The axiom of elementary sets on the edge of Peircean expressibility*

Andrea Formisano†, Eugenio G. Omodeo‡, Alberto Policriti§

25 Feb 2005

Abstract

Being able to state the principles which lie deepest in the foundations of mathematics by sentences in three variables is crucially important for a satisfactory equational rendering of set theories along the lines proposed by Alfred Tarski and Steven Givant in their monograph of 1987.

The main achievement of this paper is the proof that the 'kernel' set theory whose postulates are *extensionality*, **(E)**, and single-element *adjunction* and *removal*, **(W)** and **(L)**, cannot be axiomatized by means of three-variable sentences. This highlights a sharp edge to be crossed in order to attain an 'algebraization' of Set Theory. Indeed, one easily shows that the theory which results from the said kernel by addition of the null set axiom, **(N)**, is in its entirety expressible in three variables.

Key words: Weak set theories, n-variable expressibility, pebble games.

Introduction

Among the postulates for set theory proposed by Zermelo in his epochal paper [Zer08], the *axiom of elementary sets* states the existence of a void set and the availability of an unordered pair formation operation. This axiom comes, at the very beginning of the list of postulates, immediately after the *extensionality* axiom stating that distinct sets cannot have precisely the same elements.

Throughout this paper we will consider a first-order logic whose language involves $=$ and $\in$ as the only relators. In this context, considering a first-order

*Work partially supported by MURST/MIUR 40%, *Aggregate- and number-reasoning for computing: From decision algorithms to constraint programming with multisets, sets, and maps.* This research benefited from collaborations fostered by the European COST Action 274 (TARSKI project, see `http://www.tarski.org` or `http://www.ggediga.de/tarski`).

†Università di L'Aquila, Dip. di Informatica; email: `formisano@di.univaq.it`

‡Università di Trieste, Dip. di Matematica e Informatica; email: `eomodeo@univ.trieste.it`

§Università di Udine, Dip. di Matematica e Informatica; email: `policrit@dimi.uniud.it`

set theory which does not cater for individuals or proper classes, we can state extensionality as simply as

$$\textbf{(E)} \qquad \forall x \, \forall y \, \big(\, x \neq y \;\; \rightarrow \;\; \exists v \, (\, v \in x \, \leftrightarrow \, v \notin y \,) \big),$$

and we can decompose Zermelo's postulate of elementary sets as the conjunction of the following *null-set axiom* and *pair axiom*:

$$\textbf{(N)} \qquad \exists z \, \forall v \, v \notin z,$$
$$\textbf{(P)} \qquad \forall x \, \forall y \, \exists p \, \forall v \, \big(\, v \in p \;\; \leftrightarrow \;\; (\, v = x \;\; \vee \;\; v = y \,) \big).$$

Several studies (cf., among others, [Daw99, KV96, Hod93]) indicate the number of distinct variables as a significant measure of complexity for sentences. From this perspective, one may be led to think that **(P)** is somewhat deeper than **(E)**, because it involves 4 variables instead of 3. Alfred Tarski, however, discovered a sentence which is logically equivalent to **(P)**, involves 3 variables altogether, and explicitly states the existence of *ordered pairs* (cf. [Tar53] and [FOP04]).

In this paper we will consider a version of the postulate of elementary sets which is a bit stronger than the conjunction $\textbf{(N)} \wedge \textbf{(P)}$. Besides encompassing **(N)**, our postulate has clauses

$$\textbf{(W)} \qquad \forall x \, \forall y \, \exists w \, \forall v \, \big(\, v \in w \;\; \leftrightarrow \;\; (\, v \in x \;\; \vee \;\; v = y \,) \big),$$
$$\textbf{(L)} \qquad \forall x \, \forall y \, \exists \ell \, \forall v \, \big(\, v \in \ell \;\; \leftrightarrow \;\; (\, v \in x \;\; \wedge \;\; v \neq y \,) \big)$$

catering for the operations $x, y \mapsto x \cup \{y\}$ and $x, y \mapsto x \setminus \{y\}$ of single-element addition and removal. In [FOP04] we recast $\textbf{(N)} \wedge \textbf{(W)} \wedge \textbf{(L)}$ as a 3-variable sentence, taking advantage of the presence of **(E)**. To do this, we exploited a notion of ordered pair which slightly differs from the classical one due to Kuratowski and proceeded in a way similar (but simpler) to the way Tarski restated **(P)**.

The question then naturally arises whether the conjunction $\textbf{(W)} \wedge \textbf{(L)}$ is, by itself, expressible in 3 variables. (This question was raised, regarding **(W)** alone, in [TG87].) To get a negative answer, in this paper we construct a structure $\mathfrak{R}$ satisfying $\textbf{(W)} \wedge \textbf{(L)} \wedge \textbf{(E)}$ and another structure $\mathfrak{P}$ satisfying $\neg\textbf{(W)} \wedge \neg\textbf{(L)} \wedge \textbf{(E)}$; then we show, via a technique based on the so-called *pebble games* (see [Imm82, IK89, EF99, Daw99]), that the sentences in 3 variables that $\mathfrak{R}$ makes true and those which $\mathfrak{P}$ makes true are the same. Since **(E)** involves 3 variables only, no sentence in 3 variables can be logically equivalent to $\textbf{(W)} \wedge \textbf{(L)}$, to **(W)**, or to **(L)**.

This conclusion is, to the best of our knowledge, an original result. As a warm-up towards this main result we exploit pebble games to prove 3-variable inexpressibility of **(W)** taken in isolation from **(L)** and **(E)**. This result is provided in Sec.1.1, as an exercise in the use of pebble games preparatory to the new and more engaging result.

This paper, along with [FOP04], provides an argument in favor of treating the triad **(N)**, **(W)**, **(L)** as a single postulate: The conjunction of these three sentences can, in fact, be stated very tersely by an equivalent sentence which involves 3 variables altogether. Since **(N)** $\wedge$ **(W)** yields **(P)**, something close to Tarski's statement in 3 variables of **(N)** $\wedge$ **(P)** would be achievable for **(N)** $\wedge$ **(W)** as well; but the outcome would be somewhat lengthier and more cryptic than for the said triad.

1 Inexpressibility via games

In Sections 1.1 and 2 we will prove that the axiom **(W)** cannot be expressed, either considered in isolation or taken together with **(L)**, by means of a first-order sentence involving three variables altogether. To show this, we will rely on a model-theoretic method introduced in [Imm82] and already present, *in nuce*, in [Bar77]. This method centers around the so-called *pebble games*, which are two-player processes closely akin to the older Ehrenfeucht-Fraïssé games. We will proceed by

- first singling out two structures which disagree on the truth-values of **(W)** and **(L)** but agree in modeling **(E)**; and then
- exploiting pebble games in order to prove that these structures are indistinguishable by sentences in three variables.

The game to be played involves players, named *Duplicator* and *Spoiler*, owning three pebbles each—because this is the number of variables which we regard as the critical threshold. We must suggest the winning strategy to Duplicator, the player in favor of which our inexpressibility analysis will be inclined.

1.1 Inexpressibility of (W) in 3 variables

Preliminary to the stronger result to be discussed in Sec.2, in what follows we outline a proof that **(W)** is not expressible in 3 variables.[1] To this end, we will exhibit a *'rich'* structure $\mathfrak{R}$ and a *'poor'* structure $\mathfrak{P}$ which model **(W)** and $\neg$**(W)**, respectively. Our proof will amount to showing that these structures satisfy precisely the same sentences involving (at most) 3 variables.

Definition 1.1 *Let* $\mathbb{N} = \{0, 1, 2, \dots\}$, $\mathbb{Z} = \mathbb{N} \cup \{-i \mid i \in \mathbb{N}\}$, *and let* $|X|$ *designate the* cardinality *of any set* X.

A subset A *of* $\mathbb{Z}$ *is said to be* REPRESENTABLE *if it meets the condition*

$$0 \notin A \ \wedge\ |A \setminus \mathbb{N}| < \aleph_0 \ \wedge\ |\mathbb{N} \setminus A| < \aleph_0.$$

For $i, j \in \mathbb{Z}$ *and* $X \subseteq \mathbb{Z}$ *either representable or finite, we specify* interval*,*

[1] This question, initially raised by Tarski, is still regarded as open in [TG87, p.63], in spite of Michael K. Kwatinetz' claim that he had achieved a proof (cf. [Kwa81, pp.55–57]).

left endpoints, radius, *and* footprint *operations as follows:*

$$\begin{aligned} [i,j] &=_{Def} \{\, h \in \mathbb{Z} \mid i \leqslant h \ \wedge\ h \leqslant j \,\}, \\ \mathit{left}(X) &=_{Def} \{\, i \in X \mid i-1 \notin X \,\}, \\ \mathit{rad}(X) &=_{Def} \min\{\, i \in \mathbb{N} \mid \mathit{left}(X) \subseteq [-i,i] \,\}, \\ \mathit{foot}(X) &=_{Def} X \cap [\min(X), \max(\mathit{left}(X))]. \end{aligned}$$

□

By a little reflection, one sees that any representable set A can be uniquely decomposed in the form of a disjoint union

$$A \;=\; \bigcup_{i=1}^{\nu}[n_{2\cdot i-1}, n_{2\cdot i}] \;\cup\; \bigcup_{j=0}^{\pi-1}[p_{2\cdot j}, p_{2\cdot j+1}] \;\cup\; \{\, k \in \mathbb{Z} \mid k \geqslant p_{2\cdot\pi} \,\}\;,$$

of non-void intervals some of which may be singletons, one of which is infinite, and whose (left and right) endpoints form the set

$$\mathit{left}(A) \,\cup\, \{\, i \in A \mid i+1 \notin A \,\} = \{\, n_1, \dots, n_{2\cdot\nu}\} \cup \{\, p_0, \dots, p_{2\cdot\pi} \,\},$$

where the ns are negative integers, the ps are positive integers, and $\nu, \pi \in \mathbb{N}$. The footprint $\mathit{foot}(A)$ of such an A is a set which, despite having finite cardinality, fully characterizes A. Notice that $\mathit{rad}(A) = \mathit{rad}(\mathit{foot}(A))$ holds for any representable set A; moreover, a finite set $X \subseteq \mathbb{Z}$ equals $\mathit{foot}(A)$ for some representable set A if and only if $0 \notin X \ \wedge\ X \cap \mathbb{N} \neq \emptyset \ \wedge\ (\max(X)-1) \notin X$ holds.

Example 1.2 Let $n_1, \dots, n_6, p_0, p_1, p_2$ be integer numbers satisfying

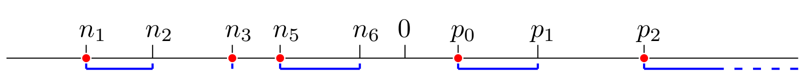
$$n_1 < n_2,\ n_2+1 < n_3 = n_4,\ n_4+1 < n_5 < n_6 < 0 < p_0 < p_1 < p_2 - 1\,,$$

and consider the representable set

$$a = [n_1, n_2] \cup [n_3, n_3] \cup [n_5, n_6] \cup [p_0, p_1] \cup \{k \in \mathbb{N} \mid k \geqslant p_2\}.$$

This set of integers can be represented graphically as follows:

A graphic rendering of a representable set a with its left endpoints • and its footprint



The respective domains of $\mathfrak{R}$ and $\mathfrak{P}$ are defined as follows:

$$\begin{array}{lll} \mathbb{Z}' & =_{\text{Def}} & \mathbb{Z} \setminus \{0\}, \\ \mathfrak{R} & =_{\text{Def}} & \mathbb{Z}' \cup \big\{ A \subseteq \mathbb{Z} \mid A \text{ is representable} \big\}, \\ \mathfrak{P} & =_{\text{Def}} & \mathbb{Z}' \cup \big\{ B \subseteq \mathbb{Z} \mid B \in \mathfrak{R} \ \wedge \ |foot(B)| \text{ is even} \big\}. \end{array}$$

The interpretation $\in^{\mathfrak{R}}$ of the membership relator in $\mathfrak{R}$ is defined to be the relation:

$$a_1 \in^{\mathfrak{R}} a_2 \quad \leftrightarrow_{\text{Def}} \ \big(a_2 \in \mathfrak{R} \setminus \mathbb{Z}' \ \rightarrow \ a_1 \in a_2 \cup (\mathfrak{R} \setminus \mathbb{Z}') \big),$$

with $a_1, a_2 \in \mathfrak{R}$. The interpretation $\in^{\mathfrak{P}}$ of membership in $\mathfrak{P}$ simply is the restriction of $\in^{\mathfrak{R}}$ to $\mathfrak{P}$.

Remark 1.3 Notice that every $a \in \mathfrak{R}$ is either an integer or a representable set, but not both. Moreover, extensionality fails to hold both in $\mathfrak{R}$ and in $\mathfrak{P}$. In fact, by definition, for any $a_1, a_2 \in \mathfrak{R}$, $a_1 \in^{\mathfrak{R}} a_2$ holds when $a_2 \in \mathbb{Z}'$ or a_1 is a representable set (and similarly for the structure $\mathfrak{P}$).

Lemma 1.4 *The structures $\mathfrak{R}$ and $\mathfrak{P}$ meet the following conditions:*

(a) $\mathfrak{R} \models$ **(W)***;*

(b) $\mathfrak{P} \models \neg$**(W)***.*

Proof.

(a) The element-addition operation $\mathsf{with}^{\mathfrak{R}}$ can be defined as follows for any pair a_1, a_2 of operands in $\mathfrak{R}$:

$$a_1 \ \mathsf{with}^{\mathfrak{R}} \ a_2 \quad =_{\text{Def}} \ \begin{cases} a_1 & \text{if } a_2 \in^{\mathfrak{R}} a_1, \\ a_1 \cup \{a_2\} & \text{otherwise.} \end{cases}$$

In fact, when $a_2 \notin^{\mathfrak{R}} a_1$, then a_1 is representable, $a_2 \in \mathbb{Z}'$, and $a_2 \notin a_1$; accordingly, $a_1 \cup \{a_2\}$ is representable, it belongs to $\mathfrak{R}$, and its $\in^{\mathfrak{R}}$-elements in $\mathbb{Z}'$ are precisely its elements.

(b) To show that $\mathfrak{P}$ is not a model of **(W)**, consider any of the sets $B_i = \{-i, -i+1\} \cup \{i, i+1, i+2\} \cup \{i+4, \dots\}$ with $i > 1$. Clearly $B_i \in \mathfrak{P}$ and $B_i \cup \{i+3\} \notin \mathfrak{P}$ because the respective cardinalities of the footprints are 6 and 3. E.g., $foot(B_2) = \{-2, -1, 2, 3, 4, 6\}$ and $foot(B_2 \cup \{5\}) = \{-2, -1, 2\}$. □

Preliminary to our next lemma, showing that Duplicator can win a 3-pebble game of any number m of rounds played over the structures $\mathfrak{R}$ and $\mathfrak{P}$, let us briefly recall how this kind of game proceeds. (For an extensive treatment, cf. [EF99, IK89], among others.)

One assumes that there are 3 different colors and two pebbles p_i^b ($b = 0, 1$) of each color i. Each round consists of two moves and develops as follows: Spoiler moves by choosing an element from one of the two structures and by placing a pebble p_i^b on it—should there be already some pebble(s) on that element, placing one more is allowed. Duplicator responds by placing the corresponding pebble p_i^{1-b} of the same color on some element of the other domain. When Spoiler runs out of pebbles, then he can simply reuse any p_i^b (shifting it from its earlier placement to the new position), in either structure. Then Duplicator must respond by placing p_i^{1-b} on an element of the opposite structure. After n rounds have been played, the colors of the pebbles lying on the structures determine a correspondence between elements of their domains. Duplicator *wins* the game if and only if this relation happens to be a partial isomorphism between $\mathfrak{R}$ and $\mathfrak{P}$. Otherwise Spoiler wins.

We will use notation and basic results on games as introduced in [EF99]. In particular, we will use the result stating that the existence of a winning strategy for Duplicator in an m-round 3-pebbles game played on two structures $\mathcal{A}, \mathcal{B}$ (denoted $G_m^3(\mathcal{A}, \mathcal{B})$) is equivalent to the fact that $\mathcal{A}$ and $\mathcal{B}$ satisfy the same first-order sentences of quantifier rank less than or equal to m.

Even though we will prove our result for arbitrary m, some of the results presented here are in fact generalizable to games with infinitely many moves. Such a generalization would allow the treatment of infinitary languages such as $\mathcal{L}^3_{\infty\omega}$. However, we will not discuss such generalizations, as for our result no infinitary logic is necessary.

Lemma 1.5 Duplicator *has a winning strategy in any m-round 3-pebble game played on the two structures $\mathfrak{R}$, $\mathfrak{P}$.*

Proof. A winning strategy for *Duplicator* is the following:

FIRST MOVE. If *Spoiler* plays his first pebble on $a_1 \in \mathfrak{R}$ (respectively, on $b_1 \in \mathfrak{P}$), then *Duplicator* can respond by choosing any $b_1 \in \mathfrak{P}$ (resp., $a_1 \in \mathfrak{R}$) such that $a_1 \in \mathbb{Z}' \leftrightarrow b_1 \in \mathbb{Z}'$.

SECOND MOVE. Suppose that *Spoiler* plays his second pebble on $a_2 \in \mathfrak{R}$, then *Duplicator* replies by choosing $b_2 \in \mathfrak{P}$ so that $a_2 \in \mathbb{Z}' \leftrightarrow b_2 \in \mathbb{Z}'$. Moreover,

- If $a_{1+\ell} \in \mathbb{Z}'$ and $a_{2-\ell} \in \mathfrak{R} \setminus \mathbb{Z}'$ (for $\ell = 0$ or $\ell = 1$), then *Duplicator* chooses b_2 so that $a_{1+\ell} \in^{\mathfrak{R}} a_{2-\ell} \leftrightarrow b_{1+\ell} \in^{\mathfrak{P}} b_{2-\ell}$.
- If $a_1, a_2 \in \mathfrak{R} \setminus \mathbb{Z}'$, then consider the Venn diagram generated by a_1 and a_2. There are exactly two regions of this diagram that have finite cardinality. Let them be X_1, X_2 (with $X_i \subsetneq a_i$ for $i = 1, 2$). Consider the two regions of the Venn diagram generated by b_1 and b_2, say Y_1, Y_2 (with $Y_i \subsetneq b_i$ for $i = 1, 2$). The choice of *Duplicator* must satisfy the condition: $X_i = \emptyset \leftrightarrow Y_i = \emptyset$ for both $i = 1, 2$. Moreover, for any $n \geqslant 0$ and for $i = 1, 2$ it must hold that: there exist $c_1, c_2, \dots, c_n \in \mathfrak{R}$ such that $a_1 \cap a_2 \subsetneq c_1 \subsetneq c_2 \subsetneq \cdots \subsetneq c_n \subsetneq a_i$ if and only if there exist $d_1, d_2, \dots, d_n \in \mathfrak{P}$ such that $b_1 \cap b_2 \subsetneq d_1 \subsetneq d_2 \subsetneq \cdots \subsetneq d_n \subsetneq b_i$.

If *Spoiler* plays his second pebble on $b_2 \in \mathfrak{P}$, then the strategy by which *Duplicator* plays his second pebble on $a_2 \in \mathfrak{R}$ is completely analogous.

THIRD MOVE. Assume that two pebbles have been played on $a_1, a_2 \in \mathfrak{R}$, and the other two have been played on $b_1, b_2 \in \mathfrak{P}$. As before, if *Spoiler* plays his third pebble on $a_3 \in \mathfrak{R}$, then *Duplicator* responds by choosing any $b_3 \in \mathfrak{P}$ satisfying $a_3 \in \mathbb{Z}' \leftrightarrow b_3 \in \mathbb{Z}'$ as well as the condition specified below. (If *Spoiler* plays in $\mathfrak{P}$, then the strategy for *Duplicator* is similar.)

The additional condition that b_3 must satisfy, depending on the objects chosen during the previous moves, is:

- If $a_1, a_2 \in \mathbb{Z}'$ and $a_3 \in \mathfrak{R} \setminus \mathbb{Z}'$, then the following must hold: $\forall i \in \{1,2\}$ $(a_i \in^{\mathfrak{R}} a_3 \leftrightarrow b_i \in^{\mathfrak{P}} b_3)$.
- If $a_3 \in \mathbb{Z}'$ and $a_{2-\ell} \in \mathfrak{R} \setminus \mathbb{Z}'$ (for $\ell = 0$ or $\ell = 1$), then b_3 must satisfy the following condition: $a_3 \in^{\mathfrak{R}} a_{2-\ell} \leftrightarrow b_3 \in^{\mathfrak{P}} b_{2-\ell}$.
- If $a_{2-\ell}, a_3 \in \mathfrak{R} \setminus \mathbb{Z}'$ (for $\ell = 0$ or $\ell = 1$), then b_3 must satisfy the following conditions: $a_{1+\ell} \in^{\mathfrak{R}} a_3 \leftrightarrow b_{1+\ell} \in^{\mathfrak{P}} b_3$; moreover, if $X_{2-\ell}$ and X_3 (resp. $Y_{2-\ell}$ and Y_3) are the regions of the Venn diagram generated by $a_{2-\ell}$ and a_3 (resp. by $b_{2-\ell}$ and b_3), then $X_i = \emptyset \leftrightarrow Y_i = \emptyset$ must hold for $i = 2-\ell, 3$. Moreover, for any $n \geqslant 0$ and for $i = 2-\ell, 3$ it must hold that: there exist $c_1, c_2, \ldots, c_n \in \mathfrak{R}$ such that $a_{2-\ell} \cap a_3 \subsetneq c_1 \subsetneq c_2 \subsetneq \cdots \subsetneq c_n \subsetneq a_i$ if and only if there exist $d_1, d_2, \ldots, d_n \in \mathfrak{P}$ such that $b_{2-\ell} \cap b_3 \subsetneq d_1 \subsetneq d_2 \subsetneq \cdots \subsetneq d_n \subsetneq b_i$.

For each one of the cases listed above, it is easy to verify that *Duplicator* can always choose some b_3 satisfying the associated constraints.

The above-outlined strategy by which *Duplicator* can respond to the third move can also be exploited to respond to every subsequent move of *Spoiler*. In fact, each subsequent round consists in removing a pair of corresponding pebbles and in playing them again on the two structures. The situation after the removal of the pair of corresponding pebbles is essentially analogous to the one holding at the end of the second round of the game. □

Proposition 1.6 *The sentence* **(W)** *cannot be expressed in 3 variables.*

Proof. By Lemma 1.5 and by virtue of the main result on pebble games (cf. [Imm82, Thm.C.1] or [EF99, Thm.3.3.5]), it follows that the two structures $\mathfrak{R}$ and $\mathfrak{P}$ defined above model the same sentences in (at most) 3 variables. Since such structures disagree on the truth value of **(W)**, there is no sentence in 3 variables which is logically equivalent to **(W)**. □

2 Inexpressibility of $(\mathbf{E}) \wedge (\mathbf{W}) \wedge (\mathbf{L})$ in 3 variables

In what follows, we will show that the conjunction $(\mathbf{E}) \wedge (\mathbf{W}) \wedge (\mathbf{L})$ cannot be stated by means of a 3-variable sentence. In analogy with the way we have treated **(W)** alone, we will design two structures, $\mathfrak{R} = \left(\mathfrak{R}, \in^{\mathfrak{R}}\right)$ and $\mathfrak{P} = \left(\mathfrak{P}, \in^{\mathfrak{P}}\right)$, such that **(W)** and **(L)** are true in $\mathfrak{R}$ and false in $\mathfrak{P}$. Both structures will satisfy **(E)**. We will see that the difficult moves for Duplicator will be the ones in which Spoiler chooses in $\mathfrak{R}$ and Duplicator must respond in $\mathfrak{P}$. To devise the strategy in those cases, we will perform a *'spiral'* construction of an embedding of $\mathfrak{R}$ into $\mathfrak{P}$.

Representable subsets of $\mathbb{Z}$, together with all pertaining notation (*rad*, *foot*, etc.), will again enter into this spiral construction. We start by putting

$$\begin{array}{lll} \mathfrak{R} & =_{\text{Def}} & \left\{ A \subseteq \mathbb{Z} \mid A \text{ is representable} \right\}, \\ \mathfrak{P} & =_{\text{Def}} & \left\{ X \in \mathfrak{R} \mid |X \cap [1, \mathit{rad}(X)]| \geqslant |\mathit{left}(X \setminus \mathbb{N})| \right\}, \end{array}$$

and, preliminary to defining the interpretations $\in^{\mathfrak{R}}, \in^{\mathfrak{P}}$ of membership in these structures, we show that a dyadic relation $\prec$ on $\mathfrak{R}$ exists such that

0) $\prec$ is a total ordering, and

the following three conditions are met by all $X, Y \in \mathfrak{R}$:

1) $\mathit{rad}(X) < \mathit{rad}(Y)$ implies $X \prec Y$;

2) there are infinitely many $\ell \in \mathbb{N}$ such that the representable set V with

$$\mathit{foot}(V) = \{-\ell\} \cup \mathit{foot}(X)$$

belongs to $\mathfrak{P}$ and is the smallest of all representable sets W with $\mathit{rad}(W) = \ell$;

3) if X is the smallest of all representable sets W satisfying $\mathit{rad}(W) = \mathit{rad}(X)$, then $-\mathit{rad}(X) \in X \in \mathfrak{P}$.

We will make substantial use of $\prec$ in specifying a winning strategy for Duplicator in what follows, since in order to model **(E)** we must work with a more limited endowment of distinguishable elements than those available in Sec.1.1. The homologous strategy developed while dealing with **(W)** alone was, indeed, comparatively easy thanks to the ample variety of elements (and relationships among them) available in the structures $\mathfrak{R}$ and $\mathfrak{P}$.

We can, for example, put

$$X \prec_g Y =_{\text{Def}} g(\mathit{foot}(X)) < g(\mathit{foot}(Y))$$

for all $X, Y \in \mathfrak{R}$, and then $\prec \equiv \prec_g$, after carrying out the construction of a suitable indexing bijection

$$g : \{ X \subseteq \mathbb{Z} \mid |X| < \aleph_0 \} \longrightarrow \mathbb{N}$$

as specified by the following algorithm:

$g := \emptyset$; $g(\emptyset) := 0$; $g(\{0\}) := 1$;
for $\ell := 1, 2, 3, \ldots$ *(ad inf.)* **loop** --*index all sets v with $rad(v) = \ell$*
  let $x \subseteq \mathbb{Z}$ and $r \in \mathbb{N}$ **be such that** $2^r \cdot (2 \cdot g(x) + 1) - 1 = \ell$;
  $m := \max(x \cup \{1\})$;
  $w := \{-\ell, m\} \cup (x \setminus \{0, m-1\})$; --*convert x into a footprint*
  if $\left|foot^{-1}(w) \cap [1, \ell]\right| < |left(w \setminus \mathbb{N})|$ **then** $w := \{-\ell, \ell\}$; **end if**;
  $g(w) := |g|$; -- *this assigns the value* $2^{2\cdot\ell-1}$
  for $v \subseteq [-\ell, \ell]$ **such that** $v \notin \mathsf{domain}(g)$ **loop**
    $g(v) := |g|$;
  end loop;
end loop.

Clearly, g induces an enumeration $a_0, a_1, a_2, \ldots$ of all representable sets via their footprints; that is, the relation $\prec_g$ is isomorphic to the standard well-order of $\mathbb{N}$.

Concerning this particular construction of g, the following conditions are satisfied for all finite subsets X, Y of $\mathbb{Z}$:

- $rad(X) < rad(Y)$ implies $g(X) < g(Y)$;
- if X if the footprint of some representable set, then
$$g\Big(\{-\ell\} \cup X\Big) \;=\; 2^{2\cdot\ell-1}\,,$$
holds for infinitely many $\ell \in \mathbb{N}$ (observe that all but a finite number of the sets $\{-\ell\} \cup X$ are footprints of representable sets in $\mathfrak{P}$);
- if $g(X) = 2^{2\cdot h-1}$ for some positive integer h, then:
$-h \in X$, $0 \notin X$, and $rad(X) = h$; moreover, $foot^{-1}(X) \in \mathfrak{P}$ (incidentally note that $foot^{-1}(X) \cap [1, h] = X \cup \{\max(X) + 1, \ldots, h\}$).

Consequently, as is easy to verify, $\prec_g$ satisfies the above conditions 1)–3).

Example 2.1 Taking into account that the overall number of finite sets $X \subseteq \mathbb{Z}$ with $rad(X) < \ell$ is $2^{2\cdot\ell-1}$, and that $r = 1, 0, 2, 0, 1$ and $n = 0, 1, 0, 2, 1$ are the values satisfying the equation $\ell = 2^r \cdot (2 \cdot n + 1) - 1$ for $\ell = 1, 2, 3, 4, 5$, respectively, one straightforwardly figures out that $g(\{-1, 1\}) = 2$, $g(\{-2, 1\}) = 8$, $g(\{-3, 1\}) = 32$, $g(\{-4, -1, 1\}) = 128$, $g(\{-5, 1\}) = 512$. Any of the sets $\{-1\}$, $\{-1, 0\}$, $\{1\}$, $\{0, 1\}$, and $\{-1, 0, 1\}$ could be taken as $g^{-1}(3)$; depending on whether or not $-1 \in g^{-1}(3)$, it will turn out either that $g^{-1}(2048) = \{-6, -1, 1\}$ or that $g^{-1}(2048) = \{-6, 1\}$. □

Next, in terms of $\prec$, we define two bijective functions

$$\mathfrak{I}_{\mathfrak{R}} : \mathfrak{R} \longrightarrow \mathbb{Z} \setminus \{0\}, \qquad\qquad \mathfrak{I}_{\mathfrak{P}} : \mathfrak{P} \longrightarrow \mathbb{Z} \setminus \{0\},$$

which associate integer *indices* with representable sets. Here are $\mathfrak{I}_{\mathfrak{R}}$ and $\mathfrak{I}_{\mathfrak{P}}$:

$$\begin{array}{lcl} \mathfrak{I}_{\mathfrak{R}}(X) & =_{\text{Def}} & \min\Big(X \setminus \big\{ \mathfrak{I}_{\mathfrak{R}}(Y) \mid Y \in \mathfrak{R} \;\wedge\; Y \prec X \big\} \Big); \\ \mathfrak{I}_{\mathfrak{P}}(X) & =_{\text{Def}} & \min\Big(X \setminus \big\{ \mathfrak{I}_{\mathfrak{P}}(Y) \mid Y \in \mathfrak{P} \;\wedge\; Y \prec X \big\} \Big). \end{array}$$

These definitions clearly make sense and ensure the injectivity of $\mathfrak{I}_{\mathfrak{R}}$ and $\mathfrak{I}_{\mathfrak{P}}$. In fact, for each X (processing all representable sets according to the ordering $\prec$), we are choosing as index $\mathfrak{I}_{\mathfrak{R}}(X)$ the least number j in X which has not been chosen as index $\mathfrak{I}_{\mathfrak{R}}(Y)$ for any $Y \prec X$—and analogously with $\mathfrak{I}_{\mathfrak{P}}$.

Membership is then interpreted in terms of $\mathfrak{I}_{\mathfrak{R}}$ in $\mathfrak{R}$, and in terms of $\mathfrak{I}_{\mathfrak{P}}$ in $\mathfrak{P}$:

$$X \in^{\mathfrak{R}} Y \text{ iff } \mathfrak{I}_{\mathfrak{R}}(X) \in Y\,, \qquad V \in^{\mathfrak{P}} Y \text{ iff } \mathfrak{I}_{\mathfrak{P}}(V) \in Y,$$

where $X \in \mathfrak{R}$, $V \in \mathfrak{P}$, and $Y \subseteq \mathbb{Z}$. The verification that both $\mathfrak{R}$ and $\mathfrak{P}$ satisfy **(E)** while **(W)** and **(L)** are true in $\mathfrak{R}$ and false in $\mathfrak{P}$ are left to the reader. Notice also that $X \in^{\mathfrak{R}} X$ and $V \in^{\mathfrak{P}} V$ hold for all $X \in \mathfrak{R}$ and $V \in \mathfrak{P}$.

As a convenient typographical convention, from now on the letters "a" and "b", with or without subscripts or superscripts, will represent a generic set in $\mathfrak{R}$ and in $\mathfrak{P}$, respectively.

Example 2.2 On the basis of the above construction of $\prec_g$, we can tabulate the beginning of $\mathfrak{I}_{\mathfrak{R}}$ and $\mathfrak{I}_{\mathfrak{P}}$ as follows, representing sets a in $\mathfrak{R}$ by their footprints $foot(a)$ and assuming specific values for the function g constructed earlier:

$foot(a)$	$rad(a)$	$g(a)$	$\mathfrak{I}_{\mathfrak{R}}(a)$	$\mathfrak{I}_{\mathfrak{P}}(a)$
$\{-1,1\}$	**1**	**2**	**−1**	**−1**
$\{1\}$	**1**	4	**1**	**1**
$\{-2,1\}$	**2**	**8**	**−2**	**−2**
$\{-2,2\}$	**2**	10	2	2
$\{2\}$	**2**	11	3	3
$\{-2,-1,2\}$	**2**	13	4	4
$\{-2,-1,1\}$	**2**	14	5	5
$\{-1,2\}$	**2**	20	6	6
$\{-3,1\}$	**3**	**32**	**−3**	**−3**
$\{-3,3\}$	**3**	33	7	7
$\{-3,2\}$	**3**	35	8	8
$\{3\}$	**3**	40	9	9
$\{-3,-1,3\}$	**3**	46	10	−−
$\{-3,-2,1,3\}$	**3**	48	11	10
⋮	⋮	⋮	⋮	⋮
$\{-4,-1,1\}$	**4**	**128**	**−4**	**−4**
⋮	⋮	⋮	⋮	⋮
$\{-5,1\}$	**5**	**512**	**−5**	**−5**
⋮	⋮	⋮	⋮	⋮
$\{-6,-1,1\}$	**6**	2048	−6	−6
⋮	⋮	⋮	⋮	⋮

(Values that are printed in boldface cannot be affected by the choices made during the construction of g. The mark "−−" indicates a missing value.)

Observe that the symmetric set-difference

$$\{a \in \mathfrak{R} \mid \mathfrak{I}_{\mathfrak{R}}(a) < 0\} \bigtriangleup \{b \in \mathfrak{P} \mid \mathfrak{I}_{\mathfrak{P}}(b) < 0\}$$

equals $\emptyset$; moreover, $\mathfrak{I}_{\mathfrak{R}}(b) = \mathfrak{I}_{\mathfrak{P}}(b)$ holds for any $b \in \{a \in \mathfrak{R} \mid \mathfrak{I}_{\mathfrak{R}}(a) < 0\}$. □

Let us now prove a few simple, yet useful, facts. Our next proposition shows, among other things, that $\mathfrak{I}_{\mathfrak{R}}$ and $\mathfrak{I}_{\mathfrak{P}}$ are surjective on $\mathbb{Z} \setminus \{0\}$ (as we have announced before):

Lemma 2.3 *For all $\ell \in \mathbb{N} \setminus \{0, 1, 2\}$, there are representable sets a, b, b' with $b, b' \in \mathfrak{P}$ such that*

$$\begin{array}{lcl} \mathfrak{I}_{\mathfrak{R}}(b') = \mathfrak{I}_{\mathfrak{P}}(b') = -\ell\,, & & \mathfrak{I}_{\mathfrak{R}}(a) = \mathfrak{I}_{\mathfrak{P}}(b) = \ell\,, \\ rad(b') = \ell\,, & \text{and} & rad(a) = rad(b) < \ell\,. \end{array}$$

Similar statements, with $rad(a) = rad(b) = \ell$, hold when $\ell = 2, 1$.

Proof. Let $\ell = 1$. The only footprints of representable sets included in $[-1, 1]$ are $\{1\}$ and $\{-1, 1\}$; the corresponding representable sets both belong to $\mathfrak{P}$. Moreover, by the properties of $\prec$, we conclude putting $b' = foot^{-1}(\{-1, 1\})$ and $a = b = foot^{-1}(\{1\})$.

If $\ell = 2$, the footprints of sets having radius ℓ are $\{-2, 2\}$, $\{-2, -1, 2\}$, $\{-2, 1\}$, $\{-2, -1, 1\}$, $\{-1, 2\}$, and $\{2\}$, all corresponding to sets in $\mathfrak{P}$ which contain 2, as well as 3, as elements. By the properties of $\prec$, one of these footprints containing -2 must correspond to the smallest set b' satisfying $rad(b') = 2$, so that clearly $\mathfrak{I}_{\mathfrak{P}}(b') = \mathfrak{I}_{\mathfrak{R}}(b') = -2$. Of course, if we let $a = b$ be the $\prec$-successor of b', $foot(b)$ will be one of the footprints just listed and will satisfy $\mathfrak{I}_{\mathfrak{P}}(b) = \mathfrak{I}_{\mathfrak{R}}(b) = 2$. In its turn the $\prec$-successor c of b, will have another of these footprints and will satisfy $\mathfrak{I}_{\mathfrak{P}}(c) = \mathfrak{I}_{\mathfrak{R}}(c) = 3$.

For $\ell > 2$, recall that the finite set $X \subseteq \mathbb{Z}$ with $rad(X) = \ell$ which comes first relative to $\prec$ satisfies $-\ell \in X$ and is the footprint of a representable set; accordingly, since $-\ell \notin Y$ for any finite set $Y \subseteq \mathbb{Z}$ with $rad(Y) < \ell$, we have $\mathfrak{I}_{\mathfrak{R}}(b') = \mathfrak{I}_{\mathfrak{P}}(b') = -\ell$ if $b' = foot^{-1}(X)$. Notice then that $\{\ell\} = foot(b_1)$ holds for $b_1 = \{\ell, \ell+1, \ldots\} \in \mathfrak{P}$, and that $rad(b_1) = \ell$. It follows from the assumption that $\mathfrak{I}_{\mathfrak{P}}(b) = \ell$ holds for some b with $rad(b) < \ell$ that $\mathfrak{I}_{\mathfrak{P}}(b_1) = \ell + 1$ must hold unless already $\mathfrak{I}_{\mathfrak{P}}(b_0) = \ell+1$ for some $b_0 \prec b_1$ (in which case $rad(b_0) \leqslant rad(b_1)$). Analogous remarks can be made concerning the structure $\mathfrak{R}$; then, to get the desired conclusion, it suffices to observe that indexing "runs faster" in $\mathfrak{R}$ because the sets in $\mathfrak{P}$ bounded by radius ℓ form a subset of those in $\mathfrak{R}$ bounded by the same radius. □

Corollary 2.4 *For each set $a \in \mathfrak{R}$, either $\mathfrak{I}_{\mathfrak{R}}(a) = -rad(a)$ or $\mathfrak{I}_{\mathfrak{R}}(a) \geqslant rad(a)$ holds; in the latter case $\mathfrak{I}_{\mathfrak{R}}(a) > rad(a)$ holds when $rad(a) > 2$. The situation with $\mathfrak{I}_{\mathfrak{P}}(b)$, $b \in \mathfrak{P}$, is entirely analogous.* □

Lemma 2.5 *Let $a, a' \in \mathfrak{R}$ and $b, b' \in \mathfrak{P}$. Then:*

i. if $a' \prec a$ *and* $rad(a') = rad(a)$*, then* $\mathfrak{I}_{\mathfrak{R}}(a) \geqslant rad(a') > 0$
(note that there are only two specific as for which $\mathfrak{I}_{\mathfrak{R}}(a) = rad(a)$ *holds);*

i′. if $b' \prec b$*,* $rad(b') = rad(b)$*, then* $\mathfrak{I}_{\mathfrak{P}}(b) \geqslant rad(b') > 0$
(note that there are only two specific bs for which $\mathfrak{I}_{\mathfrak{P}}(b) = rad(b)$ *holds);*

ii. $\mathfrak{I}_{\mathfrak{R}}(a') \geqslant rad(a)$ *implies* $a' \in^{\mathfrak{R}} a$*;* $\quad \mathfrak{I}_{\mathfrak{R}}(a') < -rad(a)$ *implies* $a' \notin^{\mathfrak{R}} a$*;*

ii′. $\mathfrak{I}_{\mathfrak{P}}(b') \geqslant rad(b)$ *implies* $b' \in^{\mathfrak{P}} b$*;* $\quad \mathfrak{I}_{\mathfrak{P}}(b') < -rad(b)$ *implies* $b' \notin^{\mathfrak{P}} b$*.*

Proof. The assertion i readily follows from Corollary 2.4, by virtue of the condition 3) imposed on $\prec$. That condition ensures in fact that the set X with $rad(X) = rad(a)$ which comes first relative to $\prec$ satisfies the equation $X = foot(a'')$ for some representable set a'' with $-rad(a) = \min(a'')$, so that $\mathfrak{I}_{\mathfrak{R}}(a'') = -rad(a)$ and $a \neq a''$ because $a'' \preccurlyeq a'$. We also have $a'' \in \mathfrak{P}$, whence the analogue i' of i follows.

The assertions ii and ii' readily follow from the remark that $\{rad(a), rad(a)+1, \dots\} \subseteq a$ and $\{\dots, -rad(a)-2, -rad(a)-1\} \cap a = \emptyset$ for every representable set a. □

The following notion will play a crucial role in the proof that $\mathfrak{R}$ and $\mathfrak{P}$ are indistinguishable by means of a 3-pebble game:

Definition 2.6 *An* EMBEDDING *of* $\mathfrak{R}$ *into* $\mathfrak{P}$ *is an injective function* $\kappa : \mathfrak{R} \longrightarrow \mathfrak{P}$ *which meets the conditions*

$$\begin{aligned} a' \in^{\mathfrak{R}} a'' \quad &\leftrightarrow \quad \kappa(a') \in^{\mathfrak{P}} \kappa(a'')\,, \\ b \in^{\mathfrak{P}} \kappa(a') \;\wedge\; b \notin^{\mathfrak{P}} \kappa(a'') \quad &\rightarrow \quad \exists\, a \in \mathfrak{R} \;\; b = \kappa(a)\,, \end{aligned}$$

for all $a', a'' \in \mathfrak{R}$ *and* $b \in \mathfrak{P}$. □

The reader may benefit from the following, equivalent, way of expressing the requirements in the definition of embedding in terms of symmetric set-difference:

$$\begin{aligned} \kappa(a'') \supseteq \{\mathfrak{I}_{\mathfrak{P}}(\kappa(a')) \mid a' \in^{\mathfrak{R}} a''\} \quad \wedge \quad & \kappa(a'') \cap \{\mathfrak{I}_{\mathfrak{P}}(\kappa(a')) \mid a' \notin^{\mathfrak{R}} a''\} = \emptyset, \\ \kappa(a') \triangle \kappa(a'') \quad = \quad & \{\mathfrak{I}_{\mathfrak{P}}(\kappa(a)) \mid a \in^{\mathfrak{R}} a' \triangle a''\}, \end{aligned}$$

where $a', a'' \in \mathfrak{R}$.

We will see below that an embedding of $\mathfrak{R}$ into $\mathfrak{P}$ exists. Before we proceed to verifying this, we must prove a very useful technical lemma:

Lemma 2.7 *The well-order* $\prec$ *satisfies the following implication for all* $a', a'' \in \mathfrak{R}$*:*

$$a'' \prec a' \quad \rightarrow \quad \big(\mathfrak{I}_{\mathfrak{R}}(a') > 0 \;\leftrightarrow\; a' \in^{\mathfrak{R}} a''\big);$$

i.e., a representable set a' *either belongs to every one of its predecessors or to none, as dictated by the sign of* $\mathfrak{I}_{\mathfrak{R}}(a')$*.*

Proof. Assume $a'' \prec a'$, so that $rad(a'') = rad(foot(a'')) \leqslant rad(foot(a')) = rad(a')$. In case $rad(a'') < rad(a')$: if $\mathfrak{I}_{\mathfrak{R}}(a') < 0$, then $\mathfrak{I}_{\mathfrak{R}}(a') = -rad(a') < -rad(a'')$ by Corollary 2.4 and hence $a' \notin^{\mathfrak{R}} a''$ by Lemma 2.5(ii); if $\mathfrak{I}_{\mathfrak{R}}(a') > 0$, then $\mathfrak{I}_{\mathfrak{R}}(a') \geqslant rad(a') > rad(a'')$ and hence $a' \in^{\mathfrak{R}} a''$ again by Corollary 2.4 and Lemma 2.5(ii). In case $rad(a') = rad(a'')$, since $a'' \prec a'$ we obtain $\mathfrak{I}_{\mathfrak{R}}(a') \geqslant rad(a'') > 0$ and $a' \in^{\mathfrak{R}} a''$ from Lemma 2.5(i,ii). □

Lemma 2.8 *Let $a, a' \in \mathfrak{R}$. If $\mathfrak{I}_{\mathfrak{R}}(a') \in [-rad(a), rad(a)]$, then it holds that $[-rad(a'), rad(a')] \subseteq [-rad(a), rad(a)]$. Consequently, one of the following three situations will take place: $a' \prec a$; $\mathfrak{I}_{\mathfrak{R}}(a') = rad(a') = rad(a)$; or $a = a'$ and $\mathfrak{I}_{\mathfrak{R}}(a') = -rad(a')$.*

Proof. Assume $\mathfrak{I}_{\mathfrak{R}}(a') \in [-rad(a), rad(a)]$. Since, by Corollary 2.4, either $\mathfrak{I}_{\mathfrak{R}}(a') = -rad(a')$ or $\mathfrak{I}_{\mathfrak{R}}(a') \geqslant rad(a')$ holds, we get $[-rad(a'), rad(a')] \subseteq [-rad(a), rad(a)]$.

If the above inclusion is strict, then $a' \prec a$ follows directly from the condition 1) to which $\prec$ is subject by definition. Otherwise $rad(a') = rad(a)$; accordingly, if $\mathfrak{I}_{\mathfrak{R}}(a') < 0$ then $a' \preccurlyeq a$, and if $\mathfrak{I}_{\mathfrak{R}}(a') > 0$ then $rad(a') \in \{1, 2\}$: in particular, if $\mathfrak{I}_{\mathfrak{R}}(a') < 0$ and $a = a'$ then $\mathfrak{I}_{\mathfrak{R}}(a') = -rad(a')$, and if $\mathfrak{I}_{\mathfrak{R}}(a') > 0$ then $\mathfrak{I}_{\mathfrak{R}}(a') = rad(a')$. □

Corollary 2.9 *Let $a, a'a'' \in \mathfrak{R}$. If $a'' \prec a$ and $a' \in^{\mathfrak{R}} a'' \triangle a$, then either $a' \prec a$, or $a' = a$ and $\mathfrak{I}_{\mathfrak{R}}(a) = -rad(a)$.*

Proof. Assume that $a'' \prec a$ and $a' \in^{\mathfrak{R}} a'' \triangle a$. It easily follows from $a'' \prec a$ that $a'' \triangle a \subseteq [-rad(a), rad(a)] \setminus \{rad(a)\}$. The desired conclusion then follows by Lemma 2.8. □

The existence of an embedding from $\mathfrak{R}$ into $\mathfrak{P}$ will be ensured (see Lemma 2.10 below) by an algorithm conforming to the generic scheme shown here, which refers to the enumeration $a_0, a_1, \ldots$ of $\mathfrak{R}$ naturally associated with $\prec$:

$\kappa := \emptyset$; $\kappa(a_0) := a_0$; -- *initialization of embedding*
for i := 1, 2, ... *(ad inf.)* **loop**
  let a_i **be** the next element of $\mathfrak{R}$ relative to $\prec$;
  pick b_i in $\mathfrak{P}$ **so that**, for all $j < i$, $\kappa(a_j) \neq b_i$ and the following conditions are met:
    i. $a_j \in^{\mathfrak{R}} a_i \;\leftrightarrow\; \kappa(a_j) \in^{\mathfrak{P}} b_i$,
    ii. $\mathfrak{I}_{\mathfrak{R}}(a_i) > 0 \;\leftrightarrow\; b_i \in^{\mathfrak{P}} \kappa(a_j)$,
    iii. $\kappa(a_j) \triangle b_i \;=\; \{\, \mathfrak{I}_{\mathfrak{P}}(\kappa(a')) \mid a' \neq a_i \wedge (a' \in^{\mathfrak{R}} a_j \;\leftrightarrow\; a' \notin^{\mathfrak{R}} a_i)\,\} \cup (\textbf{if } a_i \notin^{\mathfrak{R}} a_j \textbf{ then } \{\mathfrak{I}_{\mathfrak{R}}(b_i)\} \textbf{ else } \emptyset \textbf{ end if})$;
  $\kappa(a_i) := b_i$;
end loop.

The analysis of this procedure is carried out with the following lemma.

Lemma 2.10 *The above endless procedure actually defines a function κ which is an embedding of $\mathfrak{R}$ in $\mathfrak{P}$.*

Proof. The crucial point is to see that at the ith iteration of the for-loop it is possible to determine an element b_i satisfying the conditions. It will, indeed, be sufficient to take, for a suitable $\ell > 0$ ensuring in particular that $\mathfrak{I}_{\mathfrak{P}}(b_i) > 0 \leftrightarrow \mathfrak{I}_{\mathfrak{R}}(a_i) > 0$ and $\mathfrak{I}_{\mathfrak{P}}(b_i) < 0 \rightarrow \mathfrak{I}_{\mathfrak{P}}(b_i) = -\ell$:

$$\begin{aligned} b_i \quad = \quad & \left\{ \mathfrak{I}_{\mathfrak{P}}(\kappa(a_m)) \mid m < i \wedge a_m \in^{\mathfrak{R}} a_i \right\} \\ & \cup \left(\bigcap_{m<i} \kappa(a_m) \setminus \left\{ \mathfrak{I}_{\mathfrak{P}}(\kappa(a_m)) \mid m < i \wedge a_m \notin^{\mathfrak{R}} a_i \right\} \right) \\ & \cup \left(\textbf{if } \ \mathfrak{I}_{\mathfrak{R}}(a_i) < 0 \ \textbf{ then } \ \{-\ell\} \ \textbf{ else } \ \emptyset \ \textbf{ end if}\right). \end{aligned}$$

We claim first that a b_i as above belongs to $\mathfrak{P}$.

In fact, in case $\mathfrak{I}_{\mathfrak{R}}(a_i) < 0$, by taking an ℓ large enough we can ensure that $rad(b_i) = \ell$, and hence that $b_i \in \mathfrak{P}$, because the value $|b_i \cap [1, rad(b_i)]|$ grows with ℓ whereas $|left(b_i) \setminus \mathbb{N}|$ is bounded (it is, in fact, less than or equal to i).

In case $\mathfrak{I}_{\mathfrak{R}}(a_i) > 0$, we reason as follows. Assuming it inductively possible to increase the radius of elements preceding b_i, also the radius of b_i can grow. Hence, again since the value $|b_i \cap [1, rad(b_i)]|$ grows with the radius of b_i whereas $|left(b_i) \setminus \mathbb{N}| \leqslant i$, we have that $b_i \in \mathfrak{P}$.

Our next claim is that b_i (which by simple inspection of its definition is seen to meet condition *i*) can be so chosen as to meet condition *ii*.

In fact, if $\mathfrak{I}_{\mathfrak{R}}(a_i) < 0$, then we can manage to have $\mathfrak{I}_{\mathfrak{P}}(b_i) = -\ell$ while choosing an ℓ larger than the radius of any $\kappa(a_m)$ with $m < i$, thanks to property 2) of $\prec$.

If $\mathfrak{I}_{\mathfrak{R}}(a_i) > 0$ instead, then condition *ii* expresses the fact that $b_i \in^{\mathfrak{P}} \kappa(a_j)$ for all $j < i$. This holds indeed, because if $b_i \notin^{\mathfrak{P}} \kappa(a_j)$ then, since we have defined $b_j = \kappa(a_j)$ as we are defining b_i and since $\mathfrak{I}_{\mathfrak{R}}(a_i) > 0$, we would have that b_i coincides with some $\kappa(a_h)$ for $h < i$; but this would contradict Corollary 2.9.

As for condition *iii*, let $x \in \kappa(a_j)$ and $x \notin b_i$. If x were an element of all $\kappa(a_m)$ with $m < i$ then, by the definition of b_i, we would have $x = \mathfrak{I}_{\mathfrak{P}}(\kappa(a_{j'}))$ with $a_{j'} \notin^{\mathfrak{R}} a_i$; moreover, thanks to the fulfillment of condition *i* at the j-th iteration of the for-loop, $a_{j'} \in^{\mathfrak{R}} a_j$ because $\kappa(b_{j'}) \in^{\mathfrak{P}} \kappa(a_j)$. If x does not belong to all $\kappa(a_m)$ with $m < i$, then let j' be such that $x \notin \kappa(a_{j'})$. Thus $x \in \kappa(a_j) \triangle \kappa(a_{j'})$ which, by Corollary 2.9, guarantees that $x = \kappa(a_{j''})$ with $a_{j''} \in^{\mathfrak{R}} a_j$ and $a_{j''} \notin^{\mathfrak{R}} a_i$. The case when $x \notin \kappa(a_j)$ and $x \in b_i$ is analogous. Finally, $a_i \notin^{\mathfrak{R}} a_j$ if and only if $\mathfrak{I}_{\mathfrak{R}}(\kappa(a_i)) < 0$, and hence if and only if $\mathfrak{I}_{\mathfrak{P}}(b_i) \notin \kappa(a_j)$ and $\mathfrak{I}_{\mathfrak{P}}(b_i) \in b_i$.

At this point, in order to verify that κ is an embedding, let us assume first that $a' \in^{\mathfrak{R}} a$ with $a' \neq a$. If $a' \prec a$, then $\kappa(a') \in^{\mathfrak{P}} \kappa(a)$ by i. Otherwise $a \prec a'$, and therefore $\mathfrak{I}_{\mathfrak{R}}(a') > 0$ by Lemma 2.7, since $a' \in^{\mathfrak{R}} a$. From this, by ii, it follows that $\kappa(a') \in^{\mathfrak{P}} \kappa(a)$.

From the above argument we can conclude that $\kappa(a) \supseteq \{\kappa(a') \mid a' \in^{\mathfrak{R}} a\}$.

The case in which $a' \notin^{\mathfrak{R}} a$ is entirely symmetric and we have that $\kappa(a) \cap \{\kappa(a') \mid a' \notin^{\mathfrak{R}} a\} = \emptyset$.

The last property characterizing the notion of embedding follows directly from iii. □

Remark 2.11 Notice that one could initialize κ at the beginning of the embedding procedure as *any* finite injective partial function from $\mathfrak{R}$ to $\mathfrak{P}$. Then, replacing the **let**-statement by:

let a_i **be** the next element of $\mathfrak{R}$ relative to $\prec$ whose $\kappa(a_i)$ is still undefined;

the conclusion of Lemma 2.10 continues to hold as long as the indices of the $\kappa(a)$s in the range of the initializing partial function are chosen *cum grano salis*: their value must guarantee that the b_is to be defined later will belong to $\mathfrak{P}$. In particular, this will be the case whenever the initializing partial function is an initial segment of a given embedding. □

In order to prove the impossibility to distinguish the 'rich' structure $\mathfrak{R}$ from the 'poor' structure $\mathfrak{P}$ by means of sentences involving only three variables, we will again use pebble games.

Before doing so, we need to introduce some definitions characterizing a sort of *partial* embedding that will be used when we must update the embedding suggesting the strategy to *Duplicator*.

Definition 2.12 *We will call* Δ-CLOSURE *of a set* $X \subseteq \mathfrak{R}$ *the minimum fixed point defined as follows:*[2] $\Delta X =_{Def} (\mu Y \supseteq X)(a, b \in Y \rightarrow a \triangle b \subseteq Y)$. □

Notice that the Δ-closure is defined as being a fixpoint of the monotone non-decreasing function $X \mapsto X \cup \bigcup\{a \triangle b \,:\, a, b \in X\}$ $(= X \cup \{c \in \mathfrak{R} \mid (\exists\, a, b \in X)\ (c \in^{\mathfrak{R}} a \triangle b)\})$. Let $\prec X$ denote the set $\{y \mid (\exists x \in X)(y \preccurlyeq x)\}$. By Corollary 2.9, we have that $\Delta X \subseteq \prec X$; hence ΔX is finite when X is a finite set, because obviously $\prec X$ is finite in this case.

In the ongoing, $\vec{a} = a_1, \ldots, a_i$ and $\vec{b} = b_1, \ldots, b_i$, where as before a_i's and b_j's designate elements of $\mathfrak{R}$ and of $\mathfrak{P}$, respectively.

Definition 2.13 *We say that* $\Delta\{\vec{a}\}$ *and* $\Delta\{\vec{b}\}$ *are* ISOMORPHIC, $\Delta\{\vec{a}\} \simeq \Delta\{\vec{b}\}$, *if there exists an* $\in$*-isomorphism from the former into the latter sending* a_j *to* b_j *for* $j \in \{1, \ldots, i\}$. □

An isomorphism between Δ-closures is a sort of partial embedding. The following lemma proves the possibility of extending partial embeddings.

[2] We will sometimes abuse notation, as here, by applying certain relations (e.g. $\supseteq$ or $=$) or operations (e.g., $\triangle, \cup$, or $\cap$) to sets whose elements should be 'de-referenced' as common-sense will suggest, by applying either $\mathfrak{I}_{\mathfrak{P}}$ or $\mathfrak{I}_{\mathfrak{R}}$.

Lemma 2.14 *If $\Delta\{\vec{a}\} \simeq \Delta\{\vec{b}\}$, then $(\forall b \notin \Delta\{\vec{b}\})(\exists a)\big(\Delta\{\vec{a}, a\} \simeq \Delta\{\vec{b}, b\}\big)$. Moreover, if $\mathfrak{I}_{\mathfrak{P}}(b) > 0$ then $\mathfrak{I}_{\mathfrak{R}}(a) > 0$, and if $\mathfrak{I}_{\mathfrak{P}}(b) < 0$, then $\mathfrak{I}_{\mathfrak{R}}(a) < 0$ and infinitely many such a are available.*

Proof. Let us prove the result by induction on $n = \left|\Delta\{\vec{b}, b\} \setminus \Delta\{\vec{b}\}\right|$. Let κ be the isomorphism understood by the notation $\Delta\{\vec{a}\} \simeq \Delta\{\vec{b}\}$.

Case $n = 1$**:** For all $b' \in \Delta\{\vec{b}\}$, it holds that $b \triangle b' \subseteq \Delta\{\vec{b}\}$. If $\mathfrak{I}_{\mathfrak{P}}(b) > 0$, then put

$$a = \Big(\{a' \mid \kappa(a') \in \Delta\{\vec{b}\} \wedge \kappa(a') \in b\} \cup \bigcap \Delta\{\vec{a}\}\Big) \\ \setminus \{a' \mid \kappa(a') \in \Delta\{\vec{b}\} \wedge \kappa(a') \notin b\},$$

else consider the following equation in the unknown x in $\mathfrak{R}$

$$x = \Big(\{a' \mid \kappa(a') \in \Delta\{\vec{b}\} \wedge \kappa(a') \in b\} \cup \bigcap \Delta\{\vec{a}\}\Big) \\ \setminus \{a' \mid \kappa(a') \in \Delta\{\vec{b}\} \wedge \kappa(a') \notin b\} \cup \{x\}.$$

and observe that, by property ii of $\prec$, this admits infinitely many solutions $x = a$ with $\mathfrak{I}_{\mathfrak{R}}(a) < 0$. First of all, it is easy to see that for every $a' \in \Delta\{\vec{a}\}$,

- $x \neq a'$ and
- $a \triangle a' \subseteq \Delta\{\vec{a}\}$.

From this we have that $\Delta\{\vec{a}, a\} \setminus \Delta\{\vec{a}\} = \{a\}$. The desired conclusion easily follows once verified that, by putting $a = \kappa^{-1}(b)$, we get an $\in$-isomorphism.

Case $n > 1$**:** Consider a generic element $b' \in \Delta\{\vec{b}, b\} \setminus \Delta\{\vec{b}\}$ such that $\left|\Delta\{\vec{b}, b'\}\right| < \left|\Delta\{\vec{b}, b\}\right|$. Since the Δ-closure is defined by a monotone non-decreasing operator, we have that $n' = \left|\Delta\{\vec{b}, b'\} \setminus \Delta\{\vec{b}\}\right| < n$. The desired conclusion easily follows from the inductive hypothesis. □

Theorem 2.15 *The structures $\mathfrak{R}$ and $\mathfrak{P}$ cannot be distinguished using an m-round 3-pebble game.*

Proof. We must prove that there exists a strategy enabling *Duplicator* to win $G^3_m(\mathfrak{R}, \mathfrak{P})$. This will be shown exhibiting a sequence $(I_i)_{i \leqslant m}$ of 3-partial isomorphisms with the 3-back-and-forth property (cf. [EF99], whose notation is adopted in what follows).

We will also prove that each element in I_i will be the restriction of some embedding $\kappa_i : \mathfrak{R} \longrightarrow \mathfrak{P}$ to be used by *Duplicator* as an oracle for its strategy (let κ_m be any fixed embedding).

If *Spoiler* puts his pebble on $a' \in \mathfrak{R}$, then *Duplicator* will respond with the element $b' = \kappa_i(a') \in \mathfrak{P}$ and the oracle κ_{i-1} will remain equal to κ_i. If *Spoiler* puts his pebble on an element $b' = \kappa_i(a') \in \mathfrak{P}$, then *Duplicator* will respond with the element $a' \in \mathfrak{R}$ and, again, $\kappa_{i-1} = \kappa_i$. The 3-forth and 3-back properties are a direct consequence of κ_i being an embedding.

In order to complete the treatment of the 3-back-and-forth property, we must consider the case in which *Spoiler* puts his pebble on an element $b' \in \mathfrak{P}$ such that $b' \neq \kappa_i(a)$ for all $a \in \mathfrak{R}$. In this hypothesis let us also suppose that two more elements $b'' = \kappa_i(a'')$ and $b''' = \kappa_i(a''')$ in $\mathfrak{P}$ have been marked with a pebble (the other cases are simpler). It will be convenient to observe that, by the second property of the notion of embedding, since b' is not the image of an element in $\mathfrak{R}$, we can conclude that $b' \notin b_1 \bigtriangleup b_2$ for any b_1, b_2 such that $b_1 = \kappa_i(a_1)$ and $b_2 = \kappa_i(a_2)$ for suitable $a_1, a_2 \in \mathfrak{R}$. Hence, b' does not belong to $\Delta\{b'', b'''\}(\simeq \Delta\{a'', a'''\})$, and we can apply Lemma 2.14 in order to determine an element $a' \in \mathfrak{R}$ corresponding to Spoiler's choice $b' \in \mathfrak{P}$. Notice that Lemma 2.14 allows us to choose, when a' has negative index, a corresponding b' with arbitrarily large (negative) index. On the ground of this fact, initializing the embedding procedure with the $\in$-isomorphism between $\Delta\{a', a'', a'''\}$ and $\Delta\{b', b'', b'''\}$, we can obtain the embedding κ_{i-1} by Remark.2.11 and Lemma 2.10. □

The previous result is proved for an arbitrary m; therefore the following corollary holds.

Corollary 2.16 *The sentence* **(E)** $\wedge$ **(W)** $\wedge$ **(L)** *cannot be expressed in 3 variables.*

Proof. Consider the two structures $\mathfrak{R}$ and $\mathfrak{P}$ defined above. It is easy to verify that $\mathfrak{R} \models$ **(E)** $\wedge$ **(W)** $\wedge$ **(L)** while $\mathfrak{P} \models$ **(E)** $\wedge \neg$**(W)** $\wedge \neg$**(L)**. By Thm.2.15 and by virtue of the main result on pebble games (cf. [Imm82, Thm.C.1] and [EF99, Thm.3.3.5]) we can conclude. □

3 Conclusions

As far as expressibility issues are concerned, the content of this paper is a contribution on the 'negative' side. On the positive side, we proved elsewhere [FOP04] that the axioms **(N)**, **(W)**, and **(L)** are expressible in three variables when taken *together*. Systematic study of this subject revealed that the 3-variable expressibility of **(W)** $\wedge$ **(L)** can be ensured not only by wholesale adoption of **(N)**, but, alternatively, by an acyclicity assumption regarding membership. This indicates that 3-variable expressibility intimately depends on reasonable restraints imposed on the structure of the universe, such as saying that there is an initial set $\emptyset$ or that membership is well-founded.

Studies of this kind contribute to the rather fine classification, undertaken by Tarski and Givant [TG87, Sect.4.6], of the conditions enabling a theory of aggregates (even a very weak one) to be formulated within map-calculus (see [FOT00] for examples of such formulations). In general, to carry out this task one needs a pair of conjugated quasi-projections (cf. [TG87, Sections 4.1 and 4.4]). If a set-theoretic approach is chosen, **(N)**, **(W)**, and **(L)** constitute one of the minimal viable collections of assumptions.

As a final remark we observe that the way in which we carried out our investigation turned out to be somewhat more 'semantic' than, for instance, the application of games in establishing the inexpressibility of the density property in structures of propositional linear time logic (see [Imm82, Tho97]). *Localization*, the approach which underlies inductive constructions of the winning strategies for exploitations of pebble games in the realm of temporal structures (either linear or not), does not appear adequate for our purposes. The reason is that models of our set-theoretic frameworks are very poor in structure compared to models of time, which are either lines or trees.

In our framework, in order to suggest the winning strategy to Duplicator, it is necessary to provide a deeper insight in the peculiarities of the structures than is necessary for temporal structures. As a matter of fact, to compensate for the lack of structure and to set the ground for the desired strategy, we needed to impose somewhat artificial orderings on the domains of the structures.

References

[Bar77] J. Barwise. On Moschovakis closure ordinals. *J. Symbolic Logic*, 42:292–296, 1977.

[Daw99] A. Dawar. Finite models and finitely many variables. *Banach Center Publications*, 46, 1999.

[EF99] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Perspectives in Mathematical Logic. Springer, 1999. Second revised and enlarged edition.

[FOP04] A. Formisano, E. G. Omodeo, and P. Policriti. Three-variable statements of set-pairing. *Theoretical Computer Science*, 322(1):147–173, August 2004.

[FOT00] A. Formisano, E. G. Omodeo, and M. Temperini. Goals and benchmarks for automated map reasoning. *J. Symbolic Computation*, 29(2), 2000. Special issue. M.-P. Bonacina and U. Furbach, eds.

[Hei77] J. van Heijenoort, editor. *From Frege to Gödel - A source book in mathematical logic, 1879–1931.* Source books in the history of the sciences. Harvard University Press, 3^{rd} printing edition, 1977.

[Hod93] I. Hodkinson. Finite variable logics. *Bulletin of the European Association for Theoretical Computer Science*, 51:111–140, October 1993. Columns: Logic in Computer Science.

[IK89] N. Immerman and D. Kozen. Definability with bounded number of bound variables. *Information and Computation*, 83(2):121–139, November 1989.

[Imm82] N. Immerman. Upper and lower bounds for first order expressibility. *J. Computer and System Sciences*, 25(1):76–98, August 1982.

[KV96] P. G. Kolaitis and M. Y. Vardi. On the expressive power of variable-confined logics. In *Proceedings, 11$^{\text{th}}$ Annual IEEE Symposium on Logic in Computer Science*, pages 348–359, New Brunswick, New Jersey, 1996. IEEE Computer Society Press.

[Kwa81] M. K. Kwatinetz. *Problems of expressibility in finite languages*. PhD thesis, University of California, Berkeley, 1981.

[Tar53] A. Tarski. Some metalogical results concerning the calculus of relations. *J. Symbolic Logic*, 18:188–189, 1953.

[TG87] A. Tarski and S. Givant. *A formalization of Set Theory without variables*, volume 41 of *Colloquium Publications*. American Mathematical Society, 1987.

[Tho97] W. Thomas. Languages, automata and logic. In G. Rozenberg and Salomaa A., editors, *Handbook of formal languages, Vol. III*, pages 389–455. Springer, 1997.

[Zer08] E. Zermelo. Untersuchungen über die Grundlagen der Mengenlehre I. In Heijenoort [Hei77], pages 199–215. (English translation).