Goodman's Extensional Isomorphism and Syntactical Interpretations

Marek POLAŃSKI

BIBLID [0495-4548 (2009) 24: 65; pp. 203-211]

ABSTRACT: The aim of the present paper is to provide a model-theoretic explication of Nelson Goodman's concept of extensional isomorphism. The term "extensional isomorphism" has been infomally introduced by Nelson Goodman in the beginning paragraph of his *The Structure of Appearance*. After some conceptual clarifications Goodman's concept of isomorphy turns out to be closely related to some variant of set-theoretic definability and some variants of syntactical interpretability.

Keywords: Goodman, extensional isomorphism, syntactical interpretations, model theory.

1. Introductory remarks

The concept of extensional isomorphism was introduced by Nelson Goodman in an informal manner in the beginning paragraph of his *The Structure of Appearance*. It was Goodman's intention to generalize the concept of explicit definability. Goodman considers the conditions of identity of the extensions of the definiendum and the definiens too strong for some philosophical purposes, especially in the context of so called constructional systems. For such purposes, he argues, the requirement of extensional identity should be replaced by a suitably formulated condition of isomorphy between the extensions of the terms in question. The aim of this paper is to reconstruct Goodman's idea in terms of model-theoretic semantics and to show how is such a reconstruction related to some widely discussed notions of set-theoretic definability and syntactic interpretability.

2. Goodman's extensional isomorphism: an informal exposition

2.1. Goodman's original definition

Extensional isomorphism was conceived by Goodman as a relation between expressions belonging to two interpreted languages. According to Goodman an expressions α is extensionally isomorphic to an expression β if there exists an injective mapping of a certain kind which transforms the extension of β onto that of α. In his *The Structure of Appearance* Goodman provides the following characterization of this special kind of isomorphy:

> A relation R is isomorphic to a relation S in the sense here intended if and only if R can be obtained by consistently replacing the ultimate factors in S. Consistent replacement requires only that each not-null ultimate factor be replaced by one and only one not-null element; that different not-null ultimate factors be always replaced by different not-null elements; and that the null-class be always replaced by itself. Since the replacing elements need not be ultimate factors (...), this sort of isomorphism is not symmetric; for if R is isomorphic to S, still there may be no way of replacing the ultimate factors in R so as to obtain S. (Goodman 1977, pp. 10–11)

One of the crucial notions used in the above passage is that of an ultimate factor of a given relation which is characterized by Goodman in the following way.

> We may think of the extensions of definienda and definientia in question as relations — that is, as classes of couples, classes of triples, and classes of longer sequences of any uniform length. (...) By the *components* of a sequence I shall mean the elements that occupy entire places in the sequence.(...) On the other hand, if we progressively dissolve each component that is a sequence into its components, and every component that is a class into its members, and continue this until we reach elements that have no further members or components, we have what I call the *ultimate factors* of the sequence.(...) The ultimate factors of a relation or other class are reached in similar fashion. (...) An ultimate factor is always either an individual or the null class. (Goodman 1977, p. 10)

This rather sketchy characterization is followed by an appealing geometrical example where geometrical points are represented as couples of intersecting lines and some basic relations between points are reconstrcuted in terms of relations between sets of lines.

The present paper attempts to develope the ideas expressed in these two quotations. Goodman's idea of extensional isomorphy gives rise to a number of philosophical questions which have been already discussed in literature (compare: Hellman (1978)). Our principial concern here will be with some purely logical problems which are widely neglected in the literature und which derive from the purported analogy between the concept of extensional isomorphy and the usual notion of an explicit definition. We should focus on these logical problems putting aside the ontological motivation behind the concepts in question.

2.2. *Steps towards a model-theoretic explication*

Let us consider two languges L_1 and L_2 which are interpreted over two collections of individuals A and B, respectively. This means that to each n-ary primitive predicate P from L_1 (resp. L_2) is assigned an n-ary relation on A (resp. on B) which is a set of n-tuples of individuals from A (resp. from B). These fixed interpretations can be represented as structures of the form $\mathcal{A} = (A, P_1^{\mathcal{A}}, \ldots, P_m^{\mathcal{A}})$ and $\mathcal{B} = (B, R_1^{\mathcal{B}}, \ldots, R_k^{\mathcal{B}})$. Let $\mathrm{SET}(A)$ and $\mathrm{SET}(B)$ be the families of all sets over A and B, respectively. Intuitively, these families consist of all sets whose ultimate factors belong to $A \cup \{\emptyset\}$ and $B \cup \{\emptyset\}$, respectively. This is far from being precise but it seems to be clear enough for the proposes of an informal exposition. We shall call structures of the form $(A, \mathrm{SET}(A), \in, P_1^{\mathcal{A}}, \ldots, P_m^{\mathcal{A}})$ *full* ϵ*-structures*.

According to the informal characterization quoted above, Goodman's notion of isomorphy can be reconstructed in terms of full ϵ-structures as a certain relation of *embeddability* between two such structures $\mathcal{A}$ and $\mathcal{B}$. Goodman's informal requirements amount to the existence of an injective function f from $A \cup \mathrm{SET}(A)$ into $B \cup \mathrm{SET}(B)$ such that:

- [G 1] f preserves the membership relation,
- [G 2] f transforms the empty set onto itself,
- [G 3] f transforms A and each $P_i^{\mathcal{A}}$ onto sets which are definable in $\mathcal{B}$ (in the usual sense of "definable").

The elements of $f[A]$ serve as counterparts of the individuals in A. Hence, the image $f[\mathcal{A}]$ of $\mathcal{A}$ (i.e. the structure $(f[A], f[\mathrm{SET}(A)], \in, \ldots)$) can be regarded as a certain *definable substructure of* $\mathcal{B}$. In the following we shall give a precise model-theoretic description of the substructure relation which corresponds to the above conditions.

It seems to be natural to expect that if $\mathcal{A}$ is embeddable in such a way into $\mathcal{B}$ then at least each sentence framed in terms of $L_1 \cup \{\in\}$ which is true in $\mathcal{A}$ (more precisely: in $(\mathcal{A}, \mathrm{SET}(A), \in)$) can be uniformly translated into a sentence fomulated in terms of $L_2 \cup \{\in\}$ which is true in $\mathcal{B}$, and that such a translation commutes with the logical opeartions and preserves to some extend the set-theoretic relations expressed by each translated sentence. In other words, if a structure is embeddable in another one in the sense indicated above then, as may be reasonably expected, the theory of $\mathcal{A}$ should be reducible to the theory $\mathcal{B}$ via a translation. All sentences about the objects in the universe of $\mathcal{A}$ can be in some sense reinterpreted as sentences about sets whose ultimate factors are elements of the universe of $\mathcal{B}$. Moreover, such a translation should be induced by the embedding f in question in the sense that for each formula φ of the translated language and any tuple of objects $\bar{a}$ from the universe of the embedded structure $\mathcal{A}$: φ is satisfied in $\mathcal{A}$ by $\bar{a}$ if and only if the corresponding tuple $f(\bar{a})$ satisfies in $\mathcal{B}$ the translation $\tau(\varphi)$. The question as to whether there is a canonical form of a translation corresponding to that kind of embedding is the second main problem to the solution of which the paper is intended to contribute. In order to give a precise answer to these questions we shall first provide a portion of conceptual background.

3. *Goodman-embeddings and syntactical interpretations*

3.1. *ϵ-structures*

In this paper we shall reconstruct Goodman's notion of isomorphy as a relation between two models. Let us begin with a precise description of the relata. The following definition generalizes of the informal notion of a full ϵ-structure.

Definition 1. *An ϵ-structure for a vocabulary $L = \{P_1, \ldots, P_m\}$ is a structure of the form $\mathcal{A} = (A, \mathrm{Set}^{\mathcal{A}}, \epsilon^{\mathcal{A}}, P_1^{\mathcal{A}}, \ldots, P_m^{\mathcal{A}})$, such that A is a non empty set, each $P_i^{\mathcal{A}}$ is a relation on A whose arity equals the arity of the corresponding predicate P_i, $\mathrm{Set}^{\mathcal{A}}$ is a subset of A, $\epsilon^{\mathcal{A}}$ is a binary relation on A, and $\mathcal{A}$ is a model the following axioms, where φ is any formula of $L \cup \{\mathrm{Set}, \epsilon\}$:*

[Set] $\forall x(\exists y(y\epsilon x) \rightarrow \mathrm{Set}(x))$

[Ext] $\forall x\forall y(\mathrm{Set}(x) \wedge \mathrm{Set}(y) \wedge \forall z(z\epsilon x \leftrightarrow z\epsilon y) \rightarrow x = y))$

[Urel] $\exists x(\mathrm{Set}(x) \wedge \forall y(y\epsilon x \leftrightarrow \neg\,\mathrm{Set}(y)))$

[Comp] $\forall x(\mathrm{Set}(x) \rightarrow \exists y(\mathrm{Set}(y) \wedge \forall z(z\epsilon y \leftrightarrow z\epsilon x \wedge \varphi(z)))), x \notin Fr(\varphi(z))$.

Intuitively, an ϵ-structure for a given vocabulary L is a semantical interpretation of the extended vocabulary $L \cup \{\mathrm{Set}, \epsilon\}$ which satisfies some fairly minimal set-theoretical postulates. The universe of each such structure $\mathcal{A}$ contains objects of two kinds: individuals called *sets of* $\mathcal{A}$ (which are elements of $\mathrm{Set}^{\mathcal{A}}$) and individuals which are called *urelements* *of* $\mathcal{A}$. $\epsilon^{\mathcal{A}}$ is called the *membership relation of* $\mathcal{A}$. The postulate [Set] says intuitively that urelements do not have any members. [Ext] requires any two sets having the same elements to be identical. By [Urel], in each ϵ-structure $\mathcal{A}$ the collection $\mathrm{Urel}^{\mathcal{A}} =_{df} A \setminus \mathrm{Set}^{\mathcal{A}}$ of all urelements in $\mathcal{A}$ corresponds to the element of $\mathrm{Set}^{\mathcal{A}}$ which satisfies in $\mathcal{A}$ the formula $\forall y(y\epsilon x \leftrightarrow \neg\,\mathrm{Set}(y))$. [Comp] implies that the extension of each formula expressible in $L \cup \{\mathrm{Set}, \epsilon\}$ corresponds in an analogous way to an element of $\mathrm{Set}^{\mathcal{A}}$. The postulates imply that each ϵ-structure $\mathcal{A}$ contains exactly one element of $\mathrm{Set}^{\mathcal{A}}$, denoted by $\emptyset^{\mathcal{A}}$, which does not have any members. Observe that we do not assume the fields of the relations $P_i^{\mathcal{A}}$ to be included in $\mathrm{Urel}^{\mathcal{A}}$. In other words, some of the relations $P_i^{\mathcal{A}}$ may hold between sets in $\mathcal{A}$ and urelements in $\mathcal{A}$. We call a k-ary relation R on A *definable in* $\mathcal{A}$ if and only if there exists a formula $\xi(\bar{x})$ in the language $L \cup \{\epsilon, \mathrm{Set}\}$ with k pairwise dictinct free variables such that $R = \{\bar{a} \in A^k \colon \mathcal{A} \models \xi(\bar{x})[\bar{x}\colon \bar{a}]\}$. An ϵ-structure $\mathcal{A} = (A, \mathrm{Set}^{\mathcal{A}}, \epsilon^{\mathcal{A}}, P_1^{\mathcal{A}}, \ldots, P_m^{\mathcal{A}})$ is said to be *definable in* an ϵ-structure $\mathcal{B}$ if and only if A, $\epsilon^{\mathcal{A}}$, $\mathrm{Set}^{\mathcal{A}}$, and $P_1^{\mathcal{A}}, \ldots, P_m^{\mathcal{A}}$ are definable in $\mathcal{B}$. The following notion of a definable substructure will play a crucial role by the clarification of Goodman's ideas.

Definition 2. $\mathcal{A}$ *is a* definable ϵ-substructure of $\mathcal{B}$ *if and only if*

1. $\mathcal{A}$ *is definable in* $\mathcal{B}$*,*
2. $\epsilon^{\mathcal{A}} = \epsilon^{\mathcal{B}} \cap A^2$*,*
3. $\emptyset^{\mathcal{A}} = \emptyset^{\mathcal{B}}$.

Corollary 3. *If* $\mathcal{A}$ *is a definable* ϵ*-substructure of* $\mathcal{B}$ *then* $\mathrm{Set}^{\mathcal{A}} \subseteq \mathrm{Set}^{\mathcal{B}}$.

Proof. Let $a \in \mathrm{Set}^{\mathcal{A}}$. Then either $a = \emptyset^{\mathcal{A}} = \emptyset^{\mathcal{B}} \in \mathrm{Set}^{\mathcal{B}}$ or for some $d \in A\colon d\epsilon^{\mathcal{A}}a$ which, by Definition 2, implies that for some $d \in B\colon d\epsilon^{\mathcal{B}}a$, and hence (by [Set]) $a \in \mathrm{Set}^{\mathcal{B}}$.

The above definition permitts some sets in a ϵ-structure $\mathcal{B}$ to become urelements in some of their definable ϵ-substructures. This feature may be regarded as quite odd, but it is of crucial importance for the purposes of the present paper for the main idea behind Goodman's notion of isomorphism and the paradigmatic examples given by him in order to illustrate his idea are exactly of this kind. The notion of a definable ϵ-substructure already defined is closely connected to the notion of an initial substructure which is used in the study of models of set theories. According to a standard definition, $\mathcal{A}$ is called an *initial substructure of* $\mathcal{B}$ if and only if $A \subseteq B$, $\mathrm{Set}^{\mathcal{A}} = \mathrm{Set}^{\mathcal{B}} \cap A$, the interpretations of all primitive predicates of L in $\mathcal{A}$ are just the restrictions of their interpretations in $\mathcal{B}$, and for all $a \in A$ and for all $b \in B\colon b\epsilon^{\mathcal{A}}a \Leftrightarrow b\epsilon^{\mathcal{B}}a$ (compare: Barwise (1975, p. 34)). The following definition generalizes this idea.

Definition 4. $\mathcal{A}$ *is a* definable initial substructure of $\mathcal{B}$ *if and only if*

1. $\mathcal{A}$ *is definable in* $\mathcal{B}$*,*
2. *for all* $a \in A$ *and for all* $b \in B\colon b\epsilon^{\mathcal{A}}a \Leftrightarrow b\epsilon^{\mathcal{B}}a$*,*
3. $\mathrm{Set}^{\mathcal{A}} \subseteq \mathrm{Set}^{\mathcal{B}}$.

Corollary 5. *If $\mathcal{A}$ is a definable initial substructure of $\mathcal{B}$ then $\emptyset^{\mathcal{A}} = \emptyset^{\mathcal{B}}$ and* $\text{Urel}^{\mathcal{A}} \subseteq \text{Urel}^{\mathcal{B}}$. (*Hence:* $\text{Set}^{\mathcal{A}} = \text{Set}^{\mathcal{B}} \cap A$ *and* $\text{Urel}^{\mathcal{A}} = \text{Urel}^{\mathcal{B}} \cap A$.)

Proof. Definition 4 implies that $\emptyset^{\mathcal{A}} \in \text{Set}^{\mathcal{B}}$ and that there is no $a \in B$ such that $a\epsilon^{\mathcal{B}}\emptyset^{\mathcal{A}}$. Hence, by [Ext], $\emptyset^{\mathcal{A}} = \emptyset^{\mathcal{B}}$. Consider now an $a \in \text{Urel}^{\mathcal{A}}$ and assume that $a \notin \text{Urel}^{\mathcal{B}}$. Clearly, $a \neq \emptyset^{\mathcal{B}}$ and $a \in \text{Set}^{\mathcal{B}}$. Hence, there is a $b \in B$ such that $b\epsilon^{\mathcal{B}}a$. By Definition 4 again there is a $b \in B$ such that $b\epsilon^{\mathcal{A}}a$. Hence $a \notin \text{Urel}^{\mathcal{A}}$, which is a contradiction.

The next lemma shows that the only difference between these two kinds of substructure amounts to the status which the urelements of the substructure have in the superstructure.

Lemma 6. *Let $\mathcal{A}$ be a definable ϵ-substructure of $\mathcal{B}$. The following conditions are equivalent.*

1. *$\mathcal{A}$ is a definable initial substructure of $\mathcal{B}$.*
2. *For all $a \in A$ and all $b \in B$: $b\epsilon^{\mathcal{B}}a \Rightarrow b \in A$.*
3. $\text{Urel}^{\mathcal{A}} \subseteq \text{Urel}^{\mathcal{B}}$.

Proof. The implication *1* $\Rightarrow$ *2* is obvious. We prove the implication *2* $\Rightarrow$ *3*. Assume *2* and suppose that $a \in \text{Urel}^{\mathcal{A}} \setminus \text{Urel}^{\mathcal{B}}$. Then $a \neq \emptyset^{\mathcal{B}}$ and so for some $b \in B$: $b\epsilon^{\mathcal{B}}a$. By *2* $b \in A$ and so, by Definition 2, $b\epsilon^{\mathcal{A}}a$ which leads to a contradiction. Now we prove the implication *3* $\Rightarrow$ *1*. Let $\mathcal{A}$ and $\mathcal{B}$ satisfy the condition *3*. By Corollary 3 we have $\text{Set}^{\mathcal{A}} \subseteq \text{Set}^{\mathcal{B}}$. Hence, it remains to show that for all $a \in A$ and all $b \in B$: $b\epsilon^{\mathcal{A}}a \Leftrightarrow b\epsilon^{\mathcal{B}}a$. The implication from left to right is an obvious consequence of the fact that $\epsilon^{\mathcal{A}} = \epsilon^{\mathcal{B}} \cap A^2$. To show the converse implication assume that $b\epsilon^{\mathcal{B}}a$ but not $b\epsilon^{\mathcal{A}}a$. Then $a \notin \text{Urel}^{\mathcal{B}} \cup \{\emptyset^{B}\}$ and hence, by $\text{Urel}^{\mathcal{A}} \subseteq \text{Urel}^{\mathcal{B}}$ and $\emptyset^{\mathcal{A}} = \emptyset^{\mathcal{B}}$, we have $a \in \text{Set}^{\mathcal{A}} \setminus \{\emptyset^{\mathcal{A}}\}$. Let, by [Comp] and by the fact that A is definable in $\mathcal{B}$, a^* be the element of $\text{Set}^{\mathcal{B}}$ such that for all $d \in B$: $d\epsilon^{\mathcal{B}}a^* \Leftrightarrow d\epsilon^{\mathcal{B}}a$ and $d \in A$. Since $b\epsilon^{\mathcal{B}}a$ and non $b\epsilon^{\mathcal{B}}a^*$, we have: $a \neq a^*$. But, clearly, for each $d \in A$: $d\epsilon^{\mathcal{A}}a^* \Leftrightarrow d\epsilon^{\mathcal{A}}a$. Since $a \in \text{Set}^{\mathcal{A}} \setminus \{\emptyset^{\mathcal{A}}\}$ we have also $a^* \in \text{Set}^{\mathcal{A}}$, and since $a \neq a^*$ we conclude that $\mathcal{A}$ does not satisfy [Ext], which is impossible.

3.2. Goodman-embeddings

Now we turn back to the problem of clarifying Goodman's original idea. Using the conceptual resources provided in the previous section we define now the concept of a Goodman-embedding which is in our opinion an adequate formal counterpart of Goodman's informal notion of isomorphism.

Definition 7. *We call any injective mapping f a* Goodman-embedding *(*G-embedding, *for short)* of $\mathcal{A}$ into $\mathcal{B}$ *if and only if the structure $f[\mathcal{A}]$ is a definable ϵ-substructure of $\mathcal{B}$. We call f a* simple *G-embedding of $\mathcal{A}$ into $\mathcal{B}$ if $f[\text{Urel}^{\mathcal{A}}] \subseteq \text{Urel}^{\mathcal{B}}$.*

It is easy to check that a mapping f is a G-embedding of an ϵ-structure $\mathcal{A}$ into an ϵ-structure $\mathcal{B}$ if and only f satisfies the conditions [G1] - [G3] (compare: section 2.2 of this paper) together with the following requirements:

[G 4] f maps $\mathrm{Set}^{\mathcal{A}}$ onto a definable subcollection of $\mathrm{Set}^{\mathcal{B}}$,

[G 5] f transforms the universe of $\mathcal{A}$ onto a colection of objects from B which is definable in $\mathcal{B}$.

The last two conditions do not follow from Goodman's informal considerations. They seem, however, to be in accordance with Goodman's ideas While the general concept of a G-embedding is closely connected to the notion of a definable ϵ-substructure, the concept of a simple G-embedding, which is a special variant of the former, turns out to correspond in an analogous way to the notion of a definable initial substructure. Observe that for each object $a \in A$ such that $f(a)$ is a set in $\mathcal{B}$, the collection $\{f(b)\colon b\epsilon^{\mathcal{A}}a\}$ (which is identical with $\{d\epsilon^{\mathcal{B}}f(a)\colon d \in \delta^{\mathcal{B}}\}$ where δ is a formula which defines in $\mathcal{B}$ the image of A under f) corresponds to a set u in $\mathcal{B}$ such that for all $z \in A$: if $z\epsilon^{\mathcal{B}}u$ then $z\epsilon^{\mathcal{B}}f(a)$. The converse impication, however, does not hold in general. In fact, it holds only for simple G-embeddings.

Lemma 8. *For each G-embedding f of $\mathcal{A}$ into $\mathcal{B}$ the following conditions are equivalent.*

1. *f is a simple G-embedding of $\mathcal{A}$ into $\mathcal{B}$.*
2. *$f[\mathcal{A}]$ is a definable initial substructure of $\mathcal{B}$.*
3. *For each $a \in A\colon f(a) \in \mathrm{Urel}^{\mathcal{B}}$ or $f(a) = \{f(b)\colon b\epsilon^{\mathcal{A}}a\}$.*

Proof. Let f be a G-embedding of $\mathcal{A}$ into $\mathcal{B}$. We prove the implication *1* $\Rightarrow$ *2*. Assume *1*. $f[\mathcal{A}]$ is then a definable ϵ-substructure of $\mathcal{B}$. Moreover, since f is an isomorphism (in the usual model-theoretic sense) from $\mathcal{A}$ onto $f[\mathcal{A}]$, we have $\mathrm{Urel}^{f[\mathcal{A}]} = f[\mathrm{Urel}^{\mathcal{A}}]$. Since f is simple we have $\mathrm{Urel}^{f[\mathcal{A}]} \subseteq \mathrm{Urel}^{\mathcal{B}}$ which, by Lemma 6, implies that $f[\mathcal{A}]$ is a definable initial substructure of $\mathcal{B}$. To show the implication *2* $\Rightarrow$ *3* assume that $f(a) \in \mathrm{Set}^{\mathcal{B}}$. Now it is sufficient to prove that for each $d \in B$: if $d\epsilon^{\mathcal{B}}f(a)$ then $d \in \{f(b)\colon b\epsilon^{\mathcal{A}}a\}$. Let $d\epsilon^{\mathcal{B}}f[A]$. Since $f[\mathcal{A}]$ is a definable ϵ-substructure of $\mathcal{B}$, we have, by Lemma 6, $d \in f[A]$. So for some $b \in A\colon d = f(b)$. By ϵ-preservation of f we have $b\epsilon^{\mathcal{A}}a$. Hence, $d \in \{f(b)\colon b\epsilon^{\mathcal{A}}a\}$. Now we prove the implication *3* $\Rightarrow$ *1*. Assume *3* and suppose that for some $a \in \mathrm{Urel}^{\mathcal{A}}\colon f(a) \in \mathrm{Set}^{\mathcal{B}}$. By definition of a G-embedding, $f(a) \neq \emptyset^{\mathcal{B}}$. But then for some $b \in B\colon b\epsilon^{\mathcal{B}}f(a)$ and, by *3*, for some $d \in A\colon d\epsilon^{\mathcal{A}}a$ and $f(d) = b$, which is impossible, since $a \in \mathrm{Urel}^{\mathcal{A}}$.

Let us turn to the main question of this paper — to the problem of a syntactic characterization of the relation of G-embeddability. We will show that the existence of a G-embedding of $\mathcal{A}$ into $\mathcal{B}$ implies the existence of a certain interesting kind of effective translation of all sentences true in $\mathcal{A}$ into sentences true in $\mathcal{B}$.

3.3. Goodman-embeddings and syntactical interpretations

Let us make again some terminological clarifications. Under a *translation code for* $L_1 \cup \{\mathrm{Set}, \epsilon\}$ *in* $L_2 \cup \{\mathrm{Set}, \epsilon\}$, where $L_1 = \{P_1, \ldots, P_m\}$, we shall understand a tuple $c = \langle \delta, \sigma, \eta, \xi_1, \ldots, \xi_m \rangle$ of formulas in $L_2 \cup \{\mathrm{Set}, \epsilon\}$ such that $Fr(\delta) = Fr(\sigma) = \{v_1\}$, $Fr(\eta) = \{v_1, v_2\}$ and $Fr(\xi_i) = \{v_1, \ldots, v_k\}$, for each k-ary ξ_i $(i \leq m)$.

The formulas δ and σ are called the domain formula and the set formula of c, respectively. Each translation code c of this kind induces a translation which is defined as the function τ_c from $Fm(L_1 \cup \{\text{Set}, \epsilon\})$ into $Fm(L_2 \cup \{\text{Set}, \epsilon\})$ such that $\tau_c(x = y) = x = y$, $\tau_c(\text{Set}(x)) = \sigma(x/v_1)$, $\tau_c(x\epsilon y) = \eta(x/v_1, y/v_2)$, $\tau_c(P_i(x_1, \ldots, x_k) = \xi(x_1/v_1, \ldots, x_k/v_k)$, $\tau_c(\neg\varphi) = \neg\tau_c(\varphi)$, $\tau_c(\varphi \circ \psi) = \tau_c(\varphi) \circ \tau_c(\psi)$ (where $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$), $\tau_c(\forall x\varphi) = \forall x(\delta(x/v_1) \rightarrow \tau_c(\varphi))$, and $\tau_c(\exists x\varphi) = \exists x(\delta(x/v_1) \wedge \tau_c(\varphi))$. On the semantical level c generates also a certain mapping Γ_c, called *the canonical construction for* c which is defined for all those ϵ-structures $\mathcal{A}$ for L_2 in which the formula $\exists v_1\delta$ is true. For each such structure $\mathcal{A}$ we define $\Gamma_c(\mathcal{A})$ as the tuple $(\delta^{\mathcal{A}}, \sigma^{\mathcal{A}}, \eta^{\mathcal{A}}, \xi_1^{\mathcal{A}}, \ldots, \xi_m^{\mathcal{A}},)$. Clearly, $\Gamma_c(\mathcal{A})$ is then a semantical interpretation for $L_1 \cup \{\text{Set}, \epsilon\}$ though not necessarily an ϵ-structure for L_1. The following lemma connects the both items in an expected way. Since it can be demonstrated by an easy induction, we state it without proof.

Lemma 9 (Translation Lemma)**.** *Let* c *be a translation code for* $L_1 \cup \{\text{Set}, \epsilon\}$ *in* $L_2 \cup \{\text{Set}, \epsilon\}$ *with a domain formula* δ*. Then for each formula* $\varphi(x_1, \ldots, x_n)$ *in* $L_1 \cup \{\text{Set}, \epsilon\}$*, each* ϵ*-structure* $\mathcal{A}$ *for* L_2 *in which* $\exists v_1\delta$ *is true, and all* $a_1, \ldots, a_n \in \delta^{\mathcal{A}}$*:*

$$\Gamma_c(\mathcal{A}) \models \varphi[x_1 : a_1, \ldots, x_n : a_n] \Leftrightarrow \mathcal{A} \models \tau_c(\varphi)[x_1 : a_1, \ldots, x_n : a_n].$$

τ_c is called a *relative interpretation of* a set of sentences T_1 in $L_1 \cup \{\text{Set}, \epsilon\}$ *into* a set of sentences T_2 in $L_2 \cup \{\text{Set}, \epsilon\}$ if $T_2 \vdash \exists x\delta$ and $T_2 \vdash \tau_c(\varphi)$, for each formula φ in $L_1 \cup \{\text{Set}, \epsilon\}$ such that $T_1 \vdash \varphi$. The following result is an obvious consequence of the previous lemma.

Proposition 10. $\mathcal{A}$ *is definable in* $\mathcal{B}$ *if and only if there exists a relative interpretation* τ_c *of* $\text{Th}(\mathcal{A})$ *into* $\text{Th}(\mathcal{B})$ *such that* $\Gamma_c(\mathcal{B}) = \mathcal{A}$.

In this paper we are mainly interested in syntactical interpretations which preserve ϵ, i.e. such that $\tau_c(x\epsilon y) = x\epsilon y$. Interpretation of this kind are called ϵ*-interpretations*. It is an immediate consequence of the Translation Lemma that if τ_c is an ϵ-interpretation then for each structure $\mathcal{A}$ in the domain of Γ_c the relation $\epsilon^{\Gamma_c(\mathcal{A})}$ is just the restriction of $\epsilon^{\mathcal{A}}$ to the universe of $\Gamma_c(\mathcal{A})$.

Let us now consider two special kinds of such interpretations which turn out to be closely related to the two concepts of a definable substructure introduced above.

Definition 11. *Let* τ_c *be an* ϵ*-interpretation of a set of sentences* T_1 *in* $L_1 \cup \{\text{Set}, \epsilon\}$ *in a set of sentences* T_2 *in* $L_2 \cup \{\text{Set}, \epsilon\}$ *with domain formula* δ *and set formula* σ*.*

1. τ_c *is called a* weakly transitive ϵ-interpretation of T_1 into T_2 *if and only if* $T_2 \vdash \forall v_1(\sigma(v_1) \rightarrow (\text{Set}(v_1) \wedge \forall v_2(v_2\epsilon v_1 \rightarrow \delta(v_2))))$.
2. τ_c *is called a* transitive ϵ-interpretation of T_1 into T_2 *if and only if* τ_c *is a weakly transitive* ϵ*-interpretation of* T_1 *into* T_2 *and* $T_2 \vdash \forall v_1(\delta(v_1) \wedge \neg\sigma(v_1)) \rightarrow \neg\,\text{Set}(v_1))$.

The following proposition is the main formal result of this paper.

Proposition 12. *Let $\mathcal{A}$ and $\mathcal{B}$ be ϵ-structures for L_1 and L_2, respectively. Let f be a function from A into B.*

1. *f is a G-embedding of $\mathcal{A}$ into $\mathcal{B}$ if and only if there exists a weakly transitive ϵ-interpretation τ_c of* $\mathrm{Th}(\mathcal{A})$ *into* $\mathrm{Th}(\mathcal{B})$ *such that f is an isomorphism from $\mathcal{A}$ onto $\Gamma_c(\mathcal{B})$.*
2. *f is a simple G-embedding of $\mathcal{A}$ into $\mathcal{B}$ if and only if there exists a transitive ϵ-interpretation τ_c of* $\mathrm{Th}(\mathcal{A})$ *into* $\mathrm{Th}(\mathcal{B})$ *such that f is an isomorphism from $\mathcal{A}$ onto $\Gamma_c(\mathcal{B})$.*

Proof. We start with *1.* First we prove the implication from right to left. Let τ_c be a weakly transitive ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$ such that f is an isomorphism of $\mathcal{A}$ onto $\Gamma_c(\mathcal{B})$ (i.e., $f[\mathcal{A}] = \Gamma_c(\mathcal{B})$). We show that the last structure is a definable ϵ-substructure of $\mathcal{B}$. Clearly, $\Gamma_c(\mathcal{B})$ is definable in $\mathcal{B}$ and $\epsilon^{\Gamma_c(\mathcal{B})} = \epsilon^{\mathcal{B}} \cap (\delta^{\mathcal{B}})^2$, where δ is the domain formula of c. It remains to show that $\emptyset^{\Gamma_c(\mathcal{B})} = \emptyset^{\mathcal{B}}$. Obviously, we have $\Gamma_c(\mathcal{B}) \models \mathrm{Set}(x) \wedge \forall y\, \neg(y\epsilon x)\ [x : \emptyset^{\Gamma_c(\mathcal{B})}]$. By the Translation Lemma and the fact that τ_c is an ϵ-interpretation we conclude that $\mathcal{B} \models \mathrm{Set}(x) \wedge \forall y\, \neg(y\epsilon x)\ [x : \emptyset^{\Gamma_c(\mathcal{B})}]$. By [Ext] we have then $\emptyset^{\Gamma_c(\mathcal{B})} = \emptyset^{\mathcal{B}}$.

We show now the implication from left to right. Let $f[\mathcal{A}]$ be a definable ϵ-substructure of $\mathcal{B}$. Since $f[\mathcal{A}]$ is definable in $\mathcal{B}$, there exists a translation code $c = \langle \delta, \sigma, \eta, \xi_1, \dots, \xi_m \rangle$ for an appropriate language for $\mathcal{A}$ in a language for $\mathcal{B}$ such that $\Gamma_c(\mathcal{B}) = f[\mathcal{A}]$. By Translation Lemma, for all φ such that $\mathrm{Th}(f[\mathcal{A}]) \vdash \varphi$ we have $\mathrm{Th}(\mathcal{B}) \vdash \tau_c(\varphi)$. Hence τ_c is an interpretation of $\mathrm{Th}(f[\mathcal{A}])$ (and hence of $\mathrm{Th}(\mathcal{A})$) into $\mathrm{Th}(\mathcal{B})$. Since $\epsilon^{\Gamma_c(\mathcal{B})} = \epsilon^{f[\mathcal{A}]} = \epsilon^{\mathcal{B}} \cap (\delta^{\mathcal{B}})^2$, we conclude, using the Translation Lemma, that $\tau_c[x\epsilon y]$ and $x\epsilon y$ define the same relation in $\mathcal{B}$. So we can assume that $\tau_c[x\epsilon y] = x\epsilon y$ which shows that τ_c is an ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$. It remains to show that $\mathrm{Th}(\mathcal{B}) \vdash \sigma(v_1) \to \mathrm{Set}(v_1) \wedge \forall v_2(v_2\epsilon v_1 \to \delta(v_2))$. If $a \in \sigma^{\mathcal{B}} \subseteq \delta^{B} = A$ then by $\Gamma_c(\mathcal{B}) = f[\mathcal{A}]$ we have: $a \in \mathrm{Set}^{f[\mathcal{A}]}$. Since $f[\mathcal{A}]$ is a definable ϵ-substructure of $\mathcal{B}$ we have, by Corollary 3, $a \in \mathrm{Set}^{\mathcal{B}}$. Let $b \in B$ be such that$\mathcal{B}$ $b\epsilon^{\mathcal{B}}a$. We show that $b \in \delta^{\mathcal{B}}$. Assume the contrary. Let, by [Comp], a^* be the element of $\mathrm{Set}^{\mathcal{B}}$ such that for all $d \in B$: $d\epsilon^{\mathcal{B}}a^* \Leftrightarrow (d\epsilon^{\mathcal{B}}a$ and $d \in \delta^{\mathcal{B}})$. Since $b\epsilon^{\mathcal{B}}a$ and non $b\epsilon^{\mathcal{B}}a^*$, we have, by [Ext]: $\neq a^*$. But $a^* \in \mathrm{Set}^{f[\mathcal{A}]}$. For if $a^* = \emptyset^{\mathcal{B}}$ then $a^* = \emptyset^{f[\mathcal{A}]}$, and if $a^* \neq \emptyset^{\mathcal{B}}$ then for some $d \in B$: $d\epsilon^{\mathcal{B}}a^*$, i.e. for some $d \in A$: $d\epsilon^{\mathcal{A}}a^*$. Moreover, for all $d \in A$: $d\epsilon^{\mathcal{A}}a^* \Leftrightarrow d\epsilon^{\mathcal{A}}a$. Hence, by [Ext]: $a = a^*$, which leads to a contradiction. Therefore for all $b \in B$: $b\epsilon^{\mathcal{B}}a \Rightarrow b \in \delta^{\mathcal{B}}$. This shows that τ_c is a weakly transitive ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$ and f is an isomorphism from $\mathcal{A}$ onto $\Gamma_c(\mathcal{B})$.

Now we prove *2.* We start with the implication from right to left. Let τ_c be a transitive ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$ such that f is an isomorphism from $\mathcal{A}$ onto $\Gamma_c(\mathcal{B})$. Since τ_c is a weakly transitive ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$ so, by *1*, f is a G-embedding of $\mathcal{A}$ into $\mathcal{B}$ We show that $f[\mathrm{Urel}^{\mathcal{A}}] \subseteq \mathrm{Urel}^{\mathcal{B}}$. Suppose that $a \in \mathrm{Urel}^{f[\mathcal{A}]} = f[\mathrm{Urel}^{\mathcal{A}}]$. Hence $a \in \delta^{\mathcal{B}} \setminus \sigma^{\mathcal{B}}$. By transitivity of τ_c we conclude that $a \notin \mathrm{Set}^{\mathcal{B}}$ and so $a \in \mathrm{Urel}^{\mathcal{B}}$. Thus f is a simple G-embedding of $\mathcal{A}$ into $\mathcal{B}$. We show now the converse implication.

Assume now that f is a simple G-embedding of $\mathcal{A}$ into $\mathcal{B}$. By Lemma 8 $f[\mathcal{A}]$ is a definable initial substructure of $\mathcal{B}$ and hence, by Lemma 6, a definable ϵ-substructure of

$\mathcal{B}$ such that $\mathrm{Urel}^{f[\mathcal{A}]} \subseteq \mathrm{Urel}^{\mathcal{B}}$. By *1* there exsits a weakly transitive ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$ such that $\Gamma_c(\mathcal{B}) = f[\mathcal{A}]$. Let δ and σ be the domain and the set formula of c, respectively. If $a \in \delta^{\mathcal{B}} \cap (\neg\sigma)^{\mathcal{B}}$, then $a \in \mathrm{Urel}^{f[\mathcal{A}]}$ and hence $a \in \mathrm{Urel}^{\mathcal{B}}$ which means that $a \notin \mathrm{Set}^{\mathcal{B}}$. Thus $\mathrm{Th}(\mathcal{B}) \vdash \forall v_1(\delta(v_1) \wedge \neg\sigma(v_1)) \rightarrow \neg\,\mathrm{Set}(v_1))$. Therefore, τ_c is a transitive ϵ-interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$.

The above considerations enable us to derive the following purely syntactical partial characterization of G-embeddings.

Corollary 13. *Let $\mathcal{A}$ and $\mathcal{B}$ be ϵ-structures for L_1 and L_2, respectively.*

1. *If $\mathcal{A}$ is G-embeddable into $\mathcal{B}$ then* $\mathrm{Th}(\mathcal{A})$ *is weakly transitively interpretable into* $\mathrm{Th}(\mathcal{B})$.
2. *If $\mathcal{A}$ is simply G-embeddable into $\mathcal{B}$ then* $\mathrm{Th}(\mathcal{A})$ *is transitively interpretable into* $\mathrm{Th}(\mathcal{B})$.

The converse implications do not hold in general. The existence of appropriate counterexamples is an easy consequence of the Upward Löwenheim-Skolem Theorem. To see this consider two ϵ-structures $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A}$ is (simply) G-embeddable into $\mathcal{B}$ and choose a cardinal $\kappa > \mathrm{card}(\mathcal{B})$. By the Upward Löwenheim-Skolem Theorem there exists an elementary extension $\mathcal{C}$ of $\mathcal{A}$ such that $\mathrm{card}(\mathcal{C}) = \kappa$. Since being an ϵ-structure is an elementary property $\mathcal{C}$ is an ϵ-structure with $\mathrm{Th}(\mathcal{C}) = \mathrm{Th}(\mathcal{A})$. By Corollary 13 $\mathrm{Th}(\mathcal{C})$ is (weakly) transitively interpretable into $\mathrm{Th}(\mathcal{B})$ but there is no G-embedding of $\mathcal{A}$ into $\mathcal{B}$ for $\mathrm{card}(\mathcal{B}) < \mathrm{card}(\mathcal{C})$. It is also easy to find ϵ-structures $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A}$ is G-embeddable into $\mathcal{B}$ and $\mathrm{card}(\mathrm{Urel}^{\mathcal{B}}) < \mathrm{card}(\mathrm{Urel}^{\mathcal{A}})$. Clearly, in such a case there is no simple G-embedding of $\mathcal{A}$ into $\mathcal{B}$ and hence no transitive interpretation τ_c of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$ such that $\Gamma_c(\mathcal{B}) = \mathcal{A}$. But this does not preclude the transitve interpretability of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$. It would be of some interest to find out whether there exist two infinite ϵ-structures $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A}$ is G-embeddable into $\mathcal{B}$ but there is no transitive interpretation of $\mathrm{Th}(\mathcal{A})$ into $\mathrm{Th}(\mathcal{B})$.

REFERENCES

Barwise, J., 1975. *Admissible Sets and Structures*. Berlin, Heidelberg, New York: Springer-Verlag.

Goodman, N., 1977. *The Structure of Appearance*. Third Edition. Dordrecht: Reidel.

Hellman, G., 1978. Accuracy and actuality. *Erkenntnis*, 12, 209–228.

Marek POLAŃSKI has received his PhD in Logic and Philosophy of Science from the University of Munich (LMU). He worked at Institute of Computer Science at the Universität der Bundeswehr in Neubiberg and at the Department of Philosophy at the LMU. His main research interests include formal logic, philosophy of language and philosophy of science. He has already published in *Theoria*.

ADDRESS: Philosophie Department, LMU München, Ludwigstrasse 31, D-80539 München Germany.
E-mail: mpolanski@yahoo.de