Translation Invariance and Miller's Weather Example

Paris, Jeff and Vencovska, Alena

2017

MIMS EPrint: **2017.19**

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

Translation Invariance and Miller's Weather Example.

J.B.Paris[*]and A.Vencovská[†‡]

School of Mathematics

The University of Manchester

Manchester M13 9PL

jeff.paris@manchester.ac.uk   alena.vencovska@manchester.ac.uk

May 17, 2017

Abstract

In his 1974 paper "Popper's Qualitative Theory of Verisimilitude" published in the *British Journal for the Philosophy of Science* David Miller gave his so called 'Weather Example' to argue that the Hamming distance between constituents is flawed as a measure of proximity to truth since the former is not, unlike the latter, translation invariant. In this present paper we generalise David Miller's Weather Example in both the unary and polyadic cases, characterising precisely which permutations of constituents/atoms can be effected by translations. In turn this suggests a meta-principle of the rational assignment of subjective probabilities, that rational principles should be preserved under translations, which we formalise and give a particular characterisation of in the context of Unary Pure Inductive Logic.

1

Introduction

In [2] (continued in [3, Chapter 11]) David Miller gave an example to show that the Hamming distance between constituents/atoms is not a good indicator of closeness to the truth, verisimilitude. Paraphrased it runs as follows.

> *Jones and Smith are closeted away in prison and try to guess the weather outside. Jones thinks it is cool and dry and still whilst Smith also thinks it is cool but otherwise differs in thinking it rainy and windy. They subsequently learn that actually it is hot, rainy and windy. So Smith is right on two scores (rainy and windy) while Jones is wrong on all scores. From this we might conclude that Smith's guess is closer to the truth than Jones'. But suppose we now replace the propositions 'hot', 'rainy' and 'windy' by the equally expressive 'hot', 'Minnesotan (meaning hot and wet or cool and dry) and Arizonan (meaning hot and windy or cool and still). In that case Smith's guess becomes cool, not Minnesotan and not Arizonan, Jones' becomes cool, Minnesotan and Arizonan. The actual situation is hot and Minnesotan and Arizonan. So now it is Jones who is right on two scores and Smith who is right on none!*

The primary observation in this paper is that there is a general result (and conclusion) behind this example. We will first show this in the case of a unary (i.e. monadic) predicate language (where it may well have no great novelty) and then extend our results to a general polyadic (relational) language. On the way we will uncover a 'meta-principle' of probability assignment which we will investigate in the unary case.

To make precise the context we shall work in let $L_{\vec{P}}$ be a predicate language[1] with unary relation (i.e. predicate) symbols $P_1, P_2, \ldots, P_q$ (or propositional variables if you want to work in Propositional Logic) and, for later applications, constant symbols $a_1, a_2, a_3, \ldots$. Let $L_{\vec{P}}^-$ be the language $L_{\vec{P}}$ without the constants a_i. The *atoms*[2] of $L_{\vec{P}}$ are the 2^q formulae $\alpha_\epsilon(x)$ of the form

$$\bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x)$$

where $\epsilon : \{1, 2, \ldots, q\} \to \{0, 1\}$ and for a formula ϕ, $\phi^1 = \phi$, $\phi^0 = \neg\phi$. For shorthand let $\Omega_{\vec{P}}$ denote the set of maps from $\{1, 2, \ldots, q\}$ to $\{0, 1\}$.

---

[1] Miller's example is set within a propositional language. Working within a predicate language will however allows us to generalize it, in particular to polyadic languages.

[2] Note that atoms are not the same thing as what is commonly called atomic formulae of the language. Carnap refers to them as *Q-predicates* and Hintika and Niiniluoto as *attributive constituents*.

An atom then is specified by a map ϵ and in such situations the Hamming distance between them is defined to be the Hamming distance between these maps, that is the number of arguments on which they give different values. Atoms correspond to *constituents of the propositional language*, propositions of the form $\bigwedge_{i=1}^{q} P_i^{\epsilon(i)}$, also specified by $\epsilon \in \Omega_{\vec{P}}$.

A *translation* of the predicate symbols $P_1, P_2, \ldots, P_q$ is a family $\phi_1(x), \phi_2(x), \ldots, \phi_q(x)$ of quantifier free formulae of $L_{\vec{P}}^{-}$ such that as δ ranges over $\Omega_{\vec{P}}$ so the 2^q formulae

$$\alpha_\delta^{\vec{\phi}}(x) = \bigwedge_{i=1}^{q} \phi_i^{\delta(i)}(x)$$

range through all of the 2^q atoms.[3]

So if we imagine that the $P_1(x), P_2(x), \ldots, P_q(x)$ describe certain features of, say, the weather at location x, then the $\phi_1(x), \phi_2(x), \ldots, \phi_q(x)$ would, as in Miller's example, provide an alternative, and exactly as descriptive, way to describe the weather.

Clearly then a translation $\phi_1(x), \phi_2(x), \ldots, \phi_q(x)$ determines a permutation of $\Omega_{\vec{P}}$, which we shall denote $\tau_{\vec{\phi}}$, given by[4]

$$\tau_{\vec{\phi}}(\epsilon) = \delta \quad \Longleftrightarrow \quad \bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x) = \bigwedge_{i=1}^{q} \phi_i^{\delta(i)}(x)$$

$$\Longleftrightarrow \quad \alpha_\epsilon(x) = \alpha_\delta^{\vec{\phi}}(x). \tag{1}$$

In this case we say that the permutation $\tau = \tau_{\vec{\phi}}$ *is supported* by the translation $\vec{\phi}$.

Conversely, any permutation τ of $\Omega_{\vec{P}}$ determines (up to logical equivalence) formulae $\phi_i(x)$, $i = 1, \ldots, q$, such that $\vec{\phi}$ supports τ since if we define

$$\phi_j(x) = \bigvee_{\tau(\epsilon)(j)=1} \bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x).$$

then, noting that

$$\neg \phi_j(x) = \phi_j^0(x) = \bigvee_{\tau(\epsilon)(j)=0} \bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x),$$

for any $\epsilon \in \Omega_{\vec{P}}$ we have

$$\bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x) \models \bigwedge_{j=1}^{q} \phi_j^{\tau(\epsilon)(j)}(x)$$

---

[3]Up to logical equivalence. Throughout we will usually, for convenience, identify formulae which are logically equivalent rather than actually syntactically identical.

[4]Recall our convention of identifying formulae even if they are formally only logically equivalent.

so since the $\bigwedge_{j=1}^{q} \phi_j^{\tau(\epsilon)(j)}(x)$ (with $\epsilon \in \Omega_{\vec{P}}$ and τ fixed) are disjoint and exhaustive,

$$\alpha_\epsilon(x) = \bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x) = \bigwedge_{j=1}^{q} \phi_j^{\tau(\epsilon)(j)}(x) = \alpha_{\tau(\epsilon)}^{\vec{\phi}}(x).$$

We have actually shown here that:

Theorem 1. *Let τ be a permutation of $\Omega_{\vec{P}}$. Then there is a translation $\vec{\phi} = \langle \phi_1(x), \phi_2(x), \ldots, \phi_q(x) \rangle$ that supports τ. Conversely, any translation $\vec{\phi}$ supports a permutation of $\Omega_{\vec{P}}$ (namely $\tau_{\vec{\phi}}$).*

Putting this another way, and more in line with Miller's weather example, if we define new unary predicates $Q_1(x), Q_2(x), \ldots, Q_q(x)$ by $Q_j(x) = \phi_j(x)$ then for each $\epsilon, \delta \in \Omega_{\vec{P}}$, the atom

$$\alpha_\epsilon(x) = \bigwedge_{i=1}^{q} P_i^{\epsilon(i)}(x) \tag{2}$$

of $L_{\vec{P}}$ is logically equivalent to the atom

$$\alpha_\delta^{\vec{Q}}(x) = \bigwedge_{i=1}^{q} Q_i^{\delta(i)}(x) \tag{3}$$

of $L_{\vec{Q}}$ just when $\tau_{\vec{\phi}}(\epsilon) = \delta$.

Thus we can change the Hamming Distance between $\alpha_\epsilon(x)$ and $\alpha_\delta(x)$ (for $\delta \neq \epsilon$) by a translation in any way we want (without making it 0). Indeed provided that it is consistent (i.e. there is *some* tuple of atoms with the desired Hamming distances between them) we can in fact make multiple distance changes simultaneously. In Miller's example, writing C, D, S and M, A for 'cool, dry, still' and Minnesotan, Arizonan respectively, (meaning the weather outside Jones and Smith's prison), the situation may be represented as

Jones	C	D	C		Jones	C	M	A
Smith	C	$\neg D$	$\neg S$		Smith	C	$\neg M$	$\neg A$
True	$\neg C$	$\neg D$	$\neg S$		True	$\neg C$	M	A

respectively. We can, for example, also arrange for Jones and Smith each to differ from the true situation and from each other on two scores although we cannot arrange that they each differ from the true situation and each other on a single score. The conclusion we may draw here, in line with Miller, is that Hamming Distance is as flawed a measure of verisimilitude for constituents (in the propositional context) or instantiated atoms (in the predicate context) as can be.

Given a permutation τ of $\Omega_{\vec{P}}$ and assuming $Q_1, \ldots, Q_q$ stand for unary predicates, we can treat τ as a *renaming* of atoms which sends (2) to

$$\alpha_{\tau(\epsilon)}^{\vec{Q}}(x) = \bigwedge_{i=1}^{q} Q_i^{\tau(\epsilon)(i)}(x).$$

4

Theorem 1 has shown that in the unary context renamings of atoms and translations are the same thing, and in turn this somewhat broadens Miller's example. This raises the question of what happens when we move to polyadic languages (i.e. containing possibly binary, ternary, etc. relation symbols). We shall consider this in the next section after which we will return again to the unary context to investigate a further issue raised by Theorem 1.

Before progressing further, note that a translation/renaming as above from $L_{\vec{P}}$ to $L_{\vec{Q}}$, that sends the atom $\alpha_\epsilon(x)$ to the atom

$$\alpha^{\vec{Q}}_{\tau_{\vec{\phi}}(\epsilon)}(x),$$

which is logically equivalent to $\alpha_\epsilon(x)$ when the Q_j are defined by $Q_j(x) = \phi_j(x)$, extends to all formulae of $L_{\vec{P}}^-$, and hence also to all sentences of $L_{\vec{P}}$. This can be seen for example by noting that a formula $\xi(x_1, \ldots, x_n)$ of $L_{\vec{P}}^-$ is logically equivalent to a disjunction of the form

$$\bigvee_{j=1}^{t} \left(\bigwedge_{k=1}^{n} \alpha_{\epsilon_{j,k}}(x_i) \wedge \bigwedge_{\epsilon \in \Omega_{\vec{P}}} (\exists x \, \alpha_\epsilon(x))^{h_{j,\epsilon}} \right) \tag{4}$$

where the $\epsilon_{j,k} \in \Omega_{\vec{P}}$ and $h_{j,\epsilon} \in \{0,1\}$ and consequently, $\xi(x_1, \ldots, x_n)$ is also logically equivalent to the formula

$$\bigvee_{j=1}^{t} \left(\bigwedge_{k=1}^{n} \alpha^{\vec{Q}}_{\tau_{\vec{\phi}}(\epsilon_{j,k})}(x_i) \wedge \bigwedge_{\epsilon \in \Omega_{\vec{P}}} \left(\exists x \, \alpha^{\vec{Q}}_{\tau_{\vec{\phi}}(\epsilon)}(x) \right)^{h_{j,\epsilon}} \right)$$

of $L_{\vec{Q}}^-$. Note that $\phi_j(x)$ translates as $Q_j(x)$.

In [4] Miller also considers a related question of invariance under 'translations' of the Hamming distance between *constituents of the predicate language* $L_{\vec{P}}^-$, that is in our notation sentences of $L_{\vec{P}}^-$ of the form

$$\bigwedge_{\epsilon \in \Omega_{\vec{P}}} (\exists x \, \alpha_\epsilon(x))^{\tau(\epsilon)} \tag{5}$$

where $\tau : \Omega_{\vec{P}} \to \{0,1\}$ is not constantly zero. Let $\Omega_{\Omega_{\vec{P}}}$ be the set of such τ. Clearly, these distances are preserved under the translations we are considering, that is, generated by permutations of atoms. However as Miller demonstrates in [4] there are 'translations' - sets of disjoint exhaustive formulae $\varphi_\epsilon(x)$ for $\epsilon \in \Omega_{\vec{P}}$ of $L_{\vec{P}}^-$ with quantifiers - such that for $\tau \in \Omega_{\Omega_{\vec{P}}}$

$$\bigwedge_{\epsilon \in \Omega_{\vec{P}}} (\exists x \, \alpha_\epsilon(x))^{\tau(\epsilon)} = \bigwedge_{\epsilon \in \Omega_{\vec{P}}} (\exists x \, \varphi_\epsilon(x))^{\nu(\tau)(\epsilon)} \tag{6}$$

where $\nu : \Omega_{\Omega_{\vec{P}}} \to \Omega_{\Omega_{\vec{P}}}$ does not preserve Hamming distance, though necessarily, by considering smallest models, ν must preserve $|\{\epsilon \,|\, \tau(\epsilon) = 1\}|)$. In fact this condition exactly characterises which ν are possible here, as we shall now show.

5

Let $\nu : \Omega_{\Omega_{\vec{P}}} \to \Omega_{\Omega_{\vec{P}}}$ be such that for each $\tau \in \Omega_{\Omega_{\vec{P}}}$

$$|\tau^{-1}\{1\}| = |\nu(\tau)^{-1}\{1\}|.$$

For each $\tau \in \Omega_{\Omega_{\vec{P}}}$ let

$$\iota_\tau : \nu(\tau)^{-1}\{1\} \to \tau^{-1}\{1\}$$

be an injection and for $\epsilon \in \Omega_{\vec{P}}$ set

$$\varphi_\epsilon(x) = \bigvee_{\nu(\tau)(\epsilon)=1} \left(\alpha_{\iota_\tau(\epsilon)}(x) \wedge \bigwedge_\delta \left(\exists y\, \alpha_\delta(y)\right)^{\tau(\delta)} \right). \tag{7}$$

Then these $\varphi_\epsilon(x)$ are disjoint (since the disjuncts in (7) for $\epsilon, \epsilon' \in \Omega_{\vec{P}}$ could only meet for the same τ but $\iota_\tau(\epsilon) \neq \iota_\tau(\epsilon')$ when $\epsilon \neq \epsilon'$) and, as will follow from (8) below by taking the disjunction over τ, exhaustive.

From (7) it follows that

$$\exists x\, \varphi_\epsilon(x) = \bigvee_{\nu(\tau)(\epsilon)=1} \bigwedge_\delta \left(\exists y\, \alpha_\delta(y)\right)^{\tau(\delta)}$$

and in turn, as required, that for each τ

$$\bigwedge_\epsilon \left(\exists x\, \varphi_\epsilon(x)\right)^{\nu(\tau)(\epsilon)} = \bigwedge_\delta \left(\exists x\, \alpha_\delta(x)\right)^{\tau(\delta)}. \tag{8}$$

The General Polyadic Case

As a simple motivating example, consider Jones and Smith again. Speculating about weather and police behaviour, Jones thinks it is hot in London but not in Belfast and that there are London-trained police dogs being deployed in Belfast but not conversely. Smith thinks that it is not hot in London but hot in Belfast and that there are Belfast-trained police dogs being deployed in London but not conversely. In truth, it is hot both in London and Belfast and there are London-trained police dogs being deployed in Belfast but not conversely. Let $H, D,$ and B stand for, respectively,

$H(x)\ \ \ldots\ \ $ *It is hot in x.*

$D(x, y)\ \ \ldots\ \ $ *There are x-trained police dogs being deployed in y.*

$B(x, y)\ \ \ldots\ \ $ *It is not both hot in x and hot in y, and there are x-trained police dogs being deployed in y; or it is both hot in x and hot in y, and there are y-trained police dogs being deployed in x.*

This means that in terms of H and D, with l and b standing for the cities, the respective positions are

Jones	$H(l)$	$\neg H(b)$	$D(l, b)$	$\neg D(b, l)$
Smith	$\neg H(l)$	$H(b)$	$\neg D(l, b)$	$D(b, l)$
True	$H(l)$	$H(b)$	$D(l, b)$	$\neg D(b, l)$

so Jones is 'closer to the truth',[5] but in terms of H and B they are

Jones	$H(l)$	$\neg H(b)$	$B(l,b)$	$\neg B(b,l)$
Smith	$\neg H(l)$	$H(b)$	$\neg B(l,b)$	$B(b,l)$
True	$H(l)$	$H(b)$	$\neg B(l,b)$	$B(b,l)$

so Smith is 'closer'. This provides an example similar to the above weather example, which however involves a binary relation rather than just unary predicates/ propositional variables. By the end of this section we shall justify in some detail why H, B is an equally expressive pair as H, D.

Let the polyadic language $L_{\vec{R}}$ contain relation symbols $R_1, \ldots, R_q$ of arities $r_1, \ldots, r_q$ respectively and constants $a_1, a_2, \ldots$ (needed later). Let $L_{\vec{R}}^-$ stand for the language $L_{\vec{R}}$ without the constants a_i. A *state formula* of $L_{\vec{R}}$ for variables $x_1, \ldots, x_n$ is a formula

$$\Theta(x_1, \ldots, x_n) = \bigwedge_{\substack{i \in \{1,2,\ldots,q\} \\ \langle j_1, \ldots, j_{r_i} \rangle \in \{1,2,\ldots,n\}^{r_i}}} R_i^{\epsilon(i,j_1,\ldots,j_{r_i})}(x_{j_1} \ldots, x_{j_{r_i}}) \qquad (9)$$

of $L_{\vec{R}}^-$, where the $\epsilon(i, j_1, \ldots, j_{r_i}) \in \{0,1\}$. If $\Theta(x_1, \ldots, x_n)$ is a state formula and the (distinct) $b_1, \ldots, b_n$ are from $a_1, a_2, \ldots$ then $\Theta(b_1, \ldots, b_n)$ is called a *state description (for $b_1, \ldots, b_n$)* .

Let $r = \max\{r_1, \ldots, r_q\}$. *An atom of $L_{\vec{R}}$* is a state formula for r variables. Hence an atom of $L_{\vec{R}}$ is determined by a map

$$\epsilon : \bigcup_{i=1}^q \left(\{i\} \times \{1, 2, \ldots, r\}^{r_i} \right) \to \{0, 1\}.$$

Mimicking the notation of the previous section let $\Omega_{\vec{R}}$ denote the set of such maps ϵ and for $\epsilon \in \Omega_{\vec{R}}$ let α_ϵ denote the atom determined by ϵ. Permutations of atoms are identified with the corresponding permutations of $\Omega_{\vec{R}}$. We shall similarly use $\epsilon, \delta, \ldots$ for elements of $\Omega_{\vec{R}}$.

We say that a q-tuple of quantifier free formulae $\psi_i(x_1, \ldots x_{r_i})$ forms a *translation* of $L_{\vec{R}}^-$ if the

$$\alpha_\epsilon^{\vec{\psi}}(x_1, \ldots, x_r) = \bigwedge_{\substack{i \in \{1,2,\ldots,q\} \\ \langle j_1, \ldots, j_{r_i} \rangle \in \{1,2,\ldots,r\}^{r_i}}} \psi_i^{\epsilon(i,j_1,\ldots,j_{r_i})}(x_{j_1} \ldots, x_{j_{r_i}}) \qquad (10)$$

run through all the atoms of $L_{\vec{R}}$ as the ϵ run through $\Omega_{\vec{R}}$.

We say that a permutation τ of atoms (equivalently of $\Omega_{\vec{R}}$) *is supported by* the translation $\psi_1, \ldots, \psi_q$ if for each $\epsilon \in \Omega_{\vec{R}}$,

$$\alpha_\epsilon(x_1, \ldots, x_r) = \alpha_{\tau(\epsilon)}^{\vec{\psi}}(x_1, \ldots, x_r).$$

---

[5]We do not introduce a precise notion of distance here wishing just to convey the intuition. It could be the Hamming distance between atoms, see below, in which case we should also incorporate the prisoners' opinion about, and the truth of, $D(b,b), D(l,l), B(b,b)$ and $B(l,l)$. However since $B(x,x)$ and $D(x,x)$ coincide, this would make no difference to comparisons of descriptions using H, D and H, B.

As in the unary case, we denote such a permutation τ by $\tau_{\vec{\psi}}$.

Note that this is the situation when if we define the new predicates $Q_i(x_1, \ldots, x_{r_i})$ for $i = 1, 2, \ldots, q$ by

$$Q_i(x_1, \ldots, x_{r_i}) = \psi_i(x_1, \ldots, x_{r_i})$$

then for each ϵ the atom

$$\alpha_\epsilon(x_1, \ldots, x_r) = \bigwedge_{\substack{i \in \{1,2,\ldots,q\} \\ \langle j_1, \ldots, j_{r_i}\rangle \in \{1,2,\ldots,r\}^{r_i}}} R_i^{\epsilon(i, j_1, \ldots, j_{r_i})}(x_{j_1} \ldots, x_{j_{r_i}})$$

is logically equivalent to the atom

$$\alpha_\delta^{\vec{Q}}(x_1, \ldots, x_r) = \bigwedge_{\substack{i \in \{1,2,\ldots,q\} \\ \langle j_1, \ldots, j_{r_i}\rangle \in \{1,2,\ldots,r\}^{r_i}}} Q_i^{\delta(i, j_1, \ldots, j_{r_i})}(x_{j_1} \ldots, x_{j_{r_i}}) \qquad (11)$$

of the language $L_{\vec{Q}}$ with relation symbols $Q_1, Q_2, \ldots, Q_q$ just when $\delta = \tau_{\vec{\psi}}(\epsilon)$.

Unlike the purely unary case however, in the polyadic it is not in general the case that *just any* permutation, or renaming, of atoms is supported by a translation. As we shall prove this will be the case just if the permutation satisfies a certain property (C) from [7] which we will define shortly. Interestingly, as we shall subsequently explain, condition (C) is also equivalent to the permutation of atoms generating an automorphism of a certain structure BL relevant in Pure Inductive Logic.

To formulate (C) we shall need the following notation.

- Let $\Theta(x_1, \ldots, x_n)$ be as in (9) and let $k_1, \ldots, k_t$ be distinct numbers from $\{1, \ldots, n\}$. Then $\Theta[x_{k_1}, \ldots, x_{k_t}]$ denotes the state formula obtained from (9) by restricting it to $x_{k_1}, \ldots, x_{k_t}$, that is, replacing

$$\langle j_1, \ldots, j_{r_i}\rangle \in \{1, 2, \ldots, n\}^{r_i}$$

  by

$$\langle j_1, \ldots, j_{r_i}\rangle \in \{k_1, \ldots, k_t\}^{r_i}.$$

- Let $\Phi(x_{k_1}, \ldots, x_{k_t})$ be a state formula, $m_1, \ldots, m_s$ distinct numbers and

$$f : \{m_1, \ldots, m_s\} \to \{k_1, \ldots, k_t\}$$

  a surjection. Then $(\Phi(x_{k_1}, \ldots, x_{k_t}))_f$ denotes the state formula $\Psi(x_{m_1}, \ldots, x_{m_s})$ for which

$$\Psi(x_{f(m_1)}, \ldots, x_{f(m_s)}) = \Phi(x_{k_1}, \ldots, x_{k_t}).$$

We can now state condition (C) for a permutation σ of $\Omega_{\vec{R}}$:

(C) *For $\epsilon, \delta \in \Omega_{\vec{R}}$, $t \leq r$ and distinct $j_1, \ldots, j_t$ from $\{1, \ldots, r\}$, if $f : \{1, \ldots, r\} \to \{j_1, \ldots, j_t\}$ is a surjection then*

$$\alpha_\epsilon(x_1, \ldots, x_r) = (\alpha_\delta[x_{j_1}, \ldots, x_{j_t}])_f \iff \alpha_{\sigma(\epsilon)}(x_1, \ldots, x_r) = (\alpha_{\sigma(\delta)}[x_{j_1}, \ldots, x_{j_t}])_f.$$

8

We shall also need a consequence of (C) which it is apposite to spell out. Assume that $m_1, \ldots, m_s \in \{1, 2, \ldots, r\}$ are distinct, $k_1, \ldots, k_t \in \{1, 2, \ldots, r\}$ are distinct,

$$f : \{m_1, \ldots, m_s\} \to \{k_1, \ldots, k_t\}$$

is a surjection and $\epsilon, \delta \in \Omega_{\vec{R}}$ are such that .

$$\alpha_\epsilon[x_{m_1}, \ldots, x_{m_s}] = (\alpha_\delta[x_{k_1}, \ldots, x_{k_t}])_f.$$

Let

$$g : \{1, \ldots, r\} \mapsto \{m_1, \ldots, m_s\}$$

be such that for $i \in \{m_1, \ldots, m_s\}$, $g(i) = i$ and for $i \notin \{m_1, \ldots, m_s\}$, $g(i) = m_1$. Hence

$$fg : \{1, \ldots, r\} \to \{k_1, \ldots, k_t\}$$

is a surjection and its restriction to $\{m_1, \ldots, m_s\}$ is f. Let $\gamma \in \Omega_{\vec{R}}$ be such that

$$(\alpha_\epsilon[x_{m_1}, \ldots, x_{m_s}])_g = \alpha_\gamma(x_1, \ldots, x_r) = (\alpha_\delta[x_{k_1}, \ldots, x_{k_t}])_{fg}.$$

Then for any σ satisfying (C) we have

$$(\alpha_{\sigma(\epsilon)}[x_{m_1}, \ldots, x_{m_s}])_g = \alpha_{\sigma(\gamma)}(x_1, \ldots, x_r) = (\alpha_{\sigma(\delta)}[x_{k_1}, \ldots, x_{k_t}])_{fg}.$$

Since g is the identity on $\{m_1, \ldots, m_s\}$, it follows that

$$\alpha_{\sigma(\epsilon)}[x_{m_1}, \ldots, x_{m_s}] = \alpha_{\sigma(\gamma)}[x_{m_1}, \ldots, x_{m_s}] = (\alpha_{\sigma(\delta)}[x_{k_1}, \ldots, x_{k_t}])_f.$$

Since we can argue conversely in the same way, any σ that satisfies (C) also satisfies the following condition:

(D) *For distinct* $m_1, \ldots, m_s \in \{1, 2, \ldots, r\}$, $k_1, \ldots, k_t \in \{1, 2, \ldots, r\}$, *surjection* $f : \{m_1, \ldots, m_s\} \to \{k_1, \ldots, k_t\}$ *and* $\epsilon, \delta \in \Omega_{\vec{R}}$,

$$\alpha_\epsilon[x_{m_1}, \ldots, x_{m_s}] = (\alpha_\delta[x_{k_1}, \ldots, x_{k_t}])_f \iff \alpha_{\sigma(\epsilon)}[x_{m_1}, \ldots, x_{m_s}] = (\alpha_{\sigma(\delta)}[x_{k_1}, \ldots, x_{k_t}])_f$$

In particular, taking f to be identity, we have

(E) *For distinct* $m_1, \ldots, m_s \in \{1, 2, \ldots, r\}$,

$$\alpha_\epsilon[x_{m_1}, \ldots, x_{m_s}] = \alpha_\delta[x_{m_1}, \ldots, x_{m_s}] \iff \alpha_{\sigma(\epsilon)}[x_{m_1}, \ldots, x_{m_s}] = \alpha_{\sigma(\delta)}[x_{m_1}, \ldots, x_{m_s}].$$

Theorem 2. *Given a permutation σ of $\Omega_{\vec{R}}$ there is a translation supporting σ just if σ satisfies (C).*

Proof. First suppose that σ is a permutation of atoms satisfying condition (C). Define

$$\psi_i(x_1, \ldots, x_{r_i}) = \bigvee_{\sigma(\epsilon)(i,1,2,\ldots,r_i)=1} \alpha_\epsilon[x_1, \ldots, x_{r_i}]. \tag{12}$$

We shall show that this provides the required translation supporting σ.

Let $j_1, \ldots, j_{r_i}$ be from $\{1, \ldots, r\}$ (not necessarily distinct). We have

$$\psi_i(x_{j_1}, \ldots, x_{j_{r_i}}) = \bigvee_{\sigma(\epsilon)(i,1,\ldots,r_i)=1} (\alpha_\epsilon[x_1, \ldots, x_{r_i}])(x_{j_1}/x_1, \ldots, x_{j_{r_i}}/x_{r_i}). \tag{13}$$

We shall now prove that this gives

$$\psi_i(x_{j_1}, \ldots, x_{j_{r_i}}) = \bigvee_{\sigma(\epsilon)(i,j_1,\ldots,j_{r_i})=1} \alpha_\epsilon[x_{j_{h_1}}, \ldots, x_{j_{h_t}}] \tag{14}$$

where $j_{h_1}, \ldots, j_{h_t}$ are the distinct numbers from amongst the $j_1, \ldots, j_{r_i}$.

Let $A_1, A_2, \ldots, A_t$ form the partition of $\{1, 2, \ldots, r_i\}$ such that for $x, y \in \{1, 2, \ldots, r_i\}$, x, y are in the same A_k just if $j_x = j_y$. Let $h_k = \min\{A_k\}$ for $k = 1, 2, \ldots, t$. A disjunct

$$(\alpha_\epsilon[x_1, \ldots, x_{r_i}])(x_{j_1}/x_1, \ldots, x_{j_{r_i}}/x_{r_i}) \tag{15}$$

of (13) is consistent just if for $1 \le m \le q$, $1 \le d_1, d_2, \ldots, d_{r_m} \le r_i$, $\epsilon(m, d_1, d_2, \ldots, d_{r_m})$ depends only on which of the classes $A_1, \ldots, A_t$ the d_n are in. Equivalently

$$\epsilon(m, d_1, d_2, \ldots, d_{r_m}) = \epsilon(m, f(d_1), f(d_2), \ldots, f(d_{r_m}))$$

where f maps the members of each A_k to h_k, the least member of that A_k. Another way of expressing this is that

$$\alpha_\epsilon[x_1, \ldots, x_{r_i}] = (\alpha_\epsilon[x_{h_1}, \ldots, x_{h_t}])_f. \tag{16}$$

Notice that when (15) is consistent then

$$(\alpha_\epsilon[x_1, \ldots, x_{r_i}])(x_{j_1}/x_1, \ldots, x_{j_{r_i}}/x_{r_i}) = (\alpha_\epsilon[x_{h_1}, \ldots, x_{h_t}])(x_{j_{h_1}}/x_{h_1}, \ldots, x_{j_{h_t}}/x_{h_t}). \tag{17}$$

Let g be a permutation of $\{1, 2, \ldots, r\}$ which, in particular, maps each h_k to j_{h_k}. Note that this means that for each $c \in \{1, \ldots, r_i\}$ we have $g(f(c)) = j_c$. For each $\epsilon \in \Omega_{\vec{R}}$ define $\epsilon' \in \Omega_{\vec{R}}$ by

$$\epsilon'(m, g(u_1), g(u_2), \ldots, g(u_{r_m})) = \epsilon(m, u_1, u_2, \ldots, u_{r_m}), \tag{18}$$

for $m = 1, 2, \ldots, q$ and $u_1, \ldots, u_{r_m} \in \{1, \ldots, r\}$, equivalently

$$\alpha_{\epsilon'}(x_1, x_2, \ldots, x_r) = \alpha_\epsilon(x_{g(1)}, x_{g(2)}, \ldots, x_{g(r)}),$$

that is,

$$\alpha_\epsilon(x_1, x_2, \ldots, x_r) = (\alpha_{\epsilon'}(x_1, x_2, \ldots, x_r))_g. \tag{19}$$

10

From (17) a consistent disjunct (15) equals

$$(\alpha_\epsilon[x_{h_1},\ldots,x_{h_t}])(x_{j_{h_1}}/x_{h_1},\ldots,,x_{j_{h_t}}/x_{h_t}), \tag{20}$$

and since $j_{h_k} = g(h_k)$, we get by (18) that

$$\alpha_\epsilon[x_{h_1},\ldots,x_{h_t}](x_{j_{h_1}}/x_{h_1},\ldots,,x_{j_{h_t}}/x_{h_t}) = \tag{21}$$

$$= \bigwedge_{\substack{m\in\{1,2,\ldots,q\} \\ \langle u_1,\ldots,u_{r_m}\rangle\in\{h_1,\ldots,h_t\}^{r_m}}} R_m^{\epsilon(m,u_1,\ldots,u_{r_m})}(x_{g(u_1)}\ldots,x_{g(u_{r_m})}) = \alpha_{\epsilon'}[x_{j_{h_1}},\ldots,x_{j_{h_t}}].$$

By (C) and (19) we have

$$\alpha_{\sigma(\epsilon)}(x_1,x_2,\ldots,x_r) = (\alpha_{\sigma(\epsilon')}(x_1,x_2,\ldots,x_r))_g,$$

that is,

$$\sigma(\epsilon')(m,g(u_1),g(u_2),\ldots,g(u_{r_m})) = \sigma(\epsilon)(m,u_1,u_2,\ldots,u_{r_m}).$$

Hence, using the fact that $j_c = g(f(c))$ for each $c \in \{1,\ldots,r_i\}$,

$$\sigma(\epsilon')(i,j_1,\ldots,j_{r_i}) = \sigma(\epsilon')(i,g(f(1)),\ldots,g(f(r_i))) = \sigma(\epsilon)(i,f(1),\ldots,f(r_i)). \tag{22}$$

From (16) and (D) we have

$$\alpha_{\sigma(\epsilon)}[x_1,\ldots,x_{r_i}] = (\alpha_{\sigma(\epsilon)}[x_{h_1},\ldots,x_{h_{r_i}}])_f$$

so

$$\sigma(\epsilon)(i,f(1),\ldots,f(r_i)) = \sigma(\epsilon)(i,1,\ldots,r_i).$$

With (22) it now follows that

$$\sigma(\epsilon')(i,j_1,\ldots,j_{r_i}) = \sigma(\epsilon)(i,1,\ldots,r_i). \tag{23}$$

Hence a consistent disjunct (15), equivalently (20) and (21), with $\sigma(\epsilon)(i,1,2,\ldots,r_i) = 1$ equals

$$\alpha_{\epsilon'}[x_{j_{h_1}},\ldots,x_{j_{h_t}}]$$

with $\sigma(\epsilon')(i,j_1,\ldots,j_{r_i}) = 1$. Consequently (13) logically implies

$$\bigvee_{\sigma(\epsilon')(i,j_1,\ldots,j_{r_i})=1} \alpha_{\epsilon'}[x_{j_{h_1}},\ldots,x_{j_{h_t}}].$$

Clearly then the disjunction in (13) logically implies the disjunction in (14) with ϵ' in place of ϵ and hence also the right hand side of (14) without this replacement.

Conversely, let $\gamma \in \Omega_{\vec{R}}$ be such that $\sigma(\gamma)(i,j_1,\ldots,j_{r_i}) = 1$ (so $\alpha_\gamma[x_{j_{h_1}},\ldots,x_{j_{h_t}}]$ contributes to the right hand side of (14)). With g as above, let $\xi \in \Omega_{\vec{R}}$ be such that

$$\alpha_\xi(x_1,x_2,\ldots,x_r) = (\alpha_\gamma(x_1,x_2,\ldots,x_r))_g,$$

11

that is,
$$\gamma(m, g(u_1), g(u_2), \ldots, g(u_{r_m})) = \xi(m, u_1, u_2, \ldots, u_{r_m}). \tag{24}$$

(Using our above notation, $\gamma = \xi'$.)

Let ϵ be such that
$$\alpha_\epsilon(x_1, \ldots, x_r) = (\alpha_\xi[x_{h_1}, \ldots, x_{h_t}, x_{r_i+1}, \ldots, x_r])_f$$

where f is an extension of the above defined f, mapping the members of each A_k to h_k and $f(h) = h$ for $h > r_i$. So
$$(\alpha_\epsilon[x_1, \ldots, x_{r_i}])(x_{j_1}/x_1, \ldots, x_{j_{r_i}}/x_{r_i})$$

is consistent (which may not have been true of ξ),
$$\alpha_\epsilon[x_{h_1}, \ldots, x_{h_t}] = \alpha_\xi[x_{h_1}, \ldots, x_{h_t}] \tag{25}$$

and ϵ satisfies (16). Let ϵ' be associated with ϵ as above, that is, via
$$\epsilon'(m, g(u_1), g(u_2), \ldots, g(u_{r_m})) = \epsilon(m, u_1, u_2, \ldots, u_{r_m}). \tag{26}$$

Noting that for $h \in \{h_1, \ldots h_t\}$ we have $g(h) = j_h$, from (24), (26),(25),
$$\alpha_{\epsilon'}[x_{j_{h_1}}, \ldots, x_{j_{h_t}}] = \alpha_\gamma[x_{j_{h_1}}, \ldots, x_{j_{h_t}}] \tag{27}$$

Since $\sigma(\gamma)(i, j_1, \ldots, j_{r_i}) = 1$, by property (E) we have also $\sigma(\epsilon')(i, j_1, \ldots, j_{r_i}) = 1$ and furthermore by virtue of (23), $\sigma(\epsilon)(i, 1, \ldots, r_i) = 1$.

Hence from (21), $\alpha_\gamma[x_{j_{h_1}}, \ldots, x_{j_{h_t}}]$ with $\sigma(\gamma)(i, j_1, \ldots, j_{r_i}) = 1$ equals (20) and (15) with $\sigma(\epsilon)(i, 1, \ldots, r_i) = 1$, and thus the disjunction from (14) logically implies the disjunction from (13) and the identity (14) is proved.

Note that by virtue of the condition (E), if for some $\epsilon, \delta \in \Omega_{\vec{R}}$ we have
$$\alpha_\epsilon[x_{j_{h_1}}, \ldots, x_{j_{h_t}}] = \alpha_\eta[x_{j_{h_1}}, \ldots, x_{j_{h_t}}]$$

then
$$\sigma(\epsilon)(i, j_1, \ldots, j_{r_i}) = \sigma(\eta)(i, j_1, \ldots, j_{r_i})$$

and if $\alpha_\epsilon[x_{j_{h_1}}, \ldots, x_{j_{h_t}}], \alpha_\eta[x_{j_{h_1}}, \ldots, x_{j_{h_t}}]$ are not equal then they are disjoint. Since the disjunction of $\alpha_\epsilon[x_{j_{h_1}}, \ldots, x_{j_{h_t}}]$ over all the $\epsilon \in \Omega_{\vec{R}}$ is a tautology, we have
$$\neg\psi_i(x_{j_1}, \ldots, x_{j_{r_i}}) = \bigvee_{\sigma(\epsilon)(i, j_1, \ldots, j_{r_i})=0} \alpha_\epsilon[x_{j_{h_1}}, \ldots, x_{j_{h_t}}].$$

It follows that for $\epsilon \in \Omega_{\vec{R}}$,
$$\sigma(\epsilon)(i, j_1, \ldots, j_{r_i}) = 1 \iff \alpha_\epsilon(x_1, \ldots, x_r) \models \psi_i(x_{j_1}, \ldots, x_{j_{r_i}}). \tag{28}$$

Hence

$$\alpha_\epsilon(x_1,\ldots,x_r) \models \bigwedge_{\substack{i\in\{1,2,\ldots,q\} \\ \langle j_1,\ldots,j_{r_i}\rangle\in\{1,2,\ldots,r\}^{r_i}}} \psi_i^{\sigma(\epsilon)(i,j_1,\ldots,j_{r_i})}(x_{j_1}\ldots,x_{j_{r_i}}) \;=\; \alpha_{\sigma(\epsilon)}^{\vec{\psi}}(x_1,\ldots,x_r).$$

and since there are as many formulae $\alpha_{\sigma(\epsilon)}^{\vec{\psi}}(x_1,\ldots,x_r)$ as there are atoms and they are exclusive, we must have

$$\alpha_\epsilon(x_1,\ldots,x_r) = \alpha_{\sigma(\epsilon)}^{\vec{\psi}}(x_1,\ldots,x_r),$$

which concludes this direction of the proof.

For the converse suppose that the translation $\vec{\psi}$ supports a permutation σ of atoms. We need to show that σ satisfies (C).

Since $\vec{\psi}$ supports σ we have that for each $\epsilon\in\Omega_{\vec{R}}$,

$$\begin{aligned}
\alpha_\epsilon(x_1,\ldots,x_r) &= \bigwedge_{\substack{i\in\{1,2,\ldots,q\} \\ \langle j_1,\ldots,j_{r_i}\rangle\in\{1,2,\ldots,r\}^{r_i}}} R_i^{\epsilon(i,j_1,\ldots,j_{r_i})}(x_{j_1}\ldots,x_{j_{r_i}}) \qquad\qquad (29) \\
&= \alpha_{\sigma(\epsilon)}^{\vec{\psi}}(x_1,\ldots,x_r) = \bigwedge_{\substack{i\in\{1,2,\ldots,q\} \\ \langle j_1,\ldots,j_{r_i}\rangle\in\{1,2,\ldots,r\}^{r_i}}} \psi_i^{\sigma(\epsilon)(i,j_1,\ldots,j_{r_i})}(x_{j_1}\ldots,x_{j_{r_i}}).
\end{aligned}$$

Note that this means that when $\gamma\in\Omega_{\vec{R}}$ and $t\leq r$, $k_1,\ldots,k_t$ from $\{1,\ldots,r\}$ are distinct and $f:\{1,\ldots,r\}\to\{k_1,\ldots,k_t\}$ is a surjection, then

$$\begin{aligned}
(\alpha_\gamma[x_{k_1},\ldots,x_{k_t}])_f &= \bigwedge_{\substack{i\in\{1,2,\ldots,q\} \\ \langle j_1,\ldots,j_{r_i}\rangle\in\{1,2,\ldots,r\}^{r_i}}} R_i^{\gamma(i,f(j_1)\ldots,f(j_{r_i}))}(x_{j_1},\ldots,x_{j_{r_i}}). \\
&= (\alpha_{\sigma(\gamma)}^{\vec{\psi}}[x_{k_1},\ldots,x_{k_t}])_f = \bigwedge_{\substack{i\in\{1,2,\ldots,q\} \\ \langle j_1,\ldots,j_{r_i}\rangle\in\{1,2,\ldots,r\}^{r_i}}} \psi_i^{\sigma(\gamma)(i,f(j_1),\ldots,f(j_{r_i}))}(x_{j_1}\ldots,x_{j_{r_i}}).
\end{aligned}$$

Hence $\alpha_\epsilon(x_1,\ldots,x_r) = (\alpha_\gamma[x_{k_1},\ldots,x_{k_t}])_f$ just when for all i and $j_1,\ldots,j_{r_i}\in\{1,\ldots,r\}$ we have

$$\epsilon(i,j_1,\ldots,j_{r_i}) = \gamma(i,f(j_1)\ldots,f(j_{r_i}))$$

and just when for all i and $j_1,\ldots,j_{r_i}\in\{1,\ldots,r\}$ we have

$$\sigma(\epsilon)(i,j_1,\ldots,j_{r_i}) = \sigma(\gamma)(i,f(j_1)\ldots,f(j_{r_i}))$$

which is equivalent to $\alpha_{\sigma(\epsilon)}(x_1,\ldots,x_r) = (\alpha_{\sigma(\gamma)}[x_{k_1},\ldots,x_{k_t}])_f$, as required.

$\square$

13

The Example continued. For a language L containing one unary predicate R_1 and one binary predicate R_2, let $\sigma : \Omega_{\vec{R}} \to \Omega_{\vec{R}}$ be defined by

$$\sigma(\epsilon) = \begin{cases} \delta & \text{if } \epsilon(2,1,2) \neq \epsilon(2,2,1) \text{ and } \epsilon(1,1) = \epsilon(1,2) = 1, \\ \epsilon & \text{otherwise,} \end{cases}$$

where δ is as ϵ except that $\delta(2,1,2) = \epsilon(2,2,1)$ and $\delta(2,2,1) = \epsilon(2,1,2)$. Considering the 3 non-identity possibilities for f in condition (C) it is easy to see that the condition is satisfied and hence σ is supported by a translation. From (12) we can see that the translation is $\psi_1(x) = R_1(x)$ and

$$\psi_2(x,y) = (\neg(R_1(x) \wedge R_1(y)) \wedge R_2(x,y)) \vee (R_1(x) \wedge R_1(y) \wedge R_2(y,x)).$$

This shows that Smith and Jones could argue in terms of H, B just as well as in terms of H, D.

As in the unary case, a translation $\vec{\psi}$ as above that maps the atom $\alpha_\epsilon(x_1, \ldots, x_r)$ to the atom

$$\alpha^{\vec{Q}}_{\tau_{\vec{\psi}}(\epsilon)}(x_1, \ldots, x_r)$$

(which is logically equivalent to $\alpha_\epsilon(x_1, \ldots, x_r)$ when the Q_j are defined by $Q_j(x_1, \ldots, x_{r_j}) = \psi_j(x_1, \ldots, x_{r_j})$) extends to all formulae of $L_{\vec{R}}^-$, and hence also to all sentences of $L_{\vec{R}}^-$. To see that, first note that any state formula $\Theta(x_1, x_2, \ldots, x_n)$ is logically equivalent to a conjunction of atoms, see [7] or [5, Chapter 41]:

$$\Theta(x_1, \ldots, x_n) = \bigwedge_{\langle k_1, \ldots, k_r \rangle \in \{1, \ldots, n\}^r} \alpha_{\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}}(x_{k_1}, \ldots, x_{k_r}), \tag{30}$$

where the $\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}$ are elements of $\Omega_{\vec{R}}$.[6] Hence with the Q_j defined as the ψ_j, $\Theta(x_1, \ldots, x_n)$ is logically equivalent to

$$\bigwedge_{\langle k_1, \ldots, k_r \rangle \in \{1, \ldots, n\}^r} \alpha^{\vec{Q}}_{\tau_{\vec{\psi}}(\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle})}(x_{k_1}, \ldots, x_{k_r}) \tag{31}$$

(which we may denote $(\tau_{\vec{\psi}}\Theta)^{\vec{Q}}(x_1, \ldots, x_n)$).

The rest follows by the Prenex and Disjunctive Normal Form Theorems because any formula $\xi(x_1, \ldots, x_n)$ of $L_{\vec{R}}^-$ is logically equivalent to a formula of $L_{\vec{R}}^-$ of the form

$$T_1 x_{n+1} T_2 x_{n+2} \ldots T_m x_{n+m} \bigvee_{j=1}^{t} \Theta_j(x_1, \ldots, x_{n+m}) \tag{32}$$

---

[6]Note that this means that the notation, if the need ever arose to write this out, would require us to talk about $\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}(i, j_1, \ldots, j_{r_i})$ which are values from $\{0, 1\}$ such that $\Theta(x_1, \ldots, x_n) \models R_i(x_{k_{j_1}}, \ldots, x_{k_{j_{r_i}}})$ just when $\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}(i, j_1, \ldots, j_{r_i}) = 1$; the $\langle k_1, \ldots, k_r \rangle$ are from $\{1, \ldots, n\}^r$, i from $\{1, \ldots, q\}$ and $\langle j_1, \ldots, j_{r_i} \rangle$ from $\{1, \ldots, r\}^{r_i}$. Note also in particular that tuples $\langle k_1, \ldots, k_r \rangle$ and $\langle j_1, \ldots, j_{r_i} \rangle$ *with repeats* are included.

where each T_i is one of $\forall$ or $\exists$, and hence also to

$$T_1 x_{n+1} T_2 x_{n+2} \ldots T_m x_{n+m} \bigvee_{j=1}^{t} (\tau_{\vec{\psi}} \Theta_j)^{\vec{Q}}(x_1, \ldots, x_{n+m}). \tag{33}$$

Verisimilitude?

The original motivation for the research in this note came from Miller's paper [2] regarding a measure of closeness of a theory to the truth, where, as explained in e.g. [4], the truth is identified with a constituent of a finite language (a complete consistent theory) and the theory with a set of constituents (possibly just one as in the Prisoner Example). First propositional languages were considered, then predicate, in particular unary languages with no constants. With predicate languages however, it appears natural to employ languages with constants. This means that where there were only quantified sentences to consider the truth of, we now also have a large class of quantifier free sentences. If we shy away from wishing to contemplate infinite domains, and propose to have a name for each individual in the finite domain we consider, we are led to the notion of the truth being a state description and a theory being a set (disjunction) of such. The propositional case is like that of a unary predicate language for a domain with just one element; state descriptions for one constant (corresponding to constituents in the propositional case) are just atoms (instantiated by that constant). In Section 1 we have seen that for unary languages and any permutation of atoms (that is of $\Omega_{\vec{P}}$) there is a translation supporting it which disqualifies Hamming distance between atoms as a measure of verisimilitude.

The analogous situation for a polyadic predicate language is when the domain has r elements (with r being the highest arity of the language). State descriptions for r constants are atoms (instantiated by these constants) and a permutation of atoms (that is of $\Omega_{\vec{R}}$) is supported by a translation just when it satisfies the condition (C). Hence the same conclusion has to be reached regarding the suitability of the Hamming distance between atoms for measuring verisimilitude.

In the general case state descriptions are state formulae instantiated by an appropriate number of distinct constants and provided we fix the constants we can concentrate on state formulae. These can be expressed as conjunctions of atoms. As we have seen in Sections 1 and 2, permutations of atoms supported by a translation do extend to state formulae. This is not surprising in the unary case where atoms for distinct variables are disjoint but it is somewhat remarkable in the polyadic case where atoms for different r-tuples of variables may well be incompatible if the r-tuples have some variables in common.

Whilst atoms can be 'renamed' by translations in various ways (subject to (C)), and the Hamming distance between them (i.e. the Hamming distance between elements of $\Omega_{\vec{R}}$) can change, translations do preserve Hamming distances between state formulae

$$\Theta(x_1, \ldots, x_n) = \bigwedge_{\langle k_1, \ldots, k_r \rangle \in \{1, \ldots, n\}^r} \alpha_{\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}}(x_{k_1}, \ldots, x_{k_r}),$$

$$\Phi(x_1, \ldots, x_n) = \bigwedge_{\langle k_1, \ldots, k_r \rangle \in \{1, \ldots, n\}^r} \alpha_{\epsilon_{\langle \Phi, \langle k_1, \ldots, k_r \rangle \rangle}}(x_{k_1}, \ldots, x_{k_r}),$$

15

as given by

$$\sum_{\langle k_1, \ldots, k_r \rangle \in \{1, \ldots, n\}^r} \#\left(\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}, \epsilon_{\langle \Phi, \langle k_1, \ldots, k_r \rangle \rangle}\right)$$

where

$$\#\left(\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle}, \epsilon_{\langle \Phi, \langle k_1, \ldots, k_r \rangle \rangle}\right) = \begin{cases} 0 & \text{if } \epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle} = \epsilon_{\langle \Phi, \langle k_1, \ldots, k_r \rangle \rangle}, \\ 1 & \text{otherwise.} \end{cases}$$

Translations also preserve the *structure* of state descriptions/formulae in the sense which we now explain. In [5, Chapter 40] the following notion of *similarity* is introduced:

State formulae $\Theta(x_1, \ldots, x_n)$, $\Phi(x_1, \ldots, x_n)$ *are similar, if for distinct* $m_1, \ldots, m_s \in \{1, 2, \ldots, n\}$, $k_1, \ldots, k_t \in \{1, 2, \ldots, n\}$ *and surjection* $f : \{m_1, \ldots, m_s\} \to \{k_1, \ldots, k_t\}$,

$$\Theta[x_{m_1}, \ldots, x_{m_s}] = (\Theta[x_{k_1}, \ldots, x_{k_t}])_f \iff \Phi[x_{m_1}, \ldots, x_{m_s}] = \Phi[x_{k_1}, \ldots, x_{k_t}])_f.$$

Informally, $\Theta(x_1, \ldots, x_n)$, $\Phi(x_1, \ldots, x_n)$ are similar if, whenever in Θ the behaviour[7] of some variables $x_{m_1}, \ldots, x_{m_s}$ exactly corresponds to the behaviour of some further variables $x_{k_1}, \ldots, x_{k_t}$ (with f specifying how the $x_{k_1}, \ldots, x_{k_t}$ are 'cloned' by the $x_{m_1}, \ldots, x_{m_s}$), the same happens in Φ, and conversely.

It follows from results in [7], [5, Chapter 41] and the above that state formulae $\Theta(x_1, \ldots, x_n)$ and $\Phi(x_1, \ldots, x_n)$ with Θ as in (30) are similar just when there is a a permutation of atoms σ that is supported by a translation such that $\Phi(x_1, \ldots, x_n)$ is equal to

$$\sigma(\Theta(x_1, \ldots, x_n)) = \bigwedge_{\langle k_1, \ldots, k_r \rangle \in \{1, \ldots, n\}^r} \alpha_{\sigma(\epsilon_{\langle \Theta, \langle k_1, \ldots, k_r \rangle \rangle})}(x_{k_1}, \ldots, x_{k_r}), \qquad (34)$$

that is, just when Φ is a 'translated' version of Θ. In other words there are quantifier free formulae $\psi_i(x_1, \ldots, x_{r_i})$ such that if we replace each R_i in Φ by ψ_i we obtain Θ.

Hence any translated version of the truth carries some information about the truth, namely the shared structure. From this viewpoint then the minimum Hamming distance between translations of state descriptions/formulae might be used to measure how far a theory is from capturing the structure of the truth.

Renaming Invariance and Rationality

Consider the problem Pure Inductive Logic aims to address: How to give a rational assignment of probabilities $w(\theta)$ to sentences θ of an entirely uninterpreted language $L_{\vec{R}}$?[8] The current modus operandi here is to propose principles which we may intuitively feel

---

[7]By the *behaviour of the variables* $x_{m_1}, \ldots, x_{m_s}$ *in* Θ we mean all the information contained in Θ that involves just these variables and no others.

[8]See for example [5, Chapter 1] for further details.

are somehow 'rational' for the probability function w to satisfy and investigate their consequences and inter relationships.[9]

The previous sections suggest a principle of 'translation invariance', that w should be unaffected by the sort of translation we have considered above. For suppose that within the remit of Pure Inductive Logic we have chosen a 'rational' probability function w on the set of sentences of the language $L_{\vec{R}}$ (denoted $SL_{\vec{R}}$), that is chosen a probability function satisfying those principles which we judge to demarcate what 'rational' means. A caviller now points out that since there is supposed to be no intended interpretation here we could equally well have based our choice on a translation of the original relations - on an equally expressive set of relation symbols of appropriate arities. So, the caviller continues, to be consistent we should still be giving the same probabilities even after making this translation (and regardless of the the names chosen for the relation symbols). In other words, at risk of otherwise being seemingly not consistent we should accept the principle that a rational assignment of probabilities should additionally be invariant under translations.

Translation Invariance Principle, TIP

If σ is a permutation of atoms supported by a translation and w_σ is the probability function on $SL_{\vec{R}}$ determined by[10]

$$w_\sigma(\Theta(b_1, \ldots, b_n)) = w(\sigma(\Theta(x_1, \ldots, x_n))(b_1/x_1, \ldots, b_n/x_n)), \tag{35}$$

then $w = w_\sigma$.

It turns out that such a principle of 'translation invariance' already exists in equivalent forms in the literature. We shall now discuss this in more detail.

The Unary Case

We start with the case of the purely unary language $L_{\vec{P}}$, adopting again the notation of that first section.[11] Let w be a probability function on the set of sentences $SL_{\vec{P}}$ of $L_{\vec{P}}$. Then w is uniquely determined (see for example [5, Chapter 7] or [8]) by its values on the state descriptions of $L_{\vec{P}}$, that is sentences of $L_{\vec{P}}$ of the form

$$\bigwedge_{i=1}^{m} \alpha_{\epsilon_i}(b_i)$$

where the $b_1, b_2, \ldots, b_m$ are distinct constants from $\{a_1, a_2, a_3, \ldots\}$.

As in Section 1 but using $\vec{P}$ also in place of $\vec{Q}$, any permutation τ of $\Omega_{\vec{P}}$ determines a permutation/renaming/translation of atoms and in turn of *state formulae,* that is formulae of $L_{\vec{P}}^-$ of the form $\bigwedge_{i=1}^{m} \alpha_{\epsilon_i}(x_i)$, by sending $\bigwedge_{i=1}^{m} \alpha_{\epsilon_i}(x_i)$ to $\bigwedge_{i=1}^{m} \alpha_{\tau(\epsilon_i)}(x_i)$.

---

[9]For the definition of a probability function in this context and general background see for example [5] or [8].

[10]Using the condition (E) and [5, Page 42] or [8], we can see that this definition does yield a probability function.

[11]So in particular $\Omega_{\vec{P}}$ is the set of maps from $\{1, 2, \ldots, q\}$ to $\{0, 1\}$.

For a probability function w, each such translation uniquely determines a further probability function w_τ on $SL_{\vec{P}}$ by setting

$$w_\tau\big(\bigwedge_{i=1}^{m} \alpha_{\epsilon_i}(b_i)\big) = w\big(\bigwedge_{i=1}^{m} \alpha_{\tau(\epsilon_i)}(b_i)\big). \tag{36}$$

The requirement that a probability function w is invariant under translation in the sense we have discussed above can be seen to be equivalent to satisfying the well know property of *Atom Exchangeability*:[12]

The Principle of Atom Exchangeability, Ax

For τ a permutation of atoms and a state description $\bigwedge_{i=1}^{n} \alpha_{\epsilon_i}(b_i)$,

$$w\big(\bigwedge_{i=1}^{n} \alpha_{\epsilon_i}(b_i)\big) = w\big(\bigwedge_{i=1}^{n} \alpha_{\tau(\epsilon_i)}(b_i)\big).$$

Atom Exchangeability follows from Johnson's Sufficientness Postulate (see for example [5, Lemma 17.1] and hence holds for the members of Carnap's Continuum of Inductive methods. It implies (but is not implied by) two other principles which are central in the field:

The Principle of Predicate Exchangeability, Px

For $\theta \in SL$ and predicate symbols P_i, P_j of L, if θ' is the result of transposing P_i, P_j throughout θ then $w(\theta) = w(\theta')$.

The Strong Negation Principle, SN

For $\theta \in SL$, $w(\theta) = w(\theta')$ where θ' is the result of replacing each occurrence of the predicate symbol P_i in θ by $\neg P_i$.

Each of the principles Px and SN have an evident claim to rationality on the grounds of *symmetry*. Namely in the completely uninterpreted situation envisaged here it would be irrational to assign probability values which broke the existing symmetries in the language. However, the reasons for the rationality of Atom Exchangeability are less easy to appreciate, as indeed the surprise inherent in Miller's example shows.

For this reason one might require the constraints imposed on one's choice of probability function to at least include Px+SN. That being the case one might argue not that we should always have $w = w_\tau$ for τ a permutation of $\{0,1\}^q$ but simply that w_τ was, as far as these constraints were concerned, an equally good choice, i.e that the w_τ should also satisfy at least Px+SN.

Thus we would be advocating here a sort of *meta-principle*, namely that if one initially proposed that adherence to principles X, Y, Z etc. determined what constituted a choice

---

[12]For more details on Atom Exchangeability see [5, Chapter 14].

of probability function w being rational, then on secondary consideration w should additionally be such that all the w_τ also satisfy X, Y, Z etc..[13]

In the particular case of Px+SN this meta-principle yields a new principle which, for $q > 2$, lies strictly between Px+SN and Ax:

The Unary Principle of Inculcated Px+SN, I(Px+SN)

For every permutation τ of $\Omega_{\vec{P}}$, w_τ satisfies Px+SN.

Theorem 3. *Let w be a probability function on $SL_{\vec{P}}$. Then the probability function w_τ satisfies Px+SN for each permutation τ of $\Omega_{\vec{P}}$ just if $q > 2$ and $w = w_\sigma$ for every even permutation σ of $\Omega_{\vec{P}}$ or $q \leq 2$ and w satisfies Ax.*

Proof. We first show this in the forward direction. Let $\mathcal{S}_{\vec{P}}$ be the group of permutations of $\Omega_{\vec{P}}$. Let H be the subgroup of $\mathcal{S}_{\vec{P}}$ of permutations τ such that $v_\tau = v$ for all probability functions v on $SL_{\vec{P}}$ which satisfy Px+SN. Equivalently, H is generated by the permutations of atoms which just transposes P_i, P_j and those which just transpose $P_i, \neg P_i$. Notice that for $q > 2$ all such permutations are even whilst for $q \leq 2$ H also contains odd permutations. Note too that a probability function v satisfies Px+SN just when $v_\tau = v$ for all $\tau \in H$.

Let K_w be the set of $\tau \in \mathcal{S}_{\vec{P}}$ such that $w = w_\tau$. Then K_w is a subgroup of $\mathcal{S}_{\vec{P}}$. For $\rho \in K_w$, and $\sigma \in \mathcal{S}_{\vec{P}}$,

$$w_\sigma\Big(\bigwedge_{i=1}^{m} \alpha_{\epsilon_i}(b_i)\Big) = w\Big(\bigwedge_{i=1}^{m} \alpha_{\sigma(\epsilon_i)}(b_i)\Big) = w_\rho\Big(\bigwedge_{i=1}^{m} \alpha_{\sigma(\epsilon_i)}(b_i)\Big)$$

$$= w\Big(\bigwedge_{i=1}^{m} \alpha_{\rho\sigma(\epsilon_i)}(b_i)\Big) = w_\sigma\Big(\bigwedge_{i=1}^{m} \alpha_{\sigma^{-1}\rho\sigma(\epsilon_i)}(b_i)\Big)$$

so $\sigma^{-1} K_w \sigma \subseteq K_{w_\sigma}$ and hence, since $|\sigma^{-1} K_w \sigma| = |K_{w_\sigma}|$, $K_{w_\sigma} = \sigma^{-1} K_w \sigma$.

Also since the w_σ all satisfy Px+SN, $H \subseteq K_{w_\sigma}$ so

$$\bigcap_{\sigma \in \mathcal{S}_{\vec{P}}} K_{w_\sigma} = \bigcap_{\sigma \in \mathcal{S}_{\vec{P}}} \sigma^{-1} K_w \sigma$$

is a subgroup of $\mathcal{S}_{\vec{P}}$ containing H (and so is non-trivial). In fact it is a normal subgroup since if

$$\gamma \in \bigcap_{\sigma \in \mathcal{S}_{\vec{P}}} \sigma^{-1} K_w \sigma$$

then

$$\tau\gamma\tau^{-1} \in \bigcap_{\sigma \in \mathcal{S}_{\vec{P}}} \tau\sigma^{-1} K_w \sigma\tau^{-1} = \bigcap_{\delta \in \mathcal{S}_{\vec{P}}} \delta^{-1} K_w \delta$$

---

[13]Another such 'meta-principle' in this area is *(Unary) Language Invariance*, see for example [5], which again applies not simply to a single probability function w but to a family of related probability functions to which w belongs.

since $\delta = \tau\sigma^{-1}$ will run through all elements of $\mathcal{S}_{\vec{P}}$ as σ does.

Now suppose that $q > 2$. Then since the only non-trivial normal subgroups of $\mathcal{S}_{\vec{P}}$ are itself and the alternating group A of even permutations it must be the case that $\bigcap_{\sigma \in \mathcal{S}_{\vec{P}}} K_{w_\sigma}$ is one of these and hence, by taking σ to be the identity permutation, it follows that $A \subseteq K_w$. Hence $w = w_\sigma$ for every $\sigma \in A$. (As we shall see later in this case of $q > 2$, H only contains even permutations and we cannot obtain a stronger result here.)

We now turn to the case $q \leq 2$. When $q = 1$, $\mathcal{S}_{\vec{P}}$ is the only non-trivial normal subgroup of $\mathcal{S}_{\vec{P}}$ so in this case we must have $K_w = \mathcal{S}_{\vec{P}}$, in other words w satisfies Ax. When $q = 2$, $\mathcal{S}_{\vec{P}}$ has two non-trivial proper normal subgroups. However both of these only contain even permutations while H contains some odd permutations so again we must have $K_w = \mathcal{S}_{\vec{P}}$ and Ax follows.

Turning to the converse direction this is clear in the case of $q \leq 2$. For $q > 2$, suppose that $w = w_\sigma$ for all $\sigma \in A$. Then certainly since $H \subseteq A$ for $q > 2$, w is invariant under permutations of predicate symbols and permutations replacing P_i by $\neg P_i$, so w satisfies Px+SN, and so does w_σ for σ an even permutation. Also since $w_\sigma = w_\tau$ for all even permutations σ, τ, it is easy to see that this must also hold for all odd permutations σ, τ and hence by the same argument as in the even case, for σ an odd permutation w_σ must satisfy Px+SN too.

$\square$

At this point one might question whether even for $q > 2$ the new principle I(Px+SN), amounting to $A \subseteq K_w$, really is strictly between Px+SN and Ax. (As Theorem 3 shows it is equivalent to Ax for $q \leq 2$.) We now construct examples of probability functions which show that this is the case.

Let $\epsilon_1, \epsilon_2, \ldots, \epsilon_{2^q}$ list the elements of $\Omega_{\vec{P}}$. For $\sigma \in \mathcal{S}_{\vec{P}}$ it will be convenient to also treat σ as a permutation of these subscripts, that is $\sigma(\epsilon_i) = \epsilon_{\sigma(i)}$.

For $\vec{c} = \langle c_1, c_2, \ldots, c_{2^q} \rangle \in \mathbb{R}^{2^q}$ such that the c_i are non-negative with sum 1 define the probability function[14] $w_{\vec{c}}$ by[15]

$$w_{\vec{c}}\left(\bigwedge_{i=1}^{m} \alpha_{\epsilon_{h_i}}(b_i) \right) = \prod_{i=1}^{m} c_{h_i}.$$

Now set

$$v = |A|^{-1} \sum_{\sigma \in A} w_{\sigma\vec{c}}$$

where $\sigma\vec{c} = \langle c_{\sigma(1)}, c_{\sigma(2)}, \ldots, c_{\sigma(2^q)} \rangle$. Then $K_v \supseteq A$ and for distinct constants $b_{i,k}$, $i = 1, 2, \ldots, 2^q$, $k = 1, 2, \ldots, i-1$,

$$v\left(\bigwedge_{i=1}^{2^q} \bigwedge_{k=1}^{i-1} \alpha_i(b_{i,k}) \right)$$

---

[14]See [5, Page 51] for more details.

[15]The $w_{\vec{c}}$, and in turn the probability functions v, u to be introduced shortly, immediately satisfy the ubiquitous *Constant Exchangeability Principle,* Ex, see for example [5, Page 33].

is the sum of the positive terms from the determinant of the Vandermonde $2^q \times 2^q$ matrix with i, j'th entry c_i^{j-1}. Since we can find suitable $\vec{c}$ yielding positive Vandermonde determinant this means that for any odd permutation τ

$$v\left(\bigwedge_{i=1}^{2^q} \bigwedge_{k-1}^{i-1} \alpha_i(b_{i,k})\right) > v\left(\bigwedge_{i=1}^{2^q} \bigwedge_{k-1}^{i-1} \alpha_{\tau(i)}(b_{i,k})\right)$$

so in this case $K_v \subseteq A$, forcing $K_v = A$.

To distinguish the 'new principle' from Px+SN in the case $q > 2$ suppose without loss of generality that

$$\alpha_{\epsilon_1}(x) \quad = \quad P_1(x) \wedge P_2(x) \wedge \bigwedge_{j=3}^{q} P_j(x)$$

$$\alpha_{\epsilon_2}(x) \quad = \quad P_1(x) \wedge \neg P_2(x) \wedge \bigwedge_{j=3}^{q} P_j(x)$$

$$\alpha_{\epsilon_3}(x) \quad = \quad \neg P_1(x) \wedge P_2(x) \wedge \bigwedge_{j=3}^{q} P_j(x)$$

and let ν be the even permutation $\epsilon_1 \mapsto \epsilon_2 \mapsto \epsilon_3 \mapsto \epsilon_1$ and leaving all remaining elements of $\Omega_{\vec{P}}$ fixed. Then $\nu \notin H$ since it does not preserve Hamming Distance between atoms whilst every permutation in H does (and conversely in fact, see [1]). Then since $\nu \notin H$ the polynomials

$$\sum_{\sigma \in H} \prod_{i=1}^{2^q} x_{\sigma(i)}^i, \quad \sum_{\sigma \in H} \prod_{i=1}^{2^q} x_{\nu\sigma(i)}^i,$$

are not formally the same and hence there must be a vector in the positive quadrant, indeed since these polynomials are homogeneous, a normalized positive vector $\langle c_1, c_2, \ldots, c_{2^q} \rangle$, on which they give different values. But that means that if

$$u = |H|^{-1} \sum_{\sigma \in H} w_{\sigma\vec{c}}$$

then u satisfies Px+SN whilst $u \neq u_\nu$ since they give different values on

$$\bigwedge_{i=1}^{2^q} \bigwedge_{k=1}^{i} \alpha_i(b_{i,k}).$$

A somewhat surprising observation on the proof of Theorem 3 is that we could have dropped SN throughout. Similarly as far as the result for $q > 2$ is concerned we could have dropped Px (while retaining SN). Notice too that all the above results (and counter-examples) hold equally well if we add to Px+SN also the ubiquitous *Principle of Constant Exchangeability*, Ex, that is that for any sentence $\theta(a_1, a_2, \ldots, a_n)$ and distinct $j_1, j_2, \ldots, j_n$,

$$w(\theta(a_1, a_2, \ldots, a_n)) = w(\theta(a_{j_1}, a_{j_2}, \ldots, a_{j_n})).$$

The Polyadic Case

We now turn to the case of the polyadic language $L_{\vec{R}}$ and w a probability function on $SL_{\vec{R}}$.

In this case too, equivalent forms of the Invariance Under Translation Principle already exist in the literature, see [5, Chapters 39, 40], [6], [7]. Most directly, the principle is equivalent to the the Permutation Invariance Principle, PIP. This is a special case of the 'ultimate' symmetry principle of Pure Inductive Logic, INV, which we will briefly explain to start with.

In [5, Chapters 23, 39] we have argued that assigning probabilities to (classes of logically equivalent) sentences of an entirely uninterpreted language $L_{\vec{R}}$ could be imagined as a task to be performed by an agent who knows that s/he is in a structure M for $L_{\vec{R}}$ with universe $\{a_1, a_2, \ldots\}$, in which each constant symbol a_i is interpreted as a_i, but having no information as to what sentences of $SL_{\vec{R}}$ hold in their ambient structure M. Since a rational agent in such a situation would presumably wish to respect symmetry, this picture clearly helps to motivate the symmetry principles which we have mentioned already. Our attempt to capture that which underlies *all* symmetry principles in Pure Inductive Logic, see the foregoing, was based on the observation that any symmetry of the (classes of logically equivalent) sentences of $SL_{\vec{R}}$ corresponds to an automorphism of the set of all possible structures as above, along with the set of its definable subsets, in the following sense:

Let $\mathcal{T}L_{\vec{R}}$ be the set of structures M for $L_{\vec{R}}$ with universe $\{a_1, a_2, a_3, \ldots\}$ where each constant symbol a_i of the language is interpreted in M by the element a_i. Let $BL_{\vec{R}}$ be the two-sorted structure with universe $\mathcal{T}L_{\vec{R}}$ together with the sets

$$[\theta] = \{\, M \in \mathcal{T}L_{\vec{R}} \,|\, M \models \theta \,\} \quad \text{for } \theta \in SL_{\vec{R}}$$

and the membership relation between elements of $\mathcal{T}L_{\vec{R}}$ and these sets.

An automorphism η of $BL_{\vec{R}}$ is a bijection of $\mathcal{T}L_{\vec{R}}$ such that for each $\theta \in SL_{\vec{R}}$ there is some $\psi \in SL_{\vec{R}}$ such that

$$\eta[\theta] = \{\, \eta(M) \,|\, M \in \mathcal{T}L_{\vec{R}}, \, M \models \theta \,\} = [\psi] \tag{37}$$

and conversely, for every $\psi \in SL_{\vec{R}}$ there is a sentence $\theta \in SL_{\vec{R}}$ satisfying (37). We write $\eta(\theta)$ for the sentence $\psi \in SL_{\vec{R}}$ for which $\eta[\theta] = [\psi]$ (up to logical equivalence).

The Invariance Principle, INV

If η is an automorphism of $BL_{\vec{R}}$ then $w(\theta) = w(\eta(\theta))$ for $\theta \in SL_{\vec{R}}$.

As discussed in [5], INV in its full generality may be too strong, possibly denying rationality to almost all probability functions. This has indeed been proved for languages $L_{\vec{P}}$ with only unary predicates: there is only one (from various points of view a not-entirely-suitable) probability function on $SL_{\vec{P}}$ satisfying INV, see [5, Chapter 23]. The situation in the polyadic remains intriguingly open.

Whilst INV in the purely unary context, after corroborating the intuition for previously known symmetry principles, has been shown to just go too far, in the polyadic context INV

has yielded a further interesting symmetry principle which obtains from INV by imposing an additional requirement on the η, namely that they map state descriptions to state descriptions. In [5, Chapter 39] this has been proved to be equivalent to the principle PIP which we will state precisely after introducing some further definitions from [5].

We say that a function F *permutes state formulae* if for each n and (distinct) variables $x_{j_1}, \ldots, x_{j_n}$, F permutes the state formulae $\Phi(x_{j_1}, \ldots, x_{j_n})$ in these variables (up to logical equivalence). Properties (A) and (B) are defined as follows:

(A) *For each state formula* $\Theta(x_{k_1}, \ldots, x_{k_t})$ *and surjective mapping* $\tau : \{m_1, \ldots, m_s\} \to \{k_1, \ldots, k_t\}$,

$$(F(\Theta(x_{k_1}, \ldots, x_{k_t}))_\tau = F(\Theta(x_{k_1}, \ldots, x_{k_t})_\tau).$$

(B) *For each state formula* $\Phi(x_{j_1}, \ldots, x_{j_n})$ *and (distinct)* $i_1, i_2, \ldots, i_k \in \{j_1, \ldots, j_n\}$

$$F(\Phi)[x_{i_1}, \ldots, x_{i_k}] = F(\Phi[x_{i_1}, \ldots, x_{i_k}]).$$

The Permutation Invariance Principle, PIP

If F *is a permutation of state formulae of* $L_{\vec{R}}$ *satisfying (A) and (B) then for a state description* $\Phi(b_1, \ldots, b_n)$,

$$w(\Phi(b_1, \ldots, b_n)) = w(F(\Phi(x_1, \ldots, x_n))(b_1/x_1, \ldots, b_n/x_n)).$$

The equivalence of PIP and TIP follows by Lemma 4 from [7] which shows that a permutation σ of atoms extends to a permutation of state formulae satisfying (A) and (B), by mapping $\Theta(x_1, \ldots, x_n)$ to $\sigma(\Theta(x_1, \ldots, x_n))$ as in (34) (and analogously for any other n-tuple of distinct variables), just when σ satisfies the condition (C). We remark that PIP, and hence TIP, is also equivalent to the principle which states that similar state descriptions get the same probability (Nathanial's Invariance Principle, NIP), cf. [5, Chapter 41].

In the purely unary context, PIP is equivalent to Ax. As Ax does in the unary case, in the polyadic PIP implies Px and SN (the general formulation of these principles are as in the unary case except that in Px we need to say that we exchange predicate symbols *of the same arities*). How PIP relates to the Inculcated Px+SN, that is, to the requirement that Px and SN hold not only for w but also for any w_σ where σ is a permutation of $\Omega_{\vec{R}}$ satisfying (C) and w_σ is defined as in (35), remains an open question. Although much of the reasoning used in the proof of Theorem 3 could be used with $\mathcal{S}_{\vec{P}}$ replaced by the group of all permutations of $\Omega_{\vec{R}}$ satisfying (C), we lack insight into the structure of this group to allow us to draw interesting conclusions.

Conclusion

Inspired by Miller's Weather Example and its underlying notion of a *translation* we have considered the extent to which a simple permutation of atoms can be formulated, or explained, in terms of a translation. It turns out that this is always the case for purely unary languages whilst for general polyadic languages it requires the permutation to also satisfy a certain property (C). This also establishes a precise connection between translations and those permutations that can be extended to automorphisms of the overlying structure since (C) is again exactly the additional ingredient needed in that case too.

A salient feature of Miller's Weather Example is that it reveals an underlying rational commitment to adopting beliefs that are *translation-proof*. Whilst the most rigid meaning one might give to that expression is that beliefs should be *translation-invariant* we argue that within the context of Pure Inductive Logic a more catholic interpretation might be that the translation preserves the *rationality* of the beliefs, rather than the actual quantitative beliefs themselves. Formally this leads to a meta-principle which we have characterised for certain rationality criteria in unary languages, showing it (for languages with at least three predicate symbols) to lie strictly between simply observance of these criteria and full preservation of belief values under translation.

References

[1] Hill, A.J. & Paris, J.B., An Analogy Principle in Inductive Logic, *Annals of Pure and Applied Logic,* 2013, **164**:1293-1321.

[2] Miller, D., Popper's Qualitative Theory of Verisimilitude, *British Journal for the Philosophy of Science,* 1974, **25**(2):166-177.

[3] Miller, D., *Out of Error,* Ashgate Publishing Company, 2006.

[4] Miller, D., The Distance between Constituents, *Synthese*, **38**(2):197-212, 1978.

[5] Paris, J.B. & Vencovská, A., *Pure Inductive Logic*, in the Association of Symbolic Logic Perspectives in Mathematical Logic Series, Cambridge University Press, April 2015.

[6] Paris, J.B. & Vencovská, A., A Note on Nathanial's Invariance Principle in Polyadic Inductive Logic, in *Logic and its Applications, ICLA, Proceedings of the 4th Indian Logic Conference*, Dehli, India, January 5-11, 2011, eds. M.Banerjee & A.Seth. pp137-146.

[7] Ronel, T. & Vencovská, A., Invariance Principles in Polyadic Inductive Logic, *Logique et Analyse,* 2014, **228**: 541-561.

[8] Vencovská, A.,*Pure Inductive Logic – Nesin Maths Village Summer School Course Notes*, available at
http://www.maths.manchester.ac.uk/ jeff/lecture-notes/NesinNotes.pdf