REASSURANCE FOR THE LOGIC OF PARADOX

MARCEL CRABBÉ
Institut supérieur de Philosophie
Université catholique de Louvain

Abstract. Counterexamples to reassurance relative to a “less inconsistent” relation between models of the logic of paradox are provided. Another relation, designed to fix the problem in logic without equality, is introduced and discussed in connection with the issue of classical recapture.

“it has . . . a certain consonance with common sense which makes it inherently credible. This, however, is not a merit upon which much stress should be laid; for common sense is far more fallible than it likes to believe.” (B. Russell)

§1. Truth and falsehood. The logic of paradox is the natural paraconsistent logic arising from classical logic by simply dropping the principle of noncontradiction. Thus, an LP-model \( \mathfrak{A} \), with nonempty universe \(|\mathfrak{A}|\), is exactly like an ordinary model, except that \( n \)-ary relation symbols are interpreted by ordered pairs of their respective extension and antiextension \( (r^+_\mathfrak{A}, r^-_\mathfrak{A}) \), such that the exhaustiveness requirement (excluded middle) \( r^+_\mathfrak{A} \cup r^-_\mathfrak{A} = |\mathfrak{A}| \) is met. Constants and function symbols are interpreted, as usual, by objects and functions. Likewise, a valuation is still a function of the set of the variables and the valuation \( v_{|\mathfrak{A}|} \) extends canonically to an interpretation \( v_{|\mathfrak{A}|} \) of the terms. The truth and falsehood in a model with respect to a valuation are defined inductively, as follows:

\[
(\mathfrak{A}, v) \models^+ r_1 \ldots r_n \iff (v_{|\mathfrak{A}|}(r_1), \ldots, v_{|\mathfrak{A}|}(r_n)) \in r^+_\mathfrak{A}
\]

\[
(\mathfrak{A}, v) \models^- r_1 \ldots r_n \iff (v_{|\mathfrak{A}|}(r_1), \ldots, v_{|\mathfrak{A}|}(r_n)) \in r^-_\mathfrak{A}
\]

\[
(\mathfrak{A}, v) \models^+ \neg A \iff (\mathfrak{A}, v) \models^- A
\]

\[
(\mathfrak{A}, v) \models^- \neg A \iff (\mathfrak{A}, v) \models^+ A
\]

\[
(\mathfrak{A}, v) \models^+ (A \land B) \iff (\mathfrak{A}, v) \models^+ A \text{ and } (\mathfrak{A}, v) \models^+ B
\]

\[
(\mathfrak{A}, v) \models^- (A \land B) \iff (\mathfrak{A}, v) \models^- A \text{ and/or } (\mathfrak{A}, v) \models^- B
\]

\[
(\mathfrak{A}, v) \models^+ \exists x A \iff (\mathfrak{A}, v[x \mapsto o]) \models^+ A, \text{ for at least one } o \text{ in } |\mathfrak{A}|
\]

\[
(\mathfrak{A}, v) \models^- \exists x A \iff (\mathfrak{A}, v[x \mapsto o]) \models^- A, \text{ for all } o \text{ in } |\mathfrak{A}|
\]

where \( v[x \mapsto o] = (v \setminus \{(x, v(x))\}) \cup \{(x, o)\} \).

Received June 25, 2011.

1 This paper will be almost self-contained. I refer to chapter 16 of Priest (2006) for additional information.
The truth relations \((\mathfrak{A}, v) \models^\pm (A \lor B), (\mathfrak{A}, v) \models^\pm (A \rightarrow B)\), and \((\mathfrak{A}, v) \models^\pm \forall x A\) are defined by 

\[ (\mathfrak{A}, v) \models p \lor q \iff (\mathfrak{A}, v) \models p \lor (\mathfrak{A}, v) \models q \]

\[ (\mathfrak{A}, v) \models p \rightarrow q \iff (\mathfrak{A}, v) \models \neg p \lor (\mathfrak{A}, v) \models q \]

\[ (\mathfrak{A}, v) \models \forall x A \iff (\mathfrak{A}, v) \models a \forall x A \]

respectively.

I will omit to mention the valuation when \(A\) is a sentence (closed formula). When \(\Sigma\) is a theory (set of sentences), I write \(\mathfrak{A} \models^+ \Sigma\), if \(\mathfrak{A} \models^+ A\), for all \(A\) in \(\Sigma\).

**Definition 1.1** The consequence relation \(\Sigma \vdash C\), between a theory and a sentence, is defined by: \(\mathfrak{A} \models^+ C\), for every LP-model \(\mathfrak{A}\) such that \(\mathfrak{A} \models^+ \Sigma\).

It is well-known that the logical truths of LP are exactly the usual classical truths: we have of course \(\vdash (A \lor \neg A)\); and, even, \(\vdash \neg (A \land \neg A)\) and \(\vdash (A \land \neg A) \rightarrow B\). But the consequence relation is not the same. E contradicetione quolibet, \((A \land \neg A) \vdash B\), is not valid, and though \(((A \rightarrow B) \land A) \rightarrow B\) is true in all LP-models, modus ponens is not correct: \(^2\) \((A \rightarrow B), A \not\vdash B\).

§2. Levels of inconsistency.

**Definition 2.1** The set of contradictory facts\(^3\) in \(\mathfrak{A}\), notated \(\mathfrak{A}!\), is the set

\[ \{ \langle p, \langle a_1, \ldots, a_n \rangle \rangle \mid p \text{ is a n-ary relation symbol and } (a_1, \ldots, a_n) \in p_{\mathfrak{A}} \cap p_{\mathfrak{A}} \} \]

The strict partial order \(\prec\) is defined by:

\[ \mathfrak{B} \prec \mathfrak{A} \iff \mathfrak{B}! \subset \mathfrak{A}! \]

I will refer to such a \(\mathfrak{B}\) as a “less inconsistent” model than \(\mathfrak{A}\), or as a “restriction” of \(\mathfrak{A}\). Henceforth \(\mathfrak{A}\) will abbreviate \(A \land \neg \neg A\).

A word of warning is in order here. \(\mathfrak{B} \prec \mathfrak{A}\) means that \((\mathfrak{B}, v) \models px_1 \ldots x_n!\) implies \((\mathfrak{A}, v) \models px_1 \ldots x_n!\), for all \(p\) and \(v\) (with the proper conditions on the valuations). It does not mean that \(\mathfrak{B}\) has less gluts (contradictory tuples) than \(\mathfrak{A}\), that is, \(\mathfrak{A}!\) is not \(\{ a \mid a \in p_{\mathfrak{A}} \cap p_{\mathfrak{A}} \}\); nor does it mean, nor entail, that all the contradictions in \(\mathfrak{B}\) are transferred to \(\mathfrak{A}\), in particular that \((\mathfrak{B}, v) \models pt_1(x) \ldots t_n(x)\), implies \((\mathfrak{A}, v) \models pt_1(x) \ldots t_n(x)\).

**Definition 2.2** A minimal inconsistent (mi) model \(\mathfrak{A}\) of \(\Sigma\) is a model of \(\Sigma\) such that if \(\mathfrak{B} \prec \mathfrak{A}\), then \(\mathfrak{B}\) is not a model of \(\Sigma\).

The associated consequence relation \(\Sigma \vdash_{mi} C\) is defined as \(\mathfrak{A} \models^+ C\), for every mi model \(\mathfrak{A}\) of \(\Sigma\).

Since a mi model of a theory need not be a model of its subtheories, this consequence relation need not be monotonic. For example, \(p, (p \rightarrow q) \vdash_{mi} q\), but \(p, (p \rightarrow q) \not\vdash_{mi} q\). Neither is it closed under substitution: \(p, (p \rightarrow q) \vdash_{mi} q\), but \(p!, (p! \rightarrow q) \not\vdash_{mi} q\).

**Definition 2.3** A theory \(\Sigma\) is LP-trivial \[LPm-trivial\] if \(\Sigma \vdash A \mid \vdash_{mi} A\), for every \(A\) in the language.

A model is trivial iff every sentence of the language is true and false in it.

---

\(^2\) When dealing with sets of sentences, I write \(\Sigma, \Pi\) for \(\Sigma \cup \Pi\) and \(A\) for \(\{A\}\).

\(^3\) Though not entirely satisfactory, I would rather prefer a phrase like “contradictory statements” here, since a fact is, as a rule, rather thought of as being made up by a relation and a tuple: \(\langle p_{\mathfrak{A}}, \langle a_1, \ldots, a_n \rangle \rangle\).
Therefore, a trivial model is a model of any theory in the language and it follows (trivially!) that every theory has a finite model. Hence, a theory is LPm-trivial iff it has no mi models or has only trivial mi models. Contrary to what happens in classical logic, a theory can be LP [LPm]-trivial in one language and not in another: the theory $p!$ is only LP [LPm]-trivial in the propositional language over $p$. More amazing, as shown in Remark 6.2, there are LPm-trivial theories that can be extended to non LPm-trivial ones, in the same language.

§3. Counterexamples to strong reassurance. ‘Strong reassurance’ for $\Sigma_1$ means that every model of $\Sigma_1$ has a mi restriction. The first counterexample to strong reassurance is to be found in Batens (2000). It is an infinite model of an infinite theory saying that the domain of the reflexive relation $\approx$ has an infinite number of contradictory objects with respect to $p$.

Example 3.1 Let $\Sigma_1$ be the set of all formulas
\[
\forall x x \approx x, \ \exists x_1 \exists x_2 \ldots \exists x_n (px_1! \land \ldots \land px_n! \land \bigwedge_{i \neq j, i, j \leq n} \neg x_i \approx x_j) \ (n \geq 1)
\]
The universe of the model $\mathfrak{A}$ is infinite. The predicate $p^+_{\mathfrak{A}}$ is the whole universe, $p^-_{\mathfrak{A}}$ is an infinite subset of it, $\approx^+_{\mathfrak{A}}$ is the identity relation, and $\approx^-_{\mathfrak{A}}$ is the nonidentity relation. This defines a model that has no mi restriction in which the sentences of $\Sigma_1$ remain true, since any less inconsistent model of $\Sigma_1$ is clearly the same kind as $\mathfrak{A}$.

Our second example is that of a finite theory saying that $<$ is a strict (partial) order without maximal element, such that, from some point onwards, every object is contradictory with respect to $p$.

Example 3.2
\[
\Pi = \left\{ \forall x \neg x < x, \ \forall x \forall y \forall z (x < y \land y < z \rightarrow x < z), \ \forall x \exists y (x < y \rightarrow py!), \ \exists x \forall y.x < y \right\}
\]
Again, $\Pi$ has a model without mi restriction among the models of $\Pi$. Indeed, a model of $\Pi$ with no $<$-glut is infinite and must have a lot of $p$-gluts, the number of which can always be reduced.

§4. Counterexample to reassurance. By reassurance is meant that $\Sigma$ is not LPm-trivial, if $\Sigma$ is not LP-trivial. Our counterexample to reassurance is a finite theory with no function symbol.

Example 4.1 Let $\Sigma$ be the theory $\forall x (px! \lor qx!), \exists x (px! \land qx!)$ in the language with exactly the two unary relation symbols mentioned. Then $\Sigma$ is not trivial, because $\Sigma \not\vdash \forall x px$. However, $\Sigma$ is LPm-trivial, since its sole (up to isomorphism!) mi model is trivial.

Notice that $\Sigma$ would not be LPm-trivial, if it was viewed as a theory in an extended language. Nonetheless, this idea to formulate a theory saying that everything is contradictory in some respect and that at least something is absolutely contradictory, can be realized in various languages, using, for example, the following scheme:
EXAMPLE 4.2
\[
\Sigma_n = \begin{cases} 
\forall x \ (p_1 \bar{x} \lor \ldots \lor p_n \bar{x}), \\
\exists x \ (p_1 \bar{x} \land \ldots \land p_n \bar{x}), \\
\forall x \ (p_{n+1} \bar{x}), \forall x \ (p_{n+2} \bar{x}), \forall x \ (p_{n+3} \bar{x}), \ldots
\end{cases}, \ n \geq 2
\]

Note that, \( \Sigma_n \) being a set of sentences, the formulas \( p_i \bar{x} \) are of the form \( p_i \bar{x} \ldots \bar{x} \).
We see that \( \Sigma_n \not\vdash \forall x p_i \bar{x}, \) for \( i = 1, \ldots, n \), though \( \Sigma_n \) is \( \text{LPm}-\text{trivial} \).

§5. Other relations. In order to draw the right conclusions from the counterexample and possibly find a way to get around the problem, I now introduce stronger relations of restriction and the associated notions of minimality. Define:

\[
\begin{align*}
\mathcal{B} & \prec = \mathcal{A} \\
\mathcal{B} & \prec \preceq \mathcal{A} \quad \text{iff} \quad \mathcal{B} \prec \mathcal{A} \quad \text{and} \quad |\mathcal{B}| = |\mathcal{A}| \\
\mathcal{B} & \prec \subseteq \mathcal{A} \quad |\mathcal{B}| \subseteq |\mathcal{A}|
\end{align*}
\]

We have

\[
\mathcal{B} \prec = \mathcal{A} \quad \mathcal{B} \prec \preceq \mathcal{A} \quad \mathcal{B} \prec \subseteq \mathcal{A} \quad \mathcal{B} \prec \mathcal{A}
\]

Thus, for the associated notions of \( \text{mi model of } \Sigma \):

\[
\mathcal{A} \text{ is a mi} = \text{ model of } \Sigma \quad \mathcal{A} \text{ is a mi} \preceq \text{ model of } \Sigma
\]

Hence, the derived consequence relations behave this way:

\[
\Sigma \vdash A \quad \text{implies} \quad \Sigma \vdash_\text{mi} A \quad \Sigma \vdash_\preceq A \quad \Sigma \vdash_\subseteq A
\]

In view of the Counterexample 4.1, the proof of reassurance for the relation \( \vdash_\text{mi} \) in Priest (2006) has a flaw. The proof consists in putting together two correct facts with an incorrect one. The first fact is that every non \( \text{LP}-\text{trivial} \) theory in a language without function symbols has a finite nontrivial model, which can actually be constructed as a homomorphic image of a given nontrivial model. The second fact is that \( \mathcal{A} \) has a \( \text{mi} \) restriction whenever \( \mathcal{A}! \) is finite, which is certainly the case when \( \mathcal{A} \) is finite and the language has only a finite set of relation symbols. If a \( \text{mi} \) restriction of a nontrivial model were not trivial—as stated in Priest (2006)—then the desired conclusion would follow. But this ought no longer be the case, when one is allowed to cut out elements of the universe. Anyway, the proof goes through with \( \prec \preceq \) in place of \( \prec \).

§6. Reassurance and equality. Reassurance, as defined in Priest (2006) and already proved in Priest (1991) for \( \vdash_\text{mi} \), holds for \( \vdash_\preceq \). However, the proof is valid only for languages with a finite number of relation symbols and without function symbols, and it is unexpectedly not the case that reassurance holds generally. In fact, it fails as soon as function symbols and equality are mixed. The following example of an algebraic theory does not enjoy reassurance.

\[4 \text{ CL stands for classical logic. By a harmless abuse of language, I do not distinguish between classical models and LP-models without gluts.} \]
EXAMPLE 6.1 Beside the equality symbol =, the language contains exactly three function symbols: one binary π and two unary p₁ and p₂. The theory Σ is made up of the four sentences ∀x∀y p₁(x, y) = x, ∀x∀y p₂π(x, y) = y, ∀x π(p₁x, p₂x) = x, and ∀x ¬π(x, x) = π(x, x).

If $\mathfrak{A}$ is an LP-model of Σ, then $\pi_{\mathfrak{A}}$ is a bijection between $|\mathfrak{A}| \times |\mathfrak{A}|$ and $|\mathfrak{A}|$. Hence although the theory Σ is not LP-trivial, all its finite models have a singleton universe and are trivial.

Let $\mathfrak{A}$ be a nontrivial LP-model of Σ and D be its infinite “diagonal” set $\{ \pi_{\mathfrak{A}}(o, o) \mid o \in |\mathfrak{A}| \}$. Let further F be a permutation of $|\mathfrak{A}|$, such that F[D] is a proper subset of D.

Define an LP-model $\mathfrak{B}$ with the same universe as follows: $\equiv_{\mathfrak{B}}$ is $\equiv_{\mathfrak{A}}$, $\equiv_{\mathfrak{B}}$ is $\{ (o_1, o_2) \mid o_1 \neq o_2 \text{ or } o_1 \text{ is in } F[D] \}$, $\pi_{\mathfrak{B}}(o_1, o_2) = F(\pi_{\mathfrak{A}}(o_1, o_2))$ and $p_1\mathfrak{B}(o) = p_1\mathfrak{A}(F^{-1}(o))$ (i = 1, 2, o₁, o₂, o ∈ |$\mathfrak{A}$|). One checks that $\mathfrak{B} \models \Sigma$ and that $\mathfrak{B} \not\subseteq \mathfrak{A}$!

In addition, for $o \in D \setminus F[D]$, we have $o \equiv_{\mathfrak{A}} o$ but not $o \equiv_{\mathfrak{B}} o$.

This counterexample is not specific to any of the less inconsistency relations so far introduced. It seems that it will work for any reasonable notion of minimal inconsistency based on $\prec$. As regards reassurance, it seems therefore unavoidable that we have to discard either equality, or function symbols.

REMARK 6.2 Example 6.1 has a remarkable property not to be found in the former ones, namely that it is a LPm-trivial theory that can be extended to a non LPm-trivial one in the same language. Just add the sentence $\exists z_1 \exists z_2 \forall x\forall y (\neg(x, y) \vee \neg(x, y) \equiv (x, y))$ and consider the nontrivial mi model $\mathfrak{M}$ of it, with the set of natural numbers as universe, in which every element, except $\pi_{\mathfrak{M}}(0, 1)$, is different from itself.

§7. Recapture. The relation $\vdash_m$ was introduced in Priest (1991), preferably to $\vdash_m$, because $\forall x (p_1 x) \vdash_m \forall x (p_1 x)$ was considered as counterintuitive. Later on, Priest changes his mind and his definition, on account of the remark by Batens that $\exists x p_1 x, \exists x \neg p_1 x$ has an inconsistent model that can be made less inconsistent—in fact consistent—only by enlarging the universe. Thus $\vdash_m$ doesn’t recapture the classical logical consequence $\vdash_{\text{CL}}$ in consistent environments: $\exists x p_1 x, \exists x \neg p_1 x \vdash_{\text{CL}} \neg \exists x (p_1 x)$, but $\exists x p_1 x, \exists x \neg p_1 x \vdash_m \neg \exists x (p_1 x)$! These difficulties can be answered by replacing $\prec$ by $\prec_{\mathfrak{A}}$. This is obvious for the first one. As to the other one, we will first of all make this claim more explicit in the presence of equality.

The simplest way to handle equality in this context seems to be the following. Let’s recall that a model is general, if the extension of the interpretation of the equality symbol is a congruence, and that it is normal, if this interpretation is the identity relation. We are now in a position to adapt Definition 2.2 to the models for equality.

DEFINITION 7.1 A general minimal inconsistent (general mi$_{\mathfrak{A}}$) model of Σ is a general model of Σ that has no mi$_{\mathfrak{A}}$, restrictions among the general models of Σ.

A normal model of Σ is a minimal inconsistent (mi$_{\mathfrak{A}}$) model of Σ if it is a general mi$_{\mathfrak{A}}$ model of Σ, that is, if there are no less inconsistent ($\prec_{\mathfrak{A}}$) general models of Σ.

REMARK 7.2 It is well-known that the usual consequence relations are unaffected by the choice of either normal or general models. It is also not very hard to show that C is true in every mi$_{\mathfrak{A}}$ general model of Σ if C is true in every mi$_{\mathfrak{A}}$ normal model of Σ. This follows
at once from the fact that if $\mathbf{A}$ is a mi general model of $\Sigma$, so is the quotient\footnote{To make life simpler, we assume that quotients are defined relative to choice functions, so that the universe of a model includes the universe of "its" quotient by an equivalence relation.} $\mathbf{A}/\mathbf{B}_{\omega_1^{\omega_1}}$ of $\mathbf{A}$. Here are some hints for a proof of this proposition. Let $\mathbf{B} \prec \mathbf{A}/\mathbf{B}_{\omega_1^{\omega_1}}$.

We can wlog suppose that $|\mathbf{B}| \cap |\mathbf{A}| = |\mathbf{A}/\mathbf{B}_{\omega_1^{\omega_1}}|$. Define the equivalence $\sim$ on $|\mathbf{A}| \cup |\mathbf{B}|$ by

$$o \sim o' \mbox{ iff } o = o' \mbox{ or } o = o'_{\Delta}.$$

Define $\mathbf{B}'$ as a model with universe $|\mathbf{A}| \cup |\mathbf{B}|$, by putting $(o_1', \ldots, o_n') \in r_{\mathbf{B}'_{\Delta}}$ iff $(o_1, \ldots, o_n) \in r_{\mathbf{B}_{\Delta}}$ and $f_{\mathbf{B}'}(o_1', \ldots, o_n') = f_{\mathbf{B}}(o_1, \ldots, o_n)$, where $o_i$ is the unique element in $|\mathbf{B}|$ such that $o_i \sim o'_i$ ($1 \leq i \leq n$). One shows straightforwardly that $\mathbf{B}' \prec \mathbf{A}$.

Thus using nonnormal models as a kind of virtual auxiliaries, we can still consider the normal models as the “real” ones. If we had prematurely limited the less inconsistency relations to normal models and defined minimality accordingly, then classical recapture would have failed.\footnote{This was pointed out by the referee, who suggested the example.} The sentence $(\forall x \forall y x = y \lor \neg \exists x x = x)$ that has only classical mi models, would have a nonclassical mi$_{\Delta}$ model.

If one insists on having a relation of less inconsistency between normal models, so that it agrees with the definition of general mi$_{\Delta}$ model, one can take the following: $\mathbf{B} \prec_{\text{new}} \mathbf{A}$ iff there is a general model $\mathbf{B}'$ such that $\mathbf{B}' \prec \mathbf{A}$ and $\mathbf{B} = \mathbf{B}'_{\Delta}/\omega_1^{\omega_1}$. Let’s see how it goes with the example just mentioned. Consider a normal nonclassical model $\mathbf{A}$ of $(\forall x \forall y x = y \lor \neg \exists x x = x)$, with $|\mathbf{A}| = \{u, o\}$, and the extension and antiextension of the equality symbol specified by $u =_\mathbf{A} o, u =_\mathbf{A} u, u =_\mathbf{A} u, o =_\mathbf{A} o$. Although no normal model of the sentence is $\prec \mathbf{A}$, the nonnormal classical model $\mathbf{B}$ with $|\mathbf{B}| = |\mathbf{A}|$ and $u =_\mathbf{B} o, u =_\mathbf{B} u, o =_\mathbf{B} o$ is such that $\mathbf{B} \prec \mathbf{A}$. So, $\mathbf{A}$ is no longer a mi$_{\Delta}$ model. Notice that the classical normal model $\mathbf{C} (|\mathbf{C}| = \{u\}, u =_\mathbf{C} u)$ such that $\mathbf{C} \prec \mathbf{A}$ is in fact $\mathbf{B}_{\Delta}/\omega_1^{\omega_1}$.

We now show classical recapture for $\vdash_{\Delta}$ by modifying slightly the proof for $\vdash_m$.

**Proposition 7.3** (Classical Recapture). If $\Sigma$ is consistent, then $\Sigma \vdash_{\Delta} A$ if and only if $\Sigma \vdash_{\text{CL}} A$.

**Proof.** The only if part is obvious, because a classical model of $\Sigma$ is clearly a mi$_{\Delta}$ model of $\Sigma$. For the converse, suppose that $\mathbf{A}$ is a mi$_{\Delta}$ model of $\Sigma$ and that $\mathbf{B}$ is a classical model of $\Sigma$. As long as general models are allowed, the upward Löwenheim–Skolem theorem applies to any model, whether infinite or not. So, we can ensure that $|\mathbf{B}| \supseteq |\mathbf{A}|$. Therefore, $\emptyset = \mathbf{B} \not\subseteq \mathbf{A}$!, which shows that $\mathbf{A}$ is classical. \hfill $\square$
Classical recapture holds trivially for \( \vdash_m \) and, by Proposition 7.3, it holds also for \( \vdash_\supset \), provided nonnormal models are called to rescue. It doesn’t hold for \( \vdash= \) and \( \vdash\subseteq \) (see Section §7, p. 483).

§9. Acknowledgments. I wish to thank Graham Priest for his valuable advice and the referee for his careful report, which helped to improve the first version of this paper.

BIBLIOGRAPHY

