - 15:27 Uhr

Berlin

Mrs Eklöf-Berliner-Mauer: The woman who married the Berlin Wall

The course of true love never did run smooth. The story of how a Swedish lady married the Berlin Wall.

It was a lover and his lass. Photo: via kotzendes-einhorn.de

On August 13, 1961, amidst rising tensions between East and West Germany, and much

More importantly for the current discussion. There are always exceptions to any logical rule.

The *downsides* of symbolic prior knowledge

- The platypus problem
- The rhinoceros problem
- The chair problem
- The spork problem

12 of 27

Let's look at four examples of how we use symbolic knowledge in everyday life that show the downsides of relying too much on it.

The platypus problem

No mammals lay eggs. Only birds have bills.

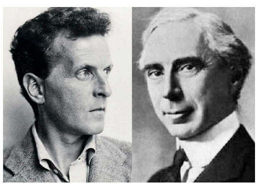
13 of 27

This doesn't mean that these rules are *useless*, just that there are occasional exceptions. More importantly, there will be occasional exceptions that we cannot account for a-priori.

We will observe them in the wild, and we will need to decide on the fly whether to trust our knowledge, or our eyes.

The rhinoceros problem

Rules require *context*.

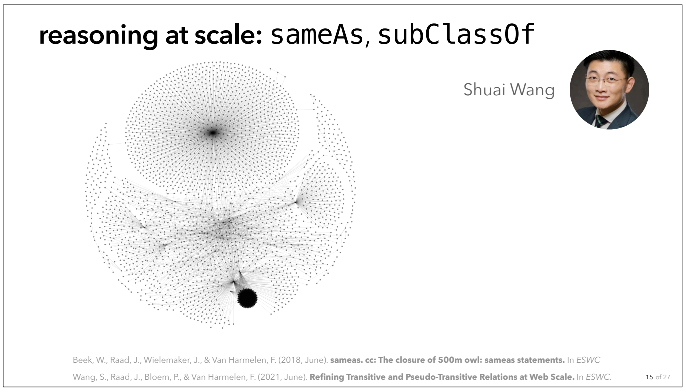


14 of 27

With a little creativity, I believe you can come up with potential counterexamples to any rule. This was a famous point of disagreement between Russell and Wittgenstein when they first met. The latter asserted that there was no such thing as a "truly knowable empirical fact". Russel suggested the statement "There is no Rhinoceros in this room." Apparently Russell even suggested looking under the desks. Wittgenstein's point appears to have been that it was merely very unlikely that was a rhinoceros in the room but not fully impossible.

I'm on Wittgenstein's side. We don't need to go so far as to image microscopic or invisible rhinoceros. With a little creativity, we can, for instance, imagine the possibility that one of the people present had a rhinoceros keychain. That would be a coincidence, but certainly not impossible.

You may argue that this is cheating. Russell was surely referring to *actual* rhinoceros. But for our purposes, at least, this is an important point. If we are talking about small probabilities, we must consider the possibility that the original statement was poorly phrased, or ambiguous. It's truth depends on our interpretation and the context in which we apply it.

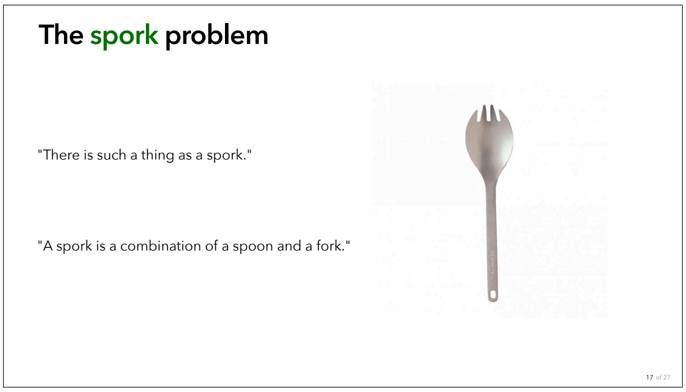


Another problem is that there are certain concepts that are simply difficult to define in simple terms. We all know when something is a chair, but when you start making rules, like "it must have legs", "you can sit on it" or so on, it becomes very easy to come up with counterexamples. Things that break the rules and are very clearly chairs, or things that satisfy all the rules and are very clearly not.

Wittgenstein used *games* as the prime example of this type of concept, and called them *family-likeness terms* (Familienänlichkeiten).

If it's so difficult to define precisely what makes something a chair, a soup or a game, why is it that

we use these concepts so easily? Probably more easily than we do concepts with very precise definitions, like "right-of-way", "finite-state-machine" or "submission deadline"? I think the answer is that we use *learning*. We see two or three examples of a chair and we get the general idea. As we go through life we see more examples and counter-examples and we refine our internal representations.

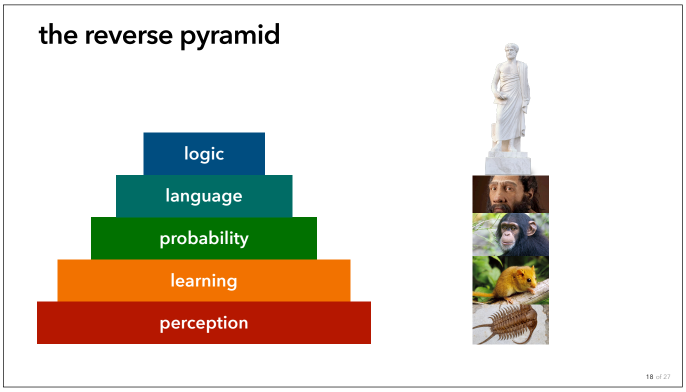


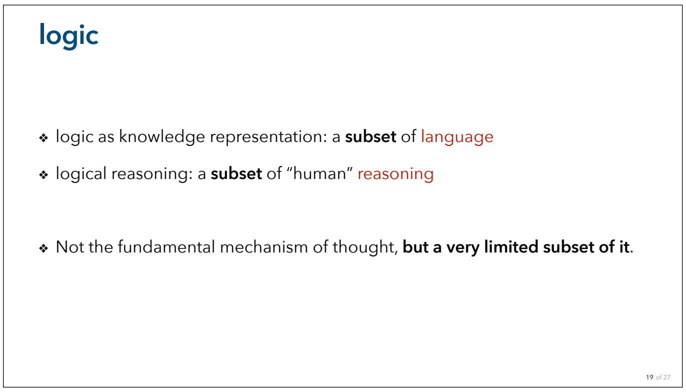
Finally, and most importantly, there's the spork problem. Imagine that you don't know what a spork is. I can tell you that there is such a thing. Even though you don't know what it is, or anything about it, you have no problem processing the information that such a thing exists. As we speak, you are creating space in your head for the concept of a spork and perhaps making some educated guesses about what it might be.

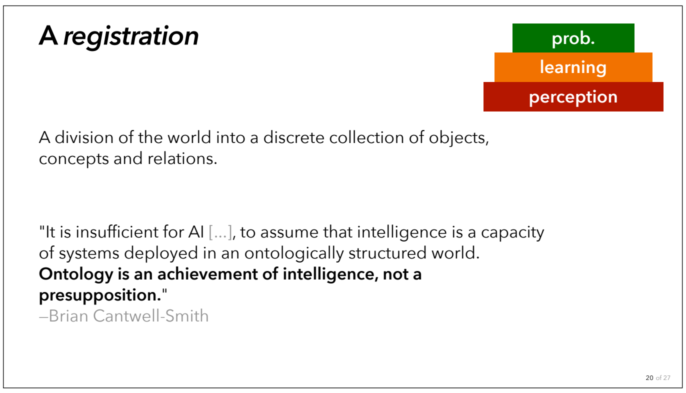
Then, as I tell that it's a combination of a spoon and a fork, you start to fill in the blanks. You now know its approximate shape and size, and you know what it's for. There are a few ways one might combine a spoon and a fork, so you still don't know exactly what it looks like, but you can already narrow it down to a small and finite number of possibilities.

Then I show you a picture and your idea of a spork is complete. Now, whenever you come across one in the rest of your life, you can recognize it. Even though you'll probably never come across one that looks exactly like this.

On the fly, with zero effort, based on almost no knowledge, you have created a new concept and tied it into the rest of your internal semantic network.







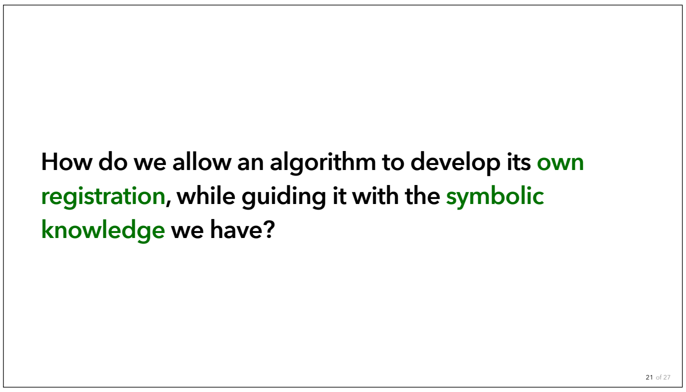
I believe all of these issues emerge from one single problem in the way symbolic knowledge is used on all neurosymbolic approaches being studied today.

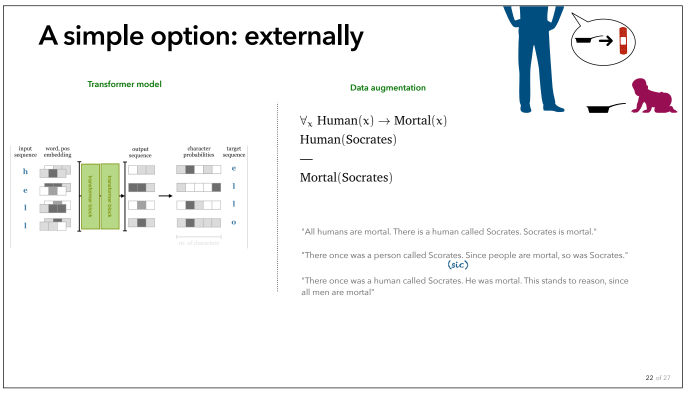
The problem of registration. This is a phrase coined by philosopher Brian Cantwell-Smith. An intelligence's registration is the way it takes its collection of raw, continuous input signal, and organizes them into a (mostly) discrete picture of the world. In short, the way it maps observations to symbols.

The point that Cantwell-Smith makes is that building a registration, including the vocabulary of symbols must be part of a true intelligence. An agent must be allowed to build its own registration, it's own collection of symbols, introducing new ones as the need arises.

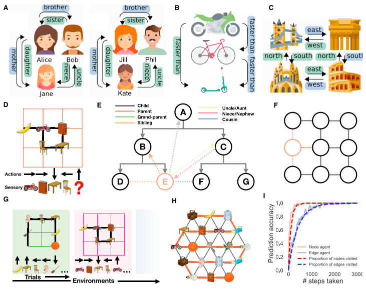
If we take our registration, our ontologies, and limit the agent to that particular registration of the world, it can never be truly intelligent, and one or all of the four problems we saw before will emerge.

That doesn't mean we can't use our own knowledge to help intelligent agents emerge, only that our knowledge can't form the internal registration of the agent. It must be outside of the

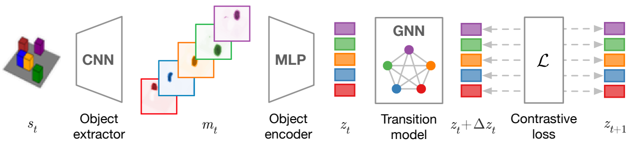




More complex: internally



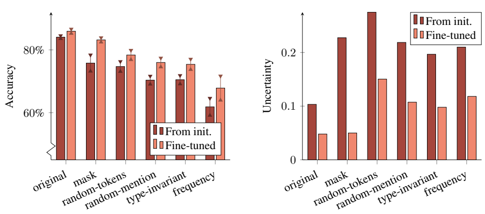
Whittington et al. (2020). **The Tolman-Eichenbaum machine: Unifying space and relational memory [...].** *Cell, 183*(5), 1249-1263.



Kipf, T., Van der Pol, E., & Welling, M. (2019). **Contrastive learning of structured world models.** *arXiv preprint arXiv:1911.12247.*

Probing the representations of named entities in Transformer-based Language Models

Stefan Schouten



(a) Accuracy for News Topic Classification.

(b) Uncertainty for News Topic Classification.

Are identities represented and are they *used*?

Reasoning about Ambiguous Definite Descriptions

Stefan F. Schouten and **Peter Bloem** and **Ilia Markov** and **Piek Vossen**
Vrije Universiteit Amsterdam
{s.f.schouten,p.bloem,i.markov,p.t.j.m.vossen}@vu.nl

Abstract

Natural language reasoning plays an increasingly important role in improving language models' ability to solve complex language understanding tasks. An interesting use case for reasoning is the resolution of context-dependent ambiguity. But no resources exist to evaluate how well Large Language Models can use explicit reasoning to resolve ambiguity in language. We propose to use ambiguous definite descriptions for this purpose and create and publish the first benchmark dataset consisting of such phrases. Our method includes all information required to resolve the ambiguity in the prompt, which means a model does not require anything but reasoning to do well. We find this to be a challenging task for recent LLMs. Code and data available at: https://github.com/sfschouten/exploiting-ambiguity

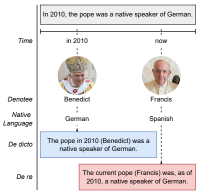


Figure 1: Example ambiguous definite description. Since a person's native language does not change over time, we know that the *de dicto* interpretation is correct.

Truth-value judgment in LLMs: 'truth directions' in language models are context sensitive

You are looking at a picture.
Q Describing it as: "Four children are playing in some water." is [in]correct.
H Saying (about the picture) that: "The children are wet." is [in]correct.

(a) SNLI (b) EntailmentBank