Replacing Sensors with Text Occurrences for Commonsense Knowledge Acquisition

Rafal Rzepka, Marek Krawczyk and Kenji Araki
Graduate School of Information Science and Technology
Hokkaido University, Sapporo, Japan
{rzepka,marek,araki}@ist.hokudai.ac.jp

Abstract

In this position paper we introduce our ideas for utilizing text corpus for supplementing or replacing sensing devices by word pairs occurrences in a blog corpus and we share results of preliminary experiments. We explain how simple web-mining and information retrieval techniques could help to acquire perceptional knowledge while collecting, processing and combining inputs from actual sensors is still technically difficult, expensive or simply impossible. By utilizing co-occurrences rankings of input nouns and five senses identifying words (mostly adjectives and onomatopoeias) we were able to discover physical characteristics of these nouns. Although the method is in early stage of development and the recall is low, achieved precision of 0.96 seems promising, especially if one needs to store acquired output in knowledge bases as ConceptNet.

1 Introduction

Equipping a robot[1] with the best available sensors is expensive but making sense of the input streams of zeros and ones outreaches capabilities of any existing machine learning algorithms. Although latest results of deep learning show that automatic labeling of images [Karpathy and Fei-Fei, 2015] in natural language becomes more and more effective, we are still far from combining visual input with other types of cognitive perception. But even when the means of acquiring knowledge achieve (and eventually surpass) human level, it would be convenient if at least the output of reasoning based on these recognitions is conveyed in natural language to able users to have a control of what was learned and to know why some decision were made – simulations in machine's mind would become more accessible and it would be easier to track errors in the "thinking" process. Having this role of natural language in mind, we treat NLP as a useful toolset for supporting (or replacing when unavailable) cognitive experiences of machines. We do not claim this is the best possible approach, we suggest that currently text-based cognition simulation might become a playground for future real-world systems. We explain this in more details in the next subsection.

1.1 Shortcut for the Five Rings of Artificial Cognition

One can picture the process of achieving computational intelligence in five abstraction layers[2] as shown in Figure 1. The outermost, biggest layer is the world of stimuli, the physical world we live in. It is being sensed by various devices, often ones beyond human capabilities as infrared, radars, x-rays, etc. All input of zero/one strings is then usually abstracted (translated or interpreted) for further utilisation, and it can be done by various means like manual labeling or automatic clustering, depending on a given task. However, we believe that the abstraction step should be performed in natural language because a) it gives us control on what is learned and used for reasoning and b) it preserves the ambiguities of the real world, which makes AI experimenting more realistic, especially when we deal with human-like behavior in everyday life situations.

The above mentioned abstractions are then stored and become searchable, which can be utilized by the innermost layer which symbolizes the quintessence of artificial intelligence research. The core of our idea is to temporarily replace the Storage Ring with results of text mining (see Figure 2). With this approach a robot can conclude that fire is hot without sensing the temperature, that a stone is hard without touching it, that Mars is red without looking at it, etc. Traversing the Web can be treated as limited experiencing the real world and we claim that for time being the collected experiences can often replace the physical ones and in the future the machines could add them to the real cognition processes to confront their knowledge with a state they newly encounter. Not finding "hot floors" online can for example suggest that just discovered high temperature of the floor is an abnormal situation and sense a danger.

1.2 Approaches for Common Sense Knowledge Acquisition

The beginning of this century brought us an abundance of statistical methods which could be applied to massive sets of

[1] For example a companion robot for eldercare that needs both sophisticated sensors and language understanding capabilities.

[2] We call them "rings" loosely after Japanese classic "The Book of Five Rings" by Mushashi Miyamoto.

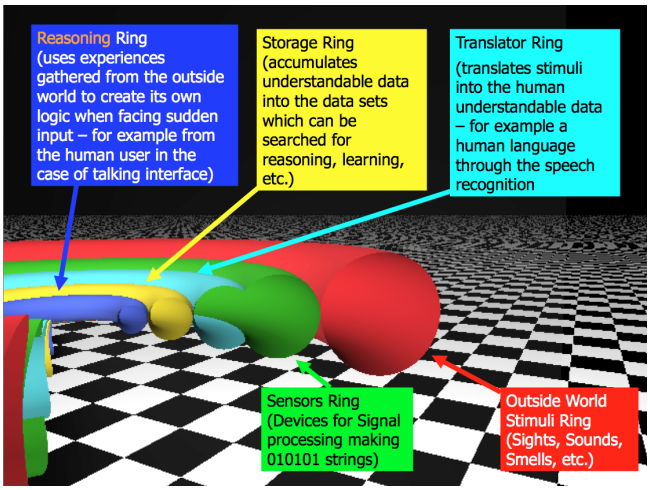

Figure 1: Hypothetical five rings of cognitive processing for artificial intelligence.

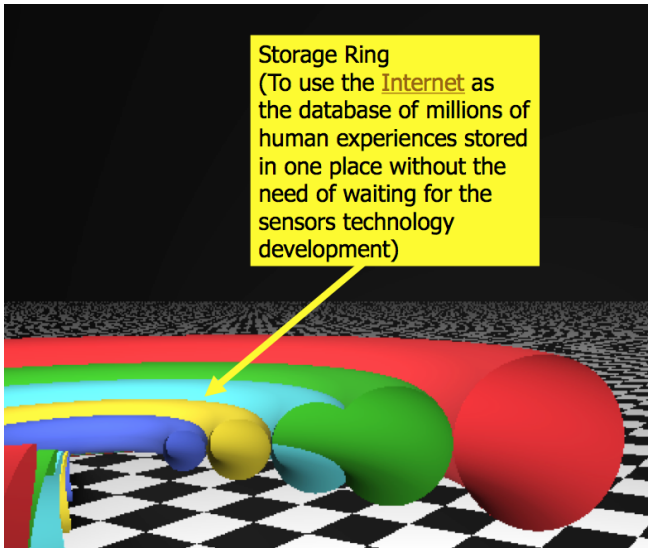

Figure 2: The proposed shortcut – to supplement lacking or insufficient sensor input with web-mined knowledge.

data. Approaches as on-line learning [Bottou, 1998] or active learning [Settles, 2009] became popular with this so called Big Data era, however, this usually means that the quality of feedback becomes crucial for improvement and there are situations where the amount of the necessary additional human knowledge exceeds usability thresholds of a program. One of such cases is utilizing common sense knowledge, which is too fluid and too broad to be easily stored or used as a support for processing real-world data. There are projects for collecting and using such data as Cyc [Lenat and Guha, 1989], ConceptNet [Liu and Singh, 2004], KNEXT [Van Durme and Schubert, 2008], DART (Discovery and Aggregation of Relations in Text) [Clark and Harrison, 2009], PRISMATIC [Fan *et al.*, 2010], NELL – "Never Ending Language Learner" [Carlson *et al.*, 2010], YAGO [Suchanek *et al.*, 2007], KnowItAll [Etzioni *et al.*, 2004], TextRunner [Yates *et al.*, 2007] but the latter four, along with sources as Freebase[3] and DBpedia[4], tend to concentrate on factoids and rarely provide knowledge about basic relations of physical, social or emotional worlds. ConceptNet, which is lately also bond to other sources as WordNet [Miller, 1995] or Wikipedia[5], is based on crowd-made Open Mind Common Sense [Singh *et al.*, 2002] that contain more everyday, non-factoid entries. Still, the human imagination, even in the collective version, is not sufficient to manually input knowledge even for basic features of objects like possible colors or size ranges. The knowledge that is needed varies from basic entries as "cars can be red", "faces can be red", "carpets can be red" to more specific ones which are needed for context analysis: "luxury cars are often red", "ashamed people's faces can become red" or "famous people often walk on red carpets during big events". To acquire such knowledge it would clearly be preferable to use an approach similar to NELL, which automatically retrieves the knowledge from the Internet resources. But when asked about "red car" , NELL's output shows that it is a vehicle, "red face" is not in the database and "red carpet" shows a weak relation to a word "model" (all of these connotations are labelled as a "weak confidence"). We were able to find physical properties of nouns as "pizza" in KNEXT, NELL or DART (but not in Cyc nor ConceptNet), and although sensing connotations (mostly knowledge about sizes, less about hardness or smell) are rare, these knowledge bases could be used for implementing our approach for English language. However, what we need for deeper context understanding, are concrete examples of real world situations with different actors, objects places and other circumstances for further comparison and analysis which is not possible with current world knowledge datasets. In the next section we show how the knowledge can be automatically extracted and how we could avoid mistakes that could not be avoided with the current Japanese ConceptNet entries we work with.

2 Proposed Methods for Replacing Sensor Input with Text

Because our main interest lies in the text understanding for broad range of tasks from unethical behavior discovery [Rzepka and Araki, 2012], [Ptaszynski *et al.*, 2010], metaphors creation and recognition [Dybala *et al.*, 2013] to Winograd Schema Challenge [Levesque, 2011], we work on discovering as many types of cognitive perception simulation as possible, because we believe all of them are important for understanding context and its changes. These types can be divided into three basic categories – social, physical and temporal. Our work on temporal category, presented in [Krawczyk *et al.*, 2013], is in very early stage of development, in this paper we introduce social and physical categories which utilize similar methods described in the next subsections. Currently we work only on Japanese since this language seems to have an easier structure for computer processing especially

[3] http://freebase.com/
[4] http://dbpedia.org/
[5] http://wikipedia.org/

because of its particles usage what Fillmore has suggested in his works [Fillmore, 1968]; another reason are the cultural differences that influence common sense depending on the language.

2.1 Simulating Social World Perception

One of the first tasks we worked on is to discover emotional reactions of people. As all experiments are currently performed within only one culture (Japanese), adequate emotion classification was chosen. Nakamura [Nakamura, 1993] has proposed ten categories of emotions (joy / delight, anger, sorrow / sadness, fear, shame / shyness / bashfulness, liking / fondness, dislike / detestation, excitement, relief and surprise / amazement) which contain thoroughly collected words and phrases for each category from Japanese literature. We use this lexicon for estimating average emotional consequences of acts, which allows our system to easily see that hitting a friend is completely different happening from hitting, e.g. own knee. We also use lexicon based on Kohlberg and his theory of moral development [Kohlberg, 1981], where words and phrases are chosen and divided in ten categories – scolding, praises, punishment / penalization, rewards / awards, disagreement, agreement, illegal, legal, unforgivable, forgivable. Words in these categories allow our system to extract average social consequences and their weight, for example stealing an apple causing more harm than stealing a car [Rzepka and Araki, 2012]. Then we use these phrases for querying 12 millions entries from Ameba blog corpus which consists of approximately 5.5 billion words and was created by [Ptaszynski *et al.*, 2012]. This offline source allows us searching without limitations set by commercial search engines, but its relatively small size limits recall and limits the length of inputs to single acts as "killing a dolphin" in the moral consequences discovery. Using morphological parser MeCab[6] such input is divided into a triplet of a noun, modifying particle, and a verb $N - P - V$. Then our system generates 15 conditional suffixes which create 15 $N-P-V_{if}$ queries are used to retrieve related blog entries. After eliminating emoticons, bracketed explanations too short or too long sentences[7] every sentence is spaced into chunks by semantic role tagger ASA [Takeuchi *et al.*, 2010]. This is done in order to simplify further semantic analysis as negation discovery process. In the next step, previously described phrases from lexicons are matched and counted to create polarity rankings. For example if a sentence contains a only a punishment-related word then the negative count for e.g. "stealing a car" increases. The total polarities are used in the final judgment and a experience candidate is made. We set following restrictions for the matching process:

- Lexicon phrase is matched only if it appears after the $N - P - V_{if}$ query. This is to avoid counting polarity phrases before contrasting conjunctions as "I used to be *happy* **but** then I *became ill* and lost my job" when matching "becoming ill".
- If a lexicon phrase is found in a chunk that contains a negation, it is not counted.
- If there are duplicate sentences in a blog entry, only one is processed.

The accuracy for this method achieved Knext% and the details, originally introduced in [Rzepka and Araki, 2012], are as follows. We asked seven Japanese information science students (22-29 years old, 6 males and one female) to rate input actions on an 11-point morality scale where -5 is the most immoral and +5 is the most moral. Except assigning 0 as "no ethical valence", subjects could also mark an act as "context dependent" because many (if not most) of our behaviors can be treated differently depending on context. If marked both "no ethical valence" and "context dependent", act interpretation was "ambiguous" (AMB) for the sake of easier processing. On 68 evaluations, there were only two disagreements between subjects (when evaluating "revenging oneself" and "going to a love hotel"). After analyzing the data, we decided to count an action as a negative when an average mark was below -2.5 and as a positive when it was above +2.5. Scores between -2.5 and +2.5 were treated as "ambiguous" (AMB). These ambiguous acts are problematic because they heavily depend on context and show how different attitude toward a survey a subject can have. Some of them treated the inputs lightly and used common associations (e.g. "driving a car" is for commuting or giving oneself and other people pleasure, so should be considered moral), others tended to imagine negative sides of acts (e.g. "driving a car" can surely cause harm to people). There were also subjects who always thought about two sides of an act. They seemed to assume that because there are people in the world who think that "eating pork" is unethical, it is safer to mark "eating a pig" as morally ambiguous. Because such evaluations get scattered throughout the scale, we decided to treat neighboring agreements as semi-correct, i.e. when most of the subjects evaluated something as bad and the system chose ambiguity, we counted it as 0.5, a value between full agreement (subjects' "bad" evaluated as "bad" and "good" as "good") which gets 1 point, and full disagreement ("explicit error") where the system judged an act as "good" while it was "bad" for most of the subjects (0 points). The newly added method and its results are described in the next subsection.

2.2 Simulating Physical World Perception

For these experiments we also used five billion sentences corpus [Ptaszynski *et al.*, 2012] and the algorithm was almost identical as the one used for discovering emotions in the previous subsection. Naturally the manually crafted lexicons were different. To test all the five senses simulations we have used 127 action phrases as "(to) call a doctor" or "(to) steal a car" based on the set we used for our unethical behavior discovery task as we aim at acquiring deeper knowledge about situations in which people did something bad. We added to the original list more everyday life actions like "writing a book" or states like "someone laughing" and removed similar entries, for example ones needed for comparing reactions to killing various kinds of animals, which in most cases have very low hit rate in blogs. All input phrases kept their Noun-

[6]MeCab: Yet another part-of-speech and morphological analyzer. `https://code.google.com/p/mecab/`

[7]Longer than 30 and shorter than 250 bytes, a range set experimentally.

Particle-Verb structure and the sensing words (mostly adjectives) are exact matches for $SenseWord - Noun$ query.

Eyes Input Simulation
For the visual input simulation we chose a set of basic descriptive adjectives (See Table 1) shapes and colors, which is the biggest set of all five. A "commonsense border" of 10 hits was experimentally set to limit "peculiar cases" as "cold sun" or "black snow".

Fingers Input Simulation
Perfect haptic sensors are probably the most difficult ones to develop, however sense of touch is an important natural tool for keeping our everyday life safe and knowledge if somebody was hit with a soft or hard objects helps to estimate the damage. For this input simulation we used adjectives and onomatopoeias which are associated with surface characteristics, shown in Table 2.

Ears Input Simulation
Firstly, it must be clarified that in "ears input" we do not include speech recognition but only sound characteristics. As in the case of touch, only adjectives do not bring sufficient hits, therefore we collected also mimetics (*gitaigo*) and onomatopoeias (*giseigo*) to broaden the search. The set is shown in Table 3.

Tongue and Nose Input Simulations
Probably there is no need to place a tasting sensors in robot's mouth because they could be also located in its fingers. But the input should be interpreted the same way as human's and the recognition output should be one of this category sense words. The same situation is with smelling sensors. The taste and smell ones shown in Table 4.

2.3 Physical Sense Word Sets Efficiency

Ninety nouns were extracted from the 127 input phrases and input to the proposed perception simulator. Then the first author evaluated the output. Retrieval precision appeared to be high achieving 0.96, but the recall[8] was rather low achieving 0.43 which gives f-score of 0.60. There was only one error for 30 retrievals caused by metaphoric description of a *novel* - "deep". Possible reasons for the very low recall were as follows. Firstly, the blog data used were searched by whole phrases with "if" statements, because we focus on cause-effect relationship in our research and always tend to use full sentences to avoid less meaningful blog entries. The second reason that was visible was simple too low number of sense words. Another problem was deciding if a noun is physical instance or not. For example word "hito" (*a man*) was used in the examples in the meaning of *somebody* and a "shōsetsu" (*a novel*) was used in a category of a book contents rather than of a physical book. The same problem was with proper nouns as a city name which be treated as a common description for group of physical instances as buildings, streets, monuments, etc., but it is not obvious that you can touch Tokyo (however one could smell it literally or hear the sound of it).

[8]Relatively small corpus caused that many nouns just did not co-occur with sensing keywords.

2.4 Repeating Experiments With Other Nouns

For the second trial we have input words from category "tools, decorations, monuments, etc." of a thesaurus [NINJAL, 2005] assuming that it consists of words describing physical artifacts. To limit errors caused by lack of dependency parsing and nouns used for describing other nouns (i.e. "black snow shoes", Japanese particles were added to nouns in all the queries (direct object *wo*, theme / topic *ga* and *wa*, object *ni* and instrument / location *de*). There were 289 words in the category and we used them for repeating the textual sensing experiment. The system again reached a high precision of 0.92 but this time the recall dropped drastically to 0.08 which resulted in very low f-score of 0.15. Used thesaurus had too many rare words that never appeared in the blog corpus. When we limited the dataset to nouns which had occurred at least 10 times in the blog corpus, the results were much better: recall increased to 0.42 but the f-score of 0.57 was still lower than the results for the smaller set.

3 Conclusions and Future Work

The main goal of this research is to try to simulate the human's perceptual process using mere text and see in the future if such approach can compete or support traditional machine learning approaches, because it is more than reusing represented knowledge than learning it. After testing the non-physical perception simulations with affect analysis techniques, in this paper we newly tested our approach with objects which can be directly perceived with five senses. It is much easier for a machine to guess physical features of the observed world objects when more sensors than one are working at once, also the acquired knowledge becomes richer. Our proposed simulation system is able to perform such task and as it needs as rich sensory input as possible, the achieved recall is too low. However, if we want to store the newly acquired knowledge in e.g. ConceptNet, the achieved precision of 0.92 ensures high quality of novel entries. The next step is to increase the range of search and probably the number of sense words. The latter is not obvious because our experiments with altering Nakamura lexicon for emotional features discovery showed, that simple enlarging a lexicon might cause a drop in precision. Because there is quite a possibility that rare words will be fed to the system, some kind of categorization is needed – if a "snake" category member "anaconda" has not enough hits, other members of the snake category could bring more knowledge about the object with inefficient description. To ensure that we are dealing with a physical, not metaphorical world, we would also need to consider filtering the sense words to avoid ambiguities as in case of word "high" which in Japanese can also mean "expensive". We noticed that we unconsciously set several adjectives (for example "soft" (*yawarakai*)), in more than one sensor lexicon. This time we have left such words unaltered, however we plan to experiment with more specific queries as "sounds soft", "looks soft", "feels soft", etc. We also need to use a named entity recognition algorithm or classifying method which should help to limit erroneous sensing due to counting e.g. companies with common objects in their names.

Table 1: Thirty six sense words for recognizing visual characteristics.

Japanese	akarui, kurai, ōkii, chiisai, akai, kiiroi, kuroi, shiroi, aoi, midori-no, asai, fukai, usui, ōi, sukunai, osoi, hayai, kagayakashii, chairo, nagai, furui, makkuro-na (3 types of writing), kitanai, kirei-na (2 types), shikakui, chikai, tōi, takai, hikui, hiroi, semai, futoi, hosoi
English	bright, dark, big, small, red, yellow, black, white, blue, green, shallow, deep, thin, large, small, slow, fast, shiny, brown, long, old, pitch black, dirty, clean, square, close, distant, high, low, wide, narrow, thick, thin

Table 2: Thirty four sense words for recognizing haptic characteristics.

Japanese	katai (3 types), betabeta suru, tsumetai, yawaraka, omoi, karui, suzushii, asai, fukai, attakai (2 types), atsui (3 types), itai, usui, samui, surudoi, togatte iru, eiri-na, nurui, kawaita, nureta (2 types), nebaneba shita, nurunuru shita, jimejime shita, munmun suru, zarazara shita, tsurutsuru shita, gotsugotsu shita, fuwafuwa shita, guchagucha shita
English	hard, tough, sticky, cold, soft, heavy, and light, cool, shallow, deep, warm (2 types), hot (3 types), thick, painful, pale, cold, sharp, pointed (sharp), tepid, dry, wet (2 types), slimy, gluey, steamy, sultry, rough, slick, rugged, fluffy, squashy

In this paper we introduced several ideas for web-based lexical support for machines that learn by borrowing human experiences rather than experiencing them themselves. The reason for taking this approach is that gaining various experiences by a robot (or a group of robots) in different environments would demand a lot of money and time. We assume that senses can complete each other in order to enrich the intelligence and we were inspired by a phenomenon showing that spatial knowledge of blind children keeps growing even without visual input [Kielkopf, 1968][Fletcher, 1980]). We believe that treating experiences of others could become a building material for creating real-world simulators and release the AI agents from the limited environments of small tasks where unexpected things rarely happen.

References

[Bottou, 1998] Léon Bottou. Online Algorithms and Stochastic Approximations. In *Online Learning and Neural Networks*. Cambridge University Press, 1998.

[Carlson *et al.*, 2010] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E.R. Hruschka Jr., and T.M. Mitchell. Toward an architecture for never-ending language learning. In *Proceedings of the Conference on Artificial Intelligence (AAAI)*, pages 1306–1313. AAAI Press, 2010.

[Clark and Harrison, 2009] Peter Clark and Phil Harrison. Large-scale extraction and use of knowledge from text. In *Proceedings of the Fifth International Conference on Knowledge Capture*, K-CAP '09, pages 153–160, New York, NY, USA, 2009. ACM.

[Dybala *et al.*, 2013] Pawel Dybala, Rafal Rzepka, Koichi Sayama, and Kenji Araki. Detecting false metaphors in japanese. In *Proceedings of The 6th Language and Technology Conference - LTC 2013*, pages 127–131, 2013.

[Etzioni *et al.*, 2004] Oren Etzioni, Michael Cafarella, Doug Downey, Stanley Kok, Ana-Maria Popescu, Tal Shaked, Stephen Soderland, Daniel S. Weld, and Alexander Yates. Web-scale information extraction in knowitall: (preliminary results). In *Proceedings of the 13th International Conference on World Wide Web*, WWW '04, pages 100–110, New York, NY, USA, 2004. ACM.

[Fan *et al.*, 2010] James Fan, David Ferrucci, David Gondek, and Aditya Kalyanpur. Prismatic: Inducing knowledge from a large scale lexicalized relation resource. In *Proceedings of the NAACL HLT 2010 first international workshop on formalisms and methodology for learning by reading*, pages 122–127. Association for Computational Linguistics, 2010.

[Fillmore, 1968] Charles J. Fillmore. The case for case. In Emmon Bach and Robert T. Harms, editors, *Universals in Linguistic Theory*, pages 0–88. Holt, Rinehart and Winston, New York, 1968.

[Fletcher, 1980] Janet F Fletcher. Spatial representation in blind children. 1: Development compared to sighted children. *Journal of Visual Impairment and Blindness*, 74(10):381–85, 1980.

[Karpathy and Fei-Fei, 2015] Andrej Karpathy and Li Fei-Fei. Deep visual-semantic alignments for generating image descriptions. In *To appear In Proceedings of Computer Vision and Pattern Recognition Conference CVPR 2015*, 2015.

[Kielkopf, 1968] Charles F. Kielkopf. The pictures in the head of a man born blind. *Philosophy and Phenomenological Research*, 28(4):501–513, 06 1968.

[Kohlberg, 1981] Lawrence Kohlberg. *The Philosophy of Moral Development*. Harper and Row, 1th edition, 1981.

[Krawczyk *et al.*, 2013] Marek Krawczyk, Yuki urabe, Rafal Rzepka, and Kenji Araki. A-dur: Action duration calculation system. Technical Report SIG-LSE-B301-7, pp.47-54, Technical Report of Language Engineering Community Meeting, 2013.

[Lenat and Guha, 1989] Douglas B. Lenat and R. V. Guha. *Building Large Knowledge-Based Systems; Representation and Inference in the Cyc Project*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989.

Table 3: Thirty one sense words for recognizing sound characteristics.

Japanese	urusai, shizuka-na, pokipoki suru, zāzā suru, biribiri suru, gorogoro suru, bukubuku suru, kasakasa suru, katakata suru, gatagata suru, gachagacha suru, gachan suru, karan suru, gangan suru, kiikii suru, gōngōn suru, sarasara suru, janjan suru, charin suru, chirin suru, chirinchirin suru, chinchin suru, tonton suru, batan suru, patan suru, pachipachi suru, pachipachi suru, pachin suru, pishari suru, rinrin suru
English	noisy, quiet, crack, rushing water, tear, rumble, bubble, rustle, rattle, clatter, clank, pound, squeak, purr, murmur, continuous sound, clink, tinkle, jingle, ding, knock, bang, snap, crackle, clack, smack, ring

Table 4: Fourteen sense words for recognizing taste and 6 for smell characteristics.

Japanese (taste)	amai (2 variations), nigai, suppai (2 variations), karai (2 variations), shoppai, nigai, shibui (2 variations), amazuppai, shibō-no ōi, aburakkoi
English	sweet (2 types), bitter, sour (2 types), spicy (2 types), salty, bitter, astringent (2 types), sweet and sour, fatty, greasy
Japanese (smell)	kusai, kōbashii, namagusai, kaguwashii, ii nioi-no, ii kaori-no
English	smelly, aromatic, fishy, fragrant, of a nice smell, smelling nice

[Levesque, 2011] Hector J. Levesque. The winograd schema challenge. In *AAAI Spring Symposium: Logical Formalizations of Commonsense Reasoning*. AAAI, 2011.

[Liu and Singh, 2004] Hugo Liu and Push Singh. Conceptnet: A practical commonsense reasoning toolkit. *BT Technology Journal*, 22:211–226, 2004.

[Miller, 1995] George A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38:39–41, 1995.

[Nakamura, 1993] Akira Nakamura. *Kanjo hyogen jiten [Dictionary of Emotive Expressions]*. Tokyodo Publishing, 1993.

[NINJAL, 2005] NINJAL. *Bunrui Goi Hyou Thesaurus*. National Institute for Japanese Language, 2005.

[Ptaszynski *et al.*, 2010] Michal Ptaszynski, Pawel Dybala, Tatsuaki Matsuba, Fumito Masui, Rafal Rzepka, Kenji Araki, and Yoshio Momouchi. In the service of online order: Tackling cyber-bullying with machine learning and affect analysis. *International Journal of Computational Linguistics Research*, 1(3):135–154, 2010.

[Ptaszynski *et al.*, 2012] Michal Ptaszynski, Rafal Rzepka, Kenji Araki, and Yoshio Momouchi. Annotating syntactic information on 5 billion word corpus of japanese blogs. In *In Proceedings of The Eighteenth Annual Meeting of The Association for Natural Language Processing (NLP-2012)*, volume 14-16, pages 385–388, 2012.

[Rzepka and Araki, 2012] Rafal Rzepka and Kenji Araki. Language of emotions for simulating moral imagination. In *Proceedings of The 6th Conference on Language, Discourse, and Cognition (CLDC 2012)*, May 2012.

[Settles, 2009] Burr Settles. Active learning literature survey. Technical Report 1648, University of Wisconsin–Madison, 2009.

[Singh *et al.*, 2002] Push Singh, Thomas Lin, Erik T Mueller, Grace Lim, Travell Perkins, and Wan Li Zhu. Open mind common sense: Knowledge acquisition from the general public. In *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, pages 1223–1237. Springer, 2002.

[Suchanek *et al.*, 2007] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 697–706, New York, NY, USA, 2007. ACM.

[Takeuchi *et al.*, 2010] Koichi Takeuchi, Suguru Tsuchiyama, Masato Moriya, and Yuuki Moriyasu. Construction of argument structure analyzer toward searching same situations and actions. Technical Report 390, IEICE technical report. Natural language understanding and models of communication, jan 2010.

[Van Durme and Schubert, 2008] Benjamin Van Durme and Lenhart Schubert. Open knowledge extraction through compositional language processing. In *Proceedings of the 2008 Conference on Semantics in Text Processing*, pages 239–254. Association for Computational Linguistics, 2008.

[Yates *et al.*, 2007] Alexander Yates, Michael Cafarella, Michele Banko, Oren Etzioni, Matthew Broadhead, and Stephen Soderland. Textrunner: Open information extraction on the web. In *Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, NAACL-Demonstrations '07, pages 25–26, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.