A note on logical description, observational orders and minimum models

Howard Barringer[*] David Rydeheard[†]

October 3, 2008

Abstract

In this note, we present an account of logical descriptions which consist of sets of formulae defining the current state of a system. Models of these descriptions are those that are minimum in an observationally-defined pre-order. We characterise these models and give examples. This provides a mathematical account of various phenomena encountered in logical developments in Artificial Intelligence, and in the foundations of revision-based logic.

1 Introduction

One way of describing the state of systems is to gather together a collection of 'observations' or 'facts' that hold in the state. These 'facts' are formulae, usually closed and often of a simple form. Such descriptions have been advocated and used widely, especially in developing systems in Artificial Intelligence [5], [8], [12], [6].

When a system develops or computes, it transforms one state to another. To describe this, the associated collection of facts is transformed; some facts are removed and new facts are added. This approach to describing state transformation is known as 'revision-based logic' [5].

Models of these state descriptions are not simply models of the formulae. The descriptions are intended to be 'full descriptions' in a sense which we make precise. Independent facts which are not included in a description are intended not to hold. This form of interpretation, related to the notion of 'circumscription' in Artificial Intelligence, is considered in, for example, [7], [4], [9], [11] and [6].

In this note, we present a logical account of observational descriptions and their models. We introduce a notion of minimum models and characterise these. Other authors, for example [4], [9], [11], [6] and [10], have considered orders on models and minimum, or minimal, models in similar contexts, but the logical settings and definitions of order differ from that here.

Mathematically, this account is straightforward. The aim is to provide a simple general account of a widely used method of logical description. We are not aware of an account of this generality in the literature. Here, the closed formulae that serve as 'facts that may be observed' are arbitrary formulae. The form of the logic and that of the axioms is unrestricted (e.g. we do not need, say, 'Horn clauses' or any other special form of axioms or logic). However, classical logic appears to be essential to the account. It is not clear what an intuitionistic version would look like.

The authors have proposed elsewhere ([1], [2]) a logical framework for defining and reasoning about evolvable computational systems using a revision-based logic. It was with a view to understanding the basis of this logical approach, and how it may be automated, that we developed the material in this note.

[*]School of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, UK. email: howard.barringer@manchester.ac.uk

[†]School of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, UK. email: david.rydeheard@manchester.ac.uk

2 Example: Blocks World

For those not familiar with observational descriptions in the form which we consider, an example may be useful. We shall use this example to illustrate definitions and results later.

The example is a form of a 'Blocks World', as introduced in [12]. A Blocks World consists of a finite collection of blocks (which we think of as of the same size) which may be placed upon each other to form towers or may reside upon a table. There is thus a predicate $on(x, y)$ stating that block x is directly on block y. Other predicates may be present, for example, the blocks may be coloured. To simplify the account, we do not consider an explicit table (this would require a typed world of blocks and tables which behave differently, but we do not need this). We consider blocks which are not on anything to be on the table (this is equivalent, in the usual account of a Blocks World, to there being a single table of unlimited capacity).

Let us illustrate a Blocks World state with four blocks A, B, C and D. Consider Figure 1, in which there are two towers of two blocks each. We may describe this state as the following

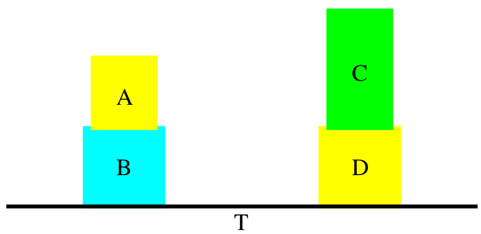

Figure 1: A Blocks World state.

set of 'facts', recording which blocks are on others:

$$\Delta = \{on(A, B), on(C, D)\}.$$

Notice that for this to be a description of this state alone, we need to impose the interpretation that facts not included in Δ, such as $on(B, C)$, do not hold. Of course, for this small example, we could include these 'negative observations' in the state description Δ, so as to circumscribe the state description precisely (as we would do in a standard interpretation). However, in general, what does not hold of a state is a large (often infinite) collection, which we do not wish to specify when the simple state description above may suffice.

Moreover, when states change, we update the state description, removing some facts and adding new ones. For example, suppose that block A is moved to the table, to result in the state depicted in Figure 2. Such a simple transformation can be described by removing the

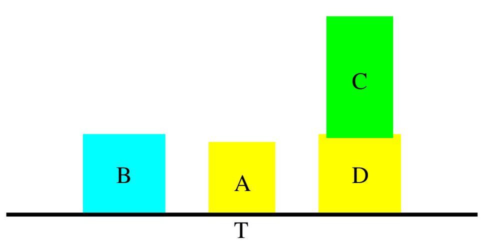

Figure 2: Another blocksworld state.

fact $on(A, B)$ from Δ to result in the state description:

$$\Delta' = \{on(C, D)\}.$$

For a full circumscription of states, this transformation becomes considerably more complex, especially in the presence of a collection of axioms (a 'theory') relating observational facts. For these reasons and others, it has become standard in many applications which use this form of state description to consider the stronger interpretation whereby unstated facts are considered as not holding ('absence as negation').

For this simplified Blocks World, we present a collection of suitable axioms (a Blocks World theory). For an axiomatisation of a fuller account of Blocks World, see [3].

PREDICATES
$$
\begin{aligned}
&\mathit{on} : \mathit{Block} \times \mathit{Block}\\
&\mathit{above} : \mathit{Block} \times \mathit{Block}
\end{aligned}
$$

AXIOMS
$$
\begin{aligned}
&\forall x, x_1, x_2, x_3 : \mathit{Block}.\\
&\qquad \neg \mathit{on}(x, x) \quad \wedge\\
&\qquad \mathit{on}(x, x_1) \wedge \mathit{on}(x, x_2) \Rightarrow (x_1 = x_2) \quad \wedge\\
&\qquad \mathit{on}(x_1, x) \wedge \mathit{on}(x_2, x) \Rightarrow (x_1 = x_2) \quad \wedge\\
&\qquad \mathit{on}(x_1, x_2) \Rightarrow \mathit{above}(x_1, x_2) \quad \wedge\\
&\qquad \mathit{above}(x_1, x_2) \wedge \mathit{above}(x_2, x_3) \Rightarrow \mathit{above}(x_1, x_3) \quad \wedge\\
&\qquad \mathit{above}(x_1, x_2) \Rightarrow \neg \mathit{above}(x_2, x_1)
\end{aligned}
$$

This theory and the soundness and completeness of the associated Blocks World logic will be used in examples below.

There is a simplicity to the above account which may be considered deceptive. The formulae that may be observed are not only of a simple form (namely, ground atomic formulae), but also are independent of each other. In the axiomatic treatment of Blocks World, new facts of this form are not derivable from others. This is important for formulating revision-based state transformations and for automating deduction. However, it is not a fundamental feature of the characterisation of models, as we show in this note, where arbitrary formulae may serve as observations.

3 Models and satisfaction

We consider a typed, first-order (classical) logic $\mathcal{L}$ with formulae $\mathsf{Form}(\mathcal{L})$ built from a signature of function and predicate symbols.

Set-theoretic models are standard:

Definition 3.1 *Let $\mathcal{L}$ be a first-order typed logic. A* model *α of $\mathcal{L}$ is an allocation of a set $\alpha(T)$ for each type T of $\mathcal{L}$, a function $\alpha(f) : \alpha(T_1) \times \cdots \times \alpha(T_n) \to \alpha(T)$ for each function symbol $f : T_1 \times \cdots \times T_n \to T$ of $\mathcal{L}$, and a relation $\alpha(r) \subseteq \alpha(T_1) \times \cdots \times \alpha(T_n)$ for each relation symbol $r : T_1 \times \cdots \times T_n$ of $\mathcal{L}$.*

For logics with equality, the equality symbol is treated as equality of elements in models. Where the logic allows *enumeration types*, these are interpreted strictly: Each declared constant c of an enumeration type T denotes a distinct element of $\alpha(T)$ and this exhausts $\alpha(T)$ i.e. $\forall x \in \alpha(T).\ \exists$ constant $c : T.\ \alpha(c) = x$.

The definition of the interpretation of a formulae $\psi \in \mathsf{Form}(\mathcal{L})$ in a model is standard. A closed first-order formula ψ is satisfied in a model α, written

$$\alpha \models \psi,$$

iff the interpretation of ψ in α is true. We extend this to sets of closed formulae Ψ: $\alpha \models \Psi$ iff for all $\psi \in \Psi$, $\alpha \models \psi$.

For a set of closed formulae Ψ and closed formula ψ, we write

$$\Psi \models \psi$$

iff for all $\mathcal{L}$-models α, $\alpha \models \Psi \implies \alpha \models \psi$.

For first-order typed theory W, we say α is a model of W (or α is a W-model) when, for all $\psi \in W$, $\alpha \models \psi$. We write $\Psi \models_W \psi$ iff for all W-models α, if $\alpha \models \Psi$ then $\alpha \models \psi$. For typed theories with equality, the standard rules for equality (equivalence and substitution) are assumed. For types described as enumeration types, the distinctness and exhaustiveness of the enumerations are assumed in the theory.

4 Observations and minimum models

Let $\mathcal{O}$ be a set of closed $\mathcal{L}$-formulae which we interpret as a collection of 'possible observations or facts' about a model ($\mathcal{O}$ may be empty).

Definition 4.1 *For theory W of $\mathcal{L}$ and $\Delta \subseteq \mathcal{O}$, a W-model α satisfies Δ iff $\alpha \models \Delta$. Let $\mathsf{Mod}_W(\Delta)$ be the set of all W-models that satisfy Δ.*

We define a pre-order $\lesssim$ on $\mathsf{Mod}_W(\Delta)$ by

$$\alpha \lesssim \beta \text{ iff } \forall \varphi \in \mathcal{O}.\ \alpha \models \varphi \implies \beta \models \varphi.$$

We now consider minimum models *in $\mathsf{Mod}_W(\Delta)$ under this pre-order, i.e. models $\alpha \in \mathsf{Mod}_W(\Delta)$ such that for all models $\beta \in \mathsf{Mod}_W(\Delta)$, $\alpha \lesssim \beta$, that is:*

$$\forall \varphi \in \mathcal{O}.\ \alpha \models \varphi \implies \forall \beta \in \mathsf{Mod}_W(\Delta).\ \beta \models \varphi.$$

Minimum models are "observationally equivalent", i.e. for two minimum models α and β,

$$\forall \varphi \in \mathcal{O}.\ \alpha \models \varphi \text{ iff } \beta \models \varphi.$$

However, minimum models need not be isomorphic or logically equivalent.

Minimum models need not exist – it depends on the theory W, the choice of observations $\mathcal{O}$ and the set Δ (see later).

Example 4.1 *Consider the Blocks World theory W with blocks A, B, C, D and observations $\mathcal{O} = \{on(x, y) \mid x, y \in Block\}$.*

For every minimum model α of $\Delta = \{on(A, B), on(C, D)\}$ (see Figure 1), we have

$$\alpha \models \neg on(B, C).$$

Proof. *If not, then $\alpha \models on(B, C)$ (because for all models α and closed formulae ψ, $\alpha \models \psi$ or $\alpha \models \neg\psi$).*

Now, because α is minimum, we have for all $\beta \in \mathsf{Mod}_W(\Delta)$, $\beta \models on(B, C)$ as $on(B, C) \in \mathcal{O}$. But because the interpretation of B is distinct from that of C, there is a model in $\mathsf{Mod}_W(\Delta)$ which does not satisfy $on(B, C)$.

We introduce another satisfaction relation, $\approx\!\!\!\models$, defined using minimum models:

Definition 4.2 *For W an $\mathcal{L}$-theory, $\Delta \subseteq \mathcal{O}$, for a set of closed $\mathcal{L}$-formulae $\mathcal{O}$, and ψ a closed $\mathcal{L}$-formula, write*

$$\Delta \mathrel{|\!\!\approx}_W \psi$$

iff for all α minimum *in $\mathsf{Mod}_W(\Delta)$, $\alpha \models \psi$.*

Note that $\Delta \models_W \psi \implies \Delta \mathrel{|\!\!\approx}_W \psi$.

Example 4.2 *In the Blocks World theory W and state $\Delta = \{on(A, B), on(C, D)\}$, we have $\Delta \mathrel{|\!\!\approx}_W \neg on(B, C)$.*

5 Characterising minimum models

We now consider the existence of minimum models, first beginning with a characterisation theorem.

Theorem 5.1 (Characterising minimum models) *Consider a theory W of $\mathcal{L}$ and $\Delta \subseteq \mathcal{O}$, for a set of closed $\mathcal{L}$-formulae $\mathcal{O}$. Define $\mathcal{T}(\Delta) \subseteq \mathsf{Form}(\mathcal{L})$, by*

$$\mathcal{T}(\Delta) = \Delta \cup \{\neg\varphi \mid \varphi \in \mathcal{O} \text{ and } \neg(\forall \beta \in \mathsf{Mod}_W(\Delta).\ \beta \models \varphi)\}.$$

Then $\alpha \in \mathsf{Mod}_W(\Delta)$ is minimum iff $\alpha \models \mathcal{T}(\Delta)$.

Proof. *If $\alpha \in \mathsf{Mod}_W(\Delta)$ is minimum, then for all $\varphi \in \mathcal{O}$,*

$$\alpha \models \varphi \Leftrightarrow \Delta \models_W \varphi$$

(by definition of minimum). Hence, for $\varphi \in \mathcal{O}$,

$$\Delta \not\models_W \varphi \Rightarrow \alpha \models \neg\varphi.$$

Thus $\alpha \models \mathcal{T}(\Delta)$.

Conversely, suppose $\alpha \models \mathcal{T}(\Delta)$. Then for all $\varphi \in \mathcal{O}$, if $\alpha \models \varphi$ then $\Delta \models_W \varphi$ (since, if not i.e. $\Delta \not\models_W \varphi$ then $\alpha \models \neg\varphi$ – contradiction). Hence, $\forall\beta \in \mathsf{Mod}_W(\Delta)$. $\beta \models \varphi$. Thus α is minimum.

Corollary 5.1 (Existence of minimum models) *Consider a theory W of $\mathcal{L}$ and $\Delta \subseteq \mathcal{O}$, for a set of closed $\mathcal{L}$-formulae $\mathcal{O}$. A minimum model in $\mathsf{Mod}_W(\Delta)$ exists iff $\mathcal{T}(\Delta)$ is W-consistent, i.e. there is a model $\gamma \in \mathsf{Mod}_W(\Delta)$ with $\gamma \models \mathcal{T}(\Delta)$.*

Example 5.1 *Revisiting the Blocks World example, consider an extension so that blocks have colours with, say, just two colours, red and green, with axiom*

$$\forall x : Block.\ red(x) \Leftrightarrow \neg green(x).$$

Let the colours be observable, i.e. for each block A, $red(A), green(A) \in \mathcal{O}$. Consider the case of two blocks A and B only. Let

$$\Delta = \{on(A, B), red(A)\}.$$

Then Δ has no minimum models since the colour of B is not specified. Hence, $\neg red(B) \in \mathcal{T}(\Delta)$ and $\neg green(B) \in \mathcal{T}(\Delta)$ and so $\mathcal{T}(\Delta)$ is inconsistent in the presence of the above axiom. However, if

$$\Delta = \{on(A, B), red(A), green(B)\},$$

then Δ has minimum models as the negations of all observations not following from Δ are consistent with Δ. In particular, $\neg green(A) \wedge \neg red(B)$ is consistent with Δ.

We now characterise the satisfaction relation $\models\!\approx$ in terms of $\models$. This provides a justification for the 'absence as negation' interpretation that may be imposed on observational descriptions.

Corollary 5.2 (Characterising satisfaction for minimum models) *Consider a theory W of $\mathcal{L}$ and $\Delta \subseteq \mathcal{O}$, for a set of closed $\mathcal{L}$-formulae $\mathcal{O}$. Define $\mathcal{T}(\Delta) \subseteq \mathsf{Form}(\mathcal{L})$ as above (Theorem 5.1). Then, for any closed $\psi \in \mathsf{Form}(\mathcal{L})$, we have*

$$\Delta \models\!\approx_W \psi \ \text{ iff } \ \mathcal{T}(\Delta) \models_W \psi.$$

Proof. Assume $\Delta \models\!\approx_W \psi$, i.e. for all $\alpha \in \mathsf{Mod}_W(\Delta)$ minimum, $\alpha \models \psi$. But by Theorem 5.1, α is minimum iff $\alpha \models \mathcal{T}(\Delta)$. Thus, for all $\alpha \in \mathsf{Mod}_W(\Delta)$ with $\alpha \models \mathcal{T}(\Delta)$, $\alpha \models \psi$, i.e. $\mathcal{T}(\Delta) \models_W \psi$.

Conversely, assume $\mathcal{T}(\Delta) \models_W \psi$, i.e. $\forall\alpha \in \mathsf{Mod}_W(\Delta)$. $\alpha \models \mathcal{T}(\Delta) \Rightarrow \alpha \models \psi$. But $\alpha \models \mathcal{T}(\Delta)$ iff α is minimum. Hence $\Delta \models\!\approx_W \psi$ as required.

A note on this account: Other authors, for example [4], [9], [11], [6] and [10], have considered orders on models and minimum, or minimal, models in similar contexts, but the logical settings and definitions of order differ from that here. When restricted to Herbrand models, the minimum models here coincide with minimal models under inclusion. Special cases of the construction of $\mathcal{T}(\Delta)$ from Δ occur in various accounts of 'circumscription' and 'augmentation' in logics for Artificial Intelligence (see, for example, in [4], [9] and [6]).

References

[1] H. Barringer and D. Rydeheard. A logical framework for monitoring and evolving software components. In *First Joint IEEE/IFIP Symposium on Theoretical Aspects of Computer Science (TASE07)*. IEEE Publications, Shanghai, 2007.

[2] H. Barringer, D. Rydeheard, B. C. Warboys, and D. Gabbay. Revision-based logical framework for evolvable software. In *IASTED International Conference on Software Engineering (SE07)*, Innsbruck, 2007.

[3] S. A. Cook and Yongmei Liu. A complete axiomatization for blocks world. *J. Logic and Computation*, 13(4), 2003.

[4] D. W. Etherington, R. E. Mercer, and R. Reiter. On the adequacy of predicate circumscription for closed-world reasoning. *Computational Intelligence*, 1:11–15, 1985.

[5] R. E. Fikes and N. J. Nilsson. Strips: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2(3–4):189–208, 1971.

[6] M. R. Geneseret and N. J. Nilson. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann Publishers, 1987.

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

[8] J. McCarthy and P. J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie, editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, 1969.

[9] D. Perlis and J. Minker. Completeness results for circumscription. *Artificial Intelligence*, 28(1):29–42, 1986.

[10] P. Rondogiannis and W. W. Wadge. Minimum model semantics for logic programs with negation-as-failure. *ACM Trans. Comput. Logic*, 6(2):441–467, 2005.

[11] J. C. Shepherdson. A sound and complete semantics for a version of negation as failure. *Theor. Comput. Sci.*, 65(3):343–371, 1989.

[12] T. Winograd. *Understanding Natural Language*. Academic Press, New York, 1972.