Coordination of 01-agents vs. coordination of worlds-based agents*

Alessandro Agostini
Università di Siena

Dick de Jongh
ILLC Amsterdam

Franco Montagna
Università di Siena

Abstract

As far as we know, a *learning to coordinate* paradigm was first introduced in *Learning Theory* [JORS99] by using the tools of recursion theory [MOss]. In this paper, we present a first-order paradigm of coordination—we call this paradigm of *sf-coordination*. The paradigm of sf-coordination is shown to extend Montagna-Osherson's paradigm, in the sense that Montagna and Osherson's "01-agents" coordinate if and only if some "first-order equivalent" agents of the first-order paradigm sf-coordinate.

1 Introduction

We start with a concrete example intended to help the reader interpret some of the abstract concepts described later. Thus, we image two "worlds-based agents", say **A**lfonso and **B**arbara, whose "background world" is represented by two nonempty classes of structures $\mathbf{A}$ and $\mathbf{B}$ of a given signature, and whose aim is to coordinate. For doing this, **A**lfonso and **B**arbara try to communicate with each other in order to respectively end up with a description of two structures $\mathcal{A} \in \mathbf{A}$ and $\mathcal{B} \in \mathbf{B}$ such that $\mathcal{A}$ is sufficiently close to $\mathbf{B}$ and $\mathcal{B}$ is sufficiently close to $\mathbf{A}$. Roughly, we might think that $\mathcal{A}$ is sufficiently close to $\mathbf{B}$ if every finite set of literals true in $\mathcal{A}$ under some full assignment is satisfiable in some $\mathcal{B} \in \mathbf{B}$, and vice versa. Thus, we expect that the more **A**lfonso is like **B**arbara, the better chance **A**lfonso and **B**arbara have of reaching coordination. To dramatize, let us suppose that each agent does not know the "background world" of the other agent, and that agents were never before in a similar situation, so they cannot rely on past experience to solve their coordination problem. **A**lfonso and **B**arbara's decisions have a strategic component. Since strategic interactions are best modeled by game theory, we image the agents' coordination problem as a form of 2-players game, which we qualify as a *coordination game*.

To start the game, **A**lfonso is conceived as choosing one member from $\mathbf{A}$ to be his "actual world"; **A**lfonso's choice is initially unknown to **B**arbara. **A**lfonso then provides a "clue" (the **A**lfonso's *behaviour*) about his world. At the same time, **B**arbara does her choice as well, and provides **A**lfonso with a clue about her

*Electronic correspondence to: agostini@wins.uva.nl, dickdj@wins.uva.nl, montagna@unisi.it.

actual world. We can assume that **A**lfonso and **B**arbara are allowed to change their actual world at each step of the game, provided that they remain coherent with the behaviour they have shown since then. Clearly, it is safe to begin with a behaviour coherent on many worlds in the class. In this way, if, say **A**lfonso, realizes that the structure he has in mind, *i.e.* his actual world, is not close enough to **B**arbara's world, he can change it, and vice versa. Agents may provide "bad clues", in principle. Thus, we can consider a paradigm of *coordination-by-failure* or "slow" coordination, according to which the agents are allowed to start the game again from the beginning whenever some disagreement occur. Of course, to reach coordination this should happen only finitely often. **A**lfonso's clues constitute the data upon which **B**arbara will base her hypotheses on **A**lfonso's background world, that eventually become themselves a clue for **A**lfonso about **B**arbara's world. And so on. Each time **A**lfonso provides a new clue, **B**arbara may produce a new hypothesis, and a new clue for **A**lfonso as well. **A**lfonso and **B**arbara *win the game* if the successive clues about their own background worlds eventually stabilize to a consistent set of clues satisfiable in one of the other agent's worlds. Both lose otherwise. As a necessary condition for winning, each agent's behaviour must be consistent with agent's own background world.

Our aim is to reduce "slow" coordination to coordination in the sense of Montagna and Osherson [MOss]. To achieve our goal, this paper is thus structured as follows. In next section we formalize the foregoing game-theoretic picture of "slow" coordination in the setting of *Formal Learning Theory* [OdJMW97] and introduce some general notation that will be used throughout the whole paper. In Section 3 we briefly review the learning to coordinate paradigm [MOss] as it serves as the base of the reduction. Section 4 then provides such a formal reduction as the main result.

2 The sf-coordination paradigm

Some concepts like 'agents', 'clues', and 'success' figure in the foregoing picture of **A**lfonso and **B**arbara coordination game. We formalize them as follows.

2.1 Notation

We fix a first-order language $\mathcal{L}_{form}$ with vocabulary $\mathcal{L} = \{=, U, s, \overline{0}\}$ and countable set of variables $Var = \{x_i \mid i \in N\}$. $\overline{0}$ and s denote, respectively, 0 and the successor function on the set of natural numbers. Thus, the term that results from $n \in N$ applications of s to 0 is denoted $\overline{n}$; U denotes a set of natural numbers. We use $\mathcal{L}_{basic}$ and $\mathcal{L}_{sen}$ to denote, respectively, the set of *basic formulas* (*i.e.*, atomic formulas and their negations) and the set of sentences of $\mathcal{L}_{form}$. We are particularly interested in the collection of all the *finite* sequences over $\mathcal{L}_{basic}$. We denote such collection by *SEQ*. The set $\{0, 1, 2, ...\}$ of natural numbers is denoted: N. If X is a set, X^ω is the set of infinite sequences over X. A sequence in X^ω is called an *ω-sequence* (over X). We use σ, η, ϱ to denote sequences in *SEQ*, and α, β

to denote elements of these sequences. Concatenation of sequences is indicated by juxtaposition, and we won't distinguish notationally between an element and the corresponding unit sequence. Thus $\alpha\sigma$ denotes the sequence with first element α and tail σ. We write $length(\eta)$ for the length of a finite sequence and η_i for the ith element of η, $0 \leq i < length(\eta)$. We write $\sigma \sqsubseteq \eta$ if σ is a prefix of η, *i.e.* $\eta = \sigma\varrho$ for some ϱ. In this case, we say that η *extends* σ.

Let ϱ be either an ω-sequence or a finite sequence with $length(\varrho) \geq i$, $i \in N$. We write $\varrho(i)$, $i < length(\varrho)$, for the finite sequence $\langle \varrho_0 \cdots \varrho_i \rangle$. For $i \leq length(\varrho)$, $\varrho|_i$ denotes the proper initial segment of length i in ϱ (so $\varrho(i) = \varrho|_{i+1}$) and $\varrho^{(i)}$ denotes the sequence obtained from ϱ by deleting its first i elements. We write $range(\eta)$ for the set of all the elements of a finite or infinite sequence. We denote the finite sequence of length zero by $\emptyset$. For η finite, η^- denotes the sequence obtained from η by dropping its last (rightmost) element, if $\eta \neq \emptyset$; η^- is $\emptyset$, otherwise.

Given a pair of sets A, B, we write $A \triangle B$ for the symmetric difference of A and B, *i.e.* the set $(A - B) \cup (B - A)$. If $A \triangle B$ is finite, we say that A is a *finite variant of* B and A and B are called *finite variants*. Observe that $A \triangle B$ is finite iff either $A = B$ or A and B differ by a finite set.

2.2 Agents and coordination criteria

Previously, we used the term "coordination" intuitively as a very general binary relation on agents. In this section we make precise this relation by extending it onto *first-order* agents. In the sequel, we shall use "structure" in place of "$\mathcal{L}$-structure".

Definition (1) *Let nonempty class* $\mathbf{A}$ *of structures and mapping* Ψ' *from SEQ to* $\mathcal{L}_{sen}$ *be given. We say that* $\Psi = \langle \Psi', \mathbf{A} \rangle$ *is a (first-order)* agent.

For all $\sigma \in SEQ$, we then write $\Psi(\sigma)$ for $\Psi'(\sigma)$. We call the first component *basic agent*, and the second component *world*. The new relation we have in mind is "slow-full" coordination. We rely on the following terminology and notation.

Let Λ denote the class of all (basic) agents. The following definition provides the formal meaning of the coordination "play" between two first-order agents.

Definition (2) *Let* $\Psi, \Phi \in \Lambda$ *be given.*

(a) The enumeration from Ψ and Φ *is the pair* $[\overline{\psi}, \overline{\phi}]$ *of* ω*-sequences inductively defined as follows.*

  (i) $\overline{\psi}_0 = \Psi(\emptyset)$ *and* $\overline{\phi}_0 = \Phi(\emptyset)$.

  (ii) Let $\overline{\psi}(n) = \langle \overline{\psi}_0 \cdots \overline{\psi}_n \rangle$ *and* $\overline{\phi}(n) = \langle \overline{\phi}_0 \cdots \overline{\phi}_n \rangle$ *be defined. Then,* $\overline{\psi}_{n+1} = \Psi(\overline{\phi}(n))$ *and* $\overline{\phi}_{n+1} = \Phi(\overline{\psi}(n))$.

(b) Let $k \in N$ *be given. The* enumeration from Ψ and Φ starting at k *is the pair* $[\overline{\psi}^{(k)}, \overline{\phi}^{(k)}]$.

Notice that $range(\overline{\psi}) = \{\overline{\psi}_n \,|\, n \in N\} = \{\Psi(\overline{\phi}|_n) \,|\, n \in N\}$ and $range(\overline{\psi}^{(k)}) = \{\overline{\psi}_n \,|\, n \geq k\}$. We say that $[\overline{\psi}, \overline{\phi}]$ is the *enumeration from worlds-based agents* $\langle \Psi_0, \mathbf{A}\rangle$ *and* $\langle \Phi_0, \mathbf{B}\rangle$ just in case $[\overline{\psi}, \overline{\phi}]$ is the enumeration from Ψ_0 and Φ_0. In the rest of this paper we will write $R\langle \Psi, \Phi\rangle$ for $\overline{\psi}$ and k-$R\langle \Psi, \Phi\rangle$ for $\overline{\psi}^{(k)}$.

Definition (3) *Let agents* $\langle \Psi_0, \mathbf{A}\rangle$ *and* $\langle \Phi_0, \mathbf{B}\rangle$ *be given.* $\langle \Psi_0, \mathbf{A}\rangle$ and $\langle \Phi_0, \mathbf{B}\rangle$ sf-coordinate *just in case for some* $k \in N$, k-$R\langle \Psi_0, \Phi_0\rangle$ *is an environment for some* $\mathcal{A} \in \mathbf{A}$, k-$R\langle \Phi_0, \Psi_0\rangle$ *is an environment for some* $\mathcal{B} \in \mathbf{B}$ *and for all* $n \geq k$, k-$R\langle \Psi_0, \Phi_0\rangle(n)$ *is satisfiable in some* $\mathcal{B} \in \mathbf{B}$ *and* k-$R\langle \Phi_0, \Psi_0\rangle(n)$ *is satisfiable in some* $\mathcal{A} \in \mathbf{A}$.

In the previous definition of coordination, agents are free to stabilize to a suitable description of their worlds after a finite number of failures (disagreements). For this reason, we qualified this paradigm of "slow" coordination. Given two agents, sf-coordination arises between them if agents fail to communicate their clues in the play. According to this paradigm, agents can restart finitely many often, but after the last failure message they must eventually coordinate.

3 The LC-paradigm

[MOss] introduce the following paradigm of coordination, called "learning to coordinate" or LC-paradigm in short. Informally, two *players* want to coordinate by repeatedly showing each other one of two possible behaviours. The problem of coordination the players are faced with follows from the shifting constraints on their behaviours. Each player tries to predict the other's behaviour, and their predictions are based on no more than the history of earlier events. One player "learns" the other's behaviour just in case his or her own behaviour matches the other's forever after. The relation of "learns" is in this case symmetric. To keep matters simple, Montagna and Osherson's paper considers only two players facing the same two options on each trial, and they denote the options by 0 and 1. A player is therefore to be identified with a function from the set of all finite binary sequences into $\{0, 1\}$, where any such sequence is conceived as the history of moves of an opposing player. From a sequence of length n, a player can reconstruct the $2 \times n$ binary matrix that includes his own responses through move n. The Montagna and Osherson paradigm is game-theoretic in spirit; there are two, fixed players in a game (a game-theorist would say that the game is a *repeated game*), where each player must discover a strategy that fits the other's play. Both the players win or lose, there is not a unique winner.

To define Montagna and Osherson's paradigm precisely, let *BISEQ* denote the collection of all finite sequences over $\{0, 1\}$. Aside from any game-theoretic paradigm, let players be simple agents of the following form.

Definition (4) *A* 01-agent *is a function from BISEQ to* $\{0, 1\}$.

Let 01-agents Ψ and Φ be given. Following [MOss], let $R\langle \Psi, \Phi\rangle$ and $R\langle \Phi, \Psi\rangle$ be finite or infinite sequences over $\{0, 1\}$ defined as follows.

Definition (5) *Let $n \in N$, and 01-agents Ψ and Φ be given. Suppose $R\langle\Psi,\Phi\rangle|_n$ and $R\langle\Phi,\Psi\rangle|_n$ be defined. Then, $R\langle\Psi,\Phi\rangle_n = \Psi(R\langle\Phi,\Psi\rangle|_n)$ and $R\langle\Phi,\Psi\rangle_n = \Phi(R\langle\Psi,\Phi\rangle|_n)$.*

In Montagna and Osherson's terminology, $R\langle\Psi,\Phi\rangle$ is Ψ*'s response to* Φ and $R\langle\Phi,\Psi\rangle$ is Φ*'s response to* Ψ.

In what follows, "η" stands for "numerical", or also "numerically", and "s" stands for as "slow" or "slowly", as required by the context in use.

Definition (6) *Let 01-agents Ψ and Φ be given.* Ψ ηs-coordinates with Φ *just in case:*

(*i*) *$R\langle\Psi,\Phi\rangle$ and $R\langle\Phi,\Psi\rangle$ are both ω-sequences, and*
(*ii*) *for all but finitely many $n \in N$, $R\langle\Psi,\Phi\rangle_n = R\langle\Phi,\Psi\rangle_n$.*

Montagna and Osherson's learning to coordinate relation is fully captured by the foregoing definitions.

4 The main result

We want to show that the first-order paradigm of coordination is more general than Montagna and Osherson's learning-to-coordinate paradigm. For doing this, recall that $\mathcal{L} = \{=, U, s, \overline{0}\}$, *i.e.* we use one constant symbol $\overline{0}$, one unary function symbol s and one unary predicate symbols U plus the symbol $=$ for equality. For $U \subseteq N$, $\mathcal{N}_U$ denotes the structure $\langle N, U, s, 0\rangle$ that interprets $\mathcal{L}$. In other words, $\mathcal{N}_U$ is the usual model $\mathcal{N} = \langle N, s, 0\rangle$ of the natural numbers plus a unary relation U on N defined as follows. For every $x \in Var$ and for some full assignment h to $\mathcal{N}_U$,

(a) if $h(x) \in U$, then $\mathcal{N}_U \models U(x)[h]$;
(b) if $h(x) \notin U$, then $\mathcal{N}_U \models \neg U(x)[h]$.

We also stipulate the following convention.

Convention (7) *We assume that $\mathcal{L}_{basic}$ is ordered according to a fixed, recursive list l. Moreover, we assume that the first element of l is: $x_0 = \overline{0}$.*

Thus, we will refer to any sequence over $\mathcal{L}_{basic}$ as ordered according to l without further explanation.

Remark. In a purely predicative framework without function and constant symbols, we may express the idea above by using an unary predicate symbol Z representing the singleton of 0 in place of $\overline{0}$, and a binary predicate S representing the graph of the successor function in place of s.

We will associate with every 01-agent Ψ a worlds-based agent $\langle\Psi^\circ, \mathbf{W}_\Psi\rangle$ so as to be able to prove our main result:

Proposition (8) *For all $\Psi, \Phi \in \Lambda_0^1$, Ψ ηs-coordinates with Φ iff $\langle\Psi^\circ, \mathbf{W}_\Psi\rangle$ and $\langle\Phi^\circ, \mathbf{W}_\Phi\rangle$ sf-coordinate.*

In fact, our construction of $\langle \Psi^\circ, \mathbf{W}_\Psi \rangle$ from Ψ and of $\langle \Phi^\circ, \mathbf{W}_\Phi \rangle$ from Φ will satisfy a stronger property, namely:

Corollary (9) *For all $\Psi, \Phi \in \Lambda^1_0$, the following holds.*

(a) If Ψ does not ηs-coordinate with Φ, then for all $k \in N$, k-$R\langle \Psi^\circ, \Phi^\circ \rangle$ and k-$R\langle \Psi^\circ, \Phi^\circ \rangle$ are not environments for any world.

(b) If Ψ ηs-coordinates with Φ, then for some $k \in N$, k-$R\langle \Psi^\circ, \Phi^\circ \rangle$ and k-$R\langle \Psi^\circ, \Phi^\circ \rangle$ are environments for a world in $\mathbf{W}_\Psi \cap \mathbf{W}_\Phi$.

Before proving the proposition, we give an informal account of Ψ°'s coordination behaviour (Φ°'s behaviour is the same). Of course, Ψ°'s behaviour is required to make sense only on worlds-based agents of the form $\langle \Phi^\circ, \mathbf{W}_\Phi \rangle$ for some $\Phi \in \Lambda^1_0$, because we make no claim in Proposition (8) about Ψ°'s behaviour in regard to other agents.

Let $U\langle \Psi, \Phi \rangle$ denote the set $\{n \in N \mid R\langle \Psi, \Phi \rangle_n = 1\}$. Thus:

- Ψ° gives some information to Φ° about Ψ-moves in a game with Φ. This information is coded by literals of the form $U(x_n)$ or $\neg U(x_n)$, under the intuition that $U(x_n)$ codes the information *at stage n Ψ plays* 1, and that $\neg U(x_n)$ codes the information *at stage n Ψ plays* 0.

- Ψ° decodes the information coming from Φ° (again, coded by literals of the form $U(x_n)$ or $\neg U(x_n)$), about Φ's response to Ψ's moves.

- Ψ° describes an environment consisting of all literals true in $\mathcal{N}_{U\langle \Psi, \Phi \rangle}$ under the assignment h defined by $h(x_i) = i$, $i \in N$. At each stage, Ψ°'s description is partial, in the sense that only a finite part of $U\langle \Psi, \Phi \rangle$ is known to Ψ°. Even though at each step Ψ°'s knowledge of $U\langle \Psi, \Phi \rangle$ is partial, Ψ° updates his information according to the moves of Φ°.

- As soon as Ψ° discovers a new disagreement between Ψ and Φ, Ψ° outputs some obviously false information, say $\neg x_0 = \overline{0}$, and restarts the description of $\mathcal{N}_{U\langle \Psi, \Phi \rangle}$ by enumerating the basic diagram of $\mathcal{N}_{U\langle \Psi, \Phi \rangle}$, using the assignment h to supply temporary names for the members of $\mathcal{N}_{U\langle \Psi, \Phi \rangle}$. The reason for this false information is that Ψ° dislikes to coordinate with Φ°; in other words, $\langle \Psi^\circ, \mathbf{W}_\Psi \rangle$ and $\langle \Phi^\circ, \mathbf{W}_\Phi \rangle$ will not sf-coordinate if Ψ does not ηs-coordinate with Φ. However, if this happens, *i.e.* if Ψ and Φ disagree infinitely often, then Ψ° repeats its false information infinitely often. The result is that Ψ°'s description is a contradictory list of literals, with infinitely many elements of the forms $x_0 = \overline{0}$, $\neg x_0 = \overline{0}$. Thus, no possibility of coordination arises. Notice that the false information $\neg x_0 = \overline{0}$ is not misleading to Φ°, because he knows everything about $\mathcal{N}$ and the assignment h.

Recall that our goal is to define Ψ° and $\mathbf{W}_\Psi$ (similarly: Φ° and $\mathbf{W}_\Phi$) such that for every $\Phi \in \Lambda^1_0$, (a) if Ψ ηs-coordinates with Φ, then for some $k \in N$, k-$R\langle \Psi^\circ, \Phi^\circ \rangle$ is an environment for $\mathcal{N}_{U\langle \Psi, \Phi \rangle} \in \mathbf{W}_\Psi$ and k-$R\langle \Phi^\circ, \Psi^\circ \rangle$ is an environment for $\mathcal{N}_{U\langle\Phi,\Psi\rangle} \in \mathbf{W}_\Phi$. Since the symmetric difference of $U\langle\Psi,\Phi\rangle$ and $U\langle\Phi,\Psi\rangle$ is finite, it follows that both k-$R\langle\Psi^\circ,\Phi^\circ\rangle$ and k-$R\langle\Phi^\circ,\Psi^\circ\rangle$ are environments for some world in $\mathbf{W}_\Psi \cap \mathbf{W}_\Phi$; (b) if Ψ does not ηs-coordinate with Φ, then for all $k \in N$, both k-$R\langle\Psi,\Phi\rangle$ and k-$R\langle\Phi,\Psi\rangle$ contain infinitely many occurrences of a literal and its negation. Hence, for all $k \in N$, k-$R\langle\Psi^\circ,\Phi^\circ\rangle$ and k-$R\langle\Phi^\circ,\Psi^\circ\rangle$ are environments for no world. We need some more notation.

Definition (10) *Let $\sigma \in BISEQ$ and $\Psi \in \Lambda_0^1$ be given. We define $\Psi[\sigma] \in BISEQ$ as follows.*

(a) $length(\Psi[\sigma]) = length(\sigma) + 1$.

(b) $\Psi[\sigma]_0 = \Psi(\emptyset)$. *If* $n < length(\sigma)$ *then* $\Psi[\sigma]_{n+1} = \Psi(\sigma(n))$.

For $\Psi \in \Lambda$ and $\sigma \in SEQ$, $\Psi[\sigma] \in SEQ$ is defined mutatis mutandis.

Example. Suppose $\sigma = \beta_0\beta_1\beta_2$. Then,
$\Psi[\sigma]_0 = \Psi(\emptyset)$;
$\Psi[\sigma]_1 = \Psi(\beta_0)$;
$\Psi[\sigma]_2 = \Psi(\beta_0\beta_1)$;
$\Psi[\sigma]_3 = \Psi(\beta_0\beta_1\beta_2)$.

The proof of Proposition (8) proceeds through the construction of worlds-based agents $\langle\Psi^\circ,\mathbf{W}_\Psi\rangle$, $\langle\Phi^\circ,\mathbf{W}_\Phi\rangle$ and a series of lemmas. We start with the construction; for $\Psi,\Phi \in \Lambda_0^1$, we first define $\mathbf{W}_\Psi$ and $\mathbf{W}_\Phi$.

Definition (11) *For $\Psi \in \Lambda_0^1$, $\mathbf{W}_\Psi$ is the class of worlds $\mathcal{N}_U$ such that for some $\Phi \in \Lambda_0^1$,*

(a) Ψ ηs-coordinates with Φ, and

(b) $U \triangle U\langle\Psi,\Phi\rangle$ is finite.

For every $\sigma \in SEQ$, we define the "environment transformation" τ, the environment e_σ, and agent Ψ°. $\tau(\sigma)$ is a compressed version of the information about the Ψ-Φ game coded by σ, while e_σ represents the partial description (according to the recursive list l in Convention (7)) produced by Ψ° upon receiving σ.

Definition (12) *Let $\sigma \in SEQ$ be given. Denote by $S(\sigma)$ the set $\{n \in N \mid \forall i \leq n\ (U(x_i) \in range(\sigma) \vee \neg U(x_i) \in range(\sigma))$. The* e-transformation *is the total function $\tau : SEQ \longrightarrow BISEQ$ such that if $S(\sigma) = \emptyset$, then $\tau(\sigma) = \emptyset$. Otherwise, let m be the maximum of $S(\sigma)$. We define:*

(a) $length(\tau(\sigma)) = m + 1$.

(b) For $i < length(\tau(\sigma))$,

$$\tau(\sigma)_i = \begin{cases} 1 & \text{if } \exists j < length(\sigma)\ (\sigma_j = U(x_i))\ \&\ \forall k \leq j\ (\sigma_k \neq \neg U(x_i)) \\ 0 & \text{otherwise.} \end{cases}$$

Roughly speaking, τ transforms the positive and negative information coded by a sequence in *SEQ* by means of a finite, binary sequence. For $\sigma \in SEQ$, observe that $\tau(\sigma)$ is generally shorter than σ.

Example. (a) Suppose $\sigma = U(x_0)\beta_1\neg U(x_1)\beta_3\beta_4$. Then, $S(\sigma) = \{0,1\}$ and $\tau(\sigma) = 10$. Let $\alpha_1 \in \mathcal{L}_{basic}$ neither of the form $U(x)$ nor $\neg U(x)$, and let $\alpha_2 = U(x_2)$. It follows that $S(\sigma\alpha_1) = \{0,1\}$ and $\tau(\sigma\alpha_1) = 10$, thus $\tau(\sigma\alpha_1) = \tau(\sigma)$; while $S(\sigma\alpha_2) = \{0,1,2\}$ and $\tau(\sigma\alpha_2) = 101$, thus $\tau(\sigma\alpha_2) = \tau(\sigma)1$, *i.e.*, $\tau(\sigma\alpha_2) = \tau(\sigma)\tau(\sigma\alpha_2)_2$.

(b) Suppose $\sigma = U(x_0)\beta_1\neg U(x_2)\beta_3\beta_4$. Then, $S(\sigma) = \{0\}$ and $\tau(\sigma) = 1$. Let $\alpha = \neg U(x_1)$. It follows that $S(\sigma\alpha) = \{0,1,2\}$ and $\tau(\sigma\alpha) = 101$, thus $\tau(\sigma\alpha) = \tau(\sigma)01$, *i.e.*, $\tau(\sigma\alpha) = \tau(\sigma)\tau(\sigma\alpha)_1\tau(\sigma\alpha)_2$.

(c) Suppose $\sigma = \beta_0\neg U(x_2)\beta_2\beta_3\neg U(x_4)\beta_5 U(x_3)U(x_0)$. Then, $S(\sigma) = \{0,1,2,3,4\}$ and $\tau(\sigma) = 11010$. Let $\alpha = U(x_4)$. It follows that $S(\sigma\alpha) = S(\sigma)$ and $\tau(\sigma\alpha) = \tau(\sigma)$.

(d) Suppose $\sigma = \beta_0 U(x_5)\beta_2 U(x_1)\neg U(x_4)\beta_5 U(x_3)\beta_7 U(x_0)$. Then, $S(\sigma) = \{0,1\}$ and $\tau(\sigma) = 11$. Let $\alpha = U(x_2)$. It follows that $S(\sigma\alpha) = \{0,1,2,3,4,5\}$ and $\tau(\sigma\alpha) = 111101$, thus $\tau(\sigma\alpha) = \tau(\sigma)1101$, *i.e.*, $\tau(\sigma\alpha) = \tau(\sigma)\tau(\sigma\alpha)_2 \ldots \tau(\sigma\alpha)_5$.

We note that case (c) of the previous example shows that if the conditional "if $U(x_i) \in range(\sigma)$" in Definition (12) were used in place of "if $\exists j < length(\sigma)\ (\sigma_j = U(x_i))\ \&\ \forall k \leq j\ (\sigma_k \neq \neg U(x_i))$", then it would be no more the case that $\tau(\sigma\alpha) = \tau(\sigma)$, because of $\tau(\sigma\alpha) = 11011$.

Definition (12) will be used in such a way that at any stage n in the coordination game between Ψ° and Φ°, when Ψ°'s last move is $R\langle\Psi^\circ, \Phi^\circ\rangle_{n-1}$, σ is the finite sequence $R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1}$ of Φ°'s previous responses to Ψ°. We will show pretty soon that in this very special case $length(\tau(\sigma^-)) = length(\tau(\sigma)) - 1$. The reader should keep this in mind in reading Lemma (13) and next two definitions.

The previous example sets some general properties of τ. We state these properties as follows.

Lemma (13) *Let $\sigma \in SEQ$ be given.*

(a) For all $\eta \in SEQ$, if $\sigma \sqsubseteq \eta$ then $\tau(\sigma) \sqsubseteq \tau(\eta)$.

(b) Let $length(\sigma) = n + 1$ with $n \in N$. Then, $\tau(\sigma)$

$= \tau(\sigma^-)$, if for all $i \in N$, $\sigma_n \neq U(x_i)$ and $\sigma_n \neq \neg U(x_i)$, or either $\sigma_n = U(x_i)$ or $\sigma_n = \neg U(x_i)$ for some $i \in N$ with $i < length(\tau(\sigma^-))$;

$= \tau(\sigma^-)\tau(\sigma)_{length(\tau(\sigma^-))} \ldots \tau(\sigma)_{length(\tau(\sigma))-1}$, otherwise.

Definition (14) *Let $\sigma \in SEQ$, $\Psi \in \Lambda_0^1$, $\mathcal{N}$ and full assignment h to $\mathcal{N}$ such that, for all $i \in N$, $h(x_i) = i$ be given. The* p-environment for $\mathcal{N}$ and h generated by Ψ on $\tau(\sigma)$ *is the sequence e_σ such that, if $\Psi[\tau(\sigma)]$ is undefined, then e_σ is undefined. Otherwise,* $range(e_\sigma) = \{\beta \in \mathcal{L}_{basic} \mid \mathcal{N} \models \beta[h]\} \cup \{U(x_i) \mid i < length(\tau(\sigma)) \wedge \Psi[\tau(\sigma)]_i = 1\} \cup \{\neg U(x_i) \mid i < length(\tau(\sigma)) \wedge \Psi[\tau(\sigma)]_i = 0\}$.

In particular, notice that e_σ strictly depends on list l (cf. Convention (7)) and assignment h. Moreover, it follows from Convention (7) that $x_0 = \overline{0}$ is the first element of e_σ. This remark is indeed crucial in the next definition.

Definition (15) *Let $\sigma \in SEQ$ and $\Psi \in \Lambda_0^1$ be given. Ψ° is the basic agent such that if $\Psi[\tau(\sigma)]$ is undefined, then $\Psi^\circ(\sigma)$ is undefined. Otherwise, let i_0 be the maximum $i < length(\Psi^\circ[\sigma^-])$ such that $\Psi^\circ[\sigma^-]_i$ is of the form $\neg x_0 = \overline{0}$; if such an i does not exist, let $i_0 = -1$. Then,*

(a) if $\Psi(\tau(\sigma^-)) = \neg x_0 = \overline{0}$, then $\Psi^\circ(\sigma)$ is of the form $x_0 = \overline{0}$;

(b) if either $\tau(\sigma) = \tau(\sigma^-)$ or $\tau(\sigma) \neq \tau(\sigma^-)$ and $\tau(\sigma)_{length(\tau(\sigma))-1} = \Psi(\tau(\sigma^-))$, $\Psi^\circ(\sigma)$ is the first element β in e_σ such that for all $i_0 < i < length(\Psi^\circ[\sigma^-])$, $\beta \neq \Psi^\circ[\sigma^-]_i$;

(c) if $\tau(\sigma) \neq \tau(\sigma^-)$ and $\tau(\sigma)_{length(\tau(\sigma))-1} \neq \Psi(\tau(\sigma^-))$, $\Psi^\circ(\sigma)$ is of the form $\neg x_0 = \overline{0}$.

We say that Ψ° is the *first-order equivalent of* Ψ. Notice that $\Psi^\circ[\sigma^-]$ is the sequence of the previous moves of Ψ°, and i_0 represents the last place where an error occurred. In this case, we can imagine that Ψ°'s description (of his world) restarts from the place $i = i_0 + 1$. Thus, if $i_0 = -1$, the previous definition suggests that agents'coordination process starts again from the beginning (place $i = 0$). In Definition (15), clause (a) says that if the last literal of σ does not contain any new information about Φ's moves, or if the last literal of σ contains information about a new move of Φ, but this move agrees with Ψ's last move, then Ψ° outputs on σ a new literal in e_σ, namely, the first element of e_σ that is not output after the last error. According to Definition (15)b, Ψ° response on σ is a failure, *i.e.* a literal of the form $\neg x_0 = \overline{0}$, if the last literal in σ contains information about a new move of Φ and this move disagrees with Ψ's last move. Finally, clause (c) ensures that $x_0 = \overline{0}$ is played by Ψ° whenever 01-agents Ψ and Φ always disagree from the beginning. Definition (15) concludes the construction of Ψ° and Φ°. Observe:

Lemma (16) *For every $\Psi \in \Lambda_0^1$:*

(a) If Ψ is total, then Ψ° is total.

(b) If Ψ is recursive, then Ψ° is recursive.

The next lemma is a reformulation of Lemma (13) for sequences of basic formulas "played" by agents Ψ° and Φ°. Before stating the lemma, we note that from the previous three definitions it follows that for every sequence σ of the form $R\langle\Psi^\circ, \Phi^\circ\rangle|_n$, there is $j < length(\sigma)$ such that both $\sigma_j = U(x_i)$ and for all $k \leq j$, $\sigma_k \neq \neg U(x_i)$ if and only if $U(x_i) \in range(\sigma)$. This remark will be indeed useful in the proof of the lemma, where we will focus for convenience on the shortest notation.

Lemma (17) *For all $\Psi, \Phi \in \Lambda_0^1$ such that $R\langle\Psi, \Phi\rangle$ is an infinite sequence, the following holds. For all $n > 0$, either*

(a) $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) = \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})$ *or*

(b) $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) = \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n)_{length(\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n))-1}$.

Proof: We have only to show that if $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) \neq \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})$, then $length(\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})) = length(\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n)) - 1$. (The lemma then follows immediately as a special case of Lemma (13)b.) Suppose $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) \neq \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})$. Then by the definition of τ, either $R\langle\Psi^\circ, \Phi^\circ\rangle_{n-1} = U(x_i)$ or $R\langle\Psi^\circ, \Phi^\circ\rangle_{n-1} = \neg U(x_i)$ with $i = length(\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1}))$. From the definition of Ψ° it follows that either $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) = \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})1$ or $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) = \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})0$. Hence, $length(\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n)) = length(\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1})) + 1$. ■

Lemma (18) *For all $\Psi, \Phi \in \Lambda_0^1$ such that $R\langle\Psi, \Phi\rangle$ is an infinite sequence, the following conditions hold.*

(a) *For all $n \in N$, $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_n) \sqsubseteq R\langle\Psi, \Phi\rangle|_n$.*

(b) *For all $n \in N$, there is $m \in N$ such that $R\langle\Psi, \Phi\rangle|_n \sqsubseteq \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_m)$.*

Roughly, Lemma (18)a says that Φ° receives only correct information from Ψ°about Ψ's moves, or also that the (positive and negative) information contained in $R\langle\Psi^\circ, \Phi^\circ\rangle$ reflects the (positive and negative) information contained in $R\langle\Psi, \Phi\rangle$. Lemma (18)b says that sooner or later Φ° receives from Ψ° all information about $R\langle\Psi, \Phi\rangle$; in other words, the (positive and negative) information contained in $R\langle\Psi, \Phi\rangle$ will appear in $R\langle\Psi^\circ, \Phi^\circ\rangle$ in finite time.

Proof of Lemma (18): (18)a By induction on n. For $n = 0$ the claim holds because of $\tau(\emptyset) = \emptyset$. For $n > 0$, assume $\tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_{n-1}) \sqsubseteq R\langle\Psi, \Phi\rangle|_{n-1}$ and $\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1}) = R\langle\Phi, \Psi\rangle|_{n-1}$ (inductive hypothesis). Let $\sigma = R\langle\Psi^\circ, \Phi^\circ\rangle|_n$ and $i = length(\tau(\sigma))$. We want to prove that $\tau(\sigma) \sqsubseteq R\langle\Psi, \Phi\rangle|_n$. By Lemma (17), either $\tau(\sigma) = \tau(\sigma^-)$ or $\tau(\sigma) = \tau(\sigma^-)\tau(\sigma)_{i-1}$. We consider the following three cases:

Case 1: $\tau(\sigma) = \tau(\sigma^-)$. It follows by the inductive hypothesis that $\tau(\sigma) \sqsubseteq R\langle\Psi, \Phi\rangle|_{n-1}$. Observe that $R\langle\Psi, \Phi\rangle|_n = R\langle\Psi, \Phi\rangle|_{n-1}R\langle\Psi, \Phi\rangle_{n-1}$, and then $R\langle\Psi, \Phi\rangle|_{n-1} \sqsubseteq R\langle\Psi, \Phi\rangle|_n$. Hence $\tau(\sigma) \sqsubseteq R\langle\Psi, \Phi\rangle|_n$.

Case 2: $\tau(\sigma) \neq \tau(\sigma^-)$ and $\tau(\sigma)_{i-1} = 1$. Then, by the definition of τ, $U(x_{i-1}) \in range(\sigma)$, that is $U(x_{i-1})$ is one of the $R\langle\Psi^\circ, \Phi^\circ\rangle_k$'s, $k = 0, 1, ..., n-1$. Because of $R\langle\Psi^\circ, \Phi^\circ\rangle_{n-1} = \Psi^\circ(R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1})$, by the definition of Ψ° it follows that $U(x_{i-1}) \in range(e_{R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1}})$ and, by the definition of $e_{R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1}}$, that $\Psi[\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1})]_{i-1} = 1$. By the inductive hypothesis and by the definition of $\Psi[\cdot]_{i-1}$, it follows that $\Psi(R\langle\Phi, \Psi\rangle|_{n-1}) = 1$ (observe: $n = i - 1$). Hence, $R\langle\Psi, \Phi\rangle_{n-1} = 1$. We have thus obtained: $\tau(\sigma) = \tau(\sigma^-)R\langle\Psi, \Phi\rangle_{n-1}$. It follows by the inductive hypothesis that $\tau(\sigma) \sqsubseteq R\langle\Psi, \Phi\rangle|_n$.

Case 3: $\tau(\sigma) \neq \tau(\sigma^-)$ and $\tau(\sigma)_{i-1} = 0$. By the definition of τ, $\tau(\sigma)_{i-1} = 0$ implies $U(x_{i-1}) \notin range(\sigma)$, that is $U(x_{i-1})$ is none of the $R\langle\Psi^\circ, \Phi^\circ\rangle_k$'s, $k \leq n-1$. Then, by the definition of Ψ°, $R\langle\Psi^\circ, \Phi^\circ\rangle_{n-1} = \neg U(x_{i-1})$ and $\neg U(x_{i-1}) \in range(e_{R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1}})$, thus $\Psi[\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{n-1})]_{i-1} = 0$. By the inductive hypothesis and by the definition of $\Psi[\cdot]_{i-1}$, it follows that $\Psi(R\langle\Phi, \Psi\rangle|_{n-1}) = 0$. Hence, $R\langle\Psi, \Phi\rangle_{n-1} = 0$. We have thus established: $\tau(\sigma) = \tau(\sigma^-)R\langle\Psi, \Phi\rangle_{n-1}$. By the inductive hypothesis, it follows $\tau(\sigma) \sqsubseteq R\langle\Psi, \Phi\rangle|_n$. This conclude the proof of (18)a.

(18)b By induction on n. For $n = 0$ the claim holds for every $m \in N$. For $n > 0$, assume that for some $m \in N$, $R\langle\Psi, \Phi\rangle|_{n-1} \sqsubseteq \tau(R\langle\Psi^\circ, \Phi^\circ\rangle|_m)$ and $R\langle\Phi, \Psi\rangle|_{n-1} = \tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})$ (inductive hypothesis). For short, let $\sigma = R\langle\Psi^\circ, \Phi^\circ\rangle|_{m+1}$ and $i = length(\tau(\sigma))$. We prove that $R\langle\Psi, \Phi\rangle|_n \sqsubseteq \tau(\sigma)$. Observe that $R\langle\Psi, \Phi\rangle|_n = R\langle\Psi, \Phi\rangle|_n R\langle\Psi, \Phi\rangle_{n-1}$. We consider two cases:

Case 1: $R\langle\Psi, \Phi\rangle_{n-1} = 1$. Then by the definition of $R\langle\Psi, \Phi\rangle$, $\Psi(R\langle\Phi, \Psi\rangle|_{n-1}) = 1$. By the inductive hypothesis, it follows that $\Psi(\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})) = 1$. Observe that $length(\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})) = n-1$, thus by the definition of $\Psi[\cdot]_n$, $\Psi[\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})]_n = 1$. This means that $U(x_n) \in range(e_{R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1}})$; by the definition of Ψ° it follows that $U(x_n) = \Psi^\circ(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})$, namely, $U(x_n) = R\langle\Psi^\circ, \Phi^\circ\rangle_m$. Then $U(x_n) \in range(\sigma)$ and, by the definition of τ, $\tau(\sigma)_n = 1$. We have thus obtained: $R\langle\Psi, \Phi\rangle_{n-1} = \tau(\sigma)_n$. Since $R\langle\Psi, \Phi\rangle|_n = R\langle\Psi, \Phi\rangle|_n R\langle\Psi, \Phi\rangle_{n-1}$, by the inductive hypothesis it follows that $R\langle\Psi, \Phi\rangle|_n \sqsubseteq \tau(\sigma^-)\tau(\sigma)_n$. Then, by Lemma (17), $R\langle\Psi, \Phi\rangle|_n \sqsubseteq \tau(\sigma)$.

Case 2: $R\langle\Psi, \Phi\rangle_{n-1} = 0$. Similarly to the proof of case 1 we obtain that $\Psi[\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})]_n \neq 1$, with $length(\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})) = n-1$. This means that $U(x_n) \notin range(e_{R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1}})$. By the definition of Ψ° it follows that $U(x_n) \neq \Psi^\circ(R\langle\Phi^\circ, \Psi^\circ\rangle|_{m+1})$, hence $U(x_n) \neq R\langle\Psi^\circ, \Phi^\circ\rangle_m$. Then $U(x_n) \notin range(\sigma)$ (because of $n-1 \leq m$), and, by the definition of τ, $\tau(\sigma)_n = 0$. We have thus obtained: $R\langle\Psi, \Phi\rangle_{n-1} = \tau(\sigma)_n$. Since $R\langle\Psi, \Phi\rangle|_n = R\langle\Psi, \Phi\rangle|_n R\langle\Psi, \Phi\rangle_{n-1}$, by the inductive hypothesis it follows that $R\langle\Psi, \Phi\rangle|_n \sqsubseteq \tau(\sigma^-)\tau(\sigma)_n$. Then, by Lemma (17), $R\langle\Psi, \Phi\rangle|_n \sqsubseteq \tau(\sigma)$. ■

Corollary (19) *For all* $\Psi, \Phi \in \Lambda_0^1$, $R\langle\Psi, \Phi\rangle = R\langle\Phi, \Psi\rangle$ *iff* $R\langle\Psi^\circ, \Phi^\circ\rangle = R\langle\Phi^\circ, \Psi^\circ\rangle$.

The next lemma says that, if $R\langle\Psi, \Phi\rangle$ is not an infinite sequence, neither is $R\langle\Psi^\circ, \Phi^\circ\rangle$.

Lemma (20) *For all* $\Psi, \Phi \in \Lambda_0^1$ *, if for some* $n \in N$ $R\langle\Psi, \Phi\rangle_n$ *is undefined, then for some* $m \in N$ $R\langle\Psi^\circ, \Phi^\circ\rangle_m$ *is undefined.*

Proof: Suppose that $R\langle\Psi, \Phi\rangle_n$ undefined for some $n \in N$ and let $R\langle\Phi, \Psi\rangle|_n = \tau(\sigma)$ for some $\sigma \in SEQ$. By Lemma (18), $\sigma = R\langle\Phi^\circ, \Psi^\circ\rangle|_m$ for some $m \in N$. Then, $R\langle\Psi, \Phi\rangle_n = \Psi(\tau(R\langle\Phi^\circ, \Psi^\circ\rangle|_m))$. By the definition of Ψ°, it follows that $\Psi^\circ(R\langle\Phi^\circ, \Psi^\circ\rangle|_m) = R\langle\Psi^\circ, \Phi^\circ\rangle_m$ is undefined. ■

Proof of Proposition (8): Let $\Psi, \Phi \in \Lambda_0^1$ be given. For the left-to-right direction, suppose that Ψ ηs-coordinates with Φ. Then $R\langle\Psi^\circ, \Phi^\circ\rangle$ and $R\langle\Phi^\circ, \Psi^\circ\rangle$ contain only

finitely many errors. (An error occurs in correspondence to a disagreement between Ψ and Φ, and only finitely many disagreements occur.) After the last error, Ψ° and Φ° start describing their worlds: Ψ°'s describes $\mathcal{N}_{U\langle\Psi,\Phi\rangle}$ by enumerating via l a subset of all literals true in $\mathcal{N}_{U\langle\Psi,\Phi\rangle}$ under the assignment h. By Lemma (18), only true literals become elements of Ψ°'s description, and all true literals enter in it. Since the order of literals is always the one induced by the list l, if a literal β is Ψ°'s description, then sooner or later it is output by Ψ°. Indeed, as soon as β enters in Ψ°'s description, by construction only literals preceeding it in the list l can be output by Ψ°. As soon as all true literals preceeding β are output, so at the next stage β is output by Ψ°. Thus, there $k \in N$, namely, the number of literals occuring in $R\langle\Psi^\circ,\Phi^\circ\rangle$ before the last error, last error included, such that k-$R\langle\Psi^\circ,\Phi^\circ\rangle$ is an environment for $\mathcal{N}_{U\langle\Psi,\Phi\rangle}$ and k-$R\langle\Phi^\circ,\Psi^\circ\rangle$ is an environment for $\mathcal{N}_{U\langle\Phi,\Psi\rangle}$. Since Ψ ηs-coordinates with Φ, the symmetric difference between $U\langle\Psi,\Phi\rangle$ and $U\langle\Phi,\Psi\rangle$ is finite. It follows that both $\mathcal{N}_{U\langle\Psi,\Phi\rangle}$ and $\mathcal{N}_{U\langle\Phi,\Psi\rangle}$ belong to $\mathbf{W}_\Psi \cap \mathbf{W}_\Phi$. Hence, $\langle\Psi^\circ,\mathbf{W}_\Psi\rangle$ and $\langle\Phi^\circ,\mathbf{W}_\Phi\rangle$ sf-coordinate.
For the right-to-left direction, we prove that if Ψ does not ηs-coordinate with Φ then $\langle\Psi^\circ,\mathbf{W}_\Psi\rangle$ and $\langle\Phi^\circ,\mathbf{W}_\Phi\rangle$ do not sf-coordinate. Suppose that Ψ does not ηs-coordinate with Φ. Two cases arise:

Case 1: Either $R\langle\Psi,\Phi\rangle$ or $R\langle\Phi,\Psi\rangle$ is not an infinite sequence. Then the claim follows directly from Lemma (20).

Case 2: $R\langle\Psi,\Phi\rangle$ and $R\langle\Phi,\Psi\rangle$ are infinite sequences and $R\langle\Psi,\Phi\rangle_k \neq R\langle\Phi,\Psi\rangle_k$ for some $k \in N$. First, observe that by Convention (7), $x_0 = \overline{0}$ is the first element of the list l. Then, the construction of $\langle\Psi^\circ,\mathbf{W}_\Psi\rangle$ and $\langle\Phi^\circ,\mathbf{W}_\Phi\rangle$ ensures that in $R\langle\Psi^\circ,\Phi^\circ\rangle$ and $R\langle\Phi^\circ,\Psi^\circ\rangle$ there are infinitely many occurrences of $x_0 = \overline{0}$ and $\neg x_0 = \overline{0}$. Thus, neither of $R\langle\Psi^\circ,\Phi^\circ\rangle$ and $R\langle\Phi^\circ,\Psi^\circ\rangle$ is an environment for any world, not even after deleting finitely many literals. ■

References

[JORS99] S. Jain, D. Osherson, J. Royer, and A. Sharma. *Systems That Learn - An Introduction to Learning Theory, 2nd edition.* The MIT Series in Learning, Development, and Conceptual Change, v. 22. MIT Press, Cambridge, MA, 1999.

[MOss] F. Montagna and D. Osherson. Learning to coordinate: A recursion theoretic perspective. *Synthese*, in press.

[OdJMW97] D. Osherson, D. de Jongh, E. Martin, and S. Weinstein. Formal Learning Theory. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, pages 737–775. Elsevier Science Publishers B.V., 1997.