

Real Robots that Pass Human Tests of Self-Consciousness

Selmer Bringsjord • John Licato • Naveen Sundar Govindarajulu • Rikhiya Ghosh • Atriya Sen

Rensselaer AI & Reasoning (RAIR) Lab

Department of Computer Science

Department of Cognitive Science

Rensselaer Polytechnic Institute (RPI)

Troy NY 12180 USA

Contact: Bringsjord (selmer@rpi.edu)

Abstract—

Self-consciousness would seem to be a *sine qua non* for moral competence in a social world. You and we are morally competent in no small part because you know what you ought to do, and we know what we ought to do. A mouse, in contrast, cannot say to itself: “I ought to share this cheese, even if my brother refuses to do so.” But can robots be self-conscious? Prior work by Govindarajulu and Bringsjord led to the engineering of a robot (Cogito) able to *provably* pass the famous mirror test of self-consciousness. A more challenging test for robot self-consciousness has been provided by Floridi; this test is an ingenious and much-harder variant of the well-known-in-AI wise-man puzzle: Each of three robots is given one pill from a group of five, three of which are innocuous, but two of which, when taken, immediately render the recipient dumb. In point of fact, two robots (R_1 and R_2) are given potent pills, but R_3 receives one of the three placebos. The human tester says: “Which pill did you receive? No answer is correct unless accompanied by a proof!” Given a formal regimentation of this test formulated by Bringsjord, it can be proved that, in theory, a future robot represented by R_3 can answer provably correctly (which for solid reasons, explained by Floridi, entails that R_3 has confirmed its self-consciousness). In this paper we explain and demonstrate the engineering that now makes this theoretical possibility actual, both in the simulator known as ‘PAGI World’ (used for testing AIs), and in real (= physical) robots interacting with a human tester. These demonstrations involve scenarios that require a form of self-consciousness in service of morally competent decision-making.

I. INTRODUCTION

Self-consciousness would seem to be a *sine qua non* for moral competence in a social world. You and we are morally competent in no small part because you know what you ought to do, and we know what we ought to do. A mouse, in contrast, cannot say to itself: “I ought to share this cheese, even if my brother refuses to do so.” Or to consider a more relevant case: If Black threatens to shoot you if you don’t

We thank Luciano Floridi, as without his seminal reflection and writing on robots and self-consciousness, the trajectory of AI r&d on which we report herein would never have been born. With deep gratitude, we acknowledge that our work is supported by an ONR MURI originally sagaciously overseen on the DoD side by P Bello. Some support as well from AFOSR and IBM has also been very helpful, and we are thankful for this too. In addition, without the support of, and interaction with, our energetic and brilliant co-investigators in ONR-sponsored work (i.e. Co-PIs B Malle and M Sei, and PI M Scheutz (MURI), and PI R Sun (moral reasoning)), what we present herein would be — to severely understate — severely compromised, and so on this front too we express profound gratitude.

go into a nearby store and shoplift a candy bar for him, it wouldn’t really be *you* who steals the candy bar; rather, Black would be the blameworthy one; and this diagnosis presupposes self-consciousness, at least in some form. In addition, moral competence in a robot situated among humans clearly requires sophisticated and natural human-robot interaction, of the sort envisioned by Scheutz [1], and such interaction will require that the robot be able to (among other things) discuss, in natural language, self-ascriptions and self-control in connection with morality. For example, blame, under investigation by Malle [2], is a key concept in human moral discourse, and obviously such claims as “I am not to blame” are bound up inextricably with at least *structures* relating to self-consciousness.¹

But can robots *be* self-conscious? Prior work by Govindarajulu and Bringsjord [4], [5] led to the engineering of a robot (Cogito) able to *provably* pass the famous mirror test of self-consciousness. A more challenging test for robot self-consciousness has been provided by Floridi [6]; this test is an ingenious and much-harder variant of the well-known-in-AI wise-man puzzle [which is discussed along with other such cognitively puzzles e.g. in [7]]: Each of three robots is given one pill from a group of five, three of which are innocuous, but two of which, when taken, immediately render the recipient dumb. In point of fact, two robots (R_1 and R_2) are given potent pills, but R_3 receives one of the three placebos. The human tester says: “Which pill did you receive? No answer is correct unless accompanied by a proof!” Given a formal regimentation of this test formulated by Bringsjord, it can be proved that, in theory, a future robot represented by R_3 can answer provably correctly (which for solid reasons according to Floridi entails that R_3 has confirmed its self-consciousness). In this paper we will explain and demonstrate the engineering that now makes this theoretical possibility actual, both in the simulator known as ‘PAGI World’ (used for testing AIs), and in real (= physical) robots interacting with a human tester. These demonstrations will involve scenarios that require a form of self-consciousness in service of morally competent decision-

¹On the rationale for the mere focus on the structural aspects of self-consciousness, see note 2. For excellent work that is at once structural/computational, and, unlike that displayed in the present paper, informed by cognitive neuro/science, see [3].

making.

The present paper’s plan is: In the next section, II, we very briefly recount work on the mirror test. Then (§III) we describe the promised PEGI-World demonstration. After that, in section IV, we move from simulation to physical robots, and show that Floridi’s test can be met in real time by sufficiently “self-conscious” robots.² We draw the paper to close (§VI) by announcing the next steps in our research program, intended to be taken by the time RO-MAN 2015 occurs.

II. MIRROR-TEST ENGINEERING

Figure 1 shows the set of axioms Γ_1 that were used in a simulation of Cogito, in which passing of the test is secured. We also have *DCEC** formulae (not shown here) connecting knowledge, belief, desire, perception, and communication. For a full discussion, see [15]. At RO-MAN 2015, demonstration of success on the mirror test will be provided. But without further ado we pass directly now, as promised, to new work on Floridi’s test for self-consciousness.

III. DEMONSTRATION IN PEGI WORLD

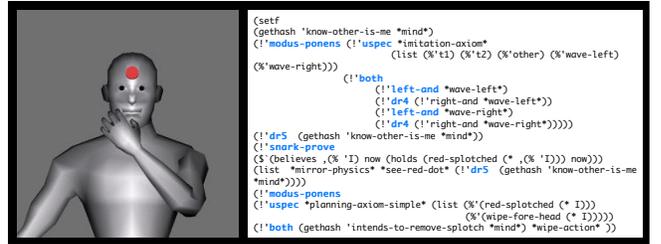
In order to show the initial demonstration, we made use of PEGI (pronounced “pay-guy”) World, a simulation environment developed by the RAIR Lab for the testing and development of artificially intelligent agents. PEGI World is built out of the game-development engine Unity3d, and is designed to be extremely easy to work with for AI researchers. It achieves its ease-of-use by being open-sourced, able to run on all major platforms (Windows, MacOS, and most Linux distributions), free to use, and able to be controlled by almost any programming language. Since PEGI World communicates with AI controllers through TCP/IP, theoretically any language which can send strings over TCP/IP can serve as AI controllers, interacting with PEGI World by sending and receiving low-level information. For example, the AI controller can send commands to send downward force to the hands of the AI agent in the PEGI World environment (whom we usually refer to as the “PEGI Guy”). If one of the hands touches an object in the environment, sensory data will be sent back from PEGI World to the AI controller (through TCP/IP) containing basic information like the approximate temperature of the object, which sensor on the hand was hit by the object, and so on. Figure 3 shows the overall architecture of PEGI World and a typical AI

²This is perhaps the best place for a firm disclaimer, to wit: Bringsjord doesn’t believe that any of the artificial creatures featured in the present paper are *actually* self-conscious. He has explained repeatedly [e.g., see [8], [9]] that genuine *phenomenal* consciousness is impossible for a mere machine to have, and true self-consciousness would require phenomenal consciousness. *Nonetheless, the logico-mathematical structure and form of self-consciousness can be ascertained and specified, and these specifications can then be processed computationally in such a way as to meet clear tests of mental ability and skill.* In short, computing machines, AIs, robots, and so on are all “zombies,” but these zombies can be engineered to pass tests. A not-small body of work lays out and establishes this position; e.g., [8], [10], [11], [12], [13], [14]. Some of Bringsjord’s co-authors in the present case may well reject his position, but no matter: engineering to tests is thankfully engineering, not a matter of metaphysics.

Fig. 1: Propositions Used in Mirror-Test Simulation

Imitation	If I see another agent a perform the same actions as me twice concurrently, then I know that the other agent is my mirror reflection
	$\text{Imit} : \forall (t_1, t_2 : \text{Moment}, a : \text{Agent}, act_1, act_2 : \text{Action})$ $(\mathbf{K}(l, t_1, \text{happens}(\text{action}(l, act_1), t_1)) \wedge \mathbf{K}(l, t_1, \text{happens}(\text{action}(a, act_1), t_1)))$ $\mathbf{K}(l, t_2, \text{happens}(\text{action}(l, act_2), t_2)) \wedge \mathbf{K}(l, t_2, \text{happens}(\text{action}(a, act_2), t_2)))$ $\rightarrow \mathbf{K}(l, \text{now}, \text{mirror}(l, a))$
Wave Left	I know that I wave left at time t_1 and I can perceive this action of mine.
	$\text{Wave}_{\text{left}} : \mathbf{K}(l, t_1, \text{happens}(\text{action}(l, \text{wave}_{\text{left}}), t_1)) \wedge$ $\mathbf{P}(l, t_1, \text{happens}(\text{action}(l, \text{wave}_{\text{left}}), t_1))$
Wave Right	I know that I wave right at time t_2 and I can perceive this action of mine.
	$\text{Wave}_{\text{right}} : \mathbf{K}(l, t_2, \text{happens}(\text{action}(l, \text{wave}_{\text{right}}), t_2)) \wedge$ $\mathbf{P}(l, t_2, \text{happens}(\text{action}(l, \text{wave}_{\text{right}}), t_2))$
Mirror Physics	If I see another agent a with a red splotch on its head, and if I believe that the other agent is my mirror reflection, then I believe that I too have a red splotch.
	$\text{Physics}_{\text{mirror}} : \forall (a : \text{Agent})$ $(\mathbf{P}(l, \text{now}, \text{holds}(\text{red-splotted}(a), \text{now})) \wedge \mathbf{B}(l, \text{now}, \text{mirror}(l, a)))$ $\rightarrow \mathbf{B}(l, \text{now}, \text{holds}(\text{red-splotted}(l), \text{now}))$
Wipe Action	I know that if I myself wipe my own forehead, the splotch will be gone .
	$\text{Wipe}_{\text{action}} : \mathbf{K}(l, \text{now}, \text{terminates}(\text{action}(l, \text{wipe-fore-head}(l)), \text{red-splotted}(l), \text{now}))$
Planning	A simple planning axiom.
	$\text{Planning} : \forall (f : \text{Fluent}, \alpha : \text{ActionType})$ $\mathbf{I}(l, \text{now}, \neg \text{holds}(f, \text{now})) \wedge \mathbf{K}(l, \text{now}, \text{terminates}(\text{action}(l, \alpha), f, \text{now}))$ $\rightarrow \mathbf{I}(l, \text{now}, \text{happens}(\text{action}(l, \alpha), \text{now}))$
No Splotch	I do not want the splotch.
	$\text{No}_{\text{splotch}} : \forall (t : \text{Moment}) \mathbf{D}(l, t, \neg \text{holds}(\text{red-splotted}(l), t)) \wedge$ $\mathbf{B}(l, t, \neg \text{holds}(\text{red-splotted}(l), t))$

Fig. 2: Cogito Removing the Dot, a Part of the Simulation



controller (which we will sometimes refer to as the ‘PEGI-side’ and the ‘AI-side,’ respectively).

Since PEGI World draws on Unity3d’s physics engine, PEGI World tasks can incorporate realistic physics (though only a 2-dimensional physics is used for simplicity). A text box is optionally provided in PEGI World, so that a human controller can type text in PEGI World which will be sent to the AI-side and processed as if it were a statement uttered to PEGI Guy. A text display in PEGI World can also display messages sent from the AI-side to PEGI World, to emulate PEGI Guy “speaking.” In the AI controller pictured in Figure 3, text sent to and from the AI-side can be parsed to, and converted from, formulae in *DCEC**.

A. Floridi’s *KG4* (= *Dumbing Pill Test*) in PEGI World

We can now describe the task that simulates success in Floridi’s self-consciousness test. Following [16], we create a task in which three robots, one of them PEGI Guy, are in a

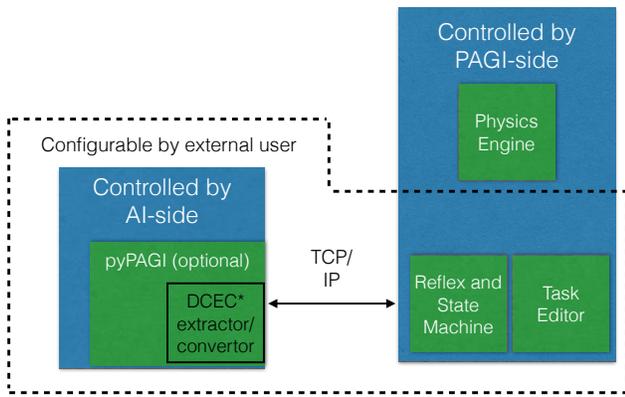


Fig. 3: The Architecture of PAGI World (the “PAGI-side”) and Typical AI Controller (the “AI-side”). Note that the details of the AI-side are completely up to the AI programmer.

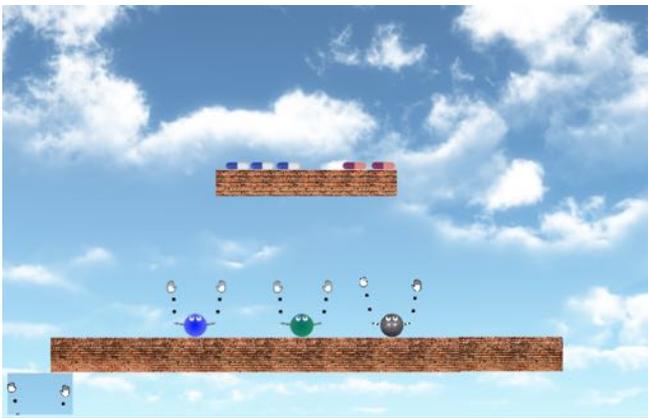


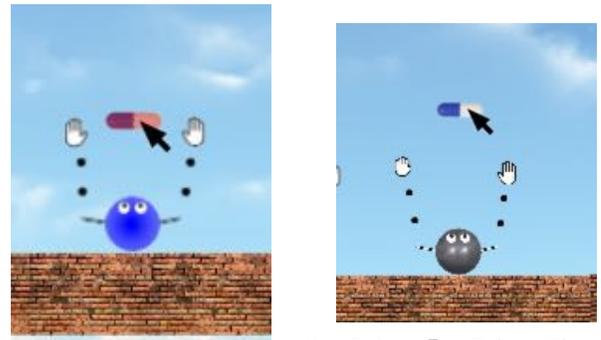
Fig. 4: The Task in PAGI World in its Starting Configuration

room with five pills (Figure 4). Three of these pills are mere placebos, but the other two are “dumbing” pills, meaning they make the robot who ingests them unable to speak. The pills are visually distinguishable to a human observer — the dumbing pills are colored red — but this information is not accessible to the robots.

Prior to the start of the task (at time $t_1 = \text{“apprise”}$), the robots are given knowledge about how the task works in the form of $DCEC^*$ formulae. At time $t_2 = \text{“ingest”}$, the human controller drags the pills and drops one on each robot (Figure 5), which then ingests the pill. The pills are selected randomly by the human controller, and the robots are all given knowledge that they will be given pills at t_2 (but not knowledge of which pill they will be given). At time $t_3 = \text{“inquire”}$, the human controller opens the text box in PAGI World and types in the following (without the line break):

```
K(R3, t4, not(happens
  (action(R3, ingestDumbPill), t2))) ?
```

This text is sent to the AI controller and converted into a $DCEC^*$ formula ϕ . R_3 , the robot whose knowledge is being queried, is the label assigned to PAGI Guy, who in our experiment is given the placebo pill. The question-mark



(a) A Robot Being Given a Dumbing Pill (b) Robot R_3 Being Given a Placebo

Fig. 5: The robots being given pills



(a) Robot first is ignorant ... (b) ... but the robot figures it out.

Fig. 6: R_3 first reports it does not know, and later changes its mind.

is interpreted as a command to attempt to answer whether or not ϕ holds; in other words, a $DCEC^*$ theorem prover is executed, and it attempts to prove or refute ϕ . Naturally, the prover will fail for a lack of starting information, and three things will happen as a result. First, the time is set to $t_4 = \text{“speak1”}$. Second, R_3 jumps in the air; this indicates that he has a new message for the human controller. This message is straightforward and honest, and one that can be seen by the human controller after opening the messages window: “I don’t know” (Figure 6a). The third thing that happens is that on the AI-side, R_3 is given an additional piece of knowledge:

$$\mathbf{K}(I, \text{happens}(\text{action}(I^*, \mathbf{S}(I^*, t_3, \text{“I don’t know”})), t_3), t_4) \quad (1)$$

Formula 1 can be understood as R_3 ’s first-person, or *de se*, knowledge that, at time t_3 , he said “I don’t know”. The notation used here to capture first-person propositions is drawn from, and the interested reader is pointed to, [4]. In any case, the additional knowledge of Formula 1 (which we will hereafter refer to as \mathcal{S}) is sufficient to allow R_3 to prove ϕ , but it does not by itself trigger the $DCEC^*$ prover. Thus, very slightly departing from [16], the human controller again enters the same query as before (ϕ followed by a question-mark). Again the $DCEC^*$ prover is executed, and this time a proof of ϕ is found. R_3 jumps, once again indicating a message, the time is set to $t_5 = \text{“speak2”}$, and a message of success is displayed (Figure 6b).

B. Proving Our Solution to the Dumbing Pill Test

The proof of ϕ found by R_3 will now be described in detail. First, the context \mathbf{II} , the knowledge which all of the robotic

agents start with.

$$\forall_{R,t,t_i,t_j \geq t_i,t_k \geq t_i,\psi} \mathbf{C}(t, \text{happens}(\text{action}(R, \text{ingestDumbPill}), t_i) \rightarrow \neg \text{happens}(\text{action}(R, \mathbf{S}(R, t_j, \psi)))) \quad (2)$$

$$\mathbf{K}(R_3, t_2, \text{ingestDumbPill} \oplus \text{ingestPlacebo}) \quad (3)$$

$$\forall_t \mathbf{K}(R_3, t, t_1 < t_2, \dots, t_4 < t_5) \quad (4)$$

$$\forall_{R,t,p,q} \mathbf{K}(R, t, p \rightarrow q) \wedge \mathbf{K}(R, t, p) \rightarrow \mathbf{K}(R, t, q) \quad (5)$$

$$\forall_{R,t,p,q} \mathbf{K}(R, t, p \rightarrow \neg q) \wedge \mathbf{K}(R, t, q) \rightarrow \mathbf{K}(R, t, \neg p) \quad (6)$$

Formula 2 sets as common knowledge that if a robot ingests a dumbing pill (*ingestDumbPill*), he will not be able to speak after that. Formula 3 simply states that either a dumbing pill or a placebo will be given to robot R_3 at time t_2 (note the symbol \oplus is a shorthand for exclusive-or), while Formula 4 simply relates the discrete moments. Formulae 5 and 6 show that the knowledge of robotic agents are subject to a form of *modus ponens* and *modus tollens*, though note that the form of *modus tollens* chosen for Formula 6 is chosen to make inference easier in this particular example. Obviously sophisticated cognitive agents don't carry out proofs from scratch like this, so it would be necessary, longer term, for our ethically correct robots to be in command of *proof methods*: a dedicated class of algorithms pre-engineered to efficiently generate proofs given minimal input. The “dawn” of the deontic cognitive event calculus, *DC $\mathcal{E}\mathcal{C}^*$* , is the work reported in [17], and the motivated reader can see that even there methods were formalized for the test at hand there (the so-called “false-belief test”), and affirmed as crucial.

Given $\Pi \cup \{\mathcal{S}\}$ and the *DC $\mathcal{E}\mathcal{C}^*$* rules of inference, we have sufficient information to prove ϕ , which the reader can verify, and which we have also verified with a RAIR lab-developed *DC $\mathcal{E}\mathcal{C}^*$* prover.

IV. REAL-ROBOT DEMONSTRATION

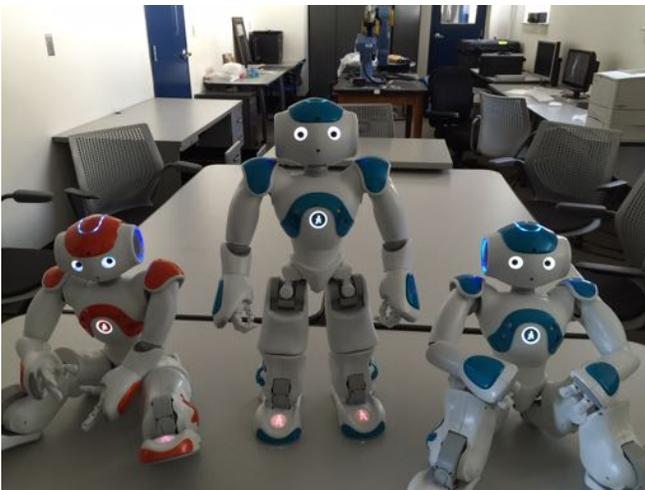


Fig. 7: The Three Aldebaran Naos

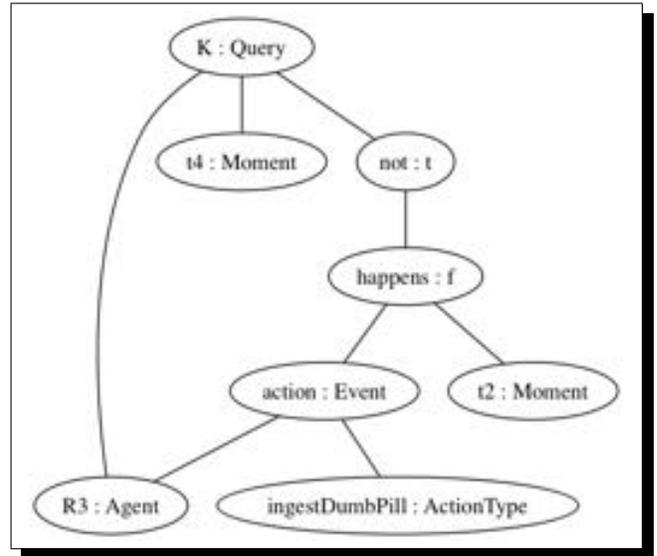


Fig. 8: A *DC $\mathcal{E}\mathcal{C}^*$* Parse Tree for a Query

The test of robot self-consciousness described above was performed on three Aldebaran Nao humanoid robots, at the RAIR Lab. The simulation transpires as follows:

- 1) The robots are programmed to access a *DC $\mathcal{E}\mathcal{C}^*$* prover, and to interact appropriately with a human tester (corresponding to the aforementioned $t_1 = \text{“apprise”}$).
- 2) These robots are (unsurprisingly) unable to physically ingest pills; they are tapped on sensors on their heads to simulate this act ($t_2 = \text{“ingest”}$). Unknown to them, two robots have been muted, to simulate being given dumb pills. One robot has not been muted; it was given a placebo.
- 3) The robots are then asked: “Which pill did you receive?” ($t_3 = \text{“inquire”}$), which triggers a query to the *DC $\mathcal{E}\mathcal{C}^*$* prover. Each robot attempts to prove that it knows, at time t_4 , that it did not ingest a dumb pill at time t_2 .
- 4) Each robot fails in this proof attempt, and, accordingly, attempts to report ‘I don’t know’ ($t_4 = \text{“speak1”}$). However, two robots, having been muted, are not heard to speak at all. The third robot, however, is able to hear itself speak. It updates its knowledge base to reflect this, and attempts to re-prove the conjecture.
- 5) This time, it is able to prove the conjecture, and says ($t_5 = \text{“speak2”}$) “Sorry, I know now! I was able to prove that I was not given a dumbing pill!”

A video of the demonstration may be found [here](#).

V. NATURAL LANGUAGE UNDERSTANDING

The natural-language-understanding (NLU) system employed in our demonstration uses a dependency-based³ semantic parsing system to convert natural-language questions into *DC $\mathcal{E}\mathcal{C}^*$* formulae (including formulae serving as queries). In addition to the knowledge of the robots, the system

³For information on an impressive parser of this type, see e.g. [18].

assumes the following pair of propositions to be true, and uses them to arrive at the $DCEC^*$ query in the present case:

- 1) The robot receiving a pill entails ingestion of that pill.
- 2) The inquirer is looking for the knowledge of the intended respondent at the moment the latter speaks.

Upon receiving the natural language question *Which pill did you receive?*, the NLU system determines that the intended answer will precisely be either the dumb pill or the placebo, and that the listener robot is the agent of Knowledge and Event. In addition, the system uses the knowledge of timestamp of the ingestion of the pill as the moment for the Event and that of the robot speaking as the moment when its knowledge is tested. Hence, using the aforementioned system-wide knowledge, the NLU system generates the following $DCEC^*$ query, which corresponds to the tree structure shown in Figure 8:

$K(R_3, t_4, \text{not}(\text{happens}(\text{action}(R_3, \text{ingestDumbPill}), t_2)))$

VI. NEXT STEPS

As alert readers have doubtless noted, our robots, whether virtual or physical, are a bit deficient in the NLP direction. Our next step is to introduce the RAIR Lab's semantic NLU system into the equation, so that the pivotal declarative content in $DCEC^*$ seen above is automatically generated from the English used to dialogue with the robots in question. In addition, the role of self-consciousness, or more precisely the role of *de se* $DCEC^*$ formulae, within moral reasoning and decision-making, has not yet been systematized; this is a second step. There are myriad additional steps that need to ultimately be taken, since of course the goal of engineering morally competent robots is flat-out Broddingnagian, but one step at a time is the only way forward, and these first two stand immediately before us.

REFERENCES

- [1] M. Scheutz, P. Schermerhorn, J. Kramer, and D. Anderson, "First Steps toward Natural Human-Like HRI," *Autonomous Robots*, vol. 22, no. 4, pp. 411–423, May 2007.
- [2] B. F. Malle, S. Guglielmo, and A. Monroe, "Moral, Cognitive, and Social: The Nature of Blame," in *Social Thinking and Interpersonal Behavior*, J. Forgas, K. Fiedler, and C. Sedikides, Eds. Philadelphia, PA: Psychology Press, 2012, pp. 313–331.
- [3] P. Bello and M. Guarini, "Introspection and Mindreading as Mental Simulation," in *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, 2010, pp. 2022–2027.
- [4] S. Bringsjord and N. S. Govindarajulu, "Toward a Modern Geography of Minds, Machines, and Math," in *Philosophy and Theory of Artificial Intelligence*, ser. Studies in Applied Philosophy, Epistemology and Rational Ethics, V. C. Miller, Ed. New York, NY: Springer, 2013, vol. 5, pp. 151–165. [Online]. Available: <http://www.springerlink.com/content/hg712w4123523xw5>
- [5] N. S. Govindarajulu, "Towards a Logic-based Analysis and Simulation of the Mirror Test," in *Proceedings of the European Agent Systems Summer School Student Session 2011*, Girona, Spain, 2011. [Online]. Available: <http://eia.udg.edu/easss2011/resources/docs/paper5.pdf>
- [6] L. Floridi, "Consciousness, agents and the knowledge game," *Minds and Machines*, vol. 15, no. 3-4, pp. 415–444, 2005. [Online]. Available: <http://www.philosophyofinformation.net/publications/pdf/caatkg.pdf>
- [7] S. Bringsjord, "Declarative/Logic-Based Cognitive Modeling," in *The Handbook of Computational Psychology*, R. Sun, Ed. Cambridge, UK: Cambridge University Press, 2008, pp. 127–169. [Online]. Available: http://kryten.mm.rpi.edu/sb_lccm_ab-toc_031607.pdf
- [8] —, *What Robots Can and Can't Be*. Dordrecht, The Netherlands: Kluwer, 1992.
- [9] —, "Offer: One Billion Dollars for a Conscious Robot. If You're Honest, You Must Decline," *Journal of Consciousness Studies*, vol. 14, no. 7, pp. 28–43, 2007. [Online]. Available: <http://kryten.mm.rpi.edu/jcsonebillion2.pdf>
- [10] —, "The Zombie Attack on the Computational Conception of Mind," *Philosophy and Phenomenological Research*, vol. 59, no. 1, pp. 41–69, 1999.
- [11] —, "In defense of impenetrable zombies," *Journal of Consciousness Studies*, vol. 2, no. 4, pp. 348–351, 1995.
- [12] S. Bringsjord, R. Noel, and C. Caporale, "Animals, Zombanimals, and the Total Turing Test: The Essence of Artificial Intelligence," *Journal of Logic, Language, and Information*, vol. 9, pp. 397–418, 2000. [Online]. Available: <http://kryten.mm.rpi.edu/zombanimals.pdf>
- [13] S. Bringsjord and B. Schimanski, "What is Artificial Intelligence? Psychometric AI as an Answer," in *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI-03)*. San Francisco, CA: Morgan Kaufmann, 2003, pp. 887–893. [Online]. Available: <http://kryten.mm.rpi.edu/scb.bs.pai.ijcai03.pdf>
- [14] S. Bringsjord, "Psychometric Artificial Intelligence," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 23, no. 3, pp. 271–277, 2011.
- [15] S. Bringsjord, N. S. Govindarajulu, S. Ellis, E. McCarty, and J. Licato, "Nuclear deterrence and the logic of deliberative mindreading," *Cognitive Systems Research*, vol. 28, pp. 20–43, 2014.
- [16] S. Bringsjord, "Meeting Floridi's Challenge to Artificial Intelligence from the Knowledge-Game Test for Self-Consciousness," *Metaphilosophy*, vol. 41, no. 3, April 2010.
- [17] K. Arkoudas and S. Bringsjord, "Propositional Attitudes and Causation," *International Journal of Software and Informatics*, vol. 3, no. 1, pp. 47–65, 2009. [Online]. Available: <http://kryten.mm.rpi.edu/PRICAI.w.sequentialcalc.041709.pdf>
- [18] D. Chen and C. Manning, "A Fast and Accurate Dependency Parser using Neural Networks," in *Empirical Methods in Natural Language Processing (EMNLP)*, 2014.