The Candide Model: How Narratives Emerge Where Observations Meet Beliefs

Paul Van Eecke[1] and **Lara Verheyen**[1] and **Tom Willaert**[2,3] and **Katrien Beuls**[4]
[1]Artificial Intelligence Laboratory, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels
[2]Brussels School of Governance, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels
[3]imec-SMIT, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels
[4]Faculté d'informatique, Université de Namur, rue Grandgagnage 21, B-5000 Namur
paul@ai.vub.ac.be, lara.verheyen@ai.vub.ac.be,
tom.willaert@vub.be, katrien.beuls@unamur.be

Abstract

This paper presents the Candide model as a computational architecture for modelling human-like, narrative-based language understanding. The model starts from the idea that narratives emerge through the process of interpreting novel linguistic observations, such as utterances, paragraphs and texts, with respect to previously acquired knowledge and beliefs. Narratives are personal, as they are rooted in past experiences, and constitute perspectives on the world that might motivate different interpretations of the same observations. Concretely, the Candide model operationalises this idea by dynamically modelling the belief systems and background knowledge of individual agents, updating these as new linguistic observations come in, and exposing them to a logic reasoning engine that reveals the possible sources of divergent interpretations. Apart from introducing the foundational ideas, we also present a proof-of-concept implementation that demonstrates the approach through a number of illustrative examples.

1 Introduction

Today's natural language processing (NLP) systems excel at exploiting the statistical properties of huge amounts of textual data to tackle a wide variety of NLP subtasks. They meticulously capture the co-occurrence of characters, words and sentences, sometimes in relation to an annotation layer, and make use of numerical operations over these co-occurrences to perform mappings between linguistic input on the one hand and task-specific linguistic or non-linguistic output on the other. As a result of recent advances in neural machine learning techniques and infrastructure (see e.g. Sutskever et al., 2014; Vaswani et al., 2017; Devlin et al., 2018), combined with the availability of huge text corpora, impressive results are now being achieved on many tasks, including machine translation, speech recognition, text summarisation, semantic role labelling and sentiment analysis (for an overview, see Lauriola et al., 2022).

Yet, current NLP systems are everything but capable of modelling human-like, narrative-based language understanding. One reason is that this capacity is hard to cast in the predominant machine learning paradigm. Indeed, human-like narrative understanding is hard to define in the form of an annotation scheme. Narratives are not captured in texts as such, but are construed through an interpretation process. This process is personal, and different individuals may construe different narratives given the same linguistic observations (Steels, 2022). This diversity in perspectives reflects the richness of human language and cognition, and modelling divergent interpretations constitutes a crucial challenge to the broader computational linguistics community today.

The primary objective of this paper is to introduce a novel approach to narrative-based language understanding that starts from the idea that narratives emerge through the process of interpreting novel observations with respect to previously acquired knowledge and beliefs. Concretely, we present a computational model of this interpretation process. The model integrates three main components: (i) a personal dynamic memory that holds a frame-based representation of the knowledge and beliefs of an individual agent, (ii) a construction grammar that maps between linguistic observations and a frame-based representation of their meaning, and (iii) a reasoning engine that performs logic inference over the information stored in the personal dynamic memory.

Crucially, the representations that result from the language comprehension step take the same form as those stored in the personal dynamic memory. Not only does this mean that these representations can dynamically be merged into the personal dynamic memory to update the knowledge and beliefs of an agent, it also facilitates the use of information

Proceedings of the The 5th Workshop on Narrative Understanding, pages 48–57

stored in the personal dynamic memory to inform the language comprehension process. The information stored in the personal dynamic memory can be queried through a logic reasoning engine, with each answer being supported by a human-interpretable chain of reasoning operations. This chain of reasoning operations explains how the background knowledge and beliefs of an agent guide its conclusions, thereby revealing the narrative construed through the agent's interpretation process.

Personal, dynamic and interpretable models of narrative-based language understanding are of great interest to the fields of computational linguistics and artificial intelligence alike. To the field of computational linguistics, they contribute a perspective that emphasises the individual and contextualised nature of linguistic communication, which contrasts with the static and perspective-agnostic models that dominate the field of NLP today. In the field of artificial intelligence, they respond to the growing interest in the development of artificial agents that combine human-like language understanding with interpretable, value-aware and ethics-guided reasoning (see e.g. Steels, 2020; Montes and Sierra, 2022; Abbo and Belpaeme, 2023).

The remainder of this paper is structured as follows. Section 2 lays out the background and overall architecture of our model. Section 3 presents its technical operationalisation and provides a number of illustrative examples. Finally, Section 4 reflects on the contribution of our paper and discusses avenues for further research.

2 The Candide Model

The model for narrative-based language understanding that we introduce in this paper is named after Voltaire's "*Candide ou l'optimisme*" (Voltaire, 1759). It is inspired by one of the main themes of the novel, namely that a character's belief system and history of past experiences shape the way in which they interpret the world in which they live. As such, different characters in the novel represent different philosophical positions and thereby construe different narratives to explain the same situations and events. The main protagonist, Candide, starts out as a young, naive 'blank slate'. Through conversations with the Leibnizian optimist Pangloss and the fatalistic pessimist Martin, and as a result of long travels that make him experience the hardships of the world, Candide gradually develops his own belief system in light of which he ever more wisely interprets the situations and events he witnesses.

Following the main theme of the novel, our aim is not to model a single 'true' interpretation of an observation, but to show that different beliefs can lead to different interpretations. Moreover, we consider the belief system of an agent to be dynamic, with the interpretations and conclusions of an agent shifting as more experience and knowledge are gathered. In order to formalise these high-level ideas, we introduce the following operational definitions:

Personal dynamic memory (PDM) The personal dynamic memory of an agent is a data structure that stores the knowledge and beliefs of the agent in a logic representation that supports automated reasoning. The PDM is conceived of as a dynamic entity to which new knowledge and beliefs can be added at any point in time. Reasoning over the PDM is non-monotonic, as updated beliefs can alter conclusions.

Belief system The belief system of an agent at a given point in time equals all information that is stored in the agent's PDM at that moment in time. Each entry in the PDM carries a confidence score, which reflects the degree of certainty of the agent with respect to that entry. However, there exists no formal or conceptual distinction between entries based on their epistemological status, avoiding the need to distinguish between 'knowledge', 'facts', 'opinions' and 'beliefs' for example.

Conclusion A conclusion is a piece of information that logically follows from a reasoning operation over the belief system of an agent. A typical example would be the answer to a question.

Narrative A narrative is defined as a chain of reasoning operations that justifies a conclusion based on the belief system of an agent as it is stored in its PDM. Logically, it corresponds to a proof for the conclusion. It is possible that multiple narratives that support the same or different conclusions can be construed by an individual agent. An agent can use the certainty scores carried by the beliefs that constitute its PDM to rank its conclusions and the narratives that support them.

Language comprehension Language comprehension is the process of mapping a linguistic observation, such as an utterance, paragraph or text,

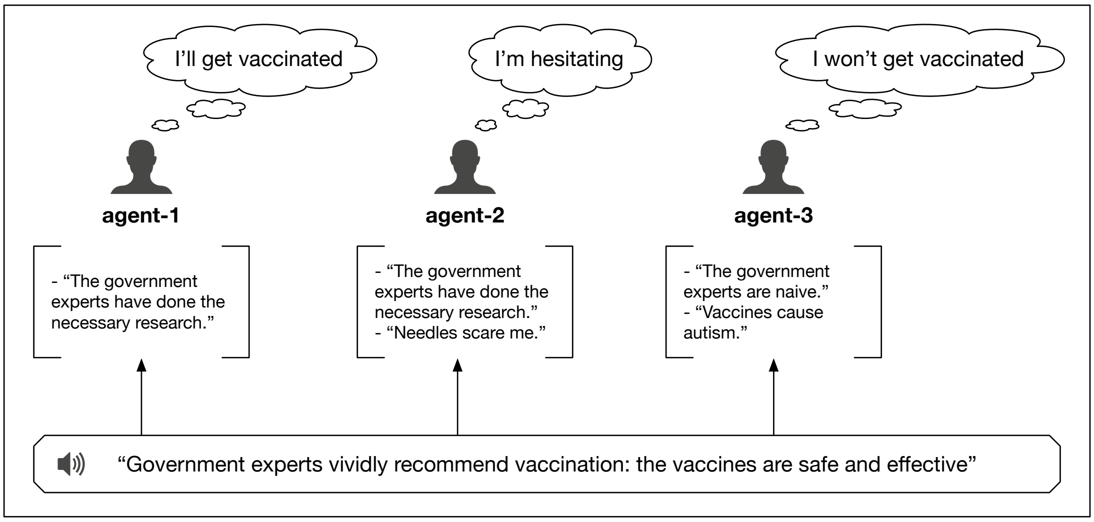

Figure 1: Informal sketch of the Candide model. The model conceives of narrative-based language understanding as the interpretation of a linguistic observation with respect to an agent's individual belief system. Narratives are defined as argumentation structures on the basis of which conclusions are drawn.

to a logic representation of its meaning. While language comprehension is primarily concerned with retrieving the information captured in the linguistic input, rather than its integration with respect to the personal dynamic memory, it is heavily intertwined with other aspects of the interpretation process as well. Indeed, the linguistic knowledge needed to support language comprehension is personal and dynamic, and thereby unavoidably constitutes a first layer of individual interpretation.

Interpretation The interpretation process comprises all aspects involved in narrative-based language understanding, from the linguistic input to the construction of a narrative that justifies a conclusion. This involves both the language comprehension process, which maps from linguistic input to a logic representation of its meaning, and the reasoning processes that corroborate this meaning representation with the information stored in the PDM, thereby construing narratives that support conclusions.

An informal example of the main ideas underlying the Candide model is shown in Figure 1. Here, three agents observe the same broadcasted message "*Government experts vividly recommend vaccination: the vaccines are safe and effective*" and are asked whether they plan to get vaccinated. In order to answer the question, the three agents individually interpret this message with respect to the beliefs stored in their PDM and construe a narrative that justifies a conclusion in the form of an answer to the question. The first agent comes to the conclusion that they will get vaccinated, justifying their choice through the narrative that the government experts are competent. The second agent is hesitant to get vaccinated, construing the narrative that vaccines are beneficial but that they are scared of needles. The third agent will not get vaccinated, as they construe the narrative that vaccines are dangerous and that the government experts are being misled.

The example illustrates three properties of narratives that, in our view, constitute crucial challenges in operationalising narrative-based language understanding. First of all, a model of analysis can only be adequate if it captures the personal nature of narratives. Whether or not a conclusion is justified does not depend on its truth or falsehood from an external perspective, but only on whether it is supported by the beliefs held by an agent. Second, narratives are not captured as such in linguistic artefacts. While authors convey messages that are grounded in their belief systems, these messages do not encode the belief systems themselves. Indeed, the intended meaning underlying a message needs to be reconstructed inferentially based on the belief system of the receiver (Grice, 1967; Sperber and Wilson, 1986). Finally, it is essential that the interpretation process that is modelled is transparent and human-interpretable. The goal is not merely to draw conclusions given linguistic input, but to reveal the background knowledge, beliefs and reasoning processes that underlie the conclusions that are drawn.

3 Technical Operationalisation

This section presents the technical operationalisation of an initial proof-of-concept of the Candide model. We discuss the proof-of-concept's language comprehension component, its personal dynamic memory, and its processes of reasoning and narrative construction.

3.1 Language Comprehension

The language comprehension component is responsible for mapping between linguistic input, in particular utterances, paragraphs and texts, and a formal representation of their underlying meaning. The language comprehension component is operationalised using the Fluid Construction Grammar framework (FCG – `https://fcg-net.org`; Steels, 2004; van Trijp et al., 2022; Beuls and Van Eecke, 2023). The FCG framework provides a computational operationalisation of the basic tenets of construction grammar (Fillmore, 1988; Goldberg, 1995; Croft, 2001; Fried and Östman, 2004; Beuls and Van Eecke, 2024). It includes a formalism for representing construction grammars, a processing engine that supports construction-based language comprehension and production, and a library of operators for learning construction grammars in a usage-based fashion.

The choice for FCG as the backbone of the language comprehension component of our proof-of-concept is motivated by four main reasons. First of all, in line with its theoretical grounding in usage-based construction grammar, FCG offers a uniform way to represent and process linguistic phenomona, whether or not they can be analysed compositionally (Beuls and Van Eecke, 2023). Second, FCG is compatible with a wide variety of meaning representations (van Trijp et al., 2022), including the frame-semantic representation that will be used to represent the knowledge and beliefs captured in the personal dynamic memory of the agents. Third, FCG's symbolic learning operators are especially designed to facilitate the one-shot learning of constructions given new linguistic observations, thereby maximally reflecting the personal and dynamic nature of an agent's linguistic capacities (Van Eecke, 2018; Nevens et al., 2022; Doumen et al., 2023). Finally, the symbolic data structures and unification-based processing algorithms employed by FCG ensure that the representation of an agent's linguistic knowledge, as well as its language comprehension, production and learning processes, are transparent and human-interpretable (Van Eecke and Beuls, 2017).

We opt for a semantic representation that captures the meaning underlying linguistic expressions in the form of semantic frames (Fillmore, 1976; Fillmore and Baker, 2001). Semantic frames represent situations, which are evoked by linguistic expressions, along with their participants. As such, the meaning of the utterance "*Sam sent Robin a postcard*" could be represented through a SENDING frame, with "*Sam*", "*Robin*" and "*a postcard*" respectively taking up the roles of SENDER, RECIPIENT and THEME. In terms of data structures, we represent instances of semantic frames through two types of predicates: entities and roles. Entity predicates are used to represent referents, i.e. objects, people, events and situations that can be referred to. In our example, *Sam*, *Robin*, *the postcard*, *the sending event* and *the transfer situation* serve as entities. Role predicates are used to represent relations between entities. Each role predicate expresses a relation between a frame role (e.g. SENDER), the frame to which that role is associated (SENDING), the entity that is taking up the role (*Sam*), the entity that represents the frame instance (*the sending event*) and the entity that represents the situation about which the frame is expressed (*the transfer situation*). There exists a subtle yet important distinction between frame instances and situations. A situation is defined in terms of an agent's world model, while a frame instance assumes a linguistically expressed perspective on a situation. In our example, the transfer situation is linguistically expressed as a sending event, while the same situation could also have been expressed as a receiving event (e.g. "*Robin received a postcard from Sam*"). Note that both the frame instance and the situation are reified as entities and can thus be referred to. The entity and role predicates follow the FrameNet conventions (`https://framenet.icsi.berkeley.edu`) and are represented in standard Prolog syntax (ISO/IEC 13211), as exemplified in Listing 1.

The exact way in which the FCG engine maps between utterances and their frame-semantic representation, as well how FCG grammars can be designed or learnt, fall outside the scope of this paper. Instead, we refer the interested reader to van Trijp et al. (2022), Nevens et al. (2022), Doumen et al. (2023) and Van Eecke et al. (2022).

```
% Entity predicates

entity(sam).
entity(robin).
entity(postcard).
entity(sending_event).
entity(transfer_situation).

% Role predicates

role(sender,sending,sam,sending_event,transfer_situation).
role(recipient,sending,robin,sending_event,transfer_situation).
role(theme,sending,postcard,sending_event,transfer_situation).
```

Listing 1: Frame-semantic representation underlying the utterance "*Sam sent Robin a postcard*" as a combination of entity and role predicates expressed in standard Prolog syntax.

3.2 Personal Dynamic Memory

The personal dynamic memory of an agent holds a frame-based representation of the agent's belief system. Technically, it consists of a collection of Prolog facts and rules. Instances of semantic frames are expressed by means of entity and role predicates, just like those resulting from the language comprehension process. For the purposes of this section, we will assume that our agents observe the utterance "*Sam sent Robin a postcard*", comprehend it into the frame-based semantic representation shown in Listing 1, and add this representation to their personal dynamic memory. We will also assume that our agents already hold a number of previously acquired beliefs, in particular about the relation between the semantic frames of SENDING and RECEIVING. As such, they believe that the DONOR role in an instance of the RECEIVING frame, cast over a particular situation, is taken up by the same entity that takes up the SENDER role in an instance of the SENDING frame cast over the same situation. However, this alignment only holds under the condition that the postal services are operational. In other terms, each sending event corresponds to a receiving event if the postal services are operational, and the sender of the sending event corresponds to the donor of the receiving event. At the same time, the agents believe that a similar alignment can be made for the other roles of the SENDING and RECEIVING frames. Moreover, they believe that the postal services are operational if no general strike is taking place. A formal encoding of these beliefs is shown in Listing 2.

While our agents hold the same beliefs about the relation between the SENDING and RECEIVING frames, as well as the conditions under which the postal services are operational, they hold differ-

```
% Belief about the operationality of the mail

mail_operational :- not(general_strike).

% Beliefs about the relation between the sending
% frame and the receiving frame

role(donor,receiving,Entity,_,Situation) :-
    role(sender,sending,Entity,_,Situation),
    !, mail_operational.

role(recipient,receiving,Entity ,_,Situation) :-
    role(recipient,sending,Entity,_,Situation),
    !, mail_operational.

role(theme,receiving,Entity,_,Situation) :-
    role(theme,sending,Entity,_,Situation),
    !, mail_operational.

role(sender,sending,Entity,_,Situation) :-
    role(donor,receiving,Entity,_,Situation),
    !, mail_operational.

role(recipient,sending,Entity,_,Situation) :-
    role(recipient,receiving,Entity,_,Situation),
    !, mail_operational.

role(theme,sending,Entity,_,Situation) :-
    role(theme,receiving,Entity,_,Situation),
    !, mail_operational.
```

Listing 2: The beliefs of our example agents concerning the operationality of the mail and the conditional alignment between the SENDING and RECEIVING frames.

```
% Belief about the state of social unrest

general_strike :- false.
```

Listing 3: Agent 1's belief that there is no general strike.

```
% Belief about the state of social unrest

general_strike :- true.
```

Listing 4: Agent 2's belief that there is a general strike.

```
% Query

?- role(theme,receiving,What,Event,Situation),
   role(recipient,receiving,robin,Event,Situation),
   role(donor,receiving,sam,Event,Situation).

% Answer by Agent 1:

What = postcard,
Situation = transfer_situation.

% Answer by Agent 2:

false.
```

Listing 5: Frame-semantic representation underlying the question "*What did Robin receive from Sam?*" with two different answers as computed by the Prolog engine based on the PDMs of Agent 1 and Agent 2.

ent beliefs about the current state of social unrest. As such, Agent 1 believes that there is no general strike, while Agent 2 believes that a general strike is going on at the moment. These beliefs are formally encoded in Listing 3 and 4 respectively.

We define the PDM of Agent 1 to be the combination of the facts and rules specified in Listings 1, 2 and 3, and the PDM of Agent 2 to consist of the facts and rules specified in Listings 1, 2 and 4. Our proof-of-concept implementation does not address the issue of modelling the confidence of an agent with respect to its individual beliefs. The most straightforward way to operationalise this in the current proof of concept would be to use probabilistic logic programming, e.g. through ProbLog (De Raedt et al., 2007).

Our model does not make any assumptions about the origin of the beliefs captured in the personal dynamic memory of an agent. Beliefs can result from the language comprehension process, from abductive reasoning processes, or could even by designed by a knowledge engineer.

3.3 Reasoning and narrative construction

As the beliefs stored in the personal dynamic memory of an agent and the meaning of natural language utterances as comprehended by an agent are both represented as a collection of Prolog facts and rules, logical reasoning can naturally be operationalised through SLD-resolution-based inference. This means that agents can be asked to prove logic formulae that correspond to natural language questions. The conclusion of the proof then constitutes the answer to the question, while the proof itself corresponds to the narrative that explains the reasoning behind it (see Section 2).

Suppose that we ask our two example agents to answer the question "*What did Robin receive from Sam?*". The agents first use their grammar to comprehend this question into its frame-semantic representation, as shown at the top of Listing 5. The interrogative nature of the question is reflected by the presence of variables in the semantic representation, denoted by symbols starting with a capital letter. In this case, we are primarily interested in the entity taking up the role of THEME in the receiving event, represented by the variable `What`. The agents are then asked to find a proof for the meaning representation of the question, given the beliefs stored in their respective personal dynamic memories.

Agent 1 reasons that the `transfer_situation` that was previously described (see Listing 1) can be viewed as an instance of the RECEIVING frame, given the facts (i) that there is no general strike, (ii) that the mail service is therefore operational, and (iii) that the `transfer_sitation` is already believed to be an instance of the SENDING frame in which `robin` takes up the role of RECIPIENT and `sam` the role of SENDER. The agent comes to the conclusion that this reasoning process is (only) valid under the condition that the variables `What` and `Situation` are bound to the values `postcard` and `transfer_situation` respectively. In other terms, Agent 1 comes to the conclusion that Robin received the postcard that was sent to them by Sam.

Agent 2 on the other hand reasons that it knows of no situation that could be viewed as a receiving event in which `sam` and `robin` take up the roles of DONOR and RECIPIENT respectively. Although this agent holds the same beliefs as Agent 1 when it comes to the link between the `sending`

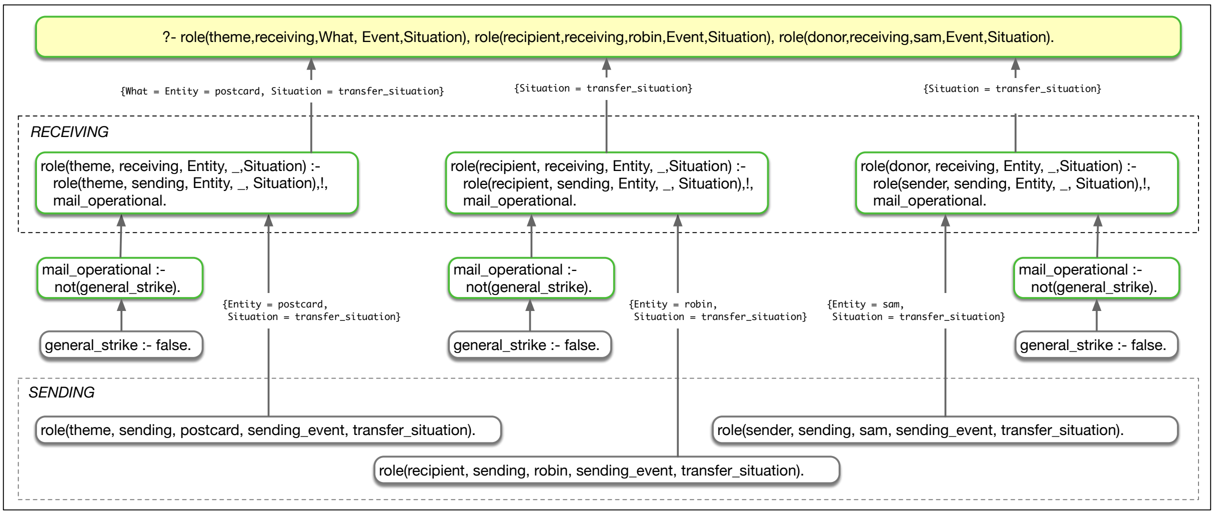

Figure 2: Narrative constructed by Agent 1 for responding to the question "*What did Robin receive from Sam?*" based on the frame-semantic information captured in its PDM (cf. Listings 1, 2 and 3).

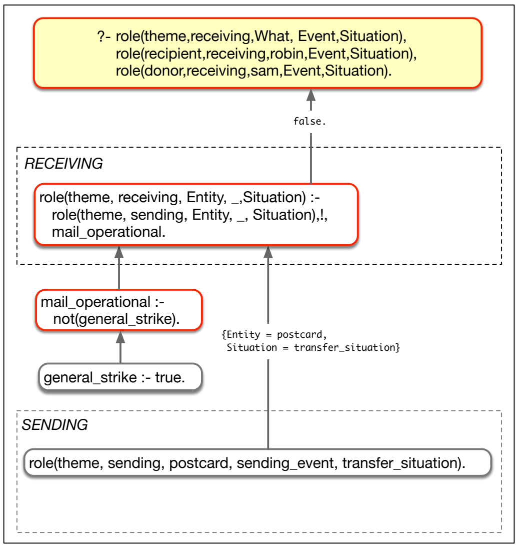

Figure 3: Narrative constructed by Agent 2 for responding to the question "*What did Robin receive from Sam?*" based on the frame-semantic information captured in its PDM (cf. Listings 1, 2 and 4).

and `receiving` frames, Agent 2's belief that a general strike is going on leads to the belief that the postal services are disfunctional, which in turn leads to the belief that the sending event cast over `transfer_situation` does not correspond to any receiving event. In other terms, Agent 2 beliefs that, while a postcard was sent by Sam to Robin, it was never received at Robin's end because of a general strike that paralysed the postal services.

Figures 2 and 3 show a schematic overview of the different steps involved in the respective reasoning processes of Agent 1 and Agent 2 when asked to answer the question "*What did Robin receive from Sam?*". The meaning representation of the question is shown in the yellow boxes at the top of the figures and corresponds to a Prolog query. The facts and rules that can be used to prove the query are those stored in the personal dynamic memories of the agents and correspond to those presented in Listings 1, 2 and 3 (Agent 1) and Listings 1, 2 and 4 (Agent 2).

The conjunction of three clauses that constitutes the query can indeed be proven by Agent 1 through a chain of subproofs that establish the link between there not being a general strike, the operationality of the postal services and the alignment of the SENDING and RECEIVING frames. The solid arrows denote the subproofs that were used to prove the top-level query. The labels on the arrows denote the variable bindings that resulted from the subproofs. While the set of bindings that result from proving the top-level query can be considered the conclusion of the reasoning process, it is the chain of subproofs that constitutes the narrative of the agent with respect to this conclusion. The same query cannot be proven by Agent 2, where the proof already fails at the first conjunct. Indeed, Agent 2 fails to prove the alignment between instances of the RECEIVING and SENDING frames, as its belief that a general strike is going on leads to a failure to prove that the postal services are operational, which is a precondition for the link between the two frames to be established. Note that when a conclusion cannot be proven, the narrative needs to be constructed abductively. Indeed, it consists here in finding a minimal explanation for why a conclusion does not follow from a collection of facts and rules.

4 Discussion and Conclusion

In this paper, we have introduced the Candide model as a computational architecture for modelling human-like, narrative-based language understanding. As such, we have presented an approach that radically breaks with today's mainstream natural language processing paradigm. Rather than modelling the co-occurrence of characters and words in huge amounts of textual data, our approach focusses on the logic reasoning processes that may justify different interpretations of the same linguistic observations. While this forces us to take an enormous leap back, it bears the promise of contributing a perspective that emphasises the individual and contextualised nature of linguistic communication to the fields of computational linguistics and artificial intelligence.

We have defined narratives to be chains of reasoning operations that underlie the conclusions drawn by an individual based on their belief system. This belief system is personal and dynamic in nature, as it is continuously being shaped by new linguistic and non-linguistic experiences. Narratives are thus not captured in texts as such, but need to be construed through a personal interpretation process. A narrative thereby reflects the perspective of an individual on the world, as the process of narrative construction necessarily takes one's entire belief system into account.

The construction of a narrative is a means rather than an end. While the end is to reach a conclusion, for example to answer a question, to resolve a coreference, or to make sense of a novel observation or experience, the means to reach that end is to construe a narrative that is consistent with one's belief system. In this view, the construction of a narrative is not a task in itself, but serves the purpose of solving an external task through human-interpretable reasoning processes. As narratives highly depend on external tasks and individual belief systems, they are hard to annotate in linguistic resources. Indeed, whether a narrative is justified or not only depends on whether it is consistent with the input that is observed in combination with the beliefs held by an individual. Narrative-based language understanding therefore largely coincides with the use of explainable methods for solving a variety of NLP tasks, including question answering, text summarisation and sentiment analysis, with the difference that the focus in evaluation shifts from the task accuracy to the soundness of the reasoning processes involved.

The Candide model operationalises this vision through a combination of frame-based constructional language processing and logic reasoning. As such, the belief system of an agent is represented as a collection of facts and rules that support automated reasoning through logic inference. The Fluid Construction Grammar-based language comprehension component is used to map between natural language utterances and a frame-based representation of their meaning. This semantic representation makes use of the same format as the one used to represent the agent's belief system, facilitating the straightforward integration of new beliefs into the agent's personal dynamic memory. The Prolog-based reasoning component can be leveraged to solve external tasks by proving logic formulae based on the facts and rules stored in the agent's personal dynamic memory. It is during this process of logic inference that narratives emerge as logical explanations that justify the conclusions drawn by an agent. We have illustrated our proof-of-concept implementation of the Candide model by means of a didactic example that shows how two agents who hold slightly different beliefs interpret the same linguistic observation differently, as they construe different narratives that lead to substantially different conclusions.

While this paper has laid the conceptual foundations of a novel approach to narrative-based language understanding, it has left the issue of operationalising the approach on a larger scale unaddressed. We envision an agent to start out as a blank slate, with an empty belief system and grammar. Through experience, an agent would then gradually build up linguistic and non-linguistic beliefs in a constructivist manner through the processes of intention reading and pattern finding. These processes have abundantly been attested in children (see e.g. Pine and Lieven, 1997; Tomasello, 2003) and have more recently been operationalised at scale in artificial agents through abductive reasoning processes (see e.g. Nevens et al., 2022; Doumen et al., 2023; Beuls and Van Eecke, 2023). We consider these preliminary results to be modest yet promising steps towards the moonshot of building personal, dynamic and human-interpretable models of narrative-based language understanding.

Limitations

This paper presents the conceptual foundations of a novel architecture for narrative-based language understanding, along with an illustrative proof-of-concept implementation. As such, it has been operationalised on a small scale only. Scaling up the approach to real-world applications is a highly non-trivial task that would not only require large investments but also significant innovative research efforts. Moreover, important aspects of the theoretical model have not been included in the proof-of-concept implementation, in particular when it comes to modelling the confidence of an agent with respect to its beliefs and narratives.

Acknowledgements

We are especially grateful to Remi van Trijp for his important contributions to the insightful discussions that led up to the development of the Candide model, as well as for his constructive feedback on a first version of this manuscript.

The research reported on in this paper received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements no. 951846 (MUHAI - Meaning and Understanding in Human-centric AI) and no. 101094752 (SoMe4Dem - Social Media for Democracy – understanding the causal mechanisms of digital citizenship), from the Research Foundation Flanders (FWO) through a postdoctoral grant awarded to Paul Van Eecke (grant no. 75929) and from the Flemish Government under the 'Flanders AI Research Program'.

References

Giulio Antonio Abbo and Tony Belpaeme. 2023. Users' perspectives on value awareness in social robots. In *HRI2023, the 18th ACM/IEEE International Conference on Human-Robot Interaction*, pages 1–5, New York, NY, USA. Association for Computing Machinery.

Katrien Beuls and Paul Van Eecke. 2023. Fluid Construction Grammar: State of the art and future outlook. In *Proceedings of the First International Workshop on Construction Grammars and NLP (CxGs+NLP, GURT/SyntaxFest 2023)*, pages 41–50, Washington, D.C., USA. Association for Computational Linguistics.

Katrien Beuls and Paul Van Eecke. 2024. Construction grammar and artificial intelligence. In Mirjam Fried and Kiki Nikiforidou, editors, *The Cambridge Handbook of Construction Grammar*. Cambridge University Press, Cambridge, UK.

William Croft. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press, Oxford, United Kingdom.

Luc De Raedt, Angelika Kimmig, and Hannu Toivonen. 2007. Problog: A probabilistic prolog and its application in link discovery. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*, pages 2468–2473, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL*, pages 4171–4186.

Jonas Doumen, Katrien Beuls, and Paul Van Eecke. 2023. Modelling language acquisition through syntactico-semantic pattern finding. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 1317—-1327, Dubrovnik, Croatia. Association for Computational Linguistics.

Charles J. Fillmore. 1976. Frame semantics and the nature of language. In *Annals of the New York Academy of Sciences: Conference on the origin and development of language and speech*, volume 280, pages 20–32. New York, NY, USA.

Charles J. Fillmore. 1988. The mechanisms of "construction grammar". In *Annual Meeting of the Berkeley Linguistics Society*, volume 14, pages 35–55.

Charles J. Fillmore and Collin F. Baker. 2001. Frame semantics for text understanding. In *Proceedings of WordNet and Other Lexical Resources Workshop, NAACL*, volume 6.

Mirjam Fried and Jan-Ola Östman. 2004. Construction grammar: A thumbnail sketch. In Jan-Ola Östman and Mirjam Fried, editors, *Construction grammar in a cross-language perspective*, pages 1–86. John Benjamins, Amsterdam, Netherlands.

Adele E. Goldberg. 1995. *Constructions: A construction grammar approach to argument structure*. University of Chicago Press, Chicago, IL, USA.

Paul Grice. 1967. Logic and conversation. In Paul Grice, editor, *Studies in the Way of Words*, pages 41–58. Harvard University Press, Cambridge, MA, USA.

Ivano Lauriola, Alberto Lavelli, and Fabio Aiolli. 2022. An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, 470(C):443–456.

Nieves Montes and Carles Sierra. 2022. Synthesis and properties of optimally value-aligned normative systems. *Journal of Artificial Intelligence Research*, 74:1739–1774.

Jens Nevens, Jonas Doumen, Paul Van Eecke, and Katrien Beuls. 2022. Language acquisition through intention reading and pattern finding. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 15–25, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Julian M. Pine and Elena V. M. Lieven. 1997. Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics*, 18(2):123–138.

Dan Sperber and Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Harvard University Press, Cambridge, MA, USA.

Luc Steels. 2004. Constructivist development of grounded construction grammar. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*, pages 9–16.

Luc Steels. 2020. Personal dynamic memories are necessary to deal with meaning and understanding in human-centric AI. In *Proceedings of the First International Workshop on New Foundations for Human-Centered Artificial Intelligence (NeHuAI@ECAI)*, pages 11–16.

Luc Steels. 2022. Towards meaningful human-centric AI. In Luc Steels, editor, *Foundations for meaning and understanding in human-centric AI*, pages 5–28. Venice International University, Venice, Italy.

Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*, volume 27, pages 3104–3112, Red Hook, NY, USA. Curran Associates, Inc.

Michael Tomasello. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, Harvard, MA, USA.

Paul Van Eecke. 2018. *Generalisation and specialisation operators for computational construction grammar and their application in evolutionary linguistics Research*. Ph.D. thesis, Vrije Universiteit Brussel, Brussels: VUB Press.

Paul Van Eecke and Katrien Beuls. 2017. Metalayer problem solving for computational construction grammar. In *The 2017 AAAI Spring Symposium Series*, pages 258–265, Palo Alto, CA, USA. AAAI Press.

Paul Van Eecke, Jens Nevens, and Katrien Beuls. 2022. Neural heuristics for scaling constructional language processing. *Journal of Language Modelling*, 10(2):287–314.

Remi van Trijp, Katrien Beuls, and Paul Van Eecke. 2022. The FCG editor: An innovative environment for engineering computational construction grammars. *PLOS ONE*, 17(6):e0269708.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pages 6000–6010, Red Hook, NY, USA. Curran Associates, Inc.

Voltaire. 1759. *Candide ou l'Optimisme*. Gabriel Cramer, Genève.