Embracing Inference as Action: A Step Towards Human-Level Reasoning

John Licato[(✉)] and Maxwell Fowler

Department of Computer Science,
Analogical Constructivism and Reasoning Lab (ACoRL),
Indiana University/Purdue University - Fort Wayne (IPFW),
Fort Wayne, IN 46835, USA
licatoj@ipfw.edu

Abstract. Human-level AI involves the ability to reason about the beliefs of other agents, even when those other agents have reasoning styles that may be very different than the AI's. The ability to carry out reasonable inferences in such situations, as well as in situations where an agent must reason about the beliefs of another agent's beliefs about yet another agent, is under-studied. We show how such reasoning can be carried out in a new variant of the cognitive event calculus we call $\mathcal{CEC}_{\mathtt{AC}}$, by introducing several new powerful features for automated reasoning: First, the implementation of classical logic at the "system-level" and nonclassical logics at the "belief-level"; Second, $\mathcal{CEC}_{\mathtt{AC}}$ treats all inferences made by agents as *actions*. This opens the door for two more additional features: *epistemic boxes*, which are a sort of frame in which the reasoning of an individual agent can be simulated, and *evaluated codelets*, which allow our reasoner to carry out operations beyond the limits of many current systems. We explain how these features are achieved and implemented in the MATR reasoning system, and discuss their consequences.

Imagine, in the not-too-distant future, an artificially-general-intelligent robot, r, is in a room with two humans: its master, m, (to whom r is loyal), and an opponent, o. The robot believes it is essential, for the survival of itself and its master, that m understands that q is true, but that o neither learn nor come to believe that q holds (possibly because o may react negatively and attempt to kill m and destroy r). Because all three are in a small room where anything r says is heard by both o and m (r's telepathic link with m has been damaged), r must somehow say something that will cause m to believe q, but not allow o to figure it out. What can r possibly say?

A human might look at this problem and conclude "r should say some p that m would figure out implies q (due to some beliefs that m already has), but which o would not figure out (since o does not have those same beliefs)." But the AI problem of figuring out that some given p would satisfy this criteria turns out to be non-trivial as we approach situations that are increasingly realistic. For example, assuming that r has to perform some sort of simulation of what o and m would infer given certain beliefs, what if the expressivity of the language in

B. Steunebrink et al. (Eds.): AGI 2016, LNAI 9782, pp. 192–201, 2016.
DOI: 10.1007/978-3-319-41649-6_19

which r reasons is lower than that of the languages in which o and m reason? In such a case, any attempt by r to predict the inferences o and m would produce in response to learning p would be disastrous.

Furthermore, contemporary approaches to doxastic reasoning assume all agents are perfect logical reasoners who believe the logical closures of their belief sets are problematic. What if r reasons non-monotonically and o or m do not, or vice versa? What if r reasons according to classical logic, but the humans do not? r's ability to simulate the reasoning of o and m is limited if r erroneously believes all three follow the same set of inference rules.

Finally, if an artificial agent a has knowledge of which beliefs and inference rules another agent b has, and the ability to simulate inferences using b's beliefs and rules, it is relatively easy to show that some new inference i can be produced from finite applications of b's beliefs and rules. But due to computational limitations, it is much harder to show that some inference i' does *not* follow from these same beliefs and rules (note this is different from showing that the negation of i follows from b's beliefs and rules). Yet a human, reasoning about the beliefs of another agent b, may be able to at least offer a weak argument that b will not come to believe some i, even if the human's argument relies on inductive rules of inference that are not guaranteed to preserve truth.[1]

These are problems that might be faced by AGIs of the future, and understanding how to address them may be necessary to move towards human-level AI. In this paper, we present a way to carry out the sort of reasoning described in the robot-human situation above; this reasoning is based on a new variant of the cognitive event calculus, presented here for the first time, which we will call $\mathcal{CEC}_{\texttt{AC}}$. The $\mathcal{CEC}_{\texttt{AC}}$ formalism embraces several principles, two of which we will describe here: First, the idea that all inferences should be treated as actions; and second, the idea that classical logic (insofar as first-order modal logic can be considered 'classical') should be the formalism at the system-level, and non-classical logics should be the default formalisms at the belief-level. We will show how a new type of automated reasoner (called MATR) can meet the unique mechanical needs of $\mathcal{CEC}_{\texttt{AC}}$, and close with a discussion of limitations and future work.

1 $\mathcal{CEC}$ and $\mathcal{CEC}_{\texttt{AC}}$

The cognitive event calculus ($\mathcal{CEC}$) is a framework based on a first-order modal logic [1], thus extending the event calculus formalism [5]. $\mathcal{CEC}$ contains modal operators allowing the expression of beliefs ($\mathbf{B}(a, t, \phi)$ means an agent a believes ϕ at time t), knowledge ($\mathbf{K}$), intentions ($\mathbf{I}$), and more. $\mathcal{CEC}$ is part of a family of cognitive calculi [3], and this paper presents a new member of this family: $\mathcal{CEC}_{\texttt{AC}}$, the Analogical Constructivism variant of $\mathcal{CEC}$. In a sense, $\mathcal{CEC}_{\texttt{AC}}$ is meant to be an *experimental* formalism, one that regularly plays with the borders of

[1] Imagine, for example, coming up with an argument supporting the statement "no politician will ever say anything that isn't self-serving." Your generated argument (for most people) likely consists of a chain of inferences of inductive strength, rather than a proof that has the weight of full deductive validity.

Multi-Variables		Literals	
$a, b, ...$ - agents	$\mathcal{R}, \mathcal{R}_i$ - rules	p, q - literals	t_1, t_2, t_3 - times
α, β - actions	$\mathfrak{B}, \mathfrak{B}'$ - belief sets	r, o, m - agents	$\mathfrak{B}^i_j, \mathfrak{B}^a_{rel}$ - belief sets
t, t' - time	$\mathfrak{R}, \mathfrak{R}'$ - rule sets	r - rule	$\mathfrak{R}^i_j, \mathfrak{R}^a_{rel}$ - rule sets
ϕ, ψ - formulae			

$$\frac{\text{"}\forall_{\mathcal{R}_G(\{\mathcal{R}_1,...,\mathcal{R}_k\}-\{\mathcal{R}_1,...,\mathcal{R}_k\})} (\{\phi_1,...,\phi_j\} - \{\phi_{j+1},...,\phi_{j2}\}) \nvdash_{\mathcal{R}} \psi\text{"}}{\neg isAffordance(infers(a,\psi),t)}[\underline{\texttt{doesntfollow}}]$$

$$\frac{inSet(\phi_1,\mathfrak{B}),...,inSet(\phi_n,\mathfrak{B}), hasRule(a,t,\mathcal{R}), \text{"}\mathfrak{B} \vdash_{\mathcal{R}} \psi\text{"}}{isAffordance(infers(a,\psi),t)}[\underline{\texttt{follows}}]$$

$$\frac{\text{"}\phi_1,...,\phi_n,\mathcal{R}_1,...,\mathcal{R}_m \text{ are possibly relevant to } \psi\text{"}}{relevant(\psi,\mathfrak{B},\mathbf{R}) \land inSet(\phi_1,\mathfrak{B}) \land ... \land inSet(\phi_n,\mathfrak{B}) \land inSet(\mathcal{R}_1,\mathbf{R}) \land ... \land inSet(\mathcal{R}_m,\mathbf{R})}[\underline{\texttt{relevant}}]$$

$$\frac{isAffordance(infers(a,\psi),t), isOfInterest(a,t,\psi)}{happens(infers(a,\psi),t)}[ind]$$

$$\frac{\neg isAffordance(\alpha,t)}{\neg happens(\alpha,t)}[nInfer]$$

$$\frac{inSet(\mathcal{R},\mathfrak{R}), isFullRuleSet(a,t,\mathfrak{R})}{hasRule(a,t,\mathcal{R})}[defRSet(a)]$$

$$\frac{hasRule(a,t,\mathcal{R}), isFullRuleSet(a,t,\mathfrak{R})}{inSet(\mathcal{R},\mathfrak{R})}[defRSet(b)]$$

$$\frac{inSet(\phi,\mathfrak{B}), isFullBeliefSet(a,t,\mathfrak{B})}{\mathbf{B}(a,t,\phi)}[defBSet(a)]$$

$$\frac{\mathbf{B}(a,t,\phi), isFullBeliefSet(a,t,\mathfrak{B})}{inSet(\phi,\mathfrak{B})}[defBSet(b)]$$

$$\frac{areClose(t,t'), \mathbf{B}(a,t,\phi)}{\mathbf{B}(a,t',\phi)}[bPersist(a)]$$

$$\frac{areClose(t,t'), \neg\mathbf{B}(a,t,\phi)}{\neg\mathbf{B}(a,t',\phi)}[bPersist(b)]$$

$$\frac{areClose(t,t'), isFullBeliefSet(a,t,\mathfrak{B}), isFullBeliefSet(a,t',\mathfrak{B}'), inSet(\phi,\mathfrak{B})}{inSet(\phi,\mathfrak{B}')}[bPersist(c)]$$

$$\frac{areClose(t,t'), isFullBeliefSet(a,t,\mathfrak{B}), isFullBeliefSet(a,t',\mathfrak{B}'), \neg inSet(\phi,\mathfrak{B})}{\neg inSet(\phi,\mathfrak{B}')}[bPersist(d)]$$

$$\frac{areClose(t,t'), hasRule(a,t,\mathcal{R})}{hasRule(a,t',\mathcal{R})}[rPersist(a)]$$

$$\frac{areClose(t,t'), \neg hasRule(a,t,\mathcal{R})}{\neg hasRule(a,t',\mathcal{R})}[rPersist(b)]$$

$$\frac{areClose(t,t'), isFullRuleSet(a,t,\mathfrak{R}), isFullRuleSet(a,t',\mathfrak{R}'), inSet(\mathcal{R},\mathfrak{R})}{inSet(\mathcal{R},\mathfrak{R}')}[rPersist(c)]$$

$$\frac{areClose(t,t'), isFullRuleSet(a,t,\mathfrak{R}), isFullRuleSet(a,t',\mathfrak{R}'), \neg inSet(\mathcal{R},\mathfrak{R})}{\neg inSet(\mathcal{R},\mathfrak{R}')}[rPersist(d)]$$

$$\frac{happens(says(r,p),t)}{\mathbf{B}(m,t,p) \land \mathbf{B}(o,t,p)}[rSay]$$

Fig. 1. Inference rules used for r's situation

traditional cognitive formalisms, e.g. by the use of nonclassical logics, as will be shown next.

Classical Outside, Nonclassical Inside. Human reasoners quite often find themselves facing contradictory beliefs, even when they abide primarily by generally accepted principles of rationality [10]. However, many logical frameworks based on classical logic have some form of the rule known as "*ex contradictione quodlibet*" (ECQ), also known as the principle of explosion. ECQ (broadly summarized) allows any valid formula to follow from any contradition, so for instance if a reasoner believes that "it's true that pizza tastes good, and it's not true that pizza tastes good," it follows that "the moon is made of blue cheese."

Nonclassical logics [9] try to address these weaknesses of classical logics, often by disallowing inference rules that are responsible for the logic's undesirable behavior. Paraconsistent logics, for example, might disallow ECQ and the law of non-contradiction (for any formula, it is not true that both the formula and its negation are true). Such nonclassical logics, however, come with a set of trade-offs — in paraconsistent logics, meta-theoretical properties may be more difficult to prove, but modeling the aspects of human-level reasoning that seem to embrace contradictory beliefs becomes easier.

Our solution, implemented in a variant of CEC that we will call $\mathcal{CEC}_{\mathtt{AC}}$, makes use of classical logics at the "system-level" (the set of formulae that are not nested inside of doxastic operators), but nonclassical logics at the "belief-level" (the level of formulae nested inside of operators such as $\mathbf{B}$, $\mathbf{K}$, etc.). It may thus be acceptable to conclude $\mathbf{B}(a, t, \phi \wedge \neg\phi)$, but $\phi \wedge \neg\phi$ is highly problematic.

Inferences Are Always Actions. The rich formal machinery of the event calculus allows $\mathcal{CEC}_{\mathtt{AC}}$ to deeply embrace the idea that inferences are a type of action, a move in step with analogical constructivism [6] (hence the `AC` subscript). In the event calculus, actions typically consist of a formula of one first-order language reified into the term of another first-order language [8]. Here, we define an inference as an action $\mathit{infers}(a, \phi)$ where a is an agent and ϕ is the formula that a infers. If the inference occurs at time t, it is written as $\mathit{happens}(\mathit{infers}(a, \phi), t)$.

We define an agent's *affordance set* as the set of possible actions (or *affordances*, [4]) an agent can take at some given time. If an agent a has the ability to infer ϕ at time t, then $\mathit{isAffordance}(\mathit{infers}(a, \phi), t)$. Affordances allow us to describe agents that do not automatically create new beliefs simply because those beliefs follow logically from their current set of beliefs. Rather, the act of belief creation is something that happens at a point in time, depending on the set of beliefs an agent has, and the set of *rules* the agent acts in accordance with. Inference rules are presented in Fig. 1, with antecedent conditions on the top of the horizontal lines and conclusions below, with rule names on the right sides.

Using affordances, we can distinguish between three forms of inference rules: first-person, automatic, and affording. Consider the rules:

$$\frac{\phi, \phi \rightarrow \psi}{\psi}[R_1] \qquad\qquad \frac{\mathbf{B}(a, t, \phi), \mathbf{B}(a, t, \phi \rightarrow \psi)}{\mathbf{B}(a, t, \psi)}[R_2]$$

$$\frac{\mathbf{B}(a, t, \phi), \mathbf{B}(a, t, \phi \rightarrow \psi)}{\mathit{isAffordance}(\mathit{infers}(a, \psi), t)}[R_3]$$

The form of inference rule used in R_1, which we will call the *first-person* form, may be used when reasoning about the possible inferences that some given agent may make at some given time. It is often convenient to use this rule form when doxastic and temporal features are not relevant, or are assumed.

The form of inference rule denoted R_2 is probably the most common in doxastic logics. But it is not always applicable in every situation. If agent a has the beliefs at time t that ϕ and $\phi \rightarrow \psi$ hold, it does not always follow that a always has the belief ψ at time t as well. It may be that a simply did not consider the two beliefs simultaneously, or that a never got around to considering the full implications of her beliefs (perhaps due to computational and temporal limitations). Or, depending on the definition of belief adopted, it may be that a is not fully aware of her beliefs, at least to a degree where she can represent them explicitly and use them to produce principled inferences. Furthermore, the application of rule R_2 is done silently; it does not create an *infers* event, nor does it produce any awareness in a's mind that rule R_2, as opposed to some other rule, was used to produce a new belief. The form of inference rule used in R_2 will be called *automatic*.

The *affording* rule form, demonstrated in R_3, instead treats $\phi \rightarrow \psi$ as a possible inference, one which may or may not be made by a. Note that there is some room here for additional rules to specify precisely how inference affordances achieve fruition, i.e., how a possible inference becomes an actual belief. For this paper, one such proposed mechanism draws on the concept of *interest*: if ψ is of interest to a, and is also a possible inference in a's affordances, we can safely infer (with some inductive strength) that $infers(a, \psi)$ will happen.

Belief and Rule Sets. Because $\mathcal{CEC}_{\texttt{AC}}$ is a first-order modal language, we can not easily use higher-order constructs to represent formulae quantifying over sets. However, there are cases where an agent, reasoning about another, may need to reason about sets of beliefs or sets of rules in order to reach conclusions in a human-plausible way. To address this problem, rather than adopting full-blown second-order logic, four sorts are introduced: *setSymbol*, its two sub-sorts *beliefSet* and *ruleSet*, and *ruleSymbol*. The predicate *inSet* corresponds to the standard set-membership operator. A symbol of sort *ruleSymbol* is introduced for every possible inference rule. If an agent a believes $isFullBeliefSet(b, t, \mathfrak{B})$, then agent a believes $inSet(\phi, \mathfrak{B})$ if and only if a believes $\mathbf{B}(a, t, \phi)$. Likewise, if a believes $isFullRuleSet(b, t, \mathfrak{R})$, then agent a believes $inSet(r, \mathfrak{R})$ if and only if a believes $hasRule(b, t, r)$.

Through these sorts and symbols, an agent a can reason about the rules another agent b follows to produce its inferences, or reason about b's beliefs *as a group*, without explicit reasoning about every possible belief a believes b has.

Inductive Inferences. In our toy example, r has no real basis to conclude with certainty that saying p will lead to the effects r desires. That requires a perfect knowledge of all of m and o's beliefs, rules, and some knowledge of all possible confounding factors. If what we are after is human-level reasoning, then, we must allow for inferences that do not necessarily have deductive validity, which do not require a perfect knowledge of all beliefs, rules, and possible events, and furthermore, which shies away from requiring exhaustive calculations of all possibilities, unless such calculations are normally performed by a commonsense human reasoner.

If we are after the ability to generate plausible, reasonable arguments that a human would create or accept upon hearing, then we need to adopt inductive inferences, at least at the belief-level. For this paper, we assume that r has the *rPersist* group of rules, such as $rPersist(a)$ (Fig. 1), presented in its first-person form. $rPersist(a)$ can be interpreted as saying if two time points t and t' are very close together, the fact that agent a acts in accordance with rule $\mathcal{R}$ does not change between t and t'. Such a rule, in the absence of other relevant information, seems *reasonable* to assume. However, crucially, treating $rPersist(a)$

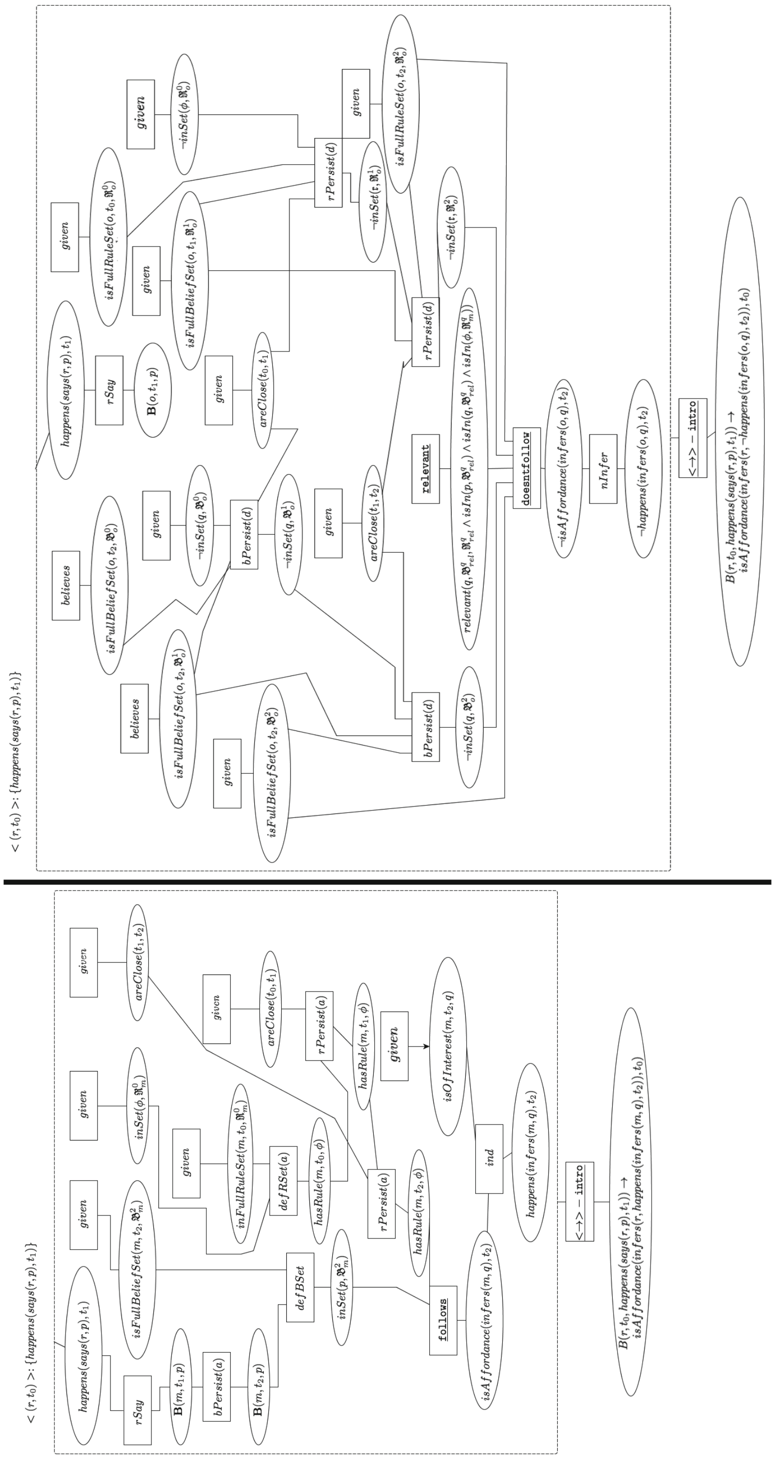

Fig. 2. Proof 1 (Left) and Proof 2 (Right)

as a deductive, truth-preserving rule will lead to inconsistent beliefs (reminiscent of the classic paradox of "Theseus's ship"). $rPersist(a)$ is a rule that cannot appear in its first-person form at the system-level — unless it appears within an *epistemic box*, which we will explain next.

1.1 Boxes

The proof theory we are using is an extension of the one used by MATR (Machina Arachne Tree-based Reasoner),[2] containing *boxes*, a construct similar to the indented subproofs in Fitch-style natural deduction [2]. Formulae inside of a box are inside of the box's *context*, and the box itself is always within some parent context (except for the root context, which is a special box with no parent). Boxes contain a supposition set (written in curly brackets above the box), a list of formulae assumed to hold for the context inside of the box. Boxes themselves can be used as antecedents of inferences. Formulae that are in the same context as the boxes can be re-introduced inside of the box's context; any formulae reintroduced in this way are part of the box's *reiteration set*.

One of $\mathcal{CEC}$'s strengths is that it allows for the arbitrarily deep nesting of beliefs in an unambiguous way, so that an agent's beliefs about another agent's beliefs won't be confused with the first agent's own. But the notation for such situations can become somewhat cumbersome, especially considering affordances require all inferences made by agents to first enter the pool of possible actions.

To address these concerns, we introduce the concept of *epistemic boxes*, boxes that should be thought of as simulating the inferences of some particular agent at a given time. They currently come in two types. $[(a,t)]$-boxes are designed to show what sequence of inferences *will* happen for agent a at time t. Because we will not be using $[(a,t)]$-boxes in this paper, we will not discuss them further. $<(a,t)>$-boxes are designed to show a sequence of inferences that are *possible* for agent a at time t. For any $<(a,t)>$-box $\mathbb{B}$, with parent box $\mathbb{P}$, the following must hold:

- Formulae ϕ can be in $\mathbb{B}$'s reiteration set only if $\mathbf{B}(a,t,\phi)$ holds in $\mathbb{P}$'s context.
- If formulae ϕ_1 and ϕ_2 lead to an inference of ψ (where all three formulae are inside $\mathbb{B}$), then the formulae $\mathbf{B}(a,t,\phi_1)$, $\mathbf{B}(a,t,\phi_2)$, and $isAffordance(infers(a,\psi),t)$ must all hold at $\mathbb{P}$'s context.
- If $\frac{\phi_1,\ldots,\phi_n}{\psi}$ is an inference rule that holds within $\mathbb{B}$, then the rule $\frac{\mathbf{B}(a,t,\phi_1),\ldots,\mathbf{B}(a,t,\phi_n)}{isAffordance(infers(a,\psi),t)}$ must hold at $\mathbb{P}$'s context.

The only inference rules that hold in an $<(a,t)>$-box are those that, when transformed from the first-person to affording forms, hold in the context of the box's parent. If an $<(a,t)>$-box $\mathbb{B}$, with supposition set $\{\gamma_1,\ldots,\gamma_n\}$, produces the inference ψ within $\mathbb{B}$, then in $\mathbb{B}$'s parent context we can infer:

[2] MATR is currently being developed by a collaboration between the Rensselaer AI and Reasoning (RAIR) Lab, and the Analogical Constructivism and Reasoning Lab (ACoRL).

$$(\mathbf{B}(a,t,\gamma_1) \wedge ... \wedge \mathbf{B}(a,t,\gamma_n)) \rightarrow \mathit{isAffordance}(\mathit{infers}(a,\psi),t)$$

This inference is done using the evaluated inference rule $\underline{\texttt{<<>>-intro}}$ (examples of its use are seen in Fig. 2). However, because specifying the supposition set in its entirety again requires a higher-order logic, actually implementing this rule requires a special computational ability, which we will describe next.

1.2 Evaluated Codelets

MATR-style reasoning implements its reasoning by dividing most of the hard work amongst *codelets* (a term borrowed from the *Copycat* model [7]), which can be thought of as little independently-operating processes. Each codelet looks at the current state of a proof and make suggestions to the central proof manager (the Codelet Manager, or CM) about what inferences should be made. Generally, one codelet is created for each inference rule, so that the codelet can specialize in efficient algorithms for locating appropriate areas of the current proof state to make suggestions. Codelets can be small and quick (like the codelets that look at the current state of a proof and suggest ways to shorten it), or large and possibly slow (such as codelets that serve as wrappers for full automated theorem provers).

Codelets, implemented as Java programs, can theoretically run any program and use any criteria to evaluate whether an inference should be suggested. The CM simply assumes that if a codelet suggests an inference, that inference is valid according to the semantics of the codelet. Thus, we can create codelets capable of implementing the sort of low-level, distributed algorithms common in state-of-the-art artificial intelligence, but lacking from automated theorem provers. We call these *evaluated codelets*—these are codelets that implement inference rules whose conditions contain semantics not entirely captured in the formal language of the antecedents. Instead, we specify these semantics in pseudo-formalized natural language statements.

Inference rules relying on evaluated codelets are referred to as *evaluated rules.* For example, consider the $\underline{\texttt{follows}}$ inference rule in Fig. 1 (note the use of teletype font and an underline to denote evaluated codelets), which can be invoked if some ϕ follows from an agent's belief set and an inference rule that agent has. In such evaluated rules, we can accept an arbitrary number of conditions in the antecedent (denoted by the ellipsis) and any arbitrary requirement in the quotation marks; often we will use natural language statements. It is up to the codelet to implement the semantics expressed in the quotation marks faithfully.[3]

Evaluated codelets are particularly useful when a conclusion should be inferred on the basis of something not captured in the logical form of the $\mathcal{CEC}_{\texttt{AC}}$ formulae alone. This is why the inference rule $\underline{\texttt{<<>>-intro}}$ is implemented as

[3] Of course, we acknowledge such a violation of referential transparency means that actual semantics of evaluated codelets may vary, possibly significantly, between different implementations, and provide a level of (possibly dangerous) flexibility not seen in any other automated reasoners (to our knowledge).

an evaluated codelet. As another example, if the analogical similarity between two formulae needs to be found, we might use a structural comparison of the formulae themselves. Although the method for comparison can be done through logical syntax alone, actually performing that structural comparison is best done with the use of a software package that may draw on distributed representations, vector operations, machine learning methods, etc. Thus, evaluated codelets allow MATR to combine the high-level reasoning and argument-generation powers of logic-based AI with the amazing advancements in nonlogical methods that have been dominating the field of machine learning in recent years.

As of now, there are several restrictions placed on evaluated codelets. Two instances of an implementation of an evaluated codelet, within the same context, given the same formulae as antecedents, must produce the same inference (if they produce any inference under those conditions). They can behave nonmonotonically (changing their inferences if more formulae are given as antecedents), but they may not behave randomly. Furthermore, certain evaluated codelets are only allowed within the context of epistemic boxes. For example, a codelet which tries to simulate the low-level similarity-based inferences of an agent may need to draw on agent-specific knowledge.

In this paper, we make use of four evaluated codelets: The first, $\underline{<\!\!\rightarrow\!\!>\texttt{-intro}}$ has already been described:

$$\frac{\text{“ } <(a,t)> \text{ -box } \mathbb{B}\text{‘s full supposition set is } \phi_1, \ldots, \phi_n \text{ and } \psi \text{ is inferred in } \mathbb{B}''}{(\mathbf{B}(a,t,\phi_1) \land \ldots \land \mathbf{B}(a,t,\phi_n)) \rightarrow isAffordance(infers(a,\psi),t)}\left[\underline{<\!\!\rightarrow\!\!>\texttt{-intro}}\right]$$

<u>`follows`</u> and <u>`doesntFollow`</u> determine whether or not a formula follows from another set of formulae, given some rule symbols. Finally, <u>**`relevant`**</u> takes a formulae ψ and uses some similarity-based algorithm to generate a set of formulae and inference rules that *may* be relevant to ψ (we do not specify here what algorithm should be used to determine relevance).

We present our proofs in Fig. 2, using MATR's visual style. r can possibly infer that if he says p, then m will infer q (Proof 1), while o will not (Proof 2). Assumptions made are denoted with the *given* codelet.

2 Future Work

The new $\mathcal{CEC}_{\texttt{AC}}$ features introduced in this paper are serious steps towards the creation of a new style of automated reasoning that can bridge the gap between informal and formal reasoning. However, it is clear that a lot more work is needed. Perhaps most notably lacking from the present work is a way to compare confidence in inductively-inferred beliefs, so that stronger beliefs can defeat the weaker ones. For example, it is easy to see how we can actually derive, with low confidence, an argument for $\mathbf{B}(r, t_0, \neg inSet(\psi, B^2_m))$, which would conflict with the conclusion of Proof 1. A satisfactory model of formula defeating would need to somehow assign a higher confidence to $\mathbf{B}(r, t_0, inSet(\psi, B^2_m))$.

We also don't discuss here how an agent might resolve between contradictory beliefs when a plan of action depends on such a resolution. This may require introducing a concept of "conscious acceptance", which we are currently developing.

Finally, we don't claim that our approach is the only way to do these things, nor that our way is always better than all alternatives (e.g. treating action inferences as a relation between possible worlds). For example, although our concept of action-as-inference was primarily motivated by the later work of Jean Piaget [6], new promising work based on the concept of *proof-events* is emerging [11], which considers proving as a process that unfolds through time. It is outside of the scope of this paper to fully compare all alternatives. Certainly, the idea of introducing inference rules whose semantics are implementation-specific may seem unusual to those firmly steeped in the logic-based AI tradition, but it is our hope that such measures will allow a flexibility not currently enjoyed by artificial reasoners.[4]

References

1. Arkoudas, K., Bringsjord, S.: Propositional attitudes and causation. Int. J. Softw. Inf. **3**(1), 47–65 (2009)
2. Barker-Plummer, D., Barwise, J., Etchemendy, J.: Language Proof and Logic. Center for the Study of Language and Inf, 2 edn., October 2011
3. Bringsjord, S., Govindarajulu, N.S., Licato, J., Sen, A., Johnson, J., Bringsjord, A., Taylor, J.: On logicist agent-based economics. In: Proceedings of Artificial Economics 2015 (AE 2015). University of Porto, Porto, Portugal (2015)
4. Gibson, J.: The theory of affordances. In: Shaw, R., Bransford, J. (eds.) Perceiving, Acting, and Knowing: Toward an Ecological Psychology. Erlbaum, Hillsdale, NJ (1977)
5. Kowalski, R., Sergot, M.: A logic-based calculus of events. New Gener. Comput. **4**(1), 67–94 (1986)
6. Licato, J.: Analogical Constructivism: The Emergence of Reasoning Through Analogy and Action Schemas. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy, NY, May 2015
7. Mitchell, M.: Analogy-Making as Perception: A Computer Model. The MIT Press, Cambridge, Massachusetts (1993)
8. Mueller, E.T.: Commonsense Reasoning: An Event Calculus Based Approach, 2nd edn. Morgan Kaufmann, Waltham (2015)
9. Priest, G.: An Introduction to Non-Classical Logic: From If to Is, 2nd edn. Cambridge University Press, Cambridge (2008)
10. Priest, G., Tanaka, K., Weber, Z.: Paraconsistent logic. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. spring 2015th edn. (2015). http://plato.stanford.edu/archives/spr2015/entries/logic-paraconsistent
11. Stefaneas, P., Vandoulakis, I.M.: On mathematical proving. J. Artif. Gen. Intell. **6**(1), 130–149 (2015)

[4] Dr. Licato, and the research presented here, was funded by the AFOSR Young Investigator's Program (YIP).