



Looking Inside the Black Box: Core Semantics Towards Accountability of Artificial Intelligence

Roberto Garigliano¹  and Luisa Mich²  

¹ Journal of Natural Language Engineering (Founding Editor), Cambridge, UK

² University of Trento, Trento, Italy

luisa.mich@unitn.it

Abstract. Recent advances in artificial intelligence raise a number of concerns. Among the challenges to be addressed by researchers, accountability of artificial intelligence solutions is one of the most critical. This paper focuses on artificial intelligence applications using natural language to investigate if the core semantics defined for a large-scale natural language processing system could assist in addressing accountability issues. Core semantics aims to obtain a full interpretation of the content of natural language texts, representing both implicit and explicit knowledge, using only ‘subj-action-(obj)’ structures and causal, temporal, spatial and personal-world links. The first part of the paper offers a summary of the difficulties to be addressed and of the reasons why representing the meaning of a natural language text is relevant for artificial intelligence accountability. In the second part, a proof-of-concept for the application of such a knowledge representation to support accountability, and a detailed example of the analysis obtained with a prototype system named CoreSystem is illustrated. While only preliminary, these results give some new insights and indicate that the provided knowledge representation can be used to support accountability, looking inside the box.

Keywords: Artificial intelligence · Natural language processing · Knowledge representation · Semantics · Rules · Accountability

1 Introduction

Thanks to more powerful hardware and to a new generation of learning algorithms [1], artificial intelligence (AI) supports the automation of a widespread number of tasks and activities, changing not only the job landscape, but also everyday life [2, 3]. Embedded in almost any device and software system, AI solutions support decisions and control systems giving advice and recommendations that may imply serious risks [4–7]. The first step to address such risks is to be able to explain why a given solution or behavior has been chosen, providing information on the data and knowledge used, and on their processing, thus including stakeholders. This feature of an AI system is named accountability [8]. The problem is that many AI systems run programs based on algorithms whose particular output cannot usually be traced back to specific parts of the input. The new generation algorithms, based on deep learning, are even more inscrutable

due to the complexity of the processing steps and the huge size of data required and produced [9]. Focusing on AI systems using natural language, whose role is relevant in a variety of areas, we assume that a ‘core’ semantic representation of the content of natural language text could assist to address accountability issues, looking inside the box. A core semantic approach aims at obtaining a full interpretation of a natural language text representing both implicit and explicit knowledge: in this way it could support explanation of the output of AI systems embedding any form of natural language processing (NLP), e.g., translators, chatbot, information extraction systems. Furthermore, a core semantic representation could be applicable also to systems which do not have natural language as their normal input or output (e.g., a medical system which takes patient data and produces a structured output), but which would benefit from being able to store their knowledge in core semantics and explain it in a comprehensible manner using natural language.

In this paper we will investigate if and how the core semantic representation defined for a large-scale domain independent NLP system could be used to support accountability of AI. The system, a prototype that in this paper is referred as CoreSystem, represents the content of natural language texts using only ‘subj-action-(obj)’ structures and causal, temporal, spatial and personal-world links, the basic elements of the ‘core semantics’. To be able to explain is the first, crucial step in ‘accounting’, which can be seen as detecting causes and then allocating responsibilities. Beyond that, the two key elements in allocating responsibilities are causal relationship and ‘personal-world’ relationships (those produced by relations which move into a person’s inner world, such as ‘think’, ‘want’, ‘need’, etc.). These are exactly the two elements which CoreSystem uses as basic links in its model.

As a proof-of-concept, a detailed example is illustrated, showing the representation of a complex sentence’s content produced by CoreSystem, and how it could be used to answer some simple questions in order to explain its interpretation. The results of the analysis, albeit preliminary, indicate that the core semantics approach produces a knowledge representation that can be understood and checked towards accountability goals.

The rest of the paper is structured as follow. Section 2 illustrates the concept of accountability and the difficulties involved in making an AI system accountable. Section 3 summarizes the problems of NLP in the light of accountability and the benefit of a core semantic representation of a text. In Sect. 4, the core semantics defined for CoreSystem is illustrated focusing on the output of the system. Section 5 gives an example of how the analysis produced by an NLP system able to implement a core semantics could be used to support accountability goals. Conclusions are drawn in Sect. 6.

2 Accountability for Artificial Intelligence

In AI, accountability is the ability to explain how a given result has been obtained from a given input, to justify why a certain decision or behaviour has been suggested and to identify roles and responsibilities. The accountability concept is connected to that of explainability, interpretability and transparency [5, 10, 11]. Accountability problems

were raised since the very first AI systems, as automatic systems and processes are based on algorithms and the problem of accountability for their output is not a new one. Many authors have underlined the risks connected with applications of algorithms in different fields. AI has elevated the complexity of algorithms and the related risks to unprecedented levels. Besides, many people are unaware of the use of the results of a software system and accept suggestions without critical reflection. For a given AI system, accountability is challenged since the first activities of definition and choice of the input data. Data and knowledge used in deep-learning AI algorithms are:

- unstructured: textual documents, audio, video, images;
- extracted from large data sets and knowledge bases;
- analysed applying data analytics or other techniques of big data analysis.

The higher risk is that of data-bias, that is of data reflecting values of the people who design and realize the data sets [12]. A number of cases have been reported in literature and in newspapers [11, 13, 14].

As regards algorithms, the need to have explanations for decisions by being able to inspect the system or the code – looking inside the box – can be defined as external accountability. Unfortunately, learning algorithms, especially unsupervised and deep learning algorithms, are based on models that do not allow tracking and understanding of the internal steps [15, 16]. Complex multilayer neural networks and large inputs cannot be described in details at the level of their inner processing and, in turn, it is difficult to understand the relationship between inputs and outputs. For expert systems, one of the first type of AI applications and usually based on if-then rules, accountability is guaranteed through ‘why?’ and ‘how?’ explanation capabilities, allowing, e.g., a doctor to know why a given diagnose was suggested for the specified symptoms, but this is not the case for the new generation of AI algorithms. Even for supervised algorithms, in which it would be possible to use (a subset of) the training sets to show the input for a given output to explain why a solution was obtained, usability problems could arise [17].

Finally, the output of an AI system comes in a variety of forms; each of them can be more or less difficult to be traced-back to the input and to justify the results.

3 Natural Language Processing and Accountability

3.1 Representing Meaning in a Text

Natural language processing is one of the main areas of AI. Natural language texts are traditionally analysed in a sequential process, starting with lexical and structural elements, parsing text to identify the most suitable parsing tree and then applying more or less complex techniques to interpret the semantic content, that is, to understand the meaning.

Parsing trees and semantic representations are typically dependent on the particular form of the sentence; in this sense, they give a surface representation. For example, the parsing and shallow semantic analysis of the following sentence:

- (a) A 59-year-old man from York has been arrested on suspicion of murdering missing chef Claudia Lawrence.

would allow identifying two instances of ‘named entity’, ‘York’ and ‘Claudia Lawrence’; the first part of the sentence as the syntactical subject of an ‘arrest’; ‘missing’ as an adjective and (possibly) ‘chef’ as a ‘role’ of the named entity ‘Claudia Lawrence’; the last part of the sentence as the object of the ‘suspicion’. The difference between a shallow and a core semantics can be illustrated comparing sentence (a) with the following:

- (b) Police have arrested a York man, aged 59, because they suspect him to be the murderer of Claudia Lawrence, the chef who has disappeared.

Phrase (b) has the same meaning for any competent native speaker as phrase (a), but it produces completely different parse tree and surface semantics. It should be noticed that there is a large amount of ways (surface structures) in which this same meaning could be expressed.

To fully understand that meaning, an NLP system needs a core semantics, that is, an approach based on an internal representation of the content of sentences in which both explicit and implicit knowledge is showed.

There are many reasons why an internal representation for natural language is necessary and the most important for an NLP system are the following:

- To deal with *problems in the natural language inputs*: an NLP system has to be able to address possible problems of incorrect data, incomplete data or skewed data, all problems quite frequent in real natural language texts.
- To show *how the NLP system reached its result*: to this end, the system has to be able to look from outside at an internal representation or ‘record’, a characteristic that is relevant to developers of the system but also to support accountability.
- To *reason on its internal representation independently from the surface form* of its NLP input.
- To implement *self-awareness* (more philosophical): in order for any system to reason and evaluate its own beliefs and actions autonomously, it needs an internal accessible representation of at least part of it [7, 18].

Focusing on the application of NLP systems to accountability goals, an example for the need of an internal representation is the recent case of the Amazon assistant Alexa, which was faced with the request from a teenager about what to do with annoying parents, and which replied “murder them”, because it had found a perfect match to the input in a ‘for-laughs’ site (<https://tinyurl.com/ybedgm6f>). The system deals with speech (i.e. natural language) input all the time in millions of home, yet has no model of what is doing or what the request-answers mean.

As regards AI systems processing different media, images or videos, when human analyze the accountability implications, they do it using natural language. For example in the government-sponsored panel in Germany to define guidelines for automated and

connected driving (<https://tinyurl.com/y3rf6mgx>), experts had to deal with questions like: “In case of possible accident, should the car prioritize the driver, the passengers or passer-by?” (with various sub-categories considered). The natural language answers then becomes suitable for analysis by a system like CoreSystem, as are the simplified output of the experts. The results can then be fed-back to the car designers, in a more formal and interactive way, helping to bridge the gap between moral experts on the panel and engineers.

3.2 Core Semantic Representation

A core semantic representation in NLP is an internal representation of the text that attempts to describe its meaning in a form that can be (very) different from the original one. Internal semantic representation can be categorised in various ways according to the supported functionalities:

- Disambiguation: does it disambiguate the text; normally disambiguation is understood to cover lexical (nouns and verbs) and some structural (e.g., attachment) elements; a core semantic approach covers also other structures such as prepositions, implicit elements (especially causal and temporal ones, events underlying nouns), redundant structures, etc.
- Normalization: does it normalize the representation, i.e., does it produce the same output for inputs that human would recognize as equivalent, even if the surface forms are very different.
- Relationships: does it make explicit all the implicit relationships: causal, temporal, spatial, and personal.
- Point of views: can the system extrapolate from the narrative point of view in the description; e.g., giving and receiving: “Tom gives a book to Mary” and “Mary receives a book from Tom” have the same deep meaning, but told from different viewpoints.
- Reasoning: does it help reasoning and query answering, by avoiding unnecessary searches, combinatorial explosions, match-failure due to surface elements, etc.

There are some difficulties in implementing such explicit representation using a deep learning approach on its own. The main problems are related to the following issues:

- Lack of data: while there are huge repositories of text translations (e.g., the EU translation repositories) [19, 20] and question answering (large repos exist in call centre databases, etc.), there are fewer such databases of text-internal representation pairs, usually of much smaller dimensions. Note that deep learning usually requires huge amount of data, from millions to billions; e.g., DeepL [21].
- While it can be used on smaller data sets, its general correct coverage on new input decreases noticeably. Parsing is a different issue, in that the parse tree is fundamentally a grammatical structure. Deep learning has been successfully used on TreeBank [22], but this has two key features: (1) all the text used is correct; (2) the representation pairs stored includes semantic decisions (e.g., on attachment) that usually cannot correctly be solved at that scale; (3) the representation is surface-based anyway.

- Existing deep semantic approaches still do not achieve the above requirements. Example of such deep semantics models are: the Mental models [18], based on instantiation, separation of possibilities in different models (cognitive bias), elimination of non-compatible models, self-centred models, etc.; the Conceptual dependency theory model [23], based on primitives chosen away from language, negation of parsing role, no explicit rules for extractions, pre-constructed scripts, difficult handling of negation, etc. Many semantic representation systems carry surface-based elements such as the absence of personal worlds (e.g., AMR, <https://amr.isi.edu>); a case-structure (e.g., FrameNet, <https://framenet.icsi.berkeley.edu>) or the inability to extract implicit causal links.

“Deep semantics” is also used to indicate a number of approaches in which ‘deep’ denotes specific characteristics of the task or of the output. For example, in [24] it is used to indicate a latent semantic model exploiting neural networks to semantic role labelling; the authors propose an approach that does not run any parsing, does not actually recover the full meaning of the sentences and is dependent on various surface elements.

4 CoreSystem

CoreSystem¹ is a prototype NLP system. Its final goal is to obtain an internal representation of the meaning of sentences independent of the surface description and able to explicit the key implicit elements. The current version of the system is based on a sequence of compositional rules. These rules tend to be linked to general semantic properties of the terms or language structures: once specific information is acquired, it can then be used to supplement the analysis with specific domain dependent information. The design is based on the principle of doing what can be done straight away, according to an economy principle, since keeping options open costs efforts (in human is also limited by working memory), while at the same time leaving open the decisions for which there is not enough information at that stage (e.g., for attachment at parsing stage). This can be done without overload by a technique that allows localizing structural ambiguities. Vice versa, where semantic information can be used efficiently early on (e.g., semantic restrictions on verbs), it is incorporated in the parsing stage. Once a surface semantic representation is achieved – i.e., one which transforms the parsing tree into a graph, in which same entities or events are unified, the process to transform it into core semantics begins.

CoreSystem deals only with English, however, given the present proficiency of automatic translation systems, for other languages it is possible to do the following: automatically translate an input in English, convert it in the internal semantics, elaborate it as required, get CoreSystem to generate an English text from such elaboration

¹ The following brief description of the prototype system is provided in order to outline what has been used to produce the analysis below. The system is at present not available for external testing; furthermore, as it is under development, no claims are made here to its coverage or efficiency with respect to other NLP systems.

as required, re-translate in the desired language. Initial experiments in this sense with Italian and Spanish have given positive results.

The analysis process is a step-by-step one, in which a set of rules has been elaborated looking at many written texts from different sources (e.g., Economist, Wikipedia, BBC, Telegraph, Mirror, Bloomberg, Reuters); these rules turn out to be by-and-large domain independent. A version for different types of input (e.g., speech transcription, dialogue, chat, etc.) is under development. CoreSystem satisfies, albeit at different levels, all the requirements described in Sect. 3.2. Its core semantic approach tackles the combinatorial explosion of meaning representations following a strongly minimalist approach that leads to a representation of the content independent of the surface description, including hidden causal, spatial and temporal connections.

The core of the present version is written in Haskell (www.haskell.org). Haskell, a purely functional, strongly typed, lazy, referentially transparent, higher order programming language, is particularly suited for representing very complex set of rules; also, because of its referential transparency, it is not dependent on side effects, which greatly help the managing of a large number of rules working together. Haskell can be run in parallel, either internally or using orchestrating systems such as Erlang. The user interface is implemented in JavaScript, with the logic controlling the display managed from Haskell.

The last version of CoreSystem has been used and preliminary validated, in a national project (Sintesys, <http://www.cerict.it/it/progetti-nazionali-conclusi/281-sintesys-ricerca.html>) and in a European project (LASIE, <http://www.lasie-project.eu>). In these projects, CoreSystem was tasked with analysing texts from similar domains (terrorism for the national project and crime for the European project), while the type of text was different (short, information rich Reuters flash news in Sintesys; long newspaper articles and blogs in LASIE). In both cases, the goal was to produce an analysis that helped investigators in the following ways:

- provide a clear representation of the information and the underlying structures;
- find common references to people, places, organizations, events, etc.;
- connect events along temporal causal and spatial chains;
- extract modal information, such as desires, beliefs, plans, likes, duties, etc.

In order to reach these objectives, the core semantic representation has proved the key feature, since it has allowed to unify apparently different entities and events and to connect them using implicit deep temporal, causal and spatial chains. It has also been essential in extracting motivations, likely actions, elements of planning and other mental structures.

As regards self-awareness, a deep neural network may embody such knowledge, beliefs, etc., but it is distributed, so there is not a part that represents it. For some scholars, such representation has to be symbolic in nature, and separated from the underlying one, in order to avoid infinite regressions [7].

5 Core Semantics and Accountability

To investigate how a core semantic representation could be used to support accountability of AI, in this section the output produced by CoreSystem for sentence (a) is analysed.

A 59-year-old man from York has been arrested on suspicion of murdering missing chef Claudia Lawrence.

In this sentence, a lot of knowledge is implicit, but a reader would be able to interpret it, understanding it as follows: Claudia Lawrence worked as a chef, she disappeared, she may have been murdered, then police suspected that a man murdered her and so they arrested him. The man had been in York before police arrested him, and he was 59 years of age when the police arrested him.

Its surface representation (parsing or semantic) is very distant from its core semantics (Figs. 1a and 1b give the output created by the NLP system split for the sake of readability).

Core semantics means that all the implicit information, as for example the events hidden inside nouns such as ‘suspicion’, adjective such as ‘missing’, or roles such as ‘chef’, has to be extracted and organized in small atomic unit, which then are put together in the correct temporal and causal sequence.

In CoreSystem, information is represented in a graph as objects and events. For the above sentence, the system creates 3 new objects (‘man’, ‘York’ – used in the event which describes the man’s position before the arrest – and ‘police’), an object created for a previous analysis (‘Claudia Lawrence’) and 14 events. Figure 2 presents an extract of the analysis obtained by the CoreSystem applying a core semantics for the sentence: the *event* ‘arrest’, the *object* ‘police’ and the event created to explicit the fact that the man is 59 years old when he is arrested by the police. Both events are interpreted according the ‘subj-action-(obj)’ structure, supplemented by other meta-level information, as the time of the action, the source of the information and the links with other objects and events used by CoreSystem to represent the content of the sentence.

The final analysis is rather distant from the original text, although it is close to how a native speaker would mentally visualize the story [18]. The full analysis is given in the appendix, and the number of the nodes used for the internal representation is reported in what follows. ‘Arrest’ (109608) is an example of a general event marked as ‘prototypical’, which allows to explicit police (172748) as the subject of the arrest. A prototype is a ‘best initial guess’ structure, based on causal models. It is not probabilistic and it can be overruled by more specific information. Other prototypes used to represent the meaning of the sentence are that of ‘murder’ (172705) and ‘suspicion’ (172745). ‘Police’ is also used as the subject of the *event* “Police suspects a man murdering Claudia Lawrence” (172745), the event representing the reason of the arrest, i.e. the causal link between suspicion and arrest. The content of the last part of the sentence is

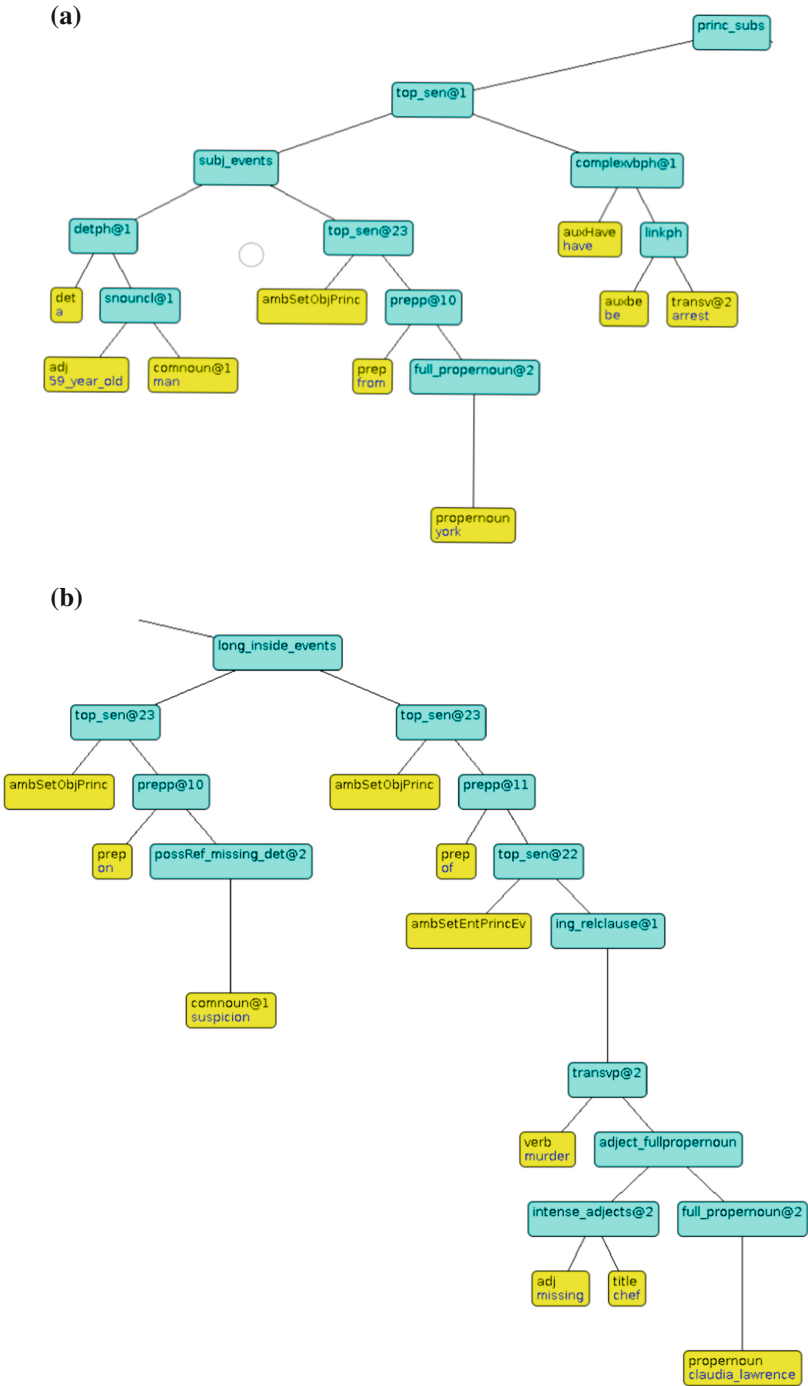


Fig. 1. Parse tree for sentence (a) first subtree (b) second subtree

represented by the following *events*: “Claudia Lawrence works as Chef” (171759), “Claudia Lawrence disappears” (172744), and “Claudia Lawrence works as Chef, before she disappears” (172780). The murder, the suspicion and the arrest are then connected by the *event* “Police suspects a man murdering Claudia Lawrence so they arrest him” (170891). The other events are needed to represent the causal, spatial and temporal relations among the events in the original sentence.

<pre>* arrest/1: 109608 * universal_: Event - 74883 - rank: universal arrest/1 - 820 - subject_: police/2 - 172748 - rank: universal action_: arrest/4 - 823 - object_: man/2 - 79018 - rank: individual time_: present_ - 248575 - source_: speaker_ - 19845 - rank: named individual object_of: Event - 170891 - rank: individual Event - 172767 - rank: individual Event - 172779 - rank: individual ***** event: Police arrests a man.</pre>	<pre>* police/2: 172748 * generalisation_: police/2 - 171402 - rank: universal subject_of: arrest/1 - 109608 - rank: individual suspicion/1 - 172745 - rank: individual ***** object: Police. * Event: 172767 * universal_: Event - 74883 - rank: universal subject_: Event - 79016 - rank: individual action_: during/2 - 61250 - object_: arrest/1 - 109608 - rank: individual ***** event: A man having age 59 during police arresting him.</pre>
--	---

Fig. 2. Examples of events and objects used for representing core semantics

Notice that the fact that the arrested man murdered Claudia Lawrence is marked as hypothetical, since it exists at present only in police’s suspicion. This, as well as other phenomena such as negation, desires, different beliefs etc., is modelled using a many-worlds semantics, some of which may be incompatible with each other.

For what concerns accountability, such representation could be used to explain some facts and actions, answering questions as “Who was arrested?”; “Where is he from?”; “Why was he arrested?”; “What was she doing for a living” but also “Who arrested him”, even if this information is not explicit in the sentence. While, what is the name of the arrested man could be answered saying “I do not know”, being that a statement about the knowledge in the AI system and not about not being able to extract such knowledge from the text. Also important is the source information given for an event (Fig. 2).

Question answering is an obvious way to achieve both explanation and accountability, as is normally done among humans. However, CoreSystem also produces a graphical view of the causal-personal structure of the model, which allows for easy understanding (and even hand-modifications if needed). Figure 3 shows a simplified screen shot for the graphical output for the text “A man, believed to be a member of an unknown Muslim militant group, planted five gasoline bombs on a bus carrying German tourists in Cairo. A guide saw the man put a bag under a seat on the bus and called the police. The man was arrested and a bomb disposal crew removed the bombs. No injuries were reported.” In the interactive version, by clicking on the various links the viewer can see the different types of causality, inspect the standard models behind etc.

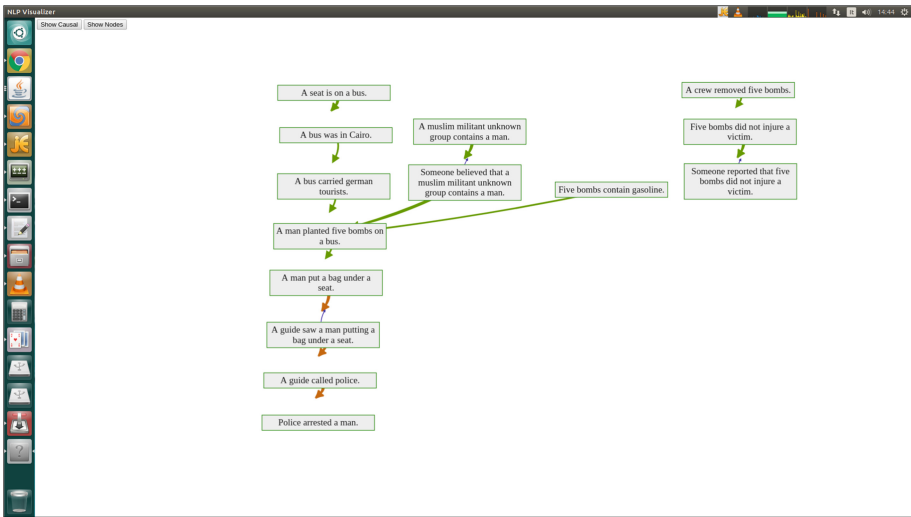


Fig. 3. Screenshot of the graphical representation produced by CoreSystem

6 Conclusions

The importance of accountability and the need for a core semantics which can fully understand the meaning of natural language processing texts have recently been underlined in two interventions by leading AI scientists [25, 26]. Focusing on AI systems embedding any form of NLP, in this paper we investigated how a core semantic approach could be used to address those concerns. In general, core semantics embedded in various AI applications would greatly help in assessing systems' performance, as well as allowing the systems themselves to have an image of their own high-level processing. The final goal of an NLP system is to extract from natural language documents core semantic version that clearly shows the crucial causal and temporal links, and this is a pre-requisite to use the NLP system to support accountability. As a proof-of-concept, the practicability of the core semantics has been tested using a prototype large-scale NLP system. For accountability goals, the example illustrated in Sect. 5 indicates that a core semantics produces textual representations that can be easily understood and checked by human. The next step is to provide also a graphical representation and the NLP query answering. To design and implement accountability functionalities as NLP system module, the large amount of knowledge used for the analysis of a single statement highlights that it is critical to be able to deal with the combinatorial explosion of the graph. Besides, according to software engineering best practices, interfaces supporting final users, and not only developers, have to satisfy usability and performance requirements.

Acknowledgments. As researchers in natural language processing and requirements engineering, authors shared a number of papers with Stefania Gnesi and her research group since the early 1990s. She is a passionate scientist, and these exchanges resulted in a fruitful and enriching relationship.

Appendix A

Representation of the Meaning of the Sentence: “A 59-year-old man from York has been arrested on suspicion of murdering missing chef Claudia Lawrence”.

```

* Event: 79016 *
universal_ :
Event - 74883 - rank: universal
subject_ :
man/2 - 79018 - rank: individual
action_ :
have_age/1 - 258771 -
object_ :
59 - 258887 - rank: universal
time_ :
present_ - 248575 -
source_ :
speaker_ - 19845 - rank: named individual
subject_of:
Event - 172767 - rank: individual
*****
event:
A man has age 59.
*****

* man/2: 79018 *
universal_ :
man/2 - 79015 - rank: universal
murderer/1 - 150961 - rank: universal
suspect/3 - 189089 - rank: universal
subject_of:
Event - 79016 - rank: individual
Event - 79021 - rank: individual
murder/2 - 172705 - rank: individual - hypothe-
sis_
object_of:
arrest/1 - 109608 - rank: individual
*****
object:
A man.
*****

* Event: 79021 *
universal_ :
Event - 74883 - rank: universal
subject_ :
man/2 - 79018 - rank: individual
action_ :
be/12 - 15902 -
location_ :
position_ - 79022 - rank: individual

```

```

time_ :
past_ - 248407 -
source_ :
speaker_ - 19845 - rank: named individual
subject_of:
Event - 172779 - rank: individual
*****
event:
A man was in York.
*****

* position_ : 79022 *
universal_ :
position_ - 11456 -
subject_of:
Event - 79023 - rank: individual
location_of:
Event - 79021 - rank: individual
*****
object:
In York.
*****

* Event: 79023 *
universal_ :
Event - 74883 - rank: universal
subject_ :
position_ - 79022 - rank: individual
action_ :
in/1 - 119448 -
object_ :
york/5 - 247198 - rank: named individual
*****
event:
Is in York.
*****

* arrest/1: 109608 *
universal_ :
Event - 74883 - rank: universal
arrest/1 - 820 -
subject_ :
police/2 - 172748 - rank: universal
action_ :
arrest/4 - 823 -
object_ :
man/2 - 79018 - rank: individual

```

time_ :
 present_ - 248575 -
 source_ :
 speaker_ - 19845 - rank: named individual
 object_of:
 Event - 170891 - rank: individual
 Event - 172767 - rank: individual
 Event - 172779 - rank: individual

 event:
 Police arrests a man.

 * Event: 170891 *
 universal_ :
 Event - 74883 - rank: universal
 subject_ :
 suspicion/1 - 172745 - rank: individual
 action_ :
 cause/3 - 33341 -
 object_ :
 arrest/1 - 109608 - rank: individual
 source_ :
 speaker_ - 19845 - rank: named individual

 event:
 Police suspects a man murdering Claudia Lawrence so they arrest him.

 * Event: 171759 *
 universal_ :
 Event - 74883 - rank: universal
 subject_ :
 claudia_lawrence/1 - 258841 - rank: named individual
 action_ :
 work_as/1 - 172933 -
 object_ :
 chef/1 - 44708 - rank: universal
 source_ :
 speaker_ - 19845 - rank: named individual
 subject_of:
 Event - 172780 - rank: individual

 event:
 Claudia Lawrence works as Chef.

 * murder/2: 172705 *
 universal_ :

murder/2 - 157791 - rank: universal
 subject_ :
 man/2 - 79018 - rank: individual
 action_ :
 murder/1 - 154816 -
 object_ :
 claudia_lawrence/1 - 258841 - rank: named individual
 source_ :
 speaker_ - 19845 - rank: named individual
 status_ :
 hypothesis_ - 248409 -
 subject_of:
 Event - 172750 - rank: individual
 object_of:
 Event - 172766 - rank: individual
 suspicion/1 - 172745 - rank: individual

 event:
 A man may murder Claudia Lawrence.

 * Event: 172744 *
 universal_ :
 Event - 74883 - rank: universal
 subject_ :
 claudia_lawrence/1 - 258841 - rank: named individual
 action_ :
 disappear/1 - 70373 -
 source_ :
 speaker_ - 19845 - rank: named individual
 subject_of:
 Event - 172766 - rank: individual
 object_of:
 Event - 172780 - rank: individual

 event:
 Claudia Lawrence disappears.

 * suspicion/1: 172745 *
 universal_ :
 suspicion/1 - 189007 - rank: universal
 subject_ :
 police/2 - 172748 - rank: universal
 action_ :
 suspect/4 - 189090 -
 object_ :

murder/2 - 172705 - rank: individual - hypothesis_
time_
present_ - 248575 -
subject_of:
Event - 170891 - rank: individual
object_of:
Event - 172750 - rank: individual

event:
Police suspects a man murdering Claudia Lawrence.

* police/2: 172748 *
generalisation_
police/2 - 171402 - rank: universal
subject_of:
arrest/1 - 109608 - rank: individual
suspicion/1 - 172745 - rank: individual

object:
Police.

* Event: 172750 *
universal_
Event - 74883 - rank: universal
subject_
murder/2 - 172705 - rank: individual - hypothesis_
action_
before/2 - 15971 -
object_
suspicion/1 - 172745 - rank: individual

event:
A man may murder Claudia Lawrence, before police suspects this.

* Event: 172766 *
universal_
Event - 74883 - rank: universal
subject_
Event - 172744 - rank: individual
action_
before/2 - 15971 -
object_:

murder/2 - 172705 - rank: individual - hypothesis_

event:
Claudia Lawrence disappears, before a man may murder her.

* Event: 172767 *
universal_
Event - 74883 - rank: universal
subject_
Event - 79016 - rank: individual
action_
during/2 - 61250 -
object_
arrest/1 - 109608 - rank: individual

event:
A man having age 59 during police arresting him.

* Event: 172779 *
universal_
Event - 74883 - rank: universal
subject_
Event - 79021 - rank: individual
action_
before/2 - 15971 -
object_
arrest/1 - 109608 - rank: individual

event:
A man was in York, before police arrests him.

* Event: 172780 *
universal_
Event - 74883 - rank: universal
subject_
Event - 171759 - rank: individual
action_
before/2 - 15971 -
object_
Event - 172744 - rank: individual

event:
Claudia Lawrence works as Chef, before she disappears.

References

1. LeCun, Y., Yoshua, B., Geoffrey, H.: Deep learning. *Nature* **521**(7553), 436 (2015)
2. Davis, A.: How artificial intelligence has crept into our everyday lives. IEEE Special Report (2016). <http://theinstitute.ieee.org/static/special-report-artificial-intelligence>
3. LinkedIn: Global recruiting trends 2018 (2018). https://business.linkedin.com/talent-solutions/recruiting-tips/2018-global-recruiting-trends?trk=bl-ba_global-recruiting-trends-launch_maria-ignatova_011018
4. Kaplan, J.: Artificial intelligence: think again. *Commun. ACM* **60**(1), 36–38 (2016). <https://doi.org/10.1145/2950039>
5. Internet Society: Artificial Intelligence and Machine Learning: Policy paper (2017). <https://www.internetsociety.org/resources/doc/2017/artificial-intelligence-and-machine-learning-policy-paper>
6. Parnas, D.L.: The real risks of artificial intelligence. *Commun. ACM* **60**(10), 27–31 (2017). <https://doi.org/10.1145/3132724>
7. Shanahan, M.: *The Technological Singularity*. MIT Press, New York (2015)
8. Anthes, G.: Artificial intelligence poised to ride a new wave. *Commun. ACM* **60**(7), 19–21 (2017). <https://doi.org/10.1145/3088342>
9. Guidotti, R., Monreale, A., Pedreschi, D.: The AI black box explanation problem. *ERCIM News* **116**, 12–13 (2019)
10. ACM U.S. Public Policy Council, ACM Europe Policy Committee: Statement on algorithmic transparency and accountability (2017). https://www.acm.org/binaries/content/assets/public-policy/2017_joint_statement_algorithms.pdf
11. Doshi-Velez, F., et al.: Accountability of AI under the law: the role of explanation. *CoRR* abs/1711.01134 (2017)
12. O’Neil, C.: *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York Time (2016)
13. Pell, D.: The 10 algorithms that rule the world and other fascinating news on the web. *Time* (2014). <http://time.com/111313/the-10-algorithms-that-rule-the-world-and-other-fascinating-news-on-the-web>
14. Barry-Jester, A.M., Casselman, B., Goldestein, D.: The Marshall Project. The new science of sentencing (2015). <https://www.themarshallproject.org/2015/08/04/the-new-science-of-sentencing#bwuhXcwqn>
15. Buregess, M.: Holding AI to account: will algorithms ever be free from bias if they’re created by humans? *Wired*, UK (2016). <http://www.wired.co.uk/article/creating-transparent-ai-algorithms-machine-learning>
16. Kirkpatrick, K.: Battling algorithmic bias: how do we ensure algorithms treat us fairly? *Commun. ACM* **59**(10), 16–17 (2016). <https://doi.org/10.1145/2983270>
17. Gaines, B.: Designing expert systems for usability. In: Shackel, B., Richardson, S.J. (eds.) *Human Factors for Informatics Usability*, pp. 207–246. Cambridge University Press, New York (1991)
18. Johnson-Laird, P.: *How We Reason*. Oxford University Press, New York (2006)
19. Wong, S.: Google Translate AI invents its own language to translate with. *Daily News* (2016). <https://www.newscientist.com/article/2114748-google-translate-ai-invents-its-own-language-to-translate-with>
20. Reynolds, M.: Google uses neural networks to translate without transcribing. *Daily News* (2017). <https://www.newscientist.com/article/2126738-google-uses-neural-networks-to-translate-without-transcribing>

21. Coldewey, D., Lardinois, F.: DeepL schools other online translators with clever machine learning. Techcrunch (2017). <https://techcrunch.com/2017/08/29/deepl-schools-other-online-translators-with-clever-machine-learning>
22. Mitchell, M.P., Santorini, B., Marcinkiewicz, M.A., Taylor, A.: Treebank-3 LDC99T42 Web Download. Linguistic Data Consortium, Philadelphia (1999)
23. Schank, S., Abelson, R.: Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures. Lawrence Erlbaum, Hillsdale (1997)
24. He, L., Lee, K., Lewis, M., Zettlemoyer, L.: Deep semantic role labeling: what works and what's next. In: Proceedings of the 55th Annual Meeting Association for Computational Linguistics, vol. 1, pp. 473–483 (2017)
25. Jordan, M.: Artificial Intelligence. The revolution hasn't happened yet. Medium (2018). <https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7>
26. Hutson, M.: AI researchers allege that machine learning is alchemy. Science **360**(6388), 861 (2018). <http://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy>