Group Solvability*

Alessandro Agostini ITC-IRST Trento
Franco Montagna Università di Siena

October 26, 2001

Abstract

We advance and compare four families of **coalitional paradigms** of solvability. A coalitional paradigm is distinguished from a "noncoalitional" paradigm primarily by its focus on what *groups* of agents can achieve, rather than on what *individual* agents can do—even if cooperating. As a criterion of group formation, our models engage a kind of pairwise, context-dependent coordination between knowledge-based "learning agents" able to communicate the *complete & local* meaning of expressions taken from the literals of a common first-order language.

Contents

---

[*] This work is [Ago01, Ch. 4] with some additions and minor changes. We are happy to acknowledge our debt to Dick de Jongh for providing several comments at any stage of the work. In particular, the notion of "success" in the paradigm of π-coordination (Definition (20)) originates from an early discussion the first author have had with Dick on Thursday, April 8 1999 at ILLC in Amsterdam. The improving of the English, Remark (57), Section 7.3, and expecially Theorems (77) and (80) owe a great deal to conversation to and corrections from Dick dated back February 2000. A consequence of a later conversation led to a question answered by Proposition (87). Electronic correspondence to: agostini@irst.itc.it, montagna@unisi.it.

1 Introduction

A coalitional paradigm is distinguished from a "noncoalitional" paradigm primarily by its focus on what *groups* of agents can achieve, rather than on what *individual* agents can do—even if cooperating. In this paper, we advance and compare three families of **coalitional paradigms** of solvability. As a criterion of group formation, the model engages a kind of pairwise, context-dependent coordination, or **cooperation** for short, between knowledge-based "learning agents" able to communicate the complete local meaning of expressions taken from the literals of a common first-order language.

Coordination and cooperation may be related with an equational *slogan*, that we hope of intuitive and immediate help for the reader.

(1) Cooperation = coordination + solvability.[1]

In words, this means that two agents cooperate iff they coordinate in solving some goal-problem. The dimensions on which our division in based are three, namely: absence of coordination, absence of solvability, presence of both. The first dimension is introduced in Section 2; the second dimension is advanced in Section 3 and extended to "groups" of agents along the third dimension in Sections 4 to 7. We conclude in Section 8 with some bibliographical notes and comments.

1.1 A Note on Notation

We denote the set $\{0, 1, 2, ...\}$ of natural numbers by N and the set $\{1, 2, ...\}$ of positive natural numbers by N^+. We denote the usual linearly ordered structure with domain N by ω. Let η be an infinite sequence over $\{0, 1\}$.[2] For $i \in N$, we write $\eta|_i$ for the proper initial sequence of length i in η. We write $|\sigma|$ for the length of a finite sequence, $\emptyset$ for the finite sequence of length zero, σ_i or also $(\sigma)_i$ for the ith element of σ, $0 \leq i < |\sigma|$, $last(\sigma)$ for the last element in σ, and σ^- for the finite sequence obtained from σ by dropping its last (rightmost) element, if $\sigma \neq \emptyset$; σ^- is $\emptyset$, otherwise. Thus, $|\sigma^-| = |\sigma| - 1$, $last(\sigma) = \sigma_{|\sigma^-|}$ and $\sigma = \sigma^- last(\sigma)$. Concatenation of sequences is indicated by juxtaposition, or also by "$\frown$" when juxtaposition might result in unclear expressions, and we won't distinguish notationally between an element and the corresponding unit sequence. Thus $\alpha\tau$ $(\alpha^\frown\tau)$ denotes the sequence with first element α and tail τ. We write $\sigma \sqsubseteq \tau$ if σ is a prefix of τ, that is, $\tau = \sigma\eta$ for some η. In this case, we say that τ **extends** σ. The set of elements in any sequence τ is denoted by $range(\tau)$. Generally, i, j, k, m, n range over N, and σ, τ, η, ν, ζ range over sequences.

2 The $\mathbf{Ex}^{\pi}$-Solvability Paradigm

The paradigm of "pure" solvability on presentation in this section is a model-theoretic paradigm. Our approach follows a first-order perspective, thereby betting that most interesting results will depend on the assumption that knowledge is best represented by formalisms and languages whose expressive power is that of *the decidable fragment of* first-order logic.

2.1 A First-Order Framework

We write $\mathcal{L}_{form}$ to denote a first-order language with equality built up from a (countable, decidable) vocabulary $\mathcal{L}$ consisting of predicates (equality symbol is:

---

[1] Here as in the title of this article and in the related paradigms' names, we prefer to use "solvability" instead of "learning" under the assumption that "learnable" denotes a wider class of objects than "solvable" does. In particular, under such assumption we mean learning to be present in both coordination and solvability, in the intuitive sense that one might learn how to coordinate as well as in problem solving. Apart from this somewhat philosophical remark, in this paper we shall always mean the two terms as synonymous.

[2] By "infinite sequence" we shall always mean an ω-sequence, or a total function defined on N.

$\doteq$) and function symbols of various arities, along with constants symbols. Language $\mathcal{L}_{form}$ contains a countably infinite set $Var = \{v_i \mid i \in N\}$ of variables.[3] The logical symbols of $\mathcal{L}_{form}$ are: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\exists$, $\forall$, all with their standard meaning. The members of $\mathcal{L}_{form}$ are called **formulas**. To avoid ambiguities, we let formulas contain parentheses as auxiliary signs. Generally, P, Q, R range over predicates; φ, ϕ, χ with their eventual decorated versions (say, supscripts and subscripts) range over formulas. We write $\mathcal{L}_{sen}$ to denote the subset of $\mathcal{L}_{form}$ containing no free variables (*i.e.* set of **sentences**). Generally, θ, ϑ, δ along with decorations range over sentences. We write $\mathcal{L}_{basic}$ to denote the subset of $\mathcal{L}_{form}$ containing no logical symbols except "$\neg$" (negation), and such that each formula in the subset contains at most one occurrence of "$\neg$". The members of $\mathcal{L}_{basic}$ are called **basic formulas**, or "literals". Generally, α, β, γ (eventually with super- and subscripts) range over basic formulas.

Despite the fact that our semantic notions are standard, we recall some of them here for future reference. Let $\mathcal{S}$ be a structure that interprets $\mathcal{L}$. Let $\models$ denote the model theoretic concept of truth in a structure. Then $\mathcal{S}$ is a **model** of $\Gamma \subseteq \mathcal{L}_{form}$, and Γ is said to be **satisfiable in** $\mathcal{S}$, if there is an assignment $h : Var \longrightarrow \mathrm{dom}(S)$ with $\mathcal{S} \models \Gamma[h]$. Γ is **satisfiable** if it is satisfiable in some structure. Let $range(\tau)$ denote the set of elements in any sequence τ. Then, in particular, a finite or infinite sequence η on $\mathcal{L}_{form}$ is satisfiable if $range(\eta)$ is satisfiable.

An assignment h to $\mathcal{S}$ is **complete** if h is a mapping onto $\mathrm{dom}(S)$. The **basic diagram** of $\mathcal{S}$ **under** complete assignment h is a 'diagram' in the sense of Abraham Robinson, that is, the subset of $\mathcal{L}_{basic}$ made true in $\mathcal{S}$ via h. We denote the basic diagram of $\mathcal{S}$ under h by $D^b_{\mathcal{S},h}$. The basic diagram of $\mathcal{S}$ is the basic diagram of $\mathcal{S}$ under some complete assignment to $\mathcal{S}$. The class of models of Γ is denoted by $MOD(\Gamma)$.

2.2 Components

The paradigm is now to be introduced in detail by stepping through its basic components.

2.2.1 Learners.

Let SEQ denote the collection of all the *finite* sequences over $\mathcal{L}_{basic}$. A **learner** (or "formal scientist") is any mapping from SEQ to $\mathcal{L}_{sen}$. We call $\mathcal{L}_{sen}$ the **learner language**. A learner thus examines finite pieces of elementary data and—on the basis of such data, advances sentential statements. Given $\sigma \in SEQ$, learner Ψ then applies to the data-stream formalized by σ and, if $\Psi(\sigma)$ is defined, outputs the statement $\Psi(\sigma) \in \mathcal{L}_{sen}$. Intuitively, $\Psi(\sigma)$ might be thought to be the "belief" of Ψ faced with σ. This intuition becomes in fact increasingly important to motivate all our paradigms of cooperation and solvability. In a game-theoretic picture of the learner against Nature, $\Psi(\sigma)$ is also to be interpreted as Ψ's hypothesis on the structure that satisfies $\bigwedge\sigma$ at position $i = |\sigma|$ of the play history. Similarly to bit

---

[3] We shall sometimes write x, y, z in place of v_0, v_1, v_2 respectively.

agents, learners can be partial or total, computable (in a precise sense given by coding *SEQ* and $\mathcal{L}_{sen}$) or uncomputable.

(2) *Example:* Sample kinds of learners $\Psi : SEQ \longrightarrow \mathcal{L}_{sen}$ are the following.

- (a) [**Conservative**]: Let structure $\mathcal{S}$ be given. For all $\sigma \in SEQ$ with $\mathcal{S} \models \bigwedge\sigma[h]$ for some complete assignment h, and for every $\beta \in \mathcal{L}_{basic}$, if $\mathcal{S} \models \Psi(\sigma)$ and $\mathcal{S}$ satisfies $range(\sigma) \cup \{\beta\}$, then $\mathcal{S} \models \Psi(\sigma\beta)$. [The learner keeps his conjecture unless it is contradicted by the next datum.][4]
- (b) [**Computable**]: First, we require language $\mathcal{L}_{form}$ to be "effectively presented", in the sense that the sets of relation, function, and constant symbols, and the sets of variable and logical symbols also are recursive sets. Second, Ψ must be a recursive function.[5]
- (c) [**Fast**]: Let function $f : SEQ \longrightarrow N$ be given. Then Ψ is f-**fast** just in case $O_r(\Psi(\sigma)) \leq O_r(f(\sigma))$ for all $\sigma \in SEQ$, where $O_r(\cdot)$ denotes the "running time of $\cdot$".
- (d) [**Unreliable**]: Let nonempty class $\mathbf{K}$ of structures be given. We require the existence of one structure $\mathcal{S} \notin \mathbf{K}$ such that for all $\sigma \in SEQ$ with $\mathcal{S} \models \bigwedge\sigma[h]$ for some complete assignment h to $\mathcal{S}$, there is $\mathcal{T} \notin \mathbf{K}$ such that $\mathcal{T} \models \Psi(\sigma)$.
- (e) [**Revision-based**]: Let mapping $\dot{+} : \mathcal{L}_{sen} \times SEQ \longrightarrow \mathcal{L}_{sen}$ and $\theta \in \mathcal{L}_{sen}$ be a given. So, $\dot{+}$ takes a sentence $\theta \in \mathcal{L}_{sen}$ and a finite sequence σ on $\mathcal{L}_{basic}$ and returns a locally revised sentence $\theta' = \theta \dot{+} \sigma$, that signifies the impact of σ on θ. Then $\lambda\sigma.\theta \dot{+} \sigma$ is a learner, that we call **revision-based**.[6] One might interpret θ as Ψ's "belief sentence" and $\dot{+}$ as a (simplified) "revision function". For simplicity, we do not restrict revision functions to any particular set of axioms.

To clarify the nature of the reality synthesized by the (finite) "belief" $\Psi(\sigma)$ of a learner Ψ on a finite data-stream σ, we need jumping to the second and third basic components of the paradigm. We do this in turn.

2.2.2 Worlds.

As the first-order perspective we adopted suggests, the underlying reality or "possible worlds" a learner is concerned about are all the countable structures that interpret $\mathcal{L}$. So, unless otherwise stated, the following convention will always be applied in this thesis.

---

[4] Conservativeness may be generalize by considering a nonempty class $\mathbf{K}$ of structures. We say that Ψ is **K-conservative** just in case for all $\mathcal{S} \in \mathbf{K}$, $\sigma \in SEQ$ with $\mathcal{T} \models \bigwedge\sigma[h]$ for some $\mathcal{T} \in \mathbf{K}$ and some complete assignment h to $\mathcal{T}$, and for all $\beta \in \mathcal{L}_{basic}$, if $\mathcal{S} \models \Psi(\sigma)$ and every structure in $\mathbf{K}$ satisfies $range(\sigma) \cup \{\beta\}$, then $\mathcal{S} \models \Psi(\sigma\beta)$.

[5] See for instance [Mil99] for background. We refer the reader to Section 7 for more on computability.

[6] We use "$\lambda\sigma.$" to denote lambda abstraction; that is, $\lambda\sigma.\theta \dot{+} \sigma$ is the function that on any input σ outputs $\theta \dot{+} \sigma$.

(3) CONVENTION: By "structure" we mean any countable **$\mathcal{L}$-structure**, that is, a structure with finite (nonempty) or denumerable domain that interpretes vocabulary $\mathcal{L}$.

Upper case script letters: $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$,..., $\mathcal{S}$, $\mathcal{T}$ are generally used as variables ranging over structures. Given structure $\mathcal{S}$, we denote the domain of $\mathcal{S}$ by $\text{dom}(S)$. We write STR for the class of all structures. Thus, $STR = MOD(\emptyset)$.

(4) *Example:* Some collections $\mathbf{W}$ of structures of interest for our theory develops follows.

(*a*) *Total Orders.* Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R plus equality symbol $\doteq$. Let T be the theory of total orders $\preceq$.[7] Let $x\overline{R}y$ be defined as $xRy \wedge \neg\, x \doteq y$. Let $\theta_0 \in \mathcal{L}_{sen}$ be of the form: $\forall x\,\forall y\,\exists z(x\overline{R}y \rightarrow (x\overline{R}z \wedge z\overline{R}y))$ ("R is dense"); $\theta_1 \in \mathcal{L}_{sen}$ of the form: $\forall x\,\exists y\,y\overline{R}x$ ("there is no least element with respect to R"); and $\theta_2 \in \mathcal{L}_{sen}$ be of the form: $\forall x\,\exists y\,x\overline{R}y$ ("there is no greatest element with respect to R"). Examples of $\mathbf{W}$ are:
$\mathbf{W}_0 = MOD(T \cup \{\theta_0\})$;
$\mathbf{W}_1 = MOD(T \cup \{\theta_1\})$;
$\mathbf{W}_2 = MOD(T \cup \{\theta_2\})$;
$\mathbf{W}_3 = MOD(T \cup \{\theta_0, \theta_1\})$;
$\mathbf{W}_4 = MOD(T \cup \{\theta_0, \theta_2\})$;
$\mathbf{W}_5 = MOD(T \cup \{\theta_1, \theta_2\})$;
$\mathbf{W}_6 = MOD(T \cup \{\neg\,\theta_1, \theta_2\})$;
$\mathbf{W}_7 = MOD(T \cup \{\theta_1 \wedge \neg\,\theta_2, \neg\,\theta_1 \wedge \theta_2\})$.

(*b*) *Strict Linear Orders.* Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let T be the theory of strict linear orders $\prec$.[8] Let $\theta_1 \in \mathcal{L}_{sen}$ of the form: $\exists x\,\forall y\,(x \doteq y \vee xRy)$ ("there is least element with respect to R"); and $\theta_2 \in \mathcal{L}_{sen}$ be of the form: $\exists x\,\forall y\,(x \doteq y \vee yRx)$ ("there is greatest element w.r.t. R").
Set $\mathbf{W} = MOD(T \cup \{\theta_1 \vee \theta_2, \neg(\theta_1 \wedge \theta_2)\})$.

(*c*) Suppose that $\mathcal{L}$ is as in (*b*). Let $\mathcal{K}_0 = \langle\, Q\,, \prec\,\rangle$ and $\mathcal{K}_1 = \langle\, Z\,, \prec\,\rangle$, where Q and Z are the integer and rational numbers, respectively, ordered by strict linear order $\prec$. Set $\mathbf{W} = \{\mathcal{K}_0, \mathcal{K}_1\}$.

(*d*) Suppose that $\mathcal{L}$ consists of a single unary predicate symbol P. Let $\mathcal{K} = \langle\, N\,, \{2n \,|\, n \in N\}\,\rangle$ and $\mathcal{K}_i = \langle\, N\,, \{2n \,|\, n < i\}\,\rangle$ for $i \in N$ ($\mathcal{K}$ interprets P). Set $\mathbf{W} = \{\mathcal{K}\} \cup \{\mathcal{K}_i \,|\, i \in N\}$.

(*e*) Suppose that $\mathcal{L}$ is limited to the binary predicate symbol R. Given $i \in N$, let structure $\mathcal{S}_i$ be defined such that:

i. $R^{\mathcal{S}_0}$ is the relation $\{(i, i+1) \,|\, i \in N\}$, and

---

[7] A **total order** is a connected, reflexive, asymmetric, and transitive binary relation (like $\leq$ on N).

[8] A **strict linear order** is a linearly connected, irreflexive, antisymmetric, and transitive binary relation (like $<$ on N).

ii. for $j > 0$, $R^{\mathcal{S}_j}$ is the finite relation $\{(i, i+1) \,|\, i < j\}$.

Set $\mathbf{W} = \{\mathcal{S}_i \,|\, i \in N\}$,

(*f*) *Infinite sets.* Suppose that $\mathcal{L} = \emptyset$. (Thus, the class of $\mathcal{L}$-structures coincides with the collection of nonempty sets.) Given $n \in N$, let $\theta_n \in \mathcal{L}_{sen}$ be of the form:

$$\exists v_0 \cdots \exists v_n \bigwedge\nolimits_{i<j\leq n} \neg\, v_i \doteq v_j$$

("there are at least $n+1$ elements"). Set $\mathbf{W} = \mathit{MOD}(\{\theta_n \,|\, n \in N\})$.

(*g*) *Algebraically closed fields.* Suppose that $\mathcal{L} = \{\doteq, +, -, \cdot, \overline{0}, \overline{1}\}$. (For economy of symbols, we represent here the function symbols of $\mathcal{L}$ in the way of their standard interpretation functions. $\overline{0}$ and $\overline{1}$ denote 0 and 1, respectively.) Given $n \in N$, let $\theta_n \in \mathcal{L}_{sen}$ be of the form:

$$\forall v_0 \cdots \forall v_n (\neg\, v_n \doteq 0 \rightarrow \exists v_{n+1}\, (v_0 + v_1 \cdot v_{n+1} + \cdots + v_n \cdot v_{n+1}^n \doteq \overline{0}))$$

("there is a root for every polynomial of degree n"). Let T be the theory of fields with respect to $+$, $-$, $\cdot$, $\overline{0}$, $\overline{1}$.
Set $\mathbf{W} = \mathit{MOD}(T \cup \{\theta_n \,|\, n \in N\})$.

(*h*) *Algebraically closed fields of characteristic p.* Suppose that $\mathcal{L}$ and T are as in (*g*). Given p prime, let $\varphi_p \in \mathcal{L}_{sen}$ be of the form:

$$\underbrace{\overline{1} + \overline{1} + \cdots + \overline{1}}_{p \text{ times}} \doteq \overline{0}.$$

Set $\mathbf{W} = \mathit{MOD}(T \cup \{\theta_n \,|\, n \in N\} \cup \{\varphi_p\})$.

2.2.3 Environments.

Learners examine data taken from collection *SEQ*. So, given $\sigma \in \mathit{SEQ}$ it is always possible that $\bigwedge\sigma$, *i.e.* the formula $\sigma_1 \wedge \sigma_2 \wedge \cdots \wedge \sigma_n$ with $n = |\sigma|$, is inconsistent: no structure is a model of it. A direct consequence of the definition of *SEQ* is then that learners' actions are not limited to consistent data. To allow data-streams to be inconsistent is in fact an interesting feature when studying a learner strategy of coordination. With this motivational remark in mind, we return now to technical issues and define an **environment** to be any infinite sequence over $\mathcal{L}_{basic}$. Thus, for all $\sigma \in \mathit{SEQ}$, there is an environment e such that $\sigma = e|_n$ with $n = |\sigma|$. In this technical sense, *SEQ* then denotes the collection of all proper initial sequences of any environment. To consider consistent data-streams, we need to relate them to a structure. We do this in the next definition.

(5) DEFINITION: Let structure $\mathcal{S}$ and complete assignment h to $\mathcal{S}$ be given.

(*a*) An environment e is **for $\mathcal{S}$ via h** just in case $\mathit{range}(e) = \{\beta \in \mathcal{L}_{basic} \,|\, \mathcal{S} \models \beta[h]\}$.

(*b*) An environment e is **for $\mathcal{S}$** just in case e is an environment for $\mathcal{S}$ via some complete assignment.

In other words, an environment for a structure $\mathcal{S}$ via complete assignment h lists the basic diagram of $\mathcal{S}$ under h. Finite initial segments of an environment for a structure thus recapitulate the *consistent* information available to a learner about the underlying structure of evidence—we called it "actual world", at a certain time of observation.

(6) *Example:* Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R plus equality symbol $\doteq$. Let $\langle N , < \rangle$ be the standard structure of natural numbers ordered according to strict linear order $<$. Let $h : Var \longrightarrow \mathrm{dom}(N)$ be defined such that for all $i \in N$, $h(v_i) = i$. Then one environment for $\langle N , < \rangle$ via h is:

$v_0 \doteq v_0 \ \neg v_0 R v_0 \ \ldots$

$\ldots \ v_1 \doteq v_1 \ \neg v_1 R v_1 \ v_0 R v_1 \ \neg v_1 R v_0 \ \ldots$

$\ldots \ v_n \doteq v_n \ \neg v_n R v_n \ v_0 R v_n \ \ldots \ v_{n-1} R v_n \ \neg v_n R v_0 \ \neg v_n R v_1 \ \ldots \neg v_n R v_{n-1} \ \ldots$

Observe:

(7) LEMMA: If some environment is for both structures $\mathcal{S}$ and $\mathcal{T}$, then $\mathcal{S}$ and $\mathcal{T}$ are isomorphic.

(8) LEMMA: All isomorphic pairs $\mathcal{S}$ and $\mathcal{T}$ of structures have an identical set of environments.

The proofs of the lemmas above is an immediate consequence of [Kei77, Prop. 3.2(i)]; the proof of Lemma (7) is also in [OW86, Lem. 3.1A].

2.2.4 Contexts.

By a (learning) **context** we mean a set of sentences, and precisely any subset of $\mathcal{L}_{sen}$. Intuitively, a context π expresses the set of "intelligible hypotheses" that a learner may output *a priori*, with the aim of learning the underlying reality eventually made "actual" by Nature. Here and in the sequel, "intelligible" means interesting in a sense that generally depends on the class of realities under investigation by an individual learner. In other words, given a collection of structures as potential realities, a learner is interested in advancing some hypotheses rather than others on such a collection, and all hypotheses must be drawn from the same learning context. A **question** for learner Ψ is thus, roughly: "What's true in world $\mathcal{W}$ in learning context π?" Clearly, there is generally a number of ways in which a learner may advance interesting guesses. We mention few below.

(9) *Example:* Some learning contexts π:

(*a*) Let $\pi = \{\theta, \neg\theta\}$, for any $\theta \in \mathcal{L}_{sen}$ of any vocabulary $\mathcal{L}$. This context is independent from any underlying structure to be learned. It leads to the paradigms of "truth-detection" (see "Notes" below).

(*b*) Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let T be the theory of strict linear orders $\prec$ with respect to R. Let $\theta_1 \in \mathcal{L}_{sen}$ of the

form: $\exists x \, \forall y \, (x \doteq y \vee xRy)$ ("there is least element with respect to R"); and $\theta_2 \in \mathcal{L}_{sen}$ be of the form: $\exists x \, \forall y \, (x \doteq y \vee yRx)$ ("there is greatest element w.r.t. R").
Set $\pi = \{\theta_1 \vee \theta_2, \neg(\theta_1 \wedge \theta_2)\}$. (Cf. Example (4)(*b*).)

(*c*) Suppose that $\mathcal{L} = \emptyset$. Given $n \in N$, let $\theta_n \in \mathcal{L}_{sen}$ be of the form:

$$\exists v_0 \cdots \exists v_n \bigwedge_{i<j\leq n} \neg \, v_i \doteq v_j$$

("there are at least $n+1$ elements"). Set $\pi = \{\theta_n \mid n \in N\}$. (Cf. Example (4)(*f*).)

Because of the relationship of reality and learning context, we sometimes bring them together in "problems". Precisely, a (learning) **problem** is a pair $(\mathbf{W}, \pi)$, where $\mathbf{W}$ is a (countable) collection of structures, and π is a learning context. Many classes of problems are possible, according to what specific relation is defined between the problem's components. We present few in the next example.

(10) *Example:* Some classes of problems $(\mathbf{W}, \pi)$ are the following.

(*a*) [**Classification Problems**]: Let $\mathbf{W} = \{MOD(\Gamma \cup \{\theta_i\}) \mid 0 \leq i \leq n\}$ for some consistent, possibly empty set $\Gamma \subseteq \mathcal{L}_{sen}$, and let $\pi = \{\theta_0, \theta_1, \ldots, \theta_n\}$ be such that for every $\mathcal{S} \in \mathbf{W}$, there is exactly one $i \in \{0 \cdots n\}$ with $\mathcal{S} \models \theta_i$. These problems admit a unique correct solution, in the strict sense that will be clarified later in this section, when we discuss "success". Because a learner faced with a problem of this form has to "classify" the models of Γ according to the partition π, we sometimes refer to these problems as **classification problems**. Classification problems are common in most model-theoretic paradigms of learning and will be ubiquitous in the thesis.

(*b*) [**One-shot Problems**]: These problems are special classification problems (*a*) for $n = 2$, $\theta_0 = \theta$, and $\theta_1 = \neg \, \theta$. Because of contexts π of this form lead to the paradigms of "truth-detection" (see Example (9)), we call the problems of this form **one-shot problems**, or also "problems of truth-detection".

Given a problem, a question is always, roughly: "Is the problem solvable?", to answer that we have first to say what the question means.

2.2.5 Success.

(11) We say that learner Ψ **solves** problem $(\mathbf{W}, \pi)$ just in case for all $\mathcal{S} \in \mathbf{W}$ and for every environment e for $\mathcal{S}$, there is $\theta \in \pi$ such that:

(*a*) for cofinitely many $n \in N$, $\Psi(e|_n) = \theta$, and
(*b*) $\mathcal{S} \models \theta$.

If some learner solves $(\mathbf{W}, \pi)$, then $(\mathbf{W}, \pi)$ is called **solvable**, otherwise **unsolvable** . We note in passing that whenever condition $(11)(a)$ holds, learner Ψ is said to **stabilize on** e **to** θ.

(12) *Remark:* We write "$\mathbf{W}$ is $\mathbf{Ex}^{\pi}$-solvable" ($\mathbf{Ex}^{\pi}$-unsolvable) in place of "$(\mathbf{W}, \pi)$ is solvable" (unsolvable). We also write "$\mathbf{W} \in \mathbf{Ex}^{\pi}(\Psi)$" in place of "$\Psi$ solves $(\mathbf{W}, \pi)$".

(13) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$, $\mathbf{Ex}^{\pi} = \{\mathbf{W} \mid \mathbf{W} \text{ is } \mathbf{Ex}^{\pi}\text{-solvable}\}$.

This definition completes the formalization of the elements that figure in the game-theoretic picture of scientific inquiry as given through our paradigm's components. What remains to be exhibited are some examples of solvable and unsolvable problems. We present one of each in the next subsection.

2.3 Solvable and Unsolvable Problems

Recall that we use the symbol ω to denote the set of natural numbers N under its standard strict order $<$, whereas ω^* denotes N under the "reverse" strict order $>$.

(14) PROPOSITION: Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let $\pi = \{\theta, \neg\theta\}$, where $\theta \in \mathcal{L}_{sen}$ denotes $\exists x \, \forall y \, xRy \vee \forall x \, \exists y \, xRy$ ("either there is a least point or there is no greatest point"). Then $\mathbf{W} = \{\omega, \omega^*\}$ is $\mathbf{Ex}^{\pi}$-solvable.

Proof: Given any $\sigma \in SEQ$, let $min\, \sigma = v_j$ and $max\, \sigma = v_k$, where:
$j = min_{i \in N} \{\text{For all } v \in Var,\ vRv_i \notin range(\sigma)\}$, and
$k = min_{i \in N} \{\text{For all } v \in Var,\ v_iRv \notin range(\sigma)\}$.
We describe a learner Ψ that $\mathbf{Ex}^{\pi}$-solves $\mathbf{W}$ as follows.

(15) For all $\sigma \in SEQ$ with $|\sigma| > 0$,

- (a) if $min\, \sigma \neq min\, \sigma^-$ then $\Psi(\sigma) = \neg\theta$;
- (b) if $min\, \sigma = min\, \sigma^-$ and $max\, \sigma \neq max\, \sigma^-$, then $\Psi(\sigma) = \theta$;
- (c) if $min\, \sigma = min\, \sigma^-$ and $max\, \sigma = max\, \sigma^-$, then $\Psi(\sigma) = \Psi(\sigma^-)$.

[In case (c) we say that there is no Ψ **mind change**.]

From (15) it follows that for all structures $\mathcal{S}$ in $\mathbf{W}$ and for every environment e for $\mathcal{S}$, either $\Psi(e|_k) = \theta$ or $\Psi(e|_k) = \neg\theta$ for cofinitely many $k \in N$, since after a finite number of steps, Ψ guesses the right sentence with respect to structure $\mathcal{S}$, and then no Ψ mind changes occur. ■

The next proposition shows an example of unsolvable problem.

(16) PROPOSITION: Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let T be the theory of strict total orders (with respect to R). Let $\pi = \{\theta, \neg\theta\}$, where $\theta \in \mathcal{L}_{sen}$ denotes $\forall x \, \forall y \, (xRy \rightarrow \exists z (xRz \wedge zRy))$ ("R is dense"). Then $\mathbf{W} = MOD(T)$ is $\mathbf{Ex}^{\pi}$-unsolvable.

Proof: See [MO98, Ex. 3.(44)]. ■

We are now ready to present a complete, finitary characterization of $\mathbf{Ex}^{\pi}$-solvability for classification problems. Recall that by the arithmetical hierarchy, the class of Σ_2^0-sentences, or "$\exists\forall$-sentences", is defined as the smallest subset of $\mathcal{L}_{sen}$ which contains the universal sentences ("Π_1^0-sentences") and is closed under propositional connectives and adding existential quantifiers at the front.

(17) THEOREM: [MO98] Let $\mathbf{W} = \{MOD(\Gamma \cup \{\theta_i\}) \,|\, 0 \leq i \leq n\}$ for some consistent, possibly empty set $\Gamma \subseteq \mathcal{L}_{sen}$, and let $\pi = \{\theta_0, \theta_1, \ldots, \theta_n\}$ be such that for every $\mathcal{S} \in \mathbf{W}$, there is exactly one $i \in \{0 \cdots n\}$ with $\mathcal{S} \models \theta_i$. Then $\mathbf{W}$ is $\mathbf{Ex}^{\pi}$-solvable if and only if for every $i \in \{0 \cdots n\}$, θ_i is equivalent in Γ to an $\exists\forall$-sentence.

The proof is a straightforward adaptation of Martin and Osherson's proof to our notation. The interested reader is referred to [MO98, Thm. 3.(55)].

3 Cooperation

To extend the paradigm of $\mathbf{Ex}^{\pi}$-solvability to cooperative problem solving, agents have to coordinate in some way. In the following, we advance and discuss a new paradigm of coordination suitable for investigating cooperative solvability. As it serves this aim, we call such a model: **cooperation**. In cooperation, the learner has not only to guess which structure the environment played by the other agent is for, but also which of the sentences in a given "context" is true in that structure. If either the environments played by the agents differ or an agent does not guess the right sentence, an agent is faced with two alternative choices: (a) he can start changing his environment; (b) he can try to convince the other agent to change her environment. In both cases coordination is relevant to agents' choices.

Cooperation, or "contextual coordination", is based on the following components.

(18) **Components of our models of cooperation:**

(*a*) a set of agents;

(*b*) a set of possible realities, or "worlds";

(*c*) for each world, a set of interaction-dynamics, or "enumerations", which provide information about the world;

(*d*) a learning context;

(*e*) a success criterion that stipulates the conditions under which the agents can be said to coordinate in a learning context (social domain);

(*f*) a success criterion that stipulates the conditions under which the agents can be said to learn a class of realities in a learning context (natural domain).

We note in passing that in the next subsection, where we state the basic element of the paradigm of contextual coordination formally, we do not consider the component (*f*) in the list (18) above, as it is an element that does not concern cooperation in its simplest meaning of "coordination without solvability". We will complete formalizing the list in Section 4 below , however, where the present paradigm is proved to be equivalent to $\mathbf{Ex}^{\pi}$-solvability on the special case of classification problems.[9]

3.1 Components

Agents are the leading concepts of the new paradigm. Once we have defined what is an "agent" in a contextual perspective, the other components will follow in quite natural way.

3.1.1 Agents.

An agent in a contextual coordination paradigm, or **learning agent**, is a pair $\langle \Psi , \mathbf{A} \rangle$, where Ψ is any mapping from *SEQ* to $\mathcal{L}_{basic} \times \mathcal{L}_{sen}$, and $\mathbf{A}$ is a nonempty (countable) class of structures. Thus, $\Psi(\sigma) = \langle\, (\Psi(\sigma))_0 , (\Psi(\sigma))_1 \,\rangle$ for all $\sigma \in$ *SEQ*.[10] We say that $(\Psi(\sigma))_0$ is Ψ's **action** (on σ), and that $(\Psi(\sigma))_1$ is Ψ's **guess** (on σ). Of the two components of any learning agent, the first is called **communication function** or "ability", while the second component is called **background world**. We say that $\langle \Psi , \mathbf{A} \rangle$ (or also Ψ) is **based on A**, Ψ is **of** $\langle \Psi , \mathbf{A} \rangle$, and $\langle \Psi , \mathbf{A} \rangle$ **has** Ψ. If Ψ is total on some $I \subseteq$ *SEQ*, we say that Ψ is a **strategy in** I. We say that Ψ is a **strategy** if Ψ is a strategy in *SEQ*. We write $\mathbf{A}_{\Psi}$ for $\mathbf{A}$ and $\Psi_{\mathbf{A}}$ for Ψ (or also $\langle \Psi , \mathbf{A} \rangle$) just in case Ψ is based on $\mathbf{A}$. We refer to $\mathcal{L}$ as the agent's vocabulary, to $\mathcal{L}_{basic}$ as the agent's **external**, or "behavioral" language, and to $\mathcal{L}_{sen}$ as the agent's **internal**, or "cognitive" language. learning agents might have partial or total, computable or noncomputable ability. Learning agents can also be based on a computable background world. We write Λ^l for the class of all learning agents and $\Lambda^l(\mathbf{A})$ for the class of all the learning agents based on $\mathbf{A}$.

3.1.2 Worlds.

Realities of the list (18) are limited to agents' preferences and beliefs. Since the collection of all learning agents involves the class of all (nonempty) structures that interpret $\mathcal{L}$, and a "world" is any of such structures, it follows that by "world" in the present paradigm we mean any structure contained in a background world of some learning agent.

3.1.3 Dynamics.

Since worlds in the present paradigm are limited to structures contained in some agent' background world, the interactions in (18)(c) are among "active" agents.[11]

---

[9] Classification problems are introduced and discussed in Example (10).

[10] The notation $()_0$ and $()_1$ here and in similar circumstances should not be confused with the usage of $(\sigma)_i$ we make to denote the ith element of a sequence σ of appropriate length.

[11] In contrast, an example of "passive" agent is Nature in the paradigm of $\mathbf{Ex}^{\pi}$-solvability.

So, we now limit our paradigm to the interaction among learning agents. For simplicity, we restrict attention to dynamics based on simultaneous moves, that is, the agents make decisions at the same time, and to pairwise communication only, that is, interactions involve just two agents. Our presentation may be generalized to sequential moves and full communication with minor, technical changes.

Let finite or infinite sequence ζ of length $n > k$, $k \in N$ be given. $\zeta[k]$ denotes the finite sequence $\langle \zeta_0 \cdots \zeta_k \rangle$ and ${}_k|\zeta$ denotes the sequence obtained from ζ by deleting its first $k+1$ elements.

(19) Definition: Let learning agents $\langle \Psi, \mathbf{A} \rangle$ and $\langle \Phi, \mathbf{B} \rangle$ be given.

(a) The **interaction sequence** (or "play") of $\langle \Psi, \mathbf{A} \rangle$ and $\langle \Phi, \mathbf{B} \rangle$ is the infinite sequence

$$D_{\Psi,\Phi} = (\langle \overline{\Psi}_i, \overline{\Phi}_i \rangle \ : \ i \in N),$$

where $\overline{\Psi}_i$ is the ith move of Ψ and $\overline{\Phi}_i$ is the ith move of Φ, defined by induction as follows.

i. $\overline{\Psi}_{00} = (\Psi(\emptyset))_0$ and $\overline{\Psi}_{10} = (\Psi(\emptyset))_1$;
$\overline{\Phi}_{00} = (\Phi(\emptyset))_0$ and $\overline{\Phi}_{10} = (\Phi(\emptyset))_1$.
ii. $\overline{\Psi}_{0n+1} = (\Psi(\overline{\Phi}_0[n]))_0$ and $\overline{\Psi}_{1n+1} = (\Psi(\overline{\Phi}_0[n]))_1$;
$\overline{\Phi}_{0n+1} = (\Phi(\overline{\Psi}_0[n]))_0$ and $\overline{\Phi}_{1n+1} = (\Phi(\overline{\Psi}_0[n]))_1$.

(b) Let $k \in N$ be given. The **interaction sequence** of $\langle \Psi, \mathbf{A} \rangle$ and $\langle \Phi, \mathbf{B} \rangle$ **starting at** k is the infinite sequence

$${}^kD_{\Psi,\Phi} = (\langle {}_k|\overline{\Psi}_i, {}_k|\overline{\Phi}_i \rangle \ : \ i \in N),$$

where ${}_k|\overline{\Psi}_i$ is the ith element in ${}_k|\overline{\Psi}$ and ${}_k|\overline{\Phi}_i$ is the ith element in ${}_k|\overline{\Phi}$.

We then say that ${}_k|\overline{\Psi}_i$ is the ith move of Ψ starting at k and ${}_k|\overline{\Phi}_i$ is the ith move of Φ starting at k. Note that $D_{\Psi,\Phi} = \langle \overline{\Psi}_0, \overline{\Phi}_0 \rangle ^\frown {}^0D_{\Psi,\Phi}$. Also note that definition of interaction sequence ${}^kD_{\Psi,\Phi}$ depends only on the agents' abilities Ψ and Φ; no background worlds are involved. We shall see later in this section that learning agents' background worlds are nevertheless relevant to determine the criterion of contextual coordination success.

The **response sequence**

$$R(\Psi, \Phi) = (\overline{\Psi}_i \ : \ i \in N)$$

is the (finite or infinite) sequence of moves by learning agent with ability Ψ in response to learning agent with ability Φ, and the **response sequence**

$$R(\Phi, \Psi) = (\overline{\Phi}_i \ : \ i \in N)$$

is the sequence of moves by Φ in response to Ψ. Notice that $R(\Psi, \Phi)$ is finite iff at any interaction step $i \in N$, $\Psi(\overline{\Psi}|_i)$ or $\Psi(\overline{\Phi}|_i)$ is undefined. If this is the case, it is immediate to verify that $R(\Phi, \Psi)$ is finite also. More precisely, $\overline{\Psi}_0$, $\overline{\Psi}_1$, ${}_k|\overline{\Psi}_0$, ${}_k|\overline{\Psi}_1$

are infinite sequences if and only if so are, respectively, $\overline{\Phi}_0$, $\overline{\Phi}_1$, ${}_k|\overline{\Phi}_0$, ${}_k|\overline{\Phi}_1$. We write $R(\Psi_0, \Phi)$, $R(\Psi_1, \Phi)$, ${}_k|R(\Psi_0, \Phi)$ and ${}_k|R(\Psi_1, \Phi)$ to denote $\overline{\Psi}_0$, $\overline{\Psi}_1$, ${}_k|\overline{\Psi}_0$ and ${}_k|\overline{\Psi}_1$, respectively. Consistently with the notation adopted, ${}_k|R(\Psi, \Phi)|_n$ denotes the finite initial sequence in ${}_k|R(\Psi, \Phi)$ of length n, and ${}_k|R(\Psi, \Phi)_n$, or also $({}_k|R(\Psi, \Phi))_n$ denotes the nth element of ${}_k|R(\Psi, \Phi)$. Thus, ${}_k|R(\Psi, \Phi)|_{n+1} = {}_k|R(\Psi, \Phi)|_n \overline{\Psi}_n$ and $R(\Psi, \Phi) = \overline{\Psi}_0{}^\frown{}_0|R(\Psi, \Phi)$.

3.1.4 Contexts.

Contexts of the list (18) are called **coordination contexts** henceforth, since they are at the service of coordination only. Recall that the kind of cooperation we are presenting is in fact coordination; no solvability is involved here. The formal definition of a coordination context is identical to the definition of a "natural" context of the kind that we have already encountered when discussing on pure solvability (see Section 2.2), namely, any subset of $\mathcal{L}_{sen}$. Coordination contexts determine what agents think to be an "interesting" context of coordination. Given context π and a class of learning agents, a learning agent Ψ is then interested in coordinating with each member of the class in π rather than other "less interesting" contexts. A question for Ψ is, roughly: "What's a class of learning agents to coordinate with in context π?"

3.1.5 Success.

Learning agents have to stabilize in the limit to an environment of a structure in their own background world, eventually after a finite number of failures. Each environment is an appropriate response each agent acts out by his basic, "behavioral" language. On the other hand, a second action concerns success, namely, the sentence each agent chooses to match the other agents' structure. It is for this second component that we qualified this paradigm of "contextual" coordination; thus, coordination depends on a context. According to the paradigm, the agents can restart their mutual interaction finitely many often, but after the last disagreement they must eventually coordinate in the following sense.

(20) Definition: Let $\pi \subseteq \mathcal{L}_{sen}$ and learning agents $\langle \Psi, \mathbf{A} \rangle$, $\langle \Phi, \mathbf{B} \rangle$ be given. We say that $\langle \Psi, \mathbf{A} \rangle$ **π-coordinates with** $\langle \Phi, \mathbf{B} \rangle$ (written: $\Psi \rightleftharpoons_\pi \Phi$) just in case for some $s, t \in N$, there is $\theta \in \pi$ such that:

(a) ${}_s|\overline{\Psi}_0$ is an environment for some $\mathcal{A} \in \mathbf{A}$,
(b) ${}_t|\overline{\Phi}_0$ is an environment for some $\mathcal{B} \in \mathbf{B}$,
(c) for cofinitely many $n \in N$, $({}_s|\overline{\Psi}_1)_n = ({}_t|\overline{\Phi}_1)_n = \theta$, and
(d) $\mathcal{A} \models \theta$ and $\mathcal{B} \models \theta$.

In this case, θ is said to be the **coordination sentence**.

In case of clause (a), we say that Ψ **enumerates with** Φ ${}_s|\overline{\Psi}_0$; similarly, in case of clause (b), we say that Φ **enumerates with** Ψ ${}_t|\overline{\Phi}_0$. In case of clause (c), we say that Ψ **guesses with** Φ (Φ **guesses with** Ψ) θ. Observe that relation

$\rightleftharpoons_\pi$ is symmetric, that is, for all learning agents $\langle \Psi , \mathbf{A} \rangle$, $\langle \Phi , \mathbf{B} \rangle$, if $\Psi \rightleftharpoons_\pi \Phi$ then $\Phi \rightleftharpoons_\pi \Psi$. This seems to be a natural property of coordination as meant in intuitive sense. However, π-coordination is neither reflexive nor transitive. To see the lack of reflexivity in general think, for example, to a learning agent whose action on any input is inconsistent. Of course, reflexivity holds if we limit the collection of learning agents to a special class of "self-centered" agents; for a sound definition of such a class see [Ago01, Sec. 3.4].

According to Definition (20), given a coordination context π learning agents must coordinate twice. First, the agents eventually stabilize on an environment for a structure $\mathcal{S}$ in their own background world. This step is done by the agents through their external, behavioral ability component. Second, both agents eventually stabilize to a sentence taken from π; such sentence must be true in $\mathcal{S}$. This second step is done by the agents through their internal, cognitive ability's component. Both steps are processes in the limit, which eventually start after a finite number of disagreements.

We now conclude to describe "success" by giving some game-theoretical terminology; this may help the reader relate some of the fundamental concepts of game theory to the present inductive setting. Let $\pi \subseteq \mathcal{L}_{sen}$, nonempty sets $\Sigma(\mathbf{A})$ and $\Sigma(\mathbf{B})$ of learning agents based on, respectively, class $\mathbf{A}$ and class $\mathbf{B}$ of structures be given. Let $\{D_{\Psi,\Phi}\} = \{D_{\Psi,\Phi} \,|\, \langle \Psi , \mathbf{A} \rangle \in \Sigma(\mathbf{A}), \langle \Phi , \mathbf{B} \rangle \in \Sigma(\mathbf{B})\}$. A **contextual coordination game**, or "π-coordination game", is a triple $\langle\, \Sigma(\mathbf{A}), \Sigma(\mathbf{B})\, , \{D_{\Psi,\Phi}\}\, \rangle$. An **equilibrium** of a contextual coordination game $\langle\, \Sigma(\mathbf{A}), \Sigma(\mathbf{B})\, , \{D_{\Psi,\Phi}\}\, \rangle$ is a triple $(\mathcal{A}, \mathcal{B}, \theta)$ that satisfies conditions (a)–(d) of Definition (20) for some pair $\langle \Psi , \mathbf{A} \rangle \in \Sigma(\mathbf{A})$ and $\langle \Phi , \mathbf{B} \rangle \in \Sigma(\mathbf{B})$. An equilibrium is not necessarily unique and does not depend on game stages. A **partial equilibrium** is a triple $(\mathcal{A}', \mathcal{B}', \theta)$, where θ is that of Definition (20), and $\mathcal{A}' \in \mathbf{A}$, $\mathcal{B}' \in \mathbf{B}$ are such that ${}_s|\overline{\Psi}|_n$ is satisfiable in $\mathcal{A}'$, ${}_t|\overline{\Phi}|_n$ is satisfiable in $\mathcal{B}'$ for some $n \in N$, and both $\mathcal{A}'$ and $\mathcal{B}'$ are models of θ. Again, a partial equilibrium is not unique, but it depends on game stages.

3.2 Examples

Here is an example of success in a contextual coordination paradigm.

(21) LEMMA: Let $\pi \subseteq \mathcal{L}_{sen}$ and nonempty classes $\mathbf{A}$, $\mathbf{B}$ of structures be given. Suppose that for some $\theta \in \pi$ $\mathbf{A} \cap MOD(\theta) \neq \emptyset$ and $\mathbf{B} \cap MOD(\theta) \neq \emptyset$. Then there are abilities Ψ, Φ such that $\langle \Psi , \mathbf{A} \rangle$ π-coordinates with $\langle \Phi , \mathbf{B} \rangle$.

Proof: Let $e^{\mathcal{A}}$ be an environment for some $\mathcal{A} \in \mathbf{A} \cap MOD(\theta)$, and let $e^{\mathcal{B}}$ be an environment for some $\mathcal{B} \in \mathbf{B} \cap MOD(\theta)$. Define $\Psi, \Phi : SEQ \longrightarrow \mathcal{L}_{basic} \times \mathcal{L}_{sen}$ as follows. For all $\sigma \in SEQ$ with $|\,\sigma\,| = i$, define $\Psi(\sigma) = \langle\, e_i^{\mathcal{A}} \,, \theta \,\rangle$ and $\Phi(\sigma) = \langle\, e_i^{\mathcal{B}} \,, \theta \,\rangle$. Clearly, $\langle \Psi , \mathbf{A} \rangle$ $\{\theta\}$-coordinates with $\langle \Phi , \mathbf{B} \rangle$. ■

Another example of π-coordination relies on nonempty collections of learning agents. We call these collections: **coalitions**. So, the next proposition shows that there are two infinite coalitions such that any pair of agents within a coalition coordinates, and such that whenever an agent coordinates with at least one member of one coalition he cannot coordinate with any member of the other coalition.

(22) Proposition: (Contextual Uncooperativeness Theorem) Let $\pi \subseteq \mathcal{L}_{sen}$ be given. Suppose that at least one $\theta \in \pi$ is satisfiable. Then there are two infinite coalitions Σ_0 and Σ_1 of learning agents with the following properties.

(a) Every learning agent in Σ_0 π-coordinates with Σ_0.
(b) Every learning agent in Σ_1 π-coordinates with Σ_1.
(c) Every learning agent that π-coordinates with Σ_i, $i = 0, 1$, does not π-coordinate with any member of Σ_{1-i}.

Proof: Let structure $\mathcal{S}$ be such that $\mathcal{S} \models \theta$. For $i = 0, 1$, let Σ_i be the class of learning agents $\langle \Psi, \{\mathcal{S}\} \rangle$ such that:

(a) $(\Psi(\emptyset))_0 = (v_i \doteq v_i)$.
(b) For all $\sigma \in SEQ$, $(\Psi(\sigma))_1 = \theta$.
(c) For every learning agent $\langle \Phi, \mathbf{B} \rangle$, if $(\Phi(\emptyset))_0 = (v_i \doteq v_i)$ then $R(\Psi_0, \Phi)$ is an environment for $\mathcal{S}$. Otherwise, for all $\sigma \in SEQ$ with $|\sigma| > 0$, $(\Psi(\sigma))_0 = \neg(v_0 \doteq v_0)$.

Clearly, Σ_0 and Σ_1 satisfy the properties of the proposition. ■

Note that the proposition holds for finite classes, as is easy to verify by taking a finite subclass of each class in the proposition.

A third example relies on a special class of learning agents—we call them "centered." This class of agents will also be useful elsewhere (cf. Definition (35) below). A preliminary definition is needed.

(23) Definition: Let $\sigma \in SEQ$ and learning agent $\langle \Psi, \mathbf{A} \rangle$ be given. We define the **communication sequence** $\overline{(\Psi(\sigma))_0} \in SEQ$ by induction on the length of σ as follows. *Base*: $\overline{(\Psi(\emptyset))_0} = (\Psi(\emptyset))_0$. Suppose that $\overline{(\Psi(\tau))_0}$ is defined for $\tau \in SEQ$. Given $\beta \in \mathcal{L}_{basic}$, define $\overline{(\Psi(\tau\beta))_0} = \overline{(\Psi(\tau))_0}{}^\frown(\Psi(\tau\beta))_0$.

Note that when $\overline{(\Psi(\sigma))_0}$ is defined, $|\overline{(\Psi(\sigma))_0}| > 0$. Also note that the definition does not depend on the learning agent's background world $\mathbf{A}$.

Let $SEQ_\Psi = \{\overline{(\Psi(\sigma))_0} \mid \sigma \in SEQ\}$ denote the collection of all the finite sequences produced on any input by a learning agent with ability Ψ.

(24) Definition: Let $\pi \subseteq \mathcal{L}_{sen}$, learning agent $\langle \Psi, \mathbf{A} \rangle$ and nonempty class $\mathbf{K}$ of structures be given.

(a) $\langle \Psi, \mathbf{A} \rangle$ is **K-centered in** π (shorten: $\mathbf{K}^\pi$-centered) just in case for all strategies Φ in SEQ_Ψ, there are $\mathcal{S} \in \mathbf{K}$ and $\theta \in \pi$ such that:
  i. $R(\Psi_0, \Phi)$ is an environment for $\mathcal{S}$;
  ii. for all but finitely many $n \in N$, $(R(\Psi_1, \Phi))_n = \theta$, and
  iii. $\mathcal{S} \models \theta$.

(b) $\langle \Psi , \mathbf{A} \rangle$ is **weakly K-centered in** π just in case for all strategies Φ in SEQ_{Ψ}, there are $t \in N$, $\mathcal{S} \in \mathbf{K}$ and $\theta \in \pi$ such that:

i. ${}_{t}|R(\Psi_0, \Phi)$ is an environment for $\mathcal{S}$;
ii. for all but finitely many $n \in N$, $({}_{t}|R(\Psi_1, \Phi))_n = \theta$, and
iii. $\mathcal{S} \models \theta$.

In this case, Ψ is said to be ${}_{t}$**K-centered in** π.

(c) $\langle \Psi , \mathbf{A} \rangle$ is **(weakly) self-centered in** π (shorten: $self^{\pi}$-centered) just in case $\langle \Psi , \mathbf{A} \rangle$ is (weakly) $\mathbf{A}$-centered in π.

We also say that learning agent $\langle \Psi , \mathbf{A} \rangle$ is **world-centered** if $\langle \Psi , \mathbf{A} \rangle$ is (weakly) $\mathbf{W}$-centered in π for some $\mathbf{W}$ and some π. For any singleton class $\{\mathcal{K}\}$ of structures, we sometimes write "$\mathcal{K}$-centered" in place of "$\{\mathcal{K}\}$-centered." Observe:

(25) Lemma: Let $\pi \subseteq \mathcal{L}_{sen}$ be given. For all learning agents $\langle \Psi , \mathbf{A} \rangle$, $\langle \Phi , \mathbf{B} \rangle$,

(a) if $\langle \Psi , \mathbf{A} \rangle$ is weakly $self^{\pi}$-centered and total, then $\Psi \rightleftharpoons_{\pi} \Psi$;
(b) if $\Psi \rightleftharpoons_{\pi} \Phi$ then $\Phi \rightleftharpoons_{\pi} \Psi$.

However, $\rightleftharpoons_{\pi}$ is not an equivalence relation even if restricted to agents which are $self^{\pi}$-centered, since it is not transitive. Indeed:

(26) *Example:* Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let $\pi = \{\theta_0, \theta_1\}$, where $\theta_0 \in \mathcal{L}_{sen}$ denotes $\exists x \forall y\, xRy \vee \forall x \exists y\, xRy$ ("there is either a least point or no greatest point with respect to R") and $\theta_1 \in \mathcal{L}_{sen}$ denotes $\exists x \forall y\, yRx \vee \forall x \exists y\, yRx$ ("there is either a greatest point or no least point with respect to R"). Let $\mathcal{S}_0 = \langle N , < \rangle$, $\mathcal{S}_1 = \langle Z , < \rangle$ and $\mathcal{S}_2 = \langle Z - N , < \rangle$, where $<$ interprets R. Clearly, $\mathcal{S}_0, \mathcal{S}_1 \models \theta_0$ and $\mathcal{S}_1, \mathcal{S}_2 \models \theta_1$. For $i = 0, 1, 2$, define learning agent $\langle \Phi_i , \{\mathcal{S}_i\} \rangle$ such that:

(a) For $j = 1, 2$, Φ_0 enumerates with Φ_j an environment for $\mathcal{S}_0$ and guesses θ_0.
(b) Φ_1 enumerates with Φ_0 an environment for $\mathcal{S}_1$ and guesses θ_0; Φ_1 enumerates with Φ_2 an environment for $\mathcal{S}_1$ and guesses θ_1.
(c) For $j = 0, 1$, Φ_2 enumerates with Φ_j an environment for $\mathcal{S}_2$ and guesses θ_1.

It is easy to verify that $\langle \Phi_0 , \{\mathcal{S}_0\} \rangle$ π-coordinates with $\langle \Phi_1 , \{\mathcal{S}_1\} \rangle$ and that $\langle \Phi_1 , \{\mathcal{S}_1\} \rangle$ π-coordinates with $\langle \Phi_2 , \{\mathcal{S}_2\} \rangle$. On the other hand, from (a) with $j = 2$ and (c) with $j = 0$ it follows that $\langle \Phi_0 , \{\mathcal{S}_0\} \rangle$ does not π-coordinate with $\langle \Phi_2 , \{\mathcal{S}_2\} \rangle$ since $\theta_0 \neq \theta_1$.

3.3 Explicit Guesses

The paradigm of π-coordination concerns agents whose input's domain is defined on sequences of basic formulas only; the sentences in π are therefore not in the agents' input. As a consequence, if a learning agent wants to communicate his guesses to an agent, he must use an indirect communication "protocol". For example, the agent could use a coded message by Gödel numbering together with the following strategy: at each even stage of the interaction sequence, guess "$v_n \doteq v_n$" by the external ability and sentence θ by the internal ability, where n is the Gödel number of θ; at each odd stage of the interaction sequence, act and guess as context situation requires.[12] By following this strategy, in this subsection we show that an agent who communicates his guesses to the opponent at each step of the interaction sequence may be simulated by a learning agent that does not manage explicit guesses. On the other hand, it is quite clear that any agent who "hides" his guesses—this is the case of learning agents, may be simulated by an agent that always shows his own guesses. This can be done simply by requiring the agent to ignore in input the guesses of the opponent. It follows that the two paradigms with and without explicit guesses are equivalent. However, the paradigm with no explicit guess is essentially simpler, so that our choice of a paradigm of coordination without explicit guesses as more fundamental is justified in this way.

In the rest of this subsection we state the equivalence of the paradigms with and without explicit guesses formally. For doing this, we modify both the definition of an agent and of the interaction sequence between two agents. For simplicity, we limit the paradigm to pairwaise interactions and simultaneous moves. We are particularly interested in the collection of all the *finite* sequences over $\mathcal{L}_{sen}$. We let *SEN* denote the collection of such finite sequences.

(27) DEFINITION: A **learning agent with explicit guess** is a pair $\langle \Psi , \mathbf{A} \rangle$, where Ψ is any mapping from *SEQ* $\times$ *SEN* to $\mathcal{L}_{basic} \times \mathcal{L}_{sen}$, and $\mathbf{A}$ is a nonempty (countable) class of structures.

Note that for fixed $\tau \in$ *SEN*, $\langle \lambda\sigma.\Psi(\sigma,\tau) , \mathbf{A} \rangle$ is a learning agent. Similarly to learning agents, of the two components of any learning agent with explicit guess, the first is called **communication function** or "ability", while the second component is called **background world**. We say that $\langle \Psi , \mathbf{A} \rangle$ (or also Ψ) is **based on** $\mathbf{A}$, Ψ is **of** $\langle \Psi , \mathbf{A} \rangle$, and $\langle \Psi , \mathbf{A} \rangle$ **has** Ψ. If Ψ is total on some $I \subseteq$ *SEQ* $\times$ *SEN*, we say that Ψ is a **strategy in** I. We say that Ψ is a **strategy** if Ψ is a strategy in *SEQ* $\times$ *SEN*. We write Λ^{lg} for the class of all learning agents with explicit guess.

Let learning agents $\langle \Psi , \mathbf{A} \rangle$, $\langle \Phi , \mathbf{B} \rangle$ with explicit guess be given. The **interaction sequence of** $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ is the infinite sequence

$$D^g_{\Psi,\Phi} = (\langle \overline{\Psi}_i , \overline{\Phi}_i \rangle \ : \ i \in N),$$

where $\overline{\Psi}_i$ is the ith move of Ψ and $\overline{\Phi}_i$ is the ith move of Φ, defined by induction as follows.

i. $\overline{\Psi}_{00} = (\Psi(\emptyset,\emptyset))_0$ and $\overline{\Psi}_{10} = (\Psi(\emptyset,\emptyset))_1$;
$\overline{\Phi}_{00} = (\Phi(\emptyset,\emptyset))_0$ and $\overline{\Phi}_{10} = (\Phi(\emptyset,\emptyset))_1$.

---

[12] For background on arithmetization and Gödel numbers for an arbitrary first-order theory see for instance [Men87, Ch. 3].

ii. $\overline{\Psi}_{0n+1} = (\Psi(\overline{\Phi}_0[n], \overline{\Phi}_1[n]))_0$ and $\overline{\Psi}_{1n+1} = (\Psi(\overline{\Phi}_0[n], \overline{\Phi}_1[n]))_1$;
$\overline{\Phi}_{0n+1} = (\Phi(\overline{\Psi}_0[n], \overline{\Psi}_1[n]))_0$ and $\overline{\Phi}_{1n+1} = (\Phi(\overline{\Psi}_0[n], \overline{\Psi}_1[n]))_1$.

Let $k \in N$ be given. The **interaction sequence** of $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ **starting at** k is the infinite sequence

$$^{k}D^{s}_{\Psi,\Phi} = (\langle\, {}_{k}|\overline{\Psi}_i \,, {}_{k}|\overline{\Phi}_i \,\rangle \,:\, i \in N),$$

where ${}_{k}|\overline{\Psi}_i$ is the ith move of Ψ **after** k and ${}_{k}|\overline{\Phi}_i$ is the ith move of Φ **after** k. The **response sequence**

$$R(\Psi, \Phi) = (\overline{\Psi}_i \,:\, i \in N)$$

is the (finite or infinite) sequence of moves by learning agent with explicit guess and ability Ψ in response to learning agent with explicit guess and ability Φ. The **response sequence** $R(\Phi, \Psi) = (\overline{\Phi}_i \,:\, i \in N)$ is the same sequence with the roles of Ψ and Ψ reversed.

A **paradigm of π-coordination and explicit guesses** is a paradigm of π-coordination such that “agents” and “dynamics” components in the list (18) are learning agents with explicit guess and dynamics are modeled via interaction sequences defined above. We state the following definition of “success”, which is a natural extension of Definition (20) to the present paradigm.

(28) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$ and learning agents $\langle \Psi , \mathbf{A} \rangle$, $\langle \Phi , \mathbf{B} \rangle$ with explicit guess be given. We say that $\langle \Psi , \mathbf{A} \rangle$ **π-coordinates with $\langle \Phi , \mathbf{B} \rangle$ by explicit guesses** (written: $\Psi \rightleftharpoons_{\pi^g} \Phi$) just in case for some $s, t \in N$, there is $\theta \in \pi$ such that:

- (a) ${}_{s}|\overline{\Psi}_0$ is an environment for some $\mathcal{A} \in \mathbf{A}$,
- (b) ${}_{t}|\overline{\Phi}_0$ is an environment for some $\mathcal{B} \in \mathbf{B}$,
- (c) for cofinitely many $n \in N$, $({}_{s}|\overline{\Psi}_1)_n = ({}_{t}|\overline{\Phi}_1)_n = \theta$, and
- (d) $\mathcal{A} \models \theta$ and $\mathcal{B} \models \theta$.

Below is an example of π^g-coordination extended to sets of learning agents. Our example concerns a special set of “self-centered learning agent” with explicit guess, to state which some further terminology is needed.

(29) DEFINITION: Let $\sigma \in SEQ$, $\sigma' \in SEN$ and learning agent with explicit guess $\langle \Psi , \mathbf{A} \rangle$ be given. We define the **communication sequence** $\overline{(\Psi(\sigma, \sigma'))_0} \in SEQ$ by induction on the length of σ and σ' as follows. $\overline{(\Psi(\emptyset, \emptyset))_0} = (\Psi(\emptyset, \emptyset))_0$. Suppose that $\overline{(\Psi(\tau, \eta))_0}$ is defined for $\tau \in SEQ$ and $\eta \in SEN$. Given $\beta \in \mathcal{L}_{basic}$ and $\theta \in \mathcal{L}_{sen}$, define $\overline{(\Psi(\tau\beta, \eta\theta))_0} = \overline{(\Psi(\tau, \eta))_0}{}^\frown (\Psi(\tau\beta, \eta\theta))_0$.

Note that when $\overline{(\Psi(\sigma, \sigma'))_0}$ is defined, $|\,\overline{(\Psi(\sigma, \sigma'))_0}\,| > 0$.

Let $SEQ_\Psi = \{\overline{(\Psi(\sigma, \sigma'))_0} \,|\, \sigma \in SEQ, \sigma' \in SEN\}$ denote the collection of all the finite sequences produced on any input by a learning agent with explicit guess and ability Ψ. Then:

(30) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$ be given. We say that learning agent with explicit guess $\langle \Psi , \mathbf{A} \rangle$ is $self^{\pi}$**-centered** just in case for all strategies Φ in SEQ_{Ψ}, there are $\mathcal{A} \in \mathbf{A}$ and $\theta \in \pi$ such that:

(a) $R(\Psi_0, \Phi)$ is an environment for $\mathcal{A}$;
(b) for all but finitely many $n \in N$, $(R(\Psi_1, \Phi))_n = \theta$, and
(c) $\mathcal{A} \models \theta$.

So a $self^{\pi}$-centered learning agent with explicit guess behaves exactly as a $self^{\pi}$-centered learning agent, that is, enumerates with any other learning agents an environment for some structure $\mathcal{S}$ in the background world the agent is based on, and guesses with the same agent a sentence true in $\mathcal{S}$.

(31) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$. For every nonempty set Σ of $self^{\pi}$-centered learning agents with explicit guess, there is a learning agent with explicit guess who π^g-coordinates with each member of Σ.

Proof: Let Σ be a nonempty set of $self^{\pi}$-centered learning agents with explicit guess. We define a learning agent $\langle \Psi , \mathbf{A} \rangle$ with explicit guess such that:

(a) $\mathbf{A}$ is the set of structures $\mathcal{S}$ such that there are $\langle \Phi , \mathbf{B} \rangle, \langle \Phi' , \mathbf{B}' \rangle \in \Sigma$ and $\theta \in \pi$ with $\mathcal{S} \models \theta$ such that $R(\Phi_0, \Phi')$ is an environment for $\mathcal{S}$ and for all $n \in N$, $(R(\Phi_1, \Phi'))_n = \theta$. Observe that $\mathbf{A}$ is nonempty.
(b) $\Psi(\emptyset, \emptyset) = \langle v_0 \doteq v_0 , \theta' \rangle$, for some $\theta' \in \pi$ such that $\mathcal{A} \models \theta'$ with $\mathcal{A} \in \mathbf{A}$. For all $\sigma \in SEQ$ and for all $\sigma' \in SEN$ with $|\sigma| = |\sigma'|$, if $|\sigma| > 0$ then $\Psi(\sigma, \sigma') = \langle \sigma_{|\sigma|-1} , \sigma'_{|\sigma|-1} \rangle$.

[Thus, when playing with learning agents $\langle \Phi , \mathbf{B} \rangle \in \Sigma$, Ψ starts by moving "safe" and then copies Φ's last move one step later.] Then, $R(\Psi_0, \Phi) = v_0 \doteq v_0{}^\frown R(\Phi_0, \Psi)$ and $R(\Psi_1, \Phi) = \forall v_0\, (v_0 \doteq v_0)^\frown R(\Phi_1, \Psi)$. Since $\langle \Phi , \mathbf{B} \rangle$ is $self^{\pi}$-centered, $R(\Psi_0, \Phi)$ is an environment for some structure $\mathcal{A} \in \mathbf{A}$ and there is $\theta \in \pi$ such that for all $n \in N$, $(R(\Psi_1, \Phi))_n = \theta$ with $\mathcal{A} \models \theta$. From the definition of Σ it follows that $\mathcal{A} \in \mathbf{A}$. So, it is easy to verify that $\langle \Psi , \mathbf{A} \rangle$ π^g-coordinates with $\langle \Phi , \mathbf{B} \rangle$. ■

We now wish to compare π-coordination and π^g-coordination. The following definition is useful.

(32) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$ and nonempty classes $\mathbf{A}$ and $\mathbf{B}$ of structures be given. We say that $\mathbf{A}$ π^g**-matches** to $\mathbf{B}$ just in case there are learning agents with explicit guess $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ such that $\langle \Psi , \mathbf{A} \rangle$ π^g-coordinates with $\langle \Phi , \mathbf{B} \rangle$. Similarly, we say that $\mathbf{A}$ π**-matches** to $\mathbf{B}$ just in case there are learning agents $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ such that $\langle \Psi , \mathbf{A} \rangle$ π-coordinates with $\langle \Phi , \mathbf{B} \rangle$.

We write "$\pi^g \approx \pi$" just in case for whatever vocabulary $\mathcal{L}$ and for all nonempty classes $\mathbf{A}$, $\mathbf{B}$ of structures for $\mathcal{L}$, $\mathbf{A}$ π^g-matches $\mathbf{B}$ iff $\mathbf{A}$ π-matches $\mathbf{B}$. We have:

(33) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$. $\pi^g \approx \pi$.

Proof: For the left-to-right direction, suppose that learning agents with explicit guess $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ π-coordinate with guesses. We define mappings Ψ',Φ' from *SEQ* to $\mathcal{L}_{basic} \times \mathcal{L}_{sen}$ as follows. For all $\sigma \in SEQ$ and all $\theta \in \mathcal{L}_{sen}$, if $\Psi(\sigma, \theta) = \langle \alpha , \theta' \rangle$, then define $\Psi'(\sigma) = \langle v_n \doteq v_n , \theta' \rangle$ with $n = \ulcorner \theta' \urcorner$ (the Gödel number of θ'), if $\alpha = (v_i \doteq v_i)$ for some $i \in N$; define $\Psi'(\sigma) = \Psi(\sigma, \theta)$ otherwise. Similarly, for all $\sigma \in SEQ$ and all $\theta \in \mathcal{L}_{sen}$, if $\Phi(\sigma, \theta) = \langle \beta , \theta'' \rangle$, then define $\Phi'(\sigma) = \langle v_m \doteq v_m , \theta' \rangle$ with $m = \ulcorner \theta'' \urcorner$ if $\beta = (v_j \doteq v_j)$ for some $j \in N$; define $\Phi'(\sigma) = \Phi(\sigma, \theta)$ otherwise. Cleary, $\langle \Psi' , \mathbf{A} \rangle$ and $\langle \Phi' , \mathbf{B} \rangle$ are learning agents and π-coordinate.

For the right-to-left direction, suppose that learning agents guess $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ π-coordinate. We define $\Psi', \Phi' : SEQ \times \mathcal{L}_{sen} \longrightarrow \mathcal{L}_{basic} \times \mathcal{L}_{sen}$ as follows. For all $\sigma \in SEQ$ and all $\theta \in \mathcal{L}_{sen}$, $\Psi'(\sigma, \theta) = \Psi(\sigma)$ and $\Phi'(\sigma, \theta) = \Phi(\sigma)$. Cleary, $\langle \Psi' , \mathbf{A} \rangle$ and $\langle \Phi' , \mathbf{B} \rangle$ are learning agents with guess and π-coordinate with guesses. ■

A stronger version of Proposition (33) is the next proposition. The proposition is stronger as it gives the exact sense by which a learning agent with explicit guess can be simulated by a learning agent (with no explicit guess), and vice versa. Recall that Λ^l and Λ^{lg} denote, respectively, the class of learning agents and the class of learning agents with explicit guess.

(34) PROPOSITION: Learning agents with explicit guess can be simulated by learning agents (without explicit guess) and vice versa in the following sense.

(a) There is an uniform map Γ from Λ^{lg} to Λ^l such that for all learning agents with explicit guess $\langle \Psi , \mathbf{A} \rangle$, $\langle \Phi , \mathbf{B} \rangle$ and for every $\pi \subseteq \mathcal{L}_{sen}$, $\langle \Psi , \mathbf{A} \rangle$ π^g-coordinates with $\langle \Phi , \mathbf{B} \rangle$ iff $\Gamma(\langle \Psi , \mathbf{A} \rangle)$ π-coordinates with $\Gamma(\langle \Phi , \mathbf{B} \rangle)$.

(b) There is an uniform map Γ' from Λ^l to Λ^{lg} such that for all learning agents $\langle \Psi , \mathbf{A} \rangle$, $\langle \Phi , \mathbf{B} \rangle$ and for every $\pi \subseteq \mathcal{L}_{sen}$, $\langle \Psi , \mathbf{A} \rangle$ π-coordinates with $\langle \Phi , \mathbf{B} \rangle$ iff $\Gamma'(\langle \Psi , \mathbf{A} \rangle)$ π^g-coordinates with $\Gamma'(\langle \Phi , \mathbf{B} \rangle)$.

Proof: We prove (a). Let $\sigma \in SEQ$, $i \in N^+$, and learning agents with explicit guess $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ be given. We define $\Psi'(\sigma) \in \mathcal{L}_{basic} \times \mathcal{L}_{sen}$ from Ψ and $\Phi'(\sigma) \in \mathcal{L}_{basic} \times \mathcal{L}_{sen}$ from Φ as follows.

(a) $\Psi'(\emptyset) = \Psi(\emptyset, \emptyset)$.

(b) If $|\sigma| = 2i$, let for $j = 0, 1, \ldots, i-1$ $\theta_j \in \mathcal{L}_{sen}$ be such that $\ulcorner \theta_j \urcorner = n$ [*i.e.* the Gödel number of θ_j] if $\sigma_{2j} = (v_n \doteq v_n)$, and let θ_j be any $\mathcal{L}$-sentence otherwise. Then $(\Psi'(\sigma))_0 = (\Psi(\langle \sigma_1 \sigma_3 \cdots \sigma_{2i-3} \rangle, \langle \theta_0 \cdots \theta_{i-1} \rangle))_0$.

(c) If $|\sigma| = 2i + 1$ then $(\Psi'(\sigma))_1 = (\Psi(\langle \sigma_1 \sigma_3 \cdots \sigma_{2i-1} \rangle, \langle \theta_0 \cdots \theta_i \rangle))_1$.

(d) If $\sigma_{2j} \neq (v_n \doteq v_n)$ for all $n \in N$, $\Psi'(\sigma)$ is arbitrary.

In similar way we define Φ'. It follows from the definitions of Ψ' and Φ' that $(R(\Psi_0, \Phi))_i = (R(\Psi'_0, \Phi'))_{2i}$ and $(R(\Psi_1, \Phi))_i = (R(\Psi'_1, \Phi'))_{2i+1}$. Define $\Gamma(\langle \Psi , \mathbf{A} \rangle) = \langle \Psi' , \mathbf{A} \rangle$ and $\Gamma(\langle \Phi , \mathbf{B} \rangle) = \langle \Phi' , \mathbf{B} \rangle$. Hence $\langle \Psi , \mathbf{A} \rangle$ π^g-coordinates with $\langle \Phi , \mathbf{B} \rangle$ iff $\Gamma(\langle \Psi , \mathbf{A} \rangle)$ π-coordinates with $\Gamma(\langle \Phi , \mathbf{B} \rangle)$.

We prove (b). Let learning agents (without explicit guess) $\langle \Psi , \mathbf{A} \rangle$ and $\langle \Phi , \mathbf{B} \rangle$ be given. For all $\sigma \in \mathit{SEQ}$ and all $\sigma' \in \mathit{SEN}$, we define $\Psi'(\sigma, \sigma') = \Psi(\sigma)$ and $\Phi'(\sigma, \sigma') = \Phi(\sigma)$. [Thus Ψ' and Φ' ignore the explicit guess and act as Ψ and Φ, respectively.] Let $\Gamma'(\langle \Psi , \mathbf{A} \rangle) = \langle \Psi' , \mathbf{A} \rangle$ and $\Gamma'(\langle \Phi , \mathbf{B} \rangle) = \langle \Phi' , \mathbf{B} \rangle$. Clearly $\langle \Psi , \mathbf{A} \rangle$ π-coordinates with $\langle \Phi , \mathbf{B} \rangle$ iff $\Gamma'(\langle \Psi , \mathbf{A} \rangle)$ π^g-coordinates with $\Gamma'(\langle \Phi , \mathbf{B} \rangle)$. ■

4 Transition to Solvability

How coordination relates to learning of structures? In this section we present our first formal answer to the question. We show how problem solving emerges from cooperation by studying a new paradigm of learning based on the contextual coordination. We will call such paradigm $\mathbf{Co}^{\pi}$-solvability. An interesting feature of the paradigm is that in the most relevant case of classification problems, the learning of a class $\mathbf{K}$ of structures in a given context reduces to a problem of coordination in such a context with an appropriate collection of learning agents based on $\mathbf{K}$. The $\mathbf{Co}^{\pi}$-solvability paradigm thus provides a link between $\mathbf{Ex}^{\pi}$-solvability and π-coordination.

4.1 $\mathbf{Co}^{\pi}$-solvability

Consider the case of the isolated agent facing Nature: Nature provides full information to the agent over the reality he is investigating, and the agent has a non null payoff function. The agent wins his game if he makes in the limit the right conjecture on the reality Nature made actual. A $\mathbf{Ex}^{\pi}$-solvability paradigm as sketched in Section 2 aims at studying how such a "learner" will decide when facing different circumstances in the natural context π, and what realities he learns under which conditions. He could have preferences and beliefs and is eventually rational according to some principle of rationality.

Suppose now we introduce other agents into the learner's environment and make them interact. Is a paradigm of learning based on their interaction reducible to a paradigm of the isolated agent? One might wonder why there should be any difficulty here. After all, the only difference between a natural context and a social context is just the presence of other people; agents' choices look the same in both cases. In this section we prove the truth of the last assertion for $\mathbf{Ex}^{\pi}$-solvability as the paradigm of learning by an isolated agent. Of course, to do this we have first to say what we mean by "interaction" and "learning".

By interaction we mean any interaction sequence of two learning agents. An agent in the present paradigm is thus a learning agent in the sense of the paradigm of contextual coordination. We model learning by restricting the π-coordination paradigm to collections of learning agents based on a fixed background world; we called such learning agents: "world-centered" (Definition (24)). Three of the six components of cooperation are already defined: worlds, contexts and the success criterion that stipulates the conditions under which learning agents can be said to coordinate in a given context are still in use in the present paradigm. To complete the list we have to introduce formally the unique, remaining component, that is, a success criterion that stipulates the conditions under which learning agents can be said to learn a class of worlds in a given context. The next definition is such a remaining component. (Recall that $\Lambda^l(\mathbf{A})$ denotes the class of all the learning agents based on $\mathbf{A}$.)

(35) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$, learning agent $\langle \Phi , \mathbf{B} \rangle$ and class $\mathbf{A}$ of structures be given.

- (*a*) $\langle \Phi , \mathbf{B} \rangle$ **$\mathbf{Co}^{\pi}$-solves** $\mathbf{A}$ (written: $\mathbf{A} \in \mathbf{Co}^{\pi}(\Phi)$) just in case $\langle \Phi , \mathbf{B} \rangle$ π-coordinates with every weakly $\mathbf{A}^{\pi}$-centered learning agent. In this case $\mathbf{A}$ is said to be **$\mathbf{Co}^{\pi}(\mathbf{B})$-solvable**.
- (*b*) $\mathbf{Co}^{\pi}(\mathbf{B}) = \{\mathbf{A} \,|\, \mathbf{A} \text{ is } \mathbf{Co}^{\pi}(\mathbf{B})\text{-solvable}\}$.
- (*c*) $\mathbf{Co}^{\pi} = \{\mathbf{A} \,|\, \mathbf{A} \text{ is } \mathbf{Co}^{\pi}(\mathbf{B})\text{-solvable for some } \mathbf{B}\}$.

We say that $\mathbf{A}$ is **$\mathbf{Co}^{\pi}$-solvable** just in case $\mathbf{A}$ is $\mathbf{Co}^{\pi}(\mathbf{B})$-solvable for some $\mathbf{B}$. "**Co**" stands for "**Co**-ordination." The prefix "$\mathbf{Co}^{\pi}$" should help to remember that a fundamental use of π-coordination is made in the criterion of success above.

Here are two examples, one of $\mathbf{Co}^{\pi}$-solvability and one of $\mathbf{Co}^{\pi}$-unsolvability. Lemmas (36) and (37) below are adapted from [OW86] and [Gol67, OW86] respectively.

(36) LEMMA: Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let $\pi = \{\theta_0, \theta_1\}$, where $\theta_0 \in \mathcal{L}_{sen}$ denotes $\exists x \forall y \; xRy$ ("there is a least point w.r.t. R) and $\theta_1 \in \mathcal{L}_{sen}$ denotes $\exists x \forall y \; yRx$ ("there is a greatest point w.r.t. R). Let $\mathcal{K}_0 = \langle \{n \in Q \,|\, n \geq 0\}, < \rangle$ and $\mathcal{K}_1 = \langle \{n \in Q \,|\, n \leq 0\}, < \rangle$, where Q is the set of rational numbers and $<$ is the standard strict order on Q which interprets R. Then $\mathbf{W} = \{\mathcal{K}_0, \mathcal{K}_1\}$ is $\mathbf{Co}^{\pi}$-solvable.

We postpone the proof of the lemma to the end of the next section.

(37) LEMMA: Suppose that $\mathcal{L}$ consist of a single unary predicate symbol P. Let $\pi = \{\theta_0\} \cup \{\theta_i \,|\, i \in N\}$, where $\theta_0 \in \mathcal{L}_{sen}$ denotes $\forall x \, Px$ ("x is even") and $\theta_i \in \mathcal{L}_{sen}$ for $i \in N$ denotes $\forall x < 2i \, Px$ ("x is even and less than $2i$"). Let $\mathcal{K} = \langle N , \{2n \,|\, n \in N\} \rangle$ and $\mathcal{K}_i = \langle N , \{2n \,|\, n < i\} \rangle$ for $i \in N$. Then $\mathbf{W} = \{\mathcal{K}\} \cup \{\mathcal{K}_i \,|\, i \in N\}$ is not $\mathbf{Co}^{\pi}$-solvable.

Again, the proof of the lemma will be presented at the end of the next section as a corollary of Theorem (40) below.

The next proposition is helpful to simplify the use of background worlds. Informally, the proposition says that a learner's background world does not reduce the power of learning in the $\mathbf{Co}^{\pi}$ paradigm.

(38) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$. For all classes $\mathbf{K}$ of structures, if $\mathbf{K} \in \mathbf{Co}^{\pi}$ then $\mathbf{K} \in \mathbf{Co}^{\pi}(\mathbf{K})$.

Proof: Suppose that learning agent $\langle \Psi , \mathbf{A} \rangle$ $\mathbf{Co}^{\pi}$-solves $\mathbf{K}$. We define a learning agent $\langle \Phi , \mathbf{K} \rangle$ that $\mathbf{Co}^{\pi}$-solves $\mathbf{K}$ as follows.

(a) For all $\sigma \in SEQ$, $(\Phi(\sigma))_1 = (\Psi(\sigma))_1$.
(b) $(\Phi(\emptyset))_0 = (v_0 \doteq v_0)$.
(c) For all $\sigma \in SEQ$ with $|\sigma| = n > 0$, $(\Phi(\sigma))_0 = \mathit{last}(\sigma)$.

So, in its second component, Φ copies all Ψ's guesses. In the first component, after a "safe" first move Φ copies the last element of the input sequence. For $\sigma \in SEQ$, suppose that $\mathbf{K}^{\pi}$-centered agent with ability $\mathcal{X}$ is such that $\sigma = R(\mathcal{X}_0, \Psi)|_n$ and suppose that Ψ has copied Ψ's moves up to $n-1$ (*). Then,

$$\begin{aligned}
&(\Phi(\sigma))_0 = (\Phi(R(\mathcal{X}_0, \Psi)|_n))_0 && \text{by assumption on } \mathcal{X};\\
&= (\Phi(R(\Psi_0, \mathcal{X})|_{n-1}(\mathcal{X}(R(\Psi_0, \mathcal{X})|_{n-1}))_0))_0 && \text{by definition of } R(.,.);\\
&= (\Phi(R(\Phi_0, \mathcal{X})|_{n-1}(\mathcal{X}(R(\Phi_0, \mathcal{X})|_{n-1}))_0))_0 && \text{by (*)};\\
&= (\mathcal{X}(R(\Phi_0, \mathcal{X})|_{n-1}))_0 && \text{by definition of } \Phi \text{ (clause } (c));\\
&= (\mathcal{X}(R(\Psi_0, \mathcal{X})|_{n-1}))_0 && \text{by (*)};\\
&= (\mathcal{X}(\sigma^-))_0 && \text{by definition of } R(.,.).
\end{aligned}$$

It is then easy to verify that $\langle \Phi\,, \mathbf{K}\rangle$ $\mathbf{Co}^{\pi}$-solves $\mathbf{K}$, hence $\mathbf{K} \in \mathbf{Co}^{\pi}(\mathbf{K})$. ■

Note that the converse of the proposition holds by definition of $\mathbf{Co}^{\pi}$.

(39) COROLLARY: Let $\pi \subseteq \mathcal{L}_{sen}$. For all classes $\mathbf{K}$ of structures, $\mathbf{K} \in \mathbf{Co}^{\pi}$ iff $\mathbf{K} \in \mathbf{Co}^{\pi}(\mathbf{K})$.

The corollary simplifies our presentation of the paradigm, since it allows us to forget a learning agents' background world. Moreover, it shows a first analogy of $\mathbf{Co}^{\pi}$ with $\mathbf{Ex}^{\pi}$-solvability, since learners (cf. Section 2) do not have a background world. A stronger analogy is given in the next subsection.

4.2 Equivalence with $\mathbf{Ex}^{\pi}$

We show the equivalence of $\mathbf{Ex}^{\pi}$ and $\mathbf{Co}^{\pi}$ for all classification problems with learning context π. To illustrate, suppose that for some $\mathbf{K}$ learning agent $\langle \Phi\,, \mathbf{K}\rangle$ π-coordinates with any self^{π}-centered learning agent. Informally, this means that we require that $\langle \Phi\,, \mathbf{K}\rangle$ π-coordinates with any "truth-teller" learning agent $\langle \Psi\,, \mathbf{K}\rangle$, even if such agent does not really wants to π-coordinate with $\langle \Phi\,, \mathbf{K}\rangle$. In this sense, $\langle \Psi\,, \mathbf{K}\rangle$ is said to be "passive". Formally, we have:

(40) THEOREM: Let $\pi \subseteq \mathcal{L}_{sen}$ and let class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is exactly one $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then, $\mathbf{A} \in \mathbf{Ex}^{\pi}$ if and only if $\mathbf{A} \in \mathbf{Co}^{\pi}(\mathbf{A})$.

Proof: For the left-to-right direction, suppose that learner Ψ $\mathbf{Ex}^{\pi}$-solves $\mathbf{A}$. Then,

(41) for every $\mathcal{A} \in \mathbf{A}$ and every environment e for $\mathcal{A}$, there is exactly one $\theta \in \pi$ with $\mathcal{A} \models \theta$ such that for all but finitely many $n \in N$, $\Psi(e|_n) = \theta$.

We define a learning agent $\langle \Phi\,, \mathbf{A}\rangle$ that $\mathbf{Co}^{\pi}$-solves $\mathbf{A}$ as follows. For all $\sigma \in SEQ$, $\Phi(\sigma)$ is defined iff $\Psi(\sigma)$ is defined. In this case, let $k < |\sigma|$ be minimal such that

$\tau(\sigma) = \langle \sigma_k \cdots \mathit{last}(\sigma)\rangle$ is consistent. Define $\Phi(\sigma) = \langle\, \mathit{last}(\tau(\sigma))\, , \Psi(\tau(\sigma))\,\rangle$, where $\tau(\sigma) = \langle \sigma_k \cdots \sigma_{|\sigma|}\rangle$. Let $\mathcal{X}$ be any learning agent weakly $\mathcal{A}$-centered in π. Then,

(42) there is $t \in N$ such that:

- (a) ${}_{t}|R(\mathcal{X}_0, \Phi)$ is an environment e for $\mathcal{A}$;
- (b) for all but finitely many $n \in N$, $({}_{t}|R(\mathcal{X}_1, \Phi))_n = \theta$.

We can assume without loss of generality that t is minimal. For sufficiently large $n \in N$, it follows that

(43) $\tau(e|_n) = e|_n$.

Then, we have the following chain of equalities:

$$\begin{array}{lr}
({}_{t}|R(\Phi_1, \mathcal{X}))_n = & \\
= (\Phi({}_{t}|R(\mathcal{X}_0, \Phi)|_n))_1 & \text{by definition of } {}_{t}|R(.,.); \\
= (\Phi(e|_n)_1 & \text{by } (42)(a); \\
= \Psi(\tau(e|_n)) & \text{by definition of } \Phi; \\
= \Psi(e|_n) & \text{by } (43); \\
= \theta & \text{by } (41).
\end{array}$$

It follows that for every weakly $\mathbf{A}^{\pi}$-centered learning agent $\langle \mathcal{X}\, , \mathbf{A}\rangle$, there is $k \in N$ such that ${}_{k}|R(\mathcal{X}_0, \Phi)$ is an environment for some $\mathcal{S} \in \mathbf{A}$, ${}_{k}|R(\Phi_0, \mathcal{X})$ is an environment for $\mathcal{S}$, and for all but finitely many $n \in N$, there is $\theta' \in \pi$ with $\mathcal{S} \models \theta'$ such that $({}_{k}|R(\Phi_1, \mathcal{X}))_n = \theta' = ({}_{k}|R(\mathcal{X}_1, \Phi))_n$. Hence, Φ π-coordinates with every weakly $\mathbf{A}^{\pi}$-centered learning agent, that is, $\mathbf{A} \in \mathbf{Co}^{\pi}(\mathbf{A})$.

For the right-to-left direction, suppose that learning agent $\langle \Phi\, , \mathbf{A}\rangle$ $\mathbf{Co}^{\pi}$-solves $\mathbf{A}$. Then,

(44) for all $\mathcal{A} \in \mathbf{A}$ and for every environment e for $\mathcal{A}$, there are weakly $\mathbf{A}^{\pi}$-centered learning agent $\mathcal{X}^{\mathcal{A},e}$ based on $\mathbf{A}$ and $t \in N$ with the following two properties.

- (a) ${}_{t}|R(\mathcal{X}_0^{\mathcal{A},e}, \Phi) = e$;
- (b) there is $\theta \in \pi$ such that for cofinitely many $n \in N$, $({}_{t}|R(\mathcal{X}_1^{\mathcal{A},e}, \Phi))_n = ({}_{t}|R(\Phi_1, \mathcal{X}^{\mathcal{A},e}))_n = \theta$ and $\mathcal{A} \models \theta$.

We define a learner Ψ that $\mathbf{Ex}^{\pi}$-solves $\mathbf{A}$ as follows. For all $\sigma \in \mathit{SEQ}$, $\Psi(\sigma)$ is defined iff $\Phi(\sigma)$ is defined. In this case, we define $\Psi(\sigma) = (\Phi(\sigma))_1$. Then for all $n \in N$, for all $\mathcal{A} \in \mathbf{A}$ and for every environment e for $\mathcal{A}$,

$$\begin{array}{lr}
\Psi(e|_n) = & \\
= (\Phi(e|_n))_1 & \text{by definition of } \Psi; \\
= (\Phi({}_{t}|R(\mathcal{X}_0^{\mathcal{A},e}, \Phi)|_n))_1 & \text{by } (44)(a); \\
= ({}_{t}|R(\Phi_1, \mathcal{X}^{\mathcal{A},e}))_n & \text{by definition of } R(.,.); \\
= \theta & \text{by } (44)(b),
\end{array}$$

Hence Ψ $\mathbf{Ex}^{\pi}$-solves every environment for each $\mathcal{A} \in \mathbf{A}$, that is, $\mathbf{A} \in \mathbf{Ex}^{\pi}$. ■

As a consequence of the theorem we have Lemmas (36) and (37). Indeed:

Proof of Lemma (36): Let vocabulary $\mathcal{L}$, $\pi \subseteq \mathcal{L}_{sen}$, and class $\mathbf{W}$ be as in Lemma (36). Observe that for all $\mathcal{S} \in \mathbf{W}$ there is exactly one $\theta \in \pi$ such that $\mathcal{S} \models \theta$. By Theorem (40) it follows that $\mathbf{W} \in \mathbf{Ex}^{\pi}$ iff $\mathbf{W} \in \mathbf{Co}^{\pi}(\mathbf{W})$. On the other hand, an easy adaptation of [OW86, Example 3.4A] shows that $\mathbf{W}$ is $\mathbf{Ex}^{\pi}$-solvable. Then $\mathbf{W}$ is $\mathbf{Co}^{\pi}(\mathbf{W})$-solvable, hence $\mathbf{Co}^{\pi}$-solvable. ■

To prove Lemma (37) we rely on the following corollary of Theorem (40).

(45) COROLLARY: Let $\pi \subseteq \mathcal{L}_{sen}$ and let class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. If $\mathbf{A} \in \mathbf{Co}^{\pi}(\mathbf{A})$ then $\mathbf{A} \in \mathbf{Ex}^{\pi}$.

Thus, the requirement on class $\mathbf{A}$ with regard to context π is less restrictive. With the corollary in hand, we may now prove Lemma (37) as follows.

Proof of Lemma (37): Let vocabulary $\mathcal{L}$, $\pi \subseteq \mathcal{L}_{sen}$, and class $\mathbf{W}$ be as in Lemma (37). An easy adaptation of [OW86, Example 3.5B] shows that $\mathbf{W}$ is not $\mathbf{Ex}^{\pi}$-solvable. On the other hand, observe that for every $\mathcal{S} \in \mathbf{W}$ there is $\theta \in \pi$ such that $\mathcal{S} \models \theta$. It follows from Corollary (45) that $\mathbf{W}$ is not $\mathbf{Co}^{\pi}(\mathbf{W})$-solvable. Hence, by Proposition (38), $\mathbf{W}$ is not $\mathbf{Co}^{\pi}$-solvable. ■

Finally, we characterize $\mathbf{Co}^{\pi}$-solvability of classification problems.

(46) PROPOSITION: Let $\mathbf{W} = \{MOD(\Gamma \cup \{\theta_i\}) \,|\, 0 \leq i \leq n\}$ for some consistent, possibly empty set $\Gamma \subseteq \mathcal{L}_{sen}$, and let $\pi = \{\theta_0, \theta_1, \ldots, \theta_n\}$ be such that for every $\mathcal{S} \in \mathbf{W}$, there is exactly one $i \in \{0 \cdots n\}$ with $\mathcal{S} \models \theta_i$. Then $\mathbf{W}$ is $\mathbf{Co}^{\pi}$-solvable if and only if for every $i \in \{0 \cdots n\}$, θ_i is equivalent in Γ to an $\exists\forall$-sentence.

Proof: It follows directly from Theorem (17) and Theorem (40). ■

5 Solvability by Restricted Coordination

Success in the $\mathbf{Co}^{\pi}$ paradigm is a strong requirement. Suppose that a learning agent $\mathbf{Co}^{\pi}$-solves class $\mathbf{K}$. Then the learning agent must π-coordinate with *all* the learning agents $\mathbf{A}$-centered in π. If we restrict the quantification on learning agents in certain way and define "success" on such restricted quantification we introduce a variation of the paradigm, where all the other components are unchanged. We do it in this section, and show that the resulting paradigm is indeed more liberal, in the sense that the class of (nonempty) collections of structures it contains is a strict superset of $\mathbf{Co}^{\pi}$ for all π.

(47) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$, learning agent $\langle \Phi, \mathbf{B} \rangle$ and nonempty class $\mathbf{A}$ of structures be given.

(*a*) $\langle \Phi, \mathbf{B} \rangle$ **$\mathbf{CoSoc}^{\pi}$-solves A** (written: $\mathbf{A} \in \mathbf{CoSoc}^{\pi}(\Phi)$) just in case there is set $\Sigma \subseteq \Lambda^{l}(\mathbf{A})$ such that:

i. $\langle \Phi, \mathbf{B} \rangle$ π-coordinates with every member of Σ;

   ii. for all $\mathcal{A} \in \mathbf{A}$, there are $\langle \Psi , \mathbf{A} \rangle \in \Sigma$ and $k \in N$ such that ${}_k|R(\Psi_0, \Phi)$ is an environment for $\mathcal{A}$.

In this case, $\mathbf{A}$ is said to be **$\mathbf{CoSoc}^{\pi}(\mathbf{B})$-solvable** (referring to Σ).

(b) $\mathbf{CoSoc}^{\pi}(\mathbf{B}) = \{\mathbf{A} \mid \mathbf{A} \text{ is } \mathbf{CoSoc}^{\pi}(\mathbf{B})\text{-solvable}\}$.

We say that $\mathbf{A}$ is **$\mathbf{CoSoc}^{\pi}$-solvable** just in case $\mathbf{A}$ is $\mathbf{CoSoc}^{\pi}(\mathbf{B})$-solvable for some $\mathbf{B}$. Then we define $\mathbf{CoSoc}^{\pi} = \{\mathbf{A} \mid \mathbf{A} \text{ is } \mathbf{CoSoc}^{\pi}\text{-solvable}\}$.

"**Soc**" stands for "**Soc**-iety." The prefix "$\mathbf{CoSoc}^{\pi}$" should help to remember that the paradigm is based on π-coordination, but it is restricted to subsets of learning agents, or "societies". Observe:

(48) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$. $\mathbf{Co}^{\pi} \subseteq \mathbf{CoSoc}^{\pi}$.

Proof: Suppose that learning agent $\langle \Phi , \mathbf{B} \rangle$ $\mathbf{Co}^{\pi}$-solves $\mathbf{A}$. Set $\Sigma = \{\langle \Psi , \mathbf{C} \rangle \in \Lambda^{l} \mid \langle \Psi , \mathbf{C} \rangle \text{ is weakly } \mathbf{A}\text{-centered in } \pi\}$ in Definition (47). Then it is easy to verify that $\langle \Phi , \mathbf{B} \rangle$ $\mathbf{CoSoc}^{\pi}$-solves $\mathbf{A}$ referring to Σ. ■

Proposition (48) raises the question if there is some context π such that $\mathbf{CoSoc}^{\pi} - \mathbf{Co}^{\pi} \neq \emptyset$. Before answering the question, we highlight the role of the learning agents' background world in the criterion of success of $\mathbf{CoSoc}^{\pi}$-solvability. The following proposition shows that limiting a learning agent to be based on the background world he wants to solve does not limit the learning ability of the agent.

(49) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$. For all classes $\mathbf{K}$ of structures, if $\mathbf{K} \in \mathbf{CoSoc}^{\pi}$ then $\mathbf{K} \in \mathbf{CoSoc}^{\pi}(\mathbf{K})$.

Proof: Suppose that learning agent $\langle \Psi , \mathbf{A} \rangle$ $\mathbf{CoSoc}^{\pi}$-solves $\mathbf{K}$. We define a learning agent $\langle \Phi , \mathbf{K} \rangle$ that $\mathbf{CoSoc}^{\pi}$-solves $\mathbf{K}$ as follows.

(a) For all $\sigma \in SEQ$, $(\Phi(\sigma))_1 = (\Psi(\sigma))_1$.
(b) $(\Phi(\emptyset))_0 = (v_0 \doteq v_0)$.
(c) For all $\sigma \in SEQ$ with $|\sigma| = n > 0$, $(\Phi(\sigma))_0 = \sigma_n$.

[Thus, Φ copies all the guesses made by $\langle \Psi , \mathbf{A} \rangle$. For the first component, after a "safe" first move Φ copies the last action (move) of each of the relevant opponent represented by the last element of σ.] It follows that $\langle \Phi , \mathbf{K} \rangle$ $\mathbf{Co}^{\pi}$-solves $\mathbf{K}$, so $\mathbf{K} \in \mathbf{Co}^{\pi}(\mathbf{K})$. ■

(50) COROLLARY: Let $\pi \subseteq \mathcal{L}_{sen}$. For all classes $\mathbf{K}$ of structures, $\mathbf{K} \in \mathbf{CoSoc}^{\pi}$ iff $\mathbf{K} \in \mathbf{CoSoc}^{\pi}(\mathbf{K})$.

By the corollary, our future discussion results can be simplified, since we consider the "class-independent" $\mathbf{CoSoc}^{\pi}$ version of the paradigm only.

Now we return to the question: Is $\mathbf{CoSoc}^{\pi} - \mathbf{Co}^{\pi} \neq \emptyset$ for some π? We give a positive answer as a corollary of the following proposition. Recall that by Convention (3) all the structures considered in this paper are assumed to be countable.

(51) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$ and let countable class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then $\mathbf{A}$ is $\mathbf{CoSoc}^{\pi}$-solvable.

Proof: Let $\pi = \{\theta_i \in \mathcal{L}_{sen} \mid i \in N\}$ and let $\{\mathcal{A}_j \mid j < \omega\}$ be an infinite-repetition enumeration of the countably many members of $\mathbf{A}$. Suppose that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Given $n, m \in N$, let learning agent $\langle \Psi_{<n,m>}, \mathbf{A} \rangle$ be $\mathcal{A}_n$-centered in $\{\theta_m\}$ and such that $\Psi_{<n,m>}(\emptyset) = \langle v_n \doteq v_n, \theta_m \rangle$. Let $\Sigma = \{\langle \Psi_{<n,m>}, \mathbf{A} \rangle \mid \mathcal{A}_n \models \theta_m\}$. Observe that Σ is nonempty. We define learning agent $\langle \Phi, \mathbf{A} \rangle$ such that:

(a) $\Phi(\emptyset) = \langle v_0 \doteq v_0, \theta_0 \rangle$.

(b) For all $\sigma \in SEQ$ with $|\sigma| > 0$, if for some $n \in N$ σ_0 is of the form $v_n \doteq v_n$, then $(\Phi(\sigma))_0 = \sigma_{|\sigma^-|}$ and $(\Phi(\sigma))_1 = \theta_n$.

(c) For all $\sigma \in SEQ$ with $|\sigma| > 0$, if for all $n \in N$ σ_0 is not of the form $v_n \doteq v_n$, then $\Phi(\sigma) = \langle v_0 \doteq v_0, \theta_0 \rangle$.

It is easy to verify that $\mathbf{A}$, Φ and Σ satisfy the conditions of Definition (47) (for $k = 0$), hence $\mathbf{A}$ is $\mathbf{CoSoc}^{\pi}$-solvable. ■

(52) COROLLARY: There is $\pi \subseteq \mathcal{L}_{sen}$ such that $\mathbf{CoSoc}^{\pi} - \mathbf{Co}^{\pi} \neq \emptyset$.

Proof: Take $\mathbf{A}$ be any countable class of structures such that (a) $\mathbf{A} \notin \mathbf{Ex}^{\pi}$ for some $\pi \subseteq \mathcal{L}_{sen}$ and (b) the pair $\mathbf{A}$, π verifies the hypothesis of Proposition (51). Such class $\mathbf{A}$ and π exist (*e.g.*, the class in Proposition (16) is countable). From (a) it follows by Theorem (40) that $\mathbf{A} \notin \mathbf{Co}^{\pi}(\mathbf{A})$ and then, by Proposition (38), $\mathbf{A} \notin \mathbf{Co}^{\pi}$. On the other hand, from (b) and Proposition (51) it follows that $\mathbf{A} \in \mathbf{CoSoc}^{\pi}$. ■

On the intuitive meaning of Corollary (52) and of the paradigm $\mathbf{CoSoc}^{\pi}$, a quite philosophical remark is on the existence of relatively small groups of learning agents that are able to coordinate pairwise and that, under some additional condition as imposed by the definition of the class $\mathbf{CoSoc}^{\pi}$, sometimes solve problems that larger groups of learning agents fail to solve.

6 $\mathbf{Gr}^{\pi}$-solvability

We now return to the problem of combining coordination and solvability (*i.e.*, cooperation; think to our equational "slogan" (1)). We introduced the matter in Section 4 by advancing the paradigm of $\mathbf{Co}^{\pi}$-solvability. There the paradigm has been proved to be equivalent to pure solvability in the most relevant cases (cf. Theorem (40)). This means that in such cases we may always define a learner who simulates (in a given learning context π) the joint actvity of a group of learning agents who coordinate in π. A consequence is that the collection of solvable problems by means of contextual coordination is not larger than the class of problems which are purely solvable, that is, the problems that can be already solved within the $\mathbf{Ex}^{\pi}$-solvability paradigm.

In the present section, we investigate the learning power of certain sets of learning agents. We call these sets: **groups**. We start in the next section by considering group solvability of countable classes of structures in a possibly uncomputable setting. In particular, we prove that every countable class $\mathbf{K}$ of structures is solvable by a suitable group of (uncomputable) learning agents in a suitable learning context. We will see that "solvable" here is taken in a "weak" sense, and that the resulting paradigm of solvability is more liberal than the paradigm of $\mathbf{Co}^{\pi}$-solvability. More precisely, our first result says that for every countable class $\mathbf{K}$ of structures and for every context π of sentences such that in every structure in $\mathbf{K}$ exactly one sentence of π is true, there is a group of learning agents such that (a) all the learning agents of the group are based on $\mathbf{K}$, (b) any pair of learning agents of the group coordinate in π, and (c) for every learning agent $\langle \Phi \, , \mathbf{K} \rangle$ of the group and for all structures $\mathcal{A}$ in $\mathbf{K}$, there is a learning agent $\langle \Psi \, , \mathbf{K} \rangle$ of the group, where $\Psi(\sigma) \neq \Phi(\sigma)$ for some $\sigma \in \mathit{SEQ}$, such that Ψ enumerates with Φ an environment for $\mathcal{A}$, eventually after a finite number of failures, and guesses with Φ a sentence in π true in $\mathcal{A}$. In other words, clause (c) means that each learning agent in the group is potentially able to solve every structure in the class provided that the agent is "helped" by a suitable other member of the group; clause (b) means that all the agents in the group are "cooperative", that is, coordination between any pair of them never fails. This result is successively generalized to uncountable classes of structures where at least one sentence in the learning context is true. A first consequence of our generalization is the separation of group solvability from $\mathbf{Co}^{\pi}$-solvability. From the side of computable solvability, we compare some of our major paradigms of solvability and exhibit a collection of problems that is solvable by a single computable learner but that a group of computable learning agents cannot solve.

To state the new paradigm, we modify the concept of learning success. Agents, worlds, dynamics, learning contexts and the criterion of successful coordination remain those of the paradigm of $\mathbf{Co}^{\pi}$-solvability.

(53) Definition: Let $\pi \subseteq \mathcal{L}_{sen}$, nonempty class $\mathbf{A}$ of structures, and set Σ of learning agents based on $\mathbf{A}$ be given.

(a) Σ $\mathbf{Gr}^{\pi}$**-solves** $\mathbf{A}$ (written: $\mathbf{A} \in \mathbf{Gr}^{\pi}(\Sigma)$) just in case:

  i. for all $\langle \Psi \, , \mathbf{A} \rangle, \langle \Phi \, , \mathbf{A} \rangle \in \Sigma$, $\langle \Psi \, , \mathbf{A} \rangle$ π-coordinates with $\langle \Phi \, , \mathbf{A} \rangle$;
  ii. for all $\langle \Phi \, , \mathbf{A} \rangle \in \Sigma$ and for all $\mathcal{A} \in \mathbf{A}$, there are $\langle \Psi \, , \mathbf{A} \rangle \in \Sigma$ and $k \in N$ such that ${}_{k}|R(\Psi_0, \Phi)$ is an environment for $\mathcal{A}$.

In this case, $\mathbf{A}$ is said to be $\mathbf{Gr}^{\pi}$**-solvable**.

(b) $\mathbf{Gr}^{\pi} = \{\mathbf{A} \mid \mathbf{A} \text{ is } \mathbf{Gr}^{\pi}\text{-solvable}\}$.

We call **group** a set Σ of learning agents that satisfies clause (a)ii of the definition. So, in particular, a group does not depend on any learning context. In the sequel we are interested in the learning power of certain infinite groups of learning agents. For any fixed context $\pi \subseteq \mathcal{L}_{sen}$, Definition (53) identifies what we refer to as (the criterion of success of a paradigm of) **group-learning** (or learning by groups). The notation "**Gr**" is and will be used in the sequel to mean that groups are taken to be the active subject of learning. This choice contrasts with our previous paradigms, where a learner is always thought to be an individual subject. Consequently, solvability as modeled by the present paradigm is a collective matter. The following definition is useful.

(54) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$ be given. We say that a group is **fully collaborative in** π just in case every pair of distinct learning agents of the group π-coordinate.

Recall that two learning agents with ability Ψ and Φ, respectively, are **distinct** just in case either $(\Psi(\sigma))_0 \neq (\Phi(\sigma))_0$ or $(\Psi(\sigma))_1 \neq (\Phi(\sigma))_1$ for some $\sigma \in SEQ$. Directly from the definition of a fully collaborative group it follows that for all $\pi \subseteq \mathcal{L}_{sen}$ and for all groups Σ of learning agents based on a nonempty class $\mathbf{A}$ of structures, Σ $\mathbf{Gr}^{\pi}$-solves $\mathbf{A}$ only if Σ is fully collaborative in π.

Here is a first example of $\mathbf{Gr}^{\pi}$-solvability.

(55) LEMMA: Let $\pi \subseteq \mathcal{L}_{sen}$ and let countable class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is exactly one $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then $\mathbf{A}$ is $\mathbf{Gr}^{\pi}$-solvable.

The proof of the lemma is based on the construction of a countable set of learning agents from a convention on the leadership requirements to be shared by all agents in the group. Informally, we may describe the proof as follows. Each agent's first move is as mainly aimed at establishing who is the leader between the agent himself and any other agent in the group. Soon after each member of the group has moved by interacting with some other agent in the group, all the agents know who is the leader in each pairwise interaction. Successively, the agent who is the leader completely and interactively describes with a partner a structure in the class $\mathbf{A}$ with index equal to the leader's own index. In other words, the leader enumerates with another agent in the group, whose role is substantially passive as he is not the leader (partner), an environment for a structure in $\mathbf{A}$. The structure is chosen by the leader looking at the leader's own index. At the same time, the leader guesses by interacting with a partner the only sentence in context π true in the structure. In fact the leader will guess such a sentence forever after. In contrast, each agent who is not the leader in the actual interaction simply copies the behavior of the leader by looking at the leader's last move. Moreover, the agent who is not the leader guesses what the leader does. Since learning agents' guesses are hidden by definition, in doing so the non-leader agent (let his index be n) looks at the first element σ_0 of the input sequence and guesses a sentence θ_m in π whenever σ_0 is of the form $v_m \doteq v_m$ for $m > n$, where we assume that m is the leader's index. Such assumption is indeed the convention the agents in the group must stipulate before any pairwaise interaction. On such a convention they will thus establish the leadership on. The formal content of this informal account is the following.

Proof of Lemma (55): Let $\{\mathcal{A}_j \mid j < \omega\}$ be an infinite-repetition enumeration of the countably many members of $\mathbf{A}$. Given $n \in N$, let $e^{\mathcal{A}_n}$ be an environment for

$\mathcal{A}_n$. We define learning agent $\langle \Phi^n , \mathbf{A} \rangle$ such that:

(a) $\Phi^n(\emptyset) = \langle\, v_n \doteq v_n \,, \theta_n \,\rangle$, where $\theta_n \in \pi$ is unique such that $\mathcal{A}_n \models \theta_n$.

(b) For all $\sigma \in SEQ$ with $|\sigma| > 0$, if $\sigma_0 = v_m \doteq v_m$ for some $m > n$, then $\Phi^n(\sigma) = \langle\, last(\sigma) \,, \theta_m \,\rangle$, where $\theta_m \in \pi$ is unique such that $\mathcal{A}_m \models \theta_m$. Otherwise, $\Phi^n(\sigma) = \langle\, (e^{\mathcal{A}_n})_{|\sigma|-1} \,, \theta_n \,\rangle$ [thus $\langle \Phi^n , \mathbf{A} \rangle$ is the **leader**].

Define $\Sigma = \{\langle \Phi^n , \mathbf{A} \rangle \,|\, n \in N\}$. For all $\mathcal{A} \in \mathbf{A}$ and for all $\langle \Phi^n , \mathbf{A} \rangle \in \Sigma$, let $m \geq n$ be such that $\mathcal{A}_m = \mathcal{A}$. Since $\{\mathcal{A}_j \,|\, j < \omega\}$ is an infinite-repetition of structures in $\mathbf{A}$ such m exists. Then $\langle \Phi^m , \mathbf{A} \rangle \in \Sigma$ [$\langle \Phi^m , \mathbf{A} \rangle$ is the leader] and $R(\Phi^m_0, \Phi^n)$ is an environment for $\mathcal{A}$, and for all but finitely many $k \in N$, $(R(\Phi^m_1, \Phi^n))_k = \theta_m$. On the other hand, it is easily verified that convention on leadership implies that each pair of learning agents in Σ π-coordinate. Hence, $\mathbf{A} \in \mathbf{Gr}^{\pi}(\Sigma)$. ■

6.1 Comparison with $\mathbf{Co}^{\pi}$

It is quite evident that $\mathbf{Gr}\pi$-solvability differs from $\mathbf{Co}^{\pi}$-solvability, in a sense that the former paradigm is more liberal than the latter. To see this informally, let us observe that a learning agent that $\mathbf{Co}^{\pi}$-solves a class $\mathbf{K}$ of structures takes as input *every* environment for all members of $\mathbf{K}$. The environments are shown to the learning agent as products of *any* learning agent who is consistent with the class as well as with the learning context π being given—we called those learning agents $\mathbf{K}^{\pi}$-centered. On the other hand, to solve class $\mathbf{K}$ in the present paradigm of $\mathbf{Gr}^{\pi}$-solvability, each learning agent in a group is concerned only with the coordination re-action of *some* other agent in the group. In other words, to be a fully collaborative group of learning agents based on the same background world in a context π is less harder than π-coordinating with every learning agent consistent in π with that background world.

In this subsection we are aiming at showing that success in $\mathbf{Gr}^{\pi}$ is in fact a more liberal affair than in the paradigm $\mathbf{Co}^{\pi}$. In a formal presentation, however, first a stronger result than Lemma (55) can be proved. The result is stronger in the sense that it refers to the weaker hypothesis that, given context $\pi \subseteq \mathcal{L}_{sen}$, each structure in the class of structures in the lemma is required to satisfy at least one sentence in π rather than exactly one sentence (in π). The next theorem thus says that there is a (countable) group of learning agents that solves any problem of a reasonable form. As in the case of Lemma (55), such a successful group is based on common knowledge of what convention regulates leadership in any pairwise interaction defined on the group members.

(56) THEOREM: Let $\pi \subseteq \mathcal{L}_{sen}$ and let countable class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then $\mathbf{A}$ is $\mathbf{Gr}^{\pi}$-solvable.

Proof: Let $\pi = \{\theta_i \in \mathcal{L}_{sen} \,|\, i \in N\}$ and let $\{\mathcal{A}_j \,|\, j < \omega\}$ be an infinite-repetition enumeration of the countably many members of $\mathbf{A}$. Let $e^{\mathcal{A}_n}$ be an environment for $\mathcal{A}_n \in \mathbf{A}$ and let $\theta_m \in \pi$ be such that $\mathcal{A}_n \models \theta_m$. We describe a learning agent $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ such that when interacting with a learning agent $\langle \Psi_{<k,j>} , \mathbf{A} \rangle$ behaves as follows.

Informally, the two agents are assumed to have some common knowledge on the convention regulating leadership. The convention we refer to is lexicographic order on the agents indexing. The leader is thus who has the lower index, so that $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ is the leader iff $<n,m>$ is lower than $<k,j>$ according to lexicographic order. The assessment of the leadership is done by $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ soon after the first two stages of the agents' interaction are played. Precisely, by the first component of communication function $\Psi_{<n,m>}$, $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ starts moving $v_n \doteq v_n$ to communicate to learning agent $\langle \Psi_{<k,j>} , \mathbf{A} \rangle$ that if he will be the leader then he will completely describe $\mathcal{A}_n$. This will be done indeed by enumerating with his partner environment $e^{\mathcal{A}_n}$. At the second interaction stage and, again, by the first component of communication function only (since guesses are "hidden" communicating them to the partner is useless), $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ moves $v_m \doteq v_m$ to communicate to $\langle \Psi_{<k,j>} , \mathbf{A} \rangle$ that if he will be the leader then he will guess (interacting with his partner) θ_m forever. From the third stage of the interaction onwards, both the agents know who is the leader. If $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ is not the leader, *i.e.* $\langle \Psi_{<k,j>} , \mathbf{A} \rangle$ results to be the leader, then $\Psi_{<n,m>}$ simply copies the moves of $\Psi_{<k,j>}$, precisely those moves $\Psi_{<k,j>}$ advances by the first component of his communication function. Also, $\Psi_{<n,m>}$ outputs the guess θ_j in π produced by $\Psi_{<k,j>}$. Because guesses are "hidden", in particular to $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$—in fact $\Psi_{<n,m>}$'s input domain is *SEQ*, θ_j is communicated by $\langle \Psi_{<k,j>} , \mathbf{A} \rangle$ to $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ via a basic formula of the form $v_j \doteq v_j$. On the other hand, if $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ is the leader then he acts consistently to what announced in advance by his first two moves, namely, he starts enumerating with $\Psi_{<k,j>}$ an environment for $\mathcal{A}_n$ (specifically environment $e^{\mathcal{A}_n}$), and starts guessing with $\Psi_{<k,j>}$ sentence θ_m forever after. Formally, we define a learning agent $\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ such that:

(*a*) $\Psi_{<n,m>}(\emptyset) = \langle\, v_n \doteq v_n \,, \theta_m \,\rangle$.

(*b*) For all $\sigma \in SEQ$ with $|\,\sigma\,| = 1$, $\Psi_{<n,m>}(\sigma) = \langle\, v_m \doteq v_m \,, \theta_m \,\rangle$.

(*c*) For all $\sigma \in SEQ$ with $|\,\sigma\,| > 1$, the following cases arise.

Case 1: $\sigma_0 = v_k \doteq v_k$ and $\sigma_1 = v_j \doteq v_j$, where either $k > n$ or $k = n$ and $j > m$. Then, $\Psi_{<n,m>}(\sigma) = \langle\, last(\sigma) \,, \theta_j \,\rangle$.
[$\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ is not the leader.]
Case 2: Not *Case 1*. Then, $\Psi_{<n,m>}(\sigma) = \langle\, (e^{\mathcal{A}_n})_{|\,\sigma\,|-2} \,, \theta_m \,\rangle$.
[$\langle \Psi_{<n,m>} , \mathbf{A} \rangle$ is the **leader**.]

Set $\Sigma = \{ \langle \Psi_{<n,m>} , \mathbf{A} \rangle \,|\, \mathcal{A}_n \models \theta_m \}$. Our claim that Σ witnesses $\mathbf{A} \in \mathbf{Gr}^{\pi}(\Sigma)$ is then easy to verify by following the argument as in the proof of Lemma (55). Hence, $\mathbf{A}$ is $\mathbf{Gr}^{\pi}$-solvable. ■

(57) *Remark:* The proof of the theorem can be made no more difficult than the proof of Lemma (55), and can indeed be simplified by fixing $\theta_n \in \pi$ such that $\mathcal{A}_n \models \theta_n$ in advance as common knowledge. However, the proof we have given is an introduction to the proof of Theorem (60) below, where the kind of simplification we mentioned does not work.

We return to the comparison of $\mathbf{Gr}^{\pi}$, $\mathbf{Ex}^{\pi}$ and $\mathbf{Co}^{\pi}$. To separate group-learning from $\mathbf{Ex}^{\pi}$-solvability we exhibit a collection of countable structures not in $\mathbf{Ex}^{\pi}$ and that satisfies the requirement of Theorem (56). Proposition (16) provides a collection of structures with these properties.

By combining Theorem (56) with Proposition (16) we obtain the following "separation" result.

(58) PROPOSITION: There is $\pi \subseteq \mathcal{L}_{sen}$ such that $\mathbf{Gr}^{\pi} - \mathbf{Ex}^{\pi} \neq \emptyset$.

Note that the pair $(\mathbf{W}, \pi)$ in Proposition (16) is a classification problem. (In fact $(\mathbf{W}, \pi)$ is a special case of classification problem in reason of the form of π—we called the family of such problems "one-shot"; see Example (10).) We can apply Theorem (40) to the previous proposition, so obtaining:

(59) PROPOSITION: There is $\pi \subseteq \mathcal{L}_{sen}$ such that $\mathbf{Gr}^{\pi} - \mathbf{Co}^{\pi} \neq \emptyset$.

The proposition gives a formal meaning to the liberal conception of coalitional solvability modeled by $\mathbf{Gr}^{\pi}$-solvability in comparison to noncoalitional learning of the paradigm of $\mathbf{Co}^{\pi}$-solvability.

6.2 Uncountable Problems

The following result generalizes Theorem (56) to possibly uncountable classes of structures. Despite the latter liberalization the hypotheses of the next theorem are those of Theorem (56).

(60) THEOREM: Let $\pi \subseteq \mathcal{L}_{sen}$ and let class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then $\mathbf{A}$ is $\mathbf{Gr}^{\pi}$-solvable.

The proof of the theorem is analogous to the proof of Theorem (56) with the exception that here we have to consider in a special way the case of two interacting learning agents with identical first and second moves. In fact, if class $\mathbf{A}$ is uncountable, then it may happen that two identical moves, say $v_m \doteq v_m$ be the first move of the two agents and $v_n \doteq v_n$ be the second move, are associated in some precise sense with two different structures $\mathcal{A}$, $\mathcal{B}$ in $\mathbf{A}$, since there are only countably many pairs of basic formulas built on countable vocabulary $\mathcal{L}$ along with countably many variables. A consequence is that the simple convention on who has to be the leader does not work in the case in question. To overcome the problem the idea for the proof is adding a further requirement to the convention we already encountered in the proof of Theorem (56). More precisely, we add the requirement that as soon as two environments for, respectively, $\mathcal{A}$ and $\mathcal{B}$, enumerated by some learning agent in the group by interacting with some other agent in the group differ, the learning agent who acts the basic formula with lowest Gödel number is the leader, while the other learning agent is the non-leader (partner). So, the agent who is not the leader copies the leader's moves starting from the first position in the play where the leader and the partner's move differ. Notice that if the two agents enumerate with each other the same environment, then there is no need of any leadership, since the agents are associated with the same structure, that is, the agents' behaviors coincide. In summary, if the agents' first and second moves coincide, then the agents face the problem of what structure they should choose and describe in the limit to the end of cooperation. To see how the agents may solve this problem, we reformulate the informal argument above as follows. For doing this, we need a technical definition.

(61) DEFINITION: Let $\sigma \in \mathit{SEQ}$ and learning agent $\langle \Psi , \mathbf{A} \rangle$ be given. We define the **communication sequence** $\overline{(\Psi(\sigma))_0} \in \mathit{SEQ}$ by induction on the length of σ as follows. *Base*: $\overline{(\Psi(\emptyset))_0} = (\Psi(\emptyset))_0$. Suppose that $\overline{(\Psi(\tau))_0}$ is defined for $\tau \in \mathit{SEQ}$. Given $\beta \in \mathcal{L}_{basic}$, define $\overline{(\Psi(\tau\beta))_0} = \overline{(\Psi(\tau))_0}{}^\frown (\Psi(\tau\beta))_0$.

Note that when $\overline{(\Psi(\sigma))_0}$ is defined, $|\overline{(\Psi(\sigma))_0}| > 0$. Also note that the definition does not depend on the learning agent's background world $\mathbf{A}$.

Proof of Theorem (60): Let $\pi = \{\theta_i \in \mathcal{L}_{sen} \,|\, i \in N\}$. Suppose that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Let $e^{\mathcal{A}}$ be an environment for $\mathcal{A} \in \mathbf{A}$ and let $\theta_m \in \pi$ be such that $\mathcal{A} \models \theta_m$. Given $n \in N$, we define learning agent $\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ as follows.

(*a*) $\Psi_{<\mathcal{A},n,m>}(\emptyset) = \langle\, v_n \doteq v_n \,, \theta_m \,\rangle$.
(*b*) For all $\sigma \in \mathit{SEQ}$ with $|\sigma| = 1$, $\Psi_{<\mathcal{A},n,m>}(\sigma) = \langle\, v_m \doteq v_m \,, \theta_m \,\rangle$.
(*c*) For all $\sigma \in \mathit{SEQ}$ with $|\sigma| > 1$, the following cases arise.

Case 1: $\sigma_0 = v_k \doteq v_k$ and $\sigma_1 = v_j \doteq v_j$, where either $k > n$ or $k = n$ and $j > m$. [In words, $<k,j>$ follows $<n,m>$ in the lexicographic order.] Then, $\Psi_{<\mathcal{A},n,m>}(\sigma) = \langle\, \mathit{last}(\sigma) \,, \theta_j \,\rangle$.

[$\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ is not the leader.]

Case 2: For all $k, j \in N$, either $\sigma_0 \neq v_k \doteq v_k$ or $\sigma_1 \neq v_j \doteq v_j$, or $\sigma_0 = v_{k'} \doteq v_{k'}$ and $\sigma_1 = v_{j'} \doteq v_{j'}$, where either $n > k'$ or $n = k'$ and $m > j'$. [In words, $<n,m>$ follows $<k',j'>$ in the lexicographic order.] Then, $\Psi_{<\mathcal{A},n,m>}(\sigma) = \langle\, (e^{\mathcal{A}})_{|\sigma|-2} \,, \theta_m \,\rangle$.

[$\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ is the **leader**.]

Case 3: $\sigma_0 = v_n \doteq v_n$ and $\sigma_1 = v_m \doteq v_m$. By the definition of sequence $\overline{(\Psi_{<\mathcal{A},n,m>}(\sigma^-))_0} \in \mathit{SEQ}$ of the finite initial moves of $\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ on σ^- (Definition (61)), we may compute $\tau = \overline{(\Psi_{<\mathcal{A},n,m>}(\sigma^-))_0}$. Observe that $|\tau| = |\sigma|$. We now consider the following two subcases:

Subcase 1: $\tau = \sigma$. Then $\Psi_{<\mathcal{A},n,m>}(\sigma) = \langle\, (e^{\mathcal{A}})_{|\sigma|-2} \,, \theta_m \,\rangle$.

[Does not matter if $\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ is the leader or not: Leadership does not influence success.]

Subcase 2: $\tau \neq \sigma$. Let $i \in N$, $i < |\tau|$ be minimal such that $\tau_i = \sigma_i$ and let $\lceil \tau_i \rceil$ and $\lceil \sigma_i \rceil$ denote the Gödel number of τ_i and σ_i, respectively. If $\lceil \tau_i \rceil < \lceil \sigma_i \rceil$ then $\Psi_{<\mathcal{A},n,m>}(\sigma) = \langle\, (e^{\mathcal{A}})_{|\sigma|-2} \,, \theta_m \,\rangle$. [So $\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ is the **leader**.] If $\lceil \tau_i \rceil > \lceil \sigma_i \rceil$ then $\Psi_{<\mathcal{A},n,m>}(\sigma) = \langle\, \sigma_{|\sigma|-i} \,, \theta_m \,\rangle$. [So $\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle$ is not the leader.]

Set $\Sigma = \{\langle \Psi_{<\mathcal{A},n,m>} , \mathbf{A} \rangle \,|\, n \in N \text{ and } \theta_m \in \pi \text{ be such that } \mathcal{A} \models \theta_m\}$. Observe that

Σ is nonempty. By a recall to the convention on leadership the members of Σ are defined on, it is easily verified that every pair of agents in Σ π-coordinate. Moreover, if we have an infinite-repetition enumeration of the members of $\mathbf{A}$, then for all $\mathcal{A} \in \mathbf{A}$ and for all $\langle \Psi_{<\mathcal{A}', n, m>}, \mathbf{A} \rangle \in \Sigma$ there is $\langle \Psi_{<\mathcal{B}, k, m>}, \mathbf{A} \rangle \in \Sigma$ with $\mathcal{B} = \mathcal{A}$ and $k > n$ [so $\langle \Psi_{<\mathcal{B}, k, m>}, \mathbf{A} \rangle$ is the leader] such that $\langle \Psi_{<\mathcal{B}, k, m>}, \mathbf{A} \rangle$ enumerates with $\Psi_{<\mathcal{A}', n, m>}$ an environment for $\mathcal{A}$. Hence $\mathbf{A}$ is $\mathbf{Gr}^{\pi}$-solvable. ■

6.3 Comparison with $\mathbf{CoSoc}^{\pi}$

We compare the "restricted coordination" variant of $\mathbf{Co}^{\pi}$-solvability with the present paradigm. We show in proposition below that the two paradigms are equivalent in the most relevant cases. To proceed, we present two lemmas.

(62) Lemma: Let $\pi \subseteq \mathcal{L}_{sen}$. $\mathbf{Gr}^{\pi} \subseteq \mathbf{CoSoc}^{\pi}$.

Proof: Let $\pi = \{\theta_i \in \mathcal{L}_{sen} \mid i \in N\}$. Suppose that set Σ of learning agents $\mathbf{Gr}^{\pi}$-solves class of structures $\mathbf{K}$. Then the following conditions hold.

(63)

(a) For all $\langle \Psi, \mathbf{K} \rangle, \langle \Phi, \mathbf{K} \rangle \in \Sigma$, $\langle \Psi, \mathbf{K} \rangle$ π-coordinates with $\langle \Phi, \mathbf{K} \rangle$.

(b) For all $\langle \Phi, \mathbf{K} \rangle \in \Sigma$ and all $\mathcal{A} \in \mathbf{K}$, there is $\langle \Psi, \mathbf{K} \rangle \in \Sigma$ and $t \in N$ such that ${}_{t}|R(\Psi_0, \Phi)$ is an environment for $\mathcal{A}$.

From (63)(a) it follows that for every pair of agents in Σ with ability respectively Ψ_1 and Ψ_2, there is $s \in N$ such that ${}_{s}|R(\Psi_{10}, \Psi_2)$ is an environment for some $\mathcal{K} \in \mathbf{K}$ and for all but finitely many $n \in N$, $({}_{s}|R(\Psi_{11}, \Psi_2))_n = \theta$ for some $\theta \in \pi$ with $\mathcal{K} \models \theta$. We define a learning agent $\langle \Psi_{12}, \mathbf{K} \rangle$ from Ψ_1 and Ψ_2 as follows.

(64) Let $e = {}_{s}|R(\Psi_{10}, \Psi_2)$. Define:

(a) $\Psi_{12}(\emptyset) = \langle v_i \doteq v_i, \theta_i \rangle$, where $i \in N$ be least such that $\theta_i = \theta$.

(b) For all $\sigma \in SEQ$, if $|\sigma| > 0$ then $\Psi_{12}(\sigma) = \langle e_{|\sigma|-1}, \theta_i \rangle$.

Let $\Sigma' = \{\langle \Psi_{12}, \mathbf{K} \rangle \in \Lambda^{l} \mid \langle \Psi_1, \mathbf{K} \rangle, \langle \Psi_2, \mathbf{K} \rangle \in \Sigma\}$.

(65) Define learning agent $\langle \Phi, \mathbf{K} \rangle$ such that:

(a) $\Phi(\emptyset) = \langle v_0 \doteq v_0, \theta_0 \rangle$.

(b) For all $\sigma \in SEQ$ with $|\sigma| > 0$, if $\sigma_0 = v_i \doteq v_i$ for some $i \in N$, then $\Phi(\sigma) = \langle \mathit{last}(\sigma), \theta_i \rangle$. Otherwise, $\Phi(\sigma) = \langle \mathit{last}(\sigma), \theta \rangle$ for arbitrary $\theta \in \mathcal{L}_{sen}$.

We claim that $\langle \Phi, \mathbf{K} \rangle$ $\mathbf{CoSoc}^{\pi}$-solves $\mathbf{K}$. From (64) and the definition of Σ' it is easy to verify that $\langle \Phi, \mathbf{K} \rangle$ π-coordinates with every member of Σ'. It remains to prove that

(66) for all $\mathcal{A} \in \mathbf{K}$, there are a learning agent in Σ' with ability $\mathcal{X}$ and $k \in N$ such that ${}_{k}|R(\mathcal{X}_0, \Phi)$ is an environment for $\mathcal{A}$.

Let $\mathcal{A} \in \mathbf{K}$ and $\langle \Psi_1 , \mathbf{K} \rangle, \langle \Psi_2 , \mathbf{K} \rangle \in \Sigma$ be such that for some $k \in N$ ${}_k|R(\Psi_{10}, \Psi_2)$ is an environment for $\mathcal{A}$. Observe that by (63)(b) such k exists. From (64)(b) it follows that there is a learning agent in Σ' with ability $\mathcal{X}$ such that ${}_k|R(\mathcal{X}_0, \Psi_1)$ is an environment for $\mathcal{A}$. It follows from (65) that ${}_k|R(\mathcal{X}_0, \Phi)$ is an environment for $\mathcal{A}$. So (66) follows. ■

A question is if $\mathbf{CoSoc}^{\pi} - \mathbf{Gr}^{\pi}$ is nonempty. The next lemma says that if we limit both paradigms' scope to problems of the form $(\mathbf{A}, \pi)$, where for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$, then every problem solvable via coordination by a single learning agent is solvable by a group of learning agents. More precisely, we have:

(67) LEMMA: Let $\pi \subseteq \mathcal{L}_{sen}$ and let class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. If $\mathbf{A} \in \mathbf{CoSoc}^{\pi}$ then $\mathbf{A} \in \mathbf{Gr}^{\pi}$.

Proof: Assume the hypothesis of the lemma. Then Theorem (60) ensures that $\mathbf{A} \in \mathbf{Gr}^{\pi}$. The claim then follows trivially. ■

From the previous lemma and Lemma (62) follows the next "equivalence" result.

(68) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$ and let class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then, $\mathbf{A} \in \mathbf{Gr}^{\pi}$ if and only if $\mathbf{A} \in \mathbf{CoSoc}^{\pi}$.

We conclude the presentation of our paradigm of solvability for groups of potentially uncomputable learning agents by a simple characterization result. The next theorem thus provides a necessary and sufficient condition for the classes of $\mathcal{L}$-structures and learning contexts $\pi \subseteq \mathcal{L}_{sen}$ to be $\mathbf{Gr}^{\pi}$-solvable.

(69) THEOREM: Let $\pi \subseteq \mathcal{L}_{sen}$ and class $\mathbf{A}$ of structures be given. Then, $\mathbf{A} \in \mathbf{Gr}^{\pi}$ if and only if for all $\mathcal{A} \in \mathbf{A}$ there is $\theta \in \pi$ such that $\mathcal{A} \models \theta$.

Proof: The "if" direction is Theorem (60). For the "only if" direction, suppose $\mathbf{A} \in \mathbf{Gr}^{\pi}$. Then there is set Σ of learning agents such that:

(a) For all $\langle \Psi , \mathbf{A} \rangle, \langle \Phi , \mathbf{A} \rangle \in \Sigma$, $\langle \Psi , \mathbf{A} \rangle$ π-coordinates with $\langle \Phi , \mathbf{A} \rangle$.

(b) For all $\langle \Phi , \mathbf{A} \rangle \in \Sigma$ and for all $\mathcal{A} \in \mathbf{A}$, there are $\langle \Psi , \mathbf{A} \rangle \in \Sigma$ and $s \in N$ such that ${}_s|R(\Psi_0, \Phi)$ is an environment for $\mathcal{A}$.

Let learning agents $\langle \Psi' , \mathbf{A} \rangle$, $\langle \Phi' , \mathbf{A} \rangle$ and $\mathcal{A}' \in \mathbf{A}$ satisfy (b). From (a) it follows that there is $\theta \in \pi$ such that for all but finitely many $n \in N$, $({}_s|R(\Psi'_1, \Phi'))_n = \theta$ with $\mathcal{A}' \models \theta$. ■

7 Computable Solvability

The focus of the present section is to begin the study of coordination by narrowing the interpretation of learning agents to computable objects. Computability thus enters our theory in two ways, as learning agents can be computable in their communication function and in their background world.

7.1 Preliminaries

Historically the phrases "computable functions" and "recursive functions" have been used synonymously and refer to the class of functions that are, intuitively speaking, effectively computable. There are many equivalent formal descriptions of this class of functions (see for instance [Kle52]; also [Soa99] for a critical, historical survey). Thus, a set is said to be **computable** (**recursive**) if its characteristic function is a computable (recursive) function. A set is set is said to be **computably enumerable** (**recursively enumerable**, or *r.e.*) if it is the range of a computable (recursive) function. To avoid duplications, we will use the terms "computable", "computably enumerable" in place of "recursive", "recursively enumerable". A first-order language is called **computable** or "effectively presented" if the sets of its relation, function, and constant symbols, and the sets of variable and logical symbols are computable sets. Moreover, the functions that associate the relation and function symbols to their arity are computable. So, we assume the following

(70) CONVENTION: From now to the end of this paper, we assume that $\mathcal{L}_{form}$ is effectively presented.

There are several possibilities for the definition of a "computable structure". We say that a structure $\mathcal{S}$ is **computable** just in case there is a computable, complete assignment h to $\mathcal{S}$ such that $D^b_{\mathcal{S},\mathrm{h}}$ is a computable set. Equivalently, structure $\mathcal{S}$ is computable if the domain of $\mathcal{S}$ is computable and constants, functions, and relations of $\mathcal{S}$ are computable objects.

7.2 c-Computable Agents

For simplicity, we limit the computability constraint to learning agents' communication function (the first component of any learning agent). The official meaning of the resulting agents is that of the following definition, after which we define the computable version of the paradigms we discussed so far. Successively, we investigate the competence of computable learning agents and groups of computable agents.

(71) DEFINITION: A learning agent $\langle \Psi , \mathbf{A} \rangle$ is ***c*-computable** just in case Ψ is a computable function.

("*c*" stands for *communication*). Note that the more general terminology "computable learning agent" is reserved to learning agents whose background world is computable also.

All the paradigms of solvability that involve learning agents or learners, as does the family of paradigms of $\mathbf{Ex}^{\pi}$-solvability, may be restricted to c-computable learning agents. We rely on the following definition.

(72) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$ be given.

(*a*) $[\mathbf{Ex}^{\pi}]^{rec} = \{\mathbf{K} \,|\, \mathbf{K} \in \mathbf{Ex}^{\pi}(\Psi) \text{ for some computable } \Psi\}$.

(*b*) $[\mathbf{Gr}^{\pi}]^{rec} = \{\mathbf{K} \,|\, \mathbf{K} \in \mathbf{Gr}^{\pi}(\Sigma)\}$, where Σ is a set of *c*-computable learning agents based on $\mathbf{K}$.

(*c*) $[\mathbf{Co}^{\pi}]^{rec} = \{\mathbf{K} \,|\, \mathbf{K} \in \mathbf{Co}^{\pi}(\Psi)$ for some computable $\Psi\}$.

(*d*) $[\mathbf{CoSoc}^{\pi}]^{rec} = \{\mathbf{K} \,|\, \mathbf{K} \in \mathbf{CoSoc}^{\pi}(\Psi)$ for some computable $\Psi\}$.

Note that the definition does not require the *c*-computability (or computability in case (*a*)) of all the communication functions, but only the *c*-computability of those communication functions that serve some learning agent (learner in case (*a*)) to solve the given class of structures. As we will see below, this restriction has some consequences on the classes of solvable structures. For now, however, we consider learners and learning agents in their full generality, and ask for the computability constraint of some learners and learning agents only. The next proposition concerns some basic relations between computable and non-computable paradigms of solvability.

(73) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$.

(*a*) $[\mathbf{Ex}^{\pi}]^{rec} \subset \mathbf{Ex}^{\pi}$.

(*b*) $[\mathbf{Gr}^{\pi}]^{rec} \subset \mathbf{Gr}^{\pi}$.

(*c*) $[\mathbf{Co}^{\pi}]^{rec} \subseteq \mathbf{Co}^{\pi}$.

(*d*) $[\mathbf{Co}^{\pi}]^{rec} \supseteq \mathbf{Co}^{\pi}$.

(*e*) $[\mathbf{CoSoc}^{\pi}]^{rec} \subseteq \mathbf{CoSoc}^{\pi}$.

Note that strict inclusions above fix a cut-off in the competence of computable agents relative to the paradigm in question. The proposition is worthy of further comments. So, (*a*) is a standard result in formal learning theory. That for some $\pi \subseteq \mathcal{L}_{sen}$ there is a $\mathbf{Ex}^{\pi}$-solvable class of structures that cannot be $\mathbf{Ex}^{\pi}$-solved by any computable learner is in fact an immediate corollary of the following theorem.

(74) THEOREM: Suppose that $\mathcal{L}$ is limited to a binary predicate symbol, a unary function symbol, and two constants symbols. Then for every countable class Σ of learners there is a problem $(\mathbf{K}, \pi)$ with the following properties.[13]

(*a*) $\mathbf{K}$ is $\mathbf{Ex}^{\pi}$-solvable.

(*b*) No member of Σ $\mathbf{Ex}^{\pi}$-solves $\mathbf{K}$.

Proof: We adapt the proof of [MO98, Thm. 3.(91)]. Let vocabulary $\mathcal{L} = \{R, s, \overline{0}, a\}$, where R denotes a binary predicate, s denotes a unary function, and $\overline{0}$, a denote 0 and a constant. The term that results from $n \in N$ applications of s to 0 is denoted by $\overline{n}$. Let $\{\vartheta\} \subseteq \mathcal{L}_{sen}$ be the theory of discrete countable total orders R without maximum such that s maps each element into the R-next element. Given $i \in N$, let θ_i^+ be $a \doteq \overline{i} \wedge \exists x \forall y \, xRy$ (the second formula of the conjunction

---

[13] Recall that by Remark (12) we say that class $\mathbf{W}$ of structures is $\mathbf{Ex}^{\pi}$-solvable iff problem $(\mathbf{W}, \pi)$ is solvable.

is “there is a least point”), and let θ_i^- be $\bar{i} \doteq a \wedge \neg\, \exists x \forall y\, xRy$. Given $X \subseteq N$, let $\pi_X = \{\theta_i^+ \mid i \in X\} \cup \{\theta_i^- \mid i \notin X\}$ and let $\mathbf{K}_X = \{MOD(\{\vartheta\} \cup \{\theta\}) \mid \theta \in \pi_X\}$. Observe that $\mathbf{K}_X$ is $\mathbf{Ex}^{\pi_X}$-solvable: if $a \doteq \bar{i} \in range(e)$ for any environment e for some $\mathcal{K} \in \mathbf{K}_X$, it is easily defined a learner who guesses θ_i^+ if $i \in X$, and θ_i^- otherwise.

To prove (b), let countable class Σ of learners be given. We have to show that there is $X \subseteq N$ such that no member of Σ $\mathbf{Ex}^{\pi_X}$-solves $\mathbf{K}_X$. Because there are uncountably many subsets of N, it sufficies to prove:

(75) If X and Y are distinct subsets of N then $\mathbf{K}_X \cup \mathbf{K}_Y$ is not $\mathbf{Ex}^{\pi_X}$-solvable.

To demonstrate (75), let $X, Y \subseteq N$ and $i \in N$ be such that $i \in X$ iff $i \notin Y$. Set $\mathbf{K}^+ = MOD(\{\vartheta\} \cup \{\theta_i^+\})$ and $\mathbf{K}^- = MOD(\{\vartheta\} \cup \{\theta_i^-\})$. Observe that $\mathbf{K}^+ \cup \mathbf{K}^- \subseteq \mathbf{K}_X \cup \mathbf{K}_Y$. Hence it suffices to show that $\mathbf{K}^+ \cup \mathbf{K}^-$ is not $\mathbf{Ex}^{\pi_X}$-solvable. Let $\mathcal{K} \in \mathbf{K}^+ \cup \mathbf{K}^-$. Clearly $\mathcal{K} \models \theta_i^+$ iff $\mathcal{K}$ has a minimum and $\mathcal{K} \models \theta_i^-$ otherwise. Then, solving $\mathbf{K}^+ \cup \mathbf{K}^-$ in the paradigm $\mathbf{Ex}^{\pi_X}$ equals solving the class $MOD(\{\vartheta\})$ of the discrete countable total orders with no greatest point in the paradigm $\mathbf{Ex}^{\pi}$, where $\pi = \{\exists x \forall y\, xRy, \neg\, \exists x \forall y\, xRy\}$. But it is a standard result that $MOD(\{\vartheta\})$ is not $\mathbf{Ex}^{\pi}$-solvable (see *e.g.* Lemma (88) below), hence $\mathbf{K}^+ \cup \mathbf{K}^- \notin \mathbf{Ex}^{\pi_X}$. ■

Proposition (73)(a) then holds for $\pi \subseteq \mathcal{L}_{sen}$ be as in the theorem. We have:

(76) Corollary: Let $\mathcal{L}$ be as in Theorem (74). Then there are $\pi \subseteq \mathcal{L}_{sen}$ and $\mathbf{Ex}^{\pi}$-solvable class $\mathbf{K}$ of structures such that $(\mathbf{K}, \pi)$ is a classification problem and $\mathbf{K} \notin [\mathbf{Ex}^{\pi}]^{rec}$.

Proof: Observe that the class of computable learners form a countable set. Then apply Theorem (74). ■

To prove Proposition (73)(b) first observe that it follows immediately from the definition that $[\mathbf{Gr}^{\pi}]^{rec} \subseteq \mathbf{Gr}^{\pi}$. From the definition of $[\mathbf{Gr}^{\pi}]^{rec}$ also follows that only countable classes of structures are in $[\mathbf{Gr}^{\pi}]^{rec}$, since each structure in a class in $[\mathbf{Gr}^{\pi}]^{rec}$ must be computable (see second clause of Definition (53)(a)) and the number of computable structures is bounded by the number of Turing Machines; but Theorem (60) ensures that there are uncomputable classes in $\mathbf{Gr}^{\pi}$. Clauses (c) and (e) of Proposition (73) are a direct consequence of the definitions. Finally, to prove (73)(d) just take $\pi \subseteq \mathcal{L}_{sen}$ be a single tautology and let $\mathbf{A} \subseteq \mathbf{Co}^{\pi}$ be a class of structures with no computable member. We leave open if the inclusion in (73)(e) is proper until the end of the next section.

7.3 Two Separation Results

The following two theorems imply that there is a class $\mathbf{K}$ of structures that are $\mathbf{Co}^{\pi}$-solvable by a c-computable learning agent but not recursively $\mathbf{Gr}^{\pi}$-solvable, that is, there is no set of c-computable learning agents that $\mathbf{Gr}^{\pi}$-solves $\mathbf{K}$. To state the theorems, we need some further definitions and notation.

7.3.1 $[\mathbf{Ex}^{\pi}]^{rec}$-solvability does not imply $[\mathbf{Gr}^{\pi}]^{rec}$-solvability.

The language of arithmetic is the first-order predicate calculus with equality and with function symbols for addition, multiplication and successor, a constant symbol for zero and a binary relation symbol for "less than". The class of **Σ_0^0-formulas** or **bounded quantifier formulas** is defined as the smallest class containing all the atomic formulas, closed under propositional connectives and under **bounded quantification**: $\forall x(x < y \rightarrow \varphi)$ and $\exists x(x < y \wedge \varphi)$.

Let $\mathbf{PA}_0$ be the subsystem of Peano arithmetic obtained by restricting induction to Σ_0^0-formulas.[14]

(77) THEOREM: Suppose that $\mathcal{L}$ is the language of $\mathbf{PA}_0$. Let $\pi = \{\theta, \neg\theta\}$, where $\theta = \exists x\, \alpha(x)$ for $\alpha(x) \in \Sigma_0^0$ such that $\mathbf{PA}_0 \not\vdash \theta$ and $\mathbf{PA}_0 \not\vdash \neg\theta$. Let $\mathbf{K} = \mathit{MOD}(\mathbf{PA}_0)$. Then $\mathbf{K} \in [\mathbf{Ex}^{\pi}]^{rec}$.

Proof: We describe a learner Ψ that $\mathbf{Ex}^{\pi}$-solves $\mathbf{K}$ as follows. For every $\sigma \in \mathit{SEQ}$, if $\bigwedge\sigma \models \alpha(\overline{a})$ for some $a \in N$, $\Psi(\sigma) = \theta$. Otherwise, $\Psi(\sigma) = \neg\theta$. Clearly, Ψ is computable. Moreover, for all $\mathcal{A} \in \mathbf{K}$ and for every environment e for $\mathcal{A}$, it is easy to see that Ψ stabilizes on e either to θ or $\neg\theta$ according to as $\mathcal{A} \models \theta$ or $\mathcal{A} \models \neg\theta$. Hence, $\mathbf{K} \in [\mathbf{Ex}^{\pi}]^{rec}$. ■

The following result is the equivalent of Theorem (40) in the present case of c-computable learning agents.

(78) PROPOSITION: Let $\pi \subseteq \mathcal{L}_{sen}$ and let class $\mathbf{A}$ of structures be such that for every $\mathcal{A} \in \mathbf{A}$ there is exactly one $\theta \in \pi$ such that $\mathcal{A} \models \theta$. Then $\mathbf{A} \in [\mathbf{Ex}^{\pi}]^{rec}$ if and only if $\mathbf{A} \in [\mathbf{Co}^{\pi}(\mathbf{A})]^{rec}$.

The proof of the proposition is similar to the proof of Theorem (40). Note that we can apply Proposition (78) to Theorem (77), thus obtaining immediately the following result, which is the equivalent of Theorem (77) where the paradigm $\mathbf{Co}^{\pi}$ is used in place of $\mathbf{Ex}^{\pi}$.

(79) THEOREM: Suppose that $\mathcal{L}$ is the language of $\mathbf{PA}_0$. Let $\pi = \{\theta, \neg\theta\}$, where $\theta = \exists x\, \alpha(x)$ for $\alpha(x) \in \Sigma_0^0$ such that $\mathbf{PA}_0 \not\vdash \theta$ and $\mathbf{PA}_0 \not\vdash \neg\theta$. Let $\mathbf{K} = \mathit{MOD}(\mathbf{PA}_0)$. Then $\mathbf{K} \in [\mathbf{Co}^{\pi}]^{rec}$.

The next theorem provides the key result for the separation of $[\mathbf{Gr}^{\pi}]^{rec}$ and $[\mathbf{Ex}^{\pi}]^{rec}$, equivalently $[\mathbf{Co}^{\pi}]^{rec}$ by the previous theorem, we are concerned.

(80) THEOREM: Let $\pi \subseteq \mathcal{L}_{sen}$ and class of structures $\mathbf{K}$ be as in Theorem (77). Then $\mathbf{K} \notin [\mathbf{Gr}^{\pi}]^{rec}$.

Our proof of Theorem (80) proceeds via some lemmas, some of which will also be useful elsewhere.

(81) LEMMA: Let countable structure $\mathcal{S}$ and environment e be given. Suppose that e is for $\mathcal{S}$. If e is recursive, then $\mathcal{S}$ is computable.

---

[14] [McA82] calls it "Σ_0^0-induction".

Proof: Suppose that recursive environment e is for $\mathcal{S}$ via complete assignment $g : Var \longrightarrow \mathrm{dom}(\mathcal{S})$. Let $Var(\sigma)$ denote the set $\{v_i \in Var \,|\, v_i \text{ appears in } \sigma\}$ for any finite sequence σ. We define a countable structure $\mathcal{S}'$ as follows.

(82) Every element $\overline{v}_i \in \mathrm{dom}(\mathcal{S}')$ has a name $v_i \in Var$; for all $v_i \in Var$, $\overline{v}_i \in \mathrm{dom}(\mathcal{S}')$ iff for all $j < i$, $\neg\, v_j \doteq v_i \in range(e)$.

Since environment e is recursive, condition (82) is recursive, namely: $\overline{v}_i \notin \mathrm{dom}(\mathcal{S}')$ iff there is $j < i$ such that $v_j \doteq v_i \in range(e)$. So $\mathrm{dom}(\mathcal{S}')$ and its complement are r.e., hence $\mathrm{dom}(\mathcal{S}')$ is recursive. For every predicate symbol P of arity n and for every $\overline{v}_1, ..., \overline{v}_n \in \mathrm{dom}(\mathcal{S}')$, we also define:

(83) $P\,\overline{v}_1, ..., \overline{v}_n$ iff there is $n \in N$ such that $e_n = P\,v_1, ..., v_n$. [For economy of symbols, we represent here the predicate and the predicate symbol in the same way.]

Note that (83) is a recursive condition, since environment e is recursive, that is: not $P\,\overline{v}_1, ..., \overline{v}_n$ iff there is $n \in N$ such that $e_n = \neg\, P\,v_1, ..., v_n$. From (82) and (83) it follows that $\mathcal{S}'$ is computable. Clearly the map : $\overline{v}_i \longrightarrow g(v_i)$ is a recursive isomorphism from $\mathcal{S}'$ into $\mathcal{S}$. Hence, $\mathcal{S}'$ is a computable structure. ■

Observe that the converse of the previous lemma is always true.

(84) LEMMA: [McA82] If $\mathcal{M}$ is a countable, nonstandard model of Σ^0_0-induction, then neither addition nor multiplication in $\mathcal{M}$ is recursive.

Proof: See [McA82], Proposition 3. ■

(85) COROLLARY: Almost all countable models of $\mathbf{PA}_0$ are not computable.

Proof of Theorem (80): Suppose for a contradiction that the set Σ of c-computable learning agents $\mathbf{Gr}^{\pi}$-solves $\mathbf{K}$. Then, for all $\langle \Phi\,, \mathbf{K}\rangle \in \Sigma$ and for all $\mathcal{A} \in \mathbf{K}$, there are $\langle \Psi\,, \mathbf{K}\rangle \in \Sigma$ and $k \in N$ such that ${}_k|R(\Psi_0, \Phi)$ is an environment for $\mathcal{A}$. Clearly, ${}_k|R(\Psi_0, \Phi)$ is recursive. From Lemma (81) it follows that $\mathcal{A}$ is computable, thus contradicting Lemma (84). ■

(86) PROPOSITION: There is $\pi \subseteq \mathcal{L}_{sen}$ such that $[\mathbf{Ex}^{\pi}]^{rec} - [\mathbf{Gr}^{\pi}]^{rec} \neq \emptyset$.

Proof: By Theorem (80), Theorem (77) and Proposition (78). ■

7.3.2 $[\mathbf{Gr}^{\pi}]^{rec}$-solvability does not imply $[\mathbf{Co}^{\pi}]^{rec}$-solvability.

The following question addresses Proposition (59) in the context of c-computable learning agents, namely: "Is there a class of models which is recursively $\mathbf{Gr}^{\pi}$-solvable but not recursively $\mathbf{Co}^{\pi}$-solvable?" We have a positive answer. Indeed,

(87) PROPOSITION: There is $\pi \subseteq \mathcal{L}_{sen}$ such that $[\mathbf{Gr}^{\pi}]^{rec} - [\mathbf{Co}^{\pi}]^{rec} \neq \emptyset$.

To prove the proposition we rely on some lemmas and a general theorem. The next lemma is adapted from [OSW91].

(88) LEMMA: Suppose that $\mathcal{L}$ is limited to a binary predicate symbol R. Let $\pi = \{\theta, \neg\theta\}$, where $\theta \in \mathcal{L}_{sen}$ denotes $\exists x \forall y\, xRy$ ("there is a least point with respect to R"), and let $\mathbf{K} = \{\langle N, < \rangle, \langle Z, < \rangle\}$, where "<" interprets R. Then, $\mathbf{K}$ is not $\mathbf{Ex}^{\pi}$-solvable.

For a proof of the lemma see [OSW91, Example 5].

(89) LEMMA: Let $\pi \subseteq \mathcal{L}_{sen}$ and class $\mathbf{K}$ of structures be as in the Lemma (88). Then $\mathbf{K} \notin [\mathbf{Co}^{\pi}]^{rec}$.

Proof: Observe that pair $(\mathbf{K}, \pi)$ is a classification problem so we may apply Theorem (40). From Theorem (40) and Lemma (88) it follows that $\mathbf{K} \notin \mathbf{Co}^{\pi}(\mathbf{K})$. So, by Proposition (38) $\mathbf{K} \notin \mathbf{Co}^{\pi}$ and, by Proposition (73)(c), $\mathbf{K} \notin [\mathbf{Co}^{\pi}]^{rec}$. ■

The next lemma will also be useful.

(90) LEMMA: Let $\mathcal{S}$ be any of $\langle N, < \rangle$ and $\langle Z, < \rangle$. There is a recursive environment for $\mathcal{S}$ via some computable complete assignment to $\mathcal{S}$.

The proof of the lemma is omitted. More interesting is the next theorem. To state the theorem we rely on the following definition.

(91) DEFINITION: Let $(\mathcal{A}_n : n \in N)$ be a countable class of structures. We say that $(\mathcal{A}_n : n \in N)$ is **uniformly recursive** just in case there is a total computable function E from $N \times N$ to the set of infinite sequences over $\mathcal{L}_{basic}$ such that for every $n \in N$, $\lambda i \,.\, E(n,i)$ is a (recursive) environment for $\mathcal{A}_n$.

Mimicking the proof of Lemma (55) and taking $\lambda i \,.\, E(n,i)$ to be an environment for $\mathcal{A}_n$ we have:

(92) THEOREM: Suppose that $(\mathcal{A}_n : n \in N)$ is a uniformly recursive class of structures with infinite repetitions. Let $\pi = \{\theta_i \in \mathcal{L}_{sen} \mid i \in N\}$ be such that there is a total computable function $f : N \longrightarrow N$ such that for every $i \in N$, $\mathcal{A}_i \models \theta_{f(i)}$. Then $(\mathcal{A}_n : n \in N) \in [\mathbf{Gr}^{\pi}]^{rec}$.

Proof: Let $\mathbf{A} = (\mathcal{A}_n : n \in N)$ be a uniformly recursive class of structures with infinite repetition. Observe that a uniformly recursive class may be transformed in a uniformly recursive class with infinite repetition of its members. Then there is a total computable function E from $N \times N$ to the set of infinite sequences over $\mathcal{L}_{basic}$ such that for every $n \in N$, $\lambda i \,.\, E(n,i)$ is a recursive environment for $\mathcal{A}_n$.

(93) Define learning agent $\langle \Phi^n, \mathbf{A} \rangle$ such that:

- (a) $\Phi^n(\emptyset) = \langle v_n \doteq v_n, \theta_{f(n)} \rangle$. [$\theta_{f(n)} \in \pi$ is unique such that $\mathcal{A}_n \models \theta_{f(n)}$.]
- (b) For all $\sigma \in SEQ$ with $|\sigma| > 0$, if $\sigma_0 = v_m \doteq v_m$ for some $m > n$, then $\Phi^n(\sigma) = \langle last(\sigma), \theta_{f(m)} \rangle$. [$\theta_{f(m)} \in \pi$ is unique such that $\mathcal{A}_m \models \theta_{f(m)}$.] Otherwise, $\Phi^n(\sigma) = \langle (E(n,m))_{|\sigma|-1}, \theta_{f(n)} \rangle$ [thus $\langle \Phi^n, \mathbf{A} \rangle$ is the **leader**].

Since f is total computable and $E(n,m)$ is recursive it follows from (93) that Φ^n is computable, hence $\langle \Phi^n, \mathbf{A} \rangle$ is c-computable. Now define $\Sigma = \{\langle \Phi^n, \mathbf{A} \rangle \mid n \in N\}$.

For all $\mathcal{A} \in \mathbf{A}$ and for all $\langle \Phi^n, \mathbf{A} \rangle \in \Sigma$, let $m \geq n$ be such that $E(m,n)$ is a recursive environment for $\mathcal{A}$. Since $\mathbf{A}$ is a class with infinite-repetition of its members such m exists. Then $\langle \Phi^m, \mathbf{A} \rangle \in \Sigma$ [$\langle \Phi^m, \mathbf{A} \rangle$ is the leader] and $R(\Phi_0^m, \Phi^n)$ is an environment for $\mathcal{A}$, and for all but finitely many $k \in N$, $(R(\Phi_1^m, \Phi^n))_k = \theta_{f(n)}$. So $\mathbf{A} \in \mathbf{Gr}^{\pi}(\Sigma)$ and Σ is a set of c-computable learning agents, that is, $\mathbf{A} \in [\mathbf{Gr}^{\pi}]^{rec}$. $\blacksquare$

Proof of Proposition (87): Let $\pi = \{\theta, \neg\theta\}$, $\theta \in \mathcal{L}_{sen}$ and class $\mathbf{K}$ of structures as in the Lemma (88). By Lemma (89) $\mathbf{K} \notin [\mathbf{Co}^{\pi}]^{rec}$, so that it remains to prove that $\mathbf{K} \in [\mathbf{Gr}^{\pi}]^{rec}$. From Lemma (90) it follows that $\mathbf{K}$ is uniformly recursive. Let $\theta_1 = \theta$ and $\theta_2 = \neg\,\theta$. Since the map $f(i) = \theta_i$ for $i = 1, 2$ is finite hence computable, by Theorem (92) it follows that $\mathbf{K} \in [\mathbf{Gr}^{\pi}]^{rec}$. $\blacksquare$

At the end of this subsection we return briefly on Proposition (73)(e). The next proposition thus shows that there is a class of $\mathbf{CoSoc}^{\pi}$-solvable structures for some learning context π that cannot be $\mathbf{CoSoc}^{\pi}$-solved by any c-computable learning agent.

(94) PROPOSITION: Suppose that $\mathcal{L}$ is the language of $\mathbf{PA}_0$. Let $\pi = \{\theta, \neg\,\theta\}$, where $\theta = \exists x\, \alpha(x)$ for $\alpha(x) \in \Sigma_0^0$ such that $\mathbf{PA}_0 \not\vdash \theta$ and $\mathbf{PA}_0 \not\vdash \neg\,\theta$. Then $\mathbf{CoSoc}^{\pi} - [\mathbf{CoSoc}^{\pi}]^{rec} \neq \emptyset$.

Proof: Let $\mathbf{K} = \mathit{MOD}(\mathbf{PA}_0)$. Then by Theorem (69) $\mathbf{K} \in \mathbf{Gr}^{\pi}$ and, by Lemma (62), $\mathbf{K} \in \mathbf{CoSoc}^{\pi}$. So it remains to show that $\mathbf{K} \notin [\mathbf{CoSoc}^{\pi}]^{rec}$. However, this is easily shown by applying in order Corollary (85), Lemma (81) and observing that no c-computable learning agent can enumerate with any learning agent an environment for an uncomputable structure. $\blacksquare$

7.4 Recursive Environments

Suppose that *all* the learning agents we consider are c-computable. Then, by force of Lemma (81) each of such agents eventually enumerates by interacting with some distinct agent the basic diagram of a computable structure. A consequence of c-computability extended to every agent then would be that some class of computable models is $[\mathbf{Co}^{\pi}]^{rec}$-solvable but not solvable in this extended paradigm.

In this subsection we rework some of Definition (72) and advance and discuss some variants of the paradigms introduced therein. In particular, it will be shown that a new paradigm based on π-coordination where *all* the learning agents are c-computable is properly included in the paradigm of $[\mathbf{Co}^{\pi}]^{rec}$-solvability, in the sense that for given $\pi \subseteq \mathcal{L}_{sen}$, the collection of structures solvable in the former is solvable in the latter, but the converse does not hold in some cases. In doing this, we will assume that all the basic components of the new paradigm are those of older paradigm of $[\mathbf{Co}^{\pi}]^{rec}$-solvability.

The next definition introduce the variant we intend to investigate. Recall that Λ^l denotes the class of all learning agents.

(95) DEFINITION: Let $\pi \subseteq \mathcal{L}_{sen}$ be given.

(a) $[\mathbf{Ex}^{\pi}]^{rec-strong} = \{\mathbf{K} \mid \mathbf{K} \in [\mathbf{Ex}^{\pi}]^{rec}$ and each learner is computable$\}$.
(b) $[\mathbf{Gr}^{\pi}]^{rec-strong} = \{\mathbf{K} \mid \mathbf{K} \in [\mathbf{Gr}^{\pi}]^{rec}$ and each $\Psi \in \Lambda^{l}$ is c-computable$\}$.
(c) $[\mathbf{Co}^{\pi}]^{rec-strong} = \{\mathbf{K} \mid \mathbf{K} \in [\mathbf{Co}^{\pi}]^{rec}$ and each $\Psi \in \Lambda^{l}$ is c-computable$\}$.

Thus, the only difference between the paradigms $[\cdot]^{rec-strong}$ and $[\cdot]^{rec}$ is this: "learning agent", or "learner" in the case of paradigm $\mathbf{Ex}^{\pi}$, now means a *c-computable* learning agent (computable learner in the case of $\mathbf{Ex}^{\pi}$). All other aspects of the paradigms are the same as before, including the criterion of success given for the uncomputable version of each paradigm.

It is easy to see that

$$[\mathbf{Ex}^{\pi}]^{rec-strong} = [\mathbf{Ex}^{\pi}]^{rec},$$
$$[\mathbf{Gr}^{\pi}]^{rec-strong} = [\mathbf{Gr}^{\pi}]^{rec}.$$

Moreover, equally easy is to verify that

$$[\mathbf{Co}^{\pi}]^{rec-strong} \subseteq [\mathbf{Ex}^{\pi}]^{rec}.$$

The next proposition shows that the inclusion is sometimes proper.

(96) PROPOSITION: Suppose that $\mathcal{L}$ is the language of $\mathbf{PA}_0$. Let $\pi = \{\theta, \neg\theta\}$, where $\theta = \exists x\, \alpha(x)$ for $\alpha(x) \in \Sigma_0^0$ such that $\mathbf{PA}_0 \not\vdash \theta$ and $\mathbf{PA}_0 \not\vdash \neg\theta$. Then $[\mathbf{Ex}^{\pi}]^{rec} - [\mathbf{Co}^{\pi}]^{rec-strong} \neq \emptyset$.

Proof: Let $\mathbf{K} = MOD(\mathbf{PA}_0)$. By Theorem (77) $\mathbf{K} \in [\mathbf{Ex}^{\pi}]^{rec}$, so it is suffices to show that:

(97) $\mathbf{K} \notin [\mathbf{Co}^{\pi}]^{rec-strong}$.

To demonstrate (97), observe that there is no learning agent who is $\mathbf{K}$-centered in π, since each environment for each structure in $\mathbf{K}$ is not recursive by Lemma (81) along with Corollary (85), and hence there is no c-computable learning agent that enumerates an environment for some structure in $\mathbf{K}$ in a recursive way. It follows that there is no learning agent (c-computable or not) that π-coordinates with all $\mathbf{K}^{\pi}$-centered learning agents. ■

(98) COROLLARY: Let $\pi \subseteq \mathcal{L}_{sen}$ be as in Proposition (96). Then $[\mathbf{Co}^{\pi}]^{rec} - [\mathbf{Co}^{\pi}]^{rec-strong} \neq \emptyset$.

Proof: By Theorem (79) $\mathbf{K} \in [\mathbf{Co}^{\pi}]^{rec}$. ■

8 Final Notes

The paradigm of $\mathbf{Ex}^{\pi}$-solvability in Section 2 is with minor exceptions the model of *X-solvability* introduced in [OdJMW97] and generalized in [MO98]. In particular, all formal definitions of our pure solvability paradigm (see Section 2.2) are those of [OdJMW97, Sections 6 and 7] with minor changes. Classification problems (see Example (10)(a)) are problems of the form $(T, \{\theta_1 \ldots \theta_n\})$ in [MO98]. Paradigms of "truth-detection" of the kind mentioned in Example (9)(a) are advanced and discussed, among others, by [KG89] and [OW89, OSW91].

Equilibria in a contextual coordination game are similar to sequential equilibria in game theory; see *e.g.* [OR94, Ch. 12] for a sample link to related literature. However, to what formal extent there exists similarity is an interesting question we leave open to future work.

The strategy used by the agent in the proof of Proposition (31) is TIT FOR TAT: The agent starts with an action that does not limit future coordination; then he copies the preceding move of the partner. The condition of "self-centering" imposed by the proposition's hypothesis then makes the agent's strategy winning.

The class of structures in Lemma (36) is proved to be learnable in an equivalent sense of $\mathbf{Ex}^{\pi}$-solvability in [OW86, Example 3.4A]; Lemma (37) reformulates [OW86, Example 3.5B].

The proofs of Lemma (55), Theorem (56), and Theorem (60) are examples in increasing order of generality of a winning strategy in obtaining pairwise coordination in a set of agents sharing some common knowledge. In other words, in the terminology often used in this paper, this means that any set of agents with identical background knowledge, some of which on the unique the definition of leadership to adopt—namely, the answer to a question like "who of two agents should be the leader?" is guaranteed to coordinate pairwise if each agent in the set follows the strategy suggested in the proofs in question. Similar winning strategies by **leadership** are widely studied in game theory and artificial intelligence, even if often from different perspective and formal development. Some pointers to the literature are [Sch98], [Lew69], [FHMV95, Lew69, RZ94], and [ST92, ST97]. In particular, in game theory a leadership strategy is usually referred to as a "Stackelberg strategy" [AH92]. The idea of a "Stackelberg game" as reported for instance in [OR94, Example 97.3] is that of two players interacting extensively by perfect knowledge of the game, and such that a "leader" chooses an action and a "follower", informed of the leader's choice takes an action as well. So, perfect knowledge is required to our agents so as to fix the rule which leadership arises; while extensive interaction is formalized in our framework by using interaction sequences.

The paradigm $[\mathbf{Ex}^{\pi}]^{rec}$ is a less general variant of the paradigm of computable solvability in [MO98, Ch. 3]. Specifically, see Proposition (73)(a) in relation with [MO98, Cor. 3.(93)]; also compare Theorem (74) to Theorem 3.(91) in Martin and Osherson's work.

Lemma (84) is due to [McA82] and extends a well-known result of Tennenbaum [Ten59], which is usually stated to the effect that a nonstandard (countable) model of Peano arithmetic cannot have both recursive (computable) addition and recursive (computable) multiplication; see also [Coh66, pp. 48-49] and [How72].

The hypothesis of Section 7.4 that all the learning agents in the paradigm $[\mathbf{Co}^{\pi}]^{rec-strong}$ are c-computable leads to questions that traditional recursive theoretic learning theory of language identification answers by restricting "texts" to be recursive (see for instance [JORS99, Sec. 8.2.3]). A difference is that language identification from recursive texts fails to yield any advantage over identification from arbitrary texts, and it does not show any disadvantage either (cf. [JORS99, Prop. 8.34]); while Corollary (98) says that this does not happen in comparing $[\mathbf{Co}^{\pi}]^{rec-strong}$ and $[\mathbf{Co}^{\pi}]^{rec}$. Some pointers to the literature on coalition formation among computable agents can be found among others in [SL97].

Background material on the concepts mentioned in Section 7 can be found in any textbook; our favorites are [Men87, Ch. 3] and [BM77, Ch. 7].

References

[Ago01] A. Agostini. *Paradigms of Coordination and Solvability.* PhD thesis, LOMIT-XI, Dip. di Matematica, Università di Siena, March 2001.

[AH92] R. J. Aumann and S. Hart. *Handbook of Game Theory - With Economic Applications, I.* Handbooks in Economics, v. 11. Elsevier Science B.V., 1992.

[BM77] J. L. Bell and M. Machover. *A Course in Mathematical Logic.* North-Holland Publishing Company, Amsterdam, 1977.

[Coh66] P. J. Cohen. *Set Theory and the Continuum Hypothesis.* W.A. Benjamin, MA, 1966.

[FHMV95] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge.* The MIT Press, Cambridge, MA, 1995.

[Gol67] E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[How72] P. E. Howard. A proof of a theorem of Tennenbaum. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 18:111–112, 1972.

[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. The MIT Press, Cambridge, MA, 1999.

[Kei77] H. J. Keisler. Fundamentals of model theory. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 47–104. North-Holland Publishing Company, 1977.

[KG89] K. T. Kelly and C. Glymour. Convergence to the truth and nothing but the truth. *Philosophy of Science*, 56(2):185–220, 1989.

[Kle52] S. C. Kleene. *Introduction to Metamathematics.* North-Holland Publishing Company, 1952.

[Lew69] D. K. Lewis. *Conventions. A Philosophical Study.* Harvard University Press, Cambridge, MA, 1969.

[McA82] K. McAloon. On the complexity of models of arithmetic. *Journal of Symbolic Logic*, 47:403–415, 1982.

[Men87] E. Mendelson. *Introduction to Mathematical Logic, 3rd edition.* The Wadsworth & Brooks/Cole mathematics series. Wadsworth, Monterey, CA, 1987.

[Mil99] T. S. Millar. Pure Recursive Model Theory. In E. R. Griffor, editor, *Handbook of Computability Theory*, Studies in logic and the fundations of mathematics, v. 140, pages 507–532. Elsevier Science B.V., 1999.

[MO98] E. Martin and D. Osherson. *Elements of Scientific Inquiry.* The MIT Press, Cambridge, MA, 1998.

[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 B.V., 1997.

[OR94] M. J. Osborne and A. Rubinstein. *A Course in Game Theory.* The MIT Press, Cambridge, MA, 1994.

[OSW91] D. Osherson, M. Stob, and S. Weinstein. A universal inductive inference machine. *Journal of Symbolic Logic*, 56(2):661–672, 1991.

[OW86] D. Osherson and S. Weinstein. Identification in the limit of first order structures. *Journal of Philosophical Logic*, 15:55–81, 1986.

[OW89] D. Osherson and S. Weinstein. Paradigms of truth-detection. *Journal of Philosophical Logic*, 18:1–42, 1989.

[RZ94] J. S. Rosenschein and G. Zlotkin. *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers.* The MIT Press, Cambridge, MA, 1994.

[Sch98] K. H. Schlag. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78:130–156, 1998.

[SL97] T. W. Sandholm and V. R. Lesser. Coalitions among computationally bounded agents. *Artificial Intelligence*, 94(1-2):99–137, 1997.

[Soa99] R. I. Soare. The History and Concept of Computability. In E. R. Griffor, editor, *Handbook of Computability Theory*, Studies in logic and the fundations of mathematics, v. 140, pages 3–36. Elsevier Science B.V., 1999.

[ST92] Y. Shoham and M. Tennenholtz. Emergent conventions in multi-agents systems: Initial experimental results and observations. In C. Rich, W. Swartout, and B. Nebel, editors, *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR-91)*, pages 225–231, Cambridge, 1992.

[ST97] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94(1-2):139–166, 1997.

[Ten59] S. Tennenbaum. Non-archimedean models of arithmetic. *Notices of the American Mathematical Society*, 6:207, 1959.