

Available from: http://dx.doi.org/10.1007/3-540-45656-2_45
Solving the Qualification Problem
(In the Presence of the Frame problem)

Quoc Bao Vo and Norman Y. Foo

Knowledge Systems Group
School of Computer Science and Engineering
University of New South Wales - Sydney NSW2052 Australia
{vqbao | norman}@cse.unsw.edu.au

Abstract. We present a uniform nonmonotonic solution for the problem of reasoning about action on the basis of argumentation-theoretic approach in a series of paper. This paper is the first one in which we solve the frame and the qualification problems in a simplifying setting without domain constraints or ramifications. Our theory is provably correct relative to a sensible minimisation policy introduced on top of a temporal propositional logic.

1 Motivation and Introduction

The need for a good reasoning about action formalism is apparent for research in artificial intelligence (AI). Alongside the logicist point of view to artificial intelligence, more recently, there emerges the cognitivist and situated action-based approaches (see [10] and the references therein). The latters provide some immediate and practical answers to certain issues of AI. The current problem domains for (Soccer) Robot Cup seem to be an area where these approaches promise to gain fruitful results. On the other hand, the logicist approach aims at long term solutions for the general problems of AI. From a logicist approach, formalising dynamic domains for reasoning about action can be realised within a logical knowledge representation. The general idea is that intelligent agents should be able to represent all kinds of knowledge in a uniform way such that some general problem solver can fully employ and find a solution based on their knowledge. As it turns out, there are difficulties with such a general approach to AI. Consider the task of formalising dynamic domains in some logical language. To formalise the dynamics of an action (or event) in a language with \( n \) fluents\(^1\), one will need to axiomatise not only about the fluents that are affected by the action but also about those that are not. Essentially, it requires that \( n \) axioms be asserted. Such a formalisation can hardly be considered a good representation. Hence, there is the need to solve this problem in logic-based reasoning about action formalisms. This is the well known frame problem as introduced by McCarthy and Hayes ([15]). Moreover, there is still a problem in axiomatising the effects of an action, called the effect axiom.

\(^1\) fluent is a technical term referring to functions or predicates whose values can be varied relative to time.
A logical axiomatisation requires that the conditions under which the effects will take place after executing the action be precisely specified. However, there are potentially infinitely many such conditions, some of which the reasoner may have never thought about. No realistic formalisation would ever be able to exhaustively enumerate all of those conditions. Nonetheless, to start a car, most people only worry about whether they have the key to that car. They never bother checking whether there is something blocking the tailpipe or checking all electric circuits to make sure that they are all well connected. Such a story has long been well-known within the community of common-sense reasoning, in particular reasoning about action. This is known as the qualification problem and was introduced by McCarthy (cf. [13]).

While there have been a number of solutions to the frame problem (see e.g. [19],[16] and [3]), the qualification problem has largely been ignored. Some people argue that the frame problem is already very challenging and it would be a good approach to thoroughly solve the frame problem before complicating a formalism with the qualification problem. We argue that there is a danger of approaching these problems from that point of view for (at least) two reasons: (1) it may be very hard to come up with a uniform solution for all problems; while many existing solutions for the frame problem are monotonic (e.g. [3] and [16]), the qualification problem inherently requires a non-monotonic solution; and (2) these solutions of the frame problem can only succeed under some precise assumptions:

- Actions always succeed. This is the action omniscience assumption. More precisely, this assumption dictates that the qualification problem is skipped.
- Fluents change if and only if the reasoner knows that there exists an action that possibly changes its value. This can be termed as domain omniscience assumption. It assumes that the reasoner has complete (ontological) knowledge about the domain on which he is reasoning about.

The above two reasons are of course closely related as (1) arises due to the underlying assumptions in (2) which can no longer hold once the qualification problem is taken into consideration.

In this paper, a uniform nonmonotonic solution for the two most basic problems of reasoning about action is proposed. Basically, when performing common sense reasoning, the reasoner is based on a number of plausible assumptions. E.g., assuming that an instance of birds flies, or assuming that shooting a turkey with a loaded gun causes it to die, etc. The proposed representation formalism aims at making these assumptions explicit so that an automated reasoner is conscious (at least) about what assumptions it relies on when performing reasoning. It is also the basic idea of assumption-based frameworks which are at heart of Bondarenko et al.’s (1997) argumentation-theoretic approach. As a first step towards a comprehensive framework, we show how the frame and the qualification problems are solved in the absence of domain constraints and ramifications.

2 Domain Descriptions

We introduce a propositional action description language based on a more comprehensive representation formalism proposed by Sandewall (1994). In particular,
we extend Drakensen and Bjørlend’s (1999) language so that it is possible to
describe narratives in our framework.

2.1 Syntax

Following Sandewall’s, the underlying representation of time is a (discrete) time
structure \( T = (T, <, +, -) \) consisting of

- a time domain \( T \) whose members are called timepoints which are integers in
this paper;
- \(<, +, -\) are as usual for integers.

Given a time structure \( T = (T, <, +, -) \), a signature with respect to \( T \) is
a tuple \( \sigma = (T, \mathcal{F}, \mathcal{A}) \), where \( T \) is a set of timepoint variables, \( \mathcal{F} \) is a set of
propositional fluent names, and \( \mathcal{A} \) is a set of action names. We assume that all
sets in \( \sigma \) are countable. We denote \( \mathcal{F} = \{ \neg f \mid f \in \mathcal{F} \} \). A member of \( \mathcal{F}^* = \mathcal{F} \cup \mathcal{F} \)
is a fluent literal. Moreover, \( \mathcal{A} = A_0 \cup \mathcal{D}A \), where is the set of domain dependent
action names, called basic actions, e.g. load, shoot, etc. and \( \mathcal{D}A = \{ da _\varphi \mid \varphi \in \mathcal{F}^* \} \)
is the set of dummy actions.

For each fluent literal \( \varphi \in \mathcal{F}^* \), we introduce the following two propositions:
\( AQ \varphi \) and \( FA \varphi \). \( AQ \varphi \) is associated with the assumed qualifications upon the
preconditions of an action regarding the fluent \( \varphi \). \( FA \varphi \) is associated with the frame assumptions regarding \( \varphi \). Given a set of fluent literals \( \Gamma \subseteq \mathcal{F}^* \), we denote
\( FA \Gamma \overset{def}{=} \{ FA _\varphi \mid \varphi \in \Gamma \} \) and \( AQ \Gamma \overset{def}{=} \{ AQ _\varphi \mid \varphi \in \Gamma \} \).

A timepoint expression is one of the following:

- a member of \( T \),
- a timepoint variable in \( T \),
- an expression formed from timepoint expressions using \( + \) and \( - \). For convenience, we will also write \( \tau^+ \) and \( \tau^- \) instead of \( \tau + 1 \) and \( \tau - 1 \), respectively.

We denote the set of timepoint expressions by \( TE \).

Definition 1 Let \( \sigma = (T, \mathcal{F}, \mathcal{A}) \) be a signature and \( \tau, \nu \in TE, f \in \mathcal{F}, A \in \mathcal{A}, \)
\( R \in \{=, <\}, \odot \in \{\land, \lor, \rightarrow, \leftrightarrow\} \). Define the basic (domain description) language \( \Lambda \)
over \( \sigma \) by:

\[
\Lambda ::= T \mid F \mid f \mid [\tau, \nu]A \mid \tau R \nu \mid \neg A \mid A_1 \odot A_2 \mid [\tau]A,
\]

and the assumption base \( AB \) by:

\[
AB = AB_{AQ} \cup AB_{FA}, \text{ where } AB_{AQ} = \{ [\tau, \nu]AQ _\varphi \mid \tau, \nu \in TE \text{ and } \varphi \in \mathcal{F}^* \},
\]

\[
AB_{FA} = \{ [\tau]FA _\varphi \mid \tau \in TE \text{ and } \varphi \in \mathcal{F}^* \}.
\]

The domain description language \( \mathcal{L} \) (over \( \sigma \)) is defined: \( \mathcal{L} = \Lambda \cup AB \).

\([\tau, \nu]A\) means the action \( A \) is performed during the time interval \([\tau, \nu]\). \([\tau, \nu]AQ _\varphi \)
means the fluent literal \( \varphi \) is assumed to be qualified to hold by the end of the
interval \([\tau, \nu]\). \([\tau]FA _\varphi \) means the fluent literal \( \varphi \) is assumed by default to persist
from the time point \( \tau \) to the next, i.e. the principle of inertia.

A formula that does not contain any connectives (i.e. \( \land, \lor, \rightarrow, \leftrightarrow, \neg, \) and \( [\ ]\)) is
atomic. If \( \gamma \) is atomic and \( \tau \in TE \), then the formula \( [\tau]\gamma, [\tau]\neg \gamma, \neg [\tau]\gamma, \) and
\([\tau]\neg \gamma \) are literals.

Let \( \gamma \) be a formula. A fluent \( f \in \mathcal{F} \) occurs free in \( \gamma \) iff it does not occur within
the scope of a \([\tau]\) expression in \( \gamma \). \( \tau \in TE \) binds \( f \) in \( \gamma \) if a formula \([\tau]\psi \) occurs as
a subformula of $\gamma$, and $f$ is free in $\psi$. If no fluent occurs free in $\gamma$, $\gamma$ is \textit{closed}. If $\gamma$ does not contain any occurrence of $\tau$ for any $\tau \in T \mathcal{E}$, then $\gamma$ is \textit{propositional}.

2.2 Semantics

\textbf{Definition 2} Let $\sigma = \langle T, \mathcal{F}, A \rangle$ be a signature. A \textit{state} over $\sigma$ is a function from $\mathcal{F}$ to the set $\{T, F\}$ of truth values. A \textit{history} over $\sigma$ is a function $h$ from $T$ to the set of states. A \textit{valuation} is a function $\phi$ from $T \mathcal{E}$ to $T$. A \textit{narrative assignment} is a function $\eta$ from $T \times A \times T$ to the set $\{T, F\}$. In addition, we define $\varepsilon_q : T \times AQ \times T \rightarrow \{T, F\}$ and $\varepsilon_f : T \times FA \times T \rightarrow \{T, F\}$. An interpretation over $\sigma$ is a tuple $\langle h, \phi, \eta, \varepsilon_q, \varepsilon_f \rangle$ where $h$ is a history, $\phi$ is a valuation, $\eta$ is a narrative assignment and $\varepsilon_q, \varepsilon_f$ are defined as above.

\textbf{Definition 3} Let $\gamma, \delta \in A$ and $I = \langle h, \phi, \eta, \varepsilon_q, \varepsilon_f \rangle$ an interpretation. Assume $\tau, \nu \in T \mathcal{E}, f \in \mathcal{F}, A \in A, R \in \{=, \leq\}, \varphi \in \mathcal{F}^*$, $\otimes \in \{\land, \lor, \rightarrow\}$, and $\chi \in \{T, F\}$. Define the truth value of $\gamma$ in $I$ for a timepoint $t \in T$, denoted $I(\gamma, t)$ as follows:

$I(\chi, t) = \chi$
$I(f, t) = h(t)(f)$
$I([\tau, \nu]A, t) = \eta(\tau, A, \nu)$
$I([\tau, \nu]AQ, t) = \varepsilon_q(\tau, AQ, \nu)$
$I([\tau]FA, t) = \varepsilon_f(\tau, FA)$
$I(\tau \nu, t) = \phi(\tau) R \phi(\nu)$
$I(\neg \gamma, t) = I(\gamma, t)$
$I(\gamma \otimes \delta, t) = I(\gamma, t) \otimes I(\delta, t)$
$I([\tau] \gamma, t) = I(\gamma, \phi(\alpha))$

Two formulas $\gamma$ and $\delta$ are equivalent iff $I(\gamma, t) = I(\delta, t)$ for all $I$ and $t$. An interpretation $I$ is a \textit{model} of a set $\Gamma \subseteq A$ of formulas, denoted $I \models \Gamma$, iff $I(\gamma, t) = T$ for every $t \in T$ and $\gamma \in \Gamma$. A formula $\gamma \in A$ is \textit{entailed} by a set $\Gamma \subseteq A$ of formulas, denoted $\Gamma \models \gamma$, iff $\gamma$ is true in all models of $\Gamma$.

\textbf{Definition 4} Let $I = \langle h, \phi, \eta, \varepsilon_q, \varepsilon_f \rangle$ be an interpretation. The set $\text{Occ}_I = \{(t, A, u) \in T \times A \times T | \eta(t, A, u) = T\}$ is called \textit{action occurrence denotation} of $I$. The set $\text{FA}_I = \{(t, FA) \in T \times FA | \varepsilon_f(t, FA) = T\}$ is called \textit{FA-denotation} of $I$. The set $\text{AQ}_I = \{(t, AQ, u) \in T \times AQ \times T | \varepsilon_q(t, AQ, u) = T\}$ is called $\textit{AQ-denotation}$ of $I$.

2.3 Background

Bondarenko \textit{et al.} (1997) propose a unified framework for default reasoning called argumentation-theoretic approach which we will use as the underlying inference mechanism for our system. We reproduce the relevant definitions from Bondarenko \textit{et al}.’s work for completeness.

A \textit{deductive system} is a pair $\langle \mathcal{L}, \mathcal{R} \rangle$, where

- $\mathcal{L}$ is a formal language consisting of countably many sentences, and
- $\mathcal{R}$ is a set of inference rules of the form

\[
\begin{array}{c}
\alpha_1, \ldots, \alpha_n \\
\hline
\alpha
\end{array}
\]
where \( \alpha, \alpha_1, \ldots, \alpha_n \in \mathcal{L} \) and \( n \geq 0 \).

Any set of sentences \( T \subseteq \mathcal{L} \) is called a \textit{theory}. A \textit{deduction} from a theory \( T \) is a sequence \( \beta_1, \ldots, \beta_m \), where \( m > 0 \), such that, for all \( i = 1, \ldots, m \),
- \( \beta_i \in T \), or
- there exists \( \frac{\alpha_1, \ldots, \alpha_n}{\alpha_{i-1}} \in \mathcal{R} \) such that \( \alpha_1, \ldots, \alpha_n \in \{ \beta_1, \ldots, \beta_{i-1} \} \).

\( T \vdash_{(\mathcal{L}, \mathcal{R})} \alpha \) means that there is a deduction from \( T \) whose last element is \( \alpha \). \( Th_{(\mathcal{L}, \mathcal{R})}(T) \) is the set \( \{ \alpha \in \mathcal{L} \mid T \vdash_{(\mathcal{L}, \mathcal{R})} \alpha \} \). Since the language \( \mathcal{L} \) is generally kept fixed whereas the set of inference rules \( \mathcal{R} \) is likely to vary depending on the description of the domain, when there is no possible confusion we will abbreviate \( \vdash_{(\mathcal{L}, \mathcal{R})} \) and \( Th_{(\mathcal{L}, \mathcal{R})} \) as \( \vdash_\mathcal{R} \) and \( Th_\mathcal{R} \), respectively. Thus the classical inference relation \( \vdash \) can also be written as \( \vdash_\mathcal{R} \mathcal{C} \) where \( \mathcal{R}_\mathcal{C} \) is the set of inference rules of classical propositional logic. Note also that every set of inference rules considered in this paper will be a super set of \( \mathcal{R}_\mathcal{C} \).

Given \( r = \frac{\alpha_1, \ldots, \alpha_n}{\alpha} \in \mathcal{R} \), we will also denote \( \text{prem}(r) = \{ \alpha_1, \ldots, \alpha_n \} \), the premises of \( r \), and \( \text{cons}(r) = \alpha \), the consequence of \( r \).

**Definition 5** [2] Given a deduction system \( (\mathcal{L}, \mathcal{R}) \), an \textit{assumption-based framework} with respect to \( (\mathcal{L}, \mathcal{R}) \) is a tuple \( (T, Ab, \neg) \) where

- \( T, Ab \subseteq \mathcal{L} \) and \( Ab \neq \emptyset \),
- \( \neg \) is a mapping from \( Ab \) into \( \mathcal{L} \), where \( \neg \) denotes the contrary of \( \alpha \).

**Definition 6** [2] Given an assumption-based framework \( (T, Ab, \neg) \),

- a set of assumptions \( \Delta \subseteq Ab \) attacks an assumption \( \alpha \in Ab \) iff \( T \cup \Delta \vdash_\mathcal{R} \neg \alpha \),
- a set of assumptions \( \Delta \subseteq Ab \) attacks a set of assumptions \( \Delta' \subseteq Ab \) iff \( \Delta \subseteq \Delta' \).

As assumptions are expressed in terms of usual propositions, we will replace the notion of contrariness in Bondarenko \textit{et al.}'s system with the classical negation \( \neg \) and omit it from the specification of assumption-based framework.

### 3 Reasoning about action with argumentation-theoretic approach

In the rest of this paper, we introduce a uniform framework for solving the frame and qualification problems using the frame and qualification assumptions. General solutions for the frame and the qualification problems can be obtained by computing plausible sets of assumptions which guarantee that extensions computed from plausible sets of assumptions will be consistent when the given theory is consistent.

**Definition 7** A deductive system \( (\mathcal{L}, \mathcal{R}) \) is \textit{well-defined} iff for each subset \( S \subseteq \mathcal{R} \), if the set \( \bigcup_{r \in S} \text{prem}(r) \) is consistent then the set \( \text{CONS}(S) = \{ \text{cons}(r) \mid r \in S \} \) is also consistent.

We will assume that a deductive system is well-defined. Being formalised in terms of the argumentation-theoretic approach, the representation requires an extended notion of consistency.
Definition 8 Let \( \langle \mathcal{L}, \mathcal{R} \rangle \) be a deductive system, (a) a set of sentences \( \Gamma \subseteq \mathcal{L} \) is \( \mathcal{R} \)-consistent iff \( \Gamma \not\vdash \mathcal{R} \), (b) an assumption-based framework \( \langle T, Ab, \neg \rangle \) with respect to \( \langle \mathcal{L}, \mathcal{R} \rangle \) is consistent iff \( T \) is \( \mathcal{R} \)-consistent.

Definition 9 Given an assumption-based framework \( \langle T, Ab, \neg \rangle \), a set of assumptions \( \Delta \subseteq Ab \) rejects an assumption \( \alpha \in Ab \) iff (a) \( \Delta \) does not attack itself, and (b) \( \Delta \cup \{ \alpha \} \) attacks itself.

Observation 1 Given an assumption-based framework \( \langle T, Ab, \neg \rangle \) and a set of assumptions \( \Delta \subseteq Ab \), if \( \Delta \) attacks an assumption \( \alpha \notin \Delta \) then \( \Delta \) rejects \( \alpha \).

We are interested in the assumptions which are rejected by a given set of assumptions without being attacked by that set.

Definition 10 Given an assumption-based framework \( \langle T, Ab, \neg \rangle \), a set of assumptions \( \Delta \subseteq Ab \) leniently rejects an assumption \( \alpha \in Ab \) iff (a) \( \Delta \) rejects \( \alpha \), and (b) \( \Delta \) does not attacks \( \alpha \).

We denote \( Lr(\Delta) \overset{\text{def}}{=} \{ \alpha \in Ab \mid \alpha \text{ is leniently rejected by } \Delta \} \).

The frame assumptions are the essence of the inertia problem, and their role in the argumentation approach is illustrated below by the Yale Shooting Problem (YSP) [9]. In this formalisation we intentionally ignore the qualification problem (it is addressed in the next section) to highlight how the frame problem is solved. We consider a well-worn example to motivate our approach to the frame problem.

Example 1

\[ AD_{YSP} = \{ [\tau, v]load \rightarrow \{ [v]loaded \wedge \neg [\tau]FA_{\neg loaded} \}, \]  
\[ ([\tau, v]shoot \wedge [\tau]loaded) \rightarrow (\neg [v]alive \wedge \neg [\tau]FA_{alive}) \}. \]

The following rules representing the frame assumptions are added:

\[ FR_{YSP} = \{ \frac{\neg [\tau]loaded, [\tau]FA_{\neg loaded}}{[\tau ^*]loaded}, \frac{\neg [\tau]loaded, [\tau]FA_{\neg loaded}}{[\tau ^*]alive}, \frac{[\tau]alive, [\tau]FA_{alive}}{[\tau ^*]alive}, \frac{\neg [\tau]alive, [\tau]FA_{alive}}{[\tau ^*]alive} \} \]

Given a theory \( T_{YSP} = \{ [0]alive, [0],[1]load, [1],[2]wait, [2,3]shoot \} \), the argumentation-theoretic approach will yield the following preferred set of assumptions (cf. [2]): \( Y_{YSP} = \{ [\tau]FA_{\neg loaded} \mid \tau \in Tme \} \cup \{ [\tau]FA_{\neg loaded} \mid \tau \in Tme \} \cup \{ [\tau]FA_{alive} \mid \tau \in Tme \} \cup \{ [\tau]FA_{\neg alive} \mid \tau \in Tme \} \cup \{ [0]FA_{\neg loaded}, [2]FA_{alive} \}. \) They give rise to the following preferred extension (cf. [2]): \( Th(T_{YSP} \cup \{ [\tau]loaded \mid \tau \geq 1 \} \cup \{ [1]alive, [2]alive \} \cup \{ [\neg \tau]alive \mid \tau \geq 3 \} \cup \Delta \}. \) This extension is also the stable extension and well-founded semantics (cf. [2]) of the given theory under the argumentation-theoretic approach. Note that in case one would like to be uncertain about whether the gun is still loaded after the shooting action, one just simply needs to add an axiom: \( [\tau, v]shoot \rightarrow \neg [\tau]FA_{\neg loaded} \) to dictate that the persistence of the fluent loaded after the action shooting is not guaranteed. In that case, we can still derive that \( [\tau]loaded \) for \( \tau = 1,2 \), but we can no longer give a definite assertion about \( [\tau]loaded \) for \( \tau \geq 3 \).

As the above formalisation of YSP resembles that using default logic, it may be surprising that the problem of unintended models pointed out by Hanks and McDermott for circumscription, default logic, autoepistemic logic does not happen here. The principal reason is the interaction of the inference rules and the notion
of attack in the argumentation-theoretic framework, which invalidates undesired assumptions. Notice that even if $\neg[2]loaded$ can be (magically) derived, it cannot lead to $\neg[1]FA_{loaded}$. Therefore, the set of assumptions corresponding to this case does not satisfy the conditions of preferred set of assumptions, thus ruling out this unintended model.

We adopt the following guidelines in seeking for a sensible solution for the problems of reasoning about action:
- The derived pieces of information don’t conflict with the given facts;
- Occurrences of events are minimised; and
- The inertia of fluent is maximised though the minimality of the event occurrences will be of higher priority.

However, while the preferred model semantics copes successfully with the YSP, it cannot properly account for the explanation problem, e.g. the Stanford Murder Mystery, the Stolen Car Problem. The subtlety lies in the derivation of the contrary of the frame assumption $FA_{\neg e}$. The contrary of a frame assumption is derived only when both the occurrence of the event that brings about the change (absent in the Stolen Car Problem) and the preconditions required to be satisfied for the change to actually take place (absent in the Stanford Murder Mystery) are explicitly derivable. This is where the notion of (leniently) rejected assumptions is called into service.

**Definition 11** Given an assumption-based framework $(T,Ab,\neg)$, a set of assumptions $\Delta \subseteq Ab$ is presumable iff (a) $\Delta = \{\alpha \in Ab \mid T \cup \Delta \vdash_R \alpha\}$ (in Bondarenko et al.’s terms, $\Delta$ is closed), (b) $\Delta$ does not attack itself, and (c) for each assumption $\alpha \notin \Delta$, $\alpha$ is rejected by $\Delta$.

**Definition 12** Given an assumption-based framework $(T,Ab,\neg)$, a set of assumptions $\Delta \subseteq Ab$ is plausible iff (a) $\Delta$ is presumable, and (b) there exists no $\Delta' \subseteq Ab$ such that $\Delta'$ is presumable and $Lr(\Delta') \subseteq Lr(\Delta)$.

## 4 Technical Framework

Aside from the trivial case of occurrences of events causing the frame assumptions to be rejected, two aspects of events can be distinguished:

(a) An event happens but the change it is supposed to cause does not take place. We call this expectation failure and this is more or less the qualification problem; and

(b) No events that are known to cause a change happen but the change does take place. We call this surprise and this is usually known as the explanation problem.

The following assumption represents our underlying intuition behind reasoning about action formalisms.

**Assumption 1** Intuitive models contain minimal (with respect to set inclusion) sets of surprises.

**Definition 13** Let $\sigma = \langle T, F, A \rangle$ be a signature. Assume $\tau, v \in T \in T \in T \in T$. $A \in A$, $R \in \{=, <\} \subseteq A$, and $\varphi \in F^*$. A domain description $D$ is defined to be a tuple $(A, R, Ab, I)$, where:
1. \( \mathcal{L} \) is the domain description language and \( AB \) an assumption base over \( \sigma \);
2. \( \mathcal{R} = \mathcal{R}_C \cup \mathcal{R}_F \cup \mathcal{R}_A \cup \mathcal{R}_Q \), where
   (a) \( \mathcal{R}_C \) is the set of inference rules of (classical) propositional logic;
   (b) \( \mathcal{R}_F \) is the set of \textit{frame-based inference rules} of the form: \( \frac{\varphi \land \psi}{\varphi} \), i.e. those that represent the frame axioms in terms of inference rules;
   (c) \( \mathcal{R}_A \) is the set of \textit{action descriptions} which are inference rules of the form: \( \frac{\varphi}{\varphi \land \neg \psi \land FA_{\varphi \neg \psi}} \), i.e. those that represent the conditions for an action to bring about some effect on a fluent; and
   (d) \( \mathcal{R}_Q \) is the set of \textit{qualification-based inference rules} of the form: \( \frac{\varphi}{\varphi \land \neg \psi \land QA_{\varphi \neg \psi}} \), i.e. those that represent the (dis-)qualifications regarding the fluent literal \( \varphi \).
3. The theory \( \Gamma \subseteq A \).

Given a set of assumptions \( \Delta \), we denote \( \Delta_{FA} = \Delta \cap AB_{FA} \) and \( \Delta_{AQ} = \Delta \cap AB_{AQ} \).

**Definition 14** Let \( \sigma = \langle T, \mathcal{F}, A \rangle \) be a signature and \( D = \langle \mathcal{L}, \mathcal{R}, AB, \Gamma \rangle \) a domain description over \( \sigma \). An interpretation \( I = \langle h, \phi, \eta, \varepsilon, \epsilon \rangle \) is a model of \( D \) iff
   1. \( I \) is a model of \( \Gamma \);
   2. for each \( r \in \mathcal{R} \), if \( I \models \text{prem}(r) \) then \( I \models \text{cons}(r) \).

The following definition captures one of several aspects of the (model-theoretic) solution of the frame problem. This aspect is known as the action-oriented frame problem in Lin and Shoham’s (1995) terms. The proposed minimization policy formalises the intuition that change does not happen by itself but is caused by some kind of event. Thus, for each fluent, if its value is changed between two timepoints \( \tau \) and \( \nu \), (at least) an occurrence of some event must end at \( \nu \) that brings about that change.

**Definition 15** Let \( D = \langle \mathcal{L}, \mathcal{R}, AB, \Gamma \rangle \) be a domain description and \( I \) a model of \( D \). \( I \) is a \textit{coherent model} of \( D \) iff
   1. for each basic action \( \alpha \in A_C \) and \( \tau, \nu \in T\mathcal{E} \), if \( \Gamma \models \neg [\tau, \nu] \alpha \) then \( I \models \neg [\tau, \nu] \alpha \); and
   2. for each \( \varphi \in \mathcal{F}^* \) and \( \tau \in T\mathcal{E} \), if \( I \models [\tau] \varphi \land \neg [\tau^+] \varphi \) then either (i) there is \( A \in A_C \) and \( \nu_1, \nu_2 \in T\mathcal{E} \) such that \( \nu_2 = \tau^+ \), and \( r = \frac{\varphi}{\varphi \land \neg \psi \land QA_{\varphi \neg \psi} \in \mathcal{R}, \quad I \models \text{prem}(r) \), or (ii) \( I \models [\tau, \tau^+] \text{da}_{\neg \varphi} \).

Thus, in a coherent model (1) all satisfiable basic actions must follow from the given theory, and (2) all changes are attributable to events.

Given an interpretation \( I \), we want to extract the sets of assumptions satisfiable in \( I \).

**Definition 16** Let \( \sigma = \langle T, \mathcal{F}, A \rangle \) be a signature and \( I = \langle h, \phi, \eta, \varepsilon, \epsilon \rangle \) an interpretation over \( \sigma \). The set of frame assumptions satisfiable in \( I \), denoted \( \Delta_{FA}^I \), is defined as follows: \( \Delta_{FA}^I = \{ [\tau] FA_{\varphi} \mid (\tau, FA_{\varphi}) \in FA^I \} \) and the set of qualification assumptions satisfiable in \( I \), denoted \( \Delta_{AQ}^I \), is: \( \Delta_{AQ}^I = \{ [\tau, \nu] QA_{\varphi \neg \psi} \mid (\tau, QA_{\varphi \neg \psi}, \nu) \in AQ^I \} \). We also write \( \Delta_{QF}^I = \Delta_{AQ}^I \cup \Delta_{FA}^I \).
Conversely, given a theory $\Gamma$ and a set of assumptions $\Delta$, a reasoner can also construct his models about the domain of interest.

**Definition 17** Let $D = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}\mathcal{B}, \Gamma \rangle$ be a domain description and $\Delta \subseteq \mathcal{A}\mathcal{B}$. A model $I$ of $D$ is $\Delta$-relativised iff

1. for each $\alpha \in \mathcal{A}\mathcal{B}$, $I = \alpha$ iff $\alpha \in \Delta$; and
2. $\text{Occ}^I = OA_D \cup \text{DAS}(\Delta)$, where: (i) $OA_D = \{(t, A, u) \in T \times A_0 \times T \mid \Gamma = [t, u]A\}$, and (ii) $\text{DAS}(\Delta) = \{(t, da_\varphi, t^+) \in T \times DA \times T \mid \exists [r, v]A \varphi \notin \Delta$ and there exists no action $A \in A_0$ such that the following hold: $\frac{\Phi, \Gamma}{\Delta} \models \forall [r, v]A \varphi \rightarrow \text{Occ}^\Delta \in \mathcal{R}$, and $v = t^+$, and $\Gamma \cup \Delta \models \{\Phi, [r, v]A, [r, v]A\varphi \}$.

4.1 The frame problem

Initially we address the frame problem in a simple setting viz. without qualifications, but will lift the restrictions later.

**Definition 18** Let $D = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}\mathcal{B}, \Gamma \rangle$ be a domain description. $D$ is a simple domain description, or $S$-domain, iff $\mathcal{R}_Q = \emptyset$ and $AQ$ does not occur any where in $\mathcal{R}$ or $\Gamma$.

**Definition 19** Let $D = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}\mathcal{B}, \Gamma \rangle$ a domain description. An interpretation $I = \langle h, \phi, \eta, \varepsilon_q, \varepsilon_f \rangle$ is a simple model, or $S$-model, of $D$ iff

1. $I$ is a model of $D$; and
2. $\varepsilon_q (t, AQ_\varphi, u) = T$ for every $(t, AQ_\varphi, u) \in T \times AQ_{\varphi} \times T$.

This effectively isolates the frame problem from the qualification problem. Note also that if $I$ is an $S$-model then $\Delta_{AQ} = \mathcal{A}\mathcal{B}AQ$. A coherent $S$-model is an $S$-model which is coherent.

**Example 1 (continued.)** Let $D_{Y\mathcal{S}} = \langle \mathcal{L}, FR_{Y\mathcal{S}}, \mathcal{A}\mathcal{B}, \Gamma \rangle$ be an $S$-domain formalising the YSP scenario, where $\Gamma = AD_{Y\mathcal{S}} \cup T_{Y\mathcal{S}}$. The following is part of one of the coherent models of $D_{Y\mathcal{S}}$:

$$\{[0, 1]\text{load, } \neg[0]\text{loaded, } [1]\text{loaded, } [0]\text{alive, } [1]\text{alive},$$

$$[1, 2]\text{wait, } [1, 2]\text{loaded}, \neg[2]\text{loaded, } [2]\text{alive},$$

$$[2, 3]\text{shoot, } \neg[3]\text{loaded, } [3]\text{alive}\}$$

which corresponds to one of the anomalous models of this scenario (the one pointed out by Hanks and McDermott).

But it is not desirable to admit the occurrence of an event when there is no evidence for it. Thus we need to minimise the set of action occurrences in a given action theory.

**Definition 20** Let $D$ be an $S$-domain. A coherent $S$-model $I$ of $D$ is a prioritised minimal model (or simply PMM) of $D$ iff there does not exist any coherent $S$-model $I'$ of $D$ such that $\text{Occ}^{I'} \subset \text{Occ}^I$.

Note that the above model-theoretic minimisation policy isn’t based on the frame assumptions. This solution to the frame problem is thus amenable to well-known techniques such as circumscription\(^2\), but we believe an argumentation-theoretic approach is not only more direct but has wider applicability. In order to

\(^2\) in combination with the introduction of occurrences of dummy actions.
provide the connection between the above (model-theoretic) minimisation policy and the (argument-theoretic) notion of plausible sets of assumptions we need to maximise the set of assumptions satisfiable in a PMM.

**Definition 21** Let $D$ be an S-domain. A PMM $I$ of $D$ is a **canonical prioritised minimal model** (or simply CPMM) of $D$ iff there does not exist any PMM $I'$ of $D$ such that $FA^I \subseteq FA^{I'}$.

We now want to see how the account of plausible sets of assumptions connects to this account of minimality.

**Theorem 1** Let $D$ be an S-domain. If $I$ is a CPMM of $D$ then $\Delta_{Q,F}^I$ is plausible.

We now prove that not only can we derive a plausible set of assumptions from a given CPMM but we can also construct CPMMs from a plausible set of assumptions of a given S-domain.

The set of $\Delta$-relativised models of an S-domain $D$ is denoted as $Mod_{\Delta}^S(D)$.

**Observation 2** Let $D$ be an S-domain and $\Delta$ a set of assumptions of $D$. For each $I \in Mod_{\Delta}^S(D)$, $\Delta = \Delta_{Q,F}^I$.

**Theorem 2** Let $D = \langle \mathcal{L}, \mathcal{R}, AB, \Gamma \rangle$ be an S-domain and $\Delta \subseteq AB$. $\Delta$ is plausible iff $Mod_{\Delta}^S(D) \neq \emptyset$ and for each $I \in Mod_{\Delta}^S(D)$, $I$ is a CPMM of $D$.

**Theorem 3** Let $D$ be an S-domain. Furthermore, suppose that $\text{CPMM}(D)$ is the set of CPMMs of $D$ and $\text{Plaus}(D)$ is the set of plausible sets of assumptions of $D$, then $\text{CPMM}(D) = \bigcup_{\Delta \in \text{Plaus}(D)} Mod_{\Delta}^S(D)$.

### 4.2 Solving the qualification problem (in the presence of the frame problem)

The results reported in the previous section are established in a simple setting. If we add the following observation to the theory in example 1: [3] *alive*, i.e. after the *shoot* action, the victim is still alive, then like most existing formalisms, the above account of plausibility would come up with a contradiction. In fact, it would be more reasonable that such a failure is explained as an occurrence of some (dis)qualification. In this section, we remove certain restrictions on the qualifications of actions in order to achieve a more general framework.

There are some subtleties in the way action theories are represented in our proposed assumption-based framework. Note first that there is a potential difficulty if frame assumptions and qualification assumptions are not distinguished, which can be illustrated by a version of the YSP. Consider the following action description:

\[
\begin{align*}
\{[r]alive, [r]\text{FA}_{alive}, [r]\text{loaded}, [r]\text{shoot}, [r]\text{AQ}_{alive}\} & \subseteq \mathcal{R} \\
\{[0]\text{loaded}, [0]alive, [0,1]\text{shoot}\} & \subseteq \Gamma.
\end{align*}
\]

From this, we have (at least) two stable set of assumptions: one contains the frame assumption $[0]\text{FA}_{alive}$ which rejects the qualification assumption $[0,1]\text{AQ}_{\neg alive}$.
and another contains \([0, 1]\)AQ which attacks \([0]\)FA. Only the latter is intuitive in this case but we do not have any explicit criterion to prefer one over another.

Given the presence of several kinds of assumptions, i.e. frame and qualification, we will adopt the following convention: we will write \(Lr_P(\Delta)\) instead of \(\langle Lr(\Delta)\rangle_P\) for \(P \in \{FA, AQ\}\). Since we no longer exclude qualification assumptions from our assumption-based domain descriptions, we will simply refer to assumption-based domain descriptions as Q-domains.

**Definition 22** Let \(D = \langle \mathcal{C}, \mathcal{R}, \mathcal{AB}, \Gamma \rangle\) be a Q-domain. A presumable set of assumptions \(\Delta \subseteq \mathcal{Ab}\) is semi-Q-plausible iff \(L_{\mathcal{F}A}(\Delta)\) is minimal (with respect to set inclusion).

**Definition 23** Let \(D = \langle \mathcal{C}, \mathcal{R}, \mathcal{AB}, \Gamma \rangle\) be a Q-domain. A set of assumptions \(\Delta \subseteq \mathcal{Ab}\) is Q-plausible iff (a) \(\Delta\) is semi-Q-plausible, (b) \(\Delta_{AQ}\) is maximal, i.e. there does not exist any \(\Delta' \subseteq \mathcal{Ab}\) such that \(\Delta'\) is semi-Q-plausible and \(\Delta_{AQ} \subseteq \Delta'_{AQ}\), and (c) \(\Delta_{FA}\) is maximal relative to the above two conditions, i.e. there does not exist any \(\Delta' \subseteq \mathcal{Ab}\) such that \(\Delta'\) satisfies the above two conditions and \(\Delta_{FA} \subseteq \Delta'_{FA}\).

We will now refer to models of a Q-domain as Q-models. A coherent Q-model is a Q-model which is coherent. We minimise the set of action occurrences in coherent Q-models of a given action theory.

**Definition 24** Let \(D\) be a Q-domain. A coherent Q-model \(I\) of \(D\) is a prioritised minimal Q-model (or simply PMQM) of \(D\) iff there does not exist any coherent Q-model \(I'\) of \(D\) such that \(\text{Occ}^{I'} \subset \text{Occ}^I\).

**Definition 25** Let \(D\) be an S-domain. A PMQM \(I\) of \(D\) is a canonical prioritised minimal Q-model (or simply CPMQM) of \(D\) iff (a) it does not exist any PMQM \(I'\) of \(D\) such that \(\text{AQ}^{I'} \subset \text{AQ}^I\), and (b) there does not exist any PMM \(I'\) of \(D\) such that \(\text{FA}^{I'} \subset \text{FA}^I\).

Now we can proceed to results for CPMQMs regarding Q-plausible sets of assumptions which are similar to those for CPMMs regarding plausible sets of assumptions.

**Theorem 4** Let \(D\) be a Q-domain. If \(I\) is a CPMQM of \(D\) then \(\Delta_{QF}\) is Q-plausible.

Similar to the previous section, we now prove that not only can we derive a plausible set of assumptions from a given CPMQM but we can also construct CPMQMs from a plausible set of assumptions of a given domain description. The set of \(\Delta\)-relativised models of a Q-domain \(D\) is denoted as \(\text{Mod}_{\Delta}^Q(D)\).

**Observation 3** Let \(D\) be a Q-domain \(\Delta\) a set of assumptions of \(D\). For each \(I \in \text{Mod}_{\Delta}^Q(D)\), \(\Delta = \Delta_{QF}^I\).

**Theorem 5** Let \(D = \langle \mathcal{C}, \mathcal{R}, \mathcal{AB}, \Gamma \rangle\) be a Q-domain and \(\Delta \subseteq \mathcal{AB}\). \(\Delta\) is Q-plausible iff \(\text{Mod}_{\Delta}^Q(D) \neq \emptyset\) and for each \(I \in \text{Mod}_{\Delta}^Q(D)\), \(I\) is a CPMQM of \(D\).
Theorem 6 Let $D$ be a $Q$-domain. Furthermore, suppose that $CPQM(D)$ is the set of $CPQMs$ of $D$ and $\text{Plaus}^Q(D)$ is the set of $Q$-plausible sets of assumptions of $D$, then $CPQM(D) = \bigcup_{\Delta \in \text{Plaus}^Q(D)} \text{Mod}^Q_\Delta(D)$. □

$Q$-plausible sets of assumptions allow one to overcome scenarios in which expectation failures (or, qualification surprises) arise, e.g., shooting a turkey with a loaded gun and it can be observed that the turkey is still alive. When such surprises arise, the reasoner knows who’s to blame: qualification assumptions. He can then accordingly remove the “guilty” assumptions.

5 Related work

The frame problem has been addressed in numerous research papers formalised under various frameworks for reasoning about actions, including the Situation Calculus (see [17]), the Event Calculus (see [19]), a temporal logic introduced by Sandewall (see [18]), the action language family (see [8]), the Fluent Calculus (see e.g. [20]). Attempts to solve the original version of the qualification problem (in contrast to the narrowed version of this problem as introduced by Ginsberg and Smith [7] and Lin and Reiter [11]) include Kvarnström and Doherty’s work in tackling the qualification problem in a version of the temporal logic introduced by Sandewall. The solution proposed in this work, however, is still largely fragmented from the solution to other problems of reasoning about actions such as the frame and the ramification problems. A more uniform solution to the qualification problem in accordance to other accounts of reasoning about actions is introduced by Thielerscher [21] for the Fluent Calculus. The solution proposed by Thielerscher is based on a monotonic solution to the frame problem. The idea with Thielerscher’s solution to the frame problem is similar to the idea behind the STRIPS problem solver. The fluents that hold in a state will be manipulated by rules that add (resp. delete) certain fluent (literals) from the preceding state in order to obtain the resulting state. On the other hand, the solutions to the ramification problem and the qualification problem rely on the causal expressions. The idea is to exploit the directional characteristic of causal expression to eliminate the unintended models (aka. the anomalous models). The solution to the qualification problem is non-monotonic while the solution to the ramification problem remains monotonic. Thielerscher’s argument in favour of this approach is largely due to the fact that minimisation of abnormalities in the traditional way as originally performed by McCarthy under circumscription [14] leads to anomalous models. However, as pointed out by Baker [1], a clever minimisation policy will overcome the problem. For a more formal analysis of the related issues from a system-theoretic point of view, the reader is referred to Foo et al.’s paper [6].

Our solution is distinctive from the above approaches in the sense that it offers solution to the major problems of reasoning about actions in a uniform manner. With the introduction of explicit assumptions and the use of reasonable arguments, only intended models should emerge and allow the reasoner to arrive at correct conclusions about the dynamic world.
6 Conclusion

We developed a uniform framework for reasoning about action using an argumentation-theoretic approach (more precisely, assumption-based approach) in a series of papers. The present paper is the first of this series in which we have presented how our framework copes with the frame and the qualification problems in a simple setting without indirect effects or domain constraints. We have shown how our framework can be naturally extended to become more and more expressive.

References