

# *An ASP-based Approach to Answering Natural Language Questions for Texts*

DHRUVA PENDHARKAR, KINJAL BASU, FARHAD SHAKERIN  
and GOPAL GUPTA

*Department of Computer Science, The University of Texas at Dallas, Richardson, TX 75080, USA*  
(e-mails: [Dhruva.Pendharkar@utdallas.edu](mailto:Dhruva.Pendharkar@utdallas.edu), [Kinjal.Basu@utdallas.edu](mailto:Kinjal.Basu@utdallas.edu),  
[Farhad.Shakerin@utdallas.edu](mailto:Farhad.Shakerin@utdallas.edu), [gupta@utdallas.edu](mailto:gupta@utdallas.edu))

*submitted 4 July 2020; revised 12 November 2021; accepted 13 December 2021*

---

## Abstract

An approach based on answer set programming (ASP) is proposed in this paper for representing knowledge generated from natural language texts. Knowledge in a text is modeled using a Neo Davidsonian-like formalism, which is then represented as an answer set program. Relevant commonsense knowledge is additionally imported from resources such as WordNet and represented in ASP. The resulting knowledge-base can then be used to perform reasoning with the help of an ASP system. This approach can facilitate many natural language tasks such as automated question answering, text summarization, and automated question generation. ASP-based representation of techniques such as default reasoning, hierarchical knowledge organization, preferences over defaults, etc., are used to model commonsense reasoning methods required to accomplish these tasks. In this paper, we describe the CASPR system that we have developed to automate the task of answering natural language questions given English text. CASPR can be regarded as a system that answers questions by “understanding” the text and has been tested on the SQuAD data set, with promising results.

**KEYWORDS:** ASP, commonsense reasoning, NLP, KR, natural language question answering

---

## 1 Introduction

The goal of AI is to build systems that can exhibit human-like intelligent behavior. Decision-making and the ability to reason are important attributes of intelligent behavior. Hence, AI systems must be capable of performing automated reasoning as well as responding to changing environment (e.g. changing knowledge). To exhibit such a behavior, an AI system needs to understand its environment as well as interact with it to achieve certain goals. Classical logic-based approaches have traditionally been used to build automated reasoning systems but have not led to systems that can be called truly intelligent. Humans, arguably, do not use classical logic in their day-to-day reasoning tasks. They considerably simplify their burden of reasoning by using techniques such as defaults, exceptions, and preference patterns. Also, humans use nonmonotonic reasoning and can deal with incomplete information (Gelfond and Kahl 2014; Jonson-Laird 2009). All these features need to be built into an AI system, if we want to simulate human-like intelligence. It has been shown that commonsense reasoning can be realized via a combination of (stable model semantics-based) negation as failure and classical negation

(Baral 2003; Gelfond and Kahl 2014) in ASP. ASP is a well-developed paradigm and has been applied to solving problem in planning, constraint satisfaction and optimization. There are comprehensive, well known implementations of ASP such as CLASP (Gebser et al. 2010) and DLV (Alviano et al. 2011). Scalable implementations of ASP that support predicates (i.e. do not require grounding) and are query-driven, such as s(ASP) and s(CASP), have also been developed (Marple et al. 2017; Arias et al. 2018). It is proven that ASP is well suited for representing knowledge and for modeling commonsense reasoning.

In this paper, we propose a system called CASPR<sup>1</sup> (Commons-sense ASP Reasoning) to automatically convert textual knowledge into ASP programs and use it to answer natural language questions coded as ASP queries. The problem of converting natural language text into ASP is challenging enough, however, even if we succeed in this translation task, the resulting knowledge is not enough to answer questions to the level that a human can. When we humans read a passage, we automatically draw upon a large amount of commonsense knowledge that we have acquired over the course of years in understanding the passage and in answering questions related to the passage. An automated QA system ought to do the same. CASPR, thus, resorts to resources such as WordNet (Miller 1995), that encapsulate some of the commonsense knowledge, to augment the knowledge derived from the text. CASPR also allows users to add commonsense knowledge – coded in ASP – manually as well.

CASPR uses default theories with exceptions and preferences to represent the textual knowledge as well as commonsense knowledge from external sources (e.g. WordNet). Also, the word-sense-disambiguation (WSD) mechanism used in CASPR is based on default theories (discussed elaborately in Section 5). As we know, a default theory can be used to reason even in the absence of information and to do so negation as failure (NAF) is indispensable. Additionally, commonsense knowledge that we may need for textual question answering may have to represent multiple possible worlds, which cannot be elegantly modeled in traditional logic programming languages such as Prolog. An example scenario is where we may have to reason whether an animal cartoon character can talk like a human. In such a case, we will have to represent the commonsense knowledge that animals can talk in the cartoon world but not in the real world. Such reasoning over multiple worlds is most elegantly modeled in ASP. The above considerations makes answer set programming the crucial building block for the CASPR system unlike other logic programming languages (e.g. Prolog), where the support for NAF is limited or absent.

CASPR runs on the s(ASP) answer set programming system. The s(ASP) system (Marple et al. 2017) is a query-driven predicate ASP system that is scalable, in that it can run answer set programs containing predicates with arbitrary terms. Since the s(ASP) system is query-driven, it does not require grounding, an important feature needed for building large-scale natural language-based KR applications using ASP. The proof of the query serves as a justification, allowing us to give the reasoning behind an answer to a question in CASPR. Just like human question answering mechanism, where we are able to give a logical explanation for each answer to portray our understanding of the matter. Similarly, CASPR showcases its “true” semantic understanding about the

<sup>1</sup> The code-base is available at – <https://github.com/DhruvaPendharkar/thesis-project>.

passage by providing justification for each answer. Also, due to its understanding CASPR is capable of answering related questions about the passage and that is a very important characteristic of any NLU application, such as a ChatBot or a SocialBot.

Our research makes several contributions: (i) it shows that with the help of novel, query-driven systems such as s(ASP), it is possible to build practical NLP applications that rely on true semantic “understanding”; and, (ii) traditional problems of Natural Language Understanding such as word sense disambiguation can be solved quite elegantly with ASP; (iii) finally, it demonstrates that a general (i.e. not domain-specific), scalable QA system can be built without training.

## 2 Background

We next describe some of the key technologies we employ. Our effort is based on answer set programming (ASP) technology (Gelfond and Kahl 2014), specifically, its goal-directed implementation in the s(ASP) system. ASP supports nonmonotonic reasoning through negation as failure that is crucial for modeling commonsense reasoning (via defaults, exceptions and preferences) (Gelfond and Kahl 2014). We assume that the reader is familiar with the basic notations of natural language processing (NLP). Expositions of NLP can be found elsewhere (Jurafsky 2000).

**Answer Set Programming (ASP):** An answer set program is a collection of rules of the form -

$$l_0 \leftarrow l_1, \dots, l_m, \text{not } l_{m+1}, \dots, \text{not } l_n.$$

where each  $l_i$  is a literal (Gelfond and Kahl 2014). In an ASP rule, the left-hand side is called the *head* and the right-hand side is the *body*. Constraints are ASP rules without a *head*, whereas facts are rules without a *body*. The variables start with an uppercase letter, while predicates and constants begin with a lowercase letter. We will follow this convention throughout the paper. The semantics of ASP is based on the stable model semantics of logic programming (Gelfond and Lifschitz 1988). ASP supports *negation as failure* (Gelfond and Kahl 2014) permitting it to elegantly model commonsense reasoning, default rules, exceptions, etc. which is responsible for CASPR’s sophistication.

**s(ASP) system:** s(ASP) (Marple *et al.* 2017) is a query-driven, goal-directed implementation of ASP. A query-driven implementation is indispensable for automating commonsense reasoning, as traditional grounding and SAT-solver based implementations of ASP has limited scalability. There are three major advantages of using s(ASP): 1. s(ASP) does not ground the program, which makes CASPR framework fast and scalable, 2. it only explores the parts of the knowledge-base needed to answer a query, and 3. it provides justification for an answer.

**Stanford CoreNLP tools:** Stanford CoreNLP (Manning *et al.* 2014) is a set of tools for natural language processing. CASPR uses the *Parts-of-Speech (POS) tagger*, *Dependency Parser*, and *Named Entity Recognizer (NER)* tool from this set (Jurafsky 2000). Based on the context, POS tagger generates the necessary parts of speech such as noun, verb, adjective, etc., for each sentence. It also identifies the question type for each question (e.g. what, where, how) and disambiguated words (e.g. *block* as a *verb* vs. *block* as a *noun*). The NER tool captures the entities present in the sentence such as location,

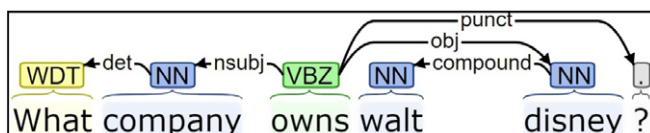


Fig. 1. Example of POS tagging and dependency graph.

- **S: (n) lion, king of beasts, Panthera leo** (large gregarious predatory feline of Africa and India having a tawny coat with a shaggy mane in the male)
- **S: (n) lion, social lion** (a celebrity who is lionized (much sought after))
- **S: (n) Leo, Lion** ((astrology) a person who is born while the sun is in Leo)
- **S: (n) Leo, Leo the Lion, Lion** (the fifth sign of the zodiac; the sun is in this sign from about July 23 to August 22)

Fig. 2. Entry for the concept of lion in WordNet.

time, etc. A dependency graph is eventually generated that shows the grammatical relations between words in the sentence. Dependency relations follow enhanced universal dependency annotation (Schuster and Manning 2016). Figure 1 shows an example of POS tagging and dependency graph of a question.

**WordNet:** WordNet is one of the most commonly used resources in English. It is a lexical database consisting of sense-relations between English words. WordNet consists of multiple databases, one each for nouns, verbs, adjectives and adverbs. WordNet also contains a set of near-synonyms called synsets. Each database contains a set of *lemmas*<sup>2</sup>, each one annotated with a set of senses. A typical entry for the noun lion in WordNet yields the different senses shown in Figure 2.

### 3 CASPR: System architecture

CASPR is composed of two main sub systems: the Knowledge Generation System and the Query Generation System. The architecture, as illustrated in Figure 3, consists of a common resource framework shared by both these systems consisting of NLP tools such as Stanford Core NLP Tools, WordNet API as well as modules for pre-processing input text.

The Knowledge Generation System is mainly responsible for extracting knowledge from a natural language text and representing it as an answer set program. For extracting the knowledge from text and to gain more information about the input text, this component uses Stanford NLP tools like the POS Tagger, Stanford Dependency Parser, and the Stanford NER Tagger. Apart from these resources it also taps into the information that is provided by WordNet (Miller 1995) and extracts knowledge from it, converting it into an answer set program. WordNet provides significant amount of commonsense knowledge about nouns. Unlike nouns, for verbs, at present, there are very few digital resources available that effectively represent the information that verbs convey and that also capture relationships between them (e.g. if we *drop* an object, it will *fall* to the

<sup>2</sup> **Lemmatization** is the task of determining that two words have the same root, despite their surface differences. For example, the word *sing* is the **lemma** for words *sang*, *sung*, and *sings* (Jurafsky 2000).

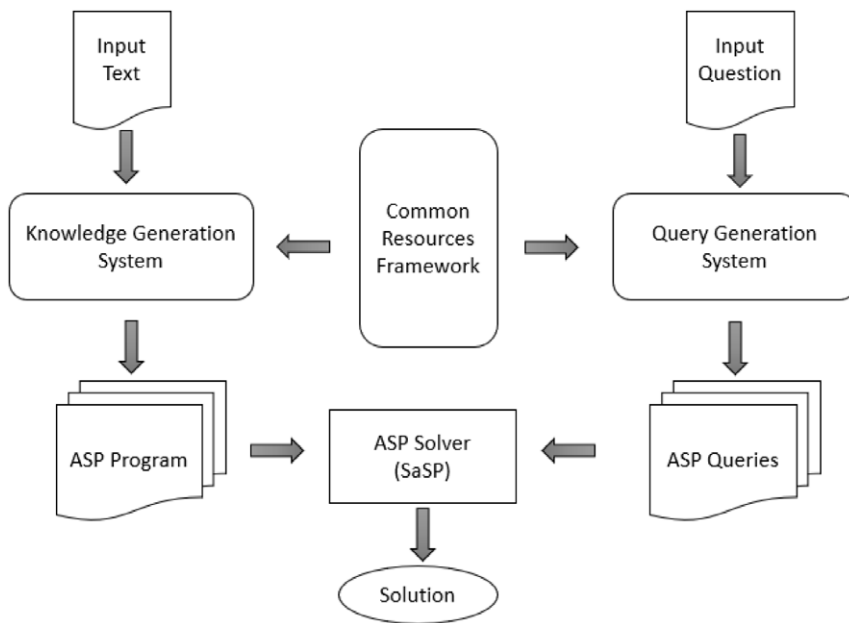


Fig. 3. System architecture.

ground). Such commonsense knowledge for verbs, thus, has to be modeled manually as an answer set program, and added to the common resource framework.

The Query Generation System automatically translates the question to an ASP query that can be executed against the ASP-coded knowledge-base generated from the textual passage augmented with ASP-coded commonsense knowledge. Execution is performed using the *s*(ASP) system. The solution found represents an accurate answer to the question. The query obtained from the natural language question is a conjunction of multiple sub-goals. If the query fails, then some of the sub-goals are systematically removed and the query re-executed to find less accurate answers (details are in Section 6). These less accurate answers are reported too, along with the level of accuracy (likely, possible, guess).

#### 4 CASPR: Knowledge generation

The Stanford Dependency Parser is used to parse the pre-processed text. A semantic graph is generated using the Stanford Typed Dependencies representation (De Marneffe and Manning 2008; De Marneffe *et al.* 2014).

**Example 1.** “NASA carried out the Apollo program.”

Following is the Stanford Dependency (SD) representation:

```

nsubj(carried-2, NASA-1), root(ROOT-0, carried-2), compound:prt(carried-2, out-3),
det(program-6, the-4), compound(program-6, Apollo-5), dobj(carried-2, program-6)

```

These dependencies map straightforwardly onto a directed graph representation in which words in the sentence are nodes in the graph and grammatical relations are edge labels. In English, most event-mentions correspond to verbs and most verbs are triggers

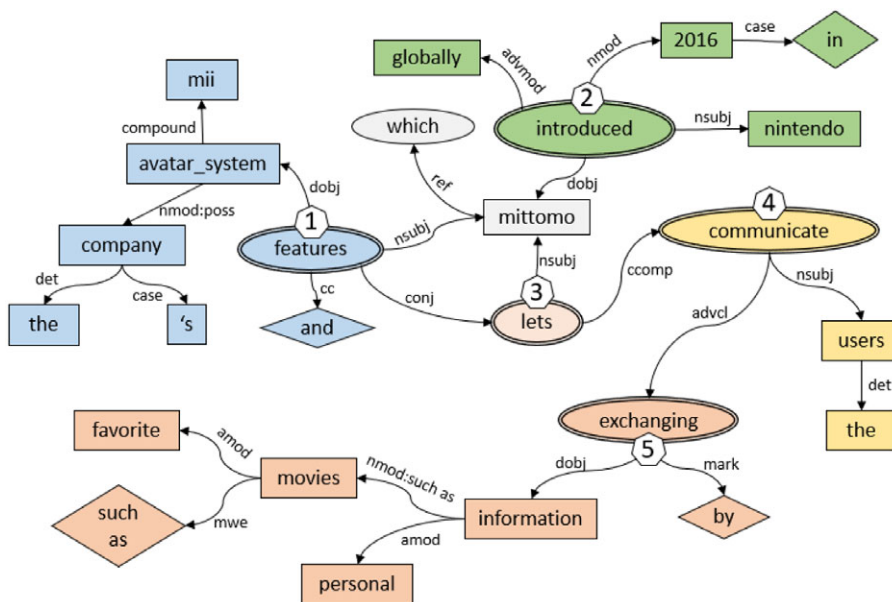


Fig. 4. Semantic graph representing event regions.

to events. Although this is true in most cases there are other word groups that can trigger events as well. The different verbs in the sentence thus define various events that take place in the sentence and how these events are connected to each other. Consider a more complex example.

**Example 2.** “Mittomo, which Nintendo introduced globally in 2016, features the company’s, Mii, avatar system and lets the users communicate by exchanging personal information such as favorite movies.”

Figure 4 shows how various words in the example passage are connected to each other in the sentence. The main verbs of the sentence are “feature” and “let” connected by a coordinating conjunction. The verbs in Figure 4 represent the head of events in the sentence and are marked using event IDs. The various color regions denote the rough boundaries of these event regions. The semantic graph along with the event regions are used to create ASP facts and rules.

#### 4.1 Predicate generation

Knowledge is represented using a collection of predefined predicates, following the neo-Davidsonian approach (Davidson 1984). Some of the predicates capture specialized concepts such as *abbreviation*, *start\_time* etc., whereas others are more generic like *mod*, *event* and so on. The generic predicates that convey information are explicitly present in the text, whereas all others model implicit information. Note that it is important to keep this predicate representation as simple as possible, so that individual pieces of knowledge can be composed using commonsense reasoning patterns. These reasoning patterns have to be kept very simple as well. Otherwise, we run the risk that we may have appropriate knowledge in our knowledge-base to answer the given question, but we fail in answering

it because we are unable to compose that knowledge due to complexity of its representation. A summary of important predicates that have been used by CASPR is given below: We explain the event predicate in detail and summarize the rest (details can be found elsewhere (Pendharkar 2018a;b)).

**Event predicate:** The event predicate defines an event that happens in the sentence. The verb marks the head of the event predicate. The *event* predicate consists of the various actors (doers) and participants involved in the event with the signature:

*event(event\_id, trigger\_verb, actor, participant)*

where the *event\_id* is an integer that uniquely identifies that event in the paragraph. The *trigger\_verb* denoted in the event predicate is the *lemma*, that is, the stem word of the actual word used in the sentence. The actors in the event predicate are the subjects to the trigger\_verb in the sentence. Subjects in the sentence can be found with the help of dependencies like *nsubj* and *nsubj:xsubj*. Just as actors can be obtained from the subject of the sentence, the participants can be determined from direct object dependency (*dobj*).

**Example 3.** “*The American\_Football\_Conference’s (AFC) champion team, Denver-Broncos, defeated the National\_Football\_Conference’s (NFC) champion team, Carolina-Panthers, by 24\_10 to earn AFC third Super\_Bowl title.*”

*event(1, defeat, denver\_broncos, carolina\_panthers)*

*event(2, earn, afc, title)*

Also, following are the *event* predicates for the passage in *Example 2* -

*event(1, feature, mittomo, avatar\_system)*

*event(2, introduce, nintendo, mittomo)*

*event(3, let, mittomo, null)*

*event(4, communicate, users, null)*

*event(5, exchange, null, information)*

Further, we can get richer information about the event by generating duplicate event predicates for each modifier for the actor as well as the participants involved in the event. To generate such event predicates, we use the *amod* or *nummod* dependencies for the actors and the participants and create compound atoms from the modifiers and their governors. Consider the following duplicate event predicate:

*event(2, earn, afc, third\_super\_bowl\_title)*

Note that in absence of information, the default value for the actor and the participant field maybe *null*. A *null* value indicates that either the term is absent for the event or the system was not able to determine it.

**Property predicate:** The property predicate elaborates on the properties of the modified noun or verb. The modifier in this case is generally a prepositional phrase in the sentence. A property predicate is coupled with an event and describes the modification only for that event.

**Example 4.** “*The game was played on February 7, 2016, at Levis\_Stadium, in the San-Francisco\_Bay\_Area, at Santa\_Clara in California*”

*\_property(2, play, on, “february\_7\_2016”)*

*\_property(2, play, at, levis\_stadium)*

*\_property(2, play, in, san\_francisco\_bay\_area)*

*\_property(2, play, at, santa\_clara)*  
*\_property(2, santa\_clara, in, california)*

**Modifier predicate:** The modifier predicate is used to model the relationship between adjectives and the nouns they modify and between verbs and their modifying adverbs.

**Example 5.** “*The Amazon\_rainforest, also known in English as Amazonia or the Amazon\_Jungle, is a moist broadleafed forest that covers most of the Amazon\_basin of South-America.*”

*\_mod(forest, broadleafed)*  
*\_mod(forest, moist)*  
*\_mod(know, also)*

**Possessive predicate:** The possessive predicate is used to model the genitive case in English. It is used to show possession or a possessive relation between two entities in the sentence.

**Example 6.** “*The American\_Football\_Conference’s (AFC) champion team, Denver-Broncos, defeated the National\_Football\_Conference’s (NFC) champion team, Carolina-Panthers, by 24\_10 to earn AFC third Super\_Bowl title*”

*\_possess(american\_football\_conference, team)*  
*\_possess(national\_football\_conference, team)*  
*\_possess(american\_football\_conference, denver\_broncos)*  
*\_possess(national\_football\_conference, carolina\_panthers)*

**Instance predicate:** The instance predicate models the concept of an instance. As an example, *red* is an instance of a *color*.

**Example 7.** “*Nikola\_Tesla was a Serbian-American inventor, electrical engineer, mechanical engineer, physicist, and futurist*”

*\_is(nikola\_tesla, inventor).*  
*\_is(nikola\_tesla, serbian\_american\_inventor).*

In the above example we see that the verb (*is*) is associated with other concepts like engineer, physicist, and futurist using the conjunction. Thus, we can extend the definition of the instance predicate to also include these other facts:

*\_is(nikola\_tesla, electrical\_engineer),*  
*\_is(nikola\_tesla, futurist),*

....

Adding these facts makes the knowledge-base richer which is now able to infer many other things about the passage. Another case, where we can generate the instance predicate is in cases where multi-word expressions like *such as*, or *like* are used to compare two concepts to be equivalent. Consider the following sentence as an example for the same.

**Example 8.** “*Mii\_tomo, which Nintendo introduced globally in 2016, features the company’s Mii, avatar system and lets the users communicate by exchanging personal information such as favorite movies.*”

*\_is(movie, personal\_information)*  
*\_is(favorite\_movie, personal\_information)*



In the above example, we use the expression such as to compare two concepts to be equivalent in the sense that one concept is a likely equivalent of the other. In such cases the generated instance predicate is shown above. Such predicates can be generated using the *nmod*, *mwe*, and the case dependencies.

**Relation predicate:** The relation predicate is used to connect two concepts in events. This predicate is generated to model mainly two relations, that is, *dependent clauses* and *conjunctions*. The signature of the relation predicate can be given as follows.

*\_relation(independent\_entity, dependent\_clause\_id, relation\_type)*

Dependent clauses can be of two types depending upon their governors, that is, adjective clauses or adverb clauses. Adjective clauses are dependent clauses that modify a noun, whereas adverb clauses modify a verb. The *independent\_entity* defined in the signature can be either the noun that is modified or the event ID of the verb that is modified. The *dependent\_clause\_id* is always the ID of the verb that is the head of the dependent clause. The *relation\_type* define the relation between the dependent and the independent clause. This system recognizes 3 types of relations viz. *\_clause*, *\_clcomplement* and *\_conj*. All these relations are discussed below with examples.

As mentioned above the *\_clause* relation is generated for adjective and adverb clauses. Examples for both these clauses is given as follows.

**Example 9.** “*The American\_Football\_Conference’s (AFC) champion team, Denver\_Broncos, defeated the National\_Football\_Conference’s (NFC) champion team, Carolina\_Panthers, by 24\_10 to earn AFC third Super\_Bowl title*”

*\_relation(1, 2, \_clause)*

*event(1, defeat, denver\_broncos, carolina\_panthers)*

*event(2, earn, afc, title)*

In the above given sentence, there are two clauses, one headed by the verb “*defeat*” and the other by the verb “*earn*”. Here, the clause headed by “*defeat*” is the main clause whereas the adverb clause headed by “*earn*” is the dependent clause. Such a relation can be found out using the dependency *advmod*. Similarly, consider the adjective clause in the following sentence.

**Example 10.** “*The American\_Broadcasting\_Company (ABC), stylized in the network’s logo as ABC since 1957, is an American commercial broadcast television network*”

*\_relation(american\_broadcasting\_company, 1, \_clause)*

*event(1, stylize, null, null)*

In this sentence, the noun “*American\_Broadcasting\_Company*” is modified by the dependent clause headed by the verb “*stylize*”. Such a predicate can be modeled using the *acl* dependency relation. One of the other relations is *\_clcomplement*. This type of relation models both the clausal complement (*comp*) as well as the open clausal complement (*xcomp*). Such a relation is used to search for the subject of the dependent clause. An example of such a relation is given as follows.

**Example 11.** “*The ideal thermodynamic cycle used to analyze the process is called the Rankine\_Cycle*”

*\_relation(1, 2, \_clcomplement)*

*event(1, use, null, null)*

*event(2, analyze, null, process)*

Conjunctions are words that connect two or more clauses. The relation predicate is also used to model this relation between any two clauses. An example of such a predicate is given below.

**Example 12.** “Water heats and transforms into steam within a boiler operating at a high pressure”

```
_relation(2, 3, _conj)
event(2, heat, water, null)
event(3, transform, water, null)
```

In the above example we use the *\_conj* relation type to model the conjunction relation between the verbs “heat” and “transform”. As seen above the *\_conj* relation type uses IDs for verb conjunctions whereas actual nouns for noun conjunctions.

**Named entity predicate:** CASPR uses the Named Entity Tagger (NER) (Finkel et al. 2005) to get information about entities in the text. The Named Entity Tagger marks various classes like LOCATION, PERSON, ORGANIZATION, MONEY, PERCENT, TIME in the text. We make use of these tags to generate facts of the form *concept(instance)*. For an example, *location(san\_francisco)* is captured and represented as a fact if a sentence has the word “San Francisco” and NER captures that it is a location name. These facts together with the rules generated by the ontology help in reasoning about the text.

**Special predicates:** Special predicates have been used in the system to model concepts that are patterns and are understood by humans implicitly. As grammatical relations in the sentence do not convey the meaning of these concepts, they must be extracted explicitly. Some of them include abbreviations, time spans and so on. Also, for each special pattern, we have added either ASP facts in the system or generic default rules, so that we need to add it once and can be used for any text without any change. The more of these patterns are learned by the system the better it can reason like humans.

- **Time span predicates:** Time spans are sometimes expressed in text using the bracketed notion. The meaning of these time spans depends upon the noun that they follow. If the noun is a person, then it may mean that the time span indicates the birth and the death date of the person. On the contrary if the noun is an organization or a product or a project then the time span indicates the start and end date for the organization or the project. As there can be multiple other possibilities, we have taken the default ones (mostly-used) unless mentioned about the exceptions and we found from the results that this works really well. A time span is generally of the format “(DATE - DATE)”, where the first date is the start date and the second is the end date. The signature for the start and the end date can be given as:

```
_start_date(entity, date)
_end_date(entity, date)
```

For an example, the time span predicates for the sentence – “Project\_Mercury was followed by the two-man Project\_Gemini (1962 1966)” is – *\_start\_date(project\_gemini, 1962)*, *\_end\_date(project\_gemini, 1966)*.

- **Date part predicates:** The time predicates that have been detected from the Named Entity Tagger, can be utilized to get the day, month, and year parts from the time. This information about the various parts of time is understood implicitly

by humans. We can use simple information extraction techniques to find out the year, month, and day predicates from the time predicate. The signature of these predicates can be given as follows:

*day(time, day)*  
*month(time, month)*  
*year(time, year)*

For an example, the date-part predicates for the time phrase “10\_november\_1483” is – *day(“10\_november\_1483”, 10)*, *month(“10\_november\_1483”, november)*, *year(“10\_november\_1483”, 1483)*.

- **Concept predicates for appositional modifiers:** Appositional modifiers directly follow the noun they describe. We could use this information to model the fact that most appositional modifiers follow the *instance\_of* pattern. This is used to generate concept predicates from the appositional modifiers. Consider the following example for creating a concept predicate.

**Example 13.** “*Jim’s brother, Sam, is coming to town today.*”

*brother(Sam)*

In the above sentences, we have a relation of an appositional modifier between the words “*brother*” and “*Sam*”. This tells us that these words may share an *instance\_of* relation between each other, which leads us to generate the above facts from the appositional modifier.

- **Abbreviation predicate:** Sometimes, texts contain abbreviation of words. It becomes very important for a system to understand the meaning of these abbreviations for querying. This is because the queries asked can sometimes contain short forms for organizations and projects rather than their long forms. Abbreviations can be detected by either doing pattern matching that is, looking for the pattern “CONCEPT (ABBREVIATION)” or by using the *appos* dependency relation. The signature for the abbreviation predicate is as follows:

*\_abbreviation(short\_form, long\_form)*

- **Number predicate:** The number predicate is used to identify numbers. Identifying a word to be a number helps in constraining the domain for certain queries. Other than that, for questions like “*how many*” and “*how much*” it is more accurate to give a number as the answer. The signature for the number predicate can be given as:

*number(value)*

For an example, the number predicates for the sentence – “*Kenya covers 581,309 sq.km. and had a population of approximately 45 million people in July 2014.*” is — *number(“581,309”)*, *number(45)*, *number(million)*.

## 4.2 Commonsense knowledge generation

We make use of additional knowledge sources such as the WordNet to gain supplementary information (commonsense knowledge) about the concepts in the input passage. CASPR uses the Hypernym relation from WordNet to build its ontology, coded as an answer set program.

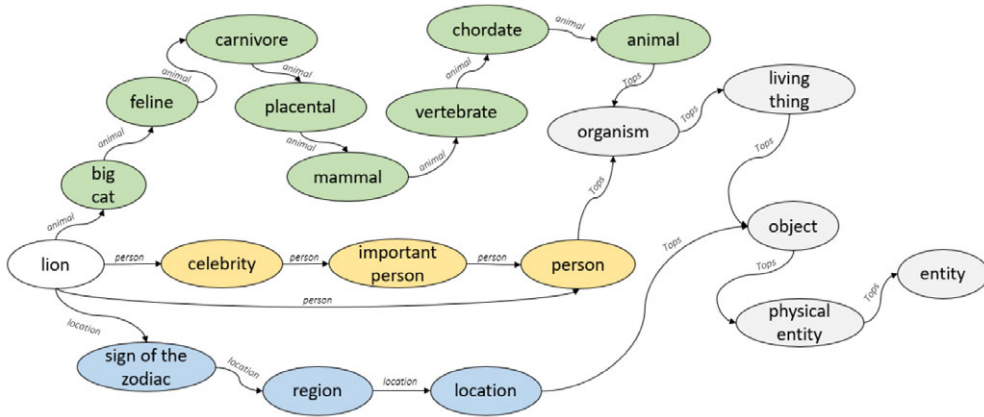


Fig. 5. Concept graph of hypernym relation.

```

big_cat(X, noun_animal) :- lion(X, noun_animal)
feline(X, noun_animal) :- big_cat(X, noun_animal)
carnivore(X, noun_animal) :- feline(X, noun_animal)
placental(X, noun_animal) :- carnivore(X, noun_animal)
mammal(X, noun_animal) :- placental(X, noun_animal)
vertebrate(X, noun_animal) :- mammal(X, noun_animal)
chordate(X, noun_animal) :- vertebrate(X, noun_animal)
animal(X, noun_tops) :- chordate(X, noun_animal)
organism(X, noun_tops) :- animal(X, noun_tops)
living_thing(X, noun_tops) :- organism(X, noun_tops)
object(X, noun_tops) :- living_thing(X, noun_tops)
physical_entity(X, noun_tops) :- object(X, noun_tops)
entity(X, noun_tops) :- physical_entity(X, noun_tops)

```

Fig. 6. ASP Rules representing the animal branch of “lion”.

To generate ontology rules, we use standard knowledge patterns from answer set programming like the preference pattern and the default reasoning pattern (Gelfond and Kahl 2014; Chen et al. 2016). In general a concept is represented in our work using the following signature:

*concept(concept\_instance, instance\_sense)*

For example: *lion(simba, noun\_animal)*.

Hypernyms can be used to infer various properties of and functions about concepts. Hypernyms make use of the generalization principle to transfer properties from more general concepts to their specific concepts. There are three steps required to do so (i) Identify the concepts from the passage and generate a hypernym graph; words may be used in multiple senses (e.g. the word *lion* has four commonly used senses); example graph for word *lion* highlighting 4 different senses is shown in Figure 5. (ii) Aggregate the concepts into common base concepts. (iii) Generate hypernym rules. Figure 6 shows ASP rules generated by CASPR for one of the senses of the concept *lion*. Due to the space constraint, we have omitted a few details, however more illustrations can be found in the thesis (Pendharkar 2018a).

**Adding Commonsense Knowledge Manually:** Note that at the moment we are only using WordNet, however, other knowledge resources, such as YAGO (Mahdisoltani et al.

2015), VerbNet (Kipper *et al.* 2006), never-ending language learner (Mitchell *et al.* 2015), etc., can be incorporated as well to extract additional commonsense knowledge. Doing so is relatively straightforward, as it mostly involves syntactic transformation to ASP syntax. Note that, currently, commonsense knowledge about verbs is entered manually. An example of adding knowledge manually is the following:

The term Nikola\_Tesla is similar to Tesla:

*\_similar(tesla, nikola.tesla).*

A team X can represent an organization Y, if X is possessed by Y:

*event(E, represent, X, Y):- possess(Y, X), organization(Y), team(X),*

*not ab\_event(E, represent, X, Y).*

These handcrafted knowledge are represented using either ASP facts or as default rules with exceptions and also it uses generic predicate names. So, these manually entered commonsense knowledge added once can reside in the knowledge-base permanently and can be used in any other application without any change. This approach is very similar to humans, where we learn knowledge once, and may use it later in another context.

## 5 Word sense disambiguation

Word sense disambiguation (WSD) is the task of selecting the best sense out of a collection of senses applicable to a concept. When we query a concept from WordNet lexicon, we get a list of senses for a specific concept ordered from the most used to the least used. WSD is a very common problem in NLP applications and has many statistical solutions that have been developed. We develop a method using negation as failure, classical negation and preferences that models WSD correctly. Essentially, contextual knowledge is used to disambiguate the word sense, in a manner that humans use. We explain it next.

**Representation of word senses:** Word senses are represented using two different logic patterns in our work. Using both these patterns together, senses are selected for the various concepts in the text which activate their hypernym relations.

In a natural language, a concept (i.e. word) can have multiple senses and one of those senses must have been used in the sentence based on the textual context. The task is to identify that sense. The pattern discussed in this section tries to discover the correct sense for a concept using the characteristics of the sense. Its template can be given as follows:

*c(X, s<sub>i</sub>) :- c(X), characteristics(s<sub>i</sub>, X), not -c(X, s<sub>i</sub>).*

In the above template, we are trying to prove that “X is an instance of concept c with sense s<sub>i</sub>”. Here every concept c has one or more senses denoted by s<sub>i</sub>. The above rule states that X is an instance of the concept c with the sense s<sub>i</sub> only if we can prove that X is some instance of concept c, X has all the characteristics required to be of sense s<sub>i</sub>, and we cannot prove that X is definitely *not* an instance of concept c with sense s<sub>i</sub>. Please note that the “characteristic” predicate becomes true for the instance X with sense s<sub>i</sub> iff it is supported by the knowledge represented from the passage (the context of the passage) in ASP form. Such rules are generated for every sense s<sub>i</sub> of the concept in the order of senses from the most used sense to the least used sense. The first term given by “c(X)” is responsible to short circuit and fail the rule if the instance X does not belong to the class c. The second term, “characteristics(s<sub>i</sub>, X)”, is the main predicate which

tries to prove that  $X$  shows the characteristics of having sense  $s_i$ . It is this predicate that can be added either manually or using other sources to prove the sense. The third term of the body is a strong exception against the head of the rule, that fails the rule if an exception is found against the application of the sense.

According to WordNet, the concept “tree” has three senses,  $S = \{plant, diagram, person\}$ , ordered according to their frequency of use. Thus, using the above-mentioned template for the sense the three rules generated for the tree concept can be given as follows

$$\begin{aligned} tree(X, plant) &:- tree(X), characteristics(plant, X), \\ &\quad not -tree(X, plant). \\ tree(X, diagram) &:- tree(X), characteristics(diagram, X), \\ &\quad not -tree(X, diagram). \\ tree(X, person) &:- tree(X), characteristics(person, X), \\ &\quad not -tree(X, person). \end{aligned}$$

**Preference patterns for senses:** We humans perform word sense disambiguation in our day-to-day life by taking cues from the context. Over a period of time we learn which senses are more common than others and develop a preference order. Consider a concept  $c$  having three senses  $s_1$ ,  $s_2$  and  $s_3$  ordered according to the frequency of their use from the most used to the least used. We first assume the sense to be  $s_1$ , unless we know that  $s_1$  is not the sense from some other source. Then we move on to the next sense  $s_2$  unless we know that both  $s_1$  and  $s_2$  cannot be applicable. Then finally we choose  $s_3$ . This process of elimination of senses and choosing senses according to preferences can be modeled as the following ASP code template.

$$\begin{aligned} c(X, s_p) &:- c(X), not -c(X, s_p), \\ &\quad -c(X, s_1), -c(X, s_2), \dots, -c(X, s_{p-1}), \\ &\quad not c(X, s_{p+1}), not c(X, s_{p+2}), \dots, not c(X, s_n). \end{aligned}$$

The above skeleton is applied for all the senses of concept  $c$  in order of preference. The above template represents the rule generated for the  $p^{th}$  sense of the concept  $c$  such that  $1 < p < n$ , where  $n$  is the total number of senses of concept  $c$ . If  $p = 1$  then the template omits the classical negation terms as follows:

$$\begin{aligned} c(X, s_1) &:- c(X), not -c(X, s_1), not c(X, s_2), \\ &\quad not c(X, s_3), \dots, not c(X, s_n). \end{aligned}$$

Similarly, if  $p = n$ , then the template omits the negation as failure terms, given as:

$$\begin{aligned} c(X, s_n) &:- c(X), not -c(X, s_n), -c(X, s_1), \\ &\quad -c(X, s_2), \dots -c(X, s_{n-1}). \end{aligned}$$

We can apply this pattern to the “tree” concept as an example. A tree can be a living object (plant), a diagram, or a person (Mr. Tree).

$$\begin{aligned} tree(X, plant) &:- tree(X), not -tree(X, plant), \\ &\quad not tree(X, diagram), not tree(X, person). \\ tree(X, diagram) &:- tree(X), not -tree(X, diagram), \\ &\quad -tree(X, plant), not tree(X, person). \\ tree(X, person) &:- tree(X), not -tree(X, person), \\ &\quad -tree(X, plant), -tree(X, diagram). \end{aligned}$$

This pattern is responsible for assigning at least one sense for every concept in the text. This preference pattern along with the property pattern mentioned previously helps disambiguate word sense just as humans do.

## 6 CASPR: Query generation from natural language question

Once knowledge has been generated from the text and auxiliary sources, our next task is to translate the question we want to answer into an ASP query.

Since a question is also a sentence, it is processed in the same way that any other sentence would be, as mentioned before in this paper. This means that a semantic graph is generated for every question, and event regions are created within the question assigning event ids to different parts of the question. To generate an ASP query from the semantic graph the following steps are applied: (i) Question understanding (ii) Generation of query predicates (iii) Applying base constraints (iv) Combining constraints. Currently, this module is only built to deal with simple interrogative sentences but can be extended to deal with more complex questions. In the following subsections, we have explained all the steps of ASP query generation process from natural language questions. For more details users are requested to refer to the thesis (chapter 7) (Pendharkar 2018a).

**Question understanding:** For analyzing questions, we obtain four kinds of information from the question, namely, the question word, the question type as mentioned in the previous section, the answer word or the focus of the question, and the answer type.

The question word is the Wh-word found in the question. In general, Wh-words can be found out by looking at the following POS Tags on the words: WDT, WP, WP\$ and WRB. If none of the tags are found, then the questions may have a copula (Jurafsky 2000) as its question word or a modal as its question word. The question type contains a set of predefined question types that help in further processing. These include WHAT, WHERE, WHO, WHICH, WHEN, HOW\_MANY, HOW\_MUCH, HOW\_LONG, HOW\_FAR, and UNKNOWN. Once the main question word is found, we can use the word and its relations to determine the exact question type. The third type of information we try to extract is the answer word or the focus of the question. The answer word is the word in a question that tells us what kind of answer is expected from the question. Not all types of questions require an answer word, so in those cases this information is null. The fourth and the last information that we extract from the question is the answer type. The answer type depends on the question type. For most of the question types the required answer type is predefined. The answer types supported by the system are as follows SUBJECT, OBJECT, PLACE, PERSON, TIME, YEAR, DAY, MONTH, NUMBER (non-negative integers) and UNKNOWN. Table 1 shows expected answer types depending on question types. Please note that, in the table, “variable” answer type means it can be any type and not bounded by a domain of answers. With the help of this basic analysis, we start generating the predicates for query generation.

**Generating query predicates:** Using the information that we gathered during Question Understanding, we start generating predicate facts for all the words in the question. We will use predicates mentioned previously in the paper as a reference. There are some changes that need to be made in a few predicates, for example, event predicate, property predicate, etc. Others (possess, mod, and named entity predicates) are generated as discussed earlier.

Event predicate: The event predicate in general can be given as follows-  
*event(event\_id, trigger\_verb, actor, participant)*

Table 1. *Expected answer types for question types*

<i>Question Type</i>	<i>Expected Answer Type</i>
WHERE	PLACE
WHO	PERSON
WHEN	TIME
HOW_MANY	NUMBER
HOW_MUCH	NUMBER
HOW_LONG	NUMBER [length]
HOW_FAR	NUMBER [distance]
WHAT	** variable **
WHICH	** variable **
UNKNOWN	UNKNOWN

In case of a sentence from a passage, we mark each verb with an event\_id and it helps to represent the knowledge around that verb by mapping them with the event\_id. Whereas, for a question, the event\_id is unknown, so we put a variable in place of the event\_id. The trigger\_verb is in the question that triggers the predicate generation. The actor acts like the subject of the verb and the participant are the object or the modifier. The participants can be obtained from the direct object (*doj*) relations of the verb. There are various ways in which we can obtain the actor or the subject in any event.

**Example 14.** Given the question “*What company owns Walt\_Disney?*”, the three possibilities for the event predicate are given below.

1.  $event(E1, own, X1, O1), \_similar(walt\_disney, O1)$ .
2.  $event(E1, own, \_, O1), \_property(E1, own, \_by, X1), \_similar(walt\_disney, O1)$ .
3.  $event(E1, own, \_, \_), \_relation(X1, E1, \_clause), \_similar(walt\_disney, O1)$ .

Here,  $X1$  variable binds the output and  $O1$  variable is used internally to prove the sub-goals.

*Property predicate:* Property predicates are generated from verbs and nouns that are modified by nominal phrases in the sentence. Here, like the event predicate, we are unaware of the event\_id and hence it will be set as a variable. The modified\_entity is the noun or the verb that triggered the predicate generation. The preposition here can either be obtained from the case relation or can be left blank (-). The modifier is the head of the nominal modifier that can be used to constraint the query. If the modifier is the answer word, then we replace the word with the answer tag ( $X_k$ ).



**Example 15.** Given “On what street is the ABC’s headquarter located?”, we generate:  $\_property(E2, locate, on, X2)$ .

*Similar predicate:* The *similar* predicate models the concept of similarity between entities in which one entity is so like the other one that they can replace each other. The similar predicate thus models one of the principles of commonsense reasoning where we as humans make use of the similarity relationship while reasoning (referring to *Albert Einstein* as just *Einstein*). Some rules for the similar predicate are as follows:

1. Abbreviations of concepts are similar to each other for example, *NYPD* and *New York Police Dept.*

$\_similar(X, Y) :- \_abbreviation(X, Y)$ .

$\_similar(X, Y) :- \_abbreviation(Y, X)$ .

2. Instances of concepts are similar to each other for example, *Mercedes* is similar to a car

$\_similar(X, Y) :- \_is(X, Y)$ .

3. Similarity follows transitivity for example, *New York, NY*, *New York City* and *NYC* are all interchangeable concepts that similar to each other.

$\_similar(X, Y) :- \_similar(X, Z), \_similar(Z, Y)$ .

**Generating base constraints:** Base constraints are generated at the end after all the constraints from the question have been generated. The term constraint here refers to constraints that arise due to the type of the answer, for example, if the expected answer is a person, then the variable holding the answer is restricted (constrained) to one of the persons mentioned in the passage. Constraints, thus, are sub-goals whose conjunction produces an answer to the question. Each sub-goal can be thought of as putting a constraint on the acceptable answers. For the following cases we generate the base constraints as follows.

TIME  $\rightarrow time(X_k)$

DAY  $\rightarrow day(T_k, X_k), time(T_k)$

MONTH  $\rightarrow month(T_k, X_k), time(T_k)$

YEAR  $\rightarrow year(T_k, X_k), time(T_k)$

PLACE  $\rightarrow location(X_k)$  or  $location(X_k, noun\_location)$

PERSON  $\rightarrow person(X_k)$  or  $person(X_k, noun\_person)$

In case the answer type is UNKNOWN, we may expect the answer to be a specific concept represented by the answer word. In such cases the base constraint comes from the answer word itself.

UNKNOWN  $\rightarrow concept(X_k)$

for example,  $company(X_k)$  or  $city(X_k)$

**Combining constraints:** After generation of the predicates from the question and the base constraints from answer type we combine all the constraints to create the final query. This can be best explained with an example.

**Example 16.** For the question “When was Nikola Tesla born?”, the following queries will be generated and automatically tried in order.

?-  $event(E2, bear, S2, O2), \_similar(nikola\_tesla, S2),$

$property(E2, bear, on, X2), time(X2)$ .

?- event(E2, bear, -, O2), \_property(E2, bear, \_by, S2),  
     \_similar(nikola\_tesla, S2), property(E2, bear, on, X2), time(X2).  
 ?- event(E2, bear, -, -), \_relation(S2, E2, \_clause),  
     similar(nikola\_tesla, S2), property(E2, bear, on, X2), time(X2).  
 ?- \_start\_date(S2, X2), \_similar(nikola\_tesla, S2), time(X2).

In this sentence, we have the special predicates `_start_date` applying the constraints of time spans on the entity “*Nikola Tesla*”.

**Query confidence classes:** Ideally, we would want to be able to answer all questions with most number of constraints applied on the query so that we have a strong justification for the answer, but this is not always the case. In natural language the same questions can be posed in multiple different ways using synonymous concepts. This makes it more difficult to answer questions. Thus, it is better to make any question answering system robust enough, so that it fails gracefully. Once such attempt has been made in the query generation module by introducing the concept of confidence classes on the generated queries and by relaxing constraints on the queries to make the query more flexible. The queries that were generated in the previous section represent the most constrained queries. If an answer is found to one of those queries, then the system has the most confidence in the answer. In the absence of an answer the system starts removing constraints on the query (i.e. starts dropping sub-goals from the query) and relaxing the constraints on the query with the aim of obtaining *some* answer. The queries are designed in such a way that it always produces an answer along with its confidence class. Currently, queries have been divided into 4 confidence classes that range from most confident Class I to the least confident Class IV:

**Certain (Class I):** These kinds of queries have all the constraints included in the natural language question. The query may contain fact predicates, subordinate constraints, answer predicates and base constraints. This class has the highest confidence level. If an answer is produced, it is a correct answer.

**Likely (Class II):** These kinds of queries are executed when a *certain* answer is not produced. They do not contain any fact predicates. Fact predicates are all the predicates in the queries that do not contain any variables. Such predicates come from the question or the named entity tagger and act as constraints.

**Possible (Class III):** These kind of queries are executed when a *likely* answer is not produced. They only contain the answer predicates and base constraints. This means that the fact predicates and any subordinate constraints that do not contain the answer tag ( $X_k$ ) are removed.

**Guess (Class IV):** These kind of queries are executed when a *possible* answer is not produced. They only contain the base constraints (rest of the sub-goals are dropped). Answers from these queries are not very reliable as they can be thought of as guesses. This class has the lowest confidence.

Confidence class-based queries generated for the question “*In what borough of New York City is ABC headquartered?*” are shown below:

Certain (Class I) Queries:

- ?- *\_property(E2, borough, of, new\_york\_city), \_property(E2, headquarter, \_by, S2),*  
*\_property(E2, headquarter, in, X2), \_similar(abc, S2), borough(X2, -),*  
*event(E2, headquarter, -, O2), organization(abc).*
- ?- *\_property(E2, borough, of, new\_york\_city), \_property(E2, headquarter, in, X2),*  
*\_relation(S2, E2, \_clause), \_similar(abc, S2),*  
*event(E2, headquarter, -, -), organization(abc), borough(X2, -).*
- ?- *\_property(E2, borough, of, new\_york\_city), \_property(E2, headquarter, in, X2),*  
*\_similar(abc, S2), event(E2, headquarter, S2, O2),*  
*organization(abc), borough(X2, -).*

Likely (Class II) Queries:

- ?- *\_property(E2, borough, of, new\_york\_city), \_property(E2, headquarter, \_by, S2),*  
*\_property(E2, headquarter, in, X2), \_similar(abc, S2),*  
*event(E2, headquarter, -, O2), borough(X2, -).*
- ?- *\_property(E2, borough, of, new\_york\_city), \_property(E2, headquarter, in, X2),*  
*\_relation(S2, E2, \_clause), \_similar(abc, S2),*  
*event(E2, headquarter, -, -), borough(X2, -).*
- ?- *\_property(E2, borough, of, new\_york\_city), \_property(E2, headquarter, in, X2),*  
*\_similar(abc, S2), event(E2, headquarter, S2, O2), borough(X2, -).*

Possible (Class III) Queries:

- ?- *\_property(E2, headquarter, in, X2), borough(X2, -).*

Guess (Class IV) Queries:

- ?- *borough(X2, -).*

These four class of queries will produce four different answers, from 100% correct answer to a guess. Class I query response will be “Manhattan,” since ABC Corp is located in the burrough of Manhattan. That’s the correct (and most specific) answer. For Class II query, the constraint “organization/1” is dropped, and so if there is another entity (not an organization) that is called ABC, then the burrough where its headquarterd will be produced as the answer. For Class III query, most constraints are dropped and a burrough containing any headquarter of any entity will be named, while for Class IV query, any arbitrary burrough located anywhere will be named. We are attempting to emulate a human who tries to give a possible answer, which could be a wild guess, if he/she does not possess adequate information.

## 7 CASPR: Results

The SQuAD Dataset (Rajpurkar *et al.* 2016) contains more than 100,000 reading comprehensions along with questions and answers for those reading passages. SQuAD dataset uses the top 500+ articles from the English Wikipedia. These articles are then divided into paragraphs. We used the Dev Set v1.1 of the SQuAD Dataset to obtain comprehension passages for building a prototype for the proposed approach. This dataset has around 48 different articles with each article having around 50 paragraphs each. Out of the 48 different articles in the SQuAD dev set, 20 articles were chosen from different domains to help build the CASPR system (these 20 passages and associated questions can be found in Pendharkar (2018b)). These 20 articles are chosen in such a way that a lot of different domains were covered. We can roughly categorize these articles into 5 different

Table 2. Results for question answering

No	Article	Result	%	No	Article	Result	%
1	ABC Corp.	3/4	100	11	Kenya	5/5	100
2	Amazon rainforest	12/14	85.7	12	Martin Luther	2/5	40
3	Apollo	4/5	80	13	Nikola Tesla	6/7	85.7
4	Chloroplasts	4/5	80	14	Normans	4/5	80
5	Computational complexity	3/3	100	15	Oxygen	8/15	53.3
6	Ctenophora	9/12	75	16	Rhine	5/8	62.5
7	European union law	13/13	100	17	Southern California	3/5	60
8	Genghis Khan	3/5	60	18	Steam engine	4/5	80
9	Geology	4/5	80	19	Super bowl 50	25/29	86.2
10	Immunesystem	13/15	86.6	20	Warsaw	3/5	60
<b>Total</b>					135/171	78.95%	
<b>Average result</b>					77.76%		

categories: People articles, Scientific articles, Project/Event articles, Region articles and Misc. articles. Diversity is important as we have used these articles to understand the relations and that helps to craft the generic predicate rules. Using the 20 different articles mentioned above, the ASP code was generated for one paragraph from each article. Then, ASP queries were generated for all the questions in the dataset for these paragraphs. The results show the percentage of questions for which the answer generated from the ASP solver was present in the list of answers specified for the question in the SQuAD dataset. The results are summarized in Table 2. Our results show that approximately 77.76% of the questions are correctly answered. This shows that most of the knowledge, if not all, has been captured successfully in the ASP program generated for the passage. The ASP queries generated for the questions are very similar to the original question and convey the same meaning.

Note that the two main reasons why a query may fail to produce an exact (certain) answer are: (i) the parser used by CASPR fails to parse the question; and, (ii) missing commonsense knowledge about concepts in the question. With the help of common framework, knowledge from the passage is converted into an answer set program and the ASP query is generated from the natural language question. Commonsense knowledge bridges the gap between these two (the ASP query and the clauses from passage and question respectively). In the process of automatic commonsense knowledge generation,

if appropriate knowledge is not generated for a concept present in the question, then CASPR will not be able to answer the question properly.

Note that our system has shown excellent execution performance on the 171 questions on 20 passages tried thus far: 80% of the questions were answered in 2 to 3 ms, while the rest were answered in a few seconds.

## 8 Contributions

The main contribution of this paper is an effective and efficient method for converting textual data into knowledge represented as an answer set program that can be processed on our query-driven s(ASP) ASP system. This includes developing a neo-davidsonian logic inspired generic calculus that helps represent knowledge, and using knowledge sources such as WordNet to acquire commonsense knowledge about terms found in the text to create a custom ontology for the problem. This automatic custom ontology generation process makes the CASPR system scalable. Yet another novelty is in showing how word sense disambiguation can be elegantly modeled using ASP. Our system is based on “*truly understanding*” the text just as humans do.

The paper also proposes a framework for converting natural language questions into ASP queries. These queries can be run on the query-driven s(ASP) system to compute answers. The query generation framework is made robust through broadening of queries by dropping constraints, thus increasing the possibility that the question will indeed be answered (even though in the worst case the answer may just be a guess). This approach to handling question answering is yet another novelty of the proposed system.

## 9 Related works

With respect to related work, Cyc ([Wikipedia contributors 2018](#)) is one of the oldest AI project that attempts to model commonsense reasoning. In Cyc, knowledge is presented in the form of a vast collection of ontologies that consist of implicit knowledge and rules about the world that represents commonsense knowledge. The ontology of Cyc as of 2017 contains around 1,500,000 terms. This includes around 500,000 collections, 50,000+ predicates and around a million well known entities. Cyc’s language CycL made it efficient to represent commonsense knowledge in the project and determined how this knowledge is represented in the project. Cyc uses a *community of agents* consisting of multiple reasoning agents that rely on more than 1000 heuristic modules to solve inference problems. Cyc, however, does not work with natural language, though recently some efforts have been started. A potential problem with Cyc is knowing which agent to apply, and which heuristic to use. For a commonsense reasoning system to be successful, *it has to be modeled in a very simple way*. Otherwise, we may possess the individual pieces of knowledge to answer a given question, but may not be able to compose these pieces together to arrive at an answer. For this reason, CASPR represents knowledge using very few generic predicates and uses simple ASP-based reasoning patterns to compute answers. Our initial experiments suggest that our approach is effective.

The application of ASP to solve NLP tasks is not new. [Vo et al. \(2015\)](#) have developed the NL2KR tool that allows natural language text to be translated to an answer set program. Similarly, the work by [Costantini and Paolucci \(2010\)](#) shows how ASP code can be

generated automatically from natural language using lambda calculus. Action language based approaches are also popular in translating natural language to ASP (Olson and Lierler 2019; Zhang et al. 2020) etc. Several other researchers have worked on applying ASP for NLP tasks and their efforts are reported in the first workshop on NLP and Automated Reasoning (Baral and Schüller 2013). Our approach has elements common with these efforts, however, our work is based on the query-driven s(ASP) predicate ASP engine, and thus is scalable and not constrained by limitations of grounding based implementations of ASP. The query-driven s(ASP) system is crucial to our success as, like the current work, our previous works in this direction show promising results in the domain of visual question answering (Basu et al. 2020) and natural language question answering (Basu et al. 2020; 2021).

There are many approaches to question answering based on machine learning (cf. SQuAD website (Rajpurkar et al. 2016)). However, they are not based on *actually understanding* the text and so can only answer questions related to data they are trained on. CASPR, in contrast, *understands* the knowledge contained in the text and has shown promising results for a subset of the SQuAD dataset. It should be noted that machine learning based approaches are based on statistical techniques and pattern matching, and do not represent the knowledge inherent in the text in any form. There is nothing that corresponds to *understanding* of the text, in contrast to our approach where this *understanding* is represented as knowledge modeled in ASP. For this reason, we do not elaborate too much here on machine learning based approaches to question answering. Also, there are some machine comprehension efforts based on semantic lexicon (Clark et al. 2018) that work really well on procedural texts, however they may not work well on non-procedural, descriptive text. In addition, they do not perform nonmonotonic reasoning. In contrast, CASPR can handle both non-procedural texts and perform nonmonotonic reasoning.

A critical component of CASPR's success is the s(ASP) query-driven, predicate ASP system that leads to three major advantages for CASPR: (i) only parts of the knowledge-base relevant to answering the question are explored during execution; (ii) justification for answers can be extracted from the justification tree produced by s(ASP); and, (iii) the question answering system is scalable, as no grounding of the program needs to be done as s(ASP) can execute predicates directly under the stable model semantics.

## 10 Conclusions and future work

In this paper, we presented a comprehensive ASP-based framework called CASPR for natural language question answering. CASPR translates natural language text and questions into ASP code and ASP query, respectively. It also imports commonsense knowledge from other sources such as WordNet. We also presented an elegant solution for word sense disambiguation based on default and classical negation. We used the query-driven s(ASP) engine to perform commonsense reasoning to obtain the answer. CASPR was applied to the SQuAD dataset and has shown promising results. Our approach always finds the most suitable answer, if adequate knowledge is present and the natural language analysis works correctly. If not, it employs a simple method for weakening the question to ensure that an answer will be found, even though it may be a guess. CASPR not only gives

promising results, it also enjoys other benefits such as explainability, generalizability, and interpretability.

CASPR is a step toward building truly intelligent systems that mimic human reasoning based on “truly understanding” the text. Our future work includes (i) extending the system to handle more complex questions (e.g. causality questions), (ii) incorporating additional knowledge resources (e.g. Wikidata (Vrandečić and Krötzsch 2014) and Dbpedia (Auer *et al.* 2007)) for importing more commonsense knowledge to handle a broader class of questions, (iii) adding more semantics in knowledge representations, such as *semantic relations* or predefined semantic predicates (e.g. VerbNet), (iv) generating justifications to a question’s answer in a more human-readable way, (v) extending the system for other NLP tasks: text summarization, question generation, etc.

### Acknowledgment

Thanks to Sarat Varanasi, Elmer Salazar, Joaquin Arias, Zhuo Chen, and Serdar Erbatur for discussions and help. Authors gratefully acknowledge support from NSF grants IIS 1910131 and IIS 1718945.

### References

- ALVIANO, M., FABER, W., LEONE, N., PERRI, S., PFEIFER, G. AND TERRACINA, G. 2011. The disjunctive datalog system dlv. In *Datalog Reloaded*. Springer, 282–301.
- ARIAS, J., CARRO, M., SALAZAR, E., MARPLE, K. AND GUPTA, G. 2018. Constraint answer set programming without grounding. *arXiv preprint arXiv:1804.11162*.
- AUER, S., BIZER, C., KOBILAROV, G., LEHMANN, J., CYGANIAK, R. AND IVES, Z. 2007. Dbpedia: A nucleus for a web of open data. In *The Semantic Web*. Springer, 722–735.
- BARAL, C. 2003. *Knowledge Representation, Reasoning and Declarative Problem Solving*. Cambridge University Press.
- BARAL, C. AND SCHÜLLER, P., Eds. 2013. *Proceedings of the 1st Workshop on Natural Language Processing and Automated Reasoning 2013*. CEUR Workshop Proceedings, vol. 1044. CEUR-WS.org.
- BASU, K., SHAKERIN, F. AND GUPTA, G. 2020. AQUA: ASP-based visual question answering. In *International Symposium on Practical Aspects of Declarative Languages*. Springer, 57–72.
- BASU, K., VARANASI, S., SHAKERIN, F., ARIAS, J. AND GUPTA, G. 2021. Knowledge-driven natural language understanding of english text and its applications. *arXiv preprint arXiv:2101.11707*.
- BASU, K., VARANASI, S. C., SHAKERIN, F. AND GUPTA, G. 2020. Square: Semantics-based question answering and reasoning engine. *arXiv preprint arXiv:2009.10239*.
- CHEN, Z., MARPLE, K., SALAZAR, E., GUPTA, G. AND TAMIL, L. 2016. A physician advisory system for chronic heart failure management based on knowledge patterns. *Theory and Practice of Logic Programming* 16, 5–6, 604–618.
- CLARK, P., DALVI, B. AND TANDON, N. 2018. What happened? leveraging verbnet to predict the effects of actions in procedural text. *arXiv preprint arXiv:1804.05435*.
- COSTANTINI, S. AND PAOLUCCI, A. 2010. Towards translating natural language sentences into asp. In *CILC*.
- DAVIDSON, D. 1984. *Inquiries into Truth and Interpretation*. Oxford University Press.

- DE MARNEFFE, M.-C., DOZAT, T., SILVEIRA, N., HAVERINEN, K., GINTER, F., NIVRE, J. AND MANNING, C. D. 2014. Universal stanford dependencies: A cross-linguistic typology. In *LREC*. Vol. 14, 4585–4592.
- DE MARNEFFE, M.-C. AND MANNING, C. D. 2008. Stanford typed dependencies manual. Tech. rep., Technical report, Stanford University.
- FINKEL, J. R., GREINER, T. AND MANNING, C. 2005. Incorporating non-local information into information extraction systems by Gibbs sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 363–370.
- GEBSER, M., KAMINSKI, R., KAUFMANN, B., OSTROWSKI, M., SCHAUB, T. AND THIELE, S. 2010. gringo, clasp, clingo, and iclingo. *User guide*.
- GELFOND, M. AND KAHL, Y. 2014. *Knowledge Representation, Reasoning, and the Design of Intelligent Agents: The Answer-Set Programming Approach*. Cambridge University Press.
- GELFOND, M. AND LIFSCHITZ, V. 1988. The stable model semantics for logic programming. In *ICLP/SLP*. Vol. 88, 1070–1080.
- JONSON-LAIRD, P. 2009. *How We Reason*. Oxford University Press.
- JURAFSKY, D. 2000. *Speech & Language Processing*. Pearson Education India.
- KIPPER, K., KORHONEN, A., RYANT, N. AND PALMER, M. 2006. Extending verbnet with novel verb classes. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation, LREC 2006, Genoa, Italy, May 22–28, 2006*, 1027–1032.
- MAHDISOLTANI, F., BIEGA, J. AND SUCHANEK, F. M. 2015. YAGO3: A knowledge base from multilingual wikipeidias. In *CIDR 2015, Seventh Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 4-7, 2015, Online Proceedings*.
- MANNING, C. D., SURDEANU, M., BAUER, J., FINKEL, J., BETHARD, S. J. AND MCCLOSKEY, D. 2014. The Stanford CoreNLP natural language processing toolkit. In *ACL System Demonstrations*, 55–60.
- MARPLE, K., SALAZAR, E. AND GUPTA, G. 2017. Computing stable models of normal logic programs without grounding. *arXiv preprint arXiv:1709.00501*.
- MILLER, G. A. 1995. WordNet: A lexical database for english. *Communications of the ACM* 38, 11, 39–41.
- 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. AND WELLING, J. 2015. Never-ending learning. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*.
- OLSON, C. AND LIERLER, Y. 2019. Information extraction tool Text2ALM: From narratives to action language system descriptions. *Electronic Proceedings in Theoretical Computer Science* 306, 87100.
- PENDHARKAR, D. 2018a. *An Answer Set Programming based Approach to Representing and Querying Textual Knowledge*. M.S. Thesis, The University of Texas at Dallas, <http://utdallas.edu/~gupta/dpthesis.pdf>.
- PENDHARKAR, D. 2018b. Caspr. <https://github.com/DhruvaPendharkar/thesis-project>.
- RAJPURKAR, P., ZHANG, J., LOPYREV, K. AND LIANG, P. 2016. SQuAD: 100,000+ questions for machine comprehension of text. *arXiv preprint arXiv:1606.05250*.
- SCHUSTER, S. AND MANNING, C. D. 2016. Enhanced english universal dependencies: An improved representation for natural language understanding tasks. In *LRED'16*, 2371–2378.
- VO, N. H., MITRA, A. AND BARAL, C. 2015. The NL2KR platform for building natural language translation systems. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language*



*Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers*, 899–908.

- VRANDEČIĆ, D. AND KRÖTZSCH, M. 2014. Wikidata: A free collaborative knowledgebase. *Communications of the ACM* 57, 10, 78–85.
- WIKIPEDIA CONTRIBUTORS. 2018. Cyc — Wikipedia, the free encyclopedia. [Accessed on May 17, 2018].
- ZHANG, Q., BENTON, C. AND INCLEZAN, D. 2020. An application of ASP theories of intentions to understanding restaurant scenarios: Insights and narrative corpus. *Theory and Practice of Logic Programming* 20, 2, 273–293.