The Qualification Problem: A Solution to the Problem of Anomalous Models

MICHAEL THIELSCHER
Department of Computer Science, Dresden University of Technology, 01062 Dresden (Germany)
e-mail: mit@inf.tu-dresden.de

Abstract. Intelligent agents in open environments inevitably face the Qualification Problem: The executability of an action can never be predicted with absolute certainty; unexpected circumstances, albeit unlikely, may at any time prevent the successful performance of an action. Reasoning agents in real-world environments rely on a solution to the Qualification Problem in order to make useful predictions but also to explain and recover from unexpected action failures. Yet the main theoretical result known today in this context is a negative one: While a solution to the Qualification Problem requires to assume away by default abnormal qualifications of actions, straightforward minimization of abnormality falls prey to the production of anomalous models. We present an approach to the Qualification Problem which resolves this anomaly. Anomalous models are shown to arise from ignoring causality, and they are avoided by appealing to just this concept. Our theory builds on the established predicate logic formalism of the Fluent Calculus as a solution to the Frame Problem and to the Ramification Problem in reasoning about actions. The monotonic Fluent Calculus is enhanced by a default theory in order to obtain the nonmonotonic approach called for by the Qualification Problem. The approach has been implemented in an action programming language based on the Fluent Calculus and successfully applied to the high-level control of robots.

1 Introduction

To program software agents and robots equipped with high-level cognitive capabilities is the enterprise of Cognitive Robotics [Lespérance *et al.*, 1994; Shanahan, 1996]. Rooted in the ability to reason—on the basis of a mental world model—about goals and means to achieve them, these cognitive capabilities free agents from pre-defined problem solutions and so are expected to lead to truly autonomous intelligent, artificial agents. An early scientific experiment in this direction was the robot Shakey [Nilsson, 1974], capable of shuffling around regular-shaped toy blocks so as to achieve certain goals like building a particular stack of blocks. This case study showed success insofar as it proved it feasible to build robots which plan ahead and use their plans to actually pursue their goals. On the other hand, Shakey acted in an artificially constrained, closed environment, in which no disturbances from unexpected sources had to be taken into account.

Some years later, a crucial problem with scaling Shakey's success up to open, real-world environments was named the *Qualification Problem* [McCarthy, 1977]. It arises from the fact that in natural environments the successful execution of actions can never be predicted with absolute certainty. Unexpected circumstances, albeit unlikely, may at any time prevent an autonomous agent from performing the intended actions. Planning and acting under this proviso requires the agent to rigorously assume away, by default, all of the numerous possible but unlikely *abnormal qualifications* of his actions, lest the agent is unable to devise plans which are perfectly reasonable although they cannot guarantee success.

Every daily-life action serves as witness to us humans constantly ignoring a raft of possible obstacles to the successful performance of an action. The classical example in the AI literature is planning to start the engine of a car without making sure that there is no potato in the tail pipe, despite the fact that a clogged tail pipe necessarily renders this action impossible.[1] This ignorance *prima facie* is rational since it is simply impossible to verify all preconditions of actions in real-world environments: Aside from the fact that besides a clogged tail pipe there could be lots of other obstacles for starting the car, how can we ensure that after checking the tail pipe it does not become clogged during us walking to the front door and taking a seat prior to turning the ignition key? Hence, while improbable preconditions must not be completely disregarded in an adequate representation of the world, a proposition like "there is no potato in the tail pipe" should not be treated as a 'regular' precondition in the formal specification of the action of starting the engine. For otherwise the reasoning agent is always forced to verify this condition before assuming that the action can be successfully executed. This is the Qualification Problem. Intelligent agents rely on a solution to this problem in order to make useful predications but also to explain and recover from unexpected action failures.[2]

Assuming away unlikely but not impossible qualifications means that if in a certain situation there are hints to the presence of such unexpected qualifications, or if to the surprise of the agent an action actually fails, then the default conclusion should no longer be adhered to. In this respect the entire process is intrinsically nonmonotonic. Consequently, McCarthy proposed circumscription [McCarthy, 1980] as a means to minimize abnormal qualifications [McCarthy, 1986]. However, a severe defect with this approach was discovered soon after [Lifschitz, 1987]. In a nutshell, this so-called problem of anomalous models arises if the successful performance of some action, say *DisableEmission*, brings about a situation which is exceptional in that another action, say *StartEngine*, is blocked. Suppose an agent considers performing *DisableEmission* followed by *StartEngine* in a situation where he has no reason to assume unusual circumstances. Then the agent can reasonably expect that *DisableEmission* will be successful, thus blocking *StartEngine*. Yet with simple minimization of abnormal qualifications in a straightforward axiomatization of this scenario along the line of [McCarthy, 1986] this does not follow. For there exists a minimal but anomalous model where *DisableEmission* is qualified in the first place. This without any actual reason at all except for the exclusively formal argument that assuming an abnormality wrt. *DisableEmission* avoids assuming an abnormality wrt. *StartEngine*. Anomalous models like this leave the logic totally impotent as regards the tasks of prediction and planning.

We recall in more detail the problem with the original approach to the Qualification Problem in the following Section 2, in which we also pin down the reason for the failure. In a nutshell again, it is the lack of a suitable notion of causality: The action *DisableEmission* causes a qualification of action *StartEngine* while no such cause can be given for the suggested qualification of *DisableEmission*. Guided by this insight, we develop a method for coping with the Qualification Problem which overcomes the problem of anomalous models by respecting causality when minimizing abnormality. Our theory builds on the axiomatization technique of the Fluent Calculus, which provides, in classical predicate logic, a solution to the basic Frame Problem [McCarthy and Hayes, 1969] by means of state update axioms [Thielscher, 1999], and to the Ramification Problem [Ginsberg and Smith, 1988a] by means of causal propagation of in-

[1] According to [Ginsberg and Smith, 1988b], this example is also due to McCarthy.

[2] Some authors, e.g., [Ginsberg and Smith, 1988b; Lin and Reiter, 1994; Shanahan, 1997] have narrowed the Qualification Problem to the problem of determining implicit preconditions of actions which derive from state constraints. Yet these are treated as regular preconditions in the above sense, that is, they must provably hold in correct plans. Hence, they are not meant by McCarthy's original conception of the problem, which concerns preconditions that can be assumed true by default.

direct effects [Thielscher, 1997; Thielscher, 1998].[3] The Fluent Calculus is recalled in Section 3. This monotonic solution to two fundamental problems in reasoning about actions is enhanced in Section 4 to account for the Qualification Problem. Minimization of abnormalities is carried out by means of default logic [Reiter, 1980] but with an important difference to the problematic standard minimization as sketched above: We assume away unjustified causes for abnormalities rather than abnormalities themselves. The crucial advantage of this approach is that the standard reasoning techniques for actions and effects apply whenever the only abnormalities that occur are justified. Our theory is shown to thus solve the problem of anomalous models. In Section 5 this basic account of the Qualification Problem is extended by the possibility to specify priorities among abnormal qualifications, by which is aided the search for reasonable explanations in case of unexpected action failure. In Section 6 we introduce into our theory the distinction between strong and weak qualification, where the latter means that performing an action is possible but fails to produce the usual effect [Gelfond *et al.*, 1991]. In Section 7 we further extend our theory so as to also cover accidents, which are non-recurring action failures. Our results are summarized, discussed, and compared to related work in Section 8.

2 The Anomalous Model Problem

2.1 A straightforward approach to the Qualification Problem

We will illustrate and analyze the occurrence of anomalous models in the context of the Qualification Problem with a formalization of a popular dynamic AI environment, the world of blocks. A number of toy blocks of equal size and shape are arranged on a table and can be stacked onto each other by a robot equipped with a gripper. Our model of this domain will, however, differ in a crucial aspect from standard models. We will not rely on the usual assumption that the action of moving a block is always guaranteed with absolute certainty to be both executable and successful provided that the block to be moved and the destination are unobstructed.

Consider, for example, a Situation Calculus [McCarthy, 1963] axiomatization which uses the two situation-dependent properties, or *fluents* as they are called, $On(x, y, s)$ and $Clear(x, s)$, meaning, respectively, that block x is on y in situation s (where y is either another block or the constant $Table$) and that block x is clear in situation s, that is, it is not obstructed by another block. Let $Move(r, u, v, w)$ denote the action of robot r moving block u away from v onto w, and consider the generic predicate $Poss(a, s)$ which shall be true if action a is possible in situation s. Then a typical axiom of the idealized blocks world is the following:[4]

$$Poss(Move(r, u, v, w), s) \equiv u \neq w \wedge v \neq w \wedge Clear(u, s) \wedge On(u, v, s) \wedge Clear(w, s) \qquad (1)$$

Yet this axiom is clearly not true in a real-world realization of the blocks world, with real blocks on a real table and a real robot shuffling the blocks around. For it may of course happen that in a certain situation, say S_{17}, one robot $Robbie$ tries to move a particular block B_9 from the table onto another block B_{34} but fails to accomplish this task although both blocks are free at the time the action is invoked. There are numerous possible reasons for such an unexpected failure: Block B_9 may somehow be stuck to the table, the robot's gripper may be stuck, or the

[3] It should be stressed that we refer exclusively to the narrow technical Frame and Ramification Problem, as opposed to the general problem of justifying assumptions of persistence [Pylyshyn, 1987].

[4] A word on the notation: Predicate and function symbols, including constants, start with a capital letter whereas variables are in lower case, sometimes with sub- or superscripts. Free variables in formulas are assumed universally quantified.

robot itself may just have run short of energy, and so on and so forth.[5] Whatever caused the non-executability of $Move(Robbie, B_9, Table, B_{34})$ in situation S_{17}, the factual observation,

$$\begin{array}{c} Clear(B_9, S_{17}) \wedge On(B_9, Table, S_{17}) \wedge Clear(B_{34}, S_{17}) \wedge B_9 \neq B_{34} \wedge Table \neq B_{34} \wedge \\ \neg Poss(Move(Robbie, B_9, Table, B_{34}), S_{17}) \end{array} \tag{2}$$

is plainly inconsistent in the light of axiom (1).

In order to obtain a realistic representation, the idealized definition of $Poss(Move(r, u, v, w), s)$ needs to be modified. A crucial obstacle towards this end is that it is usually impossible to provide in advance an exhaustive enumeration of all conceivable reasons for a particular instance of the action to turn out non-executable [McCarthy, 1977]. The only way out is to introduce one or more general propositions by which are captured all possible qualifications but which themselves are not assumed exhaustively described. Two such propositions are appropriate for describing the abnormal (strong) qualifications of a robot trying to move a block: The block may not be movable or the robot's gripper does not work (possibly due to a malfunction of the robot itself). Formally, we introduce the binary atoms $\mathsf{Ab}(Movable(x), s)$ and $\mathsf{Ab}(Functioning(Gripper\text{-}of(r)), s)$ representing, respectively, that block x cannot be moved in situation s for some abnormal reason and that the gripper of robot r does not function in situation s in the way it normally does. Then the first step towards coping with the Qualification Problem in the blocks world is to rewrite axiom (1) to,

$$\begin{array}{rl} Poss(Move(r, u, v, w), s) \equiv & u \neq w \wedge v \neq w \wedge Clear(u, s) \wedge On(u, v, s) \wedge Clear(w, s) \wedge \\ & \neg \mathsf{Ab}(Movable(u), s) \wedge \neg \mathsf{Ab}(Functioning(Gripper\text{-}of(r)), s) \end{array} \tag{3}$$

The added preconditions summarize all abnormal qualifications, that is, obstacles which are *a priori* unlikely to happen and therefore need to be assumed away by default in order to jump to the conclusion that the action is possible under normal circumstances. Hence, the extension of Ab should be minimized, e.g., by circumscription [McCarthy, 1980]. Suppose, for example, this is all that is known of the initial situation S_0:

$$Clear(A, S_0) \wedge On(A, Table, S_0) \wedge Clear(B, S_0) \tag{4}$$

Then $\mathrm{CIRC}[(3) \wedge (4); \mathsf{Ab}]$ [6] entails $Poss(Move(Robbie, A, Table, B), S_0)$, tacitly assuming that $A \neq B$ and $Table \neq B$. Under normal circumstances the new formulation thus behaves just like the idealized precondition axiom (1), allowing to predict that actions are normally successful. On the other hand, the realistic account is flexible enough to handle abnormal circumstances. For instance, it is consistent with (3) to make the observation above, (2), stating that unexpectedly our robot was unable to move B_9 onto B_{34} in situation S_{17}. Moreover, if additional knowledge hints at the presence of an abnormal qualification, then the default conclusion no longer applies. Consider, for example, the fluent $GluedToTable(x, s)$ meaning that block x is glued to the table in situation s. This new fluent relates to the existing ones in our axiomatization thus:

$$\begin{array}{l} GluedToTable(x, s) \supset On(x, Table, s) \\ GluedToTable(x, s) \supset \mathsf{Ab}(Movable(x), s) \end{array} \tag{5}$$

[5] All of these circumstances render the action of moving the block physically impossible. Following [Gelfond *et al.*, 1991] we call them *strong* qualifications of the action. A *weak* qualification, on the other hand, occurs when an action can be performed but its execution produces effects other than the expected ones. The gripper may, for instance, accidentally drop the block, or the tower on top of which the block shall be placed may be too instable to carry the additional weight, etc. Formally, a weak qualification of action a in situation s means that $Poss(a, s)$ is true but the usual effects do not materialize. In what follows, we confine ourselves to strong qualifications; the notion of weakly qualified actions is reconsidered in Section 6.

[6] By $\mathrm{CIRC}[\Psi; \mathsf{Ab}]$ we denote the formula which is obtained by circumscribing, in Ψ, predicate Ab with all other predicates allowed to vary [Lifschitz, 1994].

Then CIRC$[(3) \wedge (4) \wedge (5); \mathsf{Ab}]$ still entails $Poss(Move(Robbie, A, Table, B), S_0)$; but if the observation $GluedToTable(A, S_0)$ is added, then it can no longer be concluded that A, now known to be fixed, can be moved.

2.2 The defect of the straightforward approach

So far the basic approach to the Qualification Problem, which McCarthy already anticipated in his seminal paper [McCarthy, 1977] and formalized some 10 years later using his nonmonotonic formalism of circumscription [McCarthy, 1986]. However, the Qualification Problem turned out to resist this straightforward attack [Lifschitz, 1987]. A serious problem turns up as soon as an action is considered which causes abnormal circumstances as regards the executability of another action. Simple minimization of abnormalities then sanctions anomalous models. Take, for example, the action denoted by $GlueToTable(r, x)$ of agent r gluing a block x to the table. Realistically, this action is possible whenever robot r possesses glue, x is clear and on the table, and no abnormal qualification occurs:

$$\begin{array}{l} Poss(GlueToTable(r, x), s) \equiv \\ \quad Has(r, Glue, s) \wedge Clear(x, s) \wedge On(x, Table, s) \wedge \\ \quad \neg\mathsf{Ab}(Movable(x), s) \wedge \neg\mathsf{Ab}(Functioning(Gripper\text{-}of(r)), s) \wedge \neg\mathsf{Ab}(Usable(Glue), s) \end{array} \tag{6}$$

where fluent $Has(r, x, s)$, constant $Glue$, and abnormality atom $\mathsf{Ab}(Usable(x), s)$ bear the obvious meaning. Let the effect of the action be given by this implication:[7]

$$Poss(GlueToTable(r, x), s) \supset GluedToTable(x, Do(GlueToTable(r, x), s)) \tag{7}$$

where the generic function $Do(a, s)$ denotes the situation reached by performing action a in situation s. To see how this domain axiomatization reveals the defect with the straightforward approach to the Qualification Problem, consider the following description of the initial situation:

$$On(A, Table, S_0) \wedge Clear(A, S_0) \wedge Clear(B, S_0) \wedge Has(Robbie, Glue, S_0) \tag{8}$$

This specification does not give any reason for expecting abnormal circumstances in S_0. Therefore it is reasonable to assume that both of the two actions $Move(Robbie, A, Table, B)$ and $GlueToTable(Robbie, A)$ are possible initially since all of the respective 'regular' preconditions are known to be satisfied. Hence, we must expect that when our robot actually uses the glue, then the action's effect, (7), materializes, that is, $GluedToTable(A, S_1)$, where $S_1 = Do(GlueToTable(Robbie, A), S_0)$. In turn, this conclusion leads to the prediction that $\mathsf{Ab}(Movable(A), S_1)$ according to (5), hence $\neg Poss(Move(Robbie, A, Table, B), S_1)$ according to axiom (3). To summarize, in the light of the given information about the initial situation, reasonable predictions are,

1. The robot should succeed if he tried to move block A onto block B in S_0;
2. The robot should also succeed if he tried to glue block A to the table in S_0;
3. The robot should, however, not succeed with moving block A onto block B after having glued A to the table in S_0.

Yet none of these desirable conclusions follows by simple minimization of abnormality, as the following result shows.

[7] We do not presuppose a particular solution to the Frame Problem here since it is irrelevant at this point.

Proposition 1 *Let Σ consist of the axioms (3) and (5)–(8). There exists a minimal (wrt. the set of true instances of* Ab*) model $\mathcal{M}$ of Σ such that*

$$\begin{aligned}\mathcal{M} \models\ & \neg Poss(Move(Robbie, A, Table, B), S_0) \wedge \neg Poss(GlueToTable(Robbie, A), S_0) \\ & \wedge Poss(Move(Robbie, A, Table, B), S_1)\end{aligned}$$

where $S_1 = Do(GlueToTable(Robbie, A), S_0)$.

Proof: Let $\mathcal{M}$ be a model of Σ with $\mathsf{Ab}(Movable(A), S_0)$ the sole true instance of an abnormality predicate.[8] Then $\mathcal{M} \models \neg Poss(Move(Robbie, A, Table, B), S_0)$ according to (3) and $\mathcal{M} \models \neg Poss(GlueToTable(Robbie, A), S_0)$ according to (6). The latter moreover implies that $\mathcal{M}$ can be chosen in such a way that $\mathcal{M} \models Poss(Move(Robbie, A, Table, B), S_1)$.

To show that $\mathcal{M}$ is minimal in the above sense, it suffices to prove that Σ does not admit a model in which Ab is false for all instances. So, suppose such a model $\mathcal{M}'$ existed, then $\mathcal{M}' \models (\forall x)\, \neg\mathsf{Ab}(x, S_0)$; hence, $\mathcal{M}' \models Poss(GlueToTable(Robbie, A), S_0)$ according to (6) and (8). This implies $\mathcal{M}' \models GluedToTable(A, S_1)$ according to (7). From (5) it follows that $\mathcal{M}' \models \mathsf{Ab}(Movable(A), S_1)$, which contradicts the assumption that $\mathcal{M}'$ falsifies all instances of Ab. ■

The ease with which such anomalous models like $\mathcal{M}$ arise, prevents the Qualification Problem from admitting a straightforward solution [Lifschitz, 1987]. This is just as severe as the infamous Yale Shooting counter-example was for attempts to solve the Frame Problem [Hanks and McDermott, 1987].[9] In the above scenario, for instance, there is actually not a single action that is possible in all minimal models, which leaves our robot paralyzed. The Frame Problem being the more fundamental of the two, past research has concentrated on finding solutions which overcome the Yale Shooting problem (as documented, e.g., in [Shanahan, 1997]), and the Qualification Problem stayed in the background. Nowadays, however, more than one satisfactory solution to the Frame Problem exists, each providing a firm basis for reconsidering the Qualification Problem and in particular the problem of anomalous models.[10]

2.3 The defect's cause: ignoring causality

The intended model and the anomalous ones in our example differ in the abnormality instances they consider true: Intended is $\mathsf{Ab}(Movable(A), S_1)$, anomalous is each of $\mathsf{Ab}(Movable(A), S_0)$, $\mathsf{Ab}(Functioning(Gripper\text{-}of(Robbie)), S_0)$, and $\mathsf{Ab}(Usable(Glue), S_0)$. The crucial question to be answered towards an extensive solution to the anomalous model problem is this: Which general principle allows us to distinguish the first abnormality from the others?

The reason for $\mathsf{Ab}(Movable(A), S_1)$ being the expected conclusion reveals when one tries to find explanations for the occurrence of each of the competing abnormal circumstances—explanations which go beyond the exclusively formal argument that there is no model without making true at least one instance of the abnormality predicate. Actually it is only the intended abnormality which admits such an explanation: Successfully gluing a block to the table

[8] Other anomalous models are obtained by taking either $\mathsf{Ab}(Functioning(Gripper\text{-}of(Robbie)), S_0)$ or $\mathsf{Ab}(Usable(Glue), S_0)$, respectively, as the sole abnormality.

[9] The parallels are indeed intriguing: Both problems invalidate the nonmonotonic approaches to the Frame and Qualification Problem proposed in the very same paper [McCarthy, 1986]; both problems reveal the inadequacy of global minimization in that anomalous models are produced; and, though this is not widely known, a variant of the Yale Shooting problem was found independently by the discoverer of the problem of anomalous models in the context of the Qualification Problem [V. Lifschitz, personal communication].

[10] By "satisfactory solutions" to the Frame Problem we mean established predicate logic formalisms which allow to succinctly specify actions without the need to devise a large number of non-effect axioms, and which do not fall prey to the Yale Shooting problem.

is known to produce an effect which naturally brings about the fact that this block can no longer be moved. No such cause can be given for the three anomalous abnormalities. That is to say, while an abnormal qualification of the initial action $GlueToTable(Robbie, A)$ comes out of the blue in the anomalous models, an abnormal qualification of the subsequent action $Move(Robbie, A, Table, B)$, as is claimed in the intended models, is easily explicable. One even tends to not consider the latter truly abnormal since being unable to move a block after having glued it to the table is, after all, what one would normally expect. Here lies the obvious parallel to the Yale Shooting problem: A gun that becomes magically unloaded while waiting deserves being called abnormal, whereas causality explains the death of the turkey if being shot at with a loaded gun.[11]

2.4 The solution: respecting causality

The anomalous models sanctioned by straightforward circumscription illustrate the necessity of a minimization strategy which respects causality. Abnormalities which do not admit a causal explanation should be preferably assumed away. One way of achieving this is to not let abnormal circumstances be subject to minimization if they are the expected effect of some preceding action. This would solve the problem of anomalous models because then caused abnormal qualifications cannot be compensated for by granting other—truly unpredictable—abnormalities. In our blocks world scenario, for instance, the abnormality in the intended model, $\mathsf{Ab}(Movable(A), S_1)$, is an effect of the action $GlueToTable(Robbie, A)$ which has been performed in the preceding situation, S_0. More precisely, the abnormality is obtained as an indirect effect triggered by the direct effect $GluedToTable(A, S_1)$. This is a consequence of the second state constraint in (5), which generally gives rise to the indirect effect that $\mathsf{Ab}(Movable(x), s)$ whenever $GluedToTable(x, s)$ has been caused. Extending the Frame Problem to indirect effects is known as the Ramification Problem [Ginsberg and Smith, 1988a], and hence a solution to the problem of anomalous models presupposes a solution to the latter.

The foregoing analysis sets out a strategy for overcoming the problem of anomalous models in the context of the Qualification Problem: The fact that abnormal circumstances could arise as the normal effect of performing certain actions is accounted for by treating instances of Ab propositions as fluents which can be (indirectly) affected by actions. Abnormal qualifications, once they materialize, persist, again just like ordinary fluents do. If, for instance, the robot initially glues some block to the table, then shuffles around a number of other blocks, and finally goes back to the first block and tries to move it, then it should not come as a surprise if this action still fails. On the other hand, abnormal qualifications are only present in exceptional cases and therefore need to be minimized. This assumption of normality applies in particular to the initial situation, where the agent could not yet have caused an abnormal qualification.

To summarize, an approach to the Qualification Problem which does not fall prey to the generation of anomalous models needs to achieve the following:

1. Abnormal qualifications of actions are assumed not to hold initially and not to arise in

[11] The Yale Shooting problem goes as follows (c.f. [Hanks and McDermott, 1987]): Suppose we call abnormal any change of a proposition's truth value during the execution of an action (as suggested in [McCarthy, 1986]). Given that shooting at a turkey with a loaded gun causes the animal to drop dead, we would expect exactly this to happen when we start with the gun loaded, wait for a moment, and then shoot. Yet globally minimizing abnormalities in this example produces a second model where the gun becomes unloaded during the first action, waiting, and the turkey survives. While this magical change of the gun's status is abnormal, the turkey surviving the shot is normal in the above sense (as opposed to the change of its life status in the intended model)—hence, this second model minimizes abnormality as well, though it is obviously wrong from the perspective of causality.

later situations *unless they are caused*.

2. The foregoing assumption regarding uncaused abnormalities is made *prima facie* and therefore applies by default only.

A formal theory that satisfies these needs requires at the very least a solution to the basic Frame Problem and to the Ramification Problem in order to obtain caused abnormalities as indirect effects. Furthermore, a nonmonotonic theory is called for by the desire to assume away certain properties by default. We propose a formal account of the Qualification Problem which adds a nonmonotonic theory to the established predicate logic formalism of the Fluent Calculus, to be introduced next. This calculus constitutes an ideal basis for a realization of the above sketch since it provides a monotonic solution to both the Frame and Ramification Problem. Starting out from a monotonic theory is of advantage when integrating a solution to the Qualification Problem because there will be no interference among nonmonotonic rules for inertia and those for qualifications.

3 The Fluent Calculus

3.1 Solving the inferential Frame Problem in pure first-order logic: An informal introduction to the simple Fluent Calculus

The motivation for the development of the Fluent Calculus was to solve not only the representational but also the inferential Frame Problem. While the former means finding a succinct way of specifying all non-effects of actions, the latter means the problem of effectively computing these non-effects [Bibel, 1986; Bibel, 1998]. The inferential Frame Problem arises whenever the value of a fluent in one situation has to be derived from its value in another situation. Apparently, one-by-one and using separate instances of the relevant non-change axioms, every such fluent value needs to be carried stepwise from one situation to the other. This is done, for instance, in the Situation Calculus if successor state axioms are used, no matter whether reasoning is performed forward in time or via regression [Reiter, 1991], and in the Event Calculus where persistence needs to be proven independently for each fluent value [Shanahan, 1997]. The more fluents have to be carried unchanged through many intermediate situations or event occurrences, the more valuable is a solution to the inferential Frame Problem.

With roots in the logic programming formalism of [Hölldobler and Schneeberger, 1990], the Fluent Calculus addresses the inferential Frame Problem by specifying the effects of actions in terms of how an action modifies the state of the environment [Thielscher, 1999]. The notion of a state is therefore central to this axiomatization technique. State terms can be abstract denotations, like the generic $State(s)$ denoting the state of a world in a situation s. On the other hand, each fluent represents a concrete state, namely, the one in which just this fluent holds. Fluents are reified to this end [Quine, 1960], that is, denoted by terms like $On(A, Table)$, where On is a binary function symbol.

State terms, in particular fluents, can be composed to new states with the special binary function "$\circ$". Written in infix notation, this function maps two states into a state in which the fluents of both arguments hold. For example, the term $State(S_0) \circ GluedToTable(A)$ denotes the state which is exactly like the one in the initial situation but where block A is glued to the table. For technical reasons, the Fluent Calculus includes the pre-defined constant $\emptyset$ denoting the empty state, in which—intuitively—no fluent is true.

A fundamental notion is that of a fluent to *hold* in a state. Fluent f holds in state z just in case z can be decomposed into two states one of which is the singleton f. For notational convenience, we introduce the macro $Holds(f, z)$ as an abbreviation for the corresponding equality formula:

$$Holds(f, z) \stackrel{\text{def}}{=} (\exists z')\, z = f \circ z' \tag{9}$$

This fundamental notion of truth and falsity of fluents in states requires a special theory of state terms, by which "$\circ$" is characterized as the union operation with $\emptyset$ as the empty set of fluents (for the formal details see Section 3.3 below). Based on the standard function $State(s)$, a fluent is defined to hold in a situation just in case it holds in the corresponding state:

$$Holds(f, s) \stackrel{\text{def}}{=} Holds(f, State(s)) \tag{10}$$

As an example, suppose that of the initial state in some scenario of our world of toy blocks it is known that block A is on some block x, which in turn stands on the table; that no block y is on top of block A or block B; and that our robot *Robbie* is in possession of glue. In the Fluent Calculus, this incomplete state knowledge can be axiomatized as follows:

$$\begin{array}{l} (\exists x)\,(Holds(On(A, x), S_0) \wedge Holds(On(x, Table), S_0)) \wedge \\ (\forall y)\,(\neg Holds(On(y, A), S_0) \wedge \neg Holds(On(y, B), S_0)) \wedge Holds(Has(Robbie, Glue), S_0) \end{array} \tag{11}$$

With the help of macro definitions (9) and (10) and the foundational axioms, this specification can be transformed into an equivalent formula which specifies what is known about the 'contents' of $State(S_0)$:

$$\begin{array}{l} (\exists x, z)\,[\,State(S_0) = On(A, x) \circ On(x, Table) \circ Has(Robbie, Glue) \circ z \\ \qquad \wedge\ (\forall y)\,(\neg Holds(On(y, A), z)\ \wedge\ \neg Holds(On(y, B), z)\,)\,] \end{array} \tag{12}$$

Put in words, $State(S_0)$ contains $On(A, x)$ and $On(x, Table)$ for some x, $Has(Robbie, Glue)$, and possibly more fluents z—with the restriction that z does not include a fluent $On(y, A)$ nor a fluent $On(y, B)$, of which we know they are false in S_0 for any y.

Based on the notion of states, the Frame Problem is solved by so-called state update axioms, which specify how a state $State(Do(A(\vec{x}), s))$ after performing an action $A(\vec{x})$ relates to the original state $State(s)$ [Thielscher, 1999]. Following the classical STRIPS solution to the inferential Frame Problem [Fikes and Nilsson, 1971], positive effects are modeled by adding them to $State(s)$. This is straightforwardly specified as $State(s) \circ f_1 \circ \ldots \circ f_n$. Negative effects are modeled by removing them from the current state. We denote removal of a fluent by $z - f$. Generalizing STRIPS to incomplete state knowledge, this set operation requires a case distinction if the truth value of f is unknown in z: In case $\neg Holds(f, z)$, we have that $z - f$ is just z, else $z - f = z'$ implies $z' \circ f = z$ and $\neg Holds(f, z')$. A suitable, rigorously first-order axiomatization of removal is thus given by the following definition:

$$z' = z - f \ \stackrel{\text{def}}{=}\ \neg Holds(f, z')\ \wedge\ [z' \circ f = z \vee z' = z] \tag{13}$$

It is easy to see that this macro can be generalized to finitely many negative effects:

$$\begin{array}{rcl} z' = z - \emptyset & \stackrel{\text{def}}{=} & z' = z \\ z' = z - (f_1 \circ \ldots \circ f_n \circ f_{n+1}) & \stackrel{\text{def}}{=} & (\exists z'')\,(z'' = z - (f_1 \circ \ldots \circ f_n) \wedge z' = z'' - f_{n+1}) \end{array} \tag{14}$$

As an example, consider the action $Move(r, u, v, w)$, whose direct effect is that block u is on w and no longer on v:[12]

$$\begin{array}{l} Poss(Move(r, u, v, w), s) \supset \\ \quad State(Do(Move(r, u, v, w), s)) = State(s) \circ On(u, w) - On(u, v) \end{array} \tag{15}$$

[12] For the moment we ignore how the action affects the fluent *Clear* used above. We wish to model this as an indirect effect (see Section 3.4 below).

This state update axiom says that if $Move(r, u, v, w)$ is possible and performed in s, then the new state equals the old state except that $On(u, w)$ becomes true and $On(u, v)$ becomes false. Take, for example, the action of moving block A away from its current location onto block B, and suppose, for the sake of argument, that $(\exists x)\, Poss(Move(Robbie, A, x, B), S_0)$ where $State(S_0)$ is specified by (12). Then the instance $\{r/Robbie, u/A, v/x, w/B, s/S_0\}$ of state update axiom (15) yields

$$State(Do(Move(Robbie, A, x, B), S_0)) = State(S_0) \circ On(A, B) - On(A, x)$$

Let $S_1 = Do(Move(Robbie, A, x, B), S_0)$. Replacing $State(S_0)$ by an equal term according to (12) yields

$$(\exists x, z)\, State(S_1) = On(A, x) \circ On(x, Table) \circ Has(Robbie, Glue) \circ z \circ On(A, B) - On(A, x)$$

Since $On(A, x)$ holds in the state from which it is subtracted, macro definition (13) implies

$$\begin{array}{l} \neg Holds(On(A, x), State(S_1)) \wedge \\ (\exists x, z)\, State(S_1) \circ On(A, x) = On(A, x) \circ On(x, Table) \circ Has(Robbie, Glue) \circ z \circ On(A, B) \end{array}$$

Since $\neg Holds(On(A, x), State(S_1))$ and because state variable z in (12) can be chosen such that $\neg Holds(On(A, x), z)$, fluent $On(A, x)$ can be canceled out on both sides of the equation, which yields

$$\begin{array}{l} (\exists x, z)\, [\, State(S_1) = On(x, Table) \circ Has(Robbie, Glue) \circ On(A, B) \circ z \\ \qquad\qquad \wedge \neg Holds(On(A, x), z)\,] \end{array}$$

We have now obtained from an incomplete initial specification a still partial description of the successor state, which in particular includes the unaffected fluent terms $On(x, Table)$ and $Has(Robbie, Glue)$. These fluents have thus survived the computation of the effect of the action and so need not be carried over by separate axioms. Moreover, knowledge specified in (12) as to which fluents do not hold in z applies to the new state, which includes z, just as well. Thus all unchanged fluent values have been concluded to persist without applying extra inference steps. This is how the Fluent Calculus solves the inferential Frame Problem. Its computational value is crucially dependent on an efficient treatment of equality. While the simple addition of equality axioms may constitute a considerable handicap for theorem proving, a variety of efficient constraint solving algorithms have been developed for the special equational theory needed for the function "$\circ$" (see, e.g., [Pacholski and Podelski, 1997]). An efficient implementation of the Fluent Calculus has recently been developed using constraint logic programming [Thielscher, 2000a].

Next we will introduce the Fluent Calculus formally.

3.2 Fluent Calculus signatures

Fluent Calculus signatures [Thielscher, 1999] can be considered reified versions of standard Situation Calculus signatures Σ, which are many-sorted logic languages with equality and which include the special sorts ACTION and SIT for actions and situations, respectively [Levesque *et al.*, 1998]. Some predicate symbols in Σ are fluent denotations; these are of arity ≥ 1 with the last argument being of sort SIT. The corresponding Fluent Calculus signature is then obtained by

1. replacing each $n+1$-place predicate symbol which denotes a fluent and whose argument is of sort SORTS $\times$ SIT by an n-place function symbol whose argument is of sort SORTS;

2. adding a sort FLUENT to which belong all well-sorted terms with leading function symbol obtained in step 1, and a sort STATE of which FLUENT is a sub-sort;

3. adding the binary function symbol " $\circ$ " of type STATE $\times$ STATE $\mapsto$ STATE and the constant " $\emptyset$ " of sort STATE, which serves as a unit element of $\circ$;

4. adding the unary function *State* of type SIT $\mapsto$ STATE.

In the remainder of this paper, variables of sort ACTION will be denoted by the letter a, variables of sort SIT by s, variables of sort FLUENT by f, and variables of sort STATE by z, all possibly with sub- or superscripts.

3.3 Foundational axioms

Fundamental for any Fluent Calculus axiomatization is a set of equational axioms, denoted $\mathcal{F}_{state}$, which is a suitable subset of the Zermelo-Fraenkel axioms characterizing states as (possibly infinite) collections of fluents with "$\circ$" acting as the union operation and $\emptyset$ as the empty set of fluents:

1. Axioms ACI1 (associativity, commutativity, idempotency, unit element),

$$\begin{aligned} (z_1 \circ z_2) \circ z_3 &= z_1 \circ (z_2 \circ z_3) \\ z_1 \circ z_2 &= z_2 \circ z_1 \\ z \circ z &= z \\ z \circ \emptyset &= z \end{aligned} \tag{16}$$

2. Irreducibility and decomposition,

$$\begin{aligned} &\neg Holds(f, \emptyset) \\ &Holds(f_1, f) \supset f = f_1 \\ &Holds(f, z_1 \circ z_2) \supset Holds(f, z_1) \vee Holds(f, z_2) \end{aligned}$$

3. Equality of states,

$$(\forall f)\,(Holds(f, z_1) \equiv Holds(f, z_2)) \supset z_1 = z_2$$

4. Existence of states,

$$(\forall \Phi)(\exists z)(\forall f)\,(Holds(f, z) \equiv \Phi(f))$$

   where Φ is a second-order predicate variable of sort FLUENT.

The very last one of the axioms above stipulates the existence of a state for all possible combinations of fluents.

3.4 State update axioms

The schema $Poss(A(\vec{x}), s) \supset \Gamma[State(Do(A(\vec{x}), s)), State(s)]$ is the universal form of state update axioms. The form of the update component Γ depends on the ontological assumptions that can be made of the action in question.

The Simple Case

Deterministic actions with only direct and closed effects give rise to the simplest form of state update axioms, where the update is a mere equation relating $State(Do(A(\vec{x}), s))$ to $State(s)$. By closed effects we mean that every action has a maximal, finite number of positive and negative effects. An example of a simple state update axiom is (15) from above.

Under the provision that actions do have only direct and closed effects, simple state update axioms can be fully mechanically generated from a set of Situation Calculus-style effect axioms if the latter can be assumed to give a complete account of the relevant effects of an action. As the primary theorem of the Fluent Calculus it has been proved that a collection of thus generated state update axioms solves the technical Frame Problem, that is, it suitably reflects the basic assumption of persistence [Thielscher, 1999].[13]

Non-deterministic actions can be very elegantly specified by means of *disjunctive* state update axioms $Poss(A(\vec{x}), s) \supset \Gamma[State(Do(A(\vec{x}), s)), State(s)]$ where Γ is a disjunction of the possible effects, i.e., state updates, of the respective action [Thielscher, 2000c]. An example will be shown at the end of Section 4.

State Update Axioms with Ramifications

Only for small domains with very few fluents and actions is it possible and convenient to specify state update axioms in the simple form where all effects are explicitly enumerated. More complex domains require a solution to the Ramification Problem [Ginsberg and Smith, 1988b]. It denotes the problem of handling indirect effects of actions that follow from so-called state constraints, which describe dependencies among fluents. Consider, for example, the following state constraint, which relates the fluent $Clear(x)$ to other fluents in the blocks world:

$$Holds(Clear(x), s) \equiv x = Table \vee \neg(\exists y)\, Holds(On(y, x), s) \tag{17}$$

(This axiom implies, for instance, $Holds(Clear(Table), z)$, $Holds(Clear(A), z)$, and $Holds(Clear(B), z)$ for any z satisfying (12).) This state constraint gives rise to the indirect effect that a block x becomes unclear or clear, respectively, as soon as some other block is moved onto it or away from it. More precisely, if an action is performed with effect $On(y, x)$ for some y, then this action additionally causes $Clear(x)$ to become false whenever $x \neq Table$. Conversely, if an action is performed with effect $\neg On(y, x)$, then $Clear(x)$ becomes true as an indirect effect, provided that nothing else is on x.

In the Fluent Calculus, indirect effects are accounted for by the successive application of so-called causal relationships, which state under what conditions an effect triggers another one [Thielscher, 1997; Thielscher, 1998]. A causal relationship is formally specified with the help of the expression $Causes(\varepsilon, \varrho, z, s)$ where ε (the *triggering effect*) and ϱ (the *ramification*, i.e., indirect effect) are possibly negated atomic fluent formulas and z is a state and s a situation. The intuitive meaning is that the change to ε causes the change to ϱ in state z and situation s.[14] The following two causal relationships, for example, formalize the potential

[13] Actually, in [Thielscher, 1999] a variant of the Fluent Calculus is used where states are axiomatized as multisets of fluents using a slightly different equational foundation [Störr and Thielscher, 2000]. The result can however be straightforwardly adapted to the new axiomatization of Section 3.3, which has been introduced in [Thielscher, 2000a].

[14] The situation argument of *Causes* was not used in the original approach of [Thielscher, 1997]. We need it here because our approach to the Qualification Problem relies on causal relationships which apply only in certain situations.

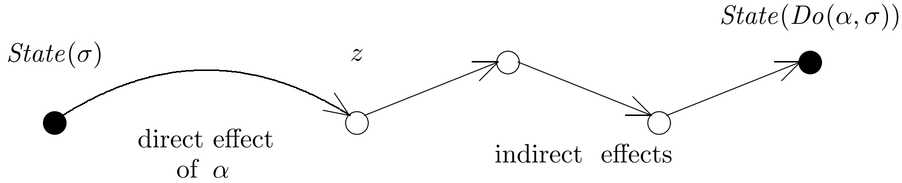

Figure 1: Ramification as causal chains: The result of the direct effect of action α, state z, is the source of a path through several intermediate states, linked by a causal relation. An example would be to move a block in $State(\sigma)$, in which case the block is on its new location in z. The further states are obtained by concluding, in any order, that the old location is now clear and the new one is no longer so (except for the special case of the table). The finally resulting state is always a fixpoint in the sense that no further causal relationship applies.

---

indirect effects on fluent $Clear(x)$:

$$
\begin{array}{l}
x \neq Table \supset Causes(On(y,x), \neg Clear(x), z, s) \\
(\forall y')\, \neg Holds(On(y',x), z) \supset Causes(\neg On(y,x), Clear(x), z, s)
\end{array}
\tag{18}
$$

Put in words, if some y is put onto x which is not the table, then $Clear(x)$ becomes false; conversely, whenever some y is removed from x, then x becomes clear if nothing else is on this block.[15]

On the basis of causal relationships, the Ramification Problem is solved by causally propagating indirect effects: Starting from the direct effects of an action, causal relationships are applied successively. The overall result of performing the action is then a fixpoint of such a chain of indirect effects. Figure 1 gives a schematic illustration of this approach.[16]

The formal axiomatization of causal propagation in the Fluent Calculus is as follows. The above usage of *Causes* being syntactic sugar, the Fluent Calculus for ramifications actually uses the predicate *Causes* with a more complex argument structure:

$$Causes : \text{STATE}^6 \times \text{SIT}$$

An instance $Causes(z, e^+, e^-, z_1, e_1^+, e_1^-, s)$ means that in situation s, if intermediate state z is the result of positive effects e^+ and negative effects e^-, then an additional effect is caused which leads to state z_1 (now the result of positive and negative effects e_1^+, e_1^-, respectively, in which the new effect is additionally recorded).[17] For example,

$$
\begin{array}{l}
Causes(z,\ On(A,B),\ On(A,x), \\
\qquad\qquad z - Clear(B),\ On(A,B),\ On(A,x) \circ Clear(B),\ s)
\end{array}
\tag{19}
$$

says that if a state z is the result of positive effect $On(A,B)$ and negative effect $On(A,x)$, then this causes $Clear(B)$ to become false in z as an additional negative effect.

[15] In [Thielscher, 1997] we have proposed a method for the automatic generation of causal relationships from a set of state constraints plus domain-dependent knowledge of 'causal influence.'

[16] For the sake of simplicity, here and in what follows we ignore the distinction between steady and stabilizing indirect effects introduced and argued for in [Thielscher, 1998].

[17] While formally collections of effects such as e^+ and e^- are terms of sort STATE, they should not be viewed as corresponding to an actual complete state of the world. In what follows, all variables e with sub- or superscripts are of sort STATE.

The macro $\mathit{Causes}(\varepsilon, \varrho, z, s)$ is defined in terms of the 7-ary Causes by distinguishing between positive and negative causes and effects:

$$\begin{array}{rcl}
\mathit{Causes}(f, f', z, s) & \stackrel{\text{def}}{=} & (\forall e^+, e^-)\, \mathit{Causes}(z, e^+ \circ f, e^-, z \circ f', e^+ \circ f \circ f', e^- - f', s) \\
\mathit{Causes}(f, \neg f', z, s) & \stackrel{\text{def}}{=} & (\forall e^+, e^-)\, \mathit{Causes}(z, e^+ \circ f, e^-, z - f', e^+ \circ f - f', e^- \circ f', s) \\
\mathit{Causes}(\neg f, f', z, s) & \stackrel{\text{def}}{=} & (\forall e^+, e^-)\, \mathit{Causes}(z, e^+, e^- \circ f, z \circ f', e^+ \circ f', e^- \circ f - f', s) \\
\mathit{Causes}(\neg f, \neg f', z, s) & \stackrel{\text{def}}{=} & (\forall e^+, e^-)\, \mathit{Causes}(z, e^+, e^- \circ f, z - f', e^+ - f', e^- \circ f \circ f', s)
\end{array} \tag{20}$$

The reader may notice how the 'momentum,' that is, the collections of positive and negative effects, is guaranteed to remain consistent: If necessary, a newly established positive (resp. negative) indirect effect is subtracted from the preceding negative (resp. positive) effects. With this definition, (19) follows from the first causal relationship in (18), given that $B \neq \mathit{Table}$.

Based on a specification of the causal relationships of a domain, the general form of updates which account for indirect effects is as follows:

$$z = \mathit{State}(s) \circ \vartheta^+ - \vartheta^- \supset \mathit{Ramify}(z, \vartheta^+, \vartheta^-, \mathit{Do}(A(\vec{x}), s)) \tag{21}$$

where

- ϑ^+ are the direct positive effects of action $A(\vec{x})$;
- ϑ^- are the direct negative effects of action $A(\vec{x})$;
- $\mathit{Ramify}(z, e^+, e^-, s)$ means that $\mathit{State}(s)$ is a fixpoint of iteratively applying causal relationships to state z and effects e^+, e^- in situation s:

  $$\mathit{Ramify}(z, e^+, e^-, s) \stackrel{\text{def}}{=} (\exists z_1, e_1^+, e_1^-)\, (\mathit{State}(s) = z_1 \wedge (z, e^+, e^-, z_1, e_1^+, e_1^-, s) \in \mu[\mathit{Causes}])$$

  where $(\vec{x}, \vec{y}, s) \in \mu[P]$ abbreviates the following formula, which is a standard second-order schema to axiomatize that $(\vec{x}, \vec{y}, s)$ belongs to the reflexive and transitive closure of predicate P with $\vec{y}$ being a fixpoint:

  $$\begin{array}{l}
  (\forall \Phi)\, \{ (\forall \vec{u})\, \Phi(\vec{u}, \vec{u}, s) \wedge (\forall \vec{u}, \vec{v}, \vec{w})\, [\, \Phi(\vec{u}, \vec{v}, s) \wedge P(\vec{v}, \vec{w}, s) \supset \Phi(\vec{u}, \vec{w}, s)\,] \;\supset\; \Phi(\vec{x}, \vec{y}, s) \} \\
  \wedge\ (\forall \vec{z})\, (P(\vec{y}, \vec{z}, s) \supset \vec{y} = \vec{z})
  \end{array}$$

The use of its transitive closure presupposes the underlying Causes relation to be completely specified. To this end, we circumscribe this predicate wrt. a given axiomatization of cause-effect pairs. If Causes occurs only as the single consequent of implications, like in (18), then second-order circumscription is equivalent to first-order completion [Lifschitz, 1994].

Consider, for example, the following state update axiom, which replaces the preliminary one, (15), in the light of indirect effects:

$$\begin{array}{l}
\mathit{Poss}(\mathit{Move}(r, u, v, w), s) \supset \\
\quad z = \mathit{State}(s) \circ \mathit{On}(u, w) - \mathit{On}(u, v) \supset \\
\qquad \mathit{Ramify}(z, \mathit{On}(u, w), \mathit{On}(u, v), \mathit{Do}(\mathit{Move}(r, u, v, w), s))
\end{array} \tag{22}$$

Suppose further that $(\exists x)\, \mathit{Poss}(\mathit{Move}(\mathit{Robbie}, A, x, B), S_0)$ be given. State update axiom (22) then implies

$$\begin{array}{l}
(\exists x)(\forall z)\, [z = \mathit{State}(S_0) \circ \mathit{On}(A, B) - \mathit{On}(A, x) \supset \\
\qquad \mathit{Ramify}(z, \mathit{On}(A, B), \mathit{On}(A, x), \mathit{Do}(\mathit{Move}(\mathit{Robbie}, A, x, B), S_0))\,]
\end{array}$$

Given $B \neq Table$, the circumscribed causal relationships, CIRC[(18); *Causes*], along with the definition of *Ramify* entail

$$(\exists x)\,(\neg Holds(Clear(B), S_1) \wedge Holds(Clear(x), S_1))$$

where $S_1 = Do(Move(Robbie, A, x, B), S_0)$.

A justification for this approach to the problem of indirect effects is given by [Sandewall, 1996], where the general concept of causal propagation has been proposed as the formal foundation for the Ramification Problem. Moreover, in a series of papers, [Peppas *et al.*, 1999; Prokopenko *et al.*, 1999; Prokopenko *et al.*, 2000], a unifying semantics is defined for a variety of approaches to the Ramification Problem, including our concept of causal relationships.

4 Qualification in the Fluent Calculus: The Basic Approach

The Fluent Calculus and its solution to the Frame and Ramification Problems shall now be extended so as to additionally address the Qualification Problem and in particular solve the problem of anomalous models. Our formal approach follows the guidelines proposed at the end of Section 2.4.

4.1 Abnormality fluents

The first step towards overcoming anomalous models is to introduce abnormality predicates as fluents so that we can appeal to ramification in order to account for abnormal qualifications which are caused by an action. To this end, the standard Fluent Calculus signatures of Section 3.2 are extended by the binary function $\mathsf{Ab}(x, y)$ whose range is the sort FLUENT. As before, the first argument, x, denotes properties like $Movable(u)$ or $Functioning(Gripper\text{-}of(r))$. The second argument, y, indicates the cause for the abnormality; e.g., fluent $\mathsf{Ab}(Movable(A), Glued)$ shall denote the abnormality that block A is not movable on account of it being glued to the table. Instances of the generic 'abnormality' fluent may occur in state constraints and, hence, in causal relationships, which then define how abnormalities could arise as indirect effects. For our blocks world formalization, for example, we introduce the following causal relationship because of the second state constraint in (5):[18]

$$Causes(GluedToTable(x), \mathsf{Ab}(Movable(x), Glued), z, s) \tag{23}$$

That is to say, whenever some block x gets glued to the table, then this causes a qualification of any action which requires x to be movable. Conversely, if the block is somehow freed, then the abnormality disappears together with its cause:

$$Causes(\neg GluedToTable(x), \neg\mathsf{Ab}(Movable(x), Glued), z, s) \tag{24}$$

It may of course happen that $\mathsf{Ab}(Movable(x), y)$ also holds for some y other than *Glued*. Such a fluent would not be affected if (24) were applied and hence x would continue to be immovable. For convenience, we use the macros $\mathsf{Ab}(x, z)$ and $\mathsf{Ab}(x, s)$ to represent that for some y, $\mathsf{Ab}(x, y)$ holds in state z resp. situation s:

$$\begin{aligned} \mathsf{Ab}(x, z) &\stackrel{\text{def}}{=} (\exists y)\, Holds(\mathsf{Ab}(x, y), z) \\ \mathsf{Ab}(x, s) &\stackrel{\text{def}}{=} \mathsf{Ab}(x, State(s)) \end{aligned} \tag{25}$$

[18] For the Fluent Calculus these state constraints need of course be rewritten using the *Holds* expression; see axioms (31) below.

4.2 Arising of abnormal qualifications

The Qualification Problem arises because in the real world any abnormality may at any time arise without being caused by the reasoning agent himself. This aspect is formally captured by allowing instances of Ab to become true during any situation transition as a side effect of the mere fact that the very transition takes place. So doing requires additional causal relationships, which, as opposed to those just added (c.f. (23)), describe exogenously caused abnormalities. The signature is further extended to this end by the predicates $ExogCaused(f, s)$ and $ExogUncaused(f, s)$. An instance $ExogCaused(\mathsf{Ab}(x, Exog), s)$ indicates that in situation s an abnormality wrt. property x arises due to an exogenous cause; conversely, an instance $ExogUncaused(\mathsf{Ab}(x, Exog), s)$ indicates that in situation s an exogenously caused abnormality wrt. x disappears.[19] The effect of exogenous causes is given by these foundational causal relationships:

$$\begin{array}{l} ExogCaused(\mathsf{Ab}(x, Exog), s) \supset \\ \quad Causes(z, e^+, e^-, z \circ \mathsf{Ab}(x, Exog), e^+ \circ \mathsf{Ab}(x, Exog), e^-, s) \\ ExogUncaused(\mathsf{Ab}(x, Exog), s) \supset \\ \quad Causes(z, e^+, e^-, z - \mathsf{Ab}(x, Exog), e^+, e^- \circ \mathsf{Ab}(x, Exog), s) \end{array} \tag{26}$$

The reader may notice two distinctive properties of these indirect effects. First, they are not conditioned on the preceding effects e^+, e^- because they describe changes that are not triggered by other effects. Second, they are conditioned on the situation s since the unusual arising or disappearance of an abnormal qualification in a particular situation does not imply that the qualification arises or disappears in all other situations as well.

4.3 Minimizing exogenously caused abnormalities

Up to this point our additions to the Fluent Calculus did not affect the monotonicity of the solution to the Frame and Ramification Problem. A nonmonotonic feature, however, is required for completing the basic account of the Qualification Problem. Abnormal qualifications are minimized whenever they are not caused by an action that has been performed. This nonmonotonic behavior is achieved by adding appropriate default rules in the sense of [Reiter, 1980], by which the Fluent Calculus gets embedded into a default theory. Formally, exogenous influence on abnormalities is minimized by default rules of the following form:

$$\begin{array}{l} \delta^+_{Exog}(x, s) = \dfrac{\alpha \ : \ \neg ExogCaused(\mathsf{Ab}(x, Exog), s)}{\neg ExogCaused(\mathsf{Ab}(x, Exog), s)} \\[2ex] \delta^-_{Exog}(x, s) = \dfrac{\alpha \ : \ \neg ExogUncaused(\mathsf{Ab}(x, Exog), s)}{\neg ExogUncaused(\mathsf{Ab}(x, Exog), s)} \end{array} \tag{27}$$

where the so-called prerequisites α can be arbitrary first-order formulas to further condition the general assumption of normality. In all examples that follow, however, we tacitly assume α to be a logical tautology, in which case it is simply omitted. As usual, a default rule with variables represents the set of its (well-sorted) ground instances [Reiter, 1980].

As has been argued in Section 2.4, an accompanying default assumption is needed for abnormalities of any kind in the initial situation. Their minimization is carried out by defaults of the

[19] With the focus on the Qualification Problem, we only let the fluent $\mathsf{Ab}(x, Exog)$ be subject to exogenous causes in this paper. In general, the new predicates $ExogCaused(f, s)$ and $ExogUncaused(f, s)$ can be used to model any kind of exogenous influence on fluents.

following form:

$$\delta_{S_0}(x,y) = \frac{\alpha \,:\, \neg Holds(\mathsf{Ab}(x,y),S_0)}{\neg Holds(\mathsf{Ab}(x,y),S_0)} \tag{28}$$

If, for instance, the observations suggest no abnormalities, then the underlying default theory has a unique extension, which includes $(\forall x)\,\neg\mathsf{Ab}(x,S_0)$. (Recall that $\neg\mathsf{Ab}(x,s)$ means $\neg Holds(\mathsf{Ab}(x,y),s)$ for any y.) On the other hand, if, say, $Holds(GluedToTable(B),S_0)$ is given, then the state constraint $Holds(GluedToTable(x),s) \supset Holds(\mathsf{Ab}(Movable(x),Glued),s)$ implies $\mathsf{Ab}(Movable(B),S_0)$ according to (25).

This completes our basic approach to the Qualification Problem by means of the Fluent Calculus. To summarize, domains are axiomatized as default theories $\Delta = (D,\Sigma)$ where D is a set of default rules of the form (27) or (28), and Σ is a set of Fluent Calculus axioms including a circumscribed causal relation $\mathrm{CIRC}[\Psi; Causes]$ where Ψ includes the foundational relationships (26). The semantics is given by the usual definition of extensions of Δ and the notion of skeptical entailment following [Reiter, 1980].[20] In what follows we prove that we have solved the anomalous model problem of Section 2. We also show why this solution is not limited to deterministic actions.

The anomalous model problem revisited

Let D_{bw} be the default rules (27) and (28), with prerequisite $\alpha = True$, and let Σ_{bw} be the Fluent Calculus theory consisting of the state update axioms,

$$\begin{aligned}
&Poss(Move(r,u,v,w),s) \supset \\
&\quad z = State(s) \circ On(u,w) - On(u,v) \supset \\
&\qquad Ramify(z, On(u,w), On(u,v), Do(Move(r,u,v,w),s)) \\[1ex]
&Poss(GlueToTable(r,x),s) \supset \\
&\quad z = State(s) \circ GluedToTable(x) \supset \\
&\qquad Ramify(z, GluedToTable(x), \emptyset, Do(GlueToTable(r,x),s))
\end{aligned} \tag{29}$$

the action precondition axioms,

$$\begin{aligned}
&Poss(Move(r,u,v,w),s) \equiv \\
&\quad u \neq w \wedge v \neq w \wedge Holds(Clear(u),s) \wedge Holds(On(u,v),s) \wedge Holds(Clear(w),s) \wedge \\
&\quad \neg\mathsf{Ab}(Movable(u),s) \wedge \neg\mathsf{Ab}(Functioning(Gripper\text{-}of(r)),s) \\[1ex]
&Poss(GlueToTable(r,x),s) \equiv \\
&\quad Holds(Has(r,Glue),s) \wedge Holds(Clear(x),s) \wedge Holds(On(x,Table),s) \wedge \\
&\quad \neg\mathsf{Ab}(Movable(x),s) \wedge \neg\mathsf{Ab}(Functioning(Gripper\text{-}of(r)),s) \wedge \neg\mathsf{Ab}(Usable(Glue),s)
\end{aligned} \tag{30}$$

[20] A formula is skeptically entailed just in case it is contained in every extension of a default theory. The concept of an extension in Default Logic underwent a number of improvements regarding arguably undesired features of the original definition (such as non-cumulativity, non-commitment to assumptions, non-existence of extensions, etc.; see, e.g., [Delgrande *et al.*, 1994]). However, the default rules needed for the Qualification Problem are *normal* in the sense of [Reiter, 1980], and none of the aforementioned undesired properties holds for normal default theories.

the state constraints,

$$\begin{array}{l} Holds(Clear(x), s) \equiv x = Table \lor \neg(\exists y)\, Holds(On(y,x), s) \\ Holds(GluedToTable(x), s) \supset Holds(On(x, Table), s) \\ Holds(GluedToTable(x), s) \equiv Holds(\mathsf{Ab}(Movable(x), Glued), s) \\ \neg Holds(On(x,x), s) \\ Holds(On(x,y), s) \land Holds(On(x,y'), s) \supset y = y' \\ Holds(On(y,x), s) \land Holds(On(y',x), s) \supset y = y' \lor x = Table \end{array} \tag{31}$$

the causal relationships $\text{CIRC}[(18) \land (23) \land (24) \land (26); Causes]$, and the unique-name axioms,[21]

$$\begin{array}{l} UNA[A, B, Table] \land UNA[On, Clear, GluedToTable, \mathsf{Ab}] \land \\ UNA[Glued, Exog] \land UNA[Movable, Functioning, Usable] \end{array}$$

plus the foundational axioms $\mathcal{F}_{state}$. We then have the following result.

Proposition 2 *Consider the initial specification,*

$$\begin{array}{l} Holds(On(A, Table), S_0) \land Holds(Clear(A), S_0) \land Holds(Clear(B), S_0) \land \\ Holds(Has(Robbie, Glue), S_0) \end{array} \tag{32}$$

Let $\Delta_{bw} = (D_{bw}, \Sigma_{bw} \cup \{(32)\})$. *Default theory* Δ_{bw} *admits a unique extension, which entails each of*

1. $Poss(Move(Robbie, A, Table, B), S_0)$
2. $Poss(GlueToTable(Robbie, A), S_0)$
3. $\neg Poss(Move(Robbie, A, Table, B), S_1)$

where $S_1 = Do(GlueToTable(Robbie, A), S_0)$.

Proof: Suppose there exists an interpretation $\mathcal{M}$ for the underlying Fluent Calculus signature which is a model of Σ_{bw}, of $(\forall x, s)\, \neg ExogCaused(\mathsf{Ab}(x), s)$ and $(\forall x, s)\, \neg ExogUncaused(\mathsf{Ab}(x), s)$, and of

$$\begin{array}{rcl} State(S_0) & = & On(A, Table) \circ Clear(A) \circ Clear(B) \circ Has(Robbie, Glue) \\ & & \circ\, On(B, Table) \circ Clear(Table) \end{array} \tag{33}$$

(Note that we have fixed an initial state in which A and B are the only blocks and both are on the table, which is consistent with (32).) Then $\Sigma_{bw} \cup \{(32)\} \cup \{\omega : \frac{:\omega}{\omega} \in D_{bw}\}$ is consistent. Hence, the unique extension E of Δ_{bw} is

$$Th\left[\Sigma_{bw} \cup \{(32)\} \cup \{\omega : \frac{:\omega}{\omega} \in D_{bw}\}\right]$$

where $Th[\Psi]$ denotes the set of logical consequences of the set of formulas Ψ. Precondition axioms (30) in conjunction with (33), the unique-name axioms, and $\mathcal{F}_{state}$ imply that

$$E \models Poss(Move(Robbie, A, Table, B), S_0) \land Poss(GlueToTable(Robbie, A), S_0)$$

[21] For convenience, we adopt from [Baker, 1989] the following notation for sets of equational axioms expressing uniqueness of names: $UNA[h_1, \ldots, h_n] \stackrel{\text{def}}{=} \bigwedge_{i<j} h_i(\vec{x}) \neq h_j(\vec{y}) \land \bigwedge_i [h_i(\vec{x}) = h_i(\vec{y}) \supset \vec{x} = \vec{y}]$.

Moreover, the second state update axiom in (29) in conjunction with causal relationship (23) implies that $E \models \mathit{Holds}(\mathsf{Ab}(\mathit{Movable}(A), \mathit{Glued}), S_1)$. From this and the first one of the precondition axioms (30), it follows that

$$E \models \neg \mathit{Poss}(\mathit{Move}(\mathit{Robbie}, A, \mathit{Table}, B), S_1)$$

It remains to be shown that an interpretation $\mathcal{M}$ exhibiting the abovementioned properties does indeed exist. Starting from a model of the given unique-name assumptions, the foundational axioms, and $\mathrm{CIRC}[(18) \wedge (23) \wedge (24) \wedge (26); \mathit{Causes}]$, such a model $\mathcal{M}$ can be obtained by inductively assigning state terms to each ground expression $\mathit{State}(\sigma)$. The base case is given by (33). This initial state satisfies axiom (32) and the underlying state constraints, (31). For the induction step, suppose $\mathit{State}(\sigma)$ has been assigned a state term τ which satisfies the state constraints, and consider any well-sorted ground instance α of the two actions $\mathit{Move}(r, u, v, w)$ and $\mathit{GlueToTable}(r, x)$, respectively. We distinguish the following cases:

1. If $\{\mathit{State}(s) = \tau\} \cup \{(30), \mathcal{F}_{\mathit{state}}\} \models \neg \mathit{Poss}(\alpha, \sigma)$, then to $\mathit{State}(\mathit{Do}(\alpha, \sigma))$ is assigned τ;
2. else if $\alpha = \mathit{GlueToTable}(\mathit{Robbie}, \zeta)$, then to $\mathit{State}(\mathit{Do}(\alpha, \sigma))$ is assigned the state term $\tau \circ \mathit{GluedToTable}(\zeta) \circ \mathsf{Ab}(\mathit{Movable}(\zeta), \mathit{Glued})$;
3. else if $\alpha = \mathit{Move}(\mathit{Robbie}, \zeta, \eta, \xi)$, then to $\mathit{State}(\mathit{Do}(\alpha, \sigma))$ is assigned
   (a) $\tau \circ \mathit{On}(\zeta, \xi) \circ \mathit{Clear}(\eta) - \mathit{On}(\zeta, \eta)$ if $\xi = \mathit{Table}$,
   (b) $\tau \circ \mathit{On}(\zeta, \xi) \circ \mathit{Clear}(\eta) - (\mathit{On}(\zeta, \eta) \circ \mathit{Clear}(\xi))$ if $\xi \neq \mathit{Table}$.

It is straightforward to verify that this assignment yields a model of the state update axioms (29) and the state constraints (31). ■

The formal proof shows how the anomalous model problem is overcome: Since all defaults are applicable, there is a unique extension, in which reasoning is performed by the standard techniques of the Fluent Calculus. This crucial feature of our general approach applies whenever the only abnormalities that occur are justified. Consider, for example, a slight modification of the anomalous model problem where each of two actions A_1 and A_2 causes an abnormality wrt. the executability of the other one. If A_1 is performed first, then A_2 will no longer be possible, whereas if A_2 is performed first, then A_1 will become qualified. In both these scenarios, all defaults apply and define a unique extension. Again the right conclusion about which action is successful is obtained solely by the reasoning techniques for actions and effects provided by the standard Fluent Calculus.

Non-deterministic actions and qualifications

Our approach to the Qualification Problem does not interfere with non-deterministic information. If an abnormal qualification is among the possible (direct or indirect) effects of a non-deterministic action, then it is not subject to minimization as it has a cause. Therefore, each extension of the underlying default theory includes the possibility that the non-deterministic action brings about the abnormality in question. In this way, uncertain information is treated in the appropriate way, that is, cautiously.

The following elaboration of the blocks world shall illustrate this way of dealing with non-deterministic actions which possibly give rise to abnormal qualifications. Consider the action of temporarily exposing glue to the air. Chances are that so doing may have the effect that the glue becomes unusable. This being a non-deterministic outcome, the action is formally described by the following disjunctive state update axiom [Thielscher, 2000c]:

$$
\begin{array}{l}
Poss(ExposeToAir(r, Glue), s) \supset \\
\quad Ramify(State(s) \circ \mathsf{Ab}(Usable(Glue), Dried), \mathsf{Ab}(Usable(Glue), Dried), \emptyset, \\
\hfill Do(ExposeToAir(r, Glue), s)) \\
\quad \vee \\
\quad Ramify(State(s), \emptyset, \emptyset, Do(ExposeToAir(r, Glue), s))
\end{array}
\tag{34}
$$

Put in words, a possible result of exposing the glue to the air is that it becomes unusable; the alternative result is that nothing changes at all. A robot can perform the new action whenever it has glue and its gripper is functioning:

$$
\begin{array}{l}
Poss(ExposeToAir(r, Glue), s) \equiv \\
\quad Holds(Has(r, Glue), s) \wedge \neg \mathsf{Ab}(Functioning(Gripper\text{-}of(r)), s)
\end{array}
\tag{35}
$$

Given this knowledge, a careful robot who intends to eventually use his glue had better not expose it to the air—despite the fact that the glue does not necessarily become unusable by doing so. In other words, if *Robbie* knows at some stage that he possesses glue and there are no hints at abnormal circumstances, then he can reasonably assume by default that he will be able to glue to the table any block which is clear and stands on the table. Yet it would be unreasonable if our robot relied on the conclusion that this will still be possible after having exposed the glue to the air. The following proposition shows that our account of the Qualification Problem, and in particular the solution to the problem of anomalous models, is correct in this respect.

Proposition 3 *Consider the initial specification,*

$$
Holds(On(A, Table), S_0) \wedge Holds(Clear(A), S_0) \wedge Holds(Has(Robbie, Glue), S_0) \tag{36}
$$

Let $\Sigma_{bw}^{nd} = \Sigma_{bw} \cup \{(34) \wedge (35)\}$ *and* $\Delta_{bw}^{nd} = (D_{bw}, \Sigma_{bw}^{nd} \cup \{(36)\})$*. Default theory* Δ_{bw}^{nd} *admits a unique extension* E*, for which the following holds:*

1. $E \models Poss(GlueToTable(Robbie, A), S_0)$
2. $E \models Poss(ExposeToAir(Robbie, Glue), S_0)$
3. $E \not\models Poss(GlueToTable(Robbie, A), Do(ExposeToAir(Robbie, Glue), S_0))$

Proof: Suppose there exists an interpretation $\mathcal{M}$ for the underlying Fluent Calculus signature which is a model of Σ_{bw}^{nd}, of $(\forall x, s)\, \neg ExogCaused(\mathsf{Ab}(x), s)$ and $(\forall x, s)\, \neg ExogUncaused(\mathsf{Ab}(x), s)$, and of both

$$
State(S_0) = On(A, Table) \circ Clear(A) \circ Has(Robbie, Glue) \circ Clear(Table) \tag{37}
$$

and

$$
\begin{array}{rcl}
State(S_1) & = & On(A, Table) \circ Clear(A) \circ Has(Robbie, Glue) \circ Clear(Table) \\
& & \circ\, \mathsf{Ab}(Usable(Glue), Dried)
\end{array}
\tag{38}
$$

where $S_1 = Do(ExposeToAir(Robbie, Glue), S_0)$. Then $\Sigma_{bw}^{nd} \cup \{(36)\} \cup \{\omega : \frac{:\omega}{\omega} \in D_{bw}\}$ is consistent. Hence, the unique extension E of Δ_{bw}^{nd} is

$$
Th\left[\Sigma_{bw}^{nd} \cup \{(36)\} \cup \{\omega : \frac{:\omega}{\omega} \in D_{bw}\}\right]
$$

The second precondition axiom in (30) in conjunction with (37) and $\mathcal{F}_{state}$ implies that

$$E \models Poss(GlueToTable(Robbie, A), S_0)$$

Likewise, precondition axiom (35) in conjunction with (37) and $\mathcal{F}_{state}$ implies that

$$E \models Poss(ExposeToAir(Robbie, Glue), S_0)$$

Moreover, from equation (38) and the second precondition axiom in (30) in conjunction with $\mathcal{F}_{state}$ it follows that

$$\mathcal{M} \models \neg Poss(GlueToTable(Robbie, A), S_1)$$

Hence, since $\mathcal{M}$ is a model of E,

$$E \not\models Poss(GlueToTable(Robbie, A), S_1)$$

It remains to be shown that an interpretation $\mathcal{M}$ with the abovementioned properties does indeed exist. Starting from a model of the given unique-name assumptions, the foundational axioms, and CIRC[(18) $\wedge$ (23) $\wedge$ (24) $\wedge$ (26); *Causes*], such a model $\mathcal{M}$ can be obtained by inductively assigning state terms to each ground expression $State(\sigma)$. The base case is given by (37). This initial state satisfies axiom (36) and the underlying state constraints, (31). For the induction step, suppose $State(\sigma)$ has been assigned a state term τ which satisfies the state constraints, and consider any well-sorted ground instance α of the three actions $Move(r, u, v, w)$, $GlueToTable(r, x)$, and $ExposeToAir(r, Glue)$, respectively. We distinguish the following cases:

1. If $\{State(\sigma) = \tau\} \cup \{(30), (35), \mathcal{F}_{state}\} \models \neg Poss(\alpha, \sigma)$, then to $State(Do(\alpha, \sigma))$ is assigned τ;

2. else if $\alpha = GlueToTable(Robbie, A)$ then to $State(Do(\alpha, \sigma))$ is assigned the state term $\tau \circ GluedToTable(A) \circ \mathsf{Ab}(Movable(A), Glued)$;

3. else if $\alpha = ExposeToAir(Robbie, Glue)$, then to $State(Do(\alpha, \sigma))$ is assigned the state term $\tau \circ \mathsf{Ab}(Usable(Glue), Dried)$.

It is straightforward to verify that this assignment yields a model of the state update axioms (29) and (34), of the state constraints (31), and of equation (38). ■

5 Explanation and Priority

Our discussion thus far was centered around the challenge raised by the problem of anomalous models in the context of the Qualification Problem. In this section we turn back to the basic issue of surprising encounters of abnormal qualifications of actions; that is, abnormal circumstances for which no cause can be found in the foregoing action sequence. In our approach to the Qualification Problem, once all regular preconditions of an action are satisfied, the default conclusion is made that the action in question can indeed be executed. In case the action surprisingly fails, the axiomatization blocks some instance of a default rule when constructing an extension of the underling default theory. In this way the theory remains consistent as it still admits one or more extensions. Incidentally, the new set of extensions entails possible explanations for the encountered failure. Explanation problems are thus solved deductively, just like prediction and planning problems are.

Suppose, for example, our robot sees block A clear and on the table, and he perceives block B as clear, too. Suppose further that *Robbie*, to his own surprise, fails to move A onto B. Then he deduces that the reason for this abnormal outcome must be that either block A was not movable or his gripper did not function: Take the initial specification,

$$Holds(On(A, Table), S_0) \wedge Holds(Clear(A), S_0) \wedge Holds(Clear(B), S_0) \tag{39}$$

along with the—unexpected—observation,

$$\neg Poss(Move(Robbie, A, Table, B), S_0) \tag{40}$$

The Fluent Calculus theory $\Sigma_{bw} \cup \{(39), (40)\}$ entails

$$\mathsf{Ab}(Movable(A), S_0) \vee \mathsf{Ab}(Functioning(Gripper\text{-}of(Robbie)), S_0)$$

according to the first one of the precondition axioms (30). Hence, $(D_{bw}, \Sigma_{bw} \cup \{(39), (40)\})$ admits two kinds of extensions, in one of which all defaults are applied except for some instance of

$$\delta_{S_0}(Movable(A), y) \tag{41}$$

while the other extension is obtained by applying all defaults except for one instance of

$$\delta_{S_0}(Functioning(Gripper\text{-}of(Robbie)), y) \tag{42}$$

The default theory thus entails[22]

$$\mathsf{Ab}(Movable(A), S_0) \oplus \mathsf{Ab}(Functioning(Gripper\text{-}of(Robbie)), S_0) \tag{43}$$

A similar result is obtained if an abnormal action qualification is observed in later states: If the robot first shuffles around a number of blocks without touching A and then turns to this very block and fails to relocate it, then the default theory entails that an abnormality either concerning $Movable(A)$ or concerning $Functioning(Gripper\text{-}of(Robbie))$ was exogenously caused at some point during the foregoing sequence of actions.

Usually, once a surprising qualification is observed, the underlying default theory gives rise to multiple extensions. Each of them determines a possible explanation for what has happened, which is exemplified by the particular default that has *not* been applied. (In our example above it was either one instance of (41) or one instance of (42) which was not applied.) Obtaining different extensions means that no preference is given to one or more explanations although some may be more likely true than others. Hence the only conclusion supported by all extensions is one possibly big exclusive disjunction of atoms of the form $Holds(\mathsf{Ab}(x, y), S_0)$ or $ExogCaused(\mathsf{Ab}(x, Exog), s)$. (In our example, we obtained formula (43) as conclusion.) Considering equal all explanation attempts might be unsatisfactory insofar as the reasoning agent often needs to know or at least conjecture what went wrong in order to re-plan the intended future course of actions. Now, by their very nature, abnormalities are *a priori* unlikely to happen. Differences among their respective likelihood seem therefore negligible. The designer of an intelligent agent may nonetheless wish to incorporate knowledge of the relative likelihood into the axiomatization in order to help the agent quickly recover from an unexpected failure during the execution of his plan. For example, the toy blocks in a real-world realization of the blocks world are presumably crafted in such a way that they all are movable. So if the robot encounters

[22] Below, by $\psi \oplus \varphi$ we denote exclusive disjunction, that is, the formula $\psi \wedge \neg\varphi \vee \neg\psi \wedge \varphi$.

a problem at some point, then chances are that the problem lies in his gripper rather than in one of the blocks. Domain knowledge of this kind can be of great help to an autonomous agent in guiding him through a possible sea of increasingly unlikely explanations.

An elegant way of expressing degrees of abnormality within our theory is given by the theory of *Prioritized Default Logic* [Brewka, 1994; Rintanen, 1995]. This extension to the classical framework supports the specification of (possibly partial) preference orderings among defaults. On this basis, a reasoner can select those extensions of a default theory which most likely correspond to the actual states of affairs in the world. In what follows, we confine ourselves to the special case of default theories with only prerequisite-free, normal default rules (which are characterized by the scheme $\frac{:\omega}{\omega}$); for the general setting see, e.g., [Brewka and Eiter, 1999].

Definition 4 [Rintanen, 1995] A *prioritized default theory* is a triple $(D, W, <)$ where D and W are as in classical Default Logic and $<$ is a partial ordering on the ground instances of the elements in D.

If E is a closed set of formulas, then a default $\frac{:\omega}{\omega}$ is said to be *applied in* E iff $\omega \in E$. Let $\Delta = (D, W, <)$ be a prioritized default theory, then an extension E of the (standard) default theory (D, W) is a *preferred extension* of Δ iff there is a total ordering $\ll$ extending $<$ such that the following holds for all extensions E' of (D, W) and all defaults $\delta' \in D$: If δ' is applied in $E' \setminus E$, then there is some $\delta \ll \delta'$ which is applied in $E \setminus E'$. ■

If $\delta < \delta'$, then default δ is said to be preferred over δ'. Thus an extension E is preferred iff for all extensions E', the defaults applied in E 'compensate' for each δ' which is applied in E' but not in E. Compensation simply means that there is some default which is applied in E but not in E' and which is preferred over δ' according to the given preference relation. Or rather, since the preference relation itself may be genuinely partial, according to a total extension $\ll$ of $<$.

Let us see how this development of classical default logic provides means to specify and reason with domain-dependent knowledge about the relative likelihood of exogenously caused abnormalities. Recall our set of defaults D_{bw} and consider this preference ordering:[23]

$$\begin{aligned}
\delta_{S_0}(\mathit{Movable}(x), y) &<_{bw} \delta_{S_0}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(r)), y') \\
\delta_{S_0}(\mathit{Movable}(x), y) &<_{bw} \delta^{+}_{\mathit{Exog}}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(r)), s) \\
\delta^{+}_{\mathit{Exog}}(\mathit{Movable}(x), s) &<_{bw} \delta_{S_0}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(r)), y) \\
\delta^{+}_{\mathit{Exog}}(\mathit{Movable}(x), s) &<_{bw} \delta^{+}_{\mathit{Exog}}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(r)), s')
\end{aligned}$$

Put in words, it is even more unlikely, *a priori*, that a block is immovable than that the gripper of a robot does not function. In other words, the latter shall always be the preferred explanation. Notice that nonetheless the ordering is genuinely partial. For there is, for example, no preference as to the situation in which an abnormal malfunction of the gripper arises if several possibilities exist in that respect. The following proposition shows that this formalization allows a reasoning agent to select preferred explanations.

Proposition 5 *Let* $\Delta^{pr}_{bw} = (D_{bw}, \Sigma_{bw} \cup \{(39), (40)\}, <_{bw})$. *Then all preferred extensions* Δ^{pr}_{bw} *entail*

$$\mathsf{Ab}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(\mathit{Robbie})), S_0)$$

Proof: Axioms (39) and (40) in conjunction with the first one of the precondition axioms in (30) imply that

$$\mathsf{Ab}(\mathit{Movable}(A), S_0) \vee \mathsf{Ab}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(\mathit{Robbie})), S_0) \tag{44}$$

[23] The following generic relations shall represent all of their ground instances.

Consequently, the (classical) extensions of $(D_{bw}, \Sigma_{bw} \cup \{(39), (40)\})$ are obtained by either applying all defaults in D_{bw} except for a single instance of $\delta_{S_0}(\mathit{Movable}(A), y)$ or by applying all defaults in D_{bw} except for a single instance of $\delta_{S_0}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(\mathit{Robbie})), y)$. Let $E(y)$ denote the extensions of the first kind and $E'(y')$ the extensions of the second kind. From $\delta_{S_0}(\mathit{Movable}(A), y) <_{bw} \delta_{S_0}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(\mathit{Robbie})), y')$ for all y, y' it follows that none of $E(y)$ is preferred. All extensions $E'(y')$ entail $\neg\mathsf{Ab}(\mathit{Movable}(A), S_0)$; hence, the claim follows by (44). ■

6 Strong vs. Weak Qualifications

The action qualifications we have considered so far were *strong* in the sense that they render an action physically impossible. A block cannot be moved at all if stuck to the table. Or, if the robot's gripper does not function it is impossible even to start gluing a block to the table. Accordingly, all instances of $\mathsf{Ab}(\mathit{Movable}(x), s)$ and $\mathsf{Ab}(\mathit{Functioning}(\mathit{Gripper\text{-}of}(r)), s)$ imply the negation of a corresponding instance of the predicate $\mathit{Poss}(a, s)$.

Actions may also have *weak* qualifications, which occur when the action can be executed but does not result in the expected outcome: Either some of the usual effects do not materialize, or unexpected additional effects are produced, or both. For example, the robot may succeed with grabbing and lifting a block u which, however, is slippery and hence soon slips off the gripper and lands on the table before it reaches the intended destination. If so, the action $\mathit{Move}(r, u, v, w)$ is possible and achieves that $\mathit{On}(u, v)$ becomes false as expected but fails to produce the other usual effect of $\mathit{On}(u, w)$ becoming true. Instead the action results in the unexpected $\mathit{On}(u, \mathit{Table})$.

Though conceptually different from strong qualifications, weak ones can be readily accommodated in our approach to the Qualification Problem as it stands. Additional instances of Ab are used to indicate abnormal circumstances regarding weak action qualifications, like, e.g., $\mathsf{Ab}(\mathit{Transportable}(x), \mathit{Slippery})$ denoting that block x cannot be transported over a longer distance on account of it being slippery. Minimization by means of defaults of the form (27) or (28) is applied in the very same fashion as in case of abnormalities leading to strong qualifications. The crucial formal difference between strong and weak qualifications is that the fluents representing the former occur in action precondition axioms while the fluents representing the latter strengthen the antecedents of state update axioms.

For example, in the light of the possibility that a weak qualification occurs, the current state update axiom for Move should be refined thus:

$$
\begin{array}{l}
\mathit{Poss}(\mathit{Move}(r,u,v,w),s) \supset \\
\quad [\,\neg\mathsf{Ab}(\mathit{Transportable}(u),s) \supset \\
\qquad z = \mathit{State}(s) \circ \mathit{On}(u,w) - \mathit{On}(u,v) \supset \\
\qquad\quad \mathit{Ramify}(z, \mathit{On}(u,w), \mathit{On}(u,v), \mathit{Do}(\mathit{Move}(r,u,v,w),s))\,] \wedge \\
\quad [\,\mathsf{Ab}(\mathit{Transportable}(u),s) \supset \\
\qquad [\,v \neq \mathit{Table} \supset \\
\qquad\quad z = \mathit{State}(s) \circ \mathit{On}(u,\mathit{Table}) - \mathit{On}(u,v) \supset \\
\qquad\qquad \mathit{Ramify}(z, \mathit{On}(u,\mathit{Table}), \mathit{On}(u,v), \mathit{Do}(\mathit{Move}(r,u,v,w),s))\,] \wedge \\
\qquad [\,v = \mathit{Table} \supset \mathit{Ramify}(\mathit{State}(s), \emptyset, \emptyset, \mathit{Do}(\mathit{Move}(r,u,v,w),s))\,]\,]
\end{array}
\tag{45}
$$

Let Σ_{bw}^{wq} be as Σ_{bw} but with the first one of the state update axioms in (29) replaced by axiom (45). Then the robot should conclude that moving a block has all of the usual effects if nothing hints at an abnormal weak qualification. On the other hand, if a Move action was possible but the robot, upon checking the new situation, sees that the block it carried is not at the destination, then it should conclude that a weak qualification occurred and that the block can be found somewhere on the table. The following proposition states that these conclusions are indeed formally supported by our axiomatization.

Proposition 6 *Consider the initial specification,*

$$Holds(On(A,B),S_0) \wedge Holds(Clear(A),S_0) \wedge Holds(Clear(C),S_0) \tag{46}$$

and let $S_1 = Do(Move(Robbie,A,B,C),S_0)$.

1. $(D_{bw}, \Sigma_{bw}^{wq} \cup \{(46)\})$ *admits a unique extension, which entails*

$$Poss(Move(Robbie,A,B,C),S_0) \wedge Holds(On(A,C),S_1)$$

2. $(D_{bw}, \Sigma_{bw}^{wq} \cup \{(46), Poss(Move(Robbie,A,B,C),S_0) \wedge \neg Holds(On(A,C),S_1)\})$ *admits a unique extension, which entails*

$$\mathsf{Ab}(Transportable(A),S_0) \wedge Holds(On(A,Table),S_1)$$

Proof:

1. Let $\Delta_{bw}^{wq} = (D_{bw}, \Sigma_{bw}^{wq} \cup \{(46)\})$. Suppose there exists an interpretation $\mathcal{M}$ which is model of Σ_{bw}^{wq}, of $(\forall x,s)\neg ExogCaused(\mathsf{Ab}(x),s)$ and $(\forall x,s)\neg ExogUncaused(\mathsf{Ab}(x),s)$, and of

$$\begin{aligned} State(S_0) \ = \ & On(A,B) \circ Clear(A) \circ Clear(C) \\ & \circ On(B,Table) \circ On(C,Table) \circ Clear(Table) \end{aligned} \tag{47}$$

   Then $\Sigma_{bw}^{wq} \cup \{(46)\} \cup \{\omega : \frac{:\omega}{\omega} \in D_{bw}\}$ is consistent. Hence, the unique extension E of Δ_{bw}^{wq} is

$$Th\left[\Sigma_{bw}^{wq} \cup \{(46)\} \cup \{\omega : \frac{:\omega}{\omega} \in D_{bw}\}\right]$$

   Precondition axioms (30) in conjunction with (47) and $\mathcal{F}_{state}$ imply that

$$E \models Poss(Move(Robbie,A,B,C),S_0)$$

   Moreover, since $E \models \neg\mathsf{Ab}(Transportable(A),S_0)$, the first part of state update axiom (45) applies to action $Move(Robbie,A,B,C)$ and situation S_0 and entails, in conjunction with CIRC$[(18) \wedge (23) \wedge (24) \wedge (26); Causes]$, that $E \models Holds(On(A,C),S_1)$. The existence of an interpretation $\mathcal{M}$ with the abovementioned properties can be proved along the line of the proof for Proposition 2.

2. Let $\Delta_{bw}^{wq} = (D_{bw}, \Sigma_{bw}^{wq} \cup \{(46), Poss(Move(Robbie,A,B,C),S_0) \wedge \neg Holds(On(A,C),S_1)\})$. From (46) and $Poss(Move(Robbie,A,B,C),S_0) \wedge \neg Holds(On(A,C),S_1)$ along with the contraposition of the first part of state update axiom (45) in conjunction with CIRC$[(18) \wedge (23) \wedge (24) \wedge (26); Causes]$, it follows that $\mathsf{Ab}(Transportable(A),S_0)$. Hence, the second part of state update axiom (45) applies to action $Move(Robbie,A,B,C)$ and situation S_0. In conjunction with the other domain axioms in Σ_{bw}^{wq} it follows that $Holds(On(A,Table),S_1)$. All of these conclusions hold without the need to apply any default rule, hence are contained in all extensions of Δ_{bw}^{wq}. The existence of a unique extension is proved in analogy to the proof above.

■

7 Accidents

Abnormal qualifications in our sense are persistent by default: Once an abnormality has been observed, the agent cannot assume that it will sort itself out. Accordingly, Ab is a fluent, which keeps its value unless it is caused to change (possibly for exogenous reasons). In this section, we extend our theory to also cover the opposite of persistent qualifications, that is, failures which do not normally recur. We call *accidents* this kind of abnormalities. An example may be that a block just accidentally drops off the gripper without being slippery in general. It may then be more reasonable, on encountering an anomaly, to first of all assume an accident and to do the action again instead of searching for a more fundamental (that is, persistent) reason.

Our concept of accidents as a non-recurring phenomenon includes the assumption that the occurrence of an accident cannot be caused by preceding actions. The problem of anomalous models, which concerned caused abnormalities, does therefore not apply to the minimization of accidents. Formally, we extend further the standard Fluent Calculus signature by the generic atom $Accident(a, s)$ to denote that an accident with action a in situation s happens. Accidents are then assumed away by the generic default rule,

$$\delta_{Acc}(a, s) = \frac{\alpha \; : \; \neg Accident(a, s)}{\neg Accident(a, s)} \tag{48}$$

On this basis, actions which can go wrong accidentally are specified by state update axioms which include a specification of the effect in case of an accident. For example, the action of moving a block (c.f. (45)) shall be specified as follows in view of a potential accident:[24]

$$\begin{array}{l}
Poss(Move(r, u, v, w), s) \supset \\
\quad [\, \neg Accident(Move(r, u, v, w), s) \wedge \neg \mathsf{Ab}(\, Transportable(u), s) \supset \\
\qquad z = State(s) \circ On(u, w) - On(u, v) \supset \\
\qquad\quad Ramify(z, On(u, w), On(u, v), Do(Move(r, u, v, w), s))\,] \wedge \\
\quad [\, Accident(Move(r, u, v, w), s) \vee \mathsf{Ab}(\, Transportable(u), s) \supset \\
\qquad [\, v \neq Table \supset \\
\qquad\quad z = State(s) \circ On(u, Table) - On(u, v) \supset \\
\qquad\qquad Ramify(z, On(u, Table), On(u, v), Do(Move(r, u, v, w), s))\,] \wedge \\
\qquad [\, v = Table \supset Ramify(State(s), \emptyset, \emptyset, Do(Move(r, u, v, w), s))\,]\,]
\end{array} \tag{49}$$

The possibility to explain unexpected effects as accidents can help a planning agent quickly recover from an observed action failure. If an accident is the best explanation, then the agent can predict that he will succeed with simply retrying the crucial action. To this end, it should be possible to specify, e.g., that a single accident is more likely than some exogenously caused persistent qualification while two such accidents in a row are less likely. For example, if our robot observes that he has dropped a block while moving it, he should first assume an accident and just try to move the block again. If, however, he fails a second time, then the most reasonable explanation shall be to consider the block not transportable.

Domain knowledge of the relative likelihood of abnormalities and accidents that takes this form goes beyond the expressiveness of Prioritized Default Logic used in Section 5. There, preferences among defaults are defined in isolation. Consequently, if a default has higher priority than another one, this preference holds regardless of which other defaults apply. We therefore

[24] For the sake of simplicity, we assume that an accident with the action *Move* has the deterministic effect that the block ends up somewhere on the table. A non-deterministic state update axiom could be used to specify that the block may also accidentally land on top of any clear block.

generalize Prioritized Default Logic to allow for context-dependent preferences among defaults as follows:

Definition 7 A *set-prioritized default theory* is a triple $\Delta = (D, W, \prec)$ where D and W are as in classical Default Logic and $\prec$ is a partial ordering on the power-set of the ground instances of the elements in D.

An extension E of the (standard) default theory (D, W) is a *preferred extension* of Δ iff there is no extension E' such that $app_D(E' \setminus E) \prec app_D(E \setminus E')$ where $app_D(E)$ denotes the defaults from D that are applied in E. ∎

Put in words, in a set-prioritized default theory a preference relation is specified on sets of defaults. An extension E is preferred just in case there is no extension E' such that the defaults applied in $E' \setminus E$ are given priority over the defaults applied in $E \setminus E'$.

As an example, consider the following set-preference ordering:[25]

$$\begin{array}{l}
\{\delta_{S_0}(\mathit{Transportable}(u), y)\} \prec_{bw} \{\delta_{Acc}(\mathit{Move}(r, u, v, w), s)\} \\
\{\delta^{+}_{Exog}(\mathit{Transportable}(u), s)\} \prec_{bw} \{\delta_{Acc}(\mathit{Move}(r, u, v, w), s')\} \\
\{\delta_{Acc}(\mathit{Move}(r_i, u, v_i, w_i), s_i)\}_{\geq 2} \prec_{bw} \{\delta_{S_0}(\mathit{Transportable}(u), y)\} \\
\{\delta_{Acc}(\mathit{Move}(r_i, u, v_i, w_i), s_i)\}_{\geq 2} \prec_{bw} \{\delta^{+}_{Exog}(\mathit{Transportable}(u), s)\}
\end{array}$$

where $\{\delta_{Acc}(\mathit{Move}(r_i, u, v_i, w_i), s_i)\}_{\geq 2}$ stands for any set of instances of $\delta_{Acc}(\mathit{Move}(r_i, u, v_i, w_i), s_i)$ with two or more elements. Hence, assuming away an initial or exogenously caused abnormality wrt. block u being transportable is preferred over assuming away a single accident when moving u. The preference is reversed in case the observations cannot be explained by a single accident with the same block.

Let Σ^{acc}_{bw} be as Σ^{wq}_{bw} but with state update axiom (45) replaced by (49), and let D^{acc}_{bw} be D_{bw} augmented by (48), with prerequisite $\alpha = \mathit{True}$. The following proposition shows that this domain axiomatization exhibits the intended behavior: After failing to move a block it can be predicted that a retry will be successful, but if the failure repeats, then the block is inferred to not being transportable.

Proposition 8 *Consider the initial specification,*

$$\mathit{Holds}(\mathit{On}(A, B), S_0) \wedge \mathit{Holds}(\mathit{Clear}(A), S_0) \wedge \mathit{Holds}(\mathit{Clear}(C), S_0) \tag{50}$$

and let $S_1 = \mathit{Do}(\mathit{Move}(\mathit{Robbie}, A, B, C), S_0)$ *and* $S_2 = \mathit{Do}(\mathit{Move}(\mathit{Robbie}, A, \mathit{Table}, C), S_1)$.

1. *Consider the observation,*

$$\mathit{Poss}(\mathit{Move}(\mathit{Robbie}, A, B, C), S_0) \wedge \mathit{Holds}(\mathit{On}(A, \mathit{Table}), S_1) \tag{51}$$

Then $\Delta_1 = (D^{acc}_{bw}, \Sigma^{acc}_{bw} \cup \{(50), (51)\}, \prec_{bw})$ *admits a unique preferred extension, which entails*

$$\mathit{Poss}(\mathit{Move}(\mathit{Robbie}, A, \mathit{Table}, C), S_1) \wedge \mathit{Holds}(\mathit{On}(A, C), S_2)$$

2. *Consider the additional observation,*

$$\mathit{Poss}(\mathit{Move}(\mathit{Robbie}, A, \mathit{Table}, C), S_1) \wedge \mathit{Holds}(\mathit{On}(A, \mathit{Table}), S_2) \tag{52}$$

Let $\Delta_2 = (D^{acc}_{bw}, \Sigma^{acc}_{bw} \cup \{(50), (51), (52)\}, \prec_{bw})$, *then all preferred extensions of* Δ_2 *entail*

$$\mathsf{Ab}(\mathit{Transportable}(A), S_0)$$

[25] As above, the following generic relations shall represent all of their ground instances.

Proof:

1. Observation (51) and state update axiom (49) in conjunction with the circumscribed causal relationships in Σ_{bw} imply that

   $$Accident(Move(Robbie, A, B, C), S_0) \lor \mathsf{Ab}(Transportable(A), S_0)$$

   Consequently, the (classical) extensions of Δ_1 are obtained by applying all elements of D_{bw}^{acc} except for either a single instance of $\delta_{S_0}(Transportable(A), y)$ or the single default $\delta_{Acc}(Move(Robbie, A, B, C), S_0)$. Let $E(y)$ denote the extensions of the first kind, then

   $$app_{D_{bw}^{acc}}(E(y)) = D_{bw} \setminus \{\delta_{S_0}(Transportable(A), y)\}$$

   Likewise, let E' denote the extension of the second kind, then

   $$app_{D_{bw}^{acc}}(E') = D_{bw} \setminus \{\delta_{Acc}(Move(Robbie, A, B, C), S_0)\}$$

   From $\{\delta_{S_0}(Transportable(A), y)\} \prec_{bw} \{\delta_{Acc}(Move(Robbie, A, B, C), S_0)\}$ for all y it follows that E' is preferred but none of $E(y)$. Moreover, E' entails

   $$\neg\mathsf{Ab}(x, s) \land [\, s \neq S_0 \supset \neg Accident(a, s)\,]$$

   The claim follows from the precondition axioms in (30) and state update axiom (49).

2. Observations (51) and (52) and state update axiom (49) in conjunction with the circumscribed causal relationships in Σ_{bw} imply that

   $$\begin{array}{l}[\, Accident(Move(Robbie, A, B, C), S_0) \lor \mathsf{Ab}(Transportable(A), S_0)\,] \land \\ [\, Accident(Move(Robbie, A, Table, C), S_1) \lor \mathsf{Ab}(Transportable(A), S_1)\,]\end{array}$$

   Consequently, the (classical) extensions of Δ_2 are $E(y)$ (for some y), E', and E'', defined by

   $$\begin{array}{lll} app_{D_{bw}^{acc}}(E(y)) & = & D_{bw} \setminus \{\delta_{S_0}(Transportable(A), y)\} \\ app_{D_{bw}^{acc}}(E') & = & D_{bw} \setminus \{\delta_{Acc}(Move(Robbie, A, B, C), S_0), \delta_{Exog}^{+}(Transportable(A), S_1)\} \\ app_{D_{bw}^{acc}}(E'') & = & D_{bw} \setminus \{\delta_{Acc}(Move(Robbie, A, B, C), S_0), \\ & & \qquad \delta_{Acc}(Move(Robbie, A, Table, C), S_1)\} \end{array}$$

   Since $\{\delta_{Acc}(Move(Robbie, A, B, C), S_0), \delta_{Acc}(Move(Robbie, A, Table, C), S_1)\}$ is given priority over $\{\delta_{S_0}(Transportable(A), y)\}$ for all y, extension E'' is not preferred. Also, since $\{\delta_{Exog}^{+}(Transportable(A), S_1)\} \prec_{bw} \{\delta_{Acc}(Move(Robbie, A, Table, C), S_1)\}$, extension E' is not preferred. The claim follows because all of the remaining extensions, $E(y)$, entail that $\mathsf{Ab}(Transportable(A), S_0)$.

■

8 Discussion

The problem of anomalous models has been the crucial barrier towards extensive approaches to the Qualification Problem, which in turn constitutes an important theoretical challenge towards the design of artificial intelligent agents for real-world environments. We have proposed a formal account of the Qualification Problem which solves the problem of anomalous models based on an established predicate logic formalism for reasoning about actions. The theory provides the formal foundations for specifying real-world agents capable of making useful predictions as well as explaining and recovering from unexpected action failures. It has been shown how the basic solution can be extended so as to deal with qualitative knowledge of the relative likelihood of the various explanations for abnormal qualifications. Furthermore, it has been illustrated how weak qualifications and accidents can be expressed, that is, unexpected effects and non-recurring action failures, respectively. We have built our theory on the Fluent Calculus as a solution to the Frame and Ramification Problem. As a result we now have a uniform formalism which successfully copes with all three classical problems in reasoning about actions. Moreover, extensions of the Fluent Calculus deal with concurrent actions and continuous change [Thielscher, 2000b; Thielscher, 2001] or with sensing actions [Thielscher, 2000d]. Staying within classical logic, these techniques are compatible with our default rules for modeling abnormal qualifications.

Based on our approach to the Qualification Problem, the logic program developed in [Martin and Thielscher, 2001] copes with the Qualification Problem in the action programming language FLUX (the *Fluent Calculus Executor*) [Thielscher, 2000a]. The system allows to plan under the default assumption that actions succeed as they normally do, and to reason about these assumptions in order to recover from unexpected action failures. The system has been successfully applied to the high-level control of robots.

The focus in this paper has been on the Fluent Calculus as a particular predicate logic formalism. The underlying principles of our theory, however, are sufficiently general to not depend on this choice. The solution to the problem of anomalous models outlined in Section 2.4 rather promises feasible in any other formalism which is sufficiently expressive in that it includes solutions to both the Frame and the Ramification Problem.

Assuming away by default abnormal qualifications of actions is an inherently nonmonotonic process. In [Lifschitz, 1993], a property called *restricted monotonicity* has been claimed generally desirable in theories of actions. A formalism possesses this property if adding observations to a domain description increases the set of entailed observations. However, when being confronted with the Qualification Problem, restricted monotonicity is not desirable, since an unexpected observation should cause the planning agent to withdraw certain normality assumptions. Consequently, our theory does not satisfy this property, thanks to the use of Default Logic.

An alternative to our solution of the problem of anomalous models might be provided by the concept of *chronological ignorance* [Shoham, 1987; Shoham, 1988]. Roughly speaking, the crucial idea is to assume away, by default, abnormal circumstances, and simultaneously to prefer minimization of abnormalities at earlier timepoints.[26] Our approach to the Qualification Problem and minimizing chronologically share the notion of directedness: By minimizing chronologically, one tends to minimize causes rather than effects—which is the right thing to do—simply because in general causes precede effects. It has been shown elsewhere (e.g., [Kautz, 1986; Sandewall, 1993; Stein and Morgenstern, 1994]) that the applicability of chronological minimization is intrinsically restricted to reasoning problems which do not involve indeterminate information, that

[26] This explains the naming: Potential abnormal qualifications are *ignored* whenever possible, and this is done in *chronological* order.

is, non-deterministic actions or incomplete state knowledge. The refined method of prioritized chronological minimization [Bell, 1998; White *et al.*, 1998] aims at overcoming these restrictions. Roughly speaking again, the crucial idea is to chronologically minimize with an additional preference ordering on atoms: First, event occurrences are minimized, then event abnormalities, and finally affectations of fluents. In particular, by minimizing potential affectations instead of actual changes the problem of non-deterministic actions is overcome [Sandewall, 1994]. There are three main conceptual differences between our framework and the theory of [Bell, 1998], which is built on the idea of prioritized chronological minimization. First, a nonmonotonic and temporal variant of Kleene's three-valued logic is used together with a special semantics tailored to chronological minimization. In contrast, our approach builds on the general framework of Default Logic and its standard semantics. Second, the tasks of prediction and explanation require different reasoning mechanisms, namely, deduction vs. abduction, while predicting, planning, and explaining are uniformly dealt with in our theory. Finally, we were interested in building our approach to the Qualification Problem on an existing solution to the technical Frame Problem.

Our account of the Qualification Problem shares with *Motivated Action Theory* (MAT) [Stein and Morgenstern, 1994; Amsterdam, 1991] the insight that an appropriate notion of causality is necessary when assuming away abnormalities. In this framework, occurrences of actions and events are assumed away by default while taking into consideration the possibility that they are caused (or, in other words, *motivated*, hence the name). This minimizing unmotivated events and our minimizing non-caused abnormal qualifications are somehow complementary while based on similar principles. Problems with MAT have been pointed out in [Ortiz, 1999] concerning the applicability to both the explanation problem and the Ramification Problem, due to the fact that effects of unmotivated events are defined as unmotivated, too. This aspect has been improved in *Explanatory Update Theory* (UAT), which combines MAT with a theory of information change to minimize information loss between states [Ortiz, 1999]. Aside from addressing other problems, a fundamental difference to our framework is that UAT solves the Frame Problem on the semantic level via a special-purpose Kripke-style semantics.

Finally, it should be mentioned that we gave emphasis only to the representational aspect of the Qualification Problem. It has been pointed out, e.g., in [Elkan, 1995], that there is also an important computational aspect to this problem. Our analysis in this paper was driven by the problem of anomalous models, which is a purely representational issue, and—to state the obvious—the computational aspect cannot be pursued without an appropriate representation of the problem. The challenge of the computational Qualification Problem is to find a computational model that enables the agent to reason without even *considering* all possible qualifying causes for his actions—unless some piece of knowledge hints at their presence. A way to tune our representation towards the computational aspect is to introduce predicates of the form $\mathsf{Norm}(A(\vec{x}), s)$, meaning that no abnormal qualification of the respective action $A(\vec{x})$ holds in situation s, along with a suitable definition like

$$\begin{aligned}&\mathsf{Norm}(Move(r,u,v,w),s) \equiv \\ &\quad \neg\mathsf{Ab}(Movable(u),s) \wedge \neg\mathsf{Ab}(Functioning(Gripper\text{-}of(r)),s)\end{aligned} \tag{53}$$

With this addition, action precondition axioms need only mention the atomic condition of normality in addition to the regular preconditions, as in

$$\begin{aligned}&Poss(Move(r,u,v,w),s) \equiv \\ &\quad u \neq w \wedge v \neq w \wedge Holds(Clear(u),s) \wedge Holds(On(u,v),s) \wedge Holds(Clear(w),s) \\ &\quad \wedge \mathsf{Norm}(Move(r,u,v,w),s)\end{aligned}$$

The generic default rule,

$$\frac{:\ \mathsf{Norm}(a,s)}{\mathsf{Norm}(a,s)}$$

then allows to jump to the conclusion that a be executable provided all regular preconditions are met. Still, on the other hand, in order that this assumption be justified, its consistency as regards the corresponding definition, like (53), must be guaranteed. In a standard automated deduction system, this in turn involves consideration (and exclusion) of all the potential qualifying abnormal circumstances. A solution to the computational aspect of the Qualification Problem thus requires a different computational model, presumably based on some parallel architecture, by which all related state constraints are ignored unless they are explicitly 'activated' by some piece of information. Although this was not an issue in this paper, the foundations have been laid.

Acknowledgments. The author is grateful to Wolfgang Bibel, Hans-Peter Störr, and the anonymous reviewers of an earlier version of this paper for helpful comments and suggestions, to Yves Martin for several improvements of the theory, and to Erik Sandewall for enlightening remarks on the subject of the paper.

References

[Amsterdam, 1991] J. B. Amsterdam. Temporal reasoning and narrative conventions. In J. F. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 15–21, Cambridge, MA, 1991.

[Baker, 1989] Andrew B. Baker. A simple solution to the Yale Shooting problem. In R. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 11–20, Toronto, Kanada, 1989. Morgan Kaufmann.

[Bell, 1998] John Bell. Chronological minimization and explanation. In *Proceedings of the Symposium on Logical Formalizations of Commonsense Reasoning*, pages 330–347, London, UK, January 1998.

[Bibel, 1986] Wolfgang Bibel. A deductive solution for plan generation. *New Generation Computing*, 4:115–132, 1986.

[Bibel, 1998] Wolfgang Bibel. Let's plan it deductively! *Artificial Intelligence*, 103(1–2):183–208, 1998.

[Brewka and Eiter, 1999] Gerhard Brewka and Thomas Eiter. Prioritizing default logic: Abridged report. In S. Hölldobler, editor, *Intellectics and Computational Logic*, pages 27–45. Kluwer Academic, 1999.

[Brewka, 1994] Gerhard Brewka. Adding priorities and specificity to default logic. In C. MacNish, D. Pearce, and L. M. Pereira, editors, *Proceedings of the European Workshop on Logics in AI (JELIA)*, volume 838 of *LNAI*, pages 247–260, York, UK, September 1994. Springer.

[Delgrande *et al.*, 1994] J. P. Delgrande, Torsten Schaub, and W. K. Jackson. Alternative approaches to default logic. *Artificial Intelligence*, 70(1–2):167–237, 1994.

[Elkan, 1995] Charles Elkan. On solving the qualification problem. In C. Boutilier and M. Goldszmidt, editors, *Extending Theories of Actions: Formal Theory and Practical Applications*, volume SS–95–07 of *AAAI Spring Symposia*, Stanford University, March 1995. AAAI Press.

[Fikes and Nilsson, 1971] Richard E. Fikes and Nils J. Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208, 1971.

[Gelfond *et al.*, 1991] Michael Gelfond, Vladimir Lifschitz, and Arkady Rabinov. What are the limitations of the situation calculus? In S. Boyer, editor, *Automated Reasoning, Essays in Honor of Woody Bledsoe*, pages 167–181. Kluwer Academic, 1991.

[Ginsberg and Smith, 1988a] Matthew L. Ginsberg and David E. Smith. Reasoning about action I: A possible worlds approach. *Artificial Intelligence*, 35:165–195, 1988.

[Ginsberg and Smith, 1988b] Matthew L. Ginsberg and David E. Smith. Reasoning about action II: The qualification problem. *Artificial Intelligence*, 35:311–342, 1988.

[Hanks and McDermott, 1987] Steve Hanks and Drew McDermott. Nonmonotonic logic and temporal projection. *Artificial Intelligence*, 33(3):379–412, 1987.

[Hölldobler and Schneeberger, 1990] Steffen Hölldobler and Josef Schneeberger. A new deductive approach to planning. *New Generation Computing*, 8:225–244, 1990.

[Kautz, 1986] Henry Kautz. The logic of persistence. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, pages 401–405, Philadelphia, PA, August 1986.

[Lespérance *et al.*, 1994] Yves Lespérance, Hector J. Levesque, Fangzhen Lin, D. Marcu, Ray Reiter, and Richard B. Scherl. A logical approach to high-level robot programming—a progress report. In B. Kuipers, editor, *Control of the Physical World by Intelligent Agents, Papers from the AAAI Fall Symposium*, pages 109–119, New Orleans, LA, November 1994.

[Levesque *et al.*, 1998] Hector Levesque, Fiora Pirri, and Ray Reiter. Foundations for a calculus of situations. *Electronic Transactions on Artificial Intelligence*, 3((1–2)):159–178, 1998. URL: `http://www.ep.liu.se/ea/cis/1998/018/`.

[Lifschitz, 1987] Vladimir Lifschitz. Formal theories of action (preliminary report). In J. McDermott, editor, *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 966–972, Milan, Italy, August 1987. Morgan Kaufmann.

[Lifschitz, 1993] Vladimir Lifschitz. Restricted monotonicity. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, pages 432–437, Washington, DC, July 1993.

[Lifschitz, 1994] Vladimir Lifschitz. Circumscription. In D. Gabbay, C. Hogger, and A. Robinson, editors, *Handbook of Logic in AI and Logic Programming*, volume 3, pages 298–352. Oxford University Press, 1994.

[Lin and Reiter, 1994] Fangzhen Lin and Ray Reiter. State constraints revisited. *Journal of Logic and Computation*, 4(5):655–678, 1994.

[Martin and Thielscher, 2001] Yves Martin and Michael Thielscher. Addressing the qualification problem in FLUX. In F. Baader, G. Brewka, and T. Eiter, editors, *Proceedings of the German Annual Conference on Artificial Intelligence (KI)*, Vienna, Austria, September 2001. Springer.

[McCarthy and Hayes, 1969] John McCarthy and Patrick J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, 4:463–502, 1969.

[McCarthy, 1963] John McCarthy. *Situations and Actions and Causal Laws.* Stanford Artificial Intelligence Project, Memo 2, Stanford University, CA, 1963.

[McCarthy, 1977] John McCarthy. Epistemological problems of artificial intelligence. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1038–1044, Cambridge, MA, 1977. MIT Press.

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

[McCarthy, 1986] John McCarthy. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28:89–116, 1986.

[Nilsson, 1974] Nils J. Nilsson. *Shakey the Robot.* SRI Technical Note 323, Stanford Research Institute, CA, 1974.

[Ortiz, 1999] Charles L. Ortiz, Jr. Explanatory update theory: applications of counterfactual reasoning to causation. *Artificial Intelligence*, 108:125–178, 1999.

[Pacholski and Podelski, 1997] Leszek Pacholski and Andreas Podelski. Set constraints: a pearl in research on constraints. In G. Smolka, editor, *Proceedings of the International Conference on Constraint Programming (CP)*, volume 1330 of *LNCS*, pages 549–561. Springer, 1997.

[Peppas *et al.*, 1999] Pavlos Peppas, Maurice Pagnucco, Mikhail Prokopenko, Norman Y. Foo, and Abhaya Nayak. Preferential semantics for causal systems. In T. Dean, editor, *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 118–123. Morgan Kaufmann, Stockholm, Sweden, 1999.

[Prokopenko *et al.*, 1999] Mikhail Prokopenko, Maurice Pagnucco, Pavlos Peppas, and Abhaya Nayak. Causal propagation semantics—a study. In N. Foo, editor, *Proceedings of the Australian Joint Conference on Artificial Intelligence*, volume 1747 of *LNAI*, pages 378–3392, Sydney, Australia, December 1999. Springer.

[Prokopenko *et al.*, 2000] Mikhail Prokopenko, Maurice Pagnucco, Pavlos Peppas, and Abhaya Nayak. A unifying semantics for causal propagation. In *Proceedings of the Pacific Rim International Conference on Artificial Intelligence*, pages 38–48, Melbourne, Australia, August 2000.

[Pylyshyn, 1987] Zenon W. Pylyshyn, editor. *The Robot's Dilemma: The Frame Problem in Artificial Intelligence.* Ablex, Norwood, New Jersey, 1987.

[Quine, 1960] Willard Quine. *Word and Object.* MIT Press, 1960.

[Reiter, 1980] Ray Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[Reiter, 1991] Ray Reiter. The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation*, pages 359–380. Academic Press, 1991.

[Rintanen, 1995] Jussi Rintanen. On specificity in default logic. In C. S. Mellish, editor, *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1474–1479, Montreal, Canada, August 1995. Morgan Kaufmann.

[Sandewall, 1993] Erik Sandewall. Systematic assessment of temporal reasoning methods for use in autonomous systems. In B. Fronhöfer, editor, *Workshop on Reasoning about Action & Change at IJCAI*, pages 21–36, Chambéry, August 1993.

[Sandewall, 1994] Erik Sandewall. *Features and Fluents. The Representation of Knowledge about Dynamical Systems*. Oxford University Press, 1994.

[Sandewall, 1996] Erik Sandewall. Assessments of ramification methods that use static domain constraints. In L. C. Aiello, J. Doyle, and S. Shapiro, editors, *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 99–110, Cambridge, MA, November 1996. Morgan Kaufmann.

[Shanahan, 1996] Murray Shanahan. Robotics and the common sense informatic situation. In W. Wahlster, editor, *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, pages 684–688, Budapest, Hungary, August 1996. John Wiley.

[Shanahan, 1997] Murray Shanahan. *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*. MIT Press, 1997.

[Shoham, 1987] Yoav Shoham. *Reasoning about Change*. MIT Press, 1987.

[Shoham, 1988] Yoav Shoham. Chronological ignorance: Experiments in nonmonotonic temporal reasoning. *Artificial Intelligence*, 36:279–331, 1988.

[Stein and Morgenstern, 1994] Lynn Andrea Stein and Leora Morgenstern. Motivated action theory: A formal theory of causal reasoning. *Artificial Intelligence*, 71:1–42, 1994.

[Störr and Thielscher, 2000] Hans-Peter Störr and Michael Thielscher. A new equational foundation for the fluent calculus. In J. Lloyd et al., editors, *Proceedings of the International Conference on Computational Logic (CL)*, volume 1861 of *LNAI*, pages 733–746, London (UK), July 2000. Springer.

[Thielscher, 1997] Michael Thielscher. Ramification and causality. *Artificial Intelligence*, 89(1–2):317–364, 1997.

[Thielscher, 1998] Michael Thielscher. Reasoning about actions: Steady versus stabilizing state constraints. *Artificial Intelligence*, 104:339–355, 1998.

[Thielscher, 1999] Michael Thielscher. From Situation Calculus to Fluent Calculus: State update axioms as a solution to the inferential frame problem. *Artificial Intelligence*, 111(1–2):277–299, 1999.

[Thielscher, 2000a] Michael Thielscher. The fluent calculus: A specification language for robots with sensors in nondeterministic, concurrent, and ramifying environments. Technical Report CL-2000-01, Artificial Intelligence Institute, Department of Computer Science, Dresden University of Technology, 2000.

[Thielscher, 2000b] Michael Thielscher. Modeling actions with ramifications in nondeterministic, concurrent, and continuous domains—and a case study. In H. Kautz and B. Porter, editors, *Proceedings of the AAAI National Conference on Artificial Intelligence*, pages 497–502, Austin, TX, July 2000. MIT Press.

[Thielscher, 2000c] Michael Thielscher. Nondeterministic actions in the fluent calculus: Disjunctive state update axioms. In S. Hölldobler, editor, *Intellectics and Computational Logic*, pages 327–345. Kluwer Academic, 2000.

[Thielscher, 2000d] Michael Thielscher. Representing the knowledge of a robot. In A. Cohn, F. Giunchiglia, and B. Selman, editors, *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 109–120, Breckenridge, CO, April 2000. Morgan Kaufmann.

[Thielscher, 2001] Michael Thielscher. The concurrent, continuous Fluent Calculus. *Studia Logica*, 67(2), 2001.

[White *et al.*, 1998] Graham White, John Bell, and Wilfrid Hodges. Building models of prediction theories. In A. G. Cohn, L. K. Schubert, and S. C. Shapiro, editors, *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 557–568, Trento, Italy, June 1998. Morgan Kaufmann.