# Modelling Assumption-Based Reasoning Using Contexts

Mark Jago*

School of Computer Science & Department of Philosophy
University of Nottingham
Nottingham, UK
mtw@cs.nott.ac.uk

**Abstract.** We propose a two-fold use of context in reasoning about agents. The first concerns the modelling of beliefs which are entertained only within the scope of some assumption. We illustrate this use by describing as logical model of an agent which reasons in a natural deduction style by making assumptions to derive new formulas. The second use of context we propose here concerns the modelling of reasoning as a step-by-step temporal process, allowing us to model agents whose resources are temporally bounded. In such a setting, the beliefs ascribed to an agent need not be closed under consequence, as they are in many traditional epistemic logics.

## 1 Introduction

This essay is concerned with the modelling of agents who make assumptions and use them in their deductive reasoning. This work differs form other formal accounts of contextual reasoning in that it deals with resource bounded agents, namely agents which requires computational resources to deduce new beliefs from old. We take as an illustration an agent who reasons in a natural deduction style, but who takes time to reach its conclusions.

We treat an assumption as a sentence entertained by an agent for a particular purpose. An assumption is not a belief, yet has the psychological effect of a belief whilst the assumption is entertained. In assuming a formula, one reasons and behaves as if it were the case (this is, then, something like the psychological notion of pretence).

We use the notion of context to model this process of making assumptions. Following [4], a context is a localised set of beliefs, connected to other contexts by inter-context *bridge rules*. Each context represents the epistemic consequences of an act of pretence (i.e., of assumption making) on behalf of the agent; for example, the context in which an agent makes the assumption that

* This research is supported by the AHRB. Thanks to Natasha Alechina, Thorsten Altenkirch, Brian Logan and Sieuwert van Otterloo for commenting on earlier drafts of the paper.

the moon is made of green cheese contains the sentences that the agent would believe, were it to consider that assumption to be true.

Contexts are a suitable tool for this purpose as they can be embedded, thus allowing one to model the making of assumptions within assumptions. In the single agent setting of this paper, we reserve one special context, the empty context, as a model of the agent's actual beliefs, i.e. those sentences entertained with no assumption whatever. In a more general multi-agent setting, we will have a unit context for each agent $i$, corresponding to each agent $i$'s assumption-free beliefs.

We should perhaps point out that this is a different use of the term *context* from that which most frequently appears in the philosophical literature, which denotes certain aspects of the world (an example is Kaplan's [5] use of such contexts in his analysis of pure indexicals). Here, contexts are comprised of psychological entities (which may be thought of as specific concepts, sentences of an internal language of thought or "mentalese", . . . ).

We argue that epistemic logics which utilise traditional possible worlds semantics are not suited to this kind of contextual reasoning. An agent may well assume a formula which contradicts its current beliefs; indeed, this is the essence of drawing conclusions by *reductio ad absurdum*. However, traditional epistemic logics based on classical consequence relations become trivial in the presence of such inconsistencies. This suggests a paraconsistent approach. Paraconsistent logics feature in epistemic frameworks that make use of *impossible* possible worlds, such as Levesque's logic of explicit and implicit belief.

However, we also wish to model resource bounded agents; we are particularly interested in agents which take time to reach their conclusions. We should therefore avoid what Perlis *et al* [3, p.1] call a "final tray approach" to modelling an agent's beliefs, which tells us what an agent will eventually believe, given its initial beliefs and the rules it reasons with, but may not at a particular point during that reasoning process.

While such approaches are applicable in many situations, there are cases in which a primary point of the model is to determine whether the agent can reach a given conclusion in a period of time. We may cite examples such as verifying security protocols and modelling the behaviour of theorem provers. In our example we discuss an who reasons in a natural deduction style, producing one new formula at a time. If we were to assume the synchronous deductive closure of an agent's beliefs within any context then our agent would be able to reach all conclusions that is possibly could immediately; this defeats the point of modelling, which is to establish whether such an agent could reach a certain conclusion in a set time.

## 2 Introducing Timed Reasoning Logics

We take an approach which differs from traditional possible worlds-based epistemic logics. We take a belief to be part of the internal state of an agent, which allows it to act given the appropriate desires and intentions (i.e., goals and plans). We represent the internal state of an agent as a finite set of sentences in some logical language. Whilst we are primarily thinking of rule-based agents, we can accommodate other styles of agent programming by providing a translation of values for program variables into logical formulas, for example. We take as our model of an agent a function which describes how the agent's beliefs change from one deductive cycle to the next. Given a set of initial beliefs, we get a sequence of sets of sentences, each set representing the internal state of the agent at that point in time (figure 1).

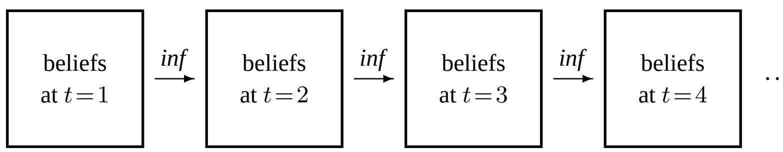


**Fig. 1.** Timeslices

We refer to each such set at a *timeslice*. In addition to modelling how the agent's actual beliefs evolve over a period of time, we can model the beliefs it would come to in the hypothetical circumstance of believing some formula $\phi$. We term these the agent's $\phi$-beliefs and we say they constitute the $\phi$-context. We combine our model with a partitioning of the agent's internal state (at a moment of time) into contexts. This allows us to separate the agent's beliefs entertained under different assumptions, and thus to model multiple assumptions concurrently. The context containing the sentences that would eventually be believed if $\phi$ were to be believed provides our model of assuming that $\phi$ is the case. By placing a total ordering on assumptions that can be made, we can view these contexts as a sequence. Since each context corresponds to assuming some formula, such an ordering simply amounts to a total order on formulas. Since assumptions can be made within assumptions, we can have contexts within contexts, as shown in figure 2.

We arrive at a model consisting of a grid of sets of sentences, with rows representing assumption contexts and columns representing timeslices of those contexts, as shown in figure 3.

We use a labelled language $\mathcal{L}^{trl}$ whose syntax is suited to describing such models. We make use of a set of *context letters* $C$ and a set of *timeslice labels*

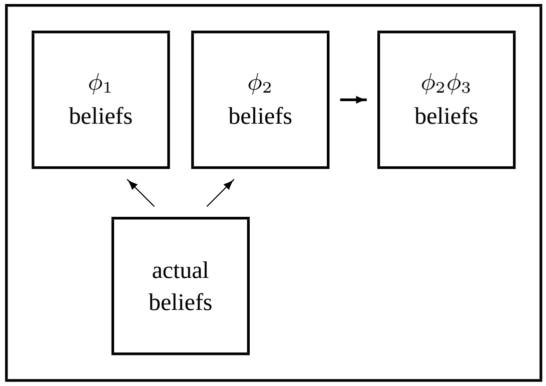


**Fig. 2.** Embedding contexts

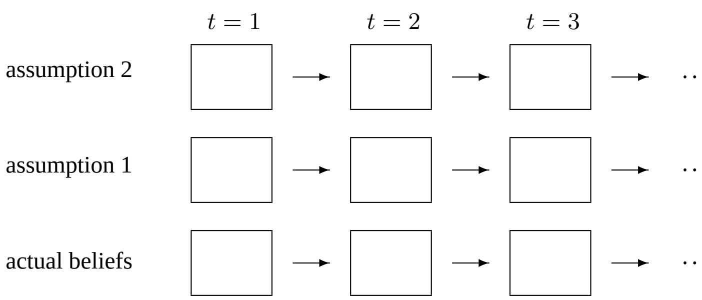


**Fig. 3.** Model of an agent

$T$. Each context is labelled with a finite sequence $\bar{c} = \langle c_1, \ldots, c_n \rangle$ where each $c_{i \leq n} \in C$ (i.e. context labels are words whose alphabet is $C$). We denote the set of all context labels by $C^*$. For convenience, we write $\bar{c}k$ to denote the result of concatenating some $k \in C$ to a sequence $\bar{c} \in C^*$.

Suppose the agent we are modelling derives formulas in a propositional language $\mathcal{L}$ over $\neg, \wedge, \vee$ and $\rightarrow$. Well-formed formulas of $\mathcal{L}^{trl}$ then consist of a *label* $l$, which describes both a context at a timepoint and a *body* which is a formula taken from the agent's language $\mathcal{L}$. A label $l$ is written as a pair $(\bar{c}, t)$ where $\bar{c} \in C^*$ and $t \in T$ (where $\bar{c}$ is some sequence $c_1, \ldots, c_n$). Well formed $\mathcal{L}^{trl}$ formulas are thus of the form $(\bar{c}, t) : \phi$. We assume that $T$ is totally ordered; for our purposes, we can take $T$ to be the set of natural numbers. We write rules connecting contexts and timeslices in the style of natural deduction inference rules as follows:

$$\frac{(\bar{c}, t) : \phi_1 \quad \ldots \quad (\bar{c}', t) : \phi_n}{(\bar{c}'', t+1) : \psi}$$

Here, $\bar{c}, \bar{c}', \bar{c}''$ are variables ranging over contexts, $t$ ranges over timeslices and $\phi_1, \ldots, \phi_n, \psi$ range over well formed formulas of the agent's language $\mathcal{L}$. Note that, while each $\phi_{i \leq n}$ and $\psi$ may belong to a distinct context, rules always connect a timeslice $t$ of the relevant context to the successor timeslice $t+1$. Thus, all rules conform to the step logic paradigm, i.e., the idea of modelling inference as a step-by-step process.

## 3 Modelling a Natural Deduction Style Reasoner

In this section, we introduce as a working example an agent which reasons in a natural deduction style. Such an agent is required to make and later cancel assumptions in order to derive new formulas. Our aim is to model this process using contexts connected by *assumption rules*, similar to the bridge rules of context logic). Moreover, we model the agent's deliberation as a step-by-step process of rule application. We again use contexts to represent this process, this time connected by *temporal rules*, similar to the rules of step logic [3].

We thus make a dual use of the notion of context to represent both temporal and assumption information. It is important to note that contexts representing assumptions and contexts representing temporal states are not distinct; rather, each context represents *both* temporal information and information about assumptions. A third use of contexts would be to use distinct contexts to represent different agents. Although we do not pursue this interpretation here, the formal apparatus it requires is already present in the logical framework we present in section 5 below.

We begin by looking how temporal rules relate timeslices. Intuitively, each timeslice label $t \in T$ corresponds to a timepoint in a linear, non-branching temporal structure and the timeslice labelled by $t$ contains the beliefs of our agent at that time (within some context). As a convenient shorthand, we will often write $t+1$ to denote the successor of $t$, although we assume no need for arithmetical operations. Temporal rules may connect a timeslice labelled with $t$ to one labelled with $t'$ only when $t < t'$. This respects the fact that causation, as a matter of metaphysical necessity, is *forwards causation* and thus our agent can only reason forwards in time.

Simple operations that our agent can perform from one timeslice to the next include: $\wedge$-introduction and elimination, $\neg\neg$-elimination and $\vee$-introduction. These inferences are simple as they operate on formulas in a single context. Example temporal rules are listed below:

$$\frac{(\bar{c}, t) : \phi \quad (\bar{c}, t) : \psi}{(\bar{c}, t+1) : \phi \wedge \psi} \wedge_{int} \qquad \frac{(\bar{c}, t) : \phi \wedge \psi}{(\bar{c}, t+1) : \phi} \wedge_{elimL}$$

$$\frac{(\bar{c},t) : \phi \wedge \psi}{(\bar{c},t+1) : \psi} \wedge_{elimR} \qquad \frac{(\bar{c},t) : \neg\neg\phi}{(\bar{c},t+1) : \phi} \neg_{elim}$$

To illustrate how these temporal rules connect timeslices, figure 4 below shows an agent's beliefs evolving from $\phi \wedge \neg\neg\psi$ at some time $t$:

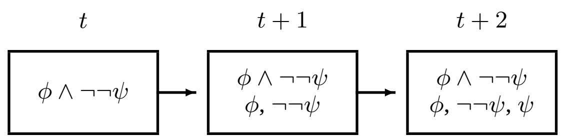


**Fig. 4.** Temporal rules

It is important to note that the formulas which hold at a timeslice only do so on the basis their appearing in the consequent of some rule. In particular, there is nothing inherently monotonic about the Timed Reasoning framework being described here. $\wedge_{int}$ tells us, for example, that $\phi \wedge \psi$ will be a $t+1$-formula when $\phi$ and $\psi$ are both $t$-formulas but we have not said whether $\phi$ or $\psi$ are themselves $t+1$ formulas. It may be the case, for example, that each timeslice (of a particular context) may only contain a fixed number of formulas, due to an agent's memory restrictions. Such a case would clearly be nonmonotonic; that is, $\Gamma \vdash \phi$ would not ensure that $\Gamma \cup \Delta \vdash \phi$ for any sets of formulas $\Gamma, \Delta$ and formula $\phi$, for formulas derived from $\Delta$ might take up space in some future timeslice required to store formulas of $\Gamma$. However, since we require the natural deduction agent we present here to reason monotonically, we need to add an extra temporal rule to ensure just that:

$$\frac{(\bar{c},t) : \phi}{(\bar{c},t+1) : \phi} \mathrm{MON}$$

We now introduce assumption contexts and the corresponding bridging rules, which we will refer to as *assumption rules*. We take as our examples reasoning by *reductio ad absurdum* and $\rightarrow$-introduction, as both require a reasoner to make an assumption which is later withdrawn. To begin with, consider using *reductio ad absurdum* to derive $\neg(\phi \wedge \neg\phi)$; one first assumes $\phi \wedge \neg\phi$, derives each conjunct using $\wedge$-elimination and, on noting the contradiction, establishes that the assumed formula cannot be the case.

We use contexts to model the making of assumptions: formulas entertained within that context are thus those within the scope of the corresponding assumption. It is therefore useful to label such contexts with the assumptions they correspond to, i.e., the very formula assumed in making the assumption. Thus, we take $C$, our set of context symbols, to be $\mathcal{L}$, the internal language of our agent. Contexts are thus labelled with sequences of well formed formulas of $\mathcal{L}$. We denote the empty context by $\cdot$.

We must ensure that, when a formula $\phi$ features in a context label $\bar{c}$, that $\phi$ also appears as a formula in that context. In addition, we must make sure that $\bar{c}$ formulas are also $\bar{c}'$ formulas whenever $\bar{c}$ is a subsequence of $\bar{c}'$. This is so we respect the fact that we may use the formulas holding in a context $\bar{c}$ in any context which we opened from within $\bar{c}$:

$$\frac{}{(\bar{c}, t) : \phi} \text{ (where } \phi \in \bar{c}) \ \textsc{con} \qquad \frac{(\bar{c}_1, t) : \phi}{(\bar{c}_2\bar{c}_1\bar{c}_3, t+1) : \phi} \ \textsc{inherit}$$

Embedded contexts are represented as sequences read left-to-right, thus if the sequence $\langle \phi, (\psi \vee \chi) \rangle$ were used to label a context, that context would represent the agent as having assumed $\phi$ and then, within the scope of that assumption, assuming $\psi \vee \chi$ as well. As mentioned above, $\bar{c}\phi$ denotes the concatenation to the right of $\phi$ to $\bar{c}$ (where $\bar{c}$ is some context label). Thus, $\bar{c}\phi$ can be used to label the context which represents the agent assuming $\phi$ from the $\bar{c}$-context. We may then write a rule for *reduction ad absurdum* as follows:

$$\frac{(\bar{c}\phi, t) : \chi, (\bar{c}\phi, t) : \neg\chi}{(\bar{c}, t+1) : \neg\phi} \ \neg_{int}$$

We may read this rule as saying: having assumed $\phi$ from a context $\bar{c}$ and having derived a contradiction (i.e., some formula $\chi$ and its negation) in $\bar{c}\phi$, we may infer $\neg\phi$ in $\bar{c}$ at the next timeslice. Figure 5 shows how we model an agent using *reductio ad absurdum* to derive $\neg(\phi \wedge \neg\phi)$.

An agent derives an implication in much the same way. Having assumed $\phi$ in a context $\bar{c}$ and having then derived $\psi$, one may infer $\phi \to \psi$ in the original context $\bar{c}$ at the next timeslice:

$$\frac{(\bar{c}\phi, t) : \psi}{(\bar{c}, t+1) : \phi \to \psi} \to_{int}$$

Figure 6 below shows a model of an agent deriving $\phi \to (\psi \to \phi)$ (formulas of the form $\phi \wedge \phi$ have been left out of the diagram to save space).

## 4 Adequacy of the Model

We now show that our framework is adequate as a model of the natural deduction agent we have described in that, for any propositional formula derivable in standard natural deduction, there is a corresponding labelled formula derivable using our labelled logic. Recall that "$\cdot$" is a context label which denotes the

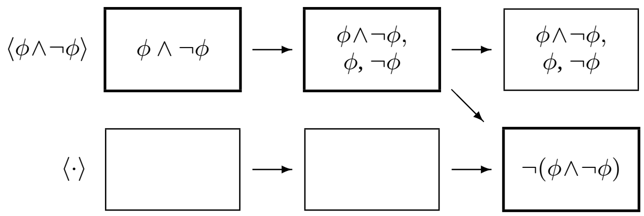


**Fig. 5.** Reductio ad absurdum

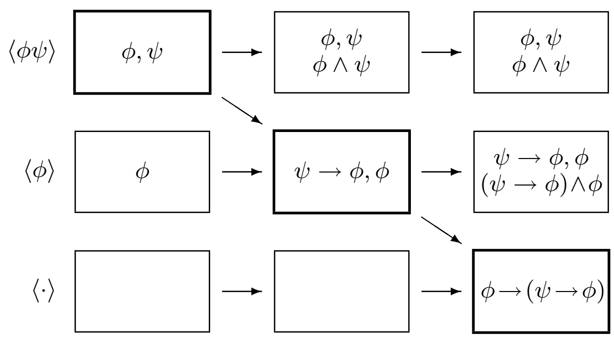


**Fig. 6.** $\rightarrow$ introduction

empty context and that the formulas appearing in a context label represent assumptions used to derive the formulas which hold in that context. Thus, formulas which hold in $\cdot$ at some timepoint $t$ hold with no assumptions being made. If we think of formulas holding in a context $\bar{c}$ to be what the agent's beliefs would be, were it to believe the formulas in $\bar{c}$, then we may think of the formulas holding in $\cdot$ as the agent's actual beliefs.

In this section, we will show that, for any (unlabelled) propositional formula $\phi$ of the agent's internal language $\mathcal{L}$ derivable from a set of (unlabelled) propositional formulas $\Gamma$ using standard natural deduction rules, it is possible to derive a labelled formula $(\cdot, t) : \phi$, for some timepoint $t$, from a labelled version of $\Gamma$. We concentrate on an agent with an internal language $\mathcal{L}^{\neg,\wedge}$ over $\neg$, $\wedge$ and a set of propositional letters $\mathcal{P}$ only (of course, the other connectives could be introduced by definition in the usual way). We will refer to the set of bridge rules we have introduced for these connectives, namely $\wedge_{int}, \wedge_{elimL}, \wedge_{elimR}, \neg_{int}$ and $\neg_{elim}$, as well as the rules MON, INHERIT and CON, as $R$.

**Definition 1 (Derivation).** *A labelled formula $(\bar{c}, t) : \phi$ is* derivable *from a set of labelled formulas $\Gamma^l$ using the rules in $R$, written $\Gamma^l \vdash_R (i, t) : \phi$, if there is a sequence of labelled formulas $(\bar{c}_1, t_1) : \phi_1, \ldots, (\bar{c}_m, t_n) : \phi_k$ such that:*

1. *each formula in the sequence is either a member of $\Gamma^l$, or is obtained from $\Gamma^l$ by applying one of the bridge rules in $R$; and*
2. *the last labelled formula in the sequence is $(\bar{c}, t) : \phi$.*

Let $\Gamma$ be a set of *unlabelled* formulas and $\phi$ an unlabelled formula; we write $\Gamma \vdash_{mp} \phi$ when $\phi$ can be derived from $\Gamma$ using the usual rules of natural deduction. We now form a set of labelled formulas $\Gamma^l = \{(\cdot, 1) : \phi \mid \phi \in \Gamma\}$ which correspond to placing each $\phi \in \Gamma$ in the empty (i.e., assumption-free) context at time 1.

**Theorem 1.** *Let $\Gamma$ be a set of unlabelled propositional formulas and $\phi$ be an unlabelled propositional formula in the agent's language $\mathcal{L}^{\neg,\wedge}$. Then $\Gamma \vdash_{nd} \phi$ iff for some $t$, $\Gamma^l \vdash_R (\cdot, t) : \phi$, where $\Gamma^l = \{(\cdot, 1) : \psi \mid \psi \in \Gamma\}$.*

The proof can be found in [6]. $\dashv$

**Corollary 1.** *For any finite set of labelled formulas $\Gamma^l$ and a labelled formula $\phi^l$, it is decidable whether $\Gamma^l \vdash_R \phi^l$.*

## 5 Semantics

In this section, we introduce a semantics with respect to which the rules $R$ presented above are sound and complete. To establish completeness, we introduce the notions of a *sufficient model* and of a *minimal model*. To keep notation uniform, we use $\phi, \psi$ to denote arbitrary unlabelled formulas and $\Gamma$ to denote a set thereof; the superscript $l$ denotes labelled versions, e.g., $\phi^l$ and $\Gamma^l$ denote an arbitrary labelled formula and a set thereof, labelled as described above.

**Definition 2 (Models).** *Let $\Sigma \subset \mathcal{L}^*$ be a finite set of finite sequences of formulas over $\mathcal{L}$ and $inf$ be a function of type $\Sigma \longrightarrow \mathbb{N} \longrightarrow \wp(\mathcal{L})$. A model $M$ is a tuple $\langle \Sigma, inf, \{m_t^{\bar{c}} \mid \bar{c} \in C^*, t \in \mathbb{N}\}\rangle$ where:*

1. *each $m_t^{\bar{c}}$ is a finite set of formulas in $\mathcal{L}$ such that $m_{t+1}^{\bar{c}} = inf(\bar{c}, t)$*
2. *$inf$ satisfies:*
   $\phi_1 \wedge \phi_2 \in m_t^{\bar{c}} \Longrightarrow \phi_1, \phi_2 \in inf(\bar{c}, t)$
   $\neg\neg\phi \in m_t^{\bar{c}} \Longrightarrow \phi \in inf(\bar{c}, t)$
   $\phi \in m_t^{\bar{c}} \Longrightarrow \phi \in inf(\bar{c}, t)$
   $\phi, \neg\phi \in m_t^{\bar{c}\psi} \Longrightarrow \neg\psi \in inf(\bar{c}, t)$
   $\phi \in \bar{c} \Longrightarrow \phi \in inf(\bar{c}, t)$
   $\phi \in m_t^{\bar{c}} \Longrightarrow \phi \in inf(\bar{c}', t)$ *where* $\bar{c}' = \cdots\bar{c}\cdots$

*When a model $M$ minimally (w.r.t. $\subseteq$) satisfies these conditions, we say that it is a minimal model.*

Intuitively, $inf(\bar{c}, t)$ looks at a timeslice $t$ and infers which formulas should be in the context $\bar{c}$ at the next timeslice.

**Definition 3 (Satisfaction).** *We say a labelled formula of the form $(\bar{c}, t) : \phi$ is satisfied by a model $M$, written $M \models (\bar{c}, t) : \phi$, iff $\phi \in m_t^{\bar{c}}$.*

**Definition 4 (Validity and Entailment).** *A labelled formula $\phi^l$ is (i) $\Sigma$-valid, written $\models_\Sigma \phi^l$ iff $M \models \phi^l$ for all models $M$ over $\Sigma$ and (ii) a $\Sigma$-consequence of a set of labelled formulas $\Gamma^l$, written $\Gamma^l \models_\Sigma \phi^l$, whenever $\phi^l$ is satisfied by all models $M$ over $\Sigma$ which also satisfy every $\psi^l \in \Gamma^l$.*

Since different derivations may rely on making different sequences of assumptions, there is a special class of model, comprising only those that ensure all such assumptions are represented. We call such models *sufficient* that particular derivation. Note that we could not simply assume that $\Sigma$ contains all possible sequences over $\mathcal{L}$ in models in this class, for this would allow the possibility of infinitely many formulas being introduced to a local state by the $inf$ condition, which would violate the definition of each $m_t^{\bar{c}}$ as a finite set. Instead, we use the following definition of sufficiency:

**Definition 5 (Sufficient Models).** *A model $M$ is sufficient for a pair $\langle \Gamma^l, \phi^l \rangle$ if either $\Gamma^l \nvdash_R \phi^l$ or else there exists a derivation $\Gamma^l \vdash_R \phi^l$ containing a sequence of labelled formulas $(\bar{c}_1, t_1) : \phi_1, \ldots, (\bar{c}_m, t_n) : \phi_k$ and each $\bar{c}_{i \leq m} \in \Sigma \in M$. Given a semantic sequent $\Gamma^l \models_\Sigma \phi^l$, we say that $\Sigma$ is sufficient (for that sequent) iff every model $M$ over $\Sigma$ is sufficient for $\langle \Gamma^l, \phi^l . \rangle$.*

We can now establish that the rules $R$ are sound and complete w.r.t. the semantics just proposed. We begin by preparing the following lemma:

**Lemma 1.** *Let $M$ be minimal for $\Gamma^l$ and sufficient for $\langle \Gamma^l, (\bar{c}, t) : \phi \rangle$. Then for every formula $\phi$, $\phi \in m_t^{\bar{c}}$ iff $\Gamma^l \vdash_R (\bar{c}, t) : \phi$.*

Again, the proof can be found in [6]. $\dashv$

**Theorem 2 (Soundness & Completeness).** *There is a sufficient $\Sigma$ such that $\Gamma^l \vdash_R \phi^l$ iff $\Gamma^l \models_\Sigma \phi^l$.*

*Proof.* Soundness is standard: clearly, the rules in $R$ preserve validity. For completeness, suppose $\Gamma^l \models_\Sigma \phi^l$. By definitions 2 and 5, there is a minimal model $M$ over $\Sigma$ of $\Gamma^l$ such that $M \models \phi^l$. By Lemma 1, $\Gamma^l \vdash_R \phi^l$. $\dashv$

**Corollary 2.** *$\phi$ is a classical consequence of $\Gamma$ iff $\Gamma^l \models_\Sigma (\cdot, t) : \phi$ for some $t$.*

**Corollary 3.** *It is decidable whether $\Gamma^l \models_\Sigma \phi^l$.*

## 6 Reasoning with Goals

Before we conclude, we briefly look at how the framework presented here can incorporate goal-centered behaviour to avoid some of the indeterminacy required in choosing assumptions. The agent we have described so far is non-deterministic in the sense that it begins by guessing which assumptions to make; our model begins with the corresponding contexts and derives the consequences. In this sense, we have a model of the consequences of making some assumption, rather than a model of the reasons an agent may have for doing so.

To incorporate this latter notion, we can add goal contexts to our framework. In the setting of our natural deduction reasoner, a goal context contains the formula or formulas which the agent would like to try to derive. For example, if a goal is to derive $\phi \rightarrow \psi$, one can begin by assuming $\phi$. Within the $\phi$-context, one's subgoal is to derive $\psi$. Thus, each context will have a goal context associated with it. We can refine the behaviour of our model by introducing bridge rules between goal contexts, labelled with a sequence ending in a unique goal symbol $\mathsf{G}$, and assumption contexts:

$$\frac{(\bar{c}\mathsf{G}, t) : \phi \rightarrow \psi}{(\phi, t{+}1) : \phi} \qquad \frac{(\bar{c}\mathsf{G}, t) : \phi \rightarrow \psi}{(\bar{c}\phi\mathsf{G}, t+1) : \psi}$$

(Note that the former replaces the earlier rule $\overline{(\bar{c}, t) : \phi}$ for $\phi \in \bar{c}$ and that the INHERIT rule (section 3) must now only apply when neither context is a goal context.)

We should stress, however, that this section is not intended as a full account of a deterministic natural deduction theorem prover. It is offered as a demonstration of how propositional attitudes other than beliefs, such as desires (goals) and intentions (plans) may be included within the Timed Reasoning Logics framework. Figure 7 below shows a model of an agent incorporating these rules. Goal contexts are represented as the smaller rectangles below the assumption contexts they correspond to — we have left out formulas such as $(\phi \wedge \phi) \wedge \phi$.

Semantically, a model will still contain a set $m_t^{\bar{c}}$ for each context in $\Sigma$ at each timeslice $t$, but we may now calculate *which* contexts are required in $\Sigma$ for a sufficient model $M$, independently of the rules in $R$. Given a goal $\alpha$, we set $\Sigma = \sigma(\alpha)$, where $\sigma : \mathcal{L}^{trl} \longrightarrow \wp(\mathcal{L}^{trl})$ satisfies the following recursive equations:

$$\sigma(p) = \emptyset \tag{1}$$

$$\sigma(\phi \rightarrow \psi) = \{\phi\} \cup \sigma(\psi) \tag{2}$$

A full specification of the semantics for goal–assumption interaction is left for future work.

| | $t=1$ | $t=2$ | $t=3$ | $t=4$ | $t=5$ |
|---|---|---|---|---|---|
| $\langle\phi\psi\rangle$ | | | $\phi \quad \psi$ | $\cdots$ | $\cdots$ |
| $\langle\phi\psi\mathsf{G}\rangle$ | | | $\phi$ | | |
| $\langle\phi\rangle$ | | $\phi$ | $\phi,\ \phi\wedge\phi$ | $\phi,\ \psi\rightarrow\phi$ | $\cdots$ |
| $\langle\phi\mathsf{G}\rangle$ | | $\psi\rightarrow\phi$ | $\psi\rightarrow\phi$ | $\psi\rightarrow\phi$ | |
| $\langle\cdot\rangle$ | | | $\phi\rightarrow\phi$ | $\phi\rightarrow\phi$ | $\phi\rightarrow(\psi\rightarrow\phi)$ |
| $\langle\cdot\mathsf{G}\rangle$ | $\phi\rightarrow(\psi\rightarrow\phi)$ | $\phi\rightarrow(\psi\rightarrow\phi)$ | $\phi\rightarrow(\psi\rightarrow\phi)$ | $\phi\rightarrow(\psi\rightarrow\phi)$ | $\phi\rightarrow(\psi\rightarrow\phi)$ |

**Fig. 7.** Reasoning with goal contexts

## 7 Summary

We have argued that contexts are suitable tools for reasoning about agents in at least two ways. Firstly, they can be used to model assumptions, thus providing a finer grained account than those in which a formula is either believed or disbelieved. Secondly, contexts can represent temporal stages, or timeslices, of an agent's reasoning process. This is a crucial notion if we aim to model resource bounded agents. A further application of context is to model multiple agents, each agent being modelled in its own context. Such an approach is explored in [1] and [2].

Future work aims to explore a full semantics for interaction between beliefs (including assumptions), goals/desires and plans/intentions and to look at real-world applications of assumption-making in reasoning, with games such as *Cluedo* as a test case.

## References


1. N. Alechina, B. Logan, and M. Whitsey (now *Jago*). A complete and decidable logic for resource-bounded agents. In *Proc. Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2004)*. ACM Press, July 2004.
2. Natasha Alechina, Brian Logan, and Mark Whitsey (now *Jago*). Modelling communicating agents in timed reasoning logics. In José Júlio Alferes and João Alexandre Leite, editors, *JELIA*, volume 3229 of *Lecture Notes in Computer Science*, pages 95–107. Springer, 2004.
3. J. Drapkin and D. Perlis. A preliminary excursion into Step-Logics. *Proceedings of the SIGART International Symposium on Methodologies for Intelligent Systems*, pages 262–269, 1986.
4. C. Ghidini and F. Giunchiglia. Local models semantics, or contextual reasoning = locality + compatability. *Artificial Intelligence*, 127(2):221–259, 2001.

5. David Kaplan. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*, chapter 17, pages 481–563. Oxford University Press, New York, 1989.
6. M. Whitsey (now *Jago*). Timed reasoning logics: An example. In *Proc. of the Logic and Communication in Multi-Agent Systems workshop (LCMAS 2004)*. Loria, 2004.