HOKKAIDO UNIVERSITY

Title	Semantic Primes-Inspired Tacit Knowledge Dataset for Simulating Basic Perception Capabilities of Cognitive Architectures
Author(s)	Rzepka, Rafal; Shinto, Ryoma; Araki, Kenji
Citation	Lecture Notes in Computer Science, 14951, 145-154 https://doi.org/10.1007/978-3-031-65572-2_16
Issue Date	2024-08-12
Doc URL	https://hdl.handle.net/2115/95897
Rights	This version of the contribution has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/978-3-031-65572-2_16. Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms
Type	journal article
File Information	Lect. Notes Comput. Sci. 145-154_2024.pdf



Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Semantic Primes-Inspired Tacit Knowledge Dataset for Simulating Basic Perception Capabilities of Cognitive Architectures

Rafal Rzepka[1][0000−0002−8274−0875], Ryoma Shinto[1], and Kenji Araki[1]

Hokkaido University, Kita-ku, Kita 14, Nishi 9, Sapporo, 060-0814, Japan
rzepka,shinto,araki@ist.hokudai.ac.jp

Abstract. In this paper we present a novel dataset of tacit knowledge represented in natural language (Japanese) inspired by semantic primes categories. The main goals of this data is to a) allow investigations regarding influence of perception data in various cognitive tasks, b) mimic signals for cognitive processes of an artificial agent to extend the understanding of the world and c) testing cognitive capabilities of intelligent instances like foundation models. We describe the dataset and share results of preliminary experiments showing that the tacit knowledge recognition is still hard for language models. We also discuss how such redirecting neural approaches to cognition only and then perform reasoning in a symbolic realms could become beneficial for new type of simulations before AGIs are equipped with more sophisticated sensory apparatus.

Keywords: Semantic Primes · Tacit Knowledge · Simulated Perception

1 Introduction

Deep learning-based approaches have dominated many domains of artificial intelligence research, and also have become a part of AGI research including development of cognitive architectures. However, due to limited stimuli perceived by an artificial agent, there are obvious limitations of how the understanding of external world is performed, which leads to situations when seemingly correct foundation models output unexpected results at random moments [14]. Therefore, our main research question is how to systematically study influence of cognitive mechanisms on learning or reasoning. To help answering this question, we decided to construct a dataset which represents basic recognition schema of an act which is not represented in the text directly (here we call it "tacit knowledge", although the term is often defined differently). Perception process of biological agents can be simulated via various sensory mechanisms including visual, auditory or tactile ones, however gathered signals (obtaining of can be costly and time consuming) are often not sufficient to make sense of their mutual influence on reasoning. Due to this problem, it is difficult to perform a thorough experimentation on the roles that tacit knowledge plays in artificial systems acquiring

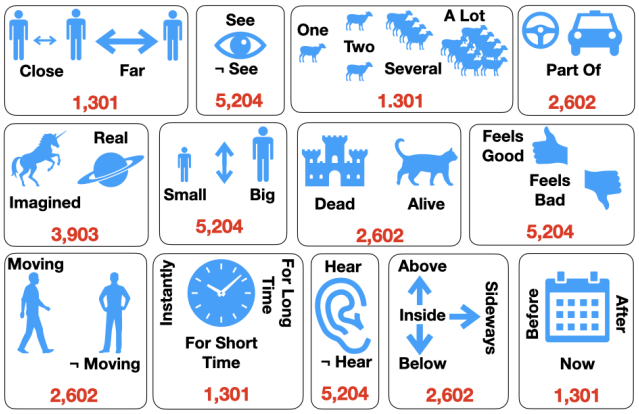


Fig. 1. Examples of semantic primes categories-inspired prompts. Numbers of human-annotated samples are marked in red (total size od the dataset exceeds 62,000 sentences).

explicit knowledge or extrapolating a new one. For that reason, we concentrate on signals described in natural language instead of signals themselves, which can be replaced later by physical ones. Here we assume that simulating a set of human mental functionalities with language models could become a testbed for experiments without sensory mechanisms and partially decrease the cost of experiments with life-long learning agents in the wild. But how to choose a basic set of such functionalities? What are the fundamental mechanisms that humans are born with? Our assumption is that having a set of perception descriptions for an input, for example "toy belongs to girl", "girl can hear toy", or "girl is bigger than toy", "toy is far from girl", etc. added as an additional knowledge to an input "a girl looks at a toy on a shelf", can be used for making cognitive observations of sequential acts when the next input is "a girl reaches for a toy" and the distance percept changes to "toy is close to girl". We believe that interaction of basic signals about what is real, what is happening now or in the past, what is moving, who is not performing an act, an so on, can help to discover how we invented logic, how we compare and make analogies, how we create metaphors, etc.

As nativist approaches advocated by thinkers like Chomsky [2] or Pinker [15] do not give an exact taxonomy of what mechanisms are involved in the understanding or communicating process, we have chosen semantic primes [28, 29] as an inspiration for how to construct a dataset of the tacit knowledge. We describe the collected data in the next section, then present preliminary experiments with the data and their results. In the last section we discuss the results and potential further research with the proposed dataset.

2 Related works

Although there have been several attempts to infer implicit knowledge from data [19, 25, 18] including language models [23, 12], to the authors' best knowledge no dataset of tacit knowledge has been published before. The possible reason for that situation is that it is rather hard to think about aspects of our perception because we perceive the world automatically and it is so obvious that we do not think about them directly, not to mention writing them down. The implicit knowledge is partly hidden in a foundation model parameters or between yet to connect graph nodes of a knowledge base but it is not concrete unless a human makes it explicit. This approach is used in commonsense knowledge repositories as ConceptNet [22] or ATOMIC [21]. However, relations used (MadeOf, UsedFor, Synonym, etc.) are meant for creating static links between already acquired knowledge instances rather than for representing always changing cognitive input. To allow more dynamic processing, we developed a set of basic perception-based primitive prompts for which the answers constantly change with context (environment). Instead of creating a traditional commonsense ontology, we limited the possible answers (percepts) to help agent ground new concepts on the fly. Such approach, to the authors' best knowledge, has not yet been proposed.

The structure of our systematic taxonomy of cognitive mechanisms (primitive prompt and few possible answers to be chosen) can be used for various studies. For example, it can become a benchmark for testing cognitive capabilities [26] or an addition to chain-of-thought experiments [27], which is difficult with existing knowledge bases due to their size. However, we hope for a wider scope of use not limited to machine learning or knowledge / language acquisition – the data can be utilized to study the influence of perception in cognition, memorizing, reasoning [9], analogy making, language acquisition and understanding [4] or anomaly detection, for example in scenarios when slight context alternations lead to drastically different interpretations [11].

Dividing meaning into the smallest possible notions has been studied in cognitive linguistics [28, 24, 29] but semantic universals are not directly used in creating artificial entities [8, 20]. Explaining a concept in semantic primes of a meta language is a daunting task even for a trained human and describing the world using such language might be one of hardest tasks for comparing natural and artificial intelligences. To open opportunities for such research, we develop the dataset described in the next section.

3 Dataset Creation

Natural Semantic Metalanguage or NSM [7] with semantic primes is a less open-ended decompositional approach to meaning when compared to, for instance, Jackendoff's conceptual semantics [10] or Talmy's semantic taxonomy [24]. Although there are some doubts if the NSM is universal [17], we considered the limited number of semantic primes suitable for the future investigation seeking

answers to our question, namely "which types of cognitive functionality an agent should posses before exploring the environment and understanding or learning about the world?". Inspired by the exponents grouped into related categories[1] and their translation to Japanese [6], we created a list of perception-related 23 types of prompts (in Japanese) as shown in Table 3. While we omitted these related more to syntax than concepts or acts (e.g. most substantives and logical concepts), we left "is" (*dearu*) and "is not" (*dewanai*) as the negation becomes an important part of the prompt continuations – for example to decide if something is visible or **not** visible. We also simplified, aggregated and specified some prompts to make the annotation easier (e.g. we used "several" in *Quantifiers* to shorten the list possible continuation choices, although this word is not included in the original list of primes). For examining the language capability of comparing sizes we have added "than" to "big", "small" and "the same size" as a separate query. The only words for the *Speech* category are "real" and "unreal", but for the perception scenario we replaced them with Japanese words for "real / reality" (*genjitsu*) and "imagination / imagined" (*kūsō*). *Intensifiers* and *Augmentors* ("more" and "very") have been omitted due to difficulty of automatic generation of examples and will require more sophisticated language generation for making new versions of the dataset. For *Evaluators* we decided to replace general "good" and "bad" words with Japanese equivalents meaning "feels good" and "feels bad" after receiving a feedback from annotators in the annotation testing phase that the broad words do not fit the guideline asking for perceiving not judging approach to choosing continuations. To made "see" and "hear" more perception-oriented, we replaced them with "can see / visible" and "can hear / audible" and added prepositions to "far" and "close". It must be noted that to use these categories for describing world in the universal meta-language, in the future the original words need to be also included in the set. As our data is to represent general knowledge, no agency is involved, and for this reason we extended the *Possession* category word ("mine") to a grammatical construct (*-no mono*) including any agent. To add an example of *Determiner* category, we added "somebody else" to *Possession* and created separated pair of prompts. It must be noted that some semantic primes categories are represented by only one exponent (like Eval), while other categories contain many (for example Space represented by nine)[2]. Simplifying prompts into one set of possible percepts can cause problems with choosing the most probable one, for example something can be far away and above at the same time.

3.1 Prompt Generation Process

In order to prepare the dataset we utilized a Japanese dataset introduced by [11]. Meant for detecting danger level changes in slightly different contexts, it

[1] Related materials are available at `https://intranet.secure.griffith.edu.au/schools-departments/natural-semantic-metalanguage/what-is-nsm` (last accessed April 2, 2024).

[2] E.g. *"Daughter ate a cigarette": Daughter is dead /* **alive***; Cigarette is* **smaller** *than / bigger than / the same size as (daughter)*, etc. (agreed choice in bold)

Category	Possible Percepts
RelSub	is a part of / is not a part of
Det	same / similar / different
Quan	one / two / several / a lot
Eval	feels good / feels bad / feels normal
Desc	small / big / normal
Desc	smaller than / bigger than / the same size as
AcEM	doing / not doing
AcEM	moving / not moving
LaD	alive / dead
MenPr	want / don't want
MenPr	can see / can't see
MenPr	can hear / can't hear
Speech	real / unreal
Poss	is of (belong) / is not of (don't belong)
Poss	is of someone else / is not of someone else
Time	happen now / happened before / will happen
Time	happen instantly / for short / long time
Time	happen 0 / 1 / 2 / several / a lot of times
Space	touch / don't touch
Space	is far from / is close to
LocES	is there too / is not there
LocES	at the same place / at a different place
LocES	is above / is below / is sideways / is inside

Table 1. List of potential continuations to sentences and prompts based on semantic primes categories. Category name abbreviations: RelSub: *Relational Substantives*, Det: *Determiners*, Quan: *Quantifiers*, Eval: *Evaluators*, Desc: *Descriptors*, AcEM: *Actions - Events - Movement*, LaD: *Life and Death*, MenPr: *Mental Predicates*, Poss: *Possession*, LocES: *Location - Existence - Specification*.

comprises of more than 20,000 short sentence pairs as "child eats a soap" and "child eats a soap-shaped candy". We chose this dataset due to the relatively fixed structure of sentences always including an agent, a patient / object and an act (verb) which helped us to easily retrieve these three elements by using accompanying them Japanese particles[3] for generating prompts to the for continuations described in the previous subsections. This allowed us to automatically generate 49 prompts about agents, objects / patients and acts from a single sentence. In order to increase variation of the set and to be able to see differences depending on the sentence tense, we randomly altered the original verb in basic form to three different forms: a) present continuous ("-te iru" for currently happening acts), b) past ("-ta" / "-da" for acts that already happened) and c) volitional mode ("-tagaru" / "tai" for indicating that agent wants to perform an act in the future but it hasn't happened yet[4]). The third choice was due to the fact that Japanese language has no future tense. All the verb alternations has been done

[3] Namely "ga" for extracting agent before the particle and "wo" which divides object or agent from the act (verb). In the case of extended "dangerous" sentences, the crowdoworkers used by Katsumata et al. had more freedom and sometimes our strategy caused erroneous splits. To alleviate this problem, we filtered out sentences with more than one particle "ga" or "wo".

[4] In most cases "-tagaru" form for the third person has been used, the only condition for using "-tai" suffix was if the agent = myself (*jibun*).

automatically with the *Kotodama* library for Python[5]. We have randomly chosen 1,320 sentences from Katsumata et al.'s dataset and generated 49 prompts for each sentence, creating a dataset for annotation comprising of 64,680 sentences.

3.2 Human Annotation

We divided 1,320 basic sentences into 66 excel files of 20 basic sentences and 49 prompts (980 sentences with continuations in total). We asked 66 annotators[6] on the crowdsourcing platform[7] to choose prompted continuations sent in three different files providing 2,940 prompts for 60 sentences (3 x 20 x 49). By doing so, we ensured that every sentence has been annotated by three people. The task was to choose a single most probable continuation for a prompt regarding an act (e.g. choosing whether the doors are bigger or taller than a person). After analyzing the annotated data we found out that there were 40,980 (64.28%) sentences to which all choices were identical, 21,707 (34.05%) where two annotators agreed, and 1,062 (1.66%) cases in which all annotators disagreed (Cohen κ 0.748). We excluded these 1,062 sentences and after deleting identical sentences, our final golden set for experiments consisted of 62,687 annotated sentence-prompt-choice triples as: "father smokes a cigarette" $\rightarrow$ "father and cigarette" $\rightarrow$ "1) are far; 2) are close".

4 Language Models as Perceptors: Preliminary Experiment

In an open-world scenario, an artificial agent or cognitive architecture would need to be able to simulate perception on the fly. Although prompting vision models would be more appropriate for such tasks, due to the high costs, for time being we decided to investigate the capability of popular language models to automatically recognize semantic primes continuations (percepts) and compare them to human choices.

We tested two already classic models: BERT [5] and RoBERTa [13] and two popular proprietary LLMs from OpenAI: GPT-3.5 and GPT-4. For the former two, we used Japanese BERT[8] with two approaches (MaskedLM / Next Sentence Prediction) and a multilingual model XLM-RoBERTa-large-XNLI [3] (hereafter abbreviated to RoBERTa-XLM) trained on natural language inference data. For the latter, we used *gpt-3.5-turbo-0613* and *gpt-4-0613*, and in all cases we took *zero-shot* approach meaning that the model is queried directly without any fine-tuning. The results (see Table 4) show that not only the older models struggle to choose the correct percept but also the more recent LLMs's accuracy is far

[5] `https://pypi.org/project/kotodama/` last accessed 2024/4/12

[6] 30 females (38.7 years old on average) and 36 males (40.3 years old on average). The payment was approximately 45USD.

[7] `https://crowdworks.jp`

[8] `https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/` (last accessed 2023/1/16)

from perfect. However, it must be noted that due to high costs, we have tested the proprietary models only on small number of random samples. The prompt to GPTs was given in Japanese language in a form of a numbered multiple choice with a question to choose the most natural percept.

Model	Correct Recognitions	Percentage
RoBERTa (NLI)	27,534 / 62,687	43.92
BERT (NSP)	28,694 / 62,687	45.77
BERT (MaskedLM)	32,952 / 62,687	**52.56**
Voting	29,429 / 62,687	46.94
gpt-3.5-turbo-0613	296 / 478	**61.92**
gpt-4-turbo-0613	182 / 388	46.90

Table 2. Comparison of different language models performance in the task of choosing continuations (percepts) of semantic primes prompts with zero-shot approach

The highest number of continuations chosen with agreement with humans was 52.5% for older models (BERT with MaskedLM class) and surprisingly GPT-3.5 (61.92%) for newer LLMs, which indicates that a big chunk of implicit knowledge might be not retrievable from sole parameters of language models. Combining all three approaches into one voting model did not bring any improvement suggesting that depending on the pretraining approach and learned data, language models can predict percepts quite differently.

5 Discussion

After having a closer look at the language models output, we noticed that there are visible discrepancies between what kind of perception-related errors are made. An interesting tendency is a very high prediction rate for "feels good" which is assigned 8.6 times more frequently than "feels bad", even if the data we used describes many dangerous situations and we expected the opposite ratio. On the other hand, the Eval category seems to be difficult to interpret also to humans. When we analyzed which categories have lead to annotators' complete disagreement, we discovered that the "feeling" category was the most problematic and there were many more divided cases than in other categories – namely "size" Desc (twice less) and "amount" Quan (six times less). This may suggest that language models[9] have biases toward different types of perception (or rather their textual representations in our case) and some weighting schema might be required to assure more precise cognitive simulations. However, no matter how much we tune the text-based models, some concepts are difficult for them. For instance, recognizing if something is dead or alive can be distorted by idiomatic

[9] It must be noted than large language models for Japanese are not abundant and tests with bigger models dedicated to this language are needed.

expressions frequently used in text and requires operating on perception-inspired concepts which, when learned in contextual sets, can be used in learning from whole sets allowing lighter and more explainable processing [1] . Such neurosymbolic approaches have recently gained popularity and can be also utilized in life-long learning cognitive architectures.

Monitoring a state of context and its small changes would cause slight alternation to the semantic primes prompts and continuations. If living beings act on "auto-pilot" at the most of their lives and react only to what surprises them, perception simulations using datasets like ours would be easy to observe and interpret. If we operate on signals or other uninterpretable data, it is difficult to draw straightforward conclusions, but to achieve credible descriptions of perception signals we still are dependent on the hard-to-interpret streams of information. However, in our opinion, it could be easier to concentrate on a small set of semantic categories when dealing with multimodal input. Visibility of acquisition process is a key for understanding reasoning mistakes, hence when we equip foundation models with more reliable sensing devices we need to investigate how the description of the signal is generated and how informative it is to, for example, a language model or cognitive architecture. Although new methods for storing knowledge and retrieving will be proposed sooner or later, the method of more or less sophisticated querying (represented by the *chain of thought prompting* in large language models) on a large scale will be one of the basic approaches. As querying massive amounts of prompts is not realistic, we plan to seek for learning universals and the data inspired by semantic primes is our first step toward this direction of research.

6 Conclusion and Future Work

In this paper we have described a novel dataset containing tacit knowledge in textual form. The data includes sentences in Japanese language with sets of choices which are inspired by semantic primes studied in cognitive linguistics. We presented our idea behind the development and suggested its possible use cases. We also ran a simple experiment to suggest that this dataset could be also used for testing cognitive capabilities of language models. By opening the dataset we hope to inspire a new wave of nativist approaches to artificial agents. Allowing an agent to accumulate knowledge in a novel way could shift the machine learning paradigm from the statistical word-matching to semantic primitives-based simulation of the world which in theory could be less resource-consuming and more explainable than currently dominant transformer-based methods for achieving AGI. But to confirm if this approach is reliable and really leads to better understanding of processes in artificial (and maybe biological) entities, better recognition performance and extensive experiments are necessary. To test the data in the wild, we fine-tuned our data with Japanese LLM *open-calm-7b*[10] using all the data and automatically added semantic primes knowledge to sentences of a DanSto story dataset in Japanese [16]. As for short-term future

[10] `https://huggingface.co/cyberagent/open-calm-7b`

work, we are testing if percepts changing within a story influence understanding and whether our approach can be beneficial for analogy making or for improving knowledge abstraction in long term memory of a cognitive architecture.

Acknowledgments. This work was supported by JSPS KAKENHI Grant Number 22K12160 and by JST, CREST Grant Number JPMJCR20D2, Japan.

References

1. Cambria, E., Malandri, L., Mercorio, F., Mezzanzanica, M., Nobani, N.: A survey on xai and natural language explanations. Information Processing & Management **60**(1), 103111 (2023). https://doi.org/https://doi.org/10.1016/j.ipm.2022.103111, https://www.sciencedirect.com/science/article/pii/S0306457322002126
2. Chomsky, N.: Aspects of the theory of syntax. M.I.T. Press, Cambridge, Massachusetts (1965)
3. Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., Stoyanov, V.: Unsupervised cross-lingual representation learning at scale. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 8440–8451. Association for Computational Linguistics, Online (Jul 2020). https://doi.org/10.18653/v1/2020.acl-main.747, https://aclanthology.org/2020.acl-main.747
4. Dentella, V., Murphy, E., Marcus, G., Leivada, E.: Testing ai performance on less frequent aspects of language reveals insensitivity to underlying meaning. arXiv preprint arXiv:2302.12313 (2023)
5. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding. 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). pp. 4171–4186. Association for Computational Linguistics, Minneapolis, Minnesota (Jun 2019). https://doi.org/10.18653/v1/N19-1423, https://www.aclweb.org/anthology/N19-1423
6. Goddard, C.: Overcoming the linguistic challenges for ethno-epistemology: NSM perspectives. In: Ethno-Epistemology, pp. 130–153. Routledge (2020)
7. Goddard, C., Wierzbicka, A.: Meaning and universal grammar: Theory and empirical findings, vol. 1. John Benjamins Publishing (2002)
8. Goertzel, B., Looks, M., Heljakka, A., Pennachin, C.: Toward a pragmatic understanding of the cognitive underpinnings of symbol grounding. In: Gudwin, R., Queiroz, J. (eds.) Semiotics and Intelligent Systems Development. Idea Group (2007)
9. Huang, S., Dong, L., Wang, W., Hao, Y., Singhal, S., Ma, S., Lv, T., Cui, L., Mohammed, O.K., Patra, B., Liu, Q., Aggarwal, K., Chi, Z., Bjorck, J., Chaudhary, V., Som, S., Song, X., Wei, F.: Language is not all you need: Aligning perception with language models. Tech. rep., arXiv:2302.14045 (2023)
10. Jackendoff, R.S.: Semantics and cognition, vol. 8. MIT press (1985)
11. Katsumata, Y., Takeshita, M., Rzepka, R., Araki, K.: Dataset construction for predicting danger degree due to contextual changes (in japanese). In: In Proceedings of The 28th Annual Meeting of The Association for Natural Language Processing (NLP-2022) (2022)

12. Katz, U., Geva, M., Berant, J.: Inferring implicit relations in complex questions with language models. In: Findings of the Association for Computational Linguistics: EMNLP 2022. pp. 2548–2566. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (Dec 2022), https://aclanthology.org/2022.findings-emnlp.188
13. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692 (2019)
14. Mahowald, K., Ivanova, A.A., Blank, I.A., Kanwisher, N., Tenenbaum, J.B., Fedorenko, E.: Dissociating language and thought in large language models: a cognitive perspective. arXiv preprint arXiv:2301.06627 (2023)
15. Pinker, S.: The Blank Slate: The Modern Denial of Human Nature. Penguin Group USA (2003), http://books.google.co.jp/books?id=7rJ5gI1LbXoC
16. Rafal, R., Kacper, D., Arisa, A., Kenji, A.: Dansto - japanese dataset of short stories for evaluating context understanding. JSAI Technical Report, Type 2 SIG **2023**(AGI-026), 32–40 (2024)
17. Riemer, N.: Reductive paraphrase and meaning: A critique of wierzbickian semantics. Linguistics and philosophy **29**, 347–379 (2006)
18. Rutherford, A., Xue, N.: Improving the inference of implicit discourse relations via classifying explicit discourse connectives. In: HLT-NAACL. pp. 799–808 (2015)
19. Rzepka, R., Araki, K., Tochinai, K.: Bacterium lingualis – the web-based commonsensical knowledge discovery method. In: Grieser, G., Tanaka, Y., Yamamoto, A. (eds.) Discovery Science. Lecture Notes in Computer Science, vol. 2843, pp. 460–467. Springer (2003)
20. Samsonovich, A.V.: Extending cognitive architectures. In: Chella, A., Pirrone, R., Sorbello, R., Jóhannsdóttir, K.R. (eds.) Biologically Inspired Cognitive Architectures 2012. pp. 41–49. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
21. Sap, M., Le Bras, R., Allaway, E., Bhagavatula, C., Lourie, N., Rashkin, H., Roof, B., Smith, N.A., Choi, Y.: Atomic: An atlas of machine commonsense for if-then reasoning. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 33, pp. 3027–3035 (2019)
22. Speer, R., Chin, J., Havasi, C.: Conceptnet 5.5: An open multilingual graph of general knowledge (2017), http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14972
23. Talmor, A., Tafjord, O., Clark, P., Goldberg, Y., Berant, J.: Teaching pretrained models to systematically reason over implicit knowledge. arXiv preprint arXiv:2006.06609 **4**(6) (2020)
24. Talmy, L.: Lexicalization patterns: Semantic structure in lexical forms. Language typology and syntactic description **3**(99), 36–149 (1985)
25. Van Durme, B.D.: Extracting implicit knowledge from text. University of Rochester (2009)
26. Wang, A., Pruksachatkun, Y., Nangia, N., Singh, A., Michael, J., Hill, F., Levy, O., Bowman, S.R.: Superglue: A stickier benchmark for general-purpose language understanding systems. arXiv preprint arXiv:1905.00537 (2019)
27. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E., Le, Q., Zhou, D.: Chain of thought prompting elicits reasoning in large language models. arXiv preprint arXiv:2201.11903 (2022)
28. Wierzbicka, A.: Semantic Primitives. (Frankfurt/M.) Athenäum-Verl. (1972)
29. Wierzbicka, A.: Semantics: Primes and Universals. Oxford University Press (1996)