Computational Context Logic and Species of ist

$[Revision{:}1.16]$

Selene Makarios Karl Heuer Richard Fikes

Abstract

Mapping context-logical theories to equivalent first-order theories allows them to be processed using existing first-order automated reasoners. Performing a "context-shift transform" of domain objects makes such mapping possible, at least for certain "species of ist", where speciation of ist is defined by distributivity of ist over the logical connectives and quantifiers.

1 Introduction

Research within Artificial Intelligence into the formalization of the human notion of *context*, a versatile, efficacious, and seemingly essential cognitive tool, has produced many different logics ([3], [8], [2]), to deal with a number of motivating examples ([6], [3], [7]).

Originally, the notion of *formalized contexts* was proposed by McCarthy [6] as a way of addressing the problem of generality in AI. The original idea was to create AI systems that never get "stuck" with the set of concepts they are using at any given moment, always being able to *transcend* any given context to a more general one.

Also, as suggested by Guha [3], contexts may offer assistance with the *qualification problem* [5]. This problem arises due to the fact that a great many useful propositions about the real world (e.g. "birds fly") tacitly assume the lack of certain *exceptional conditions*. If an artificial agent is to reason correctly in the real world, it must account for the possibility of those exceptions, but do so without requiring an explicit assertion about the status of every possible exceptional condition, that is, about every possible qualification on the truth of given proposition.

Another perspective is that, for the sake of tractability, contexts allow humans to temporarily fix a domain of discourse and a set of interpretations, for purposes of performing some cognitive or communicative task. Without this ability, that is, if humans were required to always account for all entities, relationships, and the interpretations thereof, while performing a cognitive or communicative task, all such tasks would likely be unmanageable.

Logic has been used in mathematics to state universal truths. Unlike in AI, the goal is not automated reasoning, but distilling a kernel, such as Peano's axioms, from which various mathematical theorems follow. For such a mathematical effort, it seems important that the meaning of these statements not depend on anything other than the logic in which they are stated. In particular, the circumstances of their assertion should not impact upon their meaning. Consequently, traditional logic has expressly avoided contextuality, though, some logicians and philosophers have started developing kinds of logics [1] which explicitly account for situations, which are a type of context.

In contrast, human communication exploits the situation or context of the communication, often to an extreme degree, leaving much implicit. Processing of these communications also exploits this context, i.e., we don't *completely* decontextualize what we hear into a global frame of reference before we reason with it. In fact we might argue that a complete decontextualization is not just undesirable, but impossible. We do however have an understanding of the role of the situation in determining the meaning of an utterance. We believe that logical formulas provided to AI programs by people (or sensors) are more akin to human communication than to Peano's axioms. Thus, contexts are likely to be important in the representation of knowledge in AI systems.

1.1 ist Logics

McCarthy's and Guha's approach to formalizing context [3, 7, 4] was to introduce an operator 'ist' , with the intended intuitive meaning that if c is a context, and φ is a proposition, then the expression

$$ist(c, \varphi) \tag{1}$$

is considered true exactly in case the proposition φ is true, *according to the account* of the context c. (The name 'ist' is derived from '*is true*'). As noted by Guha [3], the context object can be thought of as the *reification of all the context dependencies of the propositions associated with the context*.

The notion of "context" itself is taken as a mathematical primitive in the formalism discussed here. Regardless of what motivating conception of context one employs, (some examples will be given below), what matters is that propositions can be asserted with respect to a context, and then other propositions can be derived with respect to that context, or possibly with respect to other contexts that are related to it in certain ways.

1.2 Conceptions of Context

The following are some examples of motivating conceptions of context:

- A constraint upon time, such as *in the Middle Ages*, or *last Tuesday afternoon*;
- A constraint upon place, such as *in the Gobi desert*;
- The context of a given conversation;
- The context of a particular news report;
- The context of a fictitious universe, such as *the world of Sherlock Holmes*;

- The context of a set of beliefs;
- A linguistic context, such as a particular tongue, or a collection of domain-specific verbal efficiencies (e.g. the context in which a surgeon saying "scalpel" is understood as a request to be given the No. 3 scalpel on the tray).

The reduction of context logic theory to mechanized practice appears to be, for the most part, neglected. This is understandable, given the state of affairs of formal semantics for context logic. However, if context-logic theories can be transformed into equivalent first-order theories, then inferencing over these theories is possible using existing first-order automated reasoners. We propose such a transformation system for ist-based context logics, using the notion of a "context-shift transform", herein introduced, and explore the conditions under which context-shift transforms are valid.

2 Context-Shift Transform of Formulas

We formally introduce the concept of the context-shift transform. Consider a first-order language $\mathcal{L}$ with signature $\langle \mathcal{P}, \mathcal{F}, \mathcal{A} \rangle$, with $\mathcal{P}$, $\mathcal{F}$, and $\mathcal{A}$ giving the predicate-symbols, function-symbols, and constant-symbols respectively. We introduce a set of constant-symbols $\mathcal{C}$ that is disjoint with $\mathcal{A}$.

Definition 2.1 (Context-Shift Function)**.** For all predicates $P \in \mathcal{P}$, all functions $f \in \mathcal{F}$, and all contexts $c \in \mathcal{C}$,

$$\begin{gathered} \forall(x_1, ..., x_n)(ist(c, P(x_1, ..., x_n)) \leftrightarrow P(x_1{+}c, ..., x_n{+}c)) \ , \\ \forall(x_1, ..., x_m)(f(x_1, ..., x_m){+}c = f(x_1{+}c, ..., x_m{+}c)) \ . \end{gathered} \tag{2}$$

Definition 2.2 (Context-Shift Transform)**.** Given some formula φ in $\mathcal{L}$ and some c in $\mathcal{C}$, define the "context-shifted formula", $\langle\varphi\rangle_c$ as the *syntactic transformation*:

$$\begin{gathered} \langle \varphi_1 \circ \varphi_2 \rangle_c = \langle \varphi_1 \rangle_c \circ \langle \varphi_2 \rangle_c \ \text{ for} \circ \text{ each of } \{\wedge, \vee, \rightarrow, \leftrightarrow\} \\ \langle \neg\varphi \rangle_c = \neg \langle \varphi \rangle_c \\ \langle \forall x \varphi \rangle_c = \forall x \langle \varphi \rangle_c \quad \langle \exists x \varphi \rangle_c = \exists x \langle \varphi \rangle_c \\ \langle \varphi \rangle_c = \varphi[x_1/x_1{+}c]...[x_m/x_m{+}c][a_1/a_1{+}c]...[a_n/a_n{+}c] \ \text{ for } \varphi \text{ atomic,} \end{gathered} \tag{3}$$

thereby essentially replacing each variable x_i with $x_i{+}c$ and each constant a_i with $a_i{+}c$.

Definition 2.3 (ist-reduction)**.** We define the "ist-reduction" of a sentence $ist(c, \varphi)$, denoted by

$$\lfloor ist(c, \varphi) \rfloor \ , \tag{4}$$

to be the result of recursively distributing $ist(c, \cdot)$ syntactically over the logical connectives and quantifiers in φ, until the sentential argument of every ist is atomic:

$$\begin{gathered}\lfloor ist(c, \varphi_1 \circ \varphi_2) \rfloor = \lfloor ist(c, \varphi_1) \rfloor \circ \lfloor ist(c, \varphi_2) \rfloor \text{ (for } \circ \text{ each of } \{\wedge, \vee, \rightarrow, \leftrightarrow\}) \\ \lfloor ist(c, \neg\varphi) \rfloor = \neg \lfloor ist(c, \varphi) \rfloor \\ \lfloor ist(c, \forall x \varphi) \rfloor = \forall x \lfloor ist(c, \varphi) \rfloor \quad \lfloor ist(c, \exists x \varphi) \rfloor = \exists x \lfloor ist(c, \varphi) \rfloor \\ \lfloor ist(c, \varphi) \rfloor = ist(c, \varphi) \text{ for } \varphi \text{ atomic}\end{gathered} \tag{5}$$

Theorem 2.1 (Context-Shift Transform Theorem). *For arbitrary φ in $\mathcal{L}$, the expression produced by the syntactic operation of ist-reduction, and the expression produced by the syntactic operation of context-shift transform, are logically equivalent, that is,*

$$\lfloor ist(c, \varphi) \rfloor \leftrightarrow \langle \varphi \rangle_c \ . \tag{6}$$

Proof. Proof by induction on the nesting depth of φ, denoted $\|\varphi\|$. If $\|\varphi\| = 0$ then φ is atomic, so $\lfloor ist(c, \varphi) \rfloor$ is the same as $ist(c, \varphi)$ and our result follows from the definitions of $x+c$ and $\langle x \rangle_c$. Now, suppose that $\|\varphi_1 \circ \varphi_2\| = n$, and that $\lfloor ist(c, \varphi_1 \circ \varphi_2) \rfloor$. We know $\lfloor ist(c, \varphi_1 \circ \varphi_2) \rfloor = \lfloor ist(c, \varphi_1) \rfloor \circ \lfloor ist(c, \varphi_2) \rfloor$ by definition of ist-reduction, and since $\|\varphi_1\| = \|\varphi_2\| = n - 1$, we know that $\lfloor ist(c, \varphi_1) \rfloor \leftrightarrow \langle \varphi_1 \rangle_c$ and that $\lfloor ist(c, \varphi_2) \rfloor \leftrightarrow \langle \varphi_2 \rangle_c$ by inductive hypothesis. But this implies that $\langle \varphi_1 \rangle_c \circ \langle \varphi_2 \rangle_c$, and this is equivalent to $\langle \varphi_1 \circ \varphi_2 \rangle_c$ by definition of $\langle \cdot \rangle_c$. Now, suppose that $\|\forall x \varphi\| = n$, and that $\lfloor ist(c, \forall x \varphi) \rfloor$. We know $\lfloor ist(c, \forall x \varphi) \rfloor = \forall x \lfloor ist(c, \varphi) \rfloor$ by definition of ist-reduction, which means that for any $a \in \mathcal{A}$, $\lfloor ist(c, \varphi[x/a]) \rfloor$, and since $\|\varphi[x/a]\| = n - 1$, we know that $\lfloor ist(c, \varphi[x/a]) \rfloor \leftrightarrow \langle \varphi[x/a] \rangle_c$ by inductive hypothesis. But this implies that for any $a \in \mathcal{A}$ we have $\langle \varphi[x/a] \rangle_c$, which means that $\forall x \langle \varphi \rangle_c$, and this is equivalent to $\langle \forall x \varphi \rangle_c$ by definition of $\langle \cdot \rangle_c$. Analogous induction covers the cases for the other connectives and quantifiers. □

Corollary 2.1 (Context-Shift Transform Corollary). *If $ist(c, \varphi)$ is logically equivalent to its own ist-reduction, then it is also logically equivalent to its context-shift transform by c. That is,*

$$ist(c, \varphi) \leftrightarrow \lfloor ist(c, \varphi) \rfloor \quad \textit{implies} \quad ist(c, \varphi) \leftrightarrow \langle \varphi \rangle_c \ . \tag{7}$$

Thus, the issue of the validity of reasoning by context-shift transforms of ist-theories reduces to the issue of the conditions under which

$$ist(c, \varphi) \leftrightarrow \lfloor ist(c, \varphi) \rfloor \ . \tag{8}$$

It turns out that we can speciate ists into those for which (8) is valid for any first-order sentence φ, and those for which validity of (8) requires the satisfaction of certain constraints upon φ.

3 Species of ist

We speciate ists based upon assumptions regarding their distributivities[1]. As mentioned earlier, we assume the following distributivity postulates for any notion of context:

$$\begin{aligned} \text{(IA)} &\vdash\ ist(c, \varphi \wedge \psi) \leftrightarrow ist(c, \varphi) \wedge ist(c, \psi) \\ \text{(IU)} &\vdash\ ist(c, \forall x \varphi[x]) \leftrightarrow \forall(x) ist(c, \varphi[x]) \end{aligned} \tag{9}$$

where the domain of quantification is nonempty. Note that (IU) is *not an analog of the Barcan formula and its converse*. The treatment of the Barcan analogs involves the context logic analogs of the *necessary* and *possible* operators, which will be defined and discussed in another work.

However, the following distributivity postulates may hold or not, depending upon the notion of context that is used:

$$\begin{aligned} \text{(IO)} &\vdash\ ist(c, \varphi \vee \psi) \leftrightarrow (ist(c, \varphi) \vee ist(c, \psi)) \\ \text{(IE)} &\vdash\ ist(c, \exists x \varphi[x]) \leftrightarrow \exists(x) ist(c, \varphi[x]) \\ \text{(IN)} &\vdash\ ist(c, \neg\varphi) \leftrightarrow \neg ist(c, \varphi) \end{aligned} \tag{10}$$

Example – context as time-interval: Suppose that a context is identified with some interval of time t, and we say that $ist(t, \varphi)$ iff the time-sensitive sentence φ – sentences such as which are sometimes call "fluents" – holds during the entire interval t. For example, if φ is the fluent "Fred is asleep", and $Tuesday$ is the time-interval corresponding to the 24-hour period of some particular Tuesday, then $ist(Tuesday, \varphi)$ is the context-logical construct expressing that Fred is asleep during the entire 24-hour period of Tuesday. Then we have $ist(Tuesday, \varphi \vee \neg\varphi)$, that is, any time during Tuesday it is the case that either Fred is asleep or not. This assertion clearly does not imply, however, that Fred must either be asleep throughout all 24 hours of Tuesday, or not-asleep throughout all 24 hours, which would be $ist(Tuesday, \varphi) \vee ist(Tuesday, \neg\varphi)$. So for this notion of context, the postulate (IO) does not hold.

Example – context as closed world: Let a context be given by the material implications of some first-order sentence α, and denote this set of implications by $[\alpha]$, so we define

$$\vdash\ ist([\alpha], \varphi) \leftrightarrow (\alpha \rightarrow \varphi)\ . \tag{11}$$

Then

$$\vdash\ ist([\alpha], \varphi \vee \psi) \leftrightarrow (\alpha \rightarrow (\varphi \vee \psi))\ , \tag{12}$$

$$\vdash\ ist([\alpha], \varphi \vee \psi) \leftrightarrow ((\alpha \rightarrow \varphi) \vee (\alpha \rightarrow \psi))\ , \tag{13}$$

and so

$$\vdash\ ist([\alpha], \varphi \vee \psi) \leftrightarrow (ist([\alpha], \varphi) \vee ist([\alpha], \psi))\ . \tag{14}$$

So for this notion of context, the postulate (IO) holds, which means that for every φ we have either $ist(c, \varphi)$ or $ist(c, \neg\varphi)$. But the postulate (IN) does not: $ist([\alpha], \neg\varphi)$ does

[1] This is analogous to the speciation of modal logics based upon assumptions of various postulates regarding the interaction of modal and classical operators (e.g. $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ is a postulate of K, while K together with $\Box\varphi \rightarrow \varphi$ gives M, and M together with $\Box\varphi \rightarrow \Box\Box\varphi$ gives $S4$, etc.)

not require that $\neg ist([\alpha], \varphi)$, that is, such a context might be internally inconsistent. (This is certainly the case, for example, in the context $[p \wedge \neg p]$ for some proposition p, since for this context both $ist([p \wedge \neg p], p)$ and $ist([p \wedge \neg p], \neg p)$).

Definition 3.1 (ist_A, ist_{AO}, and ist_{AON})**.** We will call the version of ist for which just (IA) and (IU) hold, ist_A. The ist for which (IA), (IU), (IO), and (IE) all hold is called ist_{AO}, and the ist for which all five postulates hold is called ist_{AON}.

Theorem 3.1. *With* T *representing the empty conjunction, and* F *the empty disjunction, neither of*

$$\begin{aligned} ist(c, \mathrm{T}) &= \mathrm{T} \ , \\ ist(c, \mathrm{F}) &= \mathrm{F} \ , \end{aligned} \tag{15}$$

are required to hold for contexts in general, nor are their negations.

Proof. We use the "sandwich contexts", which generalize the closed-world contexts $[\alpha]$. For any two propositions α and β whose disjunction is a tautology, let

$$[\alpha, \beta] \tag{16}$$

denote the context such that

$$ist([\alpha, \beta], \varphi) \leftrightarrow (\neg\alpha \vee \varphi \wedge \beta) \ . \tag{17}$$

(It will turn out that operator precedence in the right-hand side of the bi-implication in (17) won't matter.) Let $\mathcal{C}_A$ be the collection of all contexts of the form $[\alpha, \beta]$ for first order constructs α and β, let $\mathcal{C}_{FA}$ be those contexts of the form $[\mathrm{T}, \beta]$, let $\mathcal{C}_{TA}$ be those contexts of the form $[\alpha, \mathrm{T}]$, and let $\mathcal{C}_{FTA}$ be $\{[\mathrm{T}, \mathrm{T}]\}$. Then it is straightforward to show that any of these collections satisfy (IA), however,

$$\begin{gathered} (\text{ for all } c \in \mathcal{C}_{FTA},\ ist(c, \mathrm{T}) = \mathrm{T}) \text{ and} \\ (\text{ for all } c \in \mathcal{C}_{FTA},\ ist(c, \mathrm{F}) = \mathrm{F}) \end{gathered} \tag{18}$$

and

$$\begin{gathered} (\text{ for all } c \in \mathcal{C}_{TA},\ ist(c, \mathrm{T}) = \mathrm{T}) \text{ and} \\ (\text{ it is not the case that } \quad \text{for all } c \in \mathcal{C}_{TA},\ ist(c, \mathrm{F}) = \mathrm{F}\ ) \end{gathered} \tag{19}$$

and

$$\begin{gathered} (\text{ it is not the case that } \quad \text{for all } c \in \mathcal{C}_{FA},\ ist(c, \mathrm{T}) = \mathrm{T}\ ) \text{ and} \\ (\text{ for all } c \in \mathcal{C}_{FA},\ ist(c, \mathrm{F}) = \mathrm{F}) \end{gathered} \tag{20}$$

and

$$\begin{gathered} (\text{ it is not the case that } \quad \text{for all } c \in \mathcal{C}_{A},\ ist(c, \mathrm{T}) = \mathrm{T}\ ) \text{ and} \\ (\text{ it is not the case that } \quad \text{for all } c \in \mathcal{C}_{A},\ ist(c, \mathrm{F}) = \mathrm{F}\ ) \end{gathered} \tag{21}$$

□

4 Transforming the Various Species

We discuss the constraints under which context-shift transforms of ist theories into first-order theories is valid, for the various species of ist.

4.1 ist_{AON}

Performing the context-shift transform for ist_{AON} is straightforward, since in this case

$$ist(c, \varphi) \leftrightarrow \lfloor ist(c, \varphi) \rfloor \tag{22}$$

for arbitrary first-order sentence φ, which is a direct result of the definition of ist-reduction.

4.2 ist_{AO}

For ist_{AO}, context-shift transform is of course possible when negation appears in none of the sentences in our theory. But significantly, we can *regain the use of negation* for ist_{AO} via a modified form of ist-reduction, using "ist-conjugates". This allows us to use a first-order automated reasoner to perform inferencing over any ist theory merely with the requirements of ist_{AO}, as opposed those of ist_{AON}, thereby significantly broadening the notions of context to which the context-shift transform can be applied.

Definition 4.1 (ist-conjugate)**.** We define the operator *, read "ist-conjugate", that for any context c and any first-order sentence φ, satisfies

$$ist(c^*, \varphi) \leftrightarrow \neg ist(c, \neg\varphi) \ . \tag{23}$$

Under the assumptions of ist_{AO}, it is easy to show that c^* is a legitimate context.

Definition 4.2 (Modified ist-reduction)**.** We define a modified version of ist-reduction, in which a sentence φ is first placed in negation normal form yielding φ', and then the rules of normal ist-reduction – except the rule for negation – are recursively applied to φ' until the sentential argument of every ist is a literal (either an atomic formula or the negation of one). Then, every ist-predicate of the form $ist(c, \neg\psi)$ is replaced with $\neg ist(c^*, \psi)$.

Now, in order to verify that reasoning via context-shift transform is valid under our modified form of ist-reduction, we need to show that when ψ is atomic, $\neg ist(c, \neg\psi)$ is logically equivalent to the context-shift transform of ψ by c^*.

Corollary 4.1 (Context-Conjugate Corollary)**.** *For an atomic sentence ψ,*

$$\neg ist(c, \neg\psi) \leftrightarrow \langle\psi\rangle_{c^*} \tag{24}$$

Proof. By definition of *, $\neg ist(c, \neg\psi)$ is equivalent to $ist(c^*, \psi)$, and the latter expression is equivalent to $\lfloor ist(c^*, \psi) \rfloor$ due to the definition of ist-reduction for the case of ψ atomic. This in turn is equivalent to $\langle\psi\rangle_{c^*}$ by Theorem (2.1). □

4.3 ist_A

Context-shift transform for ist_A is possible when all sentences φ involve only conjunction and universal quantification. In general, we can't regain the use of negation for ist_A via ist-conjugates as we did for ist_{AO}, since the conversion of φ to negation normal form, required by the modified ist-reduction, may introduce disjunctions, and ist cannot distribute over these disjunctions for ist_A.

References

[1] J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, 1983.

[2] Fausto Giunchiglia. Contextual reasoning. Technical Report 9211–20, Istituto per la Ricerca Scientifica e Technoligica, Trento, Italy, 1992.

[3] Ramanathan V. Guha. Contexts: a formalization and some applications. Technical Report STAN-CS-91-1399, Stanford CS Dept., Stanford, CA, 1991.

[4] R.V. Guha. Context dependence of representations in cyc. In *Colloque ICO '93*, 1993.

[5] John McCarthy. Circumscription — a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27–39, 1980.

[6] John McCarthy. Generality in artificial intelligence. *Comm. of ACM*, 30(12):1030–1035, 1987.

[7] John McCarthy. Notes on formalizing contexts. In Ruzena Bajcsy, editor, *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 555–560, San Mateo, California, 1993. Morgan Kaufmann.

[8] P. Pandurang Nayak. Representing multiple theories. In B. Hayes-Roth and R. Korf, editors, *AAAI-94*, pages 1154–1160, Menlo Park, CA, 1994.