On the Expressivity of First-Order Modal Logic with "Actually"

Alex Kocurek

For the 5th International Workshop in Logic, Rationality, and Interaction.

Abstract. Many authors have noted that a number of English modal sentences cannot be formalized into standard first-order modal logic. Some widely discussed examples include "There could have been things other than there actually are" and "Everyone who's actually rich could have been poor." In response, many authors have introduced an "actually" operator @ into the language of first-order modal logic. It is occasionally noted that some of the example sentences still cannot be formalized with @ if one allows only actualist quantifiers, and embedded versions of these example sentences cannot be formalized even with possibilist quantifiers and @. The typical justification for these claims is to observe that none of the most plausible candidate formalizations succeed. In this paper, we prove these inexpressibility results by using a modular notion of bisimulation for first-order modal logic with "actually" and other operators. In doing so, we will explain in what ways these results do or do not generalize to more expressive modal languages.

§1 Introduction

Despite all of its strengths, first-order modal logic faces fundamental limitations in expressive power. Some classic examples demonstrating this include:

(E) There could have been things other than there actually are.[1]

(R) Everyone who's actually rich could have been poor.[2]

The first says that there is a possible world where something exists that doesn't actually exist. The second, on one reading, says that there's a possible world where everyone that is rich in the actual world is poor in that world. It has been shown using (rather complicated) Henkin-style constructions that even very simple sentences like (E) and (R) cannot be expressed in first-order modal logic with actualist quantifiers (i.e., quantifiers ranging over existents) Hodes [1984b]. Using possibilist quantifiers (i.e., quantifiers ranging over all possible objects) and an existence predicate, (E) can be expressed, but (R) is still inexpressible Wehmeier [2001].

In response to these expressive limitations, a number of authors have considered introducing an "actually" operator @ into the language Crossley and Humberstone [1977]; Davies and Humberstone [1980]; Hazen [1976, 1990]; Hodes [1984a]. They then point out that in the presence of @ and possibilist quantifiers (where Π is the universal possibilist quantifier) we can formalize (R) as:

$$\Diamond \Pi x\ (@\mathsf{Rich}(x) \to \mathsf{Poor}(x))\,. \tag{1}$$

[1] Originally from [Hazen, 1976, p. 31].
[2] Originally from [Cresswell, 1990, p. 34].

However, if we replace the Π above with an actualist quantifier $\forall$, (1) would yield the wrong result Bricker [1989]; Fara and Williamson [2005]. For then (1) would only require that there is a world w where everyone in w who is actually rich is poor in w, whereas (R) requires that everyone in the actual world who's actually rich is poor in w.

It has also been noted that even with possibilist quantifiers, sentences like:

(NE) Necessarily, there could have been other things than those that existed.

(NR) Necessarily, the rich could have all been poor.

remain inexpressible Bricker [1989]; Cresswell [1990]; Hazen [1976]; Sider [2010]. For instance, on one reading, (NR) says that in all possible worlds w, there's a possible world v where everyone rich in w is poor in v. But, for instance, formalizing (NR) as

$$\Box\Diamond\Pi x\ (@\mathsf{Rich}(x) \to \mathsf{Poor}(x)) \tag{2}$$

will yield the wrong result. This says that for all worlds w, there's a world v such that everyone that's actually rich (not rich in w) is poor in v. One could try to add more operators to the language, but problems keep cropping up van Benthem [1977]; Cresswell [1990].

These inexpressibility claims are often justified in the literature by example: all of the most straightforward attempts at formalizing these English sentences fail. While this style of argument may be convincing, it does not constitute a proof of these expressive limitations. Furthermore, the only proofs known in the literature involve quite complicated and indirect Henkin constructions that are limited to specific languages. In this paper, we will provide a single proof method for generating these inexpressibility proofs for a wide variety of quantified modal languages using a suitable modular notion of bisimulation for first-order modal logic. For concreteness, we'll focus on the proofs for the inexpressibility of (R) and (NR), which have proven more difficult than (E) and (NE). In passing, we will see how these inexpressibility results do, and do not, generalize to more powerful modal languages.

§2 First-Order Modal Logic

First, we'll need to get clear about what exactly we're taking first-order modal logic to be. The details below are fairly standard, with the exception that our semantics is two-dimensional (to account for the actuality operator @). While we've picked a particularly simple formulation of first-order modal logic, these inexpressibility results apply to a wide range of formulations.[3]

The signature for our first-order modal language $\mathcal{L}^{1\mathsf{M}}$ contains:

- $\mathsf{VAR} = \{x_1, x_2, x_3, \ldots\}$ (the set of ***(object) variables***);
- $\mathsf{PRED}^n = \{P^n_1, P^n_2, P^n_3, \ldots\}$ for each $n \geqslant 1$ (the set of ***n-place predicates***);

[3]See Garson [2001] for a tree of such formulations.

The set of ***formulas in*** $\mathcal{L}^{1M}$ or $\mathcal{L}^{1M}$***-formulas*** is defined recursively:

$$\varphi ::= P^n(y_1, \ldots, y_n) \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid \forall x\, \varphi$$

where $P^n \in \mathsf{PRED}^n$ for any $n \geqslant 1$, and $x, y_1, \ldots, y_n \in \mathsf{VAR}$. The usual abbreviations for $\vee$, $\to$, $\exists$, and $\Diamond$ apply. We may drop parentheses for readability. If the free variables of φ are among $y_1, \ldots, y_n$, we may write "$\varphi(y_1, \ldots, y_n)$" to indicate this.

Let $S_1, \ldots, S_n$ be some new symbols with well-defined syntax. We'll indicate the language obtained from $\mathcal{L}^{1M}$ by adding $S_1, \ldots, S_n$ as $\mathcal{L}^{1M}(S_1, \ldots, S_n)$. Some symbols that might be added include:

$$\varphi ::= \cdots \mid y_1 \approx y_2 \mid @\varphi \mid {\downarrow}\varphi \mid \mathcal{F}\varphi \mid \forall_{@} x\, \varphi \mid \Pi x\, \varphi$$

where $\approx$ is the identity relation, $@$ is an "actually" operator, $\downarrow$ is a diagonalization operator Lewis [1973] that does the opposite of $@$, $\mathcal{F}$ is a "fixedly" operator Davies and Humberstone [1980], and $\forall_{@}$ is a quantifier over all actual objects. In what follows, $\mathcal{L}$ will just be any arbitrary $\mathcal{L}^{1M}(S_1, \ldots, S_n)$ where $S_1, \ldots, S_n$ are among the symbols above.

Definition 2.1 *(First-Order Modal Models).* An $\mathcal{L}^{1M}$***-model*** or ***modal model*** is an ordered tuple $\mathcal{M} = \langle W, R, D, \delta, I\rangle$ where:

- W is a nonempty set (the ***state space***);
- $R \subseteq W \times W$ (the ***accessibility relation***);
- D is a nonempty set (the ***(global) domain***);
- $\delta \colon W \to \wp(D)$ is a function (the ***local domain assignment***), where for each $w \in W$, $\delta(w)$ is the ***local domain of*** w;
- I is a function (the ***interpretation function***) such that for each $P^n \in \mathsf{PRED}^n$, $I(P^n, w) \subseteq D^n$.

By convention, where $\mathcal{M}$ is a modal model, we'll say that $\mathcal{M}$'s state space is $W^{\mathcal{M}}$, $\mathcal{M}$'s accessibility relation is $R^{\mathcal{M}}$, etc. We'll let $R[w] := \{v \in W \mid wRv\}$.

Let $\mathcal{M}$ be an $\mathcal{L}^{1M}$-model. A ***variable assignment for*** $\mathcal{M}$ is a function assigning members of its global domain to variables. Let the set of variable assignments on $\mathcal{M}$ be $\mathsf{VA}(\mathcal{M})$. If a variable assignment g for $\mathcal{M}$ agrees with a variable assignment g' for $\mathcal{M}$ on every variable except possibly x, then g and g' are x***-variants***, $g \sim_x g'$. The variable assignment $g[x \mapsto a]$, or g^x_a, is the x-variant of g that sends x to a.

Some notation: if $\alpha_1, \ldots, \alpha_n$ is a sequence (of terms, objects, etc.), we may write "$\overline{\alpha}$" in place of "$\alpha_1, \ldots, \alpha_n$". $\overline{\alpha}$ is assumed to be of the appropriate length, whatever that is in a given context. When f is some unary function, we may write "$f(\overline{\alpha})$" in place of "$f(\alpha_1), \ldots, f(\alpha_n)$". We'll let $|\overline{\alpha}|$ be the length of $\overline{\alpha}$.

Since we want to consider operators like $@$, our semantics will be two-dimensional (as suggested in e.g., [Davies and Humberstone, 1980, pp. 4-5]). That is, indices will have to contain two worlds. The first world is to be interpreted as the world "considered as actual", and the second as the world of evaluation.

Definition 2.2 *(Satisfaction)*. The ***satisfaction relation***, $\Vdash$, is defined recursively, for all $\mathcal{L}^{1\mathsf{M}}$-models $\mathcal{M} = \langle W, R, D, \delta, I\rangle$, all $w, v \in W$ and all $g \in \mathsf{VA}(\mathcal{M})$:

$$\begin{array}{lcl}
\mathcal{M}, w, v, g \Vdash P^n(\overline{x}) & \Leftrightarrow & \langle g(\overline{x})\rangle \in I(P^n, v) \\
\mathcal{M}, w, v, g \Vdash x \approx y & \Leftrightarrow & g(x) = g(y) \\
\mathcal{M}, w, v, g \Vdash \neg\varphi & \Leftrightarrow & \mathcal{M}, w, v, g \nVdash \varphi \\
\mathcal{M}, w, v, g \Vdash \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}, w, v, g \Vdash \varphi \text{ and } \mathcal{M}, w, v, g \Vdash \psi \\
\mathcal{M}, w, v, g \Vdash \Box\varphi & \Leftrightarrow & \forall v' \in R[v]\colon\ \mathcal{M}, w, v', g \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash @\varphi & \Leftrightarrow & \mathcal{M}, w, w, g \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \downarrow\varphi & \Leftrightarrow & \mathcal{M}, v, v, g \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \mathcal{F}\varphi & \Leftrightarrow & \forall w' \in R[w]\colon\ \mathcal{M}, w', v, g \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \forall x\,\varphi & \Leftrightarrow & \forall a \in \delta(v)\colon\ \mathcal{M}, w, v, g^x_a \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \forall_@ x\,\varphi & \Leftrightarrow & \forall a \in \delta(w)\colon\ \mathcal{M}, w, v, g^x_a \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \Pi x\,\varphi & \Leftrightarrow & \forall a \in D\colon\ \mathcal{M}, w, v, g^x_a \Vdash \varphi.
\end{array}$$

If $|\overline{x}| \leqslant |\overline{a}|$, then $\mathcal{M}, w, v \Vdash \varphi[\overline{a}]$ if for all $g \in \mathsf{VA}(\mathcal{M})$, $\mathcal{M}, w, v, g^{\overline{x}}_{\overline{a}} \Vdash \varphi(\overline{x})$.

§3 The Two-Sorted Language

In order to prove our inexpressibility results, we need to translate ordinary English sentences like (R) into a correspondence language. This language is just a two-sorted first order language: one sort for objects, and one sort for worlds.

The signature for our two-sorted first-order language $\mathcal{L}^{2\mathsf{S}}$ contains VAR plus:

- $\mathsf{SVAR} = \{s_1, s_2, s_3, \ldots\}$ (the set of ***state variables***).
- $\mathsf{PRED}^{n/m} = \left\{P^{n/m}_1, P^{n/m}_2, P^{n/m}_3, \ldots\right\}$ for each $n, m \geqslant 1$ (the set of n/m***-place predicates***).

For a predicate $P^{n/m}$, n is the object-arity, while m is the state-arity. Thus, $P^{n/m}$ takes exactly n object variables and m state variables as arguments.[4]

The set of ***formulas in*** $\mathcal{L}^{2S}$ or $\mathcal{L}^{2S}$***-formulas*** is defined recursively:

$$\varphi ::= P^{n/m}(y_1, \ldots, y_n; s_1, \ldots, s_m) \mid \mathsf{E}(x; s_1) \mid \mathsf{R}(s_1, s_2) \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \forall x\,\varphi \mid \forall s\,\varphi$$

where $P^{n/m} \in \mathsf{PRED}^{n/m}$, $x, y_1, \ldots, y_n \in \mathsf{VAR}$, and $s, s_1, \ldots, s_m \in \mathsf{SVAR}$.

For instance, here are the intended formalizations of (R) and (NR), where s^* is meant to be interpreted as the actual world:

$$\exists t\ (\mathsf{R}(s^*, t) \wedge \forall x\ (\mathsf{Rich}(x; s^*) \to \mathsf{Poor}(x; t))) \tag{3}$$

$$\forall s\,(\mathsf{R}(s^*, s) \to \exists t\ (\mathsf{R}(s, t) \to \forall x\ (\mathsf{Rich}(x; s) \to \mathsf{Poor}(x; t)))). \tag{4}$$

[4] We'll use ";" to separate object variables and state variables.

Definition 3.1 *(Two-Sorted Models).* An ***$\mathcal{L}^{2S}$-model*** or ***two-sorted model*** is an ordered tuple $\mathfrak{M} = \langle W, D, V \rangle$ where W and D are nonempty sets, and V is a function (the ***valuation function***) such that:

- for each $P^{n/m} \in \mathsf{PRED}^{n/m}$, $V(P^{n/m}) \subseteq D^n \times W^m$;
- $V(\mathsf{E}) \subseteq D \times W$;
- $V(\mathsf{R}) \subseteq W \times W$.

We are usually interested in the correspondence between $\mathcal{L}^{2\mathsf{S}}$ and $\mathcal{L}^{1\mathsf{M}}$-models.

Definition 3.2 *(Model Correspondents).* Let $\mathcal{M} = \langle W, R, D, \delta, I \rangle$ be a $\mathcal{L}^{1\mathsf{M}}$-model. A ***two-sorted correspondent of $\mathcal{M}$*** is a $\mathcal{L}^{2\mathsf{S}}$-model $\mathfrak{M} = \langle W, D, V \rangle$ such that:

- for all $P \in \mathsf{PRED}^{n/1}$, $V(P) = \{\langle \overline{a}; w \rangle \mid \langle \overline{a} \rangle \in I(P, w)\}$;
- $V(\mathsf{E}) = \{\langle a; w \rangle \in D \times W \mid a \in \delta(w)\}$;
- $V(\mathsf{R}) = R$.

The satisfaction and consequence relations $\vDash$ for $\mathcal{L}^{2\mathsf{S}}$ are just the standard ones for first-order logic with two sorts. We can now translate in the standard way every $\mathcal{L}^{1\mathsf{M}}$-formula into $\mathcal{L}^{2\mathsf{S}}$.

Definition 3.3 *(Standard Translation).* Let φ be a $\mathcal{L}$-formula, and let $s, t \in \mathsf{SVAR}$. The ***standard translation of φ wrt $\langle s, t \rangle$***, $\mathsf{ST}_{s,t}(\varphi)$, is defined recursively:

$\mathsf{ST}_{s,t}(P^n(\overline{x})) = P^n(\overline{x}; t)$	$\mathsf{ST}_{s,t}(@\varphi) = \mathsf{ST}_{s,s}(\varphi)$
$\mathsf{ST}_{s,t}(x \approx y) = x \approx y$	$\mathsf{ST}_{s,t}(\downarrow\varphi) = \mathsf{ST}_{t,t}(\varphi)$
$\mathsf{ST}_{s,t}(\neg\varphi) = \neg\mathsf{ST}_{s,t}(\varphi)$	$\mathsf{ST}_{s,t}(\forall x\, \varphi) = \forall x\, (\mathsf{E}(x; t) \to \mathsf{ST}_{s,t}(\varphi))$
$\mathsf{ST}_{s,t}(\varphi \wedge \psi) = \mathsf{ST}_{s,t}(\varphi) \wedge \mathsf{ST}_{s,t}(\psi)$	$\mathsf{ST}_{s,t}(\forall_{@} x\, \varphi) = \forall x\, (\mathsf{E}(x; s) \to \mathsf{ST}_{s,t}(\varphi))$
$\mathsf{ST}_{s,t}(\Box\varphi) = \forall t'\, (\mathsf{R}(t, t') \to \mathsf{ST}_{s,t'}(\varphi))$	$\mathsf{ST}_{s,t}(\Pi x\, \varphi) = \forall x\, \mathsf{ST}_{s,t}(\varphi)$
$\mathsf{ST}_{s,t}(\mathcal{F}\varphi) = \forall s'\, (\mathsf{R}(s, s') \to \mathsf{ST}_{s',t}(\varphi))$	

where t' is the next state variable not occurring anywhere in $\mathsf{ST}_{s,t}(\varphi)$.

Lemma 3.4 *(Translation).* Let $\mathcal{M}$ be an $\mathcal{L}^{1\mathsf{M}}$-model, $\mathfrak{M}$ a two-sorted correspondent for $\mathcal{M}$, $w, v \in W^{\mathcal{M}}$, $g \in \mathsf{VA}(\mathcal{M})$, $\mathfrak{g} \in \mathsf{VA}(\mathfrak{M})$ (where $\mathfrak{g}(x) = g(x)$ for $x \in \mathsf{VAR}$), $s, t \in \mathsf{SVAR}$, and φ an $\mathcal{L}$-formula. Then $\mathcal{M}, w, v, g \Vdash \varphi$ iff $\mathfrak{M}, \mathfrak{g}^{s,t}_{w,v} \vDash \mathsf{ST}_{s,t}(\varphi)$.

Proof: An easy induction on formulas. ■

With this result, we can define expressivity in the following manner:

Definition 3.5 *(Expressivity)*. A set of $\mathcal{L}$-formulas $\Gamma(\overline{x})$ ***expresses*** an $\mathcal{L}^{2\mathsf{S}}$-formula $\alpha(\overline{x}; s, t)$ if α is equivalent (in $\mathcal{L}^{2\mathsf{S}}$) to $\mathsf{ST}_{s,t}(\Gamma)$ ($= \{\mathsf{ST}_{s,t}(\varphi) \mid \varphi \in \Gamma\}$). A set of $\mathcal{L}$-formulas $\Gamma(\overline{x})$ ***diagonally expresses*** an $\mathcal{L}^{2\mathsf{S}}$-formula $\alpha(\overline{x}; s)$ if α is equivalent to $\mathsf{ST}_{s,s}(\Gamma)$.

In what follows, we will focus on diagonal expressivity for simplicity, noting that the results below apply equally to the more general notion of expressibility.

§4 Bisimulation

We now come to the notion of a bisimulation for ordinary first-order modal logic. This notion can be found in, e.g., van Benthem [2010]; Fine [1981]; Sturm and Wolter [2001]; Yanovich [2015]. However, we add clauses designed to ensure modal equivalence for formulas involving new symbols like @.

Definition 4.1 *(Bisimulation)*. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}^{1\mathsf{M}}$-models. An ***$\mathcal{L}^{1\mathsf{M}}$-bisimulation between $\mathcal{M}$ and $\mathcal{N}$*** is a nonempty multigrade relation Z (so without a fixed arity) such that for all $w, v \in W^{\mathcal{M}}$, all $w', v' \in W^{\mathcal{N}}$, all fininte $\overline{a} \in D^{\mathcal{M}}$, and all finite $\overline{b} \in D^{\mathcal{N}}$, where $|\overline{a}| = |\overline{b}| = n$, we have that $Z(w, v, \overline{a}; w', v', \overline{b})$ implies:

(Atomic) $\forall m \in \mathbb{N}\, \forall P^m \in \mathsf{PRED}^m\, \forall \overline{\alpha}, \overline{\beta}$ where $|\overline{\alpha}| = |\overline{\beta}| = m$, if for each i, there is a $j \leqslant n$ such that $\alpha_i = a_j$ and $\beta_i = b_j$, then: $\langle \overline{\alpha} \rangle \in I^{\mathcal{M}}(P^m, v)$ iff $\langle \overline{\beta} \rangle \in I^{\mathcal{N}}(P^m, v')$

(Zig) $\forall u \in R^{\mathcal{M}}[v]\, \exists u' \in R^{\mathcal{N}}[v']\colon\ Z(w, u, \overline{a}; w', u', \overline{b})$

(Zag) $\forall u' \in R^{\mathcal{N}}[v']\, \exists u \in R^{\mathcal{M}}[v]\colon\ Z(w, u, \overline{a}; w', u', \overline{b})$

(Forth) $\forall \alpha \in \delta^{\mathcal{M}}(v)\, \exists \beta \in \delta^{\mathcal{N}}(v')\colon\ Z(w, v, \overline{a}, \alpha; w', v', \overline{b}, \beta)$

(Back) $\forall \beta \in \delta^{\mathcal{N}}(v')\, \exists \alpha \in \delta^{\mathcal{M}}(v)\colon\ Z(w, v, \overline{a}, \alpha; w', v', \overline{b}, \beta)$.

We may write "$\mathcal{M}, w, v, \overline{a} \leftrightarrows \mathcal{N}, w', v', \overline{b}$" to indicate that there is a bisimulation Z between $\mathcal{M}$ and $\mathcal{N}$ such that $Z(w, v, \overline{a}; w', v', \overline{b})$ (where possibly $|\overline{a}| = |\overline{b}| = 0$). The notion of an ***$\mathcal{L}^{1\mathsf{M}}(S_1, \ldots, S_n)$-bisimulation between $\mathcal{M}$ and $\mathcal{N}$*** is defined similarly, except one must add the condition(s) below corresponding to each S_i:

(Eq) $\forall n, m \leqslant |\overline{a}| : a_n = a_m$ iff $b_n = b_m$

(Act) $Z(w, w, \overline{a}; w', w', \overline{b})$

(Diag) $Z(v, v, \overline{a}; v', v', \overline{b})$

(Fixedly-Zig) $\forall u \in R^{\mathcal{M}}[w]\, \exists u' \in R^{\mathcal{N}}[w']\colon\ Z(u, v, \overline{a}; u', v', \overline{b})$

(Fixedly-Zag) $\forall u' \in R^{\mathcal{N}}[w']\, \exists u \in R^{\mathcal{M}}[w]\colon\ Z(u, v, \overline{a}; u', v', \overline{b})$

($\forall_@$-Forth) $\forall\alpha \in \delta^{\mathcal{M}}(w)\,\exists\beta \in \delta^{\mathcal{N}}(w')$: $Z(w,v,\overline{a},\alpha;w',v',\overline{b},\beta)$

($\forall_@$-Back) $\forall\beta \in \delta^{\mathcal{N}}(w)\,\exists\alpha \in \delta^{\mathcal{M}}(w')$: $Z(w,v,\overline{a},\alpha;w',v',\overline{b},\beta)$

(Π-Forth) $\forall\alpha \in D^{\mathcal{M}}\,\exists\beta \in D^{\mathcal{N}}$: $Z(w,v,\overline{a},\alpha;w',v',\overline{b},\beta)$

(Π-Back) $\forall\beta \in D^{\mathcal{N}}\,\exists\alpha \in D^{\mathcal{M}}$: $Z(w,v,\overline{a},\alpha;w',v',\overline{b},\beta)$.

The (Act), for instance, can be derived as follows. Suppose we introduced a relation $R_@ \subseteq W^2 \times W^2$ into models, and that we treated @ as a normal box operator. We could derive the truth conditions for @ by restricting to the class of models where $wvR_@w'v'$ iff $w = w' = v'$. Then the usual zig-zag clauses for @ just reduce to (Act). The same method applies to the other modal operators.

The standard results regarding bisimulations all carry over straightforwardly:

Definition 4.2 *(Modal Equivalence)*. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}^{1M}$-models, where $w,v \in W^{\mathcal{M}}$, $w',v' \in W^{\mathcal{N}}$, $\overline{a} \in D^{\mathcal{M}}$, and $\overline{b} \in D^{\mathcal{N}}$ (where $|\overline{a}| = |\overline{b}|$). Then $\langle\mathcal{M},w,v,\overline{a}\rangle$ and $\langle\mathcal{N},w',v',\overline{b}\rangle$ are ***$\mathcal{L}$-equivalent*** or ***modally equivalent*** if for all $\mathcal{L}$-formulas $\varphi(\overline{x})$ (where $|\overline{x}| \leqslant |\overline{a}|$), $\mathcal{M},w,v \Vdash \varphi[\overline{a}]$ iff $\mathcal{N},w',v' \Vdash \varphi[\overline{b}]$. In such a case, we may write "$\mathcal{M},w,v,\overline{a} \equiv_{S_1,\ldots,S_n} \mathcal{N},w',v',\overline{b}$", where $\mathcal{L} = \mathcal{L}^{1M}(S_1,\ldots,S_n)$.

Theorem 4.3 *(Bisimulation Implies Modal Equivalence)*. Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}^{1M}$-models, where $w,v \in W^{\mathcal{M}}$, $w',v' \in W^{\mathcal{N}}$, $\overline{a} \in D^{\mathcal{M}}$, $\overline{b} \in D^{\mathcal{N}}$, and $\mathcal{M},w,v,\overline{a} \leftrightarrows_{S_1,\ldots,S_n} \mathcal{N},w',v',\overline{b}$. Then $\mathcal{M},w,v,\overline{a} \equiv_{S_1,\ldots,S_n} \mathcal{N},w',v',\overline{b}$.

Corollary 4.4 *(Translation Implies Invariance)*. Let $\varphi(\overline{x};s,t)$ be an $\mathcal{L}^{2S}$-formula. If φ is equivalent to the translation of some $\mathcal{L}^{1M}(S_1,\ldots,S_n)$-formula, and if we have that $\mathcal{M},w,v,\overline{a} \leftrightarrows_{S_1,\ldots,S_n} \mathcal{N},w',v',\overline{b}$, then for any two-sorted correspondents $\mathfrak{M}$ and $\mathfrak{N}$, $\mathfrak{M} \models \varphi[\overline{a};w,v]$ iff $\mathfrak{N} \models \varphi[\overline{b};w',v']$. Equivalently, if $\langle\mathcal{M},w,v,\overline{a}\rangle$ and $\langle\mathcal{N},w',v',\overline{b}\rangle$ have two-sorted correspondents that disagree on φ, then φ is not expressible as a $\mathcal{L}^{1M}(S_1,\ldots,S_n)$-formula.

§5 Inexpressibility

We now turn to showing that (R) is not expressible in $\mathcal{L}^{1M}(@)$—in fact, not even in $\mathcal{L}^{1M}(\approx,@,\downarrow,\mathcal{F})$. We'll also show that (NR) is not expressible in $\mathcal{L}^{1M}(\approx,@,\Pi)$. In both cases, we construct two bisimilar models that disagree on the two-sorted formalization of the English sentence in question, and then invoke **Corollary 4.4**. We start by presenting a proof that $\mathcal{L}^{1M}(@)$ cannot express (3).

Let $\mathbb{N}^- := \mathbb{Z} - \mathbb{N}$. Our two models $\mathcal{M}_1$ and $\mathcal{M}_2$ are pictured in Figure 1. The global domain of each model is just $\mathbb{Z}$ and the accessibility relation is universal throughout. The world w is our actual world, where every positive integer is rich (top half of circle), and every negative integer is poor (bottom half of circle). For each nonempty finite subset S of $\mathbb{N}$, there is a world v_S where the members of S don't exist, and otherwise the rich and the poor are flipped with respect to w; so at v_S, the negative integers are rich, and the positive integers not in S are poor, and the positive integers in S don't exist. The extension of all other predicates is empty. The only difference betweem $\mathcal{M}_1$ and $\mathcal{M}_2$ is that $\mathcal{M}_2$ includes an additional world $v_\varnothing$, where no integer fails to exist, and where the rich and poor are completely flipped with respect to w.

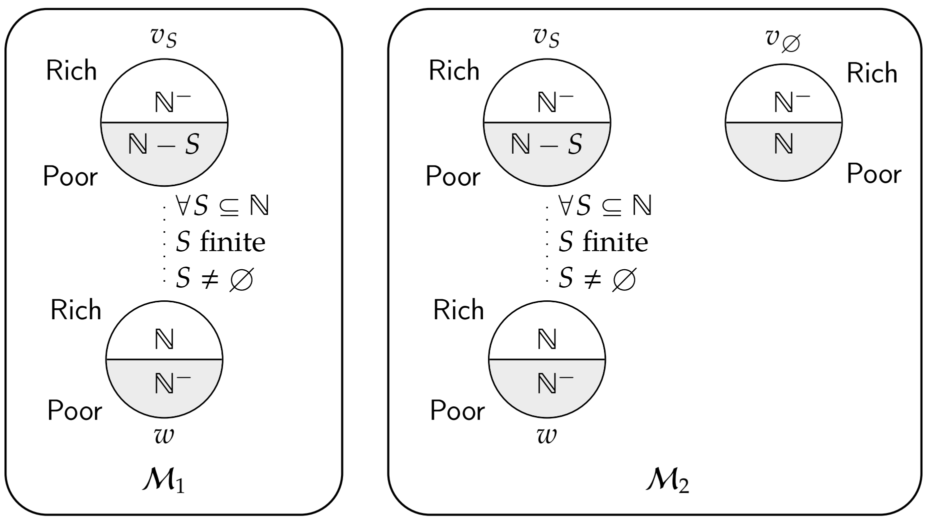

Figure 1: $\mathcal{L}^{1\mathsf{M}}(@)$-bisimilar models disagreeing on (R). The top half of each circle satisfies Rich, while the bottom half satisfies Poor; at each v_S, the members of S do not exist.

$\langle \mathcal{M}_2, w, w\rangle$ satisfies (R), but not $\langle \mathcal{M}_1, w, w\rangle$. But it turns out that $\mathcal{M}_1, w, w \equiv_@ \mathcal{M}_2, w, w$. In fact, $\mathcal{M}_1, w, w \leftrightarrows_@ \mathcal{M}_2, w, w$. The reason is that each v-world looks isomorphic relative to first-order logic to every other v-world since $\mathcal{L}^{1\mathsf{M}}(@)$ can only quantify over the existent objects. So at any given stage of construction of our bisimulation, we can treat each link between worlds and elements as if they're partial segments of an isomorphism between the two worlds considered as first-order models. Of course, we need to make sure that when we shift to new worlds, the elements linked still constitute a partial segment of an isomorphism between the new worlds. But as we'll see, this can be done.

Theorem 5.1 *(Inexpressibility of (R)).* $\mathcal{M}_1, w, w \leftrightarrows_@ \mathcal{M}_2, w, w$. But $\mathcal{M}_2, w, w \Vdash$ (3) even though $\mathcal{M}_1, w, w \nVdash$ (3). Hence, (3) is not expressible in $\mathcal{L}^{1\mathsf{M}}(@)$.

We show explicitly in the appendix how to construct a bisimulation between $\langle \mathcal{M}_1, w, w\rangle$ and $\langle \mathcal{M}_2, w, w\rangle$ in stages. Keeping track of the details is tedious, but the idea is simple. Basically, bisimulations are back-and-forth games that we might have to move to another accessible world to continue playing. So we just need to check that no matter where we move the game in one model, we can find a matching spot to move the game in the other model to keep playing.

Proof (Sketch): Our game starts at $\langle \mathcal{M}_1, w, w\rangle$ and $\langle \mathcal{M}_2, w, w\rangle$. Clearly, if we just play the back-and-forth game there, we'll eventually build an isomorphism. Let's suppose, after moving the game around a bit, we're now playing the back-and-forth game at $\langle \mathcal{M}_1, w, u_1\rangle$ and $\langle \mathcal{M}_2, w, u_2\rangle$, having linked $\overline{a} \in D_1$ to $\overline{b} \in D_2$, where a_i is positive iff b_i is. We'll show that no matter where we move the game in one model, we can move the game somewhere in the other model to keep playing. That is, we'll make sure that, wherever we move, if we want to extend the sequence of elements with a new a (that exists at our new location in $\mathcal{M}_1$), we can find a matching b (that exists at our new location in $\mathcal{M}_2$) such that a is positive iff b is (and similarly if we want to extend the sequence of elements with a new b that exists at our new location in $\mathcal{M}_2$).

Suppose first we move from u_1 to w in $\mathcal{M}_1$. It's easy to show that we can match that move in $\mathcal{M}_2$ by moving from u_2 to w.

Now suppose we move from u_1 to some v_S in $\mathcal{M}_1$. We need to match the move in $\mathcal{M}_2$ with some $v_{S'}$, but we need to do so in such a way so that $\langle \mathcal{M}_1, w, v_S\rangle$ and $\langle \mathcal{M}_2, w, v_{S'}\rangle$ don't disagree over existence between $\overline{a}$ and $\overline{b}$: we don't want a_i to exist at v_S but for b_i not to exist at $v_{S'}$. To get around this, let T be any finite set with the same cardinality as S such that $a_i \in S$ iff $b_i \in T$. Then it's straightforward to show that we can match the move to v_S in $\mathcal{M}_1$ with a move to v_T in $\mathcal{M}_2$. Similarly if we move from u_2 to either w or some v_S in $\mathcal{M}_2$.

Finally, suppose we move from u_2 to $v_\varnothing$ in $\mathcal{M}_2$. The only way to match that move in $\mathcal{M}_1$ is to move to some v_S. We can do this as long as we make sure that $\langle \mathcal{M}_1, w, v_S\rangle$ and $\langle \mathcal{M}_2, w, v_\varnothing\rangle$ don't disagree over existence between $\overline{a}$ and $\overline{b}$. But they won't disagree so long as we move to a v_S where no $a_i \in S$. So if $S \cap \{\overline{a}\} = \varnothing$, then we can match the move from u_2 to $v_\varnothing$ with a move from u_1 to v_S and continue playing. At each stage, it's easy to check that wherever we keep playing, we'll only match positives to positives, and negatives to negatives. ■

Again, the reason this strategy works is essentially because, modulo what exists, the v_S's and $v_\varnothing$ look like isomorphic first-order models, so linked elements can be treated as partial isomorphisms between the worlds. In particular, when we move to $v_\varnothing$, because only finitely many elements are linked at a time, we can always find a matching v_S where all of the linked elements exist, and just keep extending the partial isomorphism as usual. A similar strategy applies in showing that $\mathcal{M}_1, w, w \leftrightarrows_{\approx,\mathsf{E},@,\downarrow,\mathcal{F}} \mathcal{M}_2, w, w$, though the details are messier.

However, this strategy fails when we try to show that $\mathcal{M}_1, w, w \leftrightarrows_{\approx,@,\Pi} \mathcal{M}_2, w, w$. This shouldn't be surprising, since (R) can be expressed as (1). But it's instructive to see why the proof above fails. Consider what happens when we try to guarantee the Forth clause. When we move from u_2 to $v_\varnothing$ in $\mathcal{M}_2$, we try to match that move in $\mathcal{M}_1$ by moving from u_1 to some v_S where $S \cap \{\overline{a}\} = \varnothing$. But the Π-Forth clause says that for any object $a \in D_1$

that we pick, there must be a matching $b \in D_2$. But if we pick a non-existent in v_S, we can be forced to end the game. Since every integer exists at $v_\varnothing$, we must pick a b that exists at $v_\varnothing$. But then by the Back clause, if we picked b again, we would need to match that pick with an a' that exists in v_S. But by the Eq clause, $a' = a$, and a doesn't exist in v_S. So we can't match that pick, and the game is over.

Now we'll show that even $\mathcal{L}^{1M}(\approx, @, \Pi)$ can't express (NR). Consider the two models $\mathcal{N}_1$ and $\mathcal{N}_2$ pictured in Figure 2. Again, the global domain of both models is $\mathbb{Z}$, and the accessibility relation is universal. This time, however, all of $\mathbb{Z}$ exists at every world. Our actual world this time is z, where no integer is either rich or poor. For every finite set $S \subseteq \mathbb{N}$, there's a world $v_{\mathbb{N}-S}$ where all the positive integers are rich except for S, and where all other integers are poor (so our old w is now just $v_\mathbb{N}$). And for every *nonempty* finite set $S \subseteq \mathbb{N}$, there's a world v_S like before, where the rich and poor are flipped with respect to $v_{\mathbb{N}-S}$. Again, the only difference between $\mathcal{N}_1$ and $\mathcal{N}_2$ is the presence of $v_\varnothing$ in $\mathcal{N}_2$, where every negative number is rich, and every positive number is poor.

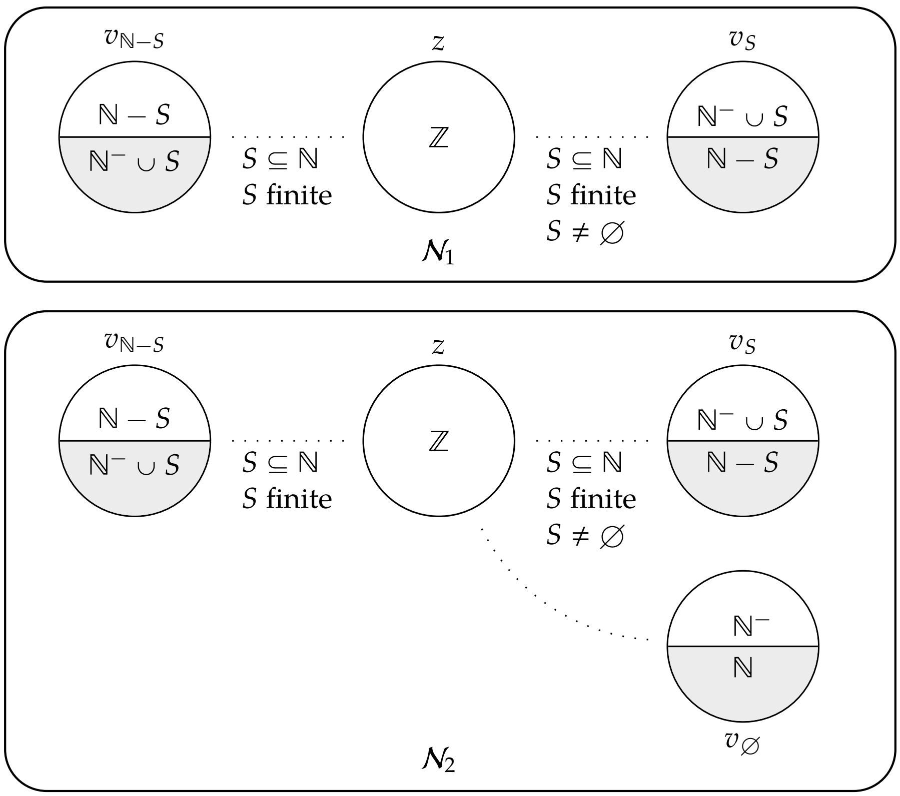

Figure 2: $\mathcal{L}^{1M}(@, \Pi)$-bisimilar models disagreeing on (NR).

$\langle \mathcal{N}_1, z, z \rangle$ and $\langle \mathcal{N}_2, z, z \rangle$ both agree that (3) is true. But they disagree on whether (4) is true; without the presence of $v_\varnothing$, there is no world for $v_\mathbb{N}$ (our old w) where everyone rich in $v_\mathbb{N}$ is poor. Furthermore, $\mathcal{N}_1, z, z \equiv_{\approx, @, \Pi} \mathcal{N}_2, z, z$. For even when we take existence into account, all of the v-worlds are isomorphic to one another. So as long as we're careful to move to the right worlds, we can always keep playing as if we're building an isomorphism between the worlds where the game is taking place. Thus:

Theorem 5.2 *(Inexpressibility of (NR)).* $\mathcal{N}_1, z, z \leftrightarrows_{\approx,@,\Pi} \mathcal{N}_2, z, z$. But $\mathcal{N}_2, z, z \Vdash$ (4) while $\mathcal{N}_1, z, z \nVdash$ (4). Hence, (4) is not expressible in $\mathcal{L}^{1M}(\approx, @, \Pi)$.

The proof is similar to the one before. For instance, suppose we're playing the back-and-forth game at $\langle \mathcal{N}_1, z, u_1 \rangle$ and $\langle \mathcal{N}_2, z, u_2 \rangle$, having linked $\overline{a}$ to $\overline{b}$, and suppose we make a move in $\mathcal{N}_2$ from u_2 to $v_\varnothing$. Then once again, the only matching move we can make in $\mathcal{N}_1$ is from u_1 to some v_S. But we can do this in general, so long as we pick an S that is disjoint from $\{\overline{a}\}$; since in that case, a_i will be in the extension of Rich at v_S iff b_i is in the extension of Rich in $v_\varnothing$. Similar reasoning as that above will show that matching moves can always be made no matter where we jump in the model.

Unlike in the case of (R), however, this inexpressibility proof doesn't extend to languages with $\downarrow$ or with $\mathcal{F}$. We can express (4) in either language with:

$$\Box{\downarrow}\Diamond\Pi x\ (@\mathsf{Rich}(x) \to \mathsf{Poor}(x)) \tag{5}$$

$$\mathcal{F}@\Diamond\Pi x\ (@\mathsf{Rich}(x) \to \mathsf{Poor}(x)). \tag{6}$$

But more complicated sentences can be constructed that reveal the expressive limitations of even languages with $\downarrow$ and $\mathcal{F}$.

§6 Conclusion

It has often been noted, without proof, that (R) and (NR) are not expressible in $\mathcal{L}^{1M}$, even when one adds an actually operator Bricker [1989]; Cresswell [1990]; Fara and Williamson [2005]; Hazen [1976]; Sider [2010]. Proofs of this claim can be found in Hodes [1984b,a], but involve rather complicated Henkin constructions that don't seem to illuminate the source of inexpressibility. In this paper, we've provided a simpler and more convenient method of proving inexpressibility results in $\mathcal{L}^{1M}$ using a modular notion of bisimulation. We've seen that inexpressibility proofs via bisimulation are illuminating as they reveal the ways in which $\mathcal{L}^{1M}$ can be insensitive to the location of certain back-and-forth games.

Some questions naturally arise from these results. First, is there a more general formal characterization of sentences like (E), (R), (NE), and (NR)? One syntactic characterization was proposed in Kocurek [2015], but it's open to debate whether this characterization is accurate, or whether there is also a nice model-theoretic characterization of this class.

Second, is there a language weaker than $\mathcal{L}^{2S}$ that can express these kinds of sentences? It has been argued by Bricker [1989] that adding second-order quantifiers suffices. In Cresswell [1990], Cresswell defined a language (that happens to be a notational variant of a quantified hybrid language without state variables as formulas) which he argued also suffices to express these kinds of sentences.[5] In both cases, heuristic arguments are given in support of the claim that these languages can express any sentence of the same kind as (E),

[5]However, Cresswell [1990] also shows that if $\approx$ is dropped and R is universal, then this language is as expressively powerful as $\mathcal{L}^{2S}$ without $\approx$.

(R), (NE), and (NR). But without an answer to the first question, no formal proof of these claims can provided.[6]

§A Proof of Theorem 5.1

To prove **Theorem 5.1**, we introduce a helpful definition:

Definition A.1 *(Partial Isomorphism)*. Let $\mathcal{M}$ and $\mathcal{N}$ be modal models, let $w, v \in W^{\mathcal{M}}$, and let $w', v' \in W^{\mathcal{N}}$. A ***partial*** $\mathcal{L}^{\mathbf{1M}}$***-isomorphism between*** $\langle \mathcal{M}, w, v\rangle$ ***and*** $\langle \mathcal{N}, w', v'\rangle$ is a finite injective map $\rho \colon D \to D'$ such that:

(Predicate) $\forall m \, \forall P^m \in \mathsf{PRED}^m \, \forall a_1, \ldots, a_m \in \operatorname{dom}(\rho) : \langle a_1, \ldots, a_m\rangle \in I^{\mathcal{M}}(P^m, v)$ iff $\langle \rho(a_1), \ldots, \rho(a_m)\rangle \in I^{\mathcal{N}}(P^m, v')$

(Existence) $\forall a \in \operatorname{dom}(\rho) : \; a \in \delta^{\mathcal{M}}(v)$ iff $\rho(a) \in \delta^{\mathcal{N}}(v')$.

The set of partial isomorphisms between $\langle \mathcal{M}, w, v\rangle$ and $\langle \mathcal{N}, w', v'\rangle$ will be $\mathrm{PAR}^{\mathcal{M},w,v}_{\mathcal{N},w',v'}$. When the $\mathcal{M}$ and $\mathcal{N}$ are clear, we'll drop mention of them.

Now, at stage 0, set $Z_0 = \{\langle w, w; w, w\rangle\}$. Next, define the following:

$$Z_i^{\mathrm{Act}} = \left\{\langle w, w, \bar{a}; w, w, \rho(\bar{a})\rangle \;\middle|\; \exists u, u' \colon \; \langle w, u, \bar{a}; w, u', \rho(\bar{a})\rangle \in Z_i \text{ and } \rho \in \mathrm{PAR}^{w,u}_{w,u'} \right\}$$

$$Z_i^{\mathrm{Zig}} = \left\{ \langle w, v_S, \bar{a}; w, v_{\rho'[S]}, \rho'(\bar{a})\rangle \;\middle|\; \begin{array}{l} \exists u, u' \colon \; \langle w, u, \bar{a}; w, u', \rho(\bar{a})\rangle \in Z_i, \text{ where} \\ \rho \in \mathrm{PAR}^{w,u}_{w,u'} \text{ and } \rho \subseteq \rho' \in \mathrm{PAR}^{w,v_S}_{w,v_{\rho'[S]}} \text{ and} \\ \operatorname{dom}(\rho') \supseteq S \end{array} \right\}$$

$$Z_i^{\mathrm{Zag}} = \left\{ \left\langle w, v_{\rho'^{-1}[S]}, \bar{a}; w, v_S, \rho'(\bar{a})\right\rangle \;\middle|\; \begin{array}{l} \exists u, u' \colon \; \langle w, u, \bar{a}; w, u', \rho(\bar{a})\rangle \in Z_i, \text{ where} \\ \rho \in \mathrm{PAR}^{w,u}_{w,u'} \text{ and } \rho \subseteq \rho' \in \mathrm{PAR}^{w,v_{\rho'^{-1}[S]}}_{w,v_S} \text{ and} \\ \operatorname{ran}(\rho') \supseteq S \end{array} \right\}$$
$$\cup \left\{ \langle w, v_S, \bar{a}; w, v_\varnothing, \rho(\bar{a})\rangle \;\middle|\; \begin{array}{l} \exists u, u' \colon \; \langle w, u, \bar{a}; w, u', \rho(\bar{a})\rangle \in Z_i \text{ and} \\ \rho \in \mathrm{PAR}^{w,u}_{w,u'}, \text{ where } S \cap \operatorname{dom}(\rho) = \varnothing \end{array} \right\}$$

$$Z_i^{\mathrm{Forth}} = \left\{ \langle w, u, \bar{a}, b; w, u', \rho'(\bar{a}), \rho'(b)\rangle \;\middle|\; \begin{array}{l} \langle w, u, \bar{a}; w, u', \rho(\bar{a})\rangle \in Z_i, \text{ where } \rho \subseteq \rho' \in \\ \mathrm{PAR}^{w,u}_{w,u'} \text{ and } b \in \delta_1(u) \cap \operatorname{dom}(\rho') \end{array} \right\}$$

$$Z_i^{\mathrm{Back}} = \left\{ \langle w, u, \bar{a}, \rho'^{-1}(b); w, u', \rho'(\bar{a}), b\rangle \;\middle|\; \begin{array}{l} \langle w, u, \bar{a}; w, u', \rho(\bar{a})\rangle \in Z_i, \text{ where } \rho \subseteq \\ \rho' \in \mathrm{PAR}^{w,u}_{w,u'} \text{ and } b \in \delta_2(u') \cap \operatorname{ran}(\rho') \end{array} \right\}.$$

Then set: $Z_{i+1} = Z_i \cup Z_i^{\mathrm{Act}} \cup Z_i^{\mathrm{Zig}} \cup Z_i^{\mathrm{Zag}} \cup Z_i^{\mathrm{Forth}} \cup Z_i^{\mathrm{Back}}$. Finally, set $Z = \bigcup_{i \in \omega} Z_i$.

Lemma A.2 *(Trivial Observations)*. If $\langle w, u, \bar{a}; w, u', \bar{b}\rangle \in Z_i$, then $u = w$ iff $u' = w$, and if $\rho \in \mathrm{PAR}^{w,u}_{w,u'}$, then $a_i \in \mathbb{N}$ iff $\rho(a_i) \in \mathbb{N}$.

[6] See Kocurek [2015] for one possible formal answer to this question.

Lemma A.3 *(Partial Isomorphisms in Z)*. For all $i \geqslant 0$, and all $\langle w, u, \overline{a}; w, u', \overline{b} \rangle \in Z_i$, there is a partial $\mathcal{L}^{1\mathsf{M}}$-isomorphism ρ between $\langle \mathcal{M}_1, w, u \rangle$ and $\langle \mathcal{M}_2, w, u' \rangle$ such that $\rho(a_k) = b_k$ for $1 \leqslant k \leqslant |\overline{a}|$.

Proof (Sketch): By induction on i. This clearly holds for the $i = 0$ case. Now suppose that every member of Z_i has the stated property. Show that for each condition C, every member of Z_i^{C} has the property, from which it will follow that every member of Z_{i+1} has the property. This is automatically guaranteed for Z_i^{Zig}, Z_i^{Back}, and Z_i^{Forth}. Z_i^{Zag} is almost immediate, but elements from the second listed set must be checked. Checking Z_i^{Act} is tedious, but straightforward. ∎

We now turn to the proof of **Theorem 5.1**.

*Proof (**Theorem 5.1**)*: Let $\langle w, u, \overline{a}; w, u', \overline{b} \rangle \in Z$. Then $\langle w, u, \overline{a}; w, u', \overline{b} \rangle \in Z_i$. By **Lemma A.3**, there's a partial isomorphism ρ between $\langle \mathcal{M}_1, w, u \rangle$ and $\langle \mathcal{M}_2, w, u' \rangle$ such that $\rho(\overline{a}) = \overline{b}$. Hence, (Atomic) is met. As for the other conditions:

Act: By definition of Z_i^{Act}, $\langle w, w, \overline{a}; w, w, \overline{b} \rangle \in Z_{i+1}$. ✓

Zig: If u moves to w, then this case is covered by the Act-case. So suppose instead u moves to some v_S. It suffices to show that $\langle w, v_S, \overline{a}; w, v_{\rho'[S]}, \rho'(\overline{a}) \rangle \in Z_i^{\mathrm{Zig}}$ for some suitable $\rho' \supseteq \rho$. If $S \subseteq \mathrm{dom}\,(\rho)$, then let $\rho' = \rho$. Otherwise, let $b_1, \ldots, b_n \in S - \mathrm{dom}\,(\rho)$. Pick the least $b'_1, \ldots, b'_n \in \mathbb{N} - \mathrm{ran}\,(\rho)$ and set $\rho' = \rho \cup \{\langle b_i, b'_i \rangle \mid 1 \leqslant i \leqslant n\}$. It suffices to show that $\rho' \in \mathrm{PAR}^{w, v_S}_{w, v_{\rho'[S]}}$. Let $a \in \mathrm{dom}\,(\rho')$. By **Lemma A.2**, $a \in I_1(\mathsf{Rich}, v_S)$ iff $\rho(a) \in I_2(\mathsf{Rich}, v_{\rho'[S]})$. As for Poor, $a \in I_1(\mathsf{Poor}, v_S)$ iff $a \in \mathbb{N} - S$ iff (by injectivity) $\rho'(a) \in \mathbb{N} - \rho'[S]$ iff $a \in I_2(\mathsf{Poor}, v_{\rho'[S]})$. ✓

Zag: We just need to check the case where u' moves to $v_\varnothing$. But by definition, for any S such that $S \cap \mathrm{dom}\,(\rho) = \varnothing$ (which will exist since $\mathrm{dom}\,(\rho)$ is finite), $\langle w, v_S, \overline{a}; w, v_\varnothing, \rho(\overline{a}) \rangle \in Z_i^{\mathrm{Zag}}$. ✓

Forth: Let $b \in \delta(u)$. WLOG, assume $b \notin \mathrm{dom}\,(\rho)$. If $b \in \mathbb{N}^-$, then just let b' be the least element in $\mathbb{N}^- - \mathrm{ran}\,(\rho)$. Otherwise, let b' be the least element in $\mathbb{N} - \mathrm{ran}\,(\rho)$. There are only three cases to consider:

(i) $u = u' = w$. By **Lemma A.2**, $b \in I_1(\mathsf{Rich}, w)$ iff $b' \in I_2(\mathsf{Rich}, w)$. ✓

(ii) $u = v_S$ and $u' = v_{\rho[S]}$. Then $b \notin S$ and thus $b' \notin \rho'[S]$ (since, according to our construction, $S \subseteq \mathrm{dom}\,(\rho)$, and so $\rho[S] = \rho'[S]$). So $b \in I_1(\mathsf{Rich}, v_S)$ iff $b' \in I_2(\mathsf{Rich}, v_{\rho[S]})$. ✓

(iii) $u = v_S$ and $u' = v_\varnothing$. Since $b \in \delta_1(u)$, (Existence) is still upheld. And again, $b \in I_1(\mathsf{Rich}, v_S)$ iff $b' \in I_2(\mathsf{Rich}, v_\varnothing)$. ✓

Back: As above, except if $b' \in \mathbb{N}$, you pick the least $b \in \mathbb{N} - \mathrm{dom}(\rho)$ in all cases except where $u = v_S$ and $u' = v_\varnothing$, in which case, you pick the least $b \in \mathbb{N} - (\mathrm{dom}(\rho) \cup S)$. ✓ ∎

References

Bricker, Phillip. 1989. "Quantified Modal Logic and the Plural De Re." *Midwest Studies in Philosophy* 14:372–394.

Cresswell, Maxwell J. 1990. *Entities and Indices*. Kluwer Academic Publishers.

Crossley, John N and Humberstone, Lloyd. 1977. "The Logic of "Actually"." *Reports on Mathematical Logic* 8:11–29.

Davies, Martin and Humberstone, Lloyd. 1980. "Two Notions of Necessity." *Philosophical Studies* 38:1–30.

Fara, Michael and Williamson, Timothy. 2005. "Counterparts and Actuality." *Mind* 114:1–30.

Fine, Kit. 1981. "Model Theory for Modal Logic - Part III Existence and Predication." *Journal of Philosophical Logic* 10:293–307.

Garson, James W. 2001. "Quantification in Modal Logic." In *Handbook of Philosophical Logic*, 267–323. Dordrecht: Springer Netherlands.

Hazen, Allen P. 1976. "Expressive Completeness in Modal Language." *Journal of Philosophical Logic* 5:25–46.

—. 1990. "Actuality and Quantification." *Notre Dame Journal of Formal Logic* 31:498–508.

Hodes, Harold T. 1984a. "On Modal Logics which Enrich First-order S5." *Journal of Philosophical Logic* 13:423–454.

—. 1984b. "Some Theorems on the Expressive Limitations of Modal Languages." *Journal of Philosophical Logic* 13:13–26.

Kocurek, Alexander W. 2015. "On the Expressivity of First-Order Modal Logic with "Actually"." In *Logic, Rationality, and Interaction*, 207–219. Berlin, Heidelberg: Springer Berlin Heidelberg.

Lewis, David K. 1973. "Counterfactuals and Comparative Possibility." *Journal of Philosophical Logic* 2:418–446.

Sider, Theodore. 2010. *Logic for Philosophy*. Oxford University Press.

Sturm, Holger and Wolter, Frank. 2001. "First-order Expressivity for S5-models: Modal vs. Two-sorted Languages." *Journal of Philosophical Logic* 30:571–591.

van Benthem, Johan. 1977. *Modal Correspondence Theory*. Ph.D. thesis.

—. 2010. "Frame Correspondences in Modal Predicate Logic." 1–14.

Wehmeier, Kai F. 2001. "World Travelling and Mood Swings." In *Trends in Logic*, 257–260. Dordrecht: Springer Netherlands.

Yanovich, Igor. 2015. "Expressive Power of "Now" and "Then" Operators." *Journal of Logic, Language and Information* 24:65–93.