



PLIN037 Sémantika a počítače

Zuzana Nevěřilová
2024

Znalost světa

Implicitní → Explicitní znalost

Objekty:

- Popis typického představitele

Situace:

- Typické situace (váleční rámce)

Encyklopedická znalost

- Německo leží v Evropě

Common sense:

- Co je větší než myš?
- Způsobí rána knihou do hlavy smrt?
- Kde se běžně nachází jídlo?



Common Sense

„Common sense includes commonsense knowledge – the kinds of facts and concepts that most of us know – but also the commonsense reasoning skills which people use for applying their knowledge. We each use terms like commonsense for the things that we expect other people to know and regard as obvious“

Marvin Minsky



Common Sense: formální popis

- **Common sense nemá přesné hranice.**
- **Common sense je sdílený (ale nevíme jak moc).**
- **Některá tvrzení jsou v rozporu s (vědeckou) skutečností (např. V noci nesvítí slunce.).**
- **Většinu se naučíme jako děti (z vyprávění nebo z ukázek od ostatních lidí, příp. vlastní zkušeností), tudíž znalosti common sense jsou málokdy zaznamenány jako text.**
- **Common sense je velmi rozsáhlý.**

Common Sense: Estimates of the Size of Human Common Sense

Turing (1950, p. 442)	10^9 bits for a 70% chance of identifying a program as not a human after 5 min of text-based questioning
von Neumann (1958, pp. 63 and 64)	2.8×10^{20} bits of "memory capacity" based on 10^{10} neurons \times 14 bits per second per neuron \times 2×10^9 s, assuming "there is no true forgetting in the nervous system"
Minsky (1968, p. 26)	10^5 - 10^7 "elements of knowledge in order to behave with reasonable sensibility in ordinary situations"
Landauer (1986)	10^9 bits of "information from experience"
Schwartz (1988, p. 126)	0.8×10^{17} bits of "long-term memory"
Marvin Minsky (Lenat & Guha, 1990, p. 21)	10^7 entries, assuming a child acquires an entry every 10 seconds from ages 0 to 8
Kurzweil (1999, p. 119)	10^8 chunks or "bits of understanding, concepts, patterns, specific skills"
Moravec (1999, p. 56)	0.8×10^{15} bits of "nervous system memory" based on 10^{14} synapses \times 1 byte per synapse
Mahoney (1999)	10^9 bits for predicting "how people respond to arbitrary input during communication"

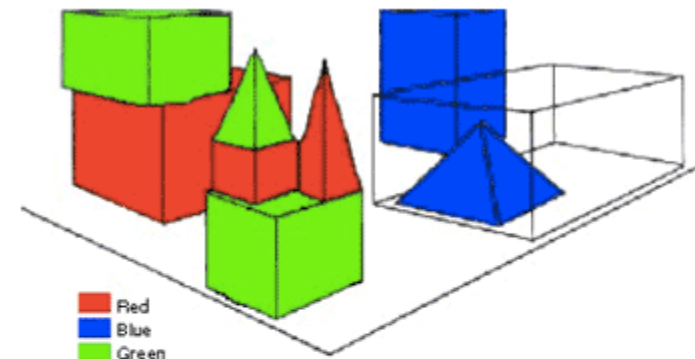
Common Sense: neobejdeme se bez něj

dlouhodobě známý problém v umělé inteligenci

SHRDLU (Terry Winograd, MIT, 1968–1970)

(mikro)svět kostek, kuželů, koulí apod.
ovládaný pomocí vět v přirozeném jazyce

<http://hci.stanford.edu/~winograd/shrdlu/>



Person: Pick up a big red block.

Computer: OK.

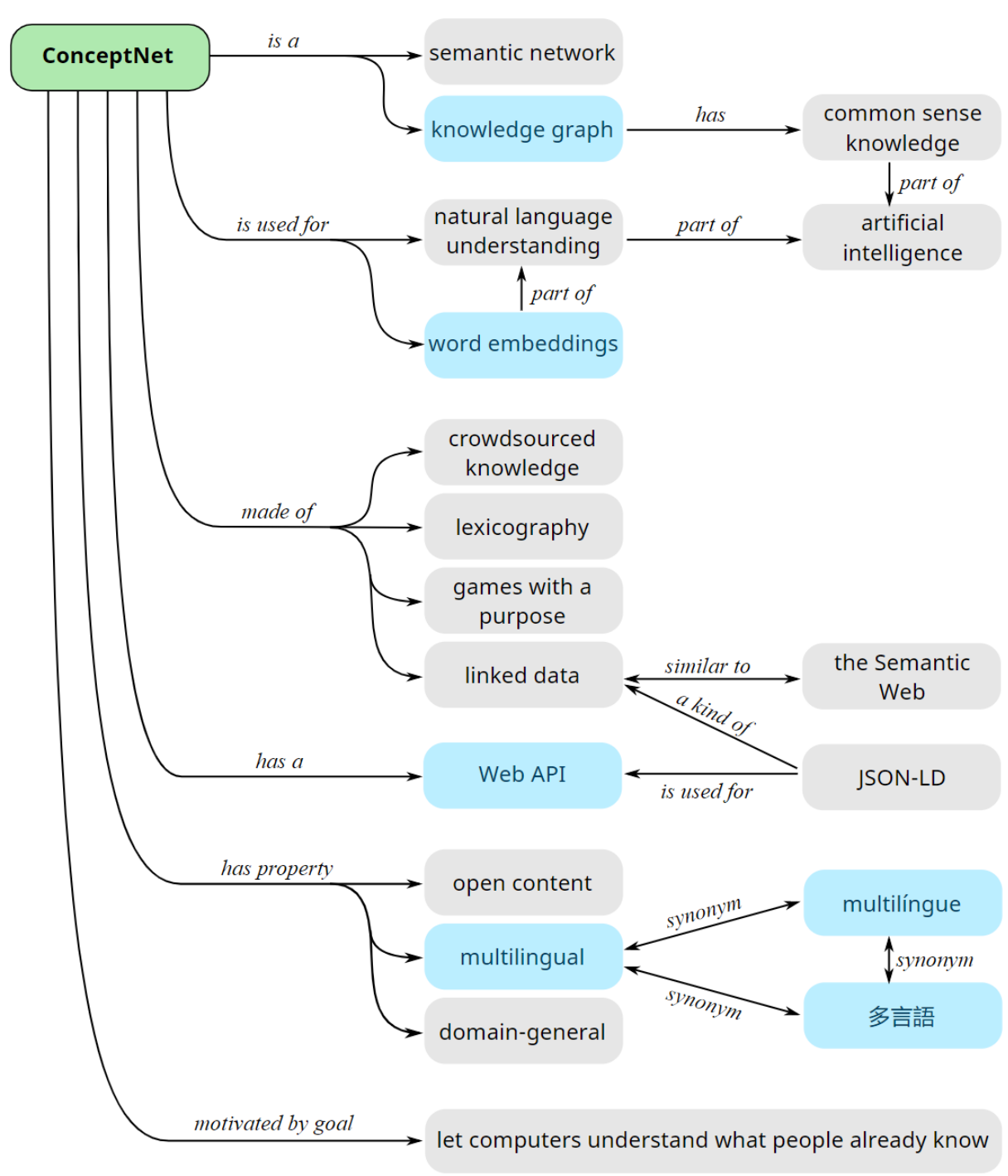
Person: Grasp the pyramid.

Computer: I don't understand which pyramid you mean.

Common Sense: SHRDLU

- **Syntaktický analyzátor a omezený slovník (doména geometrických těles: tj. umístit, velký, modrý, kostka, kvádr, krychle, kužel, posunout ...)**
- **Paměť pro uchování kontextu (např. Postav zelený kužel na červenou kostku. Přemísti ten kužel na modrou kostku.)**
- **Odvození vlastností těles na základě příkladů (např. Není možné umístit objekt na místo, kde už nějaký objekt je. Je možné postavit jednu kostku na jinou kostku. Není možné postavit jeden kužel na jiný kužel.)**
- **Schopnost naučit se pojmenovávat stavby podle jejich uspořádání (Věž je několik kvádrů nebo krychlí na sobě a na nich kužel.)**





Common Sense: Projekt

- ThoughtTreasure
- CyC (OpenCyC, ResearchCyC)
- Open Mind Common Sense → ConceptNet
- WebChild (Commonsense Knowledge from the Web)
- GLUCOSE (Generalized and Contextualized Story Explanations)
- ATOMIC (An Atlas of Machine Commonsense for If-Then Reasoning)
- hasPartKB
- ...

ThoughtTreasure

Reprezentace faktů pomocí různých formalismů:
FOPL, konečné automaty, skripty.

Přes 25 000 pojmů hierarchicky organizovaných
(synonyma ve francouzštině a angličtině).
Odvozování a porovnávání s očekáváním.

A je nepřítelem B.

A nesnáší B.

A pomáhá B.

← nezvyklé



Natural Language Processing
with ThoughtTreasure

Erik T. Mueller



CyC

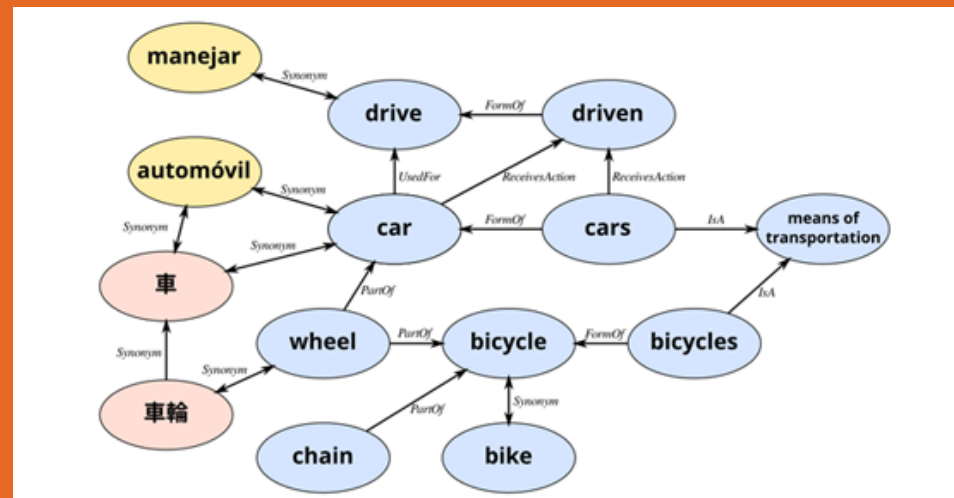
- **vývoj od 80. let (vytvořeno experty)**
- **V r. 2007 cca 3 miliony propozic (miliony dalších z nich mohou být odvozeny)**
- **různé způsoby odvozování (dedukce, dědičnost podobně jako v sémantických sítích)**
- **konzistence uvnitř mikroteorií**
- **propozice o pravděpodobnosti propozic**



Open Mind Common Sense a ConceptNet

Open Mind Common Sense sbíral „fakty“ od uživatelů internetu (ve formě jednoduchých vět)

- ConceptNet – mnohojazyčná sémantická síť zahrnující fakty
- několik druhů relací
- odvozování pomocí analogie (Divisi)



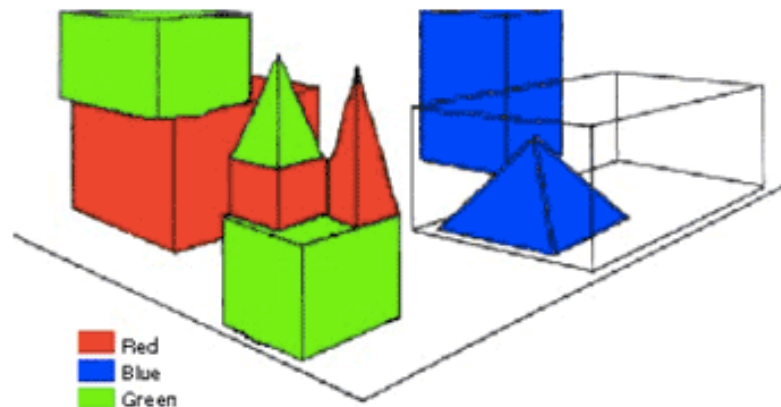
<https://opendatascience.com/conceptnet-5-5-and-conceptnet-io/>

Common sense and language models

Představ si, že jsi robot, který skládá hrací kostky různých tvarů. Popíšu ti kostky, vždy písmenem a pak specifikuji, jak kostka vypadá...

- Na některé tvary nelze stavět (lze, ale stavba hned spadne)
- ChatGPT3.5 to neví

<https://chat.openai.com/share/c98abdac-71b1-415f-8985-3d2f8d440b84>



Person: Pick up a big red block.

Computer: OK.

Person: Grasp the pyramid.

Computer: I don't understand which pyramid you mean.



Literatura

- Minsky, M. (1999).
The emotion machine: from pain to suffering.
In *C&C '99: Proceedings of the 3rd conference on Creativity & cognition*, page 7–13, New York, NY, USA. ACM.
- Erik T. Mueller, Chapter 19 - Acquisition of Commonsense Knowledge, Editor(s): Erik T. Mueller, *Commonsense Reasoning (Second Edition)*, Morgan Kaufmann, 2015,
<https://doi.org/10.1016/B978-0-12-801416-5.00019-X>.
- Mariya Yao. **4 Approaches To Natural Language Processing & Understanding**. freeCodeCamp. 2017. <https://www.freecodecamp.org/news/how-natural-language-processing-powers-chatbots-4-common-approaches-a077a4de04d4/>
- Erik Mueller: ThoughtTreasure. 2015.
<https://github.com/eriktmueller/thoughttreasure>
- <https://github.com/ElementalCognition/glucose>
- <https://homes.cs.washington.edu/~msap/atomic/>
- <https://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/commonsense/webchild>