

On Deep Computational Formalization of Natural Language

Naveen Sundar Govindarajulu • Selmer Bringsjord • John Licato

Computer Science & Cognitive Science
Rensselaer Polytechnic Institute (RPI)
Troy NY 12180 USA
govinn@rpi.edu

Keywords: Deontic Cognitive Event Calculus, formalization of representation and reasoning, Montague semantics, Discourse Representation Theory.

1 Introduction

Current AI and NLP operate on fragments of natural language and within specific application domains. Even the most successful AI/NLP system in existence today, IBM’s Watson, is highly limited when it comes to simple language processing just beyond its ken. In order to paralyze (at least the *Jeopardy!*-winning version of) Watson, one has only to ask it questions that have never been asked before, or questions to which answers have never been recorded before. Such questions are easy to formulate; for example:

- “If I have 4 foos and 5 bars, and if foos are not the same as bars, how many foos will I have if I get 3 bazes which just happen to be foos?”; or
- “What was IBM’s Sharpe ratio in the last 60 days of trading?”¹

We contend that one of the major reasons for this lack of generality in AI/NLP systems is the absence of a wide-ranging formalization of natural language that is both fully formal and rigorously computational. We herein elaborate on the requirements that such a theory should meet, and very briefly evaluate the two most prominent projects in formal semantics: the Montagovian approach and Kamp’s DRT approach (laid out under our rubric). We then encapsulate our own approach, which is rooted in formal computational logic; this approach fares markedly better than either the Montagovian or the DRT tack.

2 Requirements for a Computational Formalization of Natural Language

Formalizing language can quickly become a philosophically troubled project, but we are our humble, in that we want to formalize language just to the extent that it allows us to do certain meaningful things computationally. The requirements that drive our efforts are spelled out below.

All language formalization approaches treat language as an isolated phenomenon separate from cognitive processes that use language. From a purely scientific perspective, this view, while convenient, is incomplete. The existing general approach in formal semantics (or even pragmatics) is that sentences (or linguistic phenomena) are considered in isolation, and various formal structures are posited for their meaning. Pragmatics tries to rectify this, but there is no convincing unified framework, such as model theory for formal semantics, for formal pragmatics.

Our requirements below stem from the observation that any account of language needs to include a full account of how and where it is used. Let us denote the set of all natural-language sentences and expressions in some language by ‘ \mathcal{L} ,’ and the set of all formal expressions that we can computationally handle by ‘ \mathcal{F} .’ Though some will be philosophically inclined to reject the idea, it is not unreasonable to hold that objects in \mathcal{F} represent meaning, or, a bit more precisely, are what sentences mean. On this foundation, any formalization of language would naturally attempt to define \mathcal{F} and give us a mapping σ such that $\forall s \in \mathcal{L} \exists m \mu(s, m)$. With this background, we can distill our requirements into the following pair.

¹ The approach presented in this paper will be used in a collaboration with IBM, and with RPI’s Jim Hendler, in order to enable subsequent versions of Watson to answer such questions on the strength of robust problem-solving.

Formalization of Extensional and Intensional Representation The formalization should specify both \mathcal{F} and some general class of μ . The class \mathcal{F} should be syntactically rich enough to handle not just extensional sentences of natural language such as “*All apples are red.*” but also challenging intensional sentences such as “*Jack believes that Jane believes that all apples are red.*” (Roughly put, extensional concepts pertain to those of the physical kind and intensional concepts to cognitive concepts.) The requirement to formalize both extensional and intensional sentences rules out simple extensional languages such as those at the heart of the Semantic Web (e.g., description logics, covered e.g. in Baader, Calvanese & McGuinness 2007), or those that are the target of simple domain-specific semantic parsing techniques (Kwiatkowski, Zettlemoyer, Goldwater & Steedman 2010). For a more extensive defense of the position that extensional languages and logics are not adequate to capture intensional concepts, please consult (Bringsjord & Govindarajulu 2012), in which we look at three different ways of formalizing knowledge within first-order logic. All three approaches fail by either introducing an inconsistency or by enabling unsound inferences to be drawn.

Formalization of Reasoning The formalization should also have scope for including all the different kinds of reasoning processes that can be carried out with the aid of natural language. Such wide-ranging reasoning of course goes beyond just simple classical deduction. In general, given a set $\Gamma_{\mathcal{L}}$ of sentences in natural language from which we can deduce/produce/infer another set of sentences $\Gamma'_{\mathcal{L}}$ via some reasoning process θ , we should have formalized θ computationally as Θ (or at least be able to). That is, we should have $\Gamma_{\mathcal{F}} \rightarrow_{\Theta} \Gamma'_{\mathcal{F}}$.

3 Current Approaches to Formalizing of Natural Language

Current approaches to formalizing language can be broadly divided into the two aforementioned camps: the Montagovian approach based on Montague’s work, and the Discourse Representation Theory (DRT) framework. DRT can be considered as an offshoot of the Montagovian approach which tries to incorporate more pragmatics into its account.

3.1 Montague’s Framework

Based on the simple ontology in our requirements, we can say that this approach tries to give a formal account of the meaning of different natural-language expressions \mathcal{F} and how they relate to the syntax of natural-language expressions μ . The main shortcoming of this approach is that it fails to give a unified account, either informally and formally, of the various kinds of cognitive processes that language takes part in. Another shortcoming of this approach is that there seems to be a general absence of proof-theoretic methods. Formal Montagovian semantics leans heavily on model-theoretic methods. While model-theoretic semantics has found success in mathematical logic, the absence of proof-theoretic methods makes it hard to build computational methods. From a purely extensional standpoint, model-theoretic semantics would seem to suffice. However, using model theory to account for cognitive concepts such as knowledge becomes more problematic, a fact long reflected in the invention of intensional logics in order to model knowledge, belief, obligation, and so on (a survey of such logicist modeling is provided in Bringsjord 2008). A good overview of the general approach and a brief history of Montague’s approach can be found in (Janssen 2012).

3.2 Discourse Representation Theory

DRT (Kamp & Reyle 1993) arose to address one of the main perceived shortfalls of Montague’s approach: assigning meaning to linguistic expressions that depend on the meaning of other linguistic expressions across sentence boundaries does not happen naturally. While DRT has been successful in modeling quite a bit of pragmatics, it lacks a unified formal framework like that of Montagovian formal semantics. The lack of this framework renders it incomplete for our purposes.

4 A More Rigorous Framework

We suggest an approach based on a formal logic with an explicit proof theory rooted in computational methods. We illustrate one incarnation of this approach via the *Deontic Cognitive Event Calculus* (\mathcal{DCEC}^*). \mathcal{DCEC}^* is a quantified

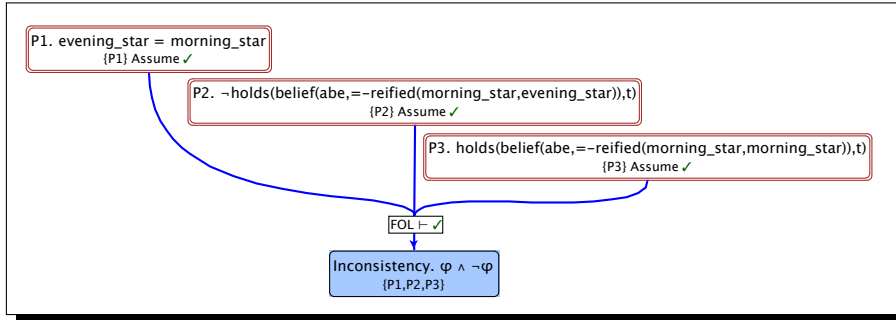
Fig. 1: Deontic Cognitive Event Calculus

Syntax	Rules of Inference
$S ::=$ Object Agent Self \square Agent ActionType Action \square Event Moment Boolean Fluent Numeric $action : Agent \times ActionType \rightarrow Action$ $initially : Fluent \rightarrow Boolean$ $holds : Fluent \times Moment \rightarrow Boolean$ $happens : Event \times Moment \rightarrow Boolean$ $clipped : Moment \times Fluent \times Moment \rightarrow Boolean$ $f ::=$ initiates : Event \times Fluent \times Moment $\rightarrow Boolean$ terminates : Event \times Fluent \times Moment $\rightarrow Boolean$ prior : Moment \times Moment $\rightarrow Boolean$ interval : Moment \times Boolean $*$: Agent \rightarrow Self payoff : Agent \times ActionType \times Moment \rightarrow Numeric $t ::= x : S \mid c : S \mid f(t_1, \dots, t_n)$ $t : Boolean \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : S. \phi \mid \exists x : S. \phi$ $P(a, t, \phi) \mid K(a, t, \phi) \mid C(t, \phi) \mid S(a, b, t, \phi) \mid S(a, t, \phi)$ $\phi ::=$ $B(a, t, \phi) \mid D(a, t, holds(f, t')) \mid I(a, t, happens(action(a^*, \alpha), t'))$ $O(a, t, \phi, happens(action(a^*, \alpha), t'))$	$\frac{}{C(t, P(a, t, \phi) \rightarrow K(a, t, \phi))} [R_1] \quad \frac{}{C(t, K(a, t, \phi) \rightarrow B(a, t, \phi))} [R_2]$ $\frac{C(t, \phi) \ t \leq t_1 \dots t \leq t_n}{K(a_1, t_1, \dots, K(a_n, t_n, \phi) \dots)} [R_3] \quad \frac{K(a, t, \phi)}{\phi} [R_4]$ $\frac{}{C(t, K(a, t_1, \phi_1 \rightarrow \phi_2) \rightarrow K(a, t_2, \phi_1) \rightarrow K(a, t_3, \phi_3))} [R_5]$ $\frac{}{C(t, B(a, t_1, \phi_1 \rightarrow \phi_2) \rightarrow B(a, t_2, \phi_1) \rightarrow B(a, t_3, \phi_3))} [R_6]$ $\frac{}{C(t, C(t_1, \phi_1 \rightarrow \phi_2) \rightarrow C(t_2, \phi_1) \rightarrow C(t_3, \phi_3))} [R_7]$ $\frac{}{C(t, \forall x. \phi \rightarrow \phi[x \mapsto i])} [R_8] \quad \frac{}{C(t, \phi_1 \leftrightarrow \phi_2 \rightarrow \neg\phi_2 \rightarrow \neg\phi_1)} [R_9]$ $\frac{}{C(t, [\phi_1 \wedge \dots \wedge \phi_n \rightarrow \phi] \rightarrow [\phi_1 \rightarrow \dots \rightarrow \phi_n \rightarrow \psi])} [R_{10}]$ $\frac{B(a, t, \phi) \ B(a, t, \phi \rightarrow \psi)}{B(a, t, \psi)} [R_{11a}] \quad \frac{B(a, t, \phi) \ B(a, t, \psi)}{B(a, t, \psi \wedge \phi)} [R_{11b}]$ $\frac{S(s, h, t, \phi)}{B(h, t, B(s, t, \phi))} [R_{12}]$ $\frac{I(a, t, happens(action(a^*, \alpha), t'))}{P(a, t, happens(action(a^*, \alpha), t))} [R_{13}]$ $\frac{B(a, t, \phi) \ B(a, t, O(a^*, t, \phi, happens(action(a^*, \alpha), t')))}{O(a, t, \phi, happens(action(a^*, \alpha), t'))} [R_{14}]$ $\frac{\phi \leftrightarrow \psi}{O(a, t, \phi, \gamma) \leftrightarrow O(a, t, \psi, \gamma)} [R_{15}]$

modal logic that builds upon on the first-order Event Calculus (EC). EC has been used quite successfully in modelling a wide range of phenomena, ranging from those that are purely physical to narratives expressed in natural-language stories (Mueller 2006).²

The EC is also a natural platform to capture natural-language semantics and especially that of tense; see (Van Lambalgen & Hamm 2005). EC has a shortcoming: it is fully extensional and hence, as explained above, has no support for capturing intensional concepts such as knowledge and belief without introducing unsoundness or inconsistencies. For example, consider the possibility of modeling changing beliefs with fluents. We can posit a “belief” fluent belief(a, f) which says whether an agent a believes another fluent f. This approach quickly leads to serious problems, as one can substitute co-referring terms into the belief term, which leads to either unsoundness or an inconsistency (see Figure 2). One can try to overcome this using more complex schemes of belief encoding in FOL, but they all seem to fail. A more detailed discussion of such schemes and how they fail can be found in the analysis in (Bringsjord & Govindarajulu 2012).

Fig. 2: An Automatic Proof of an Inconsistency Arising from Modeling Belief with Fluents



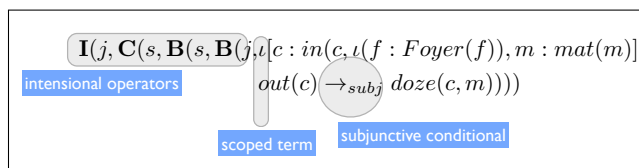
² A nice overview of EC is provided in the well-known “AIMA” textbook (Russell & Norvig 2009).

5 More Detailed Discussion

5.1 Formalization of Representation

\mathcal{DCEC}^* (deontic cognitive event calculus) is a *multi-sorted quantified modal logic* (for coverage of multi-sorted logic, see Manzano 1996), that has a formal, recursively defined syntax, and a proof calculus. The syntax of the language of \mathcal{DCEC}^* and the rules of inference for its proof calculus are shown in Figure 1. \mathcal{DCEC}^* syntax includes a system of sorts S , a signature f , a grammar for terms t , and a grammar for sentences ϕ ; these are shown on the left half of the figure. The proof calculus is based on natural deduction (Jaśkowski 1934), and includes all the introduction and elimination rules for first-order logic, as well as rules for the intensional operators; the rules are listed in the right half of the figure. The formal semantics for \mathcal{DCEC}^* is still under development; a semantic account of the wide array of cognitive and epistemic constructs found in the logic is no simple task — especially because of two self-imposed constraints: resisting fallback to the standard ammunition of possible-worlds semantics (which for reasons beyond the scope of the present paper we find manifestly implausible as a technique for formalizing the meaning of epistemic operators), and resisting the piggybacking of deontic operators on pre-established logics not expressly created and refined for the purpose of doing justice to moral reasoning in the human realm. For an introduction, see (Arkoudas & Bringsjord 2009). Figure 3 shows a simple naïve mapping of a complex English sentence, “*Jones intends to convince Smith to believe that Jones believes that were the cat, lying in the foyer now, to be let out, it would settle, dozing, on the mat.*” into \mathcal{DCEC}^* .³

Fig. 3: A Small Illustration of a Mapping



5.2 Formalization of Reasoning

\mathcal{DCEC}^* supports three different modes of reasoning. We briefly touch upon them, thereby illustrating that \mathcal{DCEC}^* (or, in fairness, any similar calculus) can be used to model a diverse array of reasoning processes.

5.3 Deductive Reasoning

\mathcal{DCEC}^* includes a deductive proof calculus Θ_d . This is shown in the right half of Figure 1. This calculus enables one to reason not only about purely extensional concepts like events, fluents and moments, which is enabled by the pure Event Calculus, but in addition enables reasoning about intensional concepts such as the beliefs, knowledge, desires, intentions, communications etc. of agents. This calculus has for instance been used to analyze the false belief task computationally. While it is possible in principle to have a monolithic proof finder to answer whether $\Gamma \vdash_{\Theta_d} \gamma$, proof search in this calculus can be broken down into cognitively plausible reusable procedures called $\lambda\mu$ -methods. For a set of methods relevant to the false-belief task, see (Arkoudas & Bringsjord 2009). This deductive calculus has also been used to solve the generalized wise-men puzzle (for any number of wise men); see (Arkoudas & Bringsjord 2004).

5.4 De Se Statements

We have proposed certain syntactic constructs in (Bringsjord & Govindarajulu 2013) by which one can:

1. distinguish *de se* statements from *de dicto* and *de re* statements; and

³ Note that this version of \mathcal{DCEC}^* has certain syntactic constructs such as scoped terms and subjunctive conditionals, under refinement now.

2. differentiate first- and third-person *de se* beliefs.

The former can be achieved by introducing an operator $*$ which can let us know if agents are referring to themselves in an irreducible manner. This construct, stemming from Castañeda’s work (e.g., Castañeda 1999), is insufficient alone, and most work in formal semantics of self-reference stops with this mechanism. This mechanism fails when we want to distinguish between third-person *de se* statements and first-person *de se* statements. What is needed is some way of recording in the reasoning process Θ a symbol for the agent doing the reasoning. This syntactic construct goes beyond what can be talked about in classical formal semantics. The table below illustrates representations for different *de se* statements in \mathcal{DCEC}^* .

Table 1: *De Se* Belief in \mathcal{DCEC}^*

NL Sentence	\mathcal{DCEC}^*	Type
Jack believes that the person named “Jack” is rich.	$\mathbf{B}(\text{jack, now, } \exists x : \text{Agent}(\text{named}(x, \text{“Jack”}) \wedge \text{rich}(x)))$	<i>De Dicto</i>
Jack believes of the person named “Jack” that he is rich.	$\exists x : \text{Agent}(\text{named}(x, \text{“Jack”}) \wedge \mathbf{B}(\text{jack, now, rich}(x)))$	<i>De Re</i>
Jack believes that he himself is rich.	$\mathbf{B}(\text{jack, now, rich}(\text{jack}*))$	Third-person <i>De Se</i>
I believe that I myself am rich.	$\vdash_1 \mathbf{B}(I, \text{now, rich}(I*))$	First-person <i>De Se</i>

5.5 Extensional Analogical Reasoning

Another focus of some of the present authors is Analogico-Deductive Reasoning (ADR) (Licato, Bringsjord & Hummel 2012), a hybrid of analogical and hypothetico-deductive reasoning that attempts to automate the generation (through analogical mapping and inference) of hypotheses, which are then subjected to deductive reasoning techniques. The use of ADR has been modeled in psychological experiments (Licato et al. 2012, Bringsjord & Licato 2012); and most recently, in Licato et al. (2013), an ADR system, Modifiable Engine for Tree-based Analogical Reasoning (META-R), was applied to the domain of theorem- and proof-discovering in cutting-edge mathematical logic. Because the surface syntax of the extensional subset of \mathcal{DCEC}^* can be represented using tree structures, it lends itself nicely to the algorithms implemented by META-R.

6 Conclusion & Next Steps

Our future work falls into two categories: work aimed at increasing the expressive power of \mathcal{DCEC}^* , and work aimed at capturing more cognitive processes.

Representation: Scoped Terms and Natural Language Connectives

The \mathcal{DCEC}^* syntax presented in this paper lacks a way to represent general noun phrases and quantifiers. All the connectives in the current syntax derive from mathematical logic; for example, the disjunctions in “*Jack fell down and Jill came tumbling after*” and “ $x \geq 3$ and $x \leq 5$ ” do not have the same meaning. This issue was briefly alluded to in Figure 3, where we use a scoped term, for representing “*the x*”, and a subjunctive conditional, for representing “*it would ...*”. Future work will incorporate such constructs and refine the proof theory to accommodate more cleanly these new elements.

Reasoning: Planning Integrated with Intensional Concepts

Another thread of research is to focus on designing a planning framework (for an agent) that integrates all the intensional operators with planning in an event calculus-based planning formalism. This work is well underway, and some demonstrations should soon be possible.

References

- Arkoudas, K. & Bringsjord, S. (2004), Metareasoning for Multi-agent Epistemic Logics, in 'Proceedings of the Fifth International Conference on Computational Logic In Multi-Agent Systems (CLIMA 2004)', Lisbon, Portugal, pp. 50–65.
URL: <http://kryten.mm.rpi.edu/arkoudas.bringsjord.clima.crc.pdf>
- Arkoudas, K. & Bringsjord, S. (2009), 'Propositional Attitudes and Causation', *International Journal of Software and Informatics* 3(1), 47–65.
URL: http://kryten.mm.rpi.edu/PRICAI-w_sequentialcalc_041709.pdf
- Baader, F., Calvanese, D. & McGuinness, D., eds (2007), *The Description Logic Handbook: Theory, Implementation (Second Edition)*, Cambridge University Press, Cambridge, UK.
- Bringsjord, S. (2008), Declarative/Logic-Based Cognitive Modeling, in R. Sun, ed., 'The Handbook of Computational Psychology', Cambridge University Press, Cambridge, UK, pp. 127–169.
URL: http://kryten.mm.rpi.edu/sb_lccm_ab-toc_031607.pdf
- Bringsjord, S. & Govindarajulu, N. (2013), Toward a Modern Geography of Minds, Machines, and Math, in V. C. Müller, ed., 'Philosophy and Theory of Artificial Intelligence', Vol. 5 of *Studies in Applied Philosophy, Epistemology and Rational Ethics*, Springer Berlin Heidelberg, pp. 151–165.
URL: <http://www.springerlink.com/content/hg712w4l23523xw5>
- Bringsjord, S. & Govindarajulu, N. S. (2012), 'Given the Web, What is Intelligence, Really?', *Metaphilosophy* 43(4), 361–532. This URL is to a preprint of the paper.
URL: http://kryten.mm.rpi.edu/SB_NSQ_Real_Intelligence_040912.pdf
- Bringsjord, S. & Licato, J. (2012), Psychometric Artificial General Intelligence: The Piaget-MacGyver Room, in P. Wang & B. Goertzel, eds, 'Theoretical Foundations of Artificial General Intelligence', Atlantis Press.
URL: http://kryten.mm.rpi.edu/Bringsjord_Licato_PAGI_071512.pdf
- Castañeda, H.-N. (1999), 'He': A Study in the Logic of Self-Consciousness, in J. G. Hart & T. Kapitan, eds, 'The Phenomeno-Logic of the I: Essays on Self-Consciousness', Indiana University Press, 601 North Morton Street, Bloomington, Indiana 4704-3797 USA.
- Janssen, T. M. V. (2012), Montague semantics, in E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', winter 2012 edn.
- Jaśkowski, S. (1934), 'On the Rules of Suppositions in Formal Logic', *Studia Logica* 1, 5–32.
- Kamp, H. & Reyle, U. (1993), *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, 1 edn, Springer.
- Kwiatkowski, T., Zettlemoyer, L., Goldwater, S. & Steedman, M. (2010), Inducing Probabilistic CCG Grammars from Logical form with Higher-order Unification, in 'Proceedings of the 2010 conference on empirical methods in natural language processing', Association for Computational Linguistics, pp. 1223–1233.
- Licato, J., Bringsjord, S. & Hummel, J. E. (2012), Exploring the Role of Analogico-Deductive Reasoning in the Balance-Beam Task, in 'Rethinking Cognitive Development: Proceedings of the 42nd Annual Meeting of the Jean Piaget Society', Toronto, Canada.
URL: <https://docs.google.com/open?id=0B1S661sacQp6NDJ0YzVXajJMwVU>
- Licato, J., Govindarajulu, N. S., Bringsjord, S., Pomeranz, M. & Gittelsohn, L. (2013), 'Analogico-deductive Generation of Gödel's First Incompleteness Theorem from the Liar Paradox', *Proceedings of the 23rd Annual International Joint Conference on Artificial Intelligence (IJCAI-13)*.
- Manzano, M. (1996), *Extensions of First Order Logic*, Cambridge University Press, Cambridge, UK.
- Mueller, E. (2006), *Commonsense Reasoning*, Morgan Kaufmann, San Francisco, CA.
- Russell, S. & Norvig, P. (2009), *Artificial Intelligence: A Modern Approach*, Prentice Hall, Upper Saddle River, NJ. Third edition.
- Van Lambalgen, M. & Hamm, F. (2005), *The Proper Treatment of Events*, Vol. 6, Blackwell Publishing.