

# Learning *Ex Nihilo*

Selmer Bringsjord<sup>1</sup>, Naveen Sundar Govindarajulu<sup>1</sup>,  
John Licato<sup>2</sup>, and Michael Giancola<sup>1</sup>

<sup>1</sup> Rensselaer AI & Reasoning (RAIR) Lab,  
Rensselaer Polytechnic Institute, Troy NY 12180 USA

<sup>2</sup> Advancing Machine and Human Reasoning (AMHR) Lab,  
University of South Florida, Tampa FL 33620 USA

{ selmer.bringsjord, naveen.sundar.g, mike.j.giancola, john.licato } @gmail.com

## Abstract

This paper introduces, philosophically and to a degree formally, the novel concept of learning *ex nihilo*, intended (obviously) to be analogous to the concept of creation *ex nihilo*. Learning *ex nihilo* is an agent’s learning “from nothing”, by the suitable employment of inference schemata for deductive and inductive reasoning. This reasoning must be in machine-verifiable accord with a formal proof/argument theory in a *cognitive calculus* (i.e., here, roughly, an intensional higher-order multi-operator quantified logic), and this reasoning is applied to percepts received by the agent, in the context of both some prior knowledge, and some prior and current interests. Learning *ex nihilo* is a challenge to contemporary forms of ML, indeed a severe one, but the challenge is here offered in the spirit of seeking to stimulate attempts, on the part of non-logicist ML researchers and engineers, to collaborate with those in possession of learning-*ex nihilo* frameworks, and eventually attempts to integrate directly with such frameworks at the implementation level. Such integration will require, among other things, the symbiotic interoperation of state-of-the-art automated reasoners and high-expressivity planners, with statistical/connectionist ML technology.

## 1 Introduction

This paper introduces, philosophically and to a degree logico-mathematically, the novel concept of learning *ex nihilo*, intended (obviously) to be analogous to the concept of creation *ex nihilo*.<sup>1</sup> Learning *ex nihilo* is an agent’s learning “from nothing,” by the suitable employment of inference schemata for deductive and inductive<sup>2</sup> (e.g., analogical, enumerative-inductive, abductive, etc.) reasoning. This reasoning must be in machine-verifiable accord with a formal proof/argument theory in a *cognitive calculus*, and this reasoning is applied to percepts received by the agent, in the context of both some prior knowledge, and some prior and current interests. While cognitive calculi are characterized later at more length in Appendices A & B, at this point we can say that, provisionally, cognitive calculi include inferential components of intensional higher-order multi-operator quantified logics, in which expressivity far outstrips off-the-shelf modal logics and possible-worlds semantics is left aside in favor of fully proof-theoretic semantics. A number of such calculi have been introduced as bases for AI that is unrelated to

---

<sup>1</sup>No such assumption as that creation *ex nihilo* is real or even formally respectable is made or needed in the present paper. The concept of creation *ex nihilo* is simply for us an intellectual inspiration — but as a matter of fact, the literature on it in analytic philosophy does provide some surprisingly rigorous accounts. In the present draft of the present paper, we don’t seek to mine these accounts.

<sup>2</sup>Not to be confused with inductive logic programming (about which more will be said later), or inductive deductive techniques and schemas (e.g. mathematical induction, the induction schema in Peano Arithmetic, etc.). As we explain later, learning *ex nihilo* is in part powered by non-deductive inference schemata seen in inductive logic. An introductory overview of inductive logic is provided in [39].

learning; e.g. see [32], where the application area is AI ethics. The very first cognitive calculus, a purely deductive one, replete with a corresponding implementation in ML, was introduced in [2].

Learning *ex nihilo* is a challenge to contemporary forms of ML, indeed a severe one, because contemporary ML, for example “deep learning,” which is based on (thoroughly non-declarative) artificial neural networks (ANNs), fails to provide any proofs or formal arguments that certify that learning of something has taken place. (As explained in [12], what is called there “real learning” ( $\mathcal{RL}$ ) requires that the agent that has learned  $X$  via  $\mathcal{RL}$  must be able to answer queries about  $X$ , and prove or at least justify by rigorous argumentation that the answers to these queries in fact hold.) However, the challenge to ML issued by the presentation herein of learning *ex nihilo* is offered in the spirit of seeking to stimulate attempts, on the part of non-logicist ML researchers and engineers, to collaborate with those in possession of learning *ex nihilo* frameworks, and to integrate directly with such frameworks at the implementation level [such spirit is in line with 9]. Such integration will require, among other things, the symbiotic use of state-of-the-art automated reasoners (such as ShadowReasoner [33], the particular reasoner that for us powers learning *ex nihilo*) with statistical/connectionist ML technology.

The plan for the paper is as follows: To start, §2 offers a rather mundane parable in which an ordinary human Robert learns *ex nihilo*. Next, we explain that learning *ex nihilo* is ubiquitous in human life (§3). In §4 we point out that learning *ex nihilo* produces in the agent that employs it knowledge, specifically *propositional* or *knowing-that* knowledge (versus *knowing-how* which is what statistical/connectionist ML can be viewed as producing in ANN-based artificial agents). Given that the new form of learning we introduce herein is intended to yield propositional knowledge, it’s incumbent upon us to defend this conception of knowledge against the famous and seemingly fatal Gettier counter-examples; we do so in §5. Next, we quickly summarize the logico-mathematics of learning *ex nihilo* (§6). We then briefly explain how we can bring the formal elements of learning *ex nihilo* to concrete life with automated-reasoning technology (§7). The following section, 8, is devoted to anticipating and rebutting a series of objections. The paper proper ends with the obligatory discussion of next steps, and then follows two appendices (A, in which a fuller account of what a cognitive calculus is is provided; and B, in which a particular cognitive calculus is presented in synoptic form).

## 2 A Real-Life Parable

Let’s now consider a scenario of a type that routinely happens in real life.<sup>3</sup> Consider Robert, a person of the human variety<sup>4</sup> who has just arrived for a black-tie dinner party at a massive and manicured stone mansion to which he has never been, hosted by a couple (who have told him they own the home) he has never met, and is soon seated at an elegant table, every seat of which is occupied by a diner Robert is now meeting for the very first time.<sup>5</sup> A thin, tall,

<sup>3</sup>Both this parable, and the next (§3), are based directly on the actual experiences of one of the authors of the present paper, for what it’s worth.

<sup>4</sup>The concept of personhood is a *mental* one that guides us. Personhood rides well above such merely biological categories as *Homo sapiens sapiens*, and for us is the target of and yardstick for measuring success in AI (and, for that matter, computational cognitive science). In this regard, we are not alone; note e.g. that in a dash of insight and eloquence, Charniak & McDermott [17] declare that “the ultimate goal of AI is to build a person” (p. 7) — from which we can deduce that personhood is in no way inseparably bound up with the particular carbon-based human case. The logico-computational modeling, in computational cognitive science, of reasoning at the level of persons, crucial for learning *ex nihilo*, along with a synoptic account of personhood itself, is provided in [13].

<sup>5</sup>Robert does know himself (and in fact self-knowledge is essential for learning *ex nihilo*), but, again, he doesn’t know any of the other diners.

crystal glass of his (arrayed among three others, each of a different shape, that are his as well) is gradually filled with liquid poured from a bottle that he notices carries the words ‘Louis Roederer,’ which have no particular meaning for him; as the pour unfolds, Robert notices tiny bubbles in the liquid in his glass, and the white-tuxedoed server says, “Your apertif, sir.” At this point, Robert is in position to learn an infinite number of propositions *ex nihilo*. He has received no large dataset, and the only direct communication with him has been composed solely of rather empty pleasantries and the one perfunctory declaration that he has been served an apertif. Yet as Robert takes his first (stunning) sip of what he has already learned *ex nihilo* is expensive Champagne,<sup>6</sup> and as he glances at the other five guests seated at the table, he is poised to learn *ex nihilo* without bound. How much new knowledge he acquires is simply a function of how much time and energy he is willing and able to devote to the form of learning in question. As his water glass is filled, first with a wafer-thin slice of lemon dropped in deftly with silver tongs, and then with the water itself, he gets started:

For example, Robert now knows that his hosts find acceptable his belief that they are quite wealthy. [They may not in fact *be* wealthy (for any number of reasons), but they know that Robert’s perceiving what they have enabled him to perceive will lead to a belief on his part that they are wealthy, and Robert knows that they know this.] Robert now also knows that the menu, on the wine side, includes at least two additional options, since otherwise his array of empty glasses wouldn’t number three, one of which he knows is for water. . . .

### 3 Learning *Ex Nihilo* is Ubiquitous: A Second Parable

Of course, where there is one parable, countless others can be found: Herman isn’t the black-tie kind of person. Given a choice between the dinner Robert had versus one under the stars in the wilderness, prepared on an open fire, Herman will take the latter, every time. Having just finished such a meal, Herman is now supine beside the fire, alone, many miles from civilization in the vast Adirondack Park, on a very chilly but crystal-clear summer evening. Looking up at the heavens, he gets to thinking — and specifically gets to learning (*ex nihilo*, of course). Herman knows next to nothing about astronomy. He sees a light traveling smoothly, steadily, and quickly across his field of vision. Herman asks himself: What is this thing? He hears no associated sound. He isn’t inclined to take seriously that this is an alien spacecraft — unless what he is seeing is a total anomaly. “Is it?” he asks himself. He waits and looks. There is another, but this one is a much smaller light. This seems routine, but if so, and if this is a UFO, the papers would routinely be filled with UFO “sightings,” and so on; but they aren’t. So, no, not a UFO. The larger light, he next notes, is traveling too quickly to be a jet at high altitude, and in the dark like this, no light pollution whatsoever, jets at high altitude are hard to see. Herman notes that the object, as it moves, blocks out his view of stars behind it. Ah! Herman now knows that he has just seen two satellites in orbit, and with this knowledge now obtained, before the night is out he will see two more. But why the amazing apparent size of the satellite he first saw? Herman surmises that it must be the international space station, which, it stands to reason, is quite large relative to unmanned satellites for communication, etc. Herman never knew that you could just lay down under these conditions and see satellites; he also never knew that there are a lot of satellites up there, going around Earth, but he reasons that since his point of view is from one particular spot on the surface of Earth, it is likely to be representative

---

<sup>6</sup>Robert perceives that his beverage is sparkling wine, that it’s likely quite dear, and knows enough about both the main countries that produce such a thing (viz. USA, Spain, France, and Italy), and linguistics, to reason to the belief that his beverage’s origin is French, and hence that it’s Champagne.

of any number of other locations, and hence there must be many of these satellites in orbit. Herman has now come to learn many things, and the night is still young.

Robert and Herman are at the tip of an infinite iceberg of cases in which agents learn *ex nihilo*, both in rather mundane fashion at fancy dinners and campfire dinners, and in the more exotic cases seen in fiction (witness e.g. the eerie ability of Sherlock Holmes to quickly learn *ex nihilo* myriad things when meeting people for the first time, a recurring and prominent phenomena in PBS’ [Sherlock](#)<sup>7</sup>). Moreover, it turns out that learning *ex nihilo* is not only ubiquitous, but is also — upon empirical investigation — a very powerful way to learn in the academic sphere, where forcing oneself to be interested enough to ask oneself questions, and then attempt to reason to their answers, can pay demonstrable academic dividends [18, 64].

## 4 Learning *Ex Nihilo* Produces Knowledge

Please note that while it may seem to the reader that learning *ex nihilo* is rather relaxed, free-wheeling, and epistemically risky, the fact is that we have *very* high standards for declaring some process, whether implemented in a person or a machine, to be learning. Put with brutal simplicity here, genuine learning of  $\phi$  by an agent, for us, must result in the acquisition of knowledge by the agent, and knowledge in turn consists in the holding of three conditions, to wit: (1) the agent must believe that  $\phi$  holds; (2) must have cogent, expressible, surveyable justification for this belief; and (3)  $\phi$  must in fact hold.<sup>8</sup> This trio constitute the doctrine that knowledge consists of justified true belief; we shall abbreviate this doctrine as ‘ $\mathbf{k=jtb}$ ’. By  $\mathbf{k=jtb}$ , which reaches back at least to Plato, most of what is called “learning” in AI today (e.g. the aforementioned so-called “deep learning”) is nothing of the sort.<sup>9</sup> But in the case of our Robert and Herman, conditions (1)–(3) obtain with respect to all the new knowledge we have ascribed to them, and this would clearly continue to be true even if we added *ad infinitum* propositions that they can come to learn *ex nihilo*, stationary physically, but moving mentally. Unfortunately,  $\mathbf{k=jtb}$  has come under considerable fire, something we now proceed to deal with.

## 5 Learning *Ex Nihilo* Includes Novel Solution to the Vexing Gettier Problem

Since Plato it was firmly held by nearly all those who thought about the nature of human knowledge that  $\mathbf{k=jtb}$  — until the sudden, seismic publication of [29], which appeared to feature clear examples in which  $\mathbf{jtb}$  holds, but not  $\mathbf{k}$ . It would be quite fair to say that since the advent of Gettier’s piece, to this very day, defenders of  $\mathbf{k=jtb}$  have been rather stymied; indeed, it wouldn’t be unfair to say that not only such defenders, but in fact all formally inclined epistemologists, have since the advent of Gettier-style counter-examples been scurrying and scrambling about, trying to pick up the pieces and somehow build up again a sturdy edifice. Our account of learning *ex nihilo* includes a formal-and-computational solution to the Gettier

<sup>7</sup>In the very first episode of Season 1, “A Study in Pink,” Watson, who at this point doesn’t know Sherlock, imprudently classifies Sherlock as an “amateur”, and in response Sherlock recounts in rapid-fire a long chain of reasoning (deductive, abductive, and analogical) that has enabled him to come to know many, many things. A stupefied Watson thus comes to know that Sherlock is no amateur.

<sup>8</sup>Note that it’s knowledge of this genuine sort that’s needed to be human-level intelligent, for see the “cognitive skills” listed and analyzed in [27].

<sup>9</sup>For an argument, with which at least the first two authors are very sympathetic, that contemporary “machine learning” fails to produce knowledge for the agent that machine-“learns,” see [12].

problem, which in turn allows AIs built with our automated-reasoning technology (described below) to acquire knowledge in accord with  $k=jtb$ . But first, what is the Gettier problem?

Gettier [29] presents a scenario in which Smith has “strong evidence” for the proposition

$f$  Jones owns a Ford.

The evidence in question, Gettier informs us, includes that “Jones has at all times in the past within Smith’s memory owned a car, and always a Ford, and that Jones has just offered Smith a ride while driving a Ford.” In addition, Smith has another friend, Brown, whose whereabouts are utterly unknown to Smith. Smith randomly picks three toponyms and “constructs the following three propositions.”

$g$  Either Jones owns a Ford, or Brown is in Boston.

$h$  Either Jones owns a Ford, or Brown is in Barcelona.

$i$  Either Jones owns a Ford, or Brown is in Brest-Litovsk.

Of course,  $\{f\} \vdash g$ ,  $\{f\} \vdash h$ ,  $\{f\} \vdash i$ .<sup>10</sup> “Imagine,” Gettier tells us, “that Smith realized the entailment of each of these propositions he has constructed by”  $f$ , and on that basis is “completely justified in believing each of these three propositions.” Two further facts in the scenario yield the apparent counter-example, to wit: Jones doesn’t own a Ford, and is currently driving a rented car; and, in a complete coincidence, Brown is in fact in Barcelona. Gettier claims, and certainly appears to be entirely correct in doing so, that Jones doesn’t know  $h$ , yet  $h$  is true, Smith believes  $h$ , and Smith is justified in believing  $h$  — which is to say that  $jtb$  appears to be clearly instantiated!

Learning *ex nihilo* includes an escape from Gettier: Encapsulated to a brutal degree here, we gladly allow that the characters like Smith in Gettier’s [29] cases *do* have knowledge on the basis of a  $k=jtb$ -style account, but the knowledge in question can be cast at any number of six levels, ranging from 1 (*more likely than not*), the weakest, to 6 (*certain*) the strongest. Specifically, we hold that Smith knows at a level of 1, because belief in these cases is itself at a strength level of 1, and that’s because the argument serving as justification for belief in these cases only supports belief at that level. To our knowledge, this proposed solution to the counter-examples in question is new, though there are echoes of it in [19].<sup>11</sup> An AI-ready inductive logic that allows Gettier to be surmounted in this fashion is presented in [31].

<sup>10</sup>We are here by the single turnstile of course denoting some standard provability relation in a standard, elementary extensional collection of inference schemata, such as that seen in first-order logic =  $\mathcal{L}_1$ , a familiar logical system discussed below. The disjunction is of course inclusive.

<sup>11</sup>Echoes only. Chisholm’s main moves are flatly inconsistent with ours. E.g., his definition of the longstanding  $jtb$ -based concept of knowledge includes not merely that the agent is *justified* in believing  $\phi$ , but the stipulation that  $\phi$  is *evident* for the agent [19, p. 102]. In our multi-valued inductive scheme, *evident* is level 5, and is reserved for propositions believed on the basis of direct perception, e.g. your belief when at your desk that there is a book in the room with you. And Chisholm’s modifications of the  $j$  condition in the  $jtb$  triad are internalist, whereas ours are externalist, inhering as they do in formal structures and methods. Somewhat amazingly, the learning-*ex nihilo* diagnosis and resolution of Gettier cases is assumed in the literature to be non-existent. E.g., here is what’s said about Gettier cases in what is supposed to be the non-controversial and comprehensive SEP:

Epistemologists who think that the JTB approach is basically on the right track must choose between two different strategies for solving the Gettier problem. The first is to strengthen the justification condition to rule out Gettier cases as cases of justified belief. This was attempted by Roderick Chisholm;<sup>12</sup> . . . The other is to amend the JTB analysis with a suitable fourth condition, a condition that succeeds in preventing justified true belief from being “gettiered.” Thus amended, the JTB analysis becomes a JTB+X account of knowledge, where the ‘X’ stands for the needed fourth condition. [37, §3, “The Gettier Problem”]

Yet the learning *ex nihilo*-solution, while retaining the  $jtb$  kernel, is based on neither of these two different strategies.

## 6 On Logico-Mathematics of Learning *ex Nihilo*

Is there a logico-mathematics of learning *ex nihilo*? If so, what is it, at least in broad strokes? The answer to the first of these questions is a resounding affirmative — but in the present paper, intended to serve as primarily an introduction to a new form of human learning driven by reasoning, and concomitantly as a challenge to learning-focused AI researchers (incl. and perhaps esp. those in AI who pursue machine learning in the absence of reasoning carried out in confirmable conformity with declarative inference schemata), the reader is accordingly asked to be willing to rest content (at least provisionally) with but an encapsulation of the logico-mathematics in question, and references (beyond those in the previous §) to prior work upon which the formal theory of learning *ex nihilo* is based. This should be appropriate, given that via the present paper we seek to place before the community a chiefly philosophical introduction to learning *ex nihilo*. We present the core of the relevant logico-mathematics, starting with the next paragraph. Our presentation presumes at least some familiarity with formal computational logic (both extensional and intensional<sup>12</sup>) and late 20th- and 21st-century automated deduction/theorem proving. (Learning *ex nihilo* is, as we shall soon see, explicitly based upon automated reasoning that is non-deductive as well, but automated non-deductive reasoning is something we can't expect readers to be familiar with.) For readers in the field of AI who are strong in statistical/connectionist ML, and/or reinforcement learning and Bayesian approaches/reasoning, but weak in formal computational logic, in either or both of its deductive and inductive forms, we recommend [7, 32], and then working backwards through readily available textbooks and papers cited in this earlier IJCAI-venue work, and in the next two sub-sections.

### 6.1 Logical Systems and Learning *Ex Nihilo*

The concept of a *logical system*, prominent in the major result known as Lindström's Theorem,<sup>13</sup> provides a detailed and rigorous way to treat logics abstractly and efficiently, so that e.g. we can examine and (at least sometimes) determine the key attributes that these logics have, relative to their expressive power. In keeping with this theorem, we take a logical system  $\mathcal{L}$  to be a triple

$$\langle \mathcal{L}, \mathcal{I}, \mathcal{S} \rangle$$

whose elements are, in turn, a formally specified language  $\mathcal{L}$  (which would customarily be organized in ways that are familiar in programming languages; e.g. types would be specified); an inference theory  $\mathcal{I}$  (which would be a proof theory in the deductive case, an argument theory in the inductive case, and best called an *inference theory* when inference schemata from both categories are mixed) that allows for precise and machine-checkable proofs/arguments, composed of inference schemata; and some sort of semantics  $\mathcal{S}$  by which the meaning of formulae in  $\mathcal{L}$  are to be interpreted. (As explained in Appendix A, in a cognitive calculus semantical conditions are folded into inference schemata.)

<sup>12</sup>Roughly, extensional logic invariably assigns a semantic value to formulae in a purely compositional way, and is ideal for formalizing mathematics itself; this is why the logical systems traditionally used in mathematical logic are such things as first-order and second-order logic. Such logical systems, in their elementary forms, are of course covered in the main AI textbooks of today, e.g. [59, 47]. In stark contrast, the meaning of a formula  $\phi$  in an intensional logic can't be computed or otherwise obtained from it and what it's composed of. For a simple example, if  $\phi$  is  $\psi \rightarrow \delta$ , and we're in (extensional) zeroth-order logic (in which  $\rightarrow$  is the material conditional), and we know that  $\psi$  is false, then we know immediately that  $\phi$  is true. But if  $\phi$  is instead  $\mathbf{B}_a\psi$ , where  $\mathbf{B}$  is an agent-indexed belief operator in epistemic logic, and  $\psi$  is what agent  $a$  believes, the falsity of  $\psi$  doesn't at all guarantee any truth-value for  $\mathbf{B}_a\psi$ . Here, the belief operator is an intensional operator, and is likely to specifically be a modal operator in some modal logic.

<sup>13</sup>Elegantly covered in [23].

Each of the elements of the abstract triple that individuates a given logical system can be vast and highly nuanced, and perhaps even a substantive part of a branch of formal logic in its own right. For example, where the logical system is standard first-order logic  $\mathcal{L}_1$ ,  $\mathcal{S}$  will include all of established model theory for first-order logic. Lindström’s Theorem tells us (roughly) that any movement to an extensional logical system whose expressive power is beyond  $\mathcal{L}_1$  will lack certain meta-attributes that many consider desirable. For instance, (standard) second-order logic  $\mathcal{L}_2$  isn’t complete, whereas  $\mathcal{L}_1$  is. Despite this, some AI researchers work productively with higher-order extensional logics.<sup>14</sup>

For learning *ex nihilo*, the most important element in the triple that makes for a logical system is  $\mathcal{I}$ , which can be viewed as a set of inference schemata.<sup>15</sup> The reason is that learning *ex nihilo* is based on reasoning in which each inference is sanctioned by some  $I \in \mathcal{I}$ , and on the *automation* of this reasoning, including when the inference schemata are non-deductive. In computer science and AI, a considerable number of people are familiar with automated *deductive* reasoners; after all, Prolog is based on automated deductive reasoning, using only one inference schema (resolution), involving formulae in a fragment of  $\mathcal{L}_1$ . Learning *ex nihilo*, in contrast, is based on automated reasoning over *any* inference schemata — not only deductive ones, but inductive ones, e.g. ones that regiment analogical reasoning, abductive reasoning, enumerative inductive reasoning, and so on.<sup>16</sup> All the reasoning patterns seen in inductive logic, in their formal expressions, are possible as inference schemata employed in learning *ex nihilo*.<sup>17</sup> This means that an inference schema is of the form

$$\frac{\text{Condition(s)}}{\text{Conclusion(s)}} \mathcal{I}$$

where both “top” and “bottom” can be built from any formal declarative statements pertaining to formulae that are seen in formal logic, period. There are no restrictions even on the mixing of object-level formal logic with formal meta-logic or meta-meta-logic. For instance, to convey the extreme openness of inference schemata for learning *ex nihilo* and cognitive calculi, consider that the following is a perfectly acceptable inference schema that sanctions inferring a simple atomic formula expressed in the language of first-order axiomatic (ZF) set theory, where ‘ $a$ ’ denotes a particular set, and the condition in this case says that the cardinality of  $a$  equals that of the natural numbers:

$$\frac{\text{card}(a) = \text{card}(\mathbb{N})}{\text{Big}(a)} \mathcal{I}'$$

<sup>14</sup>E.g. the formal verification of Gödel’s famous ontological argument for God’s existence, an argument that employs  $\mathcal{L}_3$ , has been verified by AI researchers; see e.g. [6].

<sup>15</sup>For simple logical systems, the phrase ‘inference rules’ is often used instead of the more accurate ‘inference schemata,’ and in fact there is a tradition in places of using the former. Because even an individual inference schema can be quite complex, and can involve meta-logical constructs and computation, talk of schemata is more accurate. For instance, a common inference schema in so-called natural deduction is

$$\frac{\phi(a), \text{ where } a \text{ is a constant in formula } \phi}{\exists x \phi(x)}$$

but all sorts of further restrictions can be (and sometimes are) placed on  $\phi(a)$ , such as that it must be a  $\Delta_0$  formula. As such things grow more elaborate, it quickly makes precious little sense to call them “rules,” and besides which many think of them as programs.

<sup>16</sup>When automated reasoning ranges as generally as this, the first author maintains that all of computer science can be captured in formal logic; see [14].

<sup>17</sup>For a particular example of formal, automated reasoning that blends deduction with analogical reasoning, see [43]. For a readable overview of inference patterns in inductive logic that we formalize and automate, see [39].

## 6.2 From Logical Systems to Cognitive Calculi

Because learning *ex nihilo* frequently involves the mental states of other agents (as seen e.g. in the parable regarding Robert, where iterated epistemic operators played a role), we — as announced above — employ a novel class of logical systems called *cognitive calculi*, and they form part of the singular basis of this new kind of learning. A cognitive calculus, put simply, is a logical system in which  $\mathcal{L}$  includes intensional operators (e.g. for such things as belief, desire, intention, emotional states, communication, perception, and attention);  $\mathcal{I}$  includes at least one inference schema that involves such operators; and the meaning of formulae, determined by some particular  $\mathcal{S}$ , because they can in their expressive power far outstrip any standard, off-the-shelf semantics (such as possible-worlds semantics), is generally proof-theoretic in nature. A more detailed account of cognitive calculi is given in Appendix A, and some detailed information about a particular one in Appendix B.

## 6.3 The Learning *Ex Nihilo* Loop

Learning *ex nihilo* happens when an agent loops through time as follows (put in broad strokes): Identify Interest/Goal  $\Rightarrow$  Query  $\Rightarrow$  Discover Argument/Proof to Answer Query  $\Rightarrow$  Learn  $\Rightarrow$  Identify Interest/Goal, etc. This cycle is at work in the two parables with which we began. We do not have space to detail this process. In particular, the management of the agent’s interests (or goals) requires planning — but the emphasis in the present paper, for economy, is on reasoning. Below we do discuss not only the AI technology that brings this loop to life, but some simulation of the process in our earlier parables.

# 7 The Automation of Learning *Ex Nihilo*

But how do we pass from the abstract logico-mathematics of learning *ex nihilo* to an artificial agent that can bring such a thing to life? The answer should be quite obvious, in general: We build an AI that can find the arguments that undergird the knowledge obtained by learning *ex nihilo*. In turn, this means that we need an automated reasoner of sufficient power and reach that can pursue epistemic interests, and a planner that can at least manage (e.g. prioritize) interests. This brings us to the following progression, in which we now briefly describe one such reasoner (ShadowReasoner), and then give an illustrative simulation made possible by this AI technology.

## 7.1 Automated Reasoner: ShadowReasoner

A large amount of research and engineering has gone into building first-order theorem provers in the last few decades. ShadowReasoner leverages this progress by splitting any logic into a first-order core and the “remainder,” and then calls a first-order theorem prover when needed. Briefly, ShadowReasoner partitions the inference schemata for a given  $\mathcal{L} \equiv \langle \mathcal{L}, \mathcal{I}, \mathcal{S} \rangle$  into two parts  $\mathcal{I}_1$  and  $\mathcal{I}_{>1}$ . The first part  $\mathcal{I}_1$  consists of inference schemata that can be carried out by a first-order theorem prover when the input expressions are *shadowed* down into first-order logic. The second part consists of inference schemata that cannot be reduced to first-order reasoning. Given any problem in a logic, ShadowReasoner alternates between trying out  $\mathcal{I}_{>1}$  and calling a first-order theorem prover to handle  $\mathcal{I}_1$ .

The core algorithm for ShadowReasoner has a theoretical justification based on the following theorem (which arises from the fact that first-order logic can be used to simulate Turing machines [8]):



**Theorem 1**

Given a collection of Turing-decidable inference schemata  $\mathcal{I}$ , for every inference  $\Gamma \vdash_{\mathcal{I}} \phi$ , there is a corresponding first-order inference  $\Gamma' \vdash \phi'$ .

We illustrate how ShadowReasoner works in the context of a first-order modal logic employed in [32]. Please note though there are some extant first-order modal-logic theorem provers, built upon reductions to first-order theorem provers, they have some deficiencies. Such theorem provers achieve their reduction to first-order logic via two methods. In the first method, modal operators are represented by first-order predicates. This approach is computationally fast but can quickly lead to well-known inconsistencies, as demonstrated in [11]. In the second method, the entire proof theory is implemented in first-order logic, and the reasoning is carried out *within* first-order logic. Here, the first-order theorem prover simply functions as a programming system. The second method, while accurate, can be excruciatingly slow.

ShadowReasoner uses the different approach alluded to above — one in which it alternates between calling a first-order theorem prover and applying non-first-order inference schemata. When we call the first-order prover, all non-first-order expressions are converted into propositional atoms (i.e., shadowing) to prevent substitution into non-first-order contexts, as such substitutions lead to inconsistencies [11]. This approach achieves speed without sacrificing consistency. The algorithm is briefly described below.

First we define the syntactic operation of **atomizing** a formula, denoted by  $A$ . Given any arbitrary formula  $\phi$ ,  $A_{[\phi]}$  is a unique atomic (propositional) symbol. Next, we define the **level** of a formula:  $\text{level} : \mathcal{L} \rightarrow \mathbb{N}$ .

$$\text{level}(\phi) = \begin{cases} 0; \phi \text{ is purely propositional formulae; e.g. } \textit{Rainy} \\ 1; \phi \text{ has first-order predicates or quantifiers e.g. } \textit{Sleepy}(jack) \\ 2; \phi \text{ has modal formulae e.g. } \mathbf{K}(a, t, \textit{Sleepy}(jack)) \end{cases}$$

Given the above definition, we can define the operation of **shadowing** a formula to a level:

**Shadowing**

To shadow a formula  $\chi$  to a level  $l$ , replace all sub-formulae  $\chi'$  in  $\chi$  such that  $\text{level}(\chi') > l$  with  $A_{[\chi']}$  simultaneously. We denote this by  $S[\phi, l]$ .

For a set  $\Gamma$ , the operation of shadowing all members in the set is simply denoted by  $S[\Gamma, l]$ .

Assume we have access to a first-order prover  $\mathbf{P}_F$ . For a set of pure first-order formulae  $\Gamma$  and a first-order  $\phi$ ,  $\mathbf{P}_F(\Gamma, \phi)$  gives us a proof for  $\Gamma \vdash_{\mathcal{I}_1} \phi$  if such a first-order proof exists; otherwise  $\mathbf{P}_F$  returns NO.

## 7.2 Illustrative Simulation

Figures 1 and 2 illustrate the dinner-party parable simulated in the **deontic cognitive event calculus** (*DCEC*) using ShadowReasoner within the hypergraphical argument-construction system HyperSlate<sup>®</sup>; see [16] for a similar, indeed ancestral, but less intelligent, hypergraphical system. See Appendix B for a description of syntax and inference schemata of *DCEC*. Figure 1 simulates in pure first-order logic Robert’s learning that his drink is an aperitif. Figure 2 is a proof in a cognitive calculus, viz. the one described in [32], of Robert learning the following propositions: “Robert believes that his host is wealthy”, “The host believes Robert believes that his host is wealthy”, and “Robert believes that his host believes Robert believes that his host is wealthy.”<sup>18</sup> The figures illustrate first-order and cognitive-calculus reasoners (shown as FOL  $\vdash$

<sup>18</sup> With background information  $t_i < t_j$  if  $i < j$ .

```

Input: Input Formulae  $\Gamma$ , Goal Formula  $\phi$ 
Output: A proof of  $\Gamma \vdash_{\mathcal{I}} \phi$  if such a proof exists, otherwise fail
initialization;
while goal not reached do
  answer =  $\mathbf{P}_F(S[\Gamma, 1], S[\phi, 1])$ ;
  if answer  $\neq$  NO then
    return answer ;
  else
     $\Gamma' \leftarrow$  expand  $\Gamma$  by using any applicable  $\mathcal{I}_{>1}$ ;
    if  $\Gamma' = \Gamma$  then
      /* The input cannot be expanded further
      return fail
    else
      set  $\Gamma \leftarrow \Gamma'$ 
    end
  end
end

```

Algorithm 1: ShadowReasoner Core Algorithm

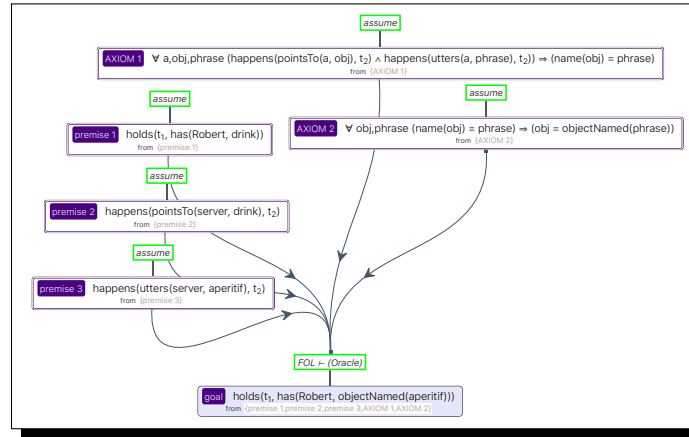


Figure 1: Dinner Party Example Part 1. (The proof here is in the HyperSlate® system (Pat. Pend.) published by Motalen in it’s Logic: A Modern Approach (LAMA)® paradigm. See [www.logicmodernapproach.com](http://www.logicmodernapproach.com).)

and  $\text{CC} \vdash$ , resp.) being employed to derive these statements. Automated discovery of the proofs in 1 took  $\sim 0.1(ms)$ , and the proofs in 2 took  $\sim 0.9(s)$ .

## 8 Objections, Rebuttals

We now anticipate and rebut six objections likely to come from skeptics, including specifically those immersed in forms of learning far removed from any notion of machine-verifiable proof or argument enabled by inference schemata.

### 8.1 Objection 1: This isn’t learning from *nothing*!

The first objection is that ‘learning *ex nihilo*’ is a misnomer. The rationale for this complaint is simply the reporting of an observation that should be clear to all: viz. that the learning in

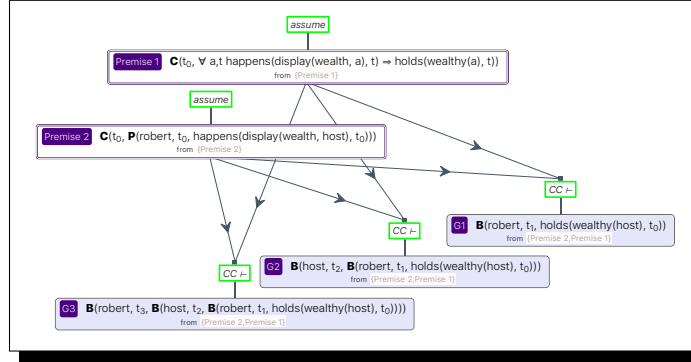


Figure 2: Dinner Party Example Part 2. (The proof here, like its predecessor, is in the HyperSlate® system (Pat. Pend.) published by Motalen in it’s Logic: A Modern Approach (LAMA)® paradigm. See [www.logicmodernapproach.com](http://www.logicmodernapproach.com).)

question undeniably makes use of pre-existing stuff; hence we’re not literally dealing with learning *from nothing*. In the parable of the dinner party, e.g., Robert has brought his pre-existing command of elementary arithmetic to the table; ditto for much other pre-known propositional content. So how then is it fair to speak of learning *ex nihilo*? It’s fair because obviously learning *ex nihilo* trades on the pre-existing concept of creation *ex nihilo*, and that millennia-old conception allows that the Almighty (by definition!) was around before the creation in question occurred. And of course this is just one part of pre-creation stuff in creation *ex nihilo*: God presumably needed to have a mental concept of a planet to create a planet. We generally suspect that learning *ex nihilo* begins to kick into “high gear” in the human sphere when children are sophisticated enough to exploit their prior knowledge of content that requires, for its underlying representation,  $\mathcal{L}_1$  and basic modal logic. From the standpoint of logicist cognitive science, rather than AI, this means that learning *ex nihilo* would be aligned with the views of Piaget and colleagues [e.g. 38], Stenning and colleagues [e.g. 62], Bringsjord and colleagues [e.g. 10, 13], and Rips and colleagues [e.g. 57, 58]. The goal of full formalization and implementation of learning *ex nihilo* would likely be regarded by these cognitive scientists as a welcome one.

## 8.2 Objection 2: Isn’t this just reasoning?

The second objection we anticipate is that learning *ex nihilo* isn’t learning at all; it’s just a form of reasoning. In reply, any process, however much it relies on reasoning, that does enable an agent running that process to acquire genuine knowledge (and our  $k=j\tau b$  definition of knowledge, note, is a very demanding one) would seem to be quite sensibly classified as a learning process. In fact, probably it strikes many as odd to say that one has a form of machine learning that does *not* result in the acquisition of any knowledge.

## 8.3 Objection 3: What about inductive logic programming?

The objection here can be encapsulated thus: “What about inductive logic programming (ILP)? Surely this established family of techniques both uses formal logic, and results in new knowledge.”

ILP is along the general lines of learning *ex nihilo*, agreed; but ILP is acutely humble and highly problematic when measured against LEN, for reasons we give momentarily. Before giving

these reasons, without loss of generality we fix a general framework  $\mathcal{F}$  [51] to fix abduction and induction in supposedly representative logicist fashion:

**Given:** Background knowledge,  $B$ , and observations (data),  $O$ , both represented as sets of formulae in first-order predicate calculus, where  $O$  is restricted to ground formulae.

**Find:** An hypothesis  $H$  (also a set of logical formulae) such that  $B \cup H \not\vdash \perp$  and  $B \cup H \vdash O$ .

From  $\mathcal{F}$  one can derive both induction and abduction, according to Mooney: For the latter,  $H$  is restricted to atomic formulae or simple formulae of the general form  $\exists x\phi$ , and  $B$  is — as Mooney says — “significantly larger” than  $H$ . For induction, Mooney says that  $H$  is to consist of universally quantified Horn clauses, and  $B$  is “relatively small” and may even be the null set.

From an explain-the-extant-literature point of view,  $\mathcal{F}$  is at least somewhat attractive. For as Mooney points out, this framework captures what many logic-oriented learning researchers in AI have taken induction and abduction to be; this includes, specifically, ILP researchers [e.g. 53, 41]. Unfortunately, despite its explanatory virtue,  $\mathcal{F}$ , when measured against human cognition of the sort involved in learning *ex nihilo*, is embarrassingly inadequate. As we have said, this can be shown effortlessly via many reasons. We quickly mention just four here.

### 8.3.1 Reason 1: $\mathcal{F}$ Runs Afoul of Non-Deductive Reasoning

Why should the combination of background knowledge and a candidate hypothesis need to *deductively entail* some observation, as  $\mathcal{F}$  says (via its use of  $\vdash$ )? Suppose that as Smith sits in his home office looking through a window he perceives ( $\omega$ ) that a uniformed man with a FEDEX cap is approaching Smith’s house, a small white box in hand. Smith has no doubt learned that ( $\delta$ ) a delivery is about to be attempted, but does it really need to be the case that, where  $B$  is background knowledge,  $B \cup \{\omega\}$  can be used to *prove*  $\delta$ ? No, it doesn’t. Maybe it’s Halloween, Smith forgot that it is, and the person approaching is in costume and playacting. Maybe the approaching man is a criminal in disguise, merely casing Smith’s domicile. And so on. All that is needed is a fairly strong *argument* in support of  $\delta$ . And that argument can make use of inferences that are deductively invalid, but valid as reasoning that is analogical, abductive, inductive, etc.<sup>19</sup>

### 8.3.2 Reason 2: $\mathcal{F}$ Leaves Aside Other Non-Deductive Reasoning

This reason was revealed in the previous sentence: that sentence refers to not just to abduction and induction, but also to reasoning that is *analogical* in nature, and such reasoning isn’t included in ILP. In fact, learning via analogical reasoning is often left aside in coverage of logicist learning of *any* textbook variety.<sup>20</sup> But if in the scenario depicted immediately above Smith had encountered not a FEDEX man, but rather a USPS mailman making a delivery to his house, he may well have believed that the man with the FEDEX cap was analogous, and hence was making a delivery.

### 8.3.3 Reason 3: $\mathcal{F}$ Runs Afoul of Robust Expressivity

A quick but decisive third reason Mooney’s  $\mathcal{F}$  explodes in the face of real human cognition is that any expressivity restriction on  $O$  and/or  $H$  is illogical, and certainly any specific restriction that  $O$  be restricted to ground formulae and/or that  $H$  be confined to Horn-clause logic (or

<sup>19</sup>Indeed, these inferences can even be *formally* valid in the inductive logics that will undergird ALML; see below.

<sup>20</sup>E.g. learning by analogy isn’t included in AI’s definitive, otherwise encyclopedic textbook: [59].

even for that matter full first-order logic) is *patently* illogical. This can be seen in middle-school classes that cover arithmetic, where even very young students cook up and affirm hypotheses that are expressed using infinitary constructions beyond even full first-order logic. For instance: Student Johnny is reminded of the definition of a prime number, and then shown that 4 is equal to 3 plus 1, that 6 is equal to 5 + 1, that 8 = 5 + 3, etc. Johnny is asked to consider whether the next two even numbers continue the pattern. He observes that 10 = 7 + 3 and that 12 = 7 + 5, and is now inclined to hypothesize ( $H'$ ) that every even integer greater than 2 is the sum of two primes. A natural form for  $H'$ , where  $e_1$  is 2 and the even numbers from there continue  $e_2, e_3, \dots$  is simply a list  $L$  like:

$$\begin{aligned} e_1 &= p_1 + p'_1 \\ e_2 &= p_2 + p'_2 \\ &\vdots \end{aligned}$$

Yet  $L$  cannot be expressed in finitary formulae,<sup>21</sup> and even if one squeezed  $L$  into finitary first-order logic, that would be done by employing the same trick as is used for instance in axiomatic set theory, where first-order logic is augmented with schematic formula that denote an infinite number of instantiations thereof.<sup>22</sup> Regardless, even if by some miracle  $H'$  could be expressed in some finitary extensional logic at or beneath first-order logic, the classroom in question wouldn't exactly operate well without the teacher's believing Johnny's believing that  $H'$ , and certainly nothing like this fact is expressible in extensional logic (let alone Horn-clause logic).

### 8.3.4 Reason 4: $\mathcal{F}$ Ignores Thinking

The dictum that truth is stranger than fiction, alas, is frequently confirmed by the oddities of contemporary AI research. The fourth reason  $\mathcal{F}$  is inadequate is a confirming example. For while the framework  $\mathcal{F}$  projects an air of being about thinking, in actuality it leaves thinking aside. Indeed there's a rather common fallacy at work in  $\mathcal{F}$ , and its promotion: the fallacy of composition. For while  $\mathcal{F}$  does include some forms of reasoning, these forms (and even, for that matter, *all* forms of reasoning put together) don't comprise thinking; thinking includes the reasoning called out in  $\mathcal{F}$  as merely a *proper part*. To be more specific, in real and powerful thinking, an hypothesis  $H$  can be wisely discarded when there is evidence against it.

The fourth flaw infecting  $\mathcal{F}$  can be easily unpacked: Real learning is intertwined with the full gamut of human-level thinking: specifically with planning, decision-making, and communication. If there are no observations to learn from, an agent can plan to get more observations. An agent can decide when to learn and what to learn. Precious little substantive learning takes place without communication between teacher and learner, including written communication. And finally, at the highest end of the spectrum of powerful learning, learners formalize learning itself, and learn more by doing so.

## 8.4 Objection 4: Isn't this just deduction?

The next objection can be expressed thus: "While you refer to non-deductive reasoning, you simply rely on deduction."

<sup>21</sup>It is naturally represented by an infinite conjunction, which the logic  $\mathcal{L}_{\omega_1\omega}$  allows.

<sup>22</sup>E.g., witness the Axiom *Schema* of Separation as a replacement for Frege's doomed Axiom V, shot down violently and decisively by the famous Russell's Paradox; see [63] for wonderfully elegant coverage.

As stated elsewhere, we allow for non-deductive forms of reasoning, as long as that reasoning can be captured by natural inference schemata as defined in in a cognitive calculus (see Appendix A). There are many inference schemata for analogical arguments that can be expressed formally in natural premise-conclusion form. For example, the schemata in Licato and Fowler (2016) govern confidence propagation between a cognitive agent’s beliefs. The many argument schemas described by Walton and colleagues [65, 48, 49] also lend themselves to inference schemata. These distinguish between familiar inference patterns, such as “Practical Reasoning from Analogy” (roughly: if the right thing to do in some scenario was  $X$ , then the right thing to do in a similar scenario is  $Y$ ) and “Argument from Precedent.”

Analogical inferences are central to human-level cognition [26, 34, 24, 35] and are essential to how humans reason in new and unfamiliar domains. Analogical reasoning has even been used to explain how children — the greatest *ex nihilo* learners of all — can so effectively bootstrap their knowledge of the world along with the concepts they use to describe it [21, 20, 22]. Our account of learning *ex nihilo* not only recognizes, but embraces this fact, allowing for analogical reasoning to be carried out seamlessly with other reasoning types (e.g., as in analogico-deductive reasoning [44, 46, 36]).

## 8.5 Objection 5: What about Lewis’ “Effusive Knowledge”?

“You claim to have a solution to the Gettier problem, one seemingly based on introducing different levels of knowledge of  $p$  (why six?), based on the different levels of justification one may have for  $p$ . This idea echoes Lewis [42] epistemic contextualism, but does it really solve the Gettier problem? And I don’t see any justification for why, say, level-5 jtb is to be equated with full knowledge?”

Actually, Lewis’ “Effusive Knowledge” paper espouses a conception that is the *opposite* of what underlies learning *ex nihilo*. Lewis holds that there is no knowledge in the Gettier scenarios, because his definition for knowledge (which marks a rejection of  $k=jtb$ ) isn’t satisfied in these cases. Learning *ex nihilo* entails that there *is* knowledge in these scenarios, but *reduced* knowledge. If the degree of belief is  $k$ , then knowledge partaking of this belief is of degree  $k$  as well. In Gettier’s original scenarios, knowledge is at the level of 2 (for reasons too far afield to articulate under current space constraints). It seems never to have occurred to Lewis that if belief comes in degrees, then knowledge (which surely includes belief<sup>23</sup>) must itself come in degrees. In the context of formal intensional logic (e.g. formal epistemic logic), Lewis’ position/paper is from our formalist point of view weak, because it hinges on the repeatedly asserted-without-argument claim that in the case of a single-ticket lottery of size  $m$ , even when the number of tickets is arbitrarily large (e.g.  $m = 1$  quadrillion), one cannot know that a given ticket  $t_k$  ( $1 \leq k \leq 1Q$ ) will not win. From the standpoint of learning *ex nihilo*, this is bizarre, because surely one knows that in the next moment one’s computer will not spontaneously combust, because such an event is preposterously unlikely. (And looking back in time, surely one knows that the computer sitting here now was there 10 seconds ago.) Moreover, Lewis rejects the very concept of justification as part and parcel of knowledge; learning *ex nihilo* by contrast is an AI-driven conception based on automated reasoning (automated deduction and automated analogical, abductive, enumerative inductive etc. reasoning).

---

<sup>23</sup>Actually, Lewis asserts that there can be knowledge without belief, because a timid student can know the answer, but has “no confidence that he has it right, and so does not believe what he knows” (p. 556). In our formal framework, the situation is that the student believes, at a low level (1, say) that he knows (at some level  $k \geq 1$ ) the answer, and as a matter of fact he *does* know the answer at level 1. This scenario is provably consistent in our relevant cognitive calculi. Not only that, but as far as we can tell, since in point of fact timidity of this type often prevents successful performance, our formal diagnosis has “real-world” plausibility.

As to levels of belief, in the case of a 13-valued inductive logic we use to undergird learning *ex nihilo*, 6 = *certain*, 5 = *evident*, 4 = *overwhelmingly likely* = *beyond reasonable doubt*, 3 = *very likely*, 2 = *likely*, 1 = *more likely than not*. 0 is *counterbalanced*, and then we have the symmetrical negative values. A belief at level 6 corresponds to knowing things that Lewis (and everyone else) agree that we know: e.g. knowing that  $2+2=4$ . We do not have space to explain that each of the values from 0 to 6 are readily found in normatively correct, first-rate human reasoning, in the formal and empirical sciences.

## 8.6 Objection 6: Frivolity?

Finally, some will doubtless declare that learning *ex nihilo* is frivolous. What good is it, really, to sit at a dinner table and learn useless tid-bits? This attitude is most illogical. The reason is that, unlike what is currently called “learning,” only persons at least at the level of humans can learn *ex nihilo*, and this gives such creatures power, for the greatest of what human persons have learned (and, we wager, *will* learn) comes via learning *ex nihilo*. In support of this we could cite countless confirmatory cases in the past, but rest content to but point out that armchair learning *ex nihilo* regarding simultaneity (Einstein) and infinitesimals (Leibniz) was rather productive.<sup>24</sup>

## 9 Conclusion and Next Steps

We have provided an introduction, philosophical in nature, but also to a degree formal and computational, to the new concept of learning *ex nihilo*, and have included enough information re. its formal foundation to allow those conversant with logicist AI to better understand this type of learning. In addition, we have explained that learning *ex nihilo* can be automated via sufficiently powerful automated-reasoning technology. Of course, this is a brief paper. Accordingly, next steps include dissemination of further details, obviously. But more importantly, what is the relationship between learning *ex nihilo* and types of machine learning that are based on artificial neural networks, Bayesian reasoning, reinforcement learning, and so on? These other currently popular types of learning are certainly *not* logicist, and hence nothing like a logical system, let alone a cognitive calculus, are present. In fact, it’s far from clear that it’s even possible to construct the needed machinery for learning *ex nihilo* out of the ingredients that go into making these non-logicist forms of learning.

**Acknowledgements** We are deeply grateful for current support from the U.S. Office of Naval Research to invent, formalize, and implement new forms of learning based on automated reasoning. Prior support from DARPA of logicist learning has also proved to be helpful, and for this too the first author expresses thanks. Support from AFOSR has helped us invent cognitive calculi, and automated-reasoning technology for them. Four anonymous reviewers provided insightful feedback, for which we are deeply grateful.

---

<sup>24</sup>And for those readers with a literary bent, it should also be pointed out that the great minds of detection, not only the aforementioned Sherlock Holmes but e.g. Poe’s seminal Le Chevalier C. Auguste Dupin, achieve success primarily because of their ability to learn *ex nihilo*.

## References

- [1] K. Arkoudas and S. Bringsjord. Toward Formalizing Common-Sense Psychology: An Analysis of the False-Belief Task. In T.-B. Ho and Z.-H. Zhou, editors, *Proceedings of the Tenth Pacific Rim International Conference on Artificial Intelligence (PRICAI 2008)*, number 5351 in Lecture Notes in Artificial Intelligence (LNAI), pages 17–29. Springer-Verlag, 2008.
- [2] K. Arkoudas and S. Bringsjord. Propositional Attitudes and Causation. *International Journal of Software and Informatics*, 3(1):47–65, 2009.
- [3] Konstantine Arkoudas and David Musser. *Fundamental Proof Methods in Computer Science: A Computer-Based Approach*. MIT Press, Cambridge, MA, 2017.
- [4] M. Ashcraft. *Human Memory and Cognition*. HarperCollins, New York, NY, 1994.
- [5] C. Benzmüller and D. Miller. Automation of Higher-Order Logic. In *Handbook of the History of Logic; Volume 9: Logic and Computation*. North Holland, Amsterdam, The Netherlands, 2014.
- [6] Christoph Benzmüller and Bruno Woltzenlogel Paleo. Automating Gödel’s Ontological Proof of Gods Existence with Higher-order Automated Theorem Provers. In Torsten Schaub, Gerhard Friedrich, and Barry O’Sullivan, editors, *Proceedings of the European Conference on Artificial Intelligence 2014 (ECAI 2014)*, pages 93–98, Amsterdam, The Netherlands, 2014. IOS Press.
- [7] Christoph Benzmüller and Bruno Woltzenlogel Paleo. The Inconsistency in Gödel’s Ontological Argument: A Success Story for AI in Metaphysics. In Subbarao Kambhampati, editor, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*, pages 936–942. AAAI Press, 2016.
- [8] George S. Boolos, John P. Burgess, and Richard C. Jeffrey. *Computability and Logic (Fifth Edition)*. Cambridge University Press, Cambridge, UK, 2003.
- [9] S. Bringsjord. Is the Connectionist-Logicist Clash one of AI’s Wonderful Red Herrings? *Journal of Experimental & Theoretical AI*, 3(4):319–349, 1991.
- [10] S. Bringsjord, E. Bringsjord, and R. Noel. In Defense of Logical Minds. In *Proceedings of the 20<sup>th</sup> Annual Conference of the Cognitive Science Society*, pages 173–178. Lawrence Erlbaum, Mahwah, NJ, 1998.
- [11] S. Bringsjord and N. S. Govindarajulu. Given the Web, What is Intelligence, Really? *Metaphilosophy*, 43(4):361–532, 2012. This URL is to a preprint of the paper.
- [12] S. Bringsjord, N.S. Govindarajulu, S. Banerjee, and J. Hummel. Do Machine-Learning Machines Learn? In V. Müller, editor, *Philosophy and Theory of Artificial Intelligence 2017*, pages 136–157. Springer SAPERE, Berlin, Germany, 2018. This book is Vol. 44 in the book series. The paper answers the question that is its title with a resounding No. A preprint of the paper can be found via the URL given here.
- [13] Selmer Bringsjord. Declarative/Logic-Based Cognitive Modeling. In Ron Sun, editor, *The Handbook of Computational Psychology*, pages 127–169. Cambridge University Press, Cambridge, UK, 2008.



- [14] Selmer Bringsjord. Computer Science as Immaterial Formal Logic. *Philosophy & Technology*, forthcoming.
- [15] Selmer Bringsjord, John Licato, and Alexander Bringsjord. The Contemporary Craft of Creating Characters Meets Today’s Cognitive Architectures: A Case Study in Expressivity. In Jeremy Turner, Michael Nixon, Ulysses Bernardet, and Steve DiPaola, editors, *Integrating Cognitive Architectures into Virtual Character Design*, pages 151–180. IGI Global, Hershey, PA, 2016.
- [16] Selmer Bringsjord, Joshua Taylor, Andrew Shilliday, Micah Clark, and Konstantine Arkoudas. Slate: An Argument-Centered Intelligent Assistant to Human Reasoners. In Floriana Grasso, Nancy Green, Rodger Kibble, and Chris Reed, editors, *Proceedings of the 8th International Workshop on Computational Models of Natural Argument (CMNA 8)*, pages 1–10, Patras, Greece, July 21 2008. University of Patras.
- [17] E. Charniak and D. McDermott. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, MA, 1985.
- [18] M. Chi, N. Leeuw, M. Chiu, and C. Lavancher. Eliciting self-explanations improves understanding. *Cognitive Science*, 18:439–477, 1994.
- [19] R. Chisholm. *Theory of Knowledge 2nd ed.* Prentice-Hall, Englewood Cliffs, NJ, 1977.
- [20] Leonidas A. Doumas and John E. Hummel. Comparison and Mapping Facilitate Relation Discovery and Predication. *PLoS ONE*, 8(6), 2013.
- [21] Leonidas A. Doumas, John E. Hummel, and C.M. Sandhofer. A Theory of the Discovery and Predication of Relational Concepts. *Psychological Review*, 115(1):1–43, 2008.
- [22] Leonidas A. A. Doumas, Guillermo Puebla, and Andrea E. Martin. How we learn things we don’t know already: A theory of learning structured representations from experience. *bioRxiv*, 2017.
- [23] H. D. Ebbinghaus, J. Flum, and W. Thomas. *Mathematical Logic (second edition)*. Springer-Verlag, New York, NY, 1994.
- [24] Kenneth D. Forbus. How Minds Will Be Built. *Advances in Cognitive Systems*, 1:47–58, 2012.
- [25] N. Francez. *Proof-theoretic Semantics*. College Publications, London, UK, 2015.
- [26] D Gentner. Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7(2):155–170, 1983.
- [27] D. Gentner. Why We’re So Smart. In D. Gentner and S. Golden-Meadow, editors, *Language in Mind: Advances in the Study of Language and Thought*, pages 195–235. MIT Press, Cambridge, MA, 2003.
- [28] Gerhard Gentzen. Investigations into Logical Deduction. In M. E. Szabo, editor, *The Collected Papers of Gerhard Gentzen*, pages 68–131. North-Holland, Amsterdam, The Netherlands, 1935. This is an English version of the well-known 1935 German version.
- [29] Edmund Gettier. Is Justified True Belief Knowledge? *Analysis*, 23:121–123, 1963.

- [30] E. B. Goldstein. *Cognitive Psychology: Connecting Mind, Research, and Everyday Experience*. Cengage Learning, Boston, MA, 2008. This is the 5th edition.
- [31] Naveen Sundar Govindarajulu and Selmer Bringsjord. Strength Factors: An Uncertainty System for Quantified Modal Logic. In V. Belle, J. Cussens, M. Finger, L. Godo, H. Prade, and G. Qi, editors, *Proceedings of the IJCAI Workshop on “Logical Foundations for Uncertainty and Machine Learning (LFU-2017)*, pages 34–40, Melbourne, Australia, August 2017. (The paper itself can also be obtained at this URL: <http://kryten.mm.rpi.edu/NSG.SB.StrengthFactorsQuantModalLogic.LFU-17.pdf>).
- [32] N.S. Govindarajulu and S. Bringsjord. On Automating the Doctrine of Double Effect. In C. Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*, pages 4722–4730. International Joint Conferences on Artificial Intelligence, 2017.
- [33] Govindarajulu, Naveen Sundar. ShadowProver, 2016.
- [34] Douglas R. Hofstadter. Epilogue: Analogy as the Core of Cognition. In Dedre Gentner, Keith J. Holyoak, and Boicho N. Kokinov, editors, *The Analogical Mind: Perspectives from Cognitive Science*, chapter 15. The MIT Press, 2001.
- [35] Douglas R. Hofstadter and Emmanuel Sander. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. Basic Books, Inc, 2013.
- [36] John E. Hummel, John Licato, and Selmer Bringsjord. Analogy, Explanation, and Proof. *Frontiers in Human Neuroscience*, 8(867), 2014.
- [37] Jonathan Ichikawa and Matthias Steup. The Analysis of Knowledge. In Edward Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2012.
- [38] B. Inhelder and J. Piaget. *The Growth of Logical Thinking from Childhood to Adolescence*. Basic Books, New York, NY, 1958.
- [39] Gregory Johnson. *Argument & Inference: An Introduction to Inductive Logic*. MIT Press, Cambridge, MA, 2016.
- [40] R. Kowalski and M. Sergot. A Logic-Based Calculus of Events. *New Generation Computing*, 4(1):67–95, 1986.
- [41] N. Lavrač and S. Džeroshki. *Inductive Logic Programming: Techniques and Applications*. Ellis Horwood, Hemel, UK, 1994.
- [42] David Lewis. Elusive Knowledge. *Australasian Journal of Philosophy*, 70(4):549–567, 1996.
- [43] J. Licato, N. S. Govindarajulu, S. Bringsjord, M. Pomeranz, and L. Gittelsohn. Analogico-Deductive Generation of Gödel’s First Incompleteness Theorem from the Liar Paradox. In Francesca Rossi, editor, *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI-13)*, pages 1004–1009, Beijing, China, 2013. Morgan Kaufmann. Proceedings are available online at <http://ijcai.org/papers13/contents.php>. The direct URL provided below is to a preprint. The published version is available at <http://ijcai.org/papers13/Papers/IJCAI13-153.pdf>.

- [44] John Licato, Selmer Bringsjord, and John E. Hummel. 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, 2012.
- [45] John Licato and Maxwell Fowler. Formalizing Confidence Propagation in Analogico-Inductive Reasoning. In *Proceedings of the 2016 Conference of the International Association for Computing and Philosophy (IACAP 2016)*, 2016.
- [46] John Licato, Naveen S. Govindarajulu, Selmer Bringsjord, Michael Pomeranz, and Logan Gittelson. 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)*, 2013.
- [47] George Luger. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving (6th Edition)*. Pearson, London, UK, 2008.
- [48] Fabrizio Macagno and Douglas Walton. Argument from Analogy in Law, the Classical Tradition, and Recent Theories. *Philosophy and Rhetoric*, 42(2), 2009.
- [49] Fabrizio Macagno, Douglas Walton, and Christopher Tindale. Analogical arguments: Inferential structures and defeasibility conditions. *Argumentation*, 31:221–243, 2017.
- [50] P. McNamara. Deontic Logic. In Edward Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2010. McNamara’s (brief) note on a paradox arising from Kant’s Law is given in an offshoot of the main entry.
- [51] Raymond Mooney. Integrating Abduction and Induction in Machine Learning. In P. Flach and A. Kakas, editors, *Abduction and Induction*, pages 181–191. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
- [52] E. Mueller. *Commonsense Reasoning: An Event Calculus Based Approach*. Morgan Kaufmann, San Francisco, CA, 2006. This is the first edition of the book. The second edition was published in 2014.
- [53] S. Muggleton, editor. *Inductive Logic Programming*. Academic Press, London, UK, 1992.
- [54] Michael Nelson. Propositional Attitude Reports. In Edward Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2015.
- [55] Dag Prawitz. The Philosophical Position of Proof Theory. In R. E. Olson and A. M. Paul, editors, *Contemporary Philosophy in Scandinavia*, pages 123–134. Johns Hopkins Press, Baltimore, MD, 1972.
- [56] A. S. Rao and M. P. Georgeff. Modeling Rational Agents Within a BDI-architecture. In R. Fikes and E. Sandewall, editors, *Proceedings of Knowledge Representation and Reasoning (KR&R-91)*, pages 473–484, San Mateo, CA, 1991. Morgan Kaufmann.
- [57] L. Rips. *The Psychology of Proof*. MIT Press, Cambridge, MA, 1994.
- [58] Lance Rips. The Psychology of Knights and Knaves. *Cognition*, 31(2):85–116, 1989.
- [59] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River, NJ, 2009. Third edition.

- [60] Peter Schroeder-Heister. Proof-Theoretic Semantics. In Edward Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2012/2018.
- [61] Stewart Shapiro. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford University Press, Oxford, UK, 1991.
- [62] K. Stenning and M. van Lambalgen. *Human Reasoning and Cognitive Science*. MIT Press, Cambridge, MA, 2008.
- [63] Patrick Suppes. *Axiomatic Set Theory*. Dover Publications, New York, NY, 1972.
- [64] K. VanLehn, R. Jones, and M. Chi. A Model of the Self-Explanation Effect. *Journal of the Learning Sciences*, 2(1):1–60, 1992.
- [65] Douglas Walton, Chris Reed, and Fabrizio Macagno. *Argumentation Schemes*. Cambridge University Press, 2008.

## A What is a Cognitive Calculus? And Why is It So Named?

What is a cognitive calculus  $\mathcal{C}$ , and why is it denoted with the two words in question? (We use ‘ $\mathcal{C}$ ’ here as an arbitrary variable ranging over (the uncountably infinite space of) all cognitive calculi). In keeping with the mathematical-logic literature [e.g. 23]<sup>25</sup>, we first take a *logical system*  $\mathcal{L}$  to be a triple  $\langle \mathcal{L}, \mathcal{I}, \mathcal{S} \rangle$  where  $\mathcal{L}$  is a (often) sorted/typed formal language (based therefore on an alphabet and a formal grammar),  $\mathcal{I}$  is a set of natural<sup>26</sup> inference schemata, and  $\mathcal{S}$  is a formal semantics of some sort. For example, the familiar propositional calculus comprises a family of simple logical systems; the same holds for first-order logic; both families are of course at the heart of AI.<sup>27</sup> In the case of both of these families, a frequently included particular inference schema is *modus ponens*, that is

$$\frac{\phi \rightarrow \psi, \phi}{\psi} \quad I_{\text{MP}}$$

And in the case of the latter family, often *universal introduction* is included in a given  $\mathcal{I}$ ; a specification of this inference schema immediately follows.<sup>28</sup>

$$\frac{\phi(a)}{\forall x \phi\left(\frac{a}{x}\right)} \quad I_{\text{UI}}$$

Note that both of the two inference schemata just shown are included in the particular cognitive calculus we use in the present paper for modeling, and as a framework for automated reasoning. Note as well that both  $\mathcal{L}_{\text{PC}}$  and  $\mathcal{L}_1$  are *extensional*, which means essentially that the meaning of any formula  $\phi$  in the relevant languages are given by compositional functions operating solely on the internal components of  $\phi$ . If we for example know that  $\phi$  is FALSE, then we know that the meaning of  $\phi \rightarrow \psi$  is TRUE, for any  $\psi$  in the language, for both of these logical systems.

Moving from the concept of a logical system to that of a cognitive calculus is straightforward, and can be viewed as taking but three steps, to wit:

S1 Expand the language of a logical system to include

- i modal operators that represent one or more mental verbs at the human level standardly covered in human-level cognitive psychology [e.g. see any standard, comprehensive textbook on human-level cognitive psychology, such as 4, 30], and regarded to be so-called “propositional attitudes” that give rise to propositional-attitude-reporting sentences, where these sentences are represented by operator-infused formulae in a cognitive calculus.<sup>29</sup> Such verbs include: *knowing*, *believing*, *deciding*, *perceiving*, *communicating*,<sup>30</sup> *desiring*, and *feeling X* where ‘ $X$ ’ denotes

<sup>25</sup>Note in particular coverage in this excellent work of Lindström’s Theorems, which pertain to the properties of certain logical systems (e.g., completeness).

<sup>26</sup>Hence when the schemata are deductive in nature, we specifically have natural deduction.

<sup>27</sup>As can be confirmed by looking to the main textbooks of the field. E.g. see [59, 47].

<sup>28</sup>The standard provisos apply here to the constant  $a$ .

<sup>29</sup>The attitudes are covered nicely in [54]. Here’s an informative quote from this work:

Propositional attitude reporting sentences concern cognitive relations people bear to propositions. A paradigm example is the sentence ‘Jill believes that Jack broke his crown.’ Arguably, ‘believes,’ ‘hopes,’ and ‘knows’ are propositional attitude verbs and, when followed by a clause that includes a full sentence expressing a proposition (a that-clause) form propositional attitude reporting sentences. [54, ¶1]

<sup>30</sup>Due to lack of space, we leave aside our approach to formal NLP on the basis of proof theory alone. For a truly excellent book on proof-theoretic semantics, including, natural language, we recommend [25].

some emotional state (e.g. possible  $X = \textit{sad}$ , and so on. Note that such verbs break the bounds of extensionality, and hence make any logic that captures them an *intensional* logic.<sup>31</sup> Step S1.i is the reason why we speak of a *cognitive calculus*.

- ii meta-logical expressions (such as that from a set  $\Phi$  of formulae a particular formula  $\phi$  can be proved:  $\Phi \vdash \phi$ ). Hence cognitive calculi are not merely object-level elements of logics, but include meta-logical elements as well. E.g. a cognitive calculus can have a meta-conditional saying that if some provability expression such as  $\Phi \vdash \phi$  holds, then  $\phi$  holds. Step S1.ii is a necessary, preparatory step for S2.

S2 Delete  $\mathcal{S}$ ; if desired, move selected elements of  $\mathcal{S}$  into  $\mathcal{I}$ , which requires casting these elements as inference schemata that employ meta-logical expressions secured by prior step S1.ii. S2 reflects the fact that cognitive calculi have purely *inferential* semantics, and hence are aligned with the tradition of *proof-theoretic semantics* [28, 55, 60]. (In particular, cognitive calculi thus do not employ possible-worlds semantics for modal operators. In possible-worlds approaches, e.g., *knows* doesn't get defined as justified true belief; but as we explained in the paper proper, knowledge in a cognitive calculus holds iff the agent in question believes the known proposition on the strength of a proof or argument.) We might for instance wish to include an inference schema that regimentes the idea that an agent knows that which is provable from what she knows. Step S2 is the reason why we speak of a cognitive *calculus* (instead of e.g. a cognitive *logic*, or cognitive logical system).

S3 Expand  $\mathcal{I}$  as needed to include inference schemata that involve the operators from S1.i. For instance, where  $\mathbf{K}$  is the modal operator for 'knows' and  $\mathbf{B}$  for 'believes,' we might (and in learning *ex nihilo*, for reasons explained in the paper proper, we do) wish to have this inference schema in a given  $\mathcal{C}$ :

$$\frac{\mathbf{K}\phi}{\mathbf{B}\phi} \quad I_{\mathbf{KB}}$$

## A.1 Regarding Related Work

Much could be said about work/systems that are related to cognitive calculi, but sustained treatment of this issue is out of scope in this brief appendix, which is merely meant to supplement the paper coming before it. We will say only a few things, and hope they are at least somewhat enlightening; here goes. The first published, implemented cognitive calculus, a multi-operator modal logic (minus, by definition, and as explained earlier in the present appendix, any conventional semantics) based on multi-sorted first-order logic, can be found in [1, 2]; the second of these publications is a refinement of the first. Implementation at that point was based upon Athena, a recent introduction to which, along with a study of proof methods in computer science, is provided in the excellent [3]. Related work as cited in this earlier work remains relevant over a decade later, and in particular, so-called "BDI logics" [e.g. 56] are related, and we applaud their advent — but as explained in 2008/2009, such logics cover very few propositional attitudes present in adult and neurobiologically normal cognition (e.g. no communication operators, and no emotional states), and are not based on purely inferential semantics. Automated reasoning in the tradition of higher-order logic (HOL) as descended from Frege, and most prominently from Church, which is masterfully chronicled in [5], is obviously

---

<sup>31</sup>This fact is discussed in some detail in [11], and is replete with relevant proofs. As an example, note that the truth or falsity of 'Jones believes that  $\phi$ ' is not determined by the truth or falsity of  $\phi$ , since humans routinely believe that falsehoods hold.

related to cognitive calculi; this is especially true since HOL is now very much on the scene in 21st-century AI [e.g. 7]. In contrast, cognitive calculi, and the automation thereof, are based on commitments guided by the study of human cognition; and as we see it, that cognition for matters formal and extensional is for the most part circumscribed by natural deduction in second-order logic in the complete absence of formal semantics [e.g. consider the raw material in the practice of mathematics that gives rise to the argument and analysis in 61] and in matters literary circumscribed by modal operators mixed with third-order logic [e.g. 15]. Traditionally, in terms of the Frege-to-Church-to... history that HOL has, HOL is extensional; in contrast, cognitive calculi by definition cannot fail to have operators that cover human cognition. The final thing we mention here is that cognitive calculi are not in any way deductive and bivalent or trivalent; they can be infused with uncertainty, and have multiple values [e.g. 31].

## B The Deontic Cognitive Event Calculus

*DC $\mathcal{E}\mathcal{C}$*  is a quantified multi-modal sorted cognitive calculus. A sorted system can be regarded analogous to a typed single-inheritance programming language. We show below some of the important sorts used in *DC $\mathcal{E}\mathcal{C}$* . Among these, the **Agent**, **Action**, and **ActionType** sorts are not native to the standard, extensional event calculus (apparently) first introduced by Kowalski & Sergot [40], and first modified and extended in the cognitive direction by Arkoudas & Bringsjord [2].

Sort	Description
<b>Agent</b>	Human and non-human actors.
<b>Time</b>	The <b>Time</b> type stands for time in the domain. E.g. simple, such as $t_i$ , or complex, such as $birthday(son(jack))$ .
<b>Event</b>	Used for events in the domain.
<b>ActionType</b>	Action types are abstract actions. They are instantiated at particular times by actors. Example: eating.
<b>Action</b>	A subtype of <b>Event</b> for events that occur as actions by agents.
<b>Fluent</b>	Used for representing states of the world in the event calculus.

Note: actions are events that are carried out by an agent. For any action type  $\alpha$  and agent  $a$ , the event corresponding to  $a$  carrying out  $\alpha$  is given by  $action(a, \alpha)$ . For instance if  $\alpha$  is “*running*” and  $a$  is “*Jack*”,  $action(a, \alpha)$  denotes “*Jack is running*”.

### B.1 Syntax

The syntax has two components: a first-order core and a modal system that builds upon this first-order core. The figures below show the syntax and inference schemata of *DC $\mathcal{E}\mathcal{C}$* . The syntax is quantified modal logic. The first-order core of *DC $\mathcal{E}\mathcal{C}$*  is the *event calculus* [52]. Commonly used function and relation symbols of the event calculus are included. Other calculi (e.g. the *situation calculus*) for modeling commonsense and physical reasoning can be easily switched out in-place of the event calculus.

The modal operators present in the calculus include the standard operators for knowledge **K**, belief **B**, desire **D**, intention **I**, etc. The general format of an intensional operator is  $\mathbf{K}(a, t, \phi)$ , which says that agent  $a$  knows at time  $t$  the proposition  $\phi$ . Here  $\phi$  can in turn be any arbitrary formula. Also, note the following modal operators: **P** for perceiving a state, **C** for common knowledge, **S** for agent-to-agent communication and public announcements, **B** for belief, **D** for desire, **I** for intention, and finally and crucially, a dyadic deontic operator **O** that states when an action is obligatory or forbidden for agents. It should be noted that *DC $\mathcal{E}\mathcal{C}$*  is one specimen in a *family* of easily extensible cognitive calculi.

The calculus also includes a dyadic (arity = 2) deontic operator **O**. It is well known that the unary ought in standard deontic logic lead to contradictions. Our dyadic version of the operator blocks the standard list of such contradictions, and beyond.<sup>32</sup>

<sup>32</sup>A overview of this list is given lucidly in [50].



Syntax	
$S ::=$	$\text{Agent} \mid \text{ActionType} \mid \text{Action} \sqsubseteq \text{Event} \mid \text{Moment} \mid \text{Fluent}$
$f ::=$	$\left\{ \begin{array}{l} \text{action} : \text{Agent} \times \text{ActionType} \rightarrow \text{Action} \\ \text{initially} : \text{Fluent} \rightarrow \text{Boolean} \\ \text{holds} : \text{Fluent} \times \text{Moment} \rightarrow \text{Boolean} \\ \text{happens} : \text{Event} \times \text{Moment} \rightarrow \text{Boolean} \\ \text{clipped} : \text{Moment} \times \text{Fluent} \times \text{Moment} \rightarrow \text{Boolean} \\ \text{initiates} : \text{Event} \times \text{Fluent} \times \text{Moment} \rightarrow \text{Boolean} \\ \text{terminates} : \text{Event} \times \text{Fluent} \times \text{Moment} \rightarrow \text{Boolean} \\ \text{prior} : \text{Moment} \times \text{Moment} \rightarrow \text{Boolean} \end{array} \right.$
$t ::=$	$x : S \mid c : S \mid f(t_1, \dots, t_n)$
$\phi ::=$	$\left\{ \begin{array}{l} q : \text{Boolean} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \\ \mathbf{C}(t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \mid \mathbf{B}(a, t, \phi) \\ \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)\text{happens}(\text{action}(a^*, \alpha), t')) \end{array} \right.$

## B.2 Inference Schemata

The figure below shows the inference schemata for *DC $\mathcal{E}\mathcal{C}$* .  $R_{\mathbf{K}}$  and  $R_{\mathbf{B}}$  are inference schemata that let us model idealized agents that have their knowledge and belief closed under the *DC $\mathcal{E}\mathcal{C}$*  proof theory. While normal humans are not deductively closed, this lets us model more closely how deliberate agents such as organizations and more strategic actors reason. (Some dialects of cognitive calculi restrict the number of iterations on intensional operators.)  $R_1$  and  $R_2$  state respectively that it is common knowledge that perception leads to knowledge, and that it is common knowledge that knowledge leads to belief.  $R_3$  lets us expand out common knowledge as unbounded iterated knowledge.  $R_4$  states that knowledge of a proposition implies that the proposition holds.  $R_5$  to  $R_{10}$  provide for a more restricted form of reasoning for propositions that are common knowledge, unlike propositions that are known or believed.  $R_{12}$  states that if an agent  $s$  communicates a proposition  $\phi$  to  $h$ , then  $h$  believes that  $s$  believes  $\phi$ .  $R_{14}$  dictates how obligations get translated into intentions.

## Inference Schemata

$$\begin{array}{c}
\frac{\mathbf{K}(a, t_1, \Gamma), \Gamma \vdash \phi, t_1 \leq t_2}{\mathbf{K}(a, t_2, \phi)} [R_{\mathbf{K}}] \quad \frac{\mathbf{B}(a, t_1, \Gamma), \Gamma \vdash \phi, t_1 \leq t_2}{\mathbf{B}(a, t_2, \phi)} [R_{\mathbf{B}}] \\
\frac{}{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \Rightarrow \mathbf{K}(a, t, \phi))} [R_1] \quad \frac{}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \Rightarrow \mathbf{B}(a, t, \phi))} [R_2] \\
\frac{\mathbf{C}(t, \phi) \ t \leq t_1 \dots t \leq t_n}{\mathbf{K}(a_1, t_1, \dots \mathbf{K}(a_n, t_n, \phi) \dots)} [R_3] \quad \frac{\mathbf{K}(a, t, \phi)}{\phi} [R_4] \\
\frac{}{\mathbf{C}(t, \mathbf{K}(a, t_1, \phi_1 \Rightarrow \phi_2)) \Rightarrow \mathbf{K}(a, t_2, \phi_1) \Rightarrow \mathbf{K}(a, t_3, \phi_2)} [R_5] \\
\frac{}{\mathbf{C}(t, \mathbf{B}(a, t_1, \phi_1 \Rightarrow \phi_2)) \Rightarrow \mathbf{B}(a, t_2, \phi_1) \Rightarrow \mathbf{B}(a, t_3, \phi_2)} [R_6] \\
\frac{}{\mathbf{C}(t, \mathbf{C}(t_1, \phi_1 \Rightarrow \phi_2)) \Rightarrow \mathbf{C}(t_2, \phi_1) \Rightarrow \mathbf{C}(t_3, \phi_2)} [R_7] \\
\frac{}{\mathbf{C}(t, \forall x. \phi \Rightarrow \phi[x \mapsto t])} [R_8] \quad \frac{}{\mathbf{C}(t, \phi_1 \Leftrightarrow \phi_2 \Rightarrow \neg\phi_2 \Rightarrow \neg\phi_1)} [R_9] \\
\frac{}{\mathbf{C}(t, [\phi_1 \wedge \dots \wedge \phi_n \Rightarrow \phi] \Rightarrow [\phi_1 \Rightarrow \dots \Rightarrow \phi_n \Rightarrow \psi])} [R_{10}] \\
\frac{\mathbf{S}(s, h, t, \phi)}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} [R_{12}] \quad \frac{\mathbf{I}(a, t, \text{happens}(\text{action}(a^*, \alpha), t'))}{\mathbf{P}(a, t, \text{happens}(\text{action}(a^*, \alpha), t))} [R_{13}] \\
\frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} [R_{14}]
\end{array}$$