Set-Theoretic Consistency via Infinitary Modal Canonicity

Tyler Greene

3 June 2005

Abstract

This paper presents Baltag's system STS along with the proof of its consistency relative to ZFC plus a large cardinal assumption. The goal is to streamline the construction of the canonical model while emphasizing its essential relation to similar methods in traditional modal logic.

Contents

1 Introduction

Sets are objects characterized uniquely by their members. We can phrase this as a sort of principle: *sets are uniquely determined up to observational equivalence.* In traditional axiomatic systems (ZFC, for example) this notion of "observational equivalence" is captured by the axiom of extensionality:

$$\forall x \forall y \forall z((z \in x \leftrightarrow z \in y) \to x = y).$$

This axiom, together with the more "constructive" axioms (like comprehension and replacement), give the principle that *every collection of sets (that is not too big) uniquely determines a set*. The recursive nature of this principle is well suited for ZFC which includes the axiom of foundation. That is, if we want to determine the structure of a set, we look at its members. Then the members of those members. And so on. But we are assured this process will stop because the universe is well-founded. For a more general theory of sets, however, this will fail. For example, if we try to analyze the set $a = \{a\}$ in this way, we will never stop. So the axioms and principles must be generalized.

This generalization was taken up by Aczel and Barwise and Moss. The appropriate generalization of observational equivalence for non-wellfounded set theory, ZFC^-, turns out to be bisimilarity. This results from thinking of the objects of the universe as pointed graphs. That is, instead of starting with the empty set and building the universe upwards synthetically, we assume we are given some complex object, and we unravel its structure downwards analytically. Then the corresponding generalized axiom is:

> **Antifoundation Axiom** (AFA): *Every formula of infinitary modal logic characterizes a unique set.*[1]

This generalization, however, is still unable to do everything one might reasonably want. For example, the universe itself is a natural object to want to talk about. But the universe is not characterisable by bisimulation (for where would this "bisimulation" exist?) At the end of [2], Barwise and Moss asked for a further generalization of set theory—one that can handle "large" sets (i.e. proper classes with respect to ZFC or $ZFA = ZFC^- + AFA$).

This generalization was taken up by Baltag in [1]. The new notion of observational equivalence becomes infinitary modal equivalence and the updated axiom is:

---

[1] This is not how AFA is usually presented. It is generally stated (as in [2]) as *every pointed graph is bisimilar to a unique set*. But these two formulations are equivalent by the following theorem of modal logic:

> *For every world s in every model M, there is a formula φ of infinitary modal logic such that:*
>
> (i) $M, s \models \varphi$
>
> (ii) *if* $N, t \models \varphi$*, then* M, s *is bisimilar to* N, t.

Our formulation given above, however, will allow for better comparison with the axiom $SAFA$ to be presented below.

Super-Antifoundation Axiom ($SAFA$): *Every weakly consistent infinitary modal theory characterizes a unique set.*

That is, two objects of the universe are identical if and only if they have the same infinitary modal theory. And every (weakly consistent) modal theory determines a set. $SAFA$ is the central axiom of Baltag's new system: STS (Structual Theory of Sets). It is clear, given the form in which we have presented AFA, that $SAFA$ is much stronger existentially. The motivation for this formulation (in particular, using weakly consistent infinitary modal theories) is being able to give partial descriptions of objects that would otherwise be too large to refer to in the system. The advantage is that we can now talk about "large" sets. Consider the set

$$U = \{\Diamond\varphi : \varphi \text{ is consistent}\}.$$

Even without the precise definitions, we can see what this means. U is a set that "sees" every "possibility." Intuitively, this represents the universal set. Once the formalization is in place, this is exactly what we get.

In this paper, the axiomatic system STS is presented. We then prove its consistency (relative to ZFC plus the existence of a large cardinal) by constructing the canonical model. Whereas Baltag's presentation emphasizes how STS is the natural generalization of ZFC^- and AFA, showing the set-theoretic evolution of the concepts involved, we will emphasize the essential modal flavor of the result. Therefore the goal of this paper is to streamline and emphasize the canonical model approach. For more background, motivation, and applications, the reader is referred to [1].

Remark Throughout this paper, it is understood that we are working with the *infinitary* modal language ML_∞. But the "infinitary" will often be dropped.

2 The Axiom System STS

In this section we present the axiom system for STS, as presented by Baltag. The language consists of two binary symbols $\in$ and $\models$ and a unary symbol V. Intuitively, these represent membership (between two sets—objects of the universe), satisfaction (of a formula by a set—both objects of the universe) and the iterative hierarchy. There are three goups of axioms: the basic axioms, the satisfaction axioms, and the super-antifoundation axiom.

Basic Axioms:

- Extensionality
- Closure under singletons and finite unions.
- V is transitive
- V is a model of *Infinity*, *Replacement*, *Union*, and *Choice*.

Satisfaction Axioms:

- $a \models \varphi$ iff $a \not\models \varphi$
- $a \models \bigwedge \Phi$ iff $a \models \varphi$ for all $\varphi \in \Phi$
- $a \models \Diamond\varphi$ iff $a' \models \varphi$ for some $a' \in a$

Super-Antifoundation Axiom:

- *Existence:* Every weakly consistent infinitary modal theory describes a set.
- *Uniqueness:* Sets are uniquely determined by their modal theories.

Much can be said about why these axioms were chosen. We limit ourselves here to just a few remarks. We assume the first two basic axioms because these are essential to our concept of set. The second two are also essential to the *traditional* conception of set, but our intuitions may deceive us with "large" sets. But by assuming these hold at least for some class V (the sets of "small" size), we retain the traditional set-theoretic models as submodels.[2] Also, by assuming the class V exists, we have an easy way to define infinitary modal formulas as sets in the system (by a version of the usual Gödel coding), so that the remaining axioms make sense. The satisfaction axioms are needed to define satisfaction for large sets (it can be defined for small sets within the system). And $SAFA$ captures our new notion of observational equivalence that characterizes the system.

3 Constructing The Canonical Model

This section and the next constitute the heart of the paper. Our goal is to prove that the new set theory STS is consistent relative to ZFC plus a large cardinal assumption. This section contains the first step in that direction: constructing a canonical model. To do this, we assume we have a model of ZFC with a weakly compact cardinal κ. We then define the infinitary modal language ML_κ within ZFC and construct a model based on sets of ML_κ formulas. The construction is guided by analogy with canonical model constructions used in the basic completeness proofs for normal modal logics.

We start by presenting some simple definitions.

Definition A cardinal κ is *weakly compact* if there is no set Γ of formulas in the language $L_{\kappa\omega}$ with the following properties:

- Γ has cardinality κ
- Γ has no model

[2] The powerset axiom can be proved from the others to hold in V. So we need not assume it.

- Every subset of Γ of cardinality $< \kappa$ has a model.

Since we are working in ZFC, we can do the usual coding of the language into the system. So we can assume that the symbols $\bot$, $\neg$, $\bigwedge$, and $\Diamond$ are sets.

Definition The *infinitary modal language* ML_κ is defined recursively as follows:

- $\bot \in ML_\kappa$
- $\neg\varphi \in ML_\kappa$ for all $\varphi \in ML_\kappa$
- $\bigwedge \Phi \in ML_\kappa$ for all $\Phi \subset ML_\kappa$ with cardinality $< \kappa$.
- $\Diamond\varphi \in ML_\kappa$ for all $\varphi \in ML_\kappa$.

We can also define, in ZFC, the notion of satisfaction for *sets* (and then extend it to a notion of satisfaction by a pointed graph—a set of sets with a relation). This is done by some coding scheme as in other set-theoretic relative consistency proofs. We can then define the following notions:

Definition

- A set $\Sigma \subseteq ML_\kappa$ is a *theory*.
- A theory Σ is *consistent* if it is satisfied by some graph.
- A theory Σ is *weakly consistent* if every subset of Σ of size $< \kappa$ is consistent.
- A theory Σ is a *maximally consistent* set (an MCS) if it is consistent and has no consistent proper extension.

With these notions in hand (and defined in ZFC), we can construct the canonical model. The construction is as usual. The domain of the model will be all maximally consistent sets of formulas of the modal language. And the relation will be that one set "sees" another if every boxed formula of the former appears "deboxed" in the latter.

Definition

- $W = \{\Sigma \subseteq ML_\kappa : \Sigma \text{ is an MCS}\}$.
- The relation $\in^M$ on W is given by:

$$\begin{aligned} w \in^M u \quad & \text{iff} \quad \forall\varphi(\varphi \in w \Rightarrow \Diamond\varphi \in u) \\ & \text{iff} \quad \forall\varphi(\Box\varphi \in u \Rightarrow \varphi \in w) \end{aligned}$$

- The *canonical model* is $M = (W, \in^M)$.[3]

[3]Note that this structure is usually defined as the *canonical frame*. But our modal language has no proposition letters, so the distinction can safely be ignored.

- $V^M = \{th_\kappa(x) : x \in V_\kappa\}$ (where $th_\kappa(x) = \{\varphi \in ML_\kappa : M, x \models \varphi\}$).[4]
- $\models^M = \{(w, \varphi) \in W \times ML_\kappa : \varphi \in w\}$.

The last two parts of this definition are needed to interpret the symbols V and $\models$ of our extended set-theoretic language. This is the only part that doesn't appear in the usual modal completeness proofs.

Now that we have constructed something, it remains to show that it really is a model for the axioms of STS. We do this in two steps. First, we prove the usual lemmas associated with canonical models. This emphasizes the modal approach. Then, once we have the truth lemma at our disposal, we can show that our new axioms are satisfied by $(M, V^M, \models^M)$.

4 The Lemmas

In the last section, we constructed the canonical model M. The goal of this section is to prove the truth lemma. We follow closely the exposition of [3]. From now on, Σ denotes a theory of ML_κ. Recall that an MCS is a *maximally consistent set* of formulas from ML_κ.

Lemma 4.1 (Properties of MCSs) *If Σ is an MCS and $\Phi \subseteq ML_\infty$ has size $< \kappa$, then:*

(i) $(\neg\varphi) \in \Sigma$ *iff* $\varphi \notin \Sigma$

(ii) $\bigwedge \Phi \in \Sigma$ *iff* $\Phi \subseteq \Sigma$

Proof The proof of (i) is as usual. The proof of (ii) proceeds in the usual way, with only a slight generalization (to handle the infinite, rather than binary, conjunction). ⊣

Lemma 4.2 (Lindenbaum) *If Σ is weakly consistent, then it is included in some MCS.*

Proof [5] Let $\Sigma \subset ML_\kappa$ be weakly consistent. By the definition of (and the fact that κ is a) weakly compact cardinal, Σ is consistent. Now we can mimic the proof of the standard version of Lindenbaum's lemma. Let $\{\varphi_\alpha : \alpha \in \kappa\}$ be an enumeration of ML_κ. We define a sequence by transfinite recursion on κ:

$$\begin{aligned} \Sigma_0 &= \Sigma \\ \Sigma_{\alpha+1} &= \begin{cases} \Sigma_\alpha \cup \{\varphi_\alpha\}, & \text{if this is consistent} \\ \Sigma_\alpha \cup \{\neg\varphi_\alpha\}, & \text{otherwise} \end{cases} \\ \Sigma_\lambda &= \bigcup_{\beta<\lambda} \Sigma_\beta \text{ for limit ordinals } \lambda \end{aligned}$$

[4] Just to be clear, V^M is not the canonical valuation, but the interpretation of the cumulative hierarchy in our new model.

[5] This proof differs from that in [1]. There, Baltag uses the notion of *pointed graph* to show that maximally consistent theories correspond with pointed graphs, and then derive Lindenbaum's lemma from this. The proof here remains closer to the spirit of the standard proofs of modal logic.

We then set $\Sigma^+ = \bigcup_{\alpha<\kappa} \Sigma_\alpha$, which clearly contains Σ and is an MCS. $\dashv$

Lemma 4.3 (Existence) *For any $w \in W, \varphi \in ML_\kappa$, if $\Diamond\varphi \in w$ then there exists $u \in W$ such that $u \in^M w$ and $\varphi \in u$.*

Proof Suppose $\Diamond\varphi \in w$. Toward the end of constructing a set $v^+ \in W$ such that $v^+ \in^M w$ and $\varphi \in v$, consider the theory $v = \{\varphi\} \cup \{\psi : \Box\psi \in w\}$. To show v is consistent, we need only show that every subset of size $< \kappa$ is (and then appeal the the weak compactness of κ). So suppose a subset $v^- \subseteq v$ of size $< \kappa$ is inconsistent. Then the formula $\bigwedge v^- \wedge \varphi$ is inconsistent. And then so is $\Diamond(\bigwedge v^- \wedge \varphi)$. But $\Diamond(\bigwedge v^- \wedge \varphi) \in w$ since w is an MCS containing $\Diamond\varphi$ and $\Box\psi$ for every $\psi \in v^-$. Then w would be inconsistent, contradicting the fact that it's an MCS.

So v is weakly consistent, hence consistent. By Lemma 4.2, it is included in some MCS v^+. Now, by construction, $\varphi \in v^+$ and $v^+ \in^M w$. $\dashv$

With the Existence Lemma, we can easily prove the Truth Lemma, which will be our main tool for proving consistency in the next section.

Lemma 4.4 (Truth) *For every $w \in W$ and every $\varphi \in ML_\kappa$, $M, w \models \varphi \Leftrightarrow \varphi \in w$.*

Proof The proof proceeds by induction on φ. The base case is $\varphi \equiv \bot$. Notice that $M, w \not\models \bot$ and $\bot \notin w$ (since w is an MCS) always. So the lemma holds. The boolean cases follow directly from Lemma 4.1:

$$\begin{aligned} M, w \models \neg\varphi &\text{ iff } M, w \not\models \varphi \\ &\text{ iff } \varphi \notin w \\ &\text{ iff } \neg\varphi \in w \end{aligned}$$

where the second equivalence follows from the induction hypothesis, while the third was established as a property of MCSs. Similarly,

$$\begin{aligned} M, w \models \bigwedge \Phi &\text{ iff } M, w \models \varphi \text{ for each } \varphi \in \Phi \\ &\text{ iff } \varphi \in w \text{ for each } \varphi \in \Phi \\ &\text{ iff } \bigwedge \Phi \in w. \end{aligned}$$

For the modal case,

$$\begin{aligned} M, w \models \Diamond\varphi &\text{ iff } \exists v(v \in^M w \wedge M, v \models \varphi) \\ &\text{ iff } \exists v(v \in^M w \wedge \varphi \in v) \\ &\text{ iff } \Diamond\varphi \in w. \end{aligned}$$

The equivalence between the second and third statements follows from the induction hypothesis. The fourth follows from the third by the definition of $\in^M$. And the third follows from the fourth by Lemma 4.3. $\dashv$

5 Consistency of STS

This section contains the main result of the paper (though maybe less interesting than the previous two from the view of modal logic). The result is one of relative consistency proved by Baltag in [1]. We have assumed that ZFC plus "there exists an infinite weakly compact cardinal" is consistent, and so has a model. We have constructed from that model another structure, $(M, V^M, \models^M)$. We now show that this new structure is a model of the axioms of STS, hence proving that STS is (relatively) consistent. Since the proof of the theorem is fairly straightfoward (and of a more set-theoretic than modal character), we prove only a few cases to show how it works.

Theorem 5.1 *$(M, V^M, \models^M)$ is a model of STS.*

Proof We first show that extensionality holds, since this is one of the more basic yet also essential characteristics of set theory.

> *Extensionality*: Let $w, u \in W$ and $\forall v \in W(v \in^M w \leftrightarrow v \in^M u)$. We want to show $w = u$ or, equivalently, $\forall \varphi \in ML_\kappa(\varphi \in w \leftrightarrow \varphi \in u)$. Notice that by the definition of $\in^M, \forall v \in W(v \in^M w \leftrightarrow v \in^M u)$ implies that $\{\varphi : \Diamond\varphi \in w\} = \{\varphi : \Box\varphi \in u\}$ and $\{\varphi : \Box\varphi \in w\} = \{\varphi : \Diamond\varphi \in u\}$. We can now easily show $\varphi \in w \leftrightarrow \varphi \in u$ by induction on φ. The base and boolean cases are straightforward. So suppose $\Diamond\varphi \in w$. Then $\Box\varphi \in u$ (by $\{\varphi : \Diamond\varphi \in w\} = \{\varphi : \Box\varphi \in u\}$). Because $\Diamond\varphi \in w$, we know there is a v such that $v \in^M w$ and $\varphi \in v$. But then $v \in^M u$ (by our original assumption that w and u have the same $\in^M$-elements). So $\Diamond\varphi \in u$. So $w \subseteq u$. The other containment is obtained similarly by using $\{\varphi : \Box\varphi \in w\} = \{\varphi : \Diamond\varphi \in u\}$.

We now show that $SAFA$ holds, since it is the axiom that distinguishes STS from other set theories.

> *$SAFA$ existence*: We want to show that given some weakly consistent modal theory (in M), it describes a set (that is, is satisfied by a set in M). So let w^- be a weakly consistent theory (in the model M). Then w^- is weakly consistent.[6] So, by Lemma 4.2, it is contained in an MCS $w \in W$. By Lemma 4.4, $M, w \models w^-$. So our weakly consistent theory describes some set (all with respect to the canonical model M).

> *$SAFA$ uniqueness*: We want to show that sets (elements $w \in W$) are uniquely defined by their modal theories. But since our sets are MCSs, it turns our that they *are* their modal theories. From this it immediately follows that if $th_\kappa(w) = th_\kappa(u)$, then $w = u$. So all that remains is to notice that

$$th_\kappa(w) \quad = \quad \{\varphi \in ML_\kappa : M, w \models \varphi\}$$

[6] This step actually requires us to prove that the notion of "weak consistency" is *absolute*. This is an obvious yet extremely tedious part of any relative consistency proof. So we omit it here.

$$= \{\varphi \in ML_\kappa : \varphi \in w\} \text{ by Lemma 4.4}$$
$$= w. \qquad \dashv$$

Corollary 5.2 *STS is consistent relative to ZFC plus the existence of an infinite weakly compact cardinal.*

Proof This follows immediately from the theorem, since we constructed $(M, V^M, \models^M)$ by assuming a model of ZFC plus "there is a weakly compact cardinal" and only used those axioms to show that $(M, V^M, \models^M)$ is indeed a model of STS. $\dashv$

6 Conclusion

Thus we have established the relative consistency of the axiomatic system STS. The theorem, though maybe superficially unrelated to modal logic, relies heavily on modal techniques. Not only does the axiomatic system ($SAFA$ in particular) depend on the infinitary modal language (which can be defined by the other axioms), but we proved its relative consistency via the construction of a canonical model—a paradigm of modal logic.

Now that we have proven the relative consistency of a new system of axioms, two obvious questions might arise:

1. What good is a relative consistency proof?

2. What do we gain with STS?

Both questions have short simple and long complicated answers. We provide the short answers here.[7]

The relative consistency proof tells us that if ZFC plus the large cardinal assumption is consistent, then so is STS. That is, insofar as talking about set theory as we usually do makes sense (as most mathematicians/logicians/philosophers believe it does), talking about sets in the way prescribed by STS makes sense also. Thus we should not be worried about STS being too powerful or strange, for it is on essentially the same footing as our beloved ZFC.

As to the second question, we gain expressive power. It has long been a somewhat undesirable aspect of ZFC (and AFA, NGB, and almost every theory of sets) that we cannot talk about certain things we'd like to talk about. The most notable example is the universal set—the set that contains everything. The set-theoretic paradoxes (which show, among other things, that the "set" of all sets cannot really be a set) are infamous in philosophy. But now it's as if we have found a loophole. We have crept up on the universal set using partial descriptions—weakly consistent infinitary modal theories—and

[7] To find the long answers for the first question, check any book on the philosophy of mathematics or set theory. For a more in depth analysis of the second question, see [1].

captured it. In particular, it is defined by the set $\{\Diamond\varphi : \varphi \text{ is consistent}\}$, as mentioned in the introduction. Being able to talk about the set of all sets is philosophical motivation enough for seriously considering STS. But there are also “real” applications presented in the second half of [1], to which the reader is now referred.

References

[1] A. Baltag. *STS: A Structual Theory of Sets*. Ph.D. Dissertation. To appear. 1998.

[2] J. Barwise and L. Moss. *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena*. CSLI Publications. 1996.

[3] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.