Mathematica Balkanica

New Series Vol. 25, 2011, Fasc. 3

Nondeduciblity on Strategies in the Temporal Logic of Knowledge

Cătălin Dima[1], *Constantin Enea*[2]

We provide a syntactic characterization of Nondeducibility on Strategies in CTL* with knowledge and past time operators, based on prior work by Halpern and O'Neill. Our characterization is provided by means of a number of axioms that have to be satisfied by formulas specifying sets of strategies.

1 Introduction

Information flow is one of the main techniques that ensure confidentiality. There are several intuitions behind information flow, and they have been informally synthesized in statements about the (non-)dependence of high-level ("confidential") activity and low-level ("public") observation and/or deduction. Recently, Halpern and O'Neill [10] have investigated the possibility to relate some models of information flow with the semantic framework of Temporal Logics of Knowledge (TLK) [9]. They have shown that Generalized Noninterference (GNI) [15], Separability (Sep) [16], as well as Nondeducibility on Strategies [21] and its probabilistic form from [12], can be expressed by using a multiagent framework. Also syntactic forms for two of these information flow models were provided, stating that a system satisfies GNI or Sep iff the low-level agent L cannot rule out that any formula ϕ (from some appropriate class of formulas ϕ in TLK) may hold – i.e. $P_L\phi$ holds – at any state of the system.

Our purpose here is to provide a syntactic formulation of *Nondeducibility on Strategies* (NDS) [21], which is a synchronous notion capturing the concept of information flow, notion that is weaker than Sep and stronger that GNI. NDS focuses on high-level strategies as means for High to interact with the system, and is inspired by previous work on computing covert channel capacity. One of the characteristics of

the synchronous model of NDS is that High cannot refuse to interact with the system contrary to the case of asynchronous and/or process-based models such as the Bisimulation-based Non-Deducibility on Compositions (BNDC) [6].

Our approach is to give axiomatic and/or syntactic characterizations of *strategy formulas* for High in a multi-agent system $\mathcal{S}$ and then, similarly to [10], to define NDS syntactically as the validity in the set of runs in the system $\mathcal{S}$ of all the formulas of the type $P_L\phi$, for any strategy formula ϕ. Strategy formulas for High are basically formulas that specify "regular" families of *pure High strategies* – i.e. strategies for High to interact with the system, in which High makes its decisions by observing only its local state. They are specified as formulas in the full Computational Tree Logic with Knowledge and Past (KCTL*P) which have to satisfy axioms for (1) *locality* (i.e. the fact that High can only observe its local state and all his decisions are the same in different runs having the same local history) (2) *independence of the future* (i.e. the impossibility for High to guess the future behavior of the system) and (3) *totality* (i.e. the fact that local states represent inputs for High, whose only opportunity to avoid a state is by choosing his *output actions*). We also give a class of formulas in the LTL with past which satisfy these axioms and characterize NDS.

As already said, Halpern & O'Neill [10] have suggested to treat NDS as a special case of their semantic framework, by including High strategies (called *protocols* in [10]) in the High local state and using an appropriate H-information function. However, they do not give a syntactic characterization for such H-information functions. Unfortunately, the semantic approach of [10] cannot be easily lifted to the syntactic level for several reasons: firstly, the trivial approach to provide a propositional symbol for each strategy – on the grounds of considering a strategy as an item that is "atomically" observable for an agent – cannot work, since it involves the manipulation of an uncountable set of propositional symbols (as there may be uncountably many strategies in a system). Secondly, putting the whole strategy in the local state at each time point i and therefore separating states that have a "common history" might also cause problems, because systems would have infinitely many states and would be difficult to specify for model-checking problems. Our approach avoids these problems by keeping the original system unchanged, and focusing on the class of formulas that can be used to specify "regular" sets of strategies in Temporal Logic.

Alternating Temporal Logic (ATL) [1] is a framework for reasoning about strategies in multi-agent systems. However ATL was not an option for specifying NDS for two reasons: firstly, NDS does not require an adversarial framework, as in ATL. In fact, between the three participants in the NDS framework (High, Low and the Environment), High and Low cooperate to produce information flow, but the Environment is not their adversary, but rather a "nondeterministic noise injector". As such, NDS is

more of a worst case scenario[1], contrary to the approaches in [11, 2], in which Environment tries to avoid information flow. Secondly, to check NDS one has to specify the existence of *two* High-strategies that can be separated by Low-observations. This does not seem to be possible in ATL in the semantic framework of the initial system. It might be possible to specify it in a *modified* system, in which transitions are labeled with pairs of H-actions – but this hides the essence of our problem: to specify syntactically what NDS means for the initial system.

The rest of the paper is divided as follows: we recall the NDS model in the next section, and the syntax and semantics of KCTL*P in Section 3. Section 4 provides the axiomatic characterization of strategy formulas, and of Nondeducibility on Strategies in KCTL*P . We also give a syntactic characterization of a class of strategy formulas. We show here that the classical formulation of NDS [21] is equivalent with the KCTL*P formulation. We end with a section of conclusions and comments.

2 Preliminaries on Nondeducibility on Strategies

Throughout the paper S^* denotes the set of finite sequences over a set S. The set of sequences over S of length m is denoted S^m while the set of sequences of length at most m is denoted $S^{\leq m}$. By ϵ we denote the empty sequence and $S^+ = S^* - \{\epsilon\}$.

A *transition system* is a tuple $Tr = (Q, R, q_0)$ where Q is the set of states, $R \subseteq Q \times Q$ is the transition relation and q_0 is the initial state. A *run of length* m in Tr is a sequence $\rho = (q_{i-1} \to q_i)_{1 \leq i \leq m}$ such that $R(q_{i-1}, q_i)$, for any $1 \leq i \leq m$, and q_0 is the initial state. We denote $\mathsf{len}(\rho) = m$ the length of the run. A *run of infinite length* is then an infinite sequence of transitions $\rho = (q_{i-1} \to q_i)_{i \geq 1}$. The set of the runs of Tr is denoted by $\mathsf{Runs}(Tr)$.

If $\rho = (q_{i-1} \to q_i)_{1 \leq i \leq m}$, then the *prefix of length* j *of* ρ, for some $1 \leq j \leq m$, is the run $\rho[1..j] = (q_{i-1} \to q_i)_{1 \leq i \leq j}$ (by an abuse of notation, we will consider that $\rho[1..0] = q_0$). We will also denote $\rho_1 \preceq \rho_2$ when ρ_1 is a prefix of length j of ρ_2, for some $0 \leq j \leq \mathsf{len}(\rho_2)$. Moreover, we will use $\rho(j)$, for any $0 \leq j \leq \mathsf{len}(\rho)$, to denote the jth state of ρ, q_j.

Definition 2.1 ([21]) *A* ***system for*** n ***agents*** *is a tuple* $\mathcal{A} = \big((I_k \mid 1 \leq k \leq n), (O_k \mid 1 \leq k \leq n), R, (i_0^k \mid 1 \leq k \leq n), (o_0^k \mid 1 \leq k \leq n)\big)$ *where* I_k *is the set of inputs of agent* k, O_k *is the set of outputs of agent* k, $i_0^k \in I_k$ *is the initial input of agent* k, $o_0^k \in O_k$ *is the initial output of agent* k, *and* $R \subseteq Q \times Q$ *is the transition relation, where* $Q = I_1 \times \cdots \times I_n \times O_1 \times \cdots \times O_n$ *is the set of global states. We require that* R *is total, that is, for any* $q \in Q$ *there exists* $q' \in Q$ *such that* $R(q, q')$.

[1] In other words, stating NDS involves no quantifier alternation and checking it reduces to a reachability problem, see [5]

Note that we consider synchronous systems in which each agent must choose an output and the inputs are received in the same time.

A system $\mathcal{A}$ is *total for k's outputs* or *k-total* if for any $q, q' \in Q$ such that $R(q, q')$, and for any $o \in O_k$, there exists $q'' \in Q$ such that $q''|_{I_k \times O_k} = (q|_{I_k}, o)$ and $R(q, q'')$. Hence, the k-totality is equivalent to the fact that agent k can at any time choose any output to continue the computation.

The *k-projection* of an $\mathcal{A}$-run $\rho = (q_{i-1} \to q_i)_{1 \leq i \leq m}$, for any $1 \leq k \leq n$, is the run $\rho|_k = (q_{i-1}|_{I_k \times O_k} \to q_i|_{I_k \times O_k})_{1 \leq i \leq m}$. This run is what agent k sees when ρ happens in the system $\mathcal{A}$. We will define also the *k-input projection* $\rho|_{I_k} = (q_{i-1}|_{I_k} \to q_i|_{I_k})_{1 \leq i \leq m}$ and the *k-output projection* $\rho|_{O_k} = (q_{i-1}|_{O_k} \to q_i|_{O_k})_{1 \leq i \leq m}$.

Definition 2.2 *A **strategy** for agent k, for any $1 \leq k \leq n$, is a mapping $s : (I_k)^* \to O_k$ such that $s(\epsilon) = o_0^k$. An m-**strategy** for agent k is a mapping $s : (I_k)^{\leq m} \to O_k$ such that $s(\epsilon) = o_0^k$.*

A strategy for agent k encodes the choices that this agent makes as a function of his observations of the system states. We assume that agent k does *not* have access to the whole system state when he makes its decisions. Note also that a strategy defines an output for k in any system state and provided any system history. By $s(\epsilon)$ we will understand the choice made by observing the initial input i_0^k, which, by definition, is o_0^k. The set of strategies for agent k is denoted by Str_k.

If s_1 is an m_1-strategy for agent k and s_2 an m_2-strategy for agent k, with $m_1 \leq m_2$, we say that s_1 is a *prefix* of s_2 if $s_2(\lambda) = s_1(\lambda)$, for any sequence $\lambda \in (I_k)^{\leq m_1}$.

Given a strategy s for k and a run $\rho = (q_{i-1} \to q_i)_{1 \leq i \leq m}$, we say that s *is compatible with* ρ if for all $i = 1, \ldots, m$,

$$s(q_1^k \ldots q_i^k) = q_i|_{O_k}.$$

If s is a j-strategy for k and $\rho = (q_{i-1} \to q_i)_{1 \leq i \leq m}$, then we say that s *is compatible with* ρ if for all $i \leq \min\{m, j\}$,

$$s(q_1^k \ldots q_i^k) = q_i|_{O_k}.$$

In other words, s is compatible with ρ if $\rho|_k$ contains the sequence of "decisions" that agent k makes when acting like s, in accordance with the part of the current state that he may observe.

The set of strategies for agent k in the system $\mathcal{A}$ is denoted $\mathsf{Str}_k(\mathcal{A})$. Formally, $\mathsf{Str}_k(\mathcal{A}) = \{s \in \mathsf{Str}_k \mid \exists \rho \in \mathsf{Runs}(\mathcal{A})$ such that s is compatible with $\rho\}$

The set of *behaviors observable by i when j acts following strategy s* is

$$\mathsf{Obs}_i(s) = \{\rho|_i \mid s \text{ is compatible with } \rho\}$$

Definition 2.3 *A system $\mathcal{A}$ satisfies* ***nondeducibility on strategies from i to j*** *(denoted* $\mathsf{NDS}(i,j)$*) if* $\mathsf{Obs}_i(s_1) = \mathsf{Obs}_i(s_2)$*, for any* $s_1, s_2 \in \mathsf{Str}_j(\mathcal{A})$.

If we consider a system with two agents, a high-level agent H and a low-level agent L, then $\mathsf{NDS}(L,H)$ is exactly the nondeduciblity on strategies from [21].

Example 2.1 *Figure 1 shows a part of a system $\mathcal{A}$ for 2 agents, named H and L, respectively. The sets of inputs and outputs for each agent are respectively:*

$$I_H = \{i_0^H, i_1^H\},\ O_H = \{o_0^H, o_1^H\},\ I_L = \{i_0^L, i_1^L, i_2^L\},\ O_L = \{o_0^L\}.$$

The initial inputs for H and L are i_0^H and i_0^L, and the initial outputs for H and L are o_0^H and o_0^L, respectively. Note also that $\mathcal{A}$ is total for H's outputs.

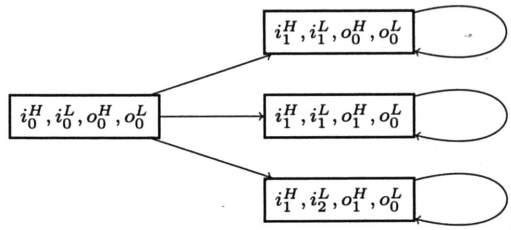

Figure 1: A system for 2 agents

Let s_1 and s_2 be two strategies for H such that $s_1(\epsilon) = o_0^H$, $s_1(i_0^H) = o_0^H$, $s_2(\epsilon) = o_0^H$ and $s_2(i_1^H) = o_1^H$. Notice that

$$\begin{aligned}\mathsf{Obs}_L(s_1) &= \{(i_0^L, o_0^L) \to (i_1^L, o_0^L) \to (i_1^L, o_0^L) \to \ldots\},\\ \mathsf{Obs}_L(s_2) &= \{(i_0^L, o_0^L) \to (i_1^L, o_0^L) \to (i_1^L, o_0^L) \to \ldots,\\ &\qquad (i_0^L, o_0^L) \to (i_2^L, o_0^L) \to (i_2^L, o_0^L) \to \ldots,\}\end{aligned}$$

which implies that $\mathcal{A}$ does not satisfy $\mathsf{NDS}(L,H)$.

3 Temporal logic of knowledge

In the following, we will enhance the systems from the previous section with labeling functions that associate to each global state a set of atomic propositions.

Definition 3.1 *Let $AP = \bigcup_{1\leq k\leq n} AP_k$ a family of atomic propositions, with $AP_k \neq \emptyset$. A **Kripke structure for** n **agents** over AP is a couple $\mathcal{K} = (\mathcal{A}, \pi)$, where $\mathcal{A}$ is a system for n agents and $\pi : Q \to 2^{AP}$ is the labeling function, such that for any two states $q, q' \in Q$ and agent k, if $q|_{I_k \times O_k} = q'|_{I_k \times O_k}$ then $\pi(q) \cap AP_k = \pi(q') \cap AP_k$.*

The semantics of a formula in the temporal logic of knowledge is given on interpreted systems which are defined as usual.

Definition 3.2 *Let $\mathcal{K}$ be a Kripke structure for n agents over some set of atomic propositions AP as above. The interpreted system corresponding to $\mathcal{K}$ is the couple $\mathcal{I}(\mathcal{K}) = (\mathsf{Runs}(\mathcal{K}), \overline{\pi})$, where $\overline{\pi} : \mathsf{Runs}(\mathcal{K}) \times \mathbb{N} \to 2^{AP}$ is the interpretation function for the atomic propositions in $\mathcal{K}$ defined by $\overline{\pi}((\rho, n)) = \pi(\rho(n))$, for any $\rho \in \mathsf{Runs}(\mathcal{K})$ and $n \in \mathbb{N}$.*

From now on, we will call a *point* any pair (ρ, n), where ρ is a run of $\mathcal{K}$ and $0 \leq n \leq \mathsf{len}(\rho)$. We will denote by $\mathsf{Points}(\mathcal{I}(\mathcal{K}))$, the set of points of the interpreted system $\mathcal{I}(\mathcal{K})$.

On the set of points $\mathsf{Points}(\mathcal{I}(\mathcal{K}))$, we will define n equivalence relations $(\sim_k | 1 \leq k \leq n)$, such that $\sim_k$ relates states being "similar" to agent k. We will adopt a *synchronous perfect-recall semantics* of knowledge and consequently, we will consider that $(\rho, m) \sim_k (\rho', m')$ whenever $m = m'$ and $\rho[1..m]|_k = \rho'[1..m]|_k$.

In the following we will use PCTL* , which is CTL* with past [17], as our supporting temporal logic. The set of formulas is defined as follows:

$$\phi = p \mid \neg\phi \mid \phi \wedge \phi \mid \bigcirc\phi \mid \Box\phi \mid \phi\mathcal{U}\phi \mid \bullet\phi \mid \blacksquare\phi \mid \phi\mathcal{S}\phi \mid \exists\phi,$$

where $p \in \bigcup_{1\leq k\leq n} AP_k$.

We define the satisfaction relation $(\rho, m) \models \phi$, where $(\rho, m) \in \mathsf{Points}(\mathcal{I}(\mathcal{K}))$ and ϕ is a temporal logic formula as follows:

1. $(\rho, m) \models \psi$ iff ψ holds by interpreting the atomic propositions in $\pi(\rho, m)$ to true and the other ones to false;
2. $(\rho, m) \models \neg\phi$ iff $(\rho, m) \not\models \phi$;
3. $(\rho, m) \models \phi_1 \wedge \phi_2$ iff $(\rho, m) \models \phi_1$ and $(\rho, m) \models \phi_2$;
4. $(\rho, m) \models \bigcirc\phi_1$ iff $(\rho, m+1) \models \phi_1$;
5. $(\rho, m) \models \phi_1\mathcal{U}\phi_2$ iff there exists $j \geq 0$ such that $(\rho, m+j) \models \phi_2$ and for every $0 \leq t < j$, $(\rho, m+t) \models \phi_1$;
6. $(\rho, m) \models \bullet\phi_1$ iff $m = 0$ or $m \geq 1 \wedge (\rho, m-1) \models \phi_1$;

7. $(\rho, m) \models \phi_1 \mathcal{S} \phi_2$ iff there exists $j \leq m$ such that $(\rho, j) \models \phi_2$ and for every $j < t \leq m$, $(\rho, t) \models \phi_1$;
8. $(\rho, m) \models \exists \phi_1$ iff there exists $\rho' \in \mathsf{Runs}(\mathcal{K})$ with $\rho'[1..m] = \rho[1..m]$ such that $(\rho', m) \models \phi_1$.

We define as syntactic sugar $\Diamond \phi = \mathsf{true} \mathcal{U} \phi$, $\Box \phi = \neg \Diamond \neg \phi$, $\phi_1 \mathcal{W} \phi_2 = \phi_1 \mathcal{U} \phi_2 \vee \Box \phi_1$, $\blacklozenge \phi_1 = \neg \blacksquare \neg \phi_1$, $\phi_1 \mathcal{B} \phi_2 = \phi_1 \mathcal{S} \phi_2 \vee \blacksquare \phi_1$ and $\forall \phi_1 = \neg \exists \neg \phi_1$, for any temporal formulas ϕ_1 and ϕ_2. Also, we will define $\mathsf{false} = \phi \wedge \neg \phi$, $\mathsf{Init} = \bullet \mathsf{false}$ and $\mathsf{true} = \neg \mathsf{false}$.

The Full Branching-Time Temporal Logic of Knowledge with Past, denoted by KCTL*P , is obtained by adding, for all agent indices $1 \leq i \leq n$, the knowledge operators K_i to the above temporal logic (e.g. [9]). The semantics of K_i is the following:

9. $(\rho, m) \models K_i \phi_1$ iff for all $\rho' \in \mathsf{Runs}(\mathcal{K})$ with $(\rho', m) \sim_i (\rho, m)$ we have that $(\rho', m) \models \phi_1$.

We say that a formula ϕ of KCTL*P is *valid* in some interpreted system $\mathcal{I}(\mathcal{K})$ and denote this by $\mathcal{I}(\mathcal{K}) \models \phi$ if $(\rho, m) \models \phi$, for all $(\rho, m) \in \mathsf{Points}(\mathcal{K})$. Also, ϕ is called *satisfiable* in $\mathcal{I}(\mathcal{K})$ if there exists some point $(\rho, m) \in \mathsf{Points}(\mathcal{I}(\mathcal{K}))$ such that $(\rho, m) \models \phi$.

We will consider a special set of atomic propositions that identify the outputs and the inputs of each agent in each state. For simplicity we will denote them in the same way as the outputs $(O_k \mid 1 \leq k \leq n)$ and the inputs $(I_k \mid 1 \leq k \leq n)$. In the following, we will consider Kripke structures $\mathcal{K} = (\mathcal{A}, \pi)$ over some set of atomic propositions that includes $\bigcup_{1 \leq k \leq n} (I_k \cup O_k)$ and such that: $\pi((i^1, \ldots i^n, o^1, \ldots, o^n)) \cap (I_k \cup O_k) = \{i^k, o^k\}$, for any global state $(i^1, \ldots i^n, o^1, \ldots, o^n) \in Q$ and $1 \leq k \leq n$.

4 Nondeducibility on strategies in the temporal logic of knowledge

In this section we give a syntactic characterization of nondeducibility on strategies in the temporal logic of knowledge. In fact, we give a syntactic characterization for the formulas that specify families of strategies for an agent j to interact with the system and then specify $\mathsf{NDS}(i, j)$ as the validity of $P_i \phi$, for any such formula ϕ. In the first subsection we give an axiomatic definition for these formulas and in the next subsection we give a class of formulas in LTL with past that completely define NDS.

4.1 A necessary condition for NDS

A PCTL* formula ϕ is *limit closed* if, whenever there exists an infinite sequence of runs $(\rho_i)_{i \geq 0}$ such that $(\rho_i, 0) \models \phi$ and $\rho_i[1..i] = \rho_{i+1}[1..i]$, then there exists a run ρ with

$\rho[1..i] = \rho_i[1..i]$ and such that $(\rho, 0) \models \phi$. Hence, ϕ defines a *safety formula*, or, in other words, the set of models of ϕ at 0 is a *closed set* in the usual topology of infinite sequences.

Definition 4.1 *Let $\mathcal{K}$ be a Kripke structure for n agents and $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system. A PCTL* formula ϕ* ***depends only on the past of agent*** *j if the following axioms hold for ϕ:*

$$\mathcal{I}(\mathcal{K}) \models \phi \to \blacksquare\Box\phi, \tag{1}$$

$$\mathcal{I}(\mathcal{K}) \models \phi \to K_j\exists\phi. \tag{2}$$

The first property (call it *extensibility*), says that once ϕ holds at a certain point on a run, then it holds throughout the whole run. Note that this implies that ϕ is limit closed.

The second property states that the choice of the truth value for ϕ at two j-similar points (ρ, m) and (ρ', m) does not depend on the set of possible continuations of ρ respectively ρ' after position m.

Example 4.1 *Let $\mathcal{K} = (\mathcal{A}, \pi)$ be a Kripke structure for 2 agents over some set of atomic propositions AP, for which a fragment is depicted in Figure 2 (the two agents will be denoted H and L, respectively). We consider that:*

- $I_H = \{i_0^H, i_1^H, i_2^H, i_3^H, i_4^H\}$, $O_H = \{o_1^H, o_2^H\}$.
- $I_L = \{i_0^L, i_1^L, i_2^L, i_3^L, i_4^L\}$, $O_L = \{o_1^L\}$.
- *The initial inputs for H and L are i_0^H and i_0^L, and the initial outputs for H and L are o_1^H and o_1^L, respectively;*
- *$\mathcal{K}$ is total for H's outputs;*

As a first example, consider the formula $\phi_1 = \blacksquare\Box\bullet(i_1^H \to (o_2^H \wedge \bullet(i_0^H \to o_1^H)))$. Intuitively, it holds in each point (ρ, k) of a run that results as (i.e., is compatible with) an application of any strategy for H in which, in the first point $(\rho, 0)$, H chooses o_1^H, and in $(\rho, 1)$, chooses o_2^H. Hence, ϕ_1 does not hold in states (ρ, k) with $k \geq 2$ and for which $(\rho, 2) \in \{(i_4^H, i_3^L, o_1^H, o_1^L), (i_4^H, i_3^L, o_2^H, o_1^L)\}$, which implies that ϕ_1 depends only on the past of H.

Example 4.2 *Another example of formula that depends only on the past of H is*

$$\phi_2 = \blacksquare\Box\big((i_1^H \to (o_2^H \wedge \bigcirc(i_2^H \to o_2^H)))\big).$$

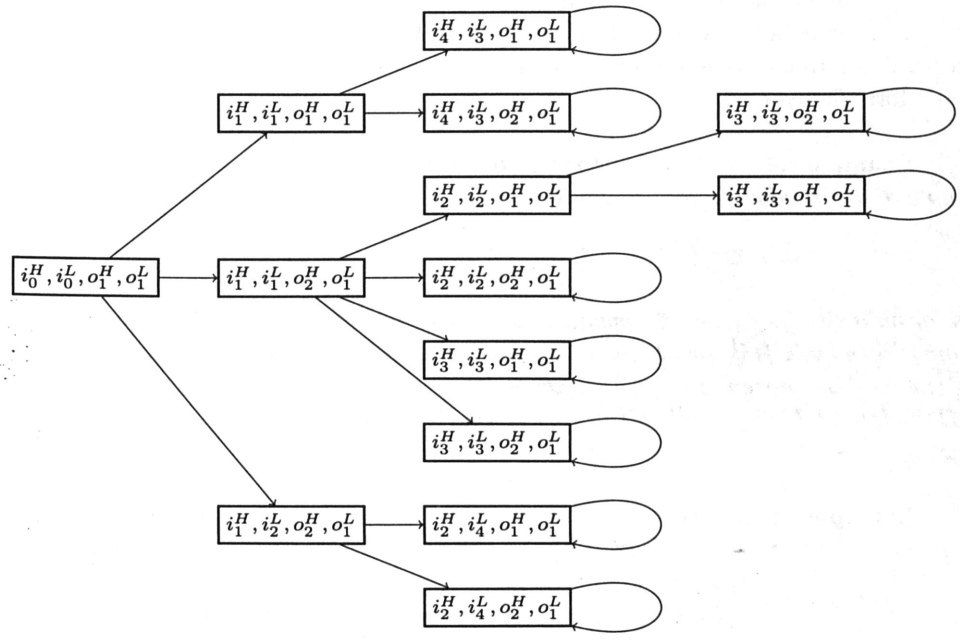

Figure 2: A Kripke structure for 2 agents

Again, it is a formula "specifying" a set of strategies in which H chooses o_1^H when he sees i_0^H, and then, if he further sees i_1^H, chooses o_2^H, and, furthermore, if he sees i_2^H, chooses once more o_2^H. This formula holds in all points (ρ, k) where ρ is either the infinite run that passes through the point $(i_2^H, i_2^L, o_2^H, o_1^L)$, or the infinite run that passes through the point $(i_2^H, i_4^L, o_2^H, o_1^L)$.

Example 4.3 *A formula that does not depend only on the past of H is $\phi_3 = \blacksquare\Box\Diamond i_3^H$. This formula may hold at a state (ρ, k) if there exists* in the future *a state (ρ, k') in which i_3^H holds, a fact which is avoided by property 2 in Definition 4.1.*

Definition 4.2 *Let $\mathcal{K}$ be a Kripke structure for n agents and $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system. A PCTL* formula ϕ is j-**admissible** if it is satisfiable, it depends only on the past of agent j and satisfies the following axiom:*

$$\mathcal{I}(\mathcal{K}) \models (\phi \wedge \exists \bigcirc \psi) \rightarrow \exists \bigcirc (\phi \wedge \psi), \tag{3}$$

for any propositional formula ψ over I_j.

A j-admissible formula describes a set of strategies for agent j, that is, a decision to issue an action (identified as a propositional formula over O_j) at some moment after passing through some sequence of states that ends in a state identified by a propositional formula over I_j.

Example 4.4 *Let $\mathcal{K}$ be the Kripke structure for 2 agents from Example 4.1. An example of an H-admissible formula is*

$$\phi_4 = \blacksquare\square\big((i_0^H \to o_1^H) \wedge (i_1^H \to o_2^H) \wedge (i_2^H \to o_1^H)\big).$$

This formula specifies a set of "memoryless" strategies for H: at each point (ρ, k), if the input seen by H is i_0^H or i_2^H, then o_1^H is chosen as the next H output, and if the input is i_1^H then o_2^H is chosen. In other words, we can associate to ϕ_4 a set S of strategies for H such that $s(\xi i_0^H) = o_1^H$, $s(\xi i_1^H) = o_2^H$ and $s(\xi i_2^H) = o_1^H$, for any $\xi \in (I_H)^*$ and $s \in S$.

Example 4.5 *An example of a formula that depends only on the past of H but it is not H-admissible is $\phi_5 = \blacksquare\square(i_3^H \to \blacklozenge i_2^H)$. To prove this take ρ_1 a run as follows: $(i_0^H, i_0^L, o_1^H, o_1^L) \to (i_1^H, i_1^L, o_2^H, o_1^L) \to (i_2^H, i_2^L, o_1^H, o_1^L) \to (i_3^H, i_3^L, o_1^H, o_1^L) \to (i_3^H, i_3^L, o_1^H, o_1^L) \to \ldots$ Then $(\rho_1, 1) \models \phi_5 \wedge \exists \bigcirc i_3^H$ but there exist no run ρ with $\rho[0..1] = \rho_1[0..1]$ such that $(\rho, 2) \models \phi_5 \wedge i_3^H$.*

Proposition 4.1 *Let $\mathcal{K}$ be a Kripke structure for n agents total for j's outputs, $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system and ϕ a j-admissible formula, for some $1 \leq j \leq n$. Then ϕ satisfies the following j-**strategy admissibility** property:*

> *For any run ρ in $\mathcal{K}$ for which $(\rho, 0) \models \phi$, there exists a strategy for j, σ which is compatible with ρ and such that for any other run ρ' in $\mathcal{K}$ with which σ is compatible, $(\rho', 0) \models \phi$.*

The converse implication also holds, in the following sense if ϕ is a PCTL formula which depends only on the past of agent j and satisfies j-strategy admissibility, then ϕ is also a j-admissible formula.*

Proof. Suppose first that ϕ is a j-admissible formula and ρ a run in $\mathcal{K}$ for which $(\rho, 0) \models \phi$. Dependence on the past implies that $(\rho, i) \models \phi$ for all $i \geq 0$. We will construct a strategy σ as required, by recursively defining all its k-length prefixes σ_k.

In fact, we will define the sequence of strategies $(\sigma_k | k \geq 0)$ such that σ_k is a k-strategy, σ_k is a prefix of σ_{k+1} and σ_k satisfies the following:

(*) σ_k is compatible with ρ and for any other run ρ' of length k with which $\sigma_k$is compatible, there exists an infinite run ρ'_1 compatible with σ_k with $\rho'_1[1..k] = \rho'$ and such that $(\rho'_1, 0) \models \phi$.

The first strategy in the sequence, σ_0, is the only strategy of length 0 that exists. Clearly, σ_0 is compatible with ρ. Now, for any run ρ' of length 0 such that σ_0 is compatible with ρ', there exists $\rho'_1 = \rho$ such that $\rho'_1[1..0] = \rho'[1..0]$ and $(\rho'_1, 0) \models \phi$.

Then, for any $k \geq 1$, suppose we have built the strategy σ_{k-1}. In the following, we will construct σ_k with the property above. Since σ_{k-1} is a prefix of σ_k, we have to define choices only for sequences of inputs of length k.

Suppose we have some sequence of inputs from I_j of length $k+1$, $\lambda = i_0 \dots i_k$, with i_0 being the initial input of agent j. If there exists no infinite run ρ in $\mathcal{K}$ such that $\rho'[1..k]\big|_{I_j} = \lambda$, then we choose randomly a value for $\sigma_k(i_1 \dots i_k)$.

On the other hand, if for all sequences of length k from O_j, $\lambda' = o_0 \dots o_{k-1}$, with o_0 being the initial output for agent j, there exists $1 \leq p \leq k-1$ such that $o_p \neq \sigma_k(i_1 \dots i_p)$ then we also choose a random value for $\sigma_k(i_1 \dots i_k)$.

Finally, in the remaining case, let $\lambda' = o_0 \dots o_{k-1}$ be the sequence of outputs from O_j such that $o_p = \sigma_k(i_1 \dots i_p)$, for all $1 \leq p \leq k-1$ (this sequence is unique because σ_k is a function). By the totality of $\mathcal{K}$ for j's outputs, there exists some run $\rho_{\lambda,\lambda'} \in \mathsf{Runs}(\mathcal{K})$ such that $\{i_p, o_p\} \subseteq \pi(\rho_{\lambda,\lambda'}(p))$, for all $0 \leq p \leq k-1$, and $i_k \in \pi(\rho_{\lambda,\lambda'}(k))$. If also, $\{i_p, o_p\} \subseteq \pi(\rho(p))$, for all $0 \leq p \leq k-1$, and $i_k \in \pi(\rho(k))$, we take $\rho_{\lambda,\lambda'} = \rho$.

From the conditions on the sequence λ', we have that σ_{k-1} is compatible with the run $\rho_{\lambda,\lambda'}[1..k-1]$ and consequently, there exists an infinite run $\rho'_{\lambda,\lambda'}$ with $\rho'_{\lambda,\lambda'}[1..k-1] = \rho_{\lambda,\lambda'}[1..k-1]$ such that $(\rho'_{\lambda,\lambda'}, 0) \models \phi$. If $\rho_{\lambda,\lambda'} = \rho$ then we choose $\rho'_{\lambda,\lambda'} = \rho$ which obviously, satisfies all the needed requirements.

Hence, $(\rho'_{\lambda,\lambda'}, k-1) \models \phi \wedge \exists \bigcirc i_k$ which by j-admissibility implies $(\rho'_{\lambda,\lambda'}, k-1) \models \exists \bigcirc (\phi \wedge i_k)$. The latter means that there exists an infinite run $\rho_{\lambda,\lambda',i_k}$ such that $\rho_{\lambda,\lambda',i_k}[1..k-1] = \rho'_{\lambda,\lambda'}[1..k-1] = \rho_{\lambda,\lambda'}[1..k-1]$, $i_k \in \pi(\rho_{\lambda,\lambda',i_k}(k))$ and $(\rho_{\lambda,\lambda',i_k}, k) \models \phi$. Again, if $\rho_{\lambda,\lambda'} = \rho'_{\lambda,\lambda'} = \rho$ then, we choose $\rho_{\lambda,\lambda',i_k} = \rho$. We will define $\sigma_k(i_1 \dots i_k) = o_k$, where $o_k \in \pi(\rho_{\lambda,\lambda',i_k}(k))$.

Now, we will prove that σ_k satisfies property (*). Take any run ρ' of length k compatible with σ_k. If we take $\lambda = \rho'[1..k-1]\big|_{I_j}$ and $\lambda' = \rho'[1..k-1]\big|_{O_j}$ and suppose that $i_k \in \pi(\rho'(k))$ then, $(\rho', k) \sim_j (\rho_{\lambda,\lambda',i_k}, k)$, where $\rho_{\lambda,\lambda',i_k}$ is the run used in the construction of σ_k.

The fact that $(\rho_{\lambda,\lambda',i_k}, k) \models \phi$ implies, by property 2, that $(\rho_{\lambda,\lambda',i_k}, k) \models K_j \exists \phi$. Hence, there exists ρ'_1 compatible with σ_k with $\rho'_1[1..k] = \rho'$ such that $(\rho'_1, k) \models \phi$ which implies, by property 1, that $(\rho'_1, 0) \models \phi$.

All that is left for proving is that the strategy σ defined by $\sigma(\lambda) = \sigma_{|\lambda|}(\lambda)$, for all $\lambda \in I_j^*$ ($|\lambda|$ denotes the length of the sequence λ) satisfies the j-strategy admissibil-

ity property. But this is an easy corollary of the fact that ϕ is invariant: suppose that ρ_3 (infinite run) is compatible with σ – therefore, for each k, $\rho_3[1..k]$ is also compatible with σ_k. By property (*), there exists an infinite run ρ_3^k with $\rho_3^k[1..k] = \rho_3[1..k]$ with $(\rho_3^k, 0) \models \phi$. Limit closure for ϕ implies then that for ρ_3, which is the unique run with $\rho_3[1..k] = \rho_3^k[1..k]$, we have that $(\rho_3, 0) \models \phi$. The fact that σ is compatible with ρ can be easily deduced from the construction above.

For the converse proof, note first that it suffices to prove the j-admissibility property only for propositional formulas of the form $\psi = i$ where $i \in I_j$. To this end, suppose that, for some $i \in I_j$, $(\rho, k) \models \phi$ and there exists ρ' with $\rho'[1..k] = \rho[1..k]$ and $(\rho', k+1) \models i$.

Note first that Property 1 implies that $(\rho, 0) \models \phi$. This implies the existence of a strategy for j, σ, compatible with ρ with the extra properties forming j-strategy admissibility.

Consider that $\pi(\rho'(p)) \cap I_j = \{i_p\}$, $\pi(\rho'(p)) \cap O_j = \{o_p\}$, for any $p \geq 0$ and $i = i_{k+1}$. Denote $o = \sigma(i_1 \dots i_{k+1})$. By the totality for j's inputs there exists an infinite run ρ'' such that $\rho''[1..k] = \rho'[1..k] = \rho[1..k]$, $\{i, o\} \subseteq \pi(\rho''(k+1))$ and σ is compatible with ρ''. By strategy admissibility, we have that $(\rho', 0) \models \phi$ which implies $(\rho'', k+1) \models \phi$. Since we also have that $(\rho', k+1) \models i$, Property 3 for ϕ is proved. ■

The following definition represents our intended restatement of NDS in KCTL*P .

Definition 4.3 *Let $\mathcal{K}$ be a Kripke structure for n agents, $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system and i, j two agents ($i \neq j$). We say that* ***agent i cannot deduce j-strategies in*** *$\mathcal{I}(\mathcal{K})$ (denoted* $\mathsf{ANS}(i,j)$*) if for any j-admissible formula ϕ we have that:*

$$\mathcal{I}(\mathcal{K}) \models P_i\phi.$$

The following theorem gives a sufficient condition for a system to satisfy the NDS restatement from Definition 4.3.

Theorem 4.1 *Let $\mathcal{K} = (\mathcal{A}, \pi)$ be a Kripke structure for n agents and $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system. If $\mathcal{A}$ satisfies* $\mathsf{NDS}(i,j)$*, for some $1 \leq i \neq j \leq n$ and it is total for j's inputs, then $\mathcal{I}(\mathcal{K})$ satisfies* $\mathsf{ANS}(i,j)$.

Proof. By means of Proposition 4.1, we will actually prove that, if there exists a formula ϕ which depends only on the past of j and satisfies j-strategy admissibility but for which $\mathcal{I}(\mathcal{K}) \not\models P_i\phi$, then $\mathcal{A}$ cannot satisfy $\mathsf{NDS}(i,j)$.

Since ϕ is satisfiable there exists some point $(\rho, m) \in \mathsf{Points}(\mathcal{I}(\mathcal{K}))$ such that $(\rho, m) \models \phi$. By the extensibility property, we have that $(\rho, 0) \models \phi$ and consequently,

by the j-strategy admissibility, there exists a strategy for j, σ, which is compatible with ρ and satisfies: for any other run ρ' in $\mathcal{K}$ such that σ is compatible with ρ', $(\rho', 0) \models \phi$.

Now, suppose that there exists a point (ρ, m') with $(\rho', m') \not\models P_i\phi$. This implies that for any infinite run ρ'' with $(\rho'', m') \sim_i (\rho', m')$ we have that $(\rho'', m') \not\models \phi$. The latter implies that σ is not compatible with ρ''. Moreover, let σ' be one of the strategies for j compatible with ρ'.

The above results say that $\rho'[1..m']|_i \cdot \xi \notin \mathsf{Obs}_i(\sigma)$, for any $\xi \in (I_i \times O_i)^*$, and $\rho'|_i \in \mathsf{Obs}_i(\sigma')$, which imply that $\mathcal{A}$ does not satisfy $\mathsf{NDS}(i, j)$. ∎

4.2 Characterizing NDS using strategy formulas

In the following, we will define the set of strategy formulas which will be used to give a characterization of nondeducibility of strategies in the temporal logic of knowledge. This characterization shows that, in fact, $\mathsf{NDS}(i, j)$ and $\mathsf{ANS}(i, j)$ are equivalent – hence, in a certain sense, giving also a proof of the converse of Theorem 4.1.

Let $\mathcal{K} = (\mathcal{A}, \pi)$ be a Kripke structure for n agents and $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system. We define the set of *strategy formulas for agent k* as follows:

$$\phi = i \rightarrow o \mid i \rightarrow (o \wedge \bigcirc\phi) \mid \phi \wedge \phi \mid \phi \vee \phi \mid \bigcirc\phi \mid \Box\phi \mid \phi \mathcal{W} i,$$

where $i \in I_k$ and $o \in O_k$. We will denote by $\mathsf{FStr}_k(\mathcal{I}(\mathcal{K}))$ the set of strategy formulas for agent k in $\mathcal{I}(\mathcal{K})$.

Strategy formulas are meant to describe strategies. For instance, in the interpreted system $\mathcal{I}(\mathcal{K})$ from Example 4.1, the formula $\phi = i_0 \rightarrow (o_1 \wedge \bigcirc(i_1 \rightarrow (o_2 \wedge \bigcirc(i_2 \rightarrow o_1))))$ describes the set of strategies for H, s, with $s(i_1) = o_2$ and $s(i_1 i_2) = o_1$.

However, there are strategy formulas that do not correspond to sets of strategies. Take, for example, the formula $\phi_{unsat} = i \rightarrow o \wedge i \rightarrow o'$ in which $o \neq o'$ which is true only in the points (ρ, m) for which $i \notin \pi((\rho, m))$. Formula ϕ_{unsat} can be translated to the fact that the agent k (whose set of inputs contains i) should in fact not receive i as input at point (ρ, m). It should be clear that there are models in which ϕ_{unsat} is not k-admissible for the agent k.

In the rest of this section, we construct a class of strategy formulas that are k-admissible. To this end, we will define two functions $\delta^k : (I_k)^+ \times \mathsf{FStr}_k(\mathcal{I}(\mathcal{K})) \rightarrow \mathcal{FP}(O_k)$ and $\xi^k : (I_k)^+ \times \mathsf{FStr}_k(\mathcal{I}(\mathcal{K})) \rightarrow \mathsf{FStr}_k(\mathcal{I}(\mathcal{K}))$, where $\mathcal{FP}(O_k)$ represents the set of propositional formulas over O_k (the first argument will be represented as an index).

Intuitively, for each $w \in (I_k)^+$, $\delta^k_w(\phi)$ represents the propositional formula that must hold in the position reached after passing through the local states of agent k that contain the inputs from w when following a strategy which is consistent with ϕ for

agent k. On the other hand, $\xi^k_w(\phi)$ gives the formula that must hold on the *next position* after passing through the local states that contain the inputs from w. These functions are inspired from the derivatives of regular expressions [3] and are defined as follows:

- $\delta^k_i(\mathsf{true}) = \mathsf{true}$ and $\delta^k_i(\mathsf{false}) = \mathsf{false}$;
- $\delta^k_i(i \to o) = o$ and $\delta^k_{i'}(i \to o) = \mathsf{true}$, for any $i' \neq i$;
- $\delta^k_i(i \to (o \wedge \bigcirc\phi)) = o$ and $\delta^k_{i'}(i \to (o \wedge \bigcirc\phi)) = \mathsf{true}$, for any $i' \neq i$;
- $\delta^k_i(\phi_1 \wedge \phi_2) = \delta^k_i(\phi_1) \wedge \delta^k_i(\phi_2)$ and $\delta^k_i(\phi_1 \vee \phi_2) = \delta^k_i(\phi_1) \vee \delta^k_i(\phi_2)$;
- $\delta^k_i(\bigcirc\phi) = \mathsf{true}$;
- $\delta^k_i(\Box\phi) = \delta^k_i(\phi)$;
- $\delta^k_i(\phi \mathcal{W} i') = \delta^k_i(\phi)$, for any $i \neq i'$ and $\delta^k_i(\phi \mathcal{W} i) = \mathsf{true}$;
- $\xi^k_i(\mathsf{true}) = \mathsf{true}$ and $\xi^k_i(\mathsf{false}) = \mathsf{false}$;
- $\xi^k_i(i' \to o) = \mathsf{true}$, for any $i, i' \in I_k$;
- $\xi^k_i(i \to (o \wedge \bigcirc\phi)) = \phi$ and $\xi^k_i(i' \to (o \wedge \bigcirc\phi)) = \mathsf{true}$, for any $i \neq i'$;
- $\xi^k_i(\phi_1 \wedge \phi_2) = \xi^k_i(\phi_1) \wedge \xi^k_i(\phi_2)$ and $\xi^k_i(\phi_1 \vee \phi_2) = \xi^k_i(\phi_1) \vee \xi^k_i(\phi_2)$;
- $\xi^k_i(\bigcirc\phi) = \phi$;
- $\xi^k_i(\Box\phi) = \Box\phi \wedge \xi^k_i(\phi)$;
- $\xi^k_i(\phi \mathcal{W} i') = \phi \mathcal{W} i' \wedge \xi^k_i(\phi)$, for any $i \neq i'$ and $\xi^k_i(\phi \mathcal{W} i) = \mathsf{true}$;
- $\delta^k_{wi}(\phi) = \delta^k_i(\xi^k_w(\phi))$, for all $w \in (I_k)^+$;
- $\xi^k_{wi}(\phi) = \xi^k_i(\xi^k_w(\phi))$, for all $w \in (I_k)^+$.

The intuition behind the two functions above is formalized in the following result:

Proposition 4.2 *Let $\mathcal{K} = (\mathcal{A}, \pi)$ be a Kripke structure for n agents, $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system and ϕ a strategy formula for agent k. Then, for all $m \geq 0$ and for any run ρ, if we take, for all $p \geq 0$, $i_p \in \pi(\rho, p) \cap I_k$, then $(\rho, m) \models \phi$ if and only if for all $j \geq 0$,*

$$(\rho, m + j) \models \delta^k_{i_m \cdots i_{m+j}}(\phi) \wedge \bigcirc\xi^k_{i_m \cdots i_{m+j}}(\phi).$$

Proof. We first address the direct proof, which follows by structural induction:

Case $\phi = i \rightarrow o$. Note first that $\xi^k_{i_m \cdots i_{m+j}}(i \rightarrow o) = \text{true}$. Then, $(\rho, m) \models \phi$ is equivalent with $i \in \pi(\rho, m) \Rightarrow o \in \pi(\rho, m)$. By definition, $\delta^k_{i_m}(\phi) = o$ for $i_m = i$ and $\delta^k_{i_m}(\phi) = \text{true}$ for $i_m \neq i$. Consequently, $\delta^k_{i_m}(\phi)$ holds in (ρ, m) iff ϕ holds there. Moreover, $\delta^k_{i_m w}(\phi) = \text{true}$, for all $w \in (I_k)^+$ and consequently $\delta^k_{i_m \cdots i_{m+j}}$ is true in $(\rho, m + j)$, for all $j \geq 0$.

Case $\phi = i \rightarrow (o \wedge \bigcirc\phi_1)$. Then, $(\rho, m) \models \phi$ is equivalent with $i \in \pi(\rho, m) \Rightarrow (o \in \pi(\rho, m) \wedge (\rho, m + 1) \models \phi_1)$. Also $\delta^k_{i_m}(\phi) = i$ if $i_m = i$ and $\delta^k_{i_m}(\phi) = \text{true}$ otherwise, which means that $(\rho, m) \models \delta^k_{i_m}(\phi)$. Also, $\xi^k_{i_m}(\phi) = \phi_1$ which means $(\rho, m) \models \bigcirc\xi^k_{i_m}(\phi)$. For the induction step, note first that, if $i_m \neq i$ then, $\delta^k_{i_m w}(\phi) = \xi^k_{i_m w} = \text{true}$, for all $w \in (I_k)^+$ and the proof is finished. Otherwise, we use the fact that $\delta^k_{i_m w}(\phi) = \delta^k_w(\phi_1)$, for all $w \in (I_k)^+$, and the inductive hypothesis for $(\rho, m + 1) \models \phi_1$, to obtain the needed result.

Case $\phi = \phi_1 \wedge \phi_2$ **and** $\phi = \phi_1 \vee \phi_2$. These cases are obtained directly from the inductive hypothesis.

Case $\phi = \bigcirc\phi_1$. Then $(\rho, m) \models \bigcirc\phi_1$ implies $(\rho, m+1) \models \phi_1$. Using $\delta^k_{i_m}(\phi) = \text{true}$ and $\xi^k_{i_m}(\phi) = \phi_1$ we obtain the base case.

For the induction step, observe first that

$$\xi^k_{i_m w}(\phi) = \xi^k_w(\phi_1), \text{ for all } w \in (I_k)^+$$

(result that can be proved by induction on the length of w). From this, if we put $w = w'j$ with $j \in I_k$, we also get

$$\delta^k_{i_m w}(\phi) = \delta^k_j(\xi^k_{i_m w}(\bigcirc\phi)) = \delta^k_j(\xi^k_w(\phi_1)) = \delta^k_w(\phi_1)$$

and the inductive hypothesis solves the case.

Case $\phi = \Box\phi_1$. Then $(\rho, m) \models \Box\phi_1$ implies $(\rho, m + p) \models \phi_1$, for all $p \geq 0$.

We will prove by induction on j that:

$$(4) \quad \xi^k_{i_m \ldots i_{m+j}}(\Box\phi_1) = \Box\phi_1 \wedge \xi^k_{i_m \ldots i_{m+j}}(\phi_1) \wedge \xi^k_{i_{m+1} \ldots i_{m+j}}(\phi_1) \wedge \cdots \wedge \xi^k_{i_{m+j}}(\phi_1)$$

The case $j = 0$ follows directly from the definition of ξ^k. Now, suppose that (4) holds

for some j. We will prove that it holds also for $j+1$:

$$\begin{aligned}\xi^k_{i_m\dots i_{m+j+1}}(\Box\phi_1) &= \xi^k_{i_{m+j+1}}(\xi^k_{i_m\dots i_{m+j}}(\Box\phi_1))\\ &= \xi^k_{i_{m+j+1}}(\Box\phi_1)\wedge\xi^k_{i_{m+j+1}}(\xi^k_{i_m\dots i_{m+j}}(\phi_1))\\ &\quad\wedge\xi^k_{i_{m+j+1}}(\xi^k_{i_{m+1}\dots i_{m+j}}(\phi_1))\wedge\dots\wedge\xi^k_{i_{m+j+1}}(\xi^k_{i_{m+j}}(\phi_1))\\ &= \Box\phi_1\wedge\xi^k_{i_{m+j+1}}(\phi_1)\wedge\xi^k_{i_m\dots i_{m+j+1}}(\phi_1)\wedge\\ &\quad\wedge\xi^k_{i_{m+1}\dots i_{m+j+1}}(\phi_1)\wedge\dots\wedge\xi^k_{i_{m+j}i_{m+j+1}}(\phi_1).\end{aligned}$$

Following an analogous procedure and using (4), we obtain:

$$\delta^k_{i_m\dots i_{m+j}}(\Box\phi_1)=\delta^k_{i_m\dots i_{m+j}}(\phi_1)\wedge\delta^k_{i_{m+1}\dots i_{m+j}}(\phi_1)\wedge\dots\wedge\delta^k_{i_{m+j}}(\phi_1) \tag{5}$$

Now, from $(\rho,m+p)\models\phi_1$, by applying the inductive hypothesis, we obtain that $(\rho,m+p+j)\models\delta^k_{i_{m+p}\dots i_{m+p+j}}(\phi_1)\wedge\bigcirc\xi^k_{i_{m+p}\dots i_{m+p+j}}(\phi_1)$, for all $p,j\geq 0$. Consequently, $(\rho,m+j)\models\delta^k_{i_m\dots i_{m+j}}(\phi)\wedge\bigcirc\delta^k_{i_m\dots i_{m+j}}(\phi)$, for all $j\geq 0$.

Case $\phi=\phi_1\mathcal{W}i$. Let us first denote $p=\min\{k\in\mathbb{N}\mid i_{m+k}=i\}$. Since there may exist runs in which i never occurs, we take this minimum over $\mathbb{N}\cup\{\infty\}$.

We may then proceed as in the previous case, and obtain, similarly with (4) and (5), the following identities, in which $l=\min\{j,p-1\}$ (with $p-1=\infty$ when $p=\infty$):

$$\delta^k_{i_m\dots i_{m+j}}(\phi_1\mathcal{W}i)=\delta^k_{i_m\dots i_{m+l}}(\phi_1)\wedge\delta^k_{i_{m+1}\dots i_{m+l}}(\phi_1)\wedge\dots\wedge\delta^k_{i_{m+l}}(\phi_1) \tag{6}$$

$$\xi^k_{i_m\dots i_{m+j}}(\phi_1\mathcal{W}i)=\phi_1\mathcal{W}i\wedge\xi^k_{i_m\dots i_{m+l}}(\phi_1)\wedge\dots\wedge\xi^k_{i_{m+l}}(\phi_1) \tag{7}$$

This ends the proof of the direct implication.

For the converse implication we may proceed analogously. ∎

Definition 4.4 *Let $\mathcal{K}=(\mathcal{A},\pi)$ be a Kripke structure for n agents, $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system and ϕ a strategy formula for agent k. The formula ϕ is called δ-**admissible** if for any $w\in(I_k)^+$, $\delta^k_w(\phi)$ is satisfiable.*

A few words on δ-admissible strategy formulas are in order. If $\delta^k_w(\phi)$ is not satisfiable for some $w\in(I_k)^+$, then the formula ϕ just forbids the sequence of outputs w and does not offer any information about some strategy. This is why we will avoid this kind of formulas. For δ-admissible formulas, we can prove that for any strategy for agent k, s, and any $m\in\mathbb{N}$ there exists some δ-admissible strategy formula $\phi_{m,s}$ that describes the choices defined by s on sequences of outputs of length at most $m+1$ (the first output is the initial output which is represented by the empty sequence when defining strategies).

Proposition 4.3 *Let $\mathcal{K} = (\mathcal{A}, \pi)$ be a Kripke structure for n agents, $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system and s a strategy for agent k. Then, for any $m \in \mathbb{N}$, there exists an δ-admissible strategy formula $\phi_{m,s}$ for agent k, such that:*

$$\forall \rho \in \mathsf{Runs}(\mathcal{K}) : s \text{ is compatible with } \rho[1..m] \Leftrightarrow (\rho, 0) \models \phi_{m,s}.$$

Proof. We define $\phi_{m,s} = \bigwedge_{w \in (I_k)^m} \phi_{m,s,w}$, where $\phi_{m,s,w}$ describes the choices implied by s on the sequence $i_0^k w$. Let $w = i_1 \ldots i_m$ and $i_j = s(i_1 \ldots i_j)$, for all $1 \leq j \leq m$. We will define $\phi_{m,s,w}$ as follows:

$$i_0^k \rightarrow (o_0^k \wedge \bigcirc(i_1 \rightarrow (o_1 \wedge \ldots (o_{m-1} \wedge \bigcirc(i_m \rightarrow o_m)) \ldots).$$

We can easily prove that $\phi_{m,s}$ is the δ-admissible strategy formula needed in the claim of the theorem. ∎

Now, we can characterize the nondeducibility of strategies in the temporal logic of knowledge.

Theorem 4.2 *Let $\mathcal{K} = (\mathcal{A}, \pi)$ be a Kripke structure for n agents total for j's inputs, for some $1 \leq j \leq n$, and $\mathcal{I}(\mathcal{K})$ its corresponding interpreted system.*

1. *For any δ-admissible strategy formula ϕ, the formula $\blacksquare\Box(\mathsf{Init} \rightarrow \phi)$ is a j-admissible formula.*
2. *$\mathcal{A}$ satisfies $\mathsf{NDS}(i, j)$, for some $1 \leq i \neq j \leq n$, iff for any δ-admissible strategy formula ϕ:*

$$\mathcal{I}(\mathcal{K}) \models P_i \blacksquare\Box(\mathsf{Init} \rightarrow \phi). \tag{8}$$

Proof. ($\Rightarrow$) We will prove that all the formulas of the form $\blacksquare\Box(\mathsf{Init} \rightarrow \phi)$ with ϕ an admissible strategy formula are j-admissible formulas. To prove that axiom 2 from Definition 4.1 holds, let $(\rho, m), (\rho', m) \in \mathsf{Points}(\mathcal{I}(\mathcal{K}))$ such that $\mathcal{I}(\mathcal{K}) \models \blacksquare\Box(\mathsf{Init} \rightarrow \phi)$ and $(\rho, m) \sim_j (\rho', m)$. We will construct a sequence of runs from $\mathcal{K}$, $(\rho_k \mid k \geq 0)$ such that $\rho_0 = \rho'$, $\rho_{k+1}[1..m+k] = \rho_k[1..m+k]$ and $(\rho_k, p) \models \delta^j_{w_p^k}(\phi)$, for any $0 \leq p \leq m + k$, where $w_p^k = \rho_k[1..p]\big|_{O_j}$.

Suppose we have build ρ_k, for some $k \geq 0$. If $(\rho_k, m + k + 1) \models \delta^j_{w^k_{m+k+1}}(\phi)$, then we take $\rho_{k+1} = \rho_k$. Otherwise, by the j-totality of $\mathcal{K}$, there exists ρ_{k+1} such that $\rho_{k+1}[1..m+k] = \rho_k[1..m+k]$, $\rho_{k+1}(m+k+1)\big|_{I_j} = \rho_k(m+k+1)\big|_{I_j}$ and $\rho_{k+1}(m+k+1)\big|_{O_j}$ is one of the outputs for agent j that appears in the disjunction equivalent to $\delta^j_{w^k_{m+k+1}}(\phi)$.

Having build the sequence above we obtain that there exists $\rho' \in \mathsf{Runs}(\mathcal{K})$ such that $\rho''[1..m] = \rho'[1..m]$ and $\rho''(m+k) = \rho_k(m+k)$, for any $k \geq 1$. Moreover, we have that $(\rho'', p) \models \delta^j_{w_p}(\phi)$, for any $p \geq 0$, where $w_p = \rho''[1..p]\big|_{I_j}$. The latter implies, by Property 4.2, that $(\rho', 0) \models \phi$ which terminates our proof for the validity of the axiom 2 from Definition 4.1.

A similar procedure can be applied to prove that the axiom 3 from Definition 4.2 holds. Consequently, we can apply Theorem 4.1 and obtain the needed result.

($\Leftarrow$) Suppose by contradiction that $\mathcal{I}_\mathcal{K}$ has the property above and $\mathcal{A}$ does not satisfy $\mathsf{NDS}(i,j)$. Consequently, there exist two strategies $s_1, s_2 \in \mathsf{Str}_j(\mathcal{A})$ such that $\mathsf{Obs}_i(s_1) \neq \mathsf{Obs}_i(s_2)$ which implies that there exists a finite run $\rho \in \mathsf{Runs}(\mathcal{K})$ of length m, for some $m \in \mathbb{N}$, such that s_1 is compatible with ρ but for any other run $\rho' \in \mathsf{Runs}(\mathcal{K})$ with $\rho'[1..m]\big|_i = \rho\big|_i$, s_2 is not compatible with ρ'.

By Proposition 4.3, for s_2 and m, we can find an admissible strategy formula ϕ_{m,s_2} such that for any $\rho_1 \in \mathsf{Runs}(\mathcal{K})$, s_2 is compatible with $\rho_1[1..m]$ iff $(\rho_1, 0) \models \phi_{m,s_2}$.

Consequently, we have that $(\rho', m) \not\models P_i \blacksquare\Box(\mathsf{Init} \to \phi_{m,s_2})$, which contradicts the hypothesis. ∎

Example 4.6 *Let $\mathcal{K}$ be a Kripke structure for 2 agents corresponding to the system $\mathcal{A}$ from Example 2.1.*

Let $\phi = i_0^H \to (o_0^H \wedge \bigcirc(i_1^H \to o_0^H))$ be an admissible strategy formula and ρ the following run:

$$(i_0^H, i_0^L, o_0^H, o_0^L) \to (i_1^H, i_2^L, o_1^H, o_0^L) \to (i_1^H, i_2^L, o_1^H, o_0^L) \to \dots$$

We can easily notice that $(\rho, 1) \not\models P_L \blacksquare\Box(\mathsf{Init} \to \phi)$ which implies

$$\mathcal{I}(\mathcal{K}) \not\models P_L \blacksquare\Box(\mathsf{Init} \to \phi).$$

Thus, by Theorem 4.2, we obtain again that $\mathcal{A}$ does not satisfy $\mathsf{NDS}(L,H)$.

Remark 4.1 *Theorem 4.2 cannot be utilized for checking whether a given Kripke structure $\mathcal{K}$ satisfies $\mathsf{NDS}(i,j)$, since the set of strategy formulas which must be model-checked in the formula 8 is not finite. However, using results from [5] or [20] we may prove that strategy formulas corresponding with* memoryless strategies *are sufficient.*

Formally, in a system for n agents $\mathcal{A} = \big((I_k \mid 1 \leq k \leq n), (O_k \mid 1 \leq k \leq n), R, (i_0^k \mid 1 \leq k \leq n), (o_0^k \mid 1 \leq k \leq n)\big)$ a strategy $s : (I_k)^* \to O_k$ is a **memoryless strategy** if for any $w, w' \in (I_k)^*$ and $a \in I_k$, we have that $s(wa) = s'(wa)$. We denote $\mathsf{Str}_k^{memless}(\mathcal{A})$ the set of memoryless strategies for agent k in system $\mathcal{A}$.

Proposition 4.4 *A system $\mathcal{A}$ satisfies* $\mathsf{NDS}(i,j)$ *if if* $\mathsf{Obs}_i(s_1) = \mathsf{Obs}_i(s_2)$, *for any* $s_1, s_2 \in \mathsf{Str}_j^{memless}(\mathcal{A})$.

Proof. Corollary of Theorem 2 (p.13) of [5], or Theorem 4.6 (p.8) of [20] (see also Theorem 2 of [4]). ■

Using then Proposition 4.3, we may restrict the application of Theorem 4.2 to a finite number of strategy formulas.

5 Conclusions

We have investigated here the possibility to define syntactically Wittbold & Johnson's notion of Nondeducibility on Strategies. We have identified a class of KCTL*P formulas which can be used to specify families of strategies in a system, and provided a formulation of NDS in KCTL*P , similar to [10].

An interesting question is whether the *dependence on the past* axiom 2 can be expressed without the knowledge operator. A negative answer would also suggest that knowledge operators strictly increase the expressivity of CTL* (with past).

References

[1] R. Alur, T. Henzinger and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, **49**(5), 2002, 672–713.

[2] Danièle Beauquier and Ruggero Lanotte. Hiding information in multi level security systems. In *Proceedings of FAST'2006*, **4691** *Lecture Notes in Computer Science*, Springer, 2007, 250–269.

[3] Janusz A. Brzozowski. Derivatives of regular expressions. *Journal of the ACM*, **11**(4), 1964, 481–494.

[4] Fr. Cassez, R. van der Meyden and Ch. Zhang. The complexity of synchronous notions of information flow security. In *Proceedings of FOSSACS'2010*, **6014** *Lecture Notes in Computer Science*, Springer Verlag, 2010, 282–296.

[5] C. Dima, C. Enea, and R. Gramatovici. A synchronous model for information flow. Technical Report TR-LACL-2006-02, LACL, Université Paris 12, 2006.

[6] R. Focardi and R. Gorrieri. A taxonomy of security properties for process algebras. *Journal of Computer Security*, **3**(1), 1995, 5–34.

[7] R. Focardi and R. Gorrieri. Classification of Security Properties (Part I: Information Flow). In *Proceedings of FOSAD'2000*, **2171** *Lecture Notes in Computer Science*, Springer, 2001, 331–396.

[8] J.A. Goguen and J. Meseguer. Security policies and security models. In *Proceedings of the IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 1982, 11–20.

[9] Joseph Y. Halpern, Ronald Fagin, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. MIT Press, 2003.

[10] Joseph Y. Halpern and Kevin R. O'Neill. Secrecy in multiagent systems. *ACM Trans. Inf. Syst. Secur.*, **12**(1), 2008.

[11] L. Hélouët —sperr, M. Zeitoun, and A. Degorre. Scenarios and covert channels: Another game... *Electr. Notes Theor. Comput. Sci.*, **119**(1), 2005, 93–116.

[12] J. W. Gray III and P. Syverson. A logical approach to multilevel security of probabilistic systems. *Distributed Computing*, **11**(2), 1998, 73–90.

[13] H. Mantel. Possibilistic definitions of security - an assembly kit. In *Proceedings of the IEEE Computer Security Foundations Workshop*, Cambridge, UK, 2000, 185–199.

[14] F. Martinelli. Partial model checking and theorem proving for ensuring security properties. In *Proceedings of CSFW'98*, 1998, 44–52.

[15] D. McCullough. Specifications for multi-level security and a hook-up property. In *Proceedings of the IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 1987, 161–166.

[16] J. McLean. A general theory of composition for trace sets closed under selective interleaving functions. In *Proceedings of the IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 1994, 79–93.

[17] Mark Reynolds. An axiomatization of PCTL*. *Inf. Comput.*, **201**(1), 2005, 72–119.

[18] Peter Y. A. Ryan and Steve A. Schneider. Process algebra and non-interference. In *Proceedings of CSFW'99*, 1999, 214–227.

[19] D. Sutherland. A model of information. In *Proceedings of the 9th National Computer Security Conference*, 1986, 175–183.

[20] R. van der Meyden and Ch. Zhang. Algorithmic verification of noninterference properties. *Electronic Notes in Theoretical Computer Science*, **168**, 2007, 61–75.

[21] J. Todd Wittbold and Dale M. Johnson. Information flow in nondeterministic systems. In *IEEE Symposium on Security and Privacy*, 1990, 144–161.

[1] *LACL,*
Université Paris-Est Créteil Val de Marne,
61 av. du Général de Gaulle,
94010 Créteil Cedex,
FRANCE

[2] *LIAFA, CNRS UMR 7089,*
Université Paris Diderot - Paris 7,
Case 7014,
75205 Paris Cedex 13,
FRANCE