Web-Mining Defeasible Knowledge from Concessional Statements

Alina Petrova[1] and Sebastian Rudolph[2]

[1] University of Oxford
[2] Technische Universität Dresden

Abstract. Mining common-sense knowledge is a vital problem of artificial intelligence that forms the basis of various tasks, from information retrieval to robotics. There have been numerous initiatives to mine common-sense facts from unstructured data, more specifically, from Web texts. However, common-sense knowledge is typically not explicitly stated in the text, as it is considered to be obvious, self-evident, and thus shared between writer and reader. We argue that certain types of *defeasible* common-sense knowledge (i.e., knowledge that holds in most but not all cases), in particular, beliefs and stereotypes, tend to appear in text in a particular manner: they are not explicitly manifested, unless the speakers encounter a situation that runs in contrast to their defeasible common-sense assumptions. For example, if a speaker believes that Spain is a very warm country, she may express a surprise when it snows in Bilbao. We further argue that such conceptual contradictions correspond to the linguistic relation of concession (e.g., *although Bilbao is in Spain, it is snowing there today*) and we present a methodology for extracting defeasible common-sense beliefs (*it is not common to snow in Spain*) from Web data using concessive linguistic markers. We illustrate the methodology by mining beliefs about persons and we show that we are able to extract new information compared to existing common-sense knowledge bases.

1 Introduction

Common-sense knowledge is a set of basic propositions of a very broad semantics that describe different classes and instances, their most common properties (e.g., shape, color, material, frequency, age) and how they relate to each other. For example, the following statements belong to the realm of common-sense knowledge: *snow is white*, *London is in England*, *July is the seventh month of a year*, *children often believe in Santa Claus*. From a human-oriented point of view, "a common-sense fact is a true statement about the world that is known to most humans" [13], while from the point of view of formal systems, a common-sense fact is a formalized statement about the world that is shared between all agents and is true across all applications.

Mining common-sense knowledge from a variety of resources is a vital problem of artificial intelligence, it is required for common-sense reasoning which

forms the basis for a plethora of tasks, from more applied ones, such as question answering and item recommendation, to more general ones, such as intelligent decision making, robotics and natural language understanding [1].

For example, let us consider recommender systems. Having information about a given user, e.g., his age, profession or personal traits, the system could use common-sense knowledge about typical preferences of users with the same characteristics and generate additional item suggestions to the user, complementing the ones based on data mining and correlation. Suppose a user states that he is energetic, yet he has not viewed or purchased any sport-related articles. A knowledge base containing a statement *energetic people tend to enjoy sports* could expand the range of items recommended to the user, thus potentially benefiting both the user and the system and overcoming the cold-start problem [6].

Yet, the acquisition and deployment of common-sense knowledge in practical scenario is still underdeveloped: "the lack of common-sense knowledge and reasoning was encountered in many if not all application areas of Artificial Intelligence" [2].

There have been numerous initiatives to collect common-sense facts and to represent them formally, e.g., in the form of RDF triples [7], [12], [13]. The facts are usually mined from unstructured data, more specifically, from Web texts. This requires, however, that common-sense knowledge has to be explicitly stated in the texts, which tends to be not the case. On the contrary: as common-sense knowledge is considered to be obvious, self-evident and shared between writer and reader anyway, it is often not stated since there is no need to convey it. This comes as a consequence of Grice's conversational maxims of quantity and manner: when communicating, one tries to be as brief and concise as possible, and does not contribute more information than is actually required [3]. Humans are barely aware of the plethora of shared common-sense knowledge their communication is implicitly based on. As Davis put it: "Since common sense consists (by definition) of knowledge and reasoning methods that are utterly obvious to us, we often overlook its astonishing scope and power" [1].

We argue that, therefore, certain types of *defeasible* common-sense knowledge (i.e., knowledge that holds in most but not all cases), in particular, beliefs and stereotypes, tend to appear in text only in an indirect manner: they are not explicitly manifested, unless the speaker encounters a situation that runs in contrast to her defeasible common-sense assumptions. For example, if a speaker believes that Spain is a very warm country, she may express a surprise when it snows in Bilbao. We further argue that this type of conceptual contradictions closely corresponds to the linguistic relation of concession (e.g., *although Bilbao is in Spain, it is snowing there today*). In this paper, we present a methodology and proof-of-concept evaluation for extracting defeasible common-sense beliefs (such as *it is not common to snow in Spain*) from Web data using concessive linguistic markers. We illustrate the methodology by mining beliefs about persons and we show that we are indeed able to extract new information complementing the one in existing common-sense knowledge bases.

2 Preliminaries

Before presenting our methodology of extracting defeasible common-sense beliefs, we provide the cognitive and linguistic justification for our general approach by looking into the nature of defeasible knowledge and the semantics of concession.

2.1 Defeasible Knowledge

An argument is called *defeasible* if it is rationally compelling but not deductively valid [5]. In deductive reasoning, a statement *if p, then q* imposes a constraint that the conclusion q must be true if the premise p is true. In defeasible reasoning, the conclusion is believed to be true given p, but it can potentially be defeated by some additional argument, while p remains to be true, which reflects the non-monotonic nature of defeasible knowledge [9]. We can formulate a defeasible conditional statement in the following way: q is commonly believed to follow from p, although in certain situations this consequence relation may not hold, or in short: *if p, then* normally *q*. Hence, defeasible reasoning may be used to model common beliefs or stereotypes: we believe that q follows from p, but this is not formally proven and may have exceptions [8]. Defeasible knowledge may or may not be supported by empirical evidence.

2.2 Concession

In natural language, there exist numerous discourse relations. A well-studied relation that is commonly encountered in texts is the *opposition relation* [4]. It is a relation with broad semantics that links two contrasting, mutually exclusive items. A particular subtype of opposition relation is *concessive relation*, or *concession*. This relation links two potentially or apparently contrasting items due to an implicit assumption, known also as *default implication*: through this assumed implication, one argument creates an expectation which is then denied in the second argument [10]. For example, "although it is summer, the weather is not warm" is a concessive sentence which relies on the common assumption that it is warm in summer (with respect to the location of the speaker). The fact that it is summer, stated in the first argument, triggers the expectation of warm weather which is then refuted by the second argument. The assumption/expectation may follow logically from the first argument (*Jack got cold* (hence it is assumed he felt it)*, but he did not realize it*), statistically correlate with it (*Nick did not buy any beer, although he was going to watch football with his friends* (and it is very common to have a beer in this situation)) or be based on a property that is commonly related to it (*although she is blonde* (and the stereotype is blonde women are not particularly smart)*, she has a degree in biochemistry*).

Winter and Rimon [14] formalized the semantics of concession in the following way, $\Diamond$ being the possibility operator from classical modal logic: $(p \wedge q) \wedge \Diamond(p \rightarrow \neg q)$, where p is the first argument, q is the second argument and $\neg q$ is the expectation that was triggered by p. As we can see, the semantics of concession

is very much aligned with that of a defeasible implication, and in fact, the default assumption that underlies concessive relation ($p \rightarrow \neg q$) is sometimes called *the defeasible rule*. Therefore, in order to find linguistic representation of common-sense knowledge that is manifested in text in a defeasible way, we will focus on concessive sentences:

▷ linguistic concession:
although p, q

▷ defeasible rule that is implicitly present in concession:
if p, then usually not q

2.3 Concessive Markers

In order to identify and extract concessive statements, it is helpful to determine what are the means of representation of concession in text. There are numerous linguistic studies of concession that analyse types of concessive relations and how they can be expressed in English. For instance, Taboada and Gómez-Gonzáles [11] present an extensive analysis of concessive relation and its *discourse markers* (elements of text that explicitly signal a particular relation). They study both spoken and written texts, in particular, book and movie reviews collected from the Web, which is in line with our setting, as we aim at analysing Web texts.

Concession can be expressed in natural language in multiple ways [11], namely, using subordinate conjunctions, coordinate conjunctions, adverbial items, phrasal expressions, and parenthetical elements. The discourse markers that accompany these means of representation are:

- conjunctions: *although, but, despite the fact that, even though* etc.;
- sentence adverbials: *nevertheless, regardless, yet* etc.;
- gerund constructions: *supposing, granting* etc.;
- prepositions: *in spite of, regardless of* etc.

Despite the heterogeneity of concessive means of representation, the vast majority of them are very rarely encountered in written texts and almost never in spoken texts. The most common concessive marker is the conjunction *although*, with the respective subordinate clause being the most common concessive grammatical construction. In addition, *although* is one of the few markers that is semantically unambiguous, i.e., it can only introduce concession and not other relations. Therefore, it is a very convenient marker to be used in automatic extraction of concessive sentences from text, hence we will use it in our methodology.

3 Methodology

In this section, we propose a way of extracting defeasible common-sense knowledge for a particular domain or topic using concessive markers. The vastly prevailing concessive construction in English is a subordinate clause with the conjunction *although*, which precedes the main sentence, e.g., *Although John studied*

hard for the exam, he failed it. In a nutshell, our approach takes **the pattern *P = "although X, Y"***, instantiates X with a particular value X_i (i.e., some property or category), queries a Web search engine with the first part of the construction *although* X_i, extracts full concessive sentences, and finds corresponding Y_is. The (X_i, Y_i) pairs are then used to generate defeasible statements *if* X_i*, then usually not* Y_i. Below is a step-by-step description of the methodology illustrated by an example that explores a personal category of gender (where the X_is are *male, female* etc.).

1. **Define a category of interest (select X_is)**
   In order to construct instances of the *although*-pattern, one needs to collect particular values for the chosen category. The selection can be done manually or using existing knowledge resources, e.g., WordNet or Wikipedia.

   For example, when interested in gender, one can search the term *gender* in WordNet[3] and collect related terms: *male*, *female*, *man*, *woman* etc.

2. **Build pattern instances P_is**
   Using the general pattern P and the chosen X_is, we construct pattern instances as follows: *although + he/she is (a) +* X_i *+ , + he/she (is)*
   - *he/she* is used as a subject, since the gender category relates to persons;
   - an article *a* is optional and depends on the grammatical category of X_i;
   - the "*, he/she (is)*" part locks the beginning of the main clause; this guarantees that the main clause of the sentence refers to the same subject as the subordinate one, and we are more likely to get a direct opposition between the content of two clauses;
   - if we add *is* to the end of the instance, we are more likely to get another class or category as Y_i; without *is*, the main clause may be of arbitrary content, e.g., action, general description, event, etc.;
   - further variations of the main clause can include *he/she is not, does, does not* etc.

   In our example one possible pattern instance is *although she is a woman, she.*

3. **Query a search engine and crawl results**
   Pattern instances are used as exact queries for a search engine, so that the Web is treated as a text corpus. The search results are then crawled and the text snippets are collected.

   One of the utterances that was crawled for the example query is *Although she is a woman, she is fighting to have high degree education.*[4].

[3] `http://wordnetweb.princeton.edu/perl/webwn?s=gender`
[4] source: `http://novrianfathi.blogspot.co.uk/`

4. **Parse results and form (X_i, Y_i) pairs**
   Using state-of-the-art linguistic processing tools (e.g., Stanford CoreNLP[5]) we extract the Y_i part from the search result snippets:
   - we locate the main clause in the snippet (*clause identification*);
   - since snippets are restricted in size, a snippet may contain only a part of the main clause, in which case it is ignored altogether; this allows us not only to improve parsing, but also to filter out sentences that are too long and are hard to transform into a common-sense statement (e.g., *although she is a woman, she has control over the men in the bar because she is able to beat them at a ...*);
   - we find the predicate of the clause and normalize it;
   - if a sentence contains clauses preceding the *although*-clause (e.g., *She said that although she is ...*), they are removed;
   - if the clause contains additional phrasal and parenthetical elements (*to tell the truth*), they are removed.

   In our example, X_i is *woman* and Y_i is *fight to have high degree education.* Currently we leave the parsed segments as is, but in future, the Y_is can potentially be modified using synonyms (*higher education*), rephrasing (*want to have a degree*), generalization (*study*) etc.

5. **Negate Y_is**
   Argument negation is done to re-construct the defeasible assumption. It can be done using antonyms (*she is brave* vs. *she is fearful*), verb negation (*she is not brave*) or an introductory clause *it is not true that.*

   From our example, the defeasible assumption is: *if she is a woman, then she does not fight for high degree education.*

4 Proof of Concept Evaluation

Suppose we are interested in common-sense facts about the *Person* category. We can collect a set of X_i values manually, or we can as well address external knowledge resources. For example, Wikipedia has a rich network of categories, *People and self* being among the top 12 categories[6]. Subcategories can be a source of relevant aspects of the chosen category (e.g., gender, ethnicity, religion, occupation), as well as of the X_i values (e.g., male, female; Cherokee, Korean; Muslim, pantheist; diplomat, engineer). From these values pattern instances P_is can be constructed.

Let us consider the gender subcategory and the example query Q from Section 3: *although she is a woman, she.* We queried `google.com` with Q[7], collected snippets from the search results, filtered out certain snippets as discussed in

[5] `http://stanfordnlp.github.io/CoreNLP/`
[6] `https://en.wikipedia.org/wiki/Portal:Contents/Categories`
[7] queried on 21.01.2016

step 4 of the methodology and saved the top 50 results. The list of extracted *although*-sentences can be found in the appendix A. With the exception of several sentences that contain references to the context of the original web page (see the discussion section), other *although*-sentences can easily be converted into default implications:

(1) women are typically fearful (antonym of *fearless*), are not skilled at riding (verb negation), poor hunters (antonym to *excellent*), and bad warriors (antonym to *fine*),

(44) women are good at cooking (verb negation),

(32) women do not have male power in their work, etc.

The two main types of sentences are: those describing atypical behavior (*...she is fighting to have high degree education*), and those mentioning uncharacteristic property (*...she is very sensible and smart*). While some sentences convey a narrower context than the others, they all are based on common accounts and stereotypes. Empirically, the shorter the second clause of the sentence, the more concise is the underlying statement and the easier it is to parse it automatically (compare: *she does not have kids* versus *she has never gone through the process of pregnancy and labor and delivery*).

Our methodology is based on utilizing the concessive pattern $P =$ *"although X, Y"*. P can be instantiated in numerous ways and proves to be quite general and versatile. While we demonstrated how it can be used to mine stereotypes about personal categories, we will now illustrate how the pattern instance "*although she is a woman, she*" can be: (a) narrowed down to target specific aspects of the category, and (b) generalized to more general categories. All examples are real-world and are queried using the specified pattern instances.

When narrowing down, we can mine particular common-sense statements that represent preferences, characteristics, actions, behavior by partially specifying the structure of Y_i:

- *...she likes/enjoys/prefers*, *...she does not like* – preferences,
- *...she attends/works/does* – actions, activities,
- *...she speaks/plays/sews* – capabilities etc.

On the other hand, we can generalize pattern instances P_is by utilizing in the X_i the wildcard operator * that acts as a placeholder for any terms: "*although she is (a) *, she*". The resulting concessional statements reflect the most common stereotypes which can involve a female (but are not necessarily bound by female gender): professional (*Although she is a highly qualified graduate, she can't find work.*), personal (*Although she is a responsible adult she has a lot of kid in her and she seems to be having as much fun as the kids.*), religious (*Although she is a devout Christian, she observes the letter, but not the spirit, of the commandment "Honor thy father.*), related to hobbies (*Although she is a slow runner, she has completed 5 marathons!*) etc.

Finally, we can go beyond personal category and mine stereotypes and common-sense beliefs about pretty much any entity, however abstract or concrete:

- *although London is *, it*"
  Example: *Although London is an expensive city, it's also one that has lots of free attractions.*)
- *although math is *, it*
  Example: *Although math is challenging, it becomes easier with practice and a little bit of fun!*
- *although inspiration is *, it*
  Example: *Although inspiration is an amazing thing, it's often temporary and wears off long before our goals are accomplished.*
- *although Pulp Fiction is *, it*
  Example: *Although Pulp Fiction is full of violence, most of it isn't directly shown.*
  etc.

5 Related Approaches

To illustrate the specificity of defeasible common-sense knowledge and the contribution of this work, we will compare our methodology with three existing approaches for acquiring common-sense knowledge.

WebChild [12] is a project that automatically mines noun-adjective pairs from text, connecting concepts represented by nouns with their typical properties represented by adjectives via a set of predefined taxonomic and non-taxonomic relations (*hasShape, hasColor, hasTaste* etc.). The resulting triples are basic, common-sense facts, e.g., *apples are round*. The system has a major drawback compared to our approach: the set of statements is strictly limited by the semantic relations, whereas we are not confined by any particular item-property structure. When queried with the word "woman"[8], WebChild returns a number of basic statements: *woman type_of female*, *woman has_substance tissue* - as well as some more involved properties: *romantic*, *emotional*, *beautiful*. The latter can be viewed as common beliefs or stereotypes, but they do not go beyond category-property scheme, whereas our approach is able to retrieve much more versatile statements, e.g., women are good at cooking.

Verbosity [13] is a common-sense knowledge acquisition approach conducted in a form of a game: one player selects a concept and describes its typical properties without naming the concept itself, while the other player tries to guess the concept. The types of hints the first player can give are restricted to a predefined set of patterns with blanks to be filled in (*it is used for X*, *it is a kind of X*). Patterns considerably facilitate processing of the input sentences and enforce high precision of the acquired statements, but they lack expressivity. Our approach

[8] `https://gate.d5.mpi-inf.mpg.de/webchild/`

is generic and results in much more diverse statements. Another advantage of our approach is its scale: when Verbosity was played by over 250 people during one week, it managed to collect less than 8,000 statements. Using an automated approach, this amount of facts can be generated in a matter of minutes. Interestingly, the authors mention as reason for using gamification rather than simply querying the Web that common-sense statements are "too obvious" to be stated explicitly. As discussed earlier, our approach is precisely designed to overcome this issue.

NELL [7] is a large-scale project of common-sense knowledge harvesting, with a knowledge base of over 80 million statements collected from Web texts. NELL is very expressive and has a massive set of categories and relations, although it should be noted that the majority of statements are made on the level of instances (*ronnie_wood is a musician who is part of rolling_stones*). As in the case of WebChild, NELL is not tailored to mine defeasible knowledge, since the two items can only be linked together if they co-occur in text. Moreover, NELL is a system that is similar to `freebase.com` or `dbpedia.org` in a sense that it collects facts rather than beliefs, thus it targets other types of assertions. We can, on the other hand, complement NELL and other projects by applying our defeasible knowledge mining pipeline over their sets of categories, augmenting their fact base with beliefs, stereotypes and popular opinions (and even common misconceptions).

6 Discussion and Future Work

A number of issues worthy of discussion correspond to particular steps of the proposed methodology. The methodology is a stepwise process, and each step of the pipeline can be customized; in particular, using a better web search crawler (more hits per query) or a better parser (more accurate clause identification and filtering) will result in better overall performance (more well-formed statements).

- **Building pattern instances**
  When a search engine is queried with a pattern *although X*, certain values for *X* tend to generate fewer results that the others. For example, *although she is a woman* returns approx. 325,000 results, while *although she is an astronaut* returns only 3 results[9]. In order to maximize the number of concessive sentences containing the chosen value, one could expand the repository of patterns used for querying, both syntactic (*he is very weak for a footballer* producing a statement *football players are strong*) and semantic ones (*it is a very surprising fact that...*). For now, we stick to *although*-patterns, since they have a very high precision in terms of yielding concessive constructions. Expanding the number of patterns, on the other hand, would increase the recall.

[9] as queried by `google.com` on Feb, 10th 2016

– **Querying a search engine**
Crawling turned out to be one of the implementation challenges of our approach. While Web data is the most comprehensive and useful text resource for the task at hand, the main access point to it is a search engine, and relying on a particular search engine means depending not only on its quality and the volume of indexed resources, but also on its data policies.

– **Parsing results and forming (X_i, Y_i) pairs**
Converting *although*-sentences into concessive pairs is not a trivial task. While we can control the structure of X_is, Y_is can come in any shape and length, and some of them are harder to process than the others. Here are the examples of types of concessional sentences, ranging from "easier" to "harder" ones, together with the issues they contain and the potential remedies:
  - Y_i is short, easy to parse and to convert
    Example: *although she is a woman, she is brave*
  - Y_i contains a reference to a broader textual context
    Example: *although she is a woman, she is prepared to do it*
    The sentence can only be understood if the reference from *it* is resolved
    Solution: Such sentences can potentially be filtered out by using a *coreference resolution* module: when an unresolved item is encountered, the sentence is ignored.
  - Y_i is long, but can be split into several predicates
    Example: *although she is a woman, she is fearless, skilled at riding, an excellent hunter, and a fine warrior*
    Solution: using syntactic parsing, we can divide the main clause into several predicates and corresponding subordinate words and negate them separately.
  - Y_i is long and has a complicated inner structure
    Example: *although she is a woman, she has only a woman's body and a woman's charm without a woman's heart*
    It is not possible to shorten the sentence and get rid of some of its parts, because its meaning can be expressed only when the sentence is taken as one unit.
    Solutions: there are two possible strategies with respect to such sentences. Either a more intricate parser procedures are implemented, or we filter out such sentences, focusing on snippets that can easily be parsed and transformed into defeasible assumptions. The latter take on common-sense knowledge mining seeks to, above all, maintain high precision rather than gain high recall and maximize the number of generated pattern instances.

– **Negating the second argument of the concessive relationship**
One of the most challenging steps in the suggested methodology is to transform an utterance of the form *although p, q* into a structured form *if p, then usually not q* and, in particular, to negate *q*. In multiple cases we need to go

beyond antonyms and verb negation and to perform generalization in order to get meaningful statements. For example, *although he is German, he lives in France* implies that most German people live in Germany, but we cannot generate this statement by simply negating *France*.

Finally, there are several aspects of a general nature. Since languages are characterized by their flexibility and variability, the linguistic relation of concession does not impose strict constraints on the types of arguments it may use. Thus, defeasible beliefs and stereotypes may be as simple as comparing two facts (*X is German, but lives in France*), or as complex as involving several steps of reasoning and further background knowledge (*Although X loves Da Vinci, he did not enjoy Louvre*; $K_{background}$ =*Louvre is a museum; Louvre contains the Mona Lisa painting; Mona Lisa is painted by Da Vinci*). The more implicit reasoning steps are required to understand the sentence, the harder it is to automatically extract the default implication behind it. On the other end of the spectrum, there are sentences that exhibit an apparent contradiction: *Although the ending was a happy one, it was also a little sad.* The *happy-sad* opposition is actually independent of the context and thus is of little interest.

Some sentences extracted by our pipeline cannot be converted into common-sense facts, since they are very context-specific and do not rely upon common-sense knowledge. For example, a sentence *although she is a woman, she does not make any efforts to understand young Hazal's sentiments* uses the default implication that a woman would understand Hazal's sentiments, which is only shared between those who is aware of the broader context. The implication does not hold in general and cannot be viewed as a common-sense knowledge.

One last point to discuss is evaluation and quality insurance of the extracted knowledge. One needs to validate whether a generated default implication is indeed an instance of common-sense knowledge. Automatic evaluation against existing knowledge bases tends to be unreliable, since defeasible knowledge is not well-represented in the latter. A more feasible approach would be human evaluation, with several participants evaluating the same fact and with inter-annotator agreement being calculated for every item.

7 Conclusions

In this paper, we proposed a novel approach to extract common-sense knowledge from textual resources on the Web. Thereby, we overcome the sparsity of explicit occurrences of this type of knowledge by focusing on cases where the common-sense knowledge – being defeasible in nature – is violated. In such cases the inherent "contrariness" of two facts requires an explicit mention.

In linguistic terms, the violation of a common-sense assumption is typically expressed by concessional statements for which a variety of linguistic markers are known.

Given a domain of interest, we are able to systematically search the web for instances of concessive lexico-syntactic patterns and to extract utterances of the

form *although p, q*, which can then be transformed into defeasible common-sense rules *if p, then usually not q*.

We gave an experimental proof of concept for our proposed methodology by extracting common beliefs and stereotypes about people, however, the suggested methodology can be adapted to various conceptual domains (e.g., organizations, events, artifacts) and types of common-sense information (e.g., typical actions, properties, relations).

The work is of exploratory nature: it serves as a proposal and first step toward accessing defeasible beliefs through the semantic relation of concession and its linguistic representation and paves the way for further research in common-sense knowledge acquisition and modelling.

References

1. Davis, E.: Representations of commonsense knowledge. Morgan Kaufmann (2014)
2. Fensel, D.: Ontologies - a silver bullet for knowledge management and electronic commerce. Springer (2001)
3. Grice, H.P.: Logic and conversation. In: Cole, P., Morgan, J. (eds.) Syntax and semantics (vol. 3): Speech acts, pp. 41–58. Academic Press, New York (1975)
4. Izutsu, M.N.: Contrast, concessive, and corrective: Toward a comprehensive study of opposition relations. Journal of Pragmatics 40(4), 646–675 (2008)
5. Koons, R.: Defeasible reasoning. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. `http://plato.stanford.edu/archives/spr2014/entries/reasoning-defeasible/`, spring 2014 edn. (2014)
6. Lam, X.N., Vu, T., Le, T.D., Duong, A.D.: Addressing cold-start problem in recommendation systems. In: Proceedings of the 2nd international conference on Ubiquitous information management and communication. pp. 208–211. ACM (2008)
7. Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Betteridge, J., Carlson, A., Dalvi, B., Gardner, M., Kisiel, B., Krishnamurthy, J., Lao, N., Mazaitis, K., Mohamed, T., Nakashole, N., Platanios, E., Ritter, A., Samadi, M., Settles, B., Wang, R., Wijaya, D., Gupta, A., Chen, X., Saparov, A., Greaves, M., Welling, J.: Never-ending learning. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15). pp. 2302–2310 (2015)
8. Moses, Y., Shoham, Y.: Belief as defeasible knowledge. Artificial intelligence 64(2), 299–321 (1993)
9. Nute, D.: Defeasible logic. In: Gabbay, D., Hogger, C., J.A.Robinson (eds.) Handbook of logic in artificial intelligence and logic programming (vol. 3): nonmonotonic reasoning and uncertain reasoning, pp. 355–395. Oxford University Press, New York (1994)
10. Robaldo, L., Miltsakaki, E.: Corpus-driven semantics of concession: Where do expectations come from? Dialogue & Discourse 5(1), 1–36 (2014)
11. Taboada, M., Gómez-González, M.d.l.A.: Discourse markers and coherence relations: Comparison across markers, languages and modalities. Linguistics and the Human Sciences 6(1-3), 17–41 (2012)
12. Tandon, N., de Melo, G., Suchanek, F., Weikum, G.: Webchild: Harvesting and organizing commonsense knowledge from the web. In: Proceedings of the 7th ACM international conference on Web search and data mining. pp. 523–532. ACM (2014)

13. Von Ahn, L., Kedia, M., Blum, M.: Verbosity: a game for collecting common-sense facts. In: Proceedings of the SIGCHI conference on Human Factors in computing systems. pp. 75–78. ACM (2006)
14. Winter, Y., Rimon, M.: Contrast and implication in natural language. Journal of Semantics 11(4), 365–406 (1994)

A Appendix

Although-sentences extracted from top 50 Google search result snippets for the query *although she is a woman, she* (incomplete sentences are ignored):

Although she is a woman, she...

1. ...is fearless, skilled at riding, an excellent hunter, and a fine warrior.
2. ...is strong and capable of keeping his secrets.
3. ...is not seen as one in the book.
4. ...is equally as capable of doing farmwork as the men are.
5. ...has some influence, and warns Krogstad to avoid offending his superiors.
6. ...displays serious proof of having “balls.”
7. ...is prepared to do it.
8. ...does not make any efforts to understand young Hazal’s sentiments
9. ...can endure the march as well as any man.
10. ...is the muscle of the family.
11. ...has the physique of a man with broad shoulders.
12. ...has the heart of a king and that the invasion by the Spanish Armada is still ”foul.”
13. ...cannot bare working with women and this is reflected through her manners.
14. ...must demonstrate the ”courage, ingenuity, and selflessness that is associated with Disney’s male heroes”.
15. ...is responsible for the tavern with her husband and she questions Falstaff without hesitation.
16. ...has nous within her.
17. ...has only a woman’s body and a woman’s charm without a woman’s heart.
18. ...is determined to surpass men.
19. ...is similar in many ways to Jack LaLanne.
20. ...has brought science, enlightenment, and ”masculine” rationality to the ”female” Orient.
21. ...is the dominant one in her relationship and is known for all her accomplishments in ”The Family”.
22. ...is not seen as one.
23. ...believes it is only herself who can achieve her own fulfilment.
24. ...is more of a man than you.
25. ...is prepared to die.
26. ...holds the same power and authority as all the men who have ruled before her.
27. ...was born a boy.

28. ...is brave.
29. ...has never gone through the process of pregnancy and labor and delivery.
30. ...has more of an ambitious like a man compared to Duncan.
31. ...is very sensible and smart.
32. ...has a male power in her work.
33. ...fights to revive her ruined homeland
34. ...is hardly worth considering to be a sex object.
35. ...still "manned up".
36. ...acts like man, so we can consider her a male.
37. ...will rule alone.
38. ...is equal to the occasion.
39. ...believes in chivalry.
40. ...has lofty aspirations.
41. ...acts much like a warrior, fighting alongside her Thenns like any other knight.
42. ...can do thing, which helps safe her family while her father loses the power as a family protector.
43. ...doesn't use the typical features of women's writing
44. ...isn't good at cooking.
45. ...fights like a man.
46. ...has much confidence
47. ...is fighting to have high degree education
48. ...has not lost the wonder and playfulness of a child.
49. ...often can be more dependable and confident than men.
50. ...is still old yet mysterious and attractive to men.