# Modal Logics Final

Yvan Quinn

July 22, 2018

**Abstract**

**Theorem 2.68 (Van Benthem Characterization Theorem)** Let $\alpha(x)$ be a first-order formula in $\mathcal{L}^1_\tau$. Then $\alpha(x)$ is invariant for bisimulations iff it is (equivalent to) the standard translation of a modal $\tau$-formula.

We begin by defining invariance over bisimulation as follows: A first order formula $\alpha(x) \in \mathcal{L}^1_\tau$ is *invariant for bisimulations* if, for all models $\mathfrak{M}, \mathfrak{N}$ and all bisimulations $Z$ between them,

$$\forall \omega \in \mathfrak{M}, \forall \nu \in \mathfrak{N} : \omega Z \nu \to (\mathfrak{M} \models \alpha(x)[\omega] \leftrightarrow \mathfrak{N} \models \alpha(x)[\nu])$$

As Theorem 2.20 establishes the invariance under bisimulation of all modal (and thus any equivalent) formulas, the right to left implication of the desired theorem is obtained. This fact is fairly obvious as the entire point of bisimulated models is to preserve the satisfaction of modal formulas. The left to right implication is perhaps less obvious and certainly more difficult to prove.

Assume $\alpha(x)$ is invariant for bisimulations. The set of modal consequences $\mathrm{MOC}(\alpha)$ is the set of formulas implied by $\alpha(x)$ that are standard translations of a modal formula. As is shown in the text, if $\mathrm{MOC}(\alpha) \models \alpha(x)$ then $\exists \phi$ such that $\alpha(x)$ is equivalent to $ST_x(\phi)$. Thus, if it can be shown that $\mathrm{MOC}(\alpha) \to \alpha(x)$ then the desired theorem has been proven.

Assume $\mathfrak{M} \models \mathrm{MOC}(\alpha)[\omega]$. Let $T(x)$ be the set of standard translations of modal formulas true at the point $\omega$, and assume for contradiction that $T(x) \cup \{\alpha(x)\}$ is inconsistent. By first-order compactness,

$$\exists T_0 \subseteq T(x) : \models \alpha(x) \to \neg \bigwedge T_0$$

As the conjunction/negation of a standard translation is still a standard translation, $\neg \bigwedge T_0$ is the standard translation of a modal formula, and as it is implied by $\alpha(x)$, must be $\in \mathrm{MOC}(\alpha)$. Thus, by our assumption that $\mathfrak{M}, \omega \models \mathrm{MOC}(\alpha)$, it must be that $\mathfrak{M} \models \neg \bigwedge T_0[\omega]$. By the definition of $T(x)$, this is a contradiction, so $T(x) \cup \{\alpha(x)\}$ is consistent.

# 1 Original Approach

Now I will diverge slightly from the text. If $\mathfrak{M}$ were a countably saturated model, then the fact that this set is consistent would guarantee that $T(x) \cup \{\alpha(x)\}$ would be realized by the expansion $\mathfrak{M}_A$. This means that $\exists \omega' : \mathfrak{M} \models T(x) \cup \{\alpha(x)\}[\omega']$, which demonstrates that the desired theorem is true for any countably saturated model if it can be shown that $\omega'$ is the same point as the $\omega$ used in defining $T(x)$, or that we can find a point $\omega'$ for which $p(x)[\omega'] \to p(x)[\omega]$.

Assume we have chosen an $\omega'$ for which this is false, so we have

$$\exists p(x) : p(x)[\omega'] \wedge \neg p(x)[\omega]$$

This means $\{\neg p(x)\} \cup P(x) \cup T(x)$ is consistent, where $P(x)$ is the (possibly empty) set of all other formulas which are true at both $\omega, \omega'$. It also must be the case that $\{\neg p(x)\} \cup P(x) \cup T(x) \cup \{\alpha(x)\}$ is consistent. Assume it is not for contradiction, so by compactness,

$$\exists T_0 \subseteq T, \exists P_0 \subseteq P(x) \models \bigwedge T_0 \to \neg[\alpha \wedge \neg p(x) \wedge \bigwedge P_0]$$

$$\models \bigwedge T_0 \to \neg\alpha \vee p(x) \vee \neg \bigwedge P_0$$

The first possibility is not possible, as $T(x) \cup \{\alpha(x)\}$ was shown to be consistent before, and the last is not possible as $T(x), P(x)$ are true at both points, so we have

$$\models \bigwedge T_0 \to p(x)$$

This means that $T_0 \cup \{\neg p(x)\}$ is inconsistent, and as $T_0 \subseteq T(x)$, this contradicts the consistency of $\{\neg p(x)\} \cup T(x)$.

Since $\{\neg p(x)\} \cup P(x) \cup T(x) \cup \{\alpha(x)\}$ is consistent, and $\mathfrak{M}$ is countably saturated, there is another point $\omega''$ such that $\mathfrak{M} \models \{\neg p(x)\} \cup P(x) \cup T(x) \cup \{\alpha(x)\}[\omega'']$. By induction (with $P(x) = \{\varnothing\}$ as the base case), we can see that for any choice of formulas true at $\omega$, we can find a point $\omega''$ for which those same formulas are true, in addition to $\alpha(x)$. Once all formulas true for $\omega$ are transferred over,

$$\forall p(x) : p(x)[\omega] \rightarrow p(x)[\omega'']$$

Now, assume for contradiction that at $\omega''$, we the reverse implication, $p(x)[\omega''] \rightarrow p(x)[\omega]$, does not hold. Therefore, we have that, as before,

$$\exists p(x) : p(x)[\omega''] \wedge \neg p(x)[\omega]$$

This would imply that $p(x)[\omega''] \wedge \neg p(x)[\omega'']$, which is a contradiction, so we have that

$$\forall p(x) : p(x)[\omega''] \leftrightarrow p(x)[\omega]$$

Therefore for any $\omega$ in a countably saturated set, we may always choose an $\omega''$ such that $\models T(x) \cup \{\alpha(x)\}[\omega'']$ and for which $\forall p(x) : p(x)[\omega''] \rightarrow p(x)[\omega]$. This proves the Van Benthem Characterization Theorem only for countably saturated models.

Unfortunately not every model is a countably saturated model, but every model can be elementarily embedded in one. If this is done, the same set of formulas is consistent in the countably saturated model $\mathfrak{M}^*$, and so, we know that $\mathfrak{M}^* \models T(x)[\omega^*]$ and $\exists \omega'^* : \mathfrak{M}^* \models T(x) \cup \{\alpha(x)\}[\omega'^*]$. Applying the previous result, an $\omega''^*$ can be found for which $\models p(x)[\omega''^*] \leftrightarrow p(x)[\omega]$ and $\mathfrak{M}^* \models \alpha(x)[\omega''^*]$. This means that $\mathfrak{M}^* \models \alpha(x)[\omega^*]$ and with the invariance of formulas over elementary embeddings, $\mathfrak{M} \models \{\alpha(x)\}[\omega]$, which proves the theorem for all models.$_\square$

# 2 Source Approach

Returning to the strategy outlined in the text, let $\mathfrak{N}$ be a model with a point $\nu$ such that $\mathfrak{N} \models T(x) \cup \alpha(x)[\nu]$. The text shows that this implies the modal equivalence of $\omega$ and $\nu$. When we embed the models in their countably saturated ultraproducts, $\omega^* \in \mathfrak{M}^*, \nu^* \in (N)^*$, their modal equivalence persists. According to the Hennessy-Millner property of m-saturated, and by extension countably saturated models, this guarantees a bisimilarity $\nu^* Z \omega^*$ exists, and so

$$\mathfrak{N} \models \alpha(x)[\nu] \Rightarrow \mathfrak{N}^* \models \alpha(x)[\nu^*] \Rightarrow \mathfrak{M}^* \models \alpha(x)[\omega^*] \Rightarrow \mathfrak{M} \models \alpha(x)[\omega]_\square$$