A strategic perspective on IF games*

Merlijn Sevenster†

June 23, 2007

Abstract

Hintikka and Sandu's Independence-friendly logic ([5] and [6]) has traditionally been associated with extensive games of imperfect information. In this paper we set up a strategic framework for the evaluation of IF logic la Hintikka and Sandu. We show that the traditional semantic interpretation of IF logic can be characterized in terms of Nash equilibria. We note that moving to the strategic framework we get rid of IF semantic games that violate the principle of perfect recall. We explore the strategic framework by replacing the notion of Nash equilibrium by other solution concepts, that are inspired by weakly dominant strategies and iterated removal thereof, charting the expressive power of IF logic under the resulting semantics.

1 Introduction

Game theory has proven to be a tool capable of covering the essentials of established subjects in research areas such as logic, mathematics, linguistics and computer science. Game-theoretic concepts have also been proposed in cases where traditional machinery broke down. In this paper we will study the game theory that functions as a verificational framework for *independence-friendly* (IF) first-order logic, which is a generalization of standard first-order logic (FOL).

As a semantics used for evaluating FOL, Tarski semantics is well-known and widely agreed upon. Yet this semantics cannot be used to evaluate Hintikka and Sandu's IF first-order logic, see [1]. IF logic abstracts away from the Fregean assumption that syntactical scope and semantical dependence of quantifier-variable pairs coincides. That is, in an IF logical formula, if '$\exists x$' is in the syntactical scope of '$\forall y$', the variable x can be made semantically independent of y by means of the slash operator. To evaluate IF logical formulae, Hintikka and Sandu (in [5, 6]) introduce the notion of a *semantic evaluation game*. The independency of two variables expressible in IF logic is typically

*This paper was presented at the *Prague Colloquium* on *"Logic, Games and Philosophy: Foundational Perspectives"*, organized September 2004. It has been accepted for publication in the colloquium's proceedings.

†I gratefully acknowledge Peter van Emde Boas, Denis Bonnay, Sieuwert van Otterloo, Gabriel Sandu, Tero Tulenheimo, and the anonymous reviewers for valuable discussion and comments. I thank the organizers of the Prague Colloquium for letting me present my research. Special thanks are due to Theo Janssen for publishing this paper in the ILLC pre-print series.

reflected by the corresponding semantic evaluation game being of *imperfect information*. This is in contrast to the evaluation games related to first-order formulae, they are of perfect information. Truth of an FOL or IF formula is defined in terms of its semantic evaluation game. This semantics was coined *game-theoretic semantics* (GTS) by Hintikka.

It has been noted in the literature ([12], [3]) that some IF evaluation games violate the game-theoretic principle of *perfect recall*. In game theory, games without perfect recall have not been studied extensively, one of the reasons being that it is hard to understand what real-life situations they capture — put loosely, they are not '*playable*'. Thereby also the playability of IF games is called into question.

In this paper, we set up a strategic game-theoretic framework in which IF games can be defined. We will see that truth of IF under GTS can be characterized in terms of Nash equilibria in the strategic framework. We observe that the playability issues, concerning perfect recall, evaporate in the strategic framework, yet we get so-called *coordination problems* in return.

We explore the strategic framework by replacing the notion of Nash equilibrium by other solution concepts. That is, we also define truth for IF logic in terms of weakly dominant strategies and iterated removal thereof. Naturally, changing semantics affects the truth conditions of IF formulae, a phenomenon we study in terms of the expressive power of IF logic w.r.t. the resulting semantics.

Section 2 recalls the basics of IF logic and GTS. In Section 3, we define the strategic framework and establish the connection between GTS and truth in terms of Nash equilibria. Sections 4 and 5 explore the notions of truth that result after replacing the Nash equilibrium solution concept by different ones, that are inspired by the game-theoretic notions of weak dominance and iterated removal of strategies in strategic games.

The formal results are mostly given without proof. We hope to make an extended version of this paper, containing full proofs, available soon.

2 IF logic and game-theoretic semantics

The program of *quantifier independence*, as founded by [4] and later [5], is concerned with abstracting away from the Fregean assumption that the syntactical scope and binding of quantifiers in first-order logic coincide. The syntax of *independence-friendly first-order logic* as proposed by [5] extends FOL, in the sense that, for example, if

$$\forall x_1 \exists x_2 \ldots \forall x_{n-1} \exists x_n \; R(x_1, \ldots, x_n)$$

is a FOL sentence containing the n-ary predicate R, then

$$(\forall x_1/X_1)(\exists x_2/X_2)\ldots(\forall x_{n-1}/X_{n-1})(\exists x_n/X_n)\; R(x_1, \ldots, x_n) \tag{1}$$

is an IF sentence, provided that $X_i \subseteq \{x_1, \ldots, x_{i-1}\}$. The variable x_i is intuitively meant to be *independent* of the variables in X_i, although it appears under their syntactical scope.

Definition 1 *In this paper* FOL *denotes the smallest set of first-order* sentences, *that are in prenex normal form and in which every variable is quantified exactly once. We* *will assume them being of the form*

$$Q_1 x_1 \dots Q_n x_n \; R(x_1, \dots, x_n), \tag{2}$$

where $Q_i \in \{\exists, \forall\}$*. If no confusion arises we will abbreviate any string of variables* $x_1, x_2, \dots$ *using* $\overline{x}$*.*

The reader has noted that the language we call FOL is really a simple version of first-order logic. This simplification streamlines notation considerably when we define the IF language, without affecting the contention of this paper. Analogous to [5] we define the syntax of IF logic in terms of FOL, as follows.

Definition 2 *The language* IF *is obtained from* FOL *by repeating the following procedure a finite number of times: if* $\phi \in$ FOL*, then*

> *If '*$Q_i x_i \; \psi$*' occurs in* ϕ*, then it may be replaced by '*$(Q_i x_i / X_i) \; \psi$*', where* $Q_i \in \{\exists, \forall\}$ *and* $X_i \subseteq \{x_1, \dots, x_{i-1}\}$*.*

Since sentences in FOL *are assumed to be as in* (2)*, sentences of* IF *will be of the form*

$$(Q_1 x_1 / X_1) \dots (Q_n x_n / X_n) \; R(x_1, \dots, x_n), \tag{3}$$

writing '$Q_i x_i$*' instead of '*$(Q_i x_i / \emptyset)$*'.*

In $\phi \in$ FOL containing the strings '$Q_i x_i$' and '$Q_j x_j$', variable x_j depends on x_i iff $i < j$. In IF this linear ordering of dependency is given up — the quantifiers of IF sentences are *partially ordered*. The first partially ordered quantifier, also known as *Henkin quantifier*, appeared in [4]. For later usage, we formalize variable dependence by means of a binary relation. To this end let the set $Var(\phi) = \{x_1, \dots, x_n\}$ denote the variables for the IF formula ϕ as in (3). Then, $B_\phi \subseteq Var(\phi) \times Var(\phi)$ is ϕ's *dependency relation*, such that for every $x_i, x_j \in Var(\phi)$

$$(x_i, x_j) \in B_\phi \quad \text{if} \quad i < j \;\text{ and }\; x_i \notin X_j.$$

Truth of an IF sentence is evaluated relative to a suitable *model* $M = (D, I, p)$ in which we distinguish a *domain* D of objects; an *interpretation function* I, that determines the extension of relation symbols; and an *assignment* p that assigns an object from the domain D to each variable. [6] associate with every $\phi \in$ IF and suitable model M a *semantic evaluation game* $g(\phi, M)$. The game is played by two players, called E and A, that control the existential and universal quantifiers in ϕ. In $g(\phi, M)$ the players and quantifiers are associated through the *player function* P, that is the function such that $P(\exists) = E$ and $P(\forall) = A$. Intuitively, $g(\phi, M)$ proceeds as follows:

> $g((Q_i x_i / X_i) \; \psi, M)$ triggers player $P(Q_i)$ choosing an object $d_i \in D$; the game continues as $g(\psi, M)$.
>
> $g(R(\overline{x}), M)$ has no moves; E receives payoff 1 if $\overline{d} \in I(R)$, and -1 otherwise. A gets E's payoff times -1.

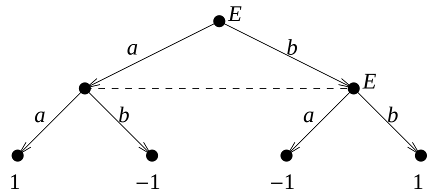

Figure 1: The game tree of $g(\theta, (\{a, b\}, =))$, containing seven histories. The top node corresponds to the empty history; the histories on the intermediate layer are denoted by $\langle a\rangle, \langle b\rangle$; and $\langle a, a\rangle, \langle a, b\rangle, \langle b, a\rangle, \langle b, b\rangle$ are the terminal nodes. The fact that $\langle a\rangle, \langle b\rangle$ sit in the same information set is reflected by the dashed line. The values 1 and -1 are payoffs for E.

The above rules regulate the behavior of the game $g(\phi, M)$. [6] do not provide a rigorous game-theoretic model for these games. However, the formal treatments provided in the literature all take an *extensive* stance towards these games, viz. [12, 9, 3] and [10] for a propositional variant. In this paper the game $g(\phi, M)$ — with a lower-case 'g' — denotes a Hintikka-Sandu style, extensive semantic game. In these games independence is modeled by means of *information sets* imposed on the *histories* of the game tree. We omit rigorous definitions, but illustrate the idea by means of the game tree of an IF sentence that reappears in our discussion below

$$\theta \quad = \quad \exists x_1(\exists x_2/\{x_1\})\,[x_1 = x_2], \tag{4}$$

evaluated on the model $(\{a, b\}, =)$, depicted in Figure 1. From a game-theoretic perspective, every node in a game tree corresponds to a history, and every leaf to a complete history. On every complete history the utility function of the players is defined.

To say that two histories are in the same information set means that the player owning the set at hand cannot distinguish between the two histories while at it. As a consequence any *pure strategy* for this player prescribes only *one* action for all the histories in the information set.

We say that E has a *winning* strategy in $g(\phi, M)$ if there exists a strategy that guarantees an outcome of 1, against every strategy played by A; and a strategy is *uniform* with respect to the game's information sets, if it assigns to every information set in which E is to move exactly one object from the domain. Note that here and henceforth we consequently mean 'pure strategy' when speaking of 'strategy'. Truth under GTS is defined in terms of winning strategies.

Definition 3 *Let $\phi \in$ IF and let M be a suitable model. Then define truth under GTS as follows:*

> *ϕ is true under GTS on M, denoted by $M \models_{\mathrm{GTS}} \phi$, if E has a winning strategy in $g(\phi, M)$.*
>
> *ϕ is false under GTS on M, if A has a winning strategy in $g(\phi, M)$.*

ϕ is undetermined under GTS on M, if neither E nor A has a winning strategy in $g(\phi, M)$.

In the realm of IF semantic evaluation games, information sets only partition histories of equal length, cf. [10]. Pure strategies in IF semantic games therefore coincide with tuples of *Skolem functions*, as we know them from logic. We introduce Skolem functions by illustrative means. Let ϕ be as in (1), then its Skolemization looks like

$$\exists f_2 \ldots \exists f_n \forall x_1 \ldots \forall x_{n-1}\ R(\overline{x}, \overline{f}),$$

where f_i is a Skolem function, being a function of type $D^{\{x_1,\ldots,x_{i-1}\}\setminus X_i} \to D$.

[13] showed that the truth condition of every formula with partially ordered quantifiers can be expressed in the Σ^1_1 fragment of second-order logic. Later, the result, applied to IF, reappears in Sandu's and Hintikka's work (for references see [6]) hinging on the fact that for ϕ as in (1)

$$M \models_{\mathrm{GTS}} \phi \qquad \text{iff} \qquad M \models_{\mathrm{Tarski}} \exists f_2 \ldots \exists f_n \forall x_1 \ldots \forall x_{n-1}\ R(\overline{x}, \overline{f}),$$

since any tuple $f_2, \ldots, f_n$ witnessing the truth of ϕ's Skolemization is a winning strategy for E in $g(\phi, M)$ and the other way around, assuming the Axiom of Choice. For [6] it is the strategies that form the heart of the game-theoretic apparatus involved.

> What is essential [about game-theoretic conceptualizations] is not the idea of competition, winning and losing. …What is essential is the notion of strategy. Game theory was born the moment John von Neumann formulated explicitly this notion.

Having read this, the thought occurs that defining IF evaluation games in a *strategic* way may be more in line with Sandu and Hintikka's thinking. In this paper we will set up such a strategic framework; discuss the 'playability' of IF games in this context; and start exploring the framework.

The issue of playability of IF games, mentioned above, arises when we actually want to play games for IF sentences ϕ. In a game for ϕ, the turn-taking is governed by ϕ's quantifier prefix and the epistemic qualities of the agents by ϕ's slash sets. However, defining the IF language, we took no special care that our formulas would give rise to playable games. In fact, it has been observed that certain IF sentences yield games that require agents with odd epistemic features. That is, games that violate the game-theoretic principles of *perfect memory* and *action memory*. Roughly speaking a game of imperfect information has perfect memory if a player learning something (in our context: a previous move), implies it knowing this piece of information for the rest of the game; and, a game has action memory if every player recalls at least it's own moves. We refer the reader to [11] for an elaborate treatment of perfect recall and IF games.

For the sake of illustration, consider the extensive game $g(\theta, (\{a, b\}, =))$, with θ as in (4).[1] It is the case that $(\{a, b\}, =) \models_{\mathrm{GTS}} \theta$, since the tuple (play a, play a) is a winning

[1]The formula θ also appears in [7], as an argument against Hintikka's claim of IF logic modeling quantifier independency. Janssen argues that, since θ holds on the domain, it must be the case that x_2 depends on x_1. However, in θ the choice for x_2 is independent of x_1, since $X_2 = \{x_1\}$. For more on IF logic and intuitions on independence consult [7].

strategy. But also we have it that the histories $\langle a \rangle$ and $\langle b \rangle$ are in E's information set indicating that these histories are *indistinguishable* for E. Thus, $g(\theta, (\{a, b\}, =))$ lacks both perfect memory and action memory.

The issue of the playability of $g(\theta, (\{a, b\}, =))$ evolves around the question *how E can understand that* (*play a, play a*) *is a winning strategy for E, despite the fact that she is uninformed at the intermediate stage.* That is, E seems to forget her own move right after playing it!

One explanation may be that E is allowed to decide beforehand on a strategy and consult it while playing the game, even if she is unsure about her own moves at the intermediary stage. (This explanation appears in [12].) In particular, that (play a, play a) is a winning strategy can then be understood as follows: First E picks a, thereafter she is uncertain about what history she is actually in: $\langle a \rangle$ or $\langle b \rangle$. By consulting here winning strategy, however, she derives that she actually is in $\langle a \rangle$ and not in $\langle b \rangle$. The imperfect information evaporates!

This explanation requires more game-theoretic structure — i.e., consulting of one's strategy — than present in its description and would imply a non-game-theoretic understanding of having imperfect information *during* the game.

Another explanation may be that E is an *existential team*, hence associating with every existential variable a member of the team. This would make $g(\theta, (\{a, b\}, =))$ a two-player cooperative game. But then the very fact that θ holds on the model at stake suggests to be interpreted in such a way that the x_1-player and the x_2-player are allowed to settle on their strategies *before* the game. Again, no such event can be found in the definition of $g(\theta, (\{a, b\}, =))$ and it seems such an event would violate the game-theoretic understanding of information sets. Because, for instance in $g(\theta, (\{a, b\}, =))$ the second player in the E-team would really know the move of the first player.

Below we shall reduce the puzzle that arises with θ to the question how Nash equilibria are supposed to arise in strategic games. First we set up a strategic framework, in which the notion of Nash equilibrium and other solution concepts can be meaningfully employed.

3 Strategic framework for IF games

In this section we define IF games as strategic games. We characterize truth of IF under GTS in terms of Nash equilibria.

Definition 4 *Let $\phi \in$ IF and let M be a suitable model. Then, define the* strategic evaluation game of ϕ and M *as*

$$G(\phi, M) \quad = \quad \Big(N_\phi, (S_{i,\phi})_{i \in N_\phi}, (u_{i,\phi,M})_{i \in N_\phi}\Big).$$

N_ϕ denotes the set of players, $S_{i,\phi}$ the set of strategies for player i, and $u_{i,\phi,M}$ is player i's utility function. We also call $G(\phi, M)$ an IF game.

Below we briefly introduce these ingredients componentwise and introduce some notation involved. Note that strings in IF are assumed to be as in (3). All definitions below are restricted to this assumption, but can be generalized without much ado.

Players. The set $N_\phi = \{i \mid x_i \in \mathit{Var}(\phi)\}$ contains the players. The set N_ϕ conveys the strong connection between variables in ϕ and players in $G(\phi, M)$. In fact, if $V \subseteq \mathit{Var}(\phi)$, then we will use $N(V) = \{i \mid x_i \in V\}$ to denote the set of players associated with the variables in V. Let E_ϕ (A_ϕ) be the set of existentially (universally) quantified variables in ϕ. We have adopted the *multi-player* view on IF games here, mainly because it is the framework that is most open to generalizations with respect to, for instance, the utility functions. Moreover, it allows for smoother terminology.

Strategies. For $x_i \in \mathit{Var}(\phi)$, define $U_{i,\phi} \subseteq \mathit{Var}(\phi)$ to be the set of variables on which x_i depends in ϕ. That is, $U_{i,\phi} = \{x_j \mid (x_j, x_i) \in B_\phi\}$. In the context of the game and player i having control over x_i, we often say that *i sees* $U_{i,\phi}$. $S_{i,\phi}$ denotes the set of all player i's strategies in $G(\phi, M)$, being (Skolem) functions of type $s_i : D^{U_{i,\phi}} \rightarrow D$. If $U_{i,\phi}$ is empty, $S_{i,\phi}$ only contains *atomic strategy*.

Manipulating strategies. Define a *profile* s in $G(\phi, M)$ as an object in

$$\bigtimes_{i \in N'} S_{i,\phi},$$

for some $N' \subseteq N_\phi$. We call s *existential* (*universal*), if $N' \subseteq E_\phi$ (A_ϕ); otherwise we call it *mixed*. We call s *complete*, if $N' = N_\phi$; otherwise we call it *partial*. If $N' = N(E_\phi)$ ($N(A_\phi)$), we call the profile *completely existential (universal)*. If no confusion arises we will drop as many of the terms as possible.

If $s \in \bigtimes_{i \in N'} S_{i,\phi}$ for some $N' \subseteq N_\phi$ and $\{1, \ldots, j\} \in N'$, then $s_{1,\ldots,j}$ denotes the strategy profile s containing only player 1 to j's strategies. We will often discuss player j changing strategies with respect to a strategy profile s. We write (s_{-j}, t_j) to denote the profile that is the result of replacing s_j by t_j. If $s \in \bigtimes_{i \in N'} S_{i,\phi}$ and $s' \in \bigtimes_{i \in N''} S_{i,\phi}$ for disjoint $N', N'' \subseteq N_\phi$, then ss' is the result of concatenating s and s'. If s_i is a strategy of type $D^{\{y_1,\ldots,y_k\}} \rightarrow D$ and assignment p is defined over $\{y_1, \ldots, y_k\}$, then we will write $s_i(p)$ instead of $s_i(p(y_1), \ldots, p(y_k))$.

Finally, every profile $s \in \bigtimes_{i \in \{1,\ldots,j\}} S_{i,\phi}$ in $G(\phi, M)$ gives rise to an assignment $[s]$ that is defined over $\{x_1, \ldots, x_j\}$ as below. Note that s_1 is an atomic strategy.

$$\begin{array}{rcl} [\mathrm{s}](x_1) & = & s_1 \\ [\mathrm{s}](x_i) & = & s_i([s_{1,\ldots,i-1}]). \end{array}$$

Utility functions. Let $i \in N_\phi$. Then, i's *utility function* in $G(\phi, (D, I, p))$ is defined over complete profiles s as follows:

$$u_{i,\phi,(D,I,p)}(s) = \begin{cases} c_i & \text{if } [s] \in I(R) \\ -c_i & \text{if } [s] \notin I(R), \end{cases}$$

where $c_i = 1$ if $i \in N(E_\phi)$, and $c_i = -1$ if $i \in N(A_\phi)$. As all utility functions of the players in $N(E_\phi)$ and $N(A_\phi)$, respectively, are equivalent and the models under consideration can be made up from the context we will simply denote them by u_E and u_A.

	play a	play b
play a	1	-1
play b	-1	1

Table 1: Every cell in the matrix corresponds to an assignment $[s]$ over $Var(\theta)$. We filled in the value $u_E([s])$ reflecting payoff for the members of the existential team.

Now that we switched from extensive to strategic semantic games, observe that the notion of winning strategy in extensive games has a respectable strategic counterpart: *Nash equilibrium*. We say that the strategy profile $\hat{s}$ is a Nash equilibrium in the strategic game G, if none of the players i gains from unilateral deviation (see also [8]):

$$u_i((\hat{s}_{-i}, s_i)) \leq u_i(\hat{s}),$$

where s_i is any other strategy for player i and u_i is player i's utility function in G. The following lemma can also be understood as a proof of effective equivalence between $g(\phi, M)$ and $G(\phi, M)$.

Lemma 5 *Let $\phi \in$ IF and let M be a suitable model. Then, the following are equivalent:*

- $M \models_{\text{GTS}} \phi$.
- *There exists a Nash equilibrium s in $G(\phi, M)$, such that $u_E(s) = 1$.*

Technically this lemma is not deep. Yet it shows us that strategic games can account for truth of IF logic. In the strategic framework the playability issues concerning perfect recall, encountered in extensive IF games, evaporate simply because the strategic games ignore the inner structure of games defined by consecutive moves by the agents. By ignoring the inner structure of the game, also the epistemic states of the agents — i.e., their information sets — are ignored.

But the issue of playability pops up in the strategic framework under a different guise. Revisit the game $G(\theta, (\{a, b\}, =))$. As is common usage in strategic games we draw its payoff matrix, see Table 1. The puzzle induced by the truth of θ on $(\{a, b\}, =)$ in extensive contexts appears in the strategic context as a coordination problem. There are two equally profitable Nash equilibria, but which one to choose, without possibility to coordinate? How to understand Nash equilibria is a problem central in game theory, see [8].

In the upcoming two sections we explore semantic interpretations for IF logic that are motivated by solution concepts that are not subject to coordination problems.

4 Weak dominance semantics

In this section, we define a semantics based on the existence of *weakly dominant* strategies. Intuitively, a strategy is weakly dominant for a player if it outperforms any other strategy independently of the other players' strategic behavior.

Definition 6 *Fix some IF game $G(\phi, M)$. Then, $\hat{s}_i$ is a* weakly dominant *strategy in $G(\phi, M)$ for player i, if $\hat{s}_i \in S_{i,\phi}$ and for every complete mixed profile s it is the case that*

$$u_E((s_{-i}, \hat{s}_i)) \geq u_E(s).$$

We call $\hat{s}_i$ weakly dominant, *because possibly it is exactly as good as player i's original strategy in s. Dually, we define strategy $t_i \in S_{i,\phi}$ to be* strictly dominated *by $\hat{s}_i$ in $G(\phi, M)$, if for every complete mixed profile s it is the case that*

$$u_E((s_{-i}, \hat{s}_i)) \geq u_E((s_{-i}, t_i)) \qquad \textit{and} \qquad u_E((r_{-i}, \hat{s}_i)) > u_E((r_{-i}, t_i))$$

for at least one complete mixed profile r.

The notion of weak dominance we employ is weaker than the one usually adopted in game theory. For comparison we refer to [8]. We now come to our definition of truth in terms of weak dominance.

Definition 7 *Let $\phi \in$ IF and let M be a suitable model. Then we define truth of ϕ on M under* weak dominance semantics *(WDS) as follows*

> *$M \models_{\text{WDS}} \phi$ iff in $G(\phi, M)$ there exists a complete existential profile $\hat{s}$ such that $\hat{s}_i$ is a weakly dominant strategy for every $i \in N(E_\phi)$, and $u_E(\hat{s}t) = 1$, for any complete universal profile t.*

Falsity and undeterminedness under WDS are defined similarly.

The question remains, of course, what remains of IF logic evaluated under WDS. It becomes clear that GTS is less restrictive a semantics for IF logic than WDS, after reformulating truth under GTS in multi-player terms, since we may simply omit $\hat{s}_i$'s constraint of being weakly dominant:

> $M \models_{\text{GTS}} \phi$ iff in $G(\phi, M)$ there exists a complete existential profile s such that $u_E(st) = 1$, for any complete universal profile t.

Formally, our claim boils down to the claim that

$$M \models_{\text{WDS}} \phi \quad \text{implies} \quad M \models_{\text{GTS}} \phi, \tag{5}$$

but not the other way around. Since it is the case that $(\{a, b\}, =) \models_{\text{GTS}} \theta$, but θ does not hold on this domain under WDS, see Table 1. As an example of WDS observe, that, surprisingly, for any model M with more than one object in its domain it is the case that for $\tau = \exists x_1 \exists x_2\ [x_1 = x_2]$:

$$M \not\models_{\text{WDS}} \tau \quad \text{whereas} \quad M \models_{\text{Tarski}} \tau.$$

That τ is true under Tarski semantics is obvious. From Table 2 it becomes clear that τ is not true under WDS on the model with two objects $\{a, b\}$. Although player 2 has a weakly dominant strategy, player 1 has none.

	s_2^a	s_2^b	s_2^{copy}	s_2^{invert}
s_1^a	1	−1	1	−1
s_1^b	−1	1	1	−1

Table 2: s_i^d is the atomic strategy for player $i \in \{1, 2\}$ assigning object $d \in \{a, b\}$. s_2^{copy} is player 2's strategy such that $s_2^{\text{copy}}(d) = d$, whereas s_2^{invert} switches the object chosen by player 1.

In the remainder of this section we will characterize the truth-conditions of IF under WDS and see that very little is left of IF's Σ_1^1-expressiveness it enjoyed under GTS. We show in Theorem 10 that truth under WDS can be expressed in a fragment of FOL (evaluated under Tarski semantics). Before we come to a rigorous formulation, let us classify an IF sentence ϕ's variables and characterize one of the resulting classes.

Recall that we defined the dependency relation of ϕ's variables as a binary relation B_ϕ. The result of taking the transitive closure of B_ϕ we denote B_ϕ^*. That is, $(x_i, x_j) \in B_\phi^*$ iff there exists a chain $z_0, \ldots, z_m$ of variables in $Var(\phi)$ such that $z_0 = x_i$, $z_m = x_j$, and for every $t \in \{0, \ldots, m-1\}$ it is the case that $(z_t, z_{t+1}) \in B_\phi$. Such a chain of variables $z_0, \ldots, z_m$ we will call a *B_ϕ-chain*. Note that B_ϕ^* is irreflexive.

For every variable $x_i \in Var(\phi)$, partition $Var(\phi)\backslash\{x_i\}$ as follows:

$$U_{i,\phi} = \{x_j \mid (x_j, x_i) \in B_\phi\} \tag{6}$$

$$W_{i,\phi} = \{x_j \mid (x_i, x_j) \in B_\phi^*\} \tag{7}$$

$$V_{i,\phi} = Var(\phi)\backslash(U_{i,\phi} \cup \{x_i\} \cup W_{i,\phi}). \tag{8}$$

We encountered $U_{i,\phi}$ before, as it contains all variables seen by player i. $W_{i,\phi}$ contains the variables that can *(in)directly* see x_i. $V_{i,\phi}$ is the set of all other variables in ϕ not containing x_i. What is meant by 'seeing (in)directly' is pinpointed by the following lemma, that characterizes the variables in $W_{i,\phi}$.

Lemma 8 *Let $\phi \in$ IF be as in* (3) *and let M be a suitable model. Let $W_{i,\phi}$ be defined as in* (7) *for some sentence ϕ and $i \in N_\phi$. Then, $x_j \in W_{i,\phi}$ iff $i \neq j$ and in $G(\phi, M)$ there exists a complete strategy profile s and a strategy $t_i \in S_i$ such that $[s](x_j) \neq [(s_{-i}, t_i)](x_j)$.*

Intuitively, $W_{i,\phi}$ is the subset of $Var(\phi)$ consisting of variables that are sensitive to x_i changing assignments. The lemma, interpreted the other way around, teaches that, if x_j is not in $W_{i,\phi}$, for every strategy profile, player i changing strategies does affect the object assigned to x_j.

Theorem 9 *Let $\phi \in$ IF as in* (3) *and let M be a suitable model. The sets $U_{i,\phi}$, $W_{i,\phi}$, $V_{i,\phi}$ are defined as in* (6), (7) *and* (8)*, respectively. We also consider the set $W'_{i,\phi} = \{x' \mid x \in W_{i,\phi}\}$. The strings of variables in these respective sets will be referred to by means of $\overline{u}, \overline{v}, \overline{w}$, and $\overline{w}'$. Then, in $G(\phi, M)$ player $i \in N_\phi$ has a weakly dominant strategy iff*

$$M \models_{\text{Tarski}} \forall\overline{u}\exists x_i'\forall x_i\forall\overline{v}\forall\overline{w}\forall\overline{w}' \quad [(i) \wedge (ii) \wedge (iii) \rightarrow (iv)], \tag{9}$$

where

$$
\begin{array}{rcl}
(i) & = & R(\overline{u}, x_i, \overline{v}, \overline{w}) \\
(ii) & = & x_i \neq x'_i \\
(iii) & = & \bigwedge\limits_{j \neq i \in N} \left(\left(\bigwedge\limits_{x_k \in U_{j,\phi}} x_k = x^*_k \right) \rightarrow x_j = x^*_j \right) \\
(iv) & = & R(\overline{u}, x'_i, \overline{v}, \overline{w}').
\end{array}
$$

If $U_{j,\phi}$ is empty, interpret $(\bigwedge_{x_k \in U_{j,\phi}} x_k = x^*_k)$ as $\top$. Note that $Var((i)) = Var(\phi)$ and that $Var((iv)) = U_{i,\phi} \cup \{x'_i\} \cup V_{i,\phi} \cup W'_{i,\phi}$. Furthermore, $\cdot^*$ is a mapping from $Var((i))$ to $Var((iv))$, as follows

$$
y^* = \begin{cases} y & \text{if } y \in U_{i,\phi} \cup V_{i,\phi} \\ y' & \text{if } y \in \{x_i\} \cup W_{i,\phi}. \end{cases}
$$

We will refer to the first-order formula in (9) *as $\alpha_i(\phi)$.*

Basically, $\alpha_i(\phi)$ states that *if* (i) there exists an assignment that satisfies R, (ii) player i changes the object assigned to x_i, but (iii) the other players j play according to a Skolem function that is uniform with respect to what they can see (i.e., the objects assigned to the variables in $U_{j,\phi}$), *then* (iv) there exists an object d_i to assign to x'_i that guarantees truth of R no matter what is played by the players that can (in)directly see to x_i. The strategy such that $\hat{s}_i(\overline{u}) = d_i$ for all $\overline{u}$ that satisfy (i) is a weakly dominant strategy. It is a weakly dominant strategy in $G(\phi, M)$ for player i, because $\hat{s}_i \in S_{i,\phi}$.

Theorem 9 characterizes the condition under which a player has a weakly dominant strategy. To be true under WDS, however, slightly more is required. The following theorem characterizes truth under WDS.

Theorem 10 *Let $\phi \in$* IF *be as in* (3) *and let M be a suitable model. Let E_ϕ and A_ϕ partition $Var(\phi)$ in such a way that E_ϕ contains the existentially quantified variables in ϕ. We abbreviate the string of all variables in E_ϕ and A_ϕ using $\overline{e}$ and $\overline{a}$. Then,*

$$
M \models_{\text{WDS}} \phi \quad \textit{iff} \quad M \models_{\text{Tarski}} \alpha(\phi) \wedge \beta(\phi),
$$

where $\alpha(\phi) = \bigwedge_{i \in N(E_\phi)} \alpha_i(\phi)$ and $\beta(\phi) = \forall\overline{a}\exists\overline{e}\, R(\overline{a}, \overline{e})$.

Formula $\alpha(\phi)$ being true on M is equivalent to every existential player i having a weakly dominant strategy $\hat{s}_i$ in $G(\phi, M)$. Yet this does not guarantee that the existential players i playing according to $\hat{s}_i$ will always get 1. For instance, in $G(\psi, M)$ every existential player has a weakly dominant strategy, if ψ's relational symbol is false for every suitable tuple of objects from M's domain. However, playing according to it will always yield an outcome of -1. Truth of $\beta(\phi)$ is a sufficient and necessary condition for avoiding the latter situations.

For future comparison we conclude this section with a meta-statement about IF logic interpreted under WDS, that follows straightforwardly from Theorem 10.

Theorem 11 IF *under WDS has less than elementary expressive power.*

	s_2^{copy}
s_1^a	1
s_1^b	1

Table 3: The payoff matrix of $G'(\tau, (\{a, b\}, =)) = (\{1, 2\}, (\{s_1^a, s_1^b\}, \{s_2^{\text{copy}}\}), \{u_E, u_A\})$.

5 Beyond WDS

From Theorem 9 we learn that for a player to have a weakly dominant strategy it does not matter what is played by his team members. Even in the case all its team members leave him and join the other team, this would not make a difference with respect to him having a weakly dominant strategy. I.e., WDS ignores the opportunities that might come with the notion of a *team*. In this section we show by example that increasing the 'powers' of the involved players in IF games increases the expressive power of IF logic on the obtained semantics, Theorem 14 as opposed to Theorem 11.

Let us revisit the sentence $\tau = \exists x_1 \exists x_2\ [x_1 = x_2]$. We observed that τ is not true under WDS on any model M that has a domain with more than one object (see Table 2). On the assumption that player 1 knows 2 is rational, player 1 may infer that 2 plays s_2^{copy}, because playing this strategy is better for it than any other strategy. That is, s_2^{copy} is weakly dominant. After this inference, player 1 choosing a strategy in $G(\tau, M)$ then effectively boils down to it choosing a strategy in the game

$$G'(\tau, (\{a, b\}, =)) \quad = \quad \left(\{1, 2\}, (\{s_1^a, s_1^b\}, \{s_2^{\text{copy}}\}), \{u_E, u_A\}\right).$$

G''s trivial payoff matrix is depicted in Table 3.

In this spirit, the following definition hard-wires the procedure of players calculating what other players will play. As such it bears strong similarity to the game-theoretic literature on *iterated removal of dominated strategies*, see [8].[2] The result of this procedure $\mathcal{P}$ as applied to some IF game will be the game that is effectively played.

Definition 12 *Let* $\phi \in$ IF *as in* (3) *and let* M *be a suitable model. Then, define*

$$\begin{aligned} G^n(\phi, M) &= G(\phi, M) = (N, (S_1, \ldots, S_n), \{u_E, u_A\}) \\ G^{i-1}(\phi, M) &= \left(N, (S_1, \ldots, S_{i-1}, S_i^{\mathcal{P}}, S_{i+1}^{\mathcal{P}}, \ldots, S_n^{\mathcal{P}}), \{u_E, u_A\}\right) \end{aligned}$$

where $S_1, \ldots, S_{i-1}, S_{i+1}^{\mathcal{P}} \ldots, S_n^{\mathcal{P}}$ *are copied from* $G^i(\phi, M)$ *and*

$$S_i^{\mathcal{P}} \quad = \quad \{s_i \in S_i \mid s_i \text{ weakly dominant in } G^i(\phi, M)\}.$$

Finally, put the strategic evaluation game $G^{\mathcal{P}}(\phi, M) = G^0(\phi, M)$.

[2] It is tempting to clarify the inferences of the players by assuming *common knowledge of rationality*. (In fact a weaker concept of knowledge would do to trigger the procedure.) In this paper we consider the procedures simply as formal objects, leaving us space to define procedures that are not epistemologically justified (such as $\mathcal{ND}$, defined below). For much more on epistemological characterizations of game-theoretic solution concepts we refer to [2].

This vehicle we employ to define a semantics 'on top' of WDS.

Definition 13 *Let $\phi \in$ IF and let M be a suitable model. Then we define truth of ϕ on M under* weak dominance semantics plus $\mathcal{P}$ *as follows:*

$M \models^{\mathcal{P}}_{\text{WDS}} \phi$ *iff in $G^{\mathcal{P}}(\phi, M)$ for every complete profile $\hat{s}$ it is the case that $u_E(\hat{s}) =$ 1.*

We thus state the truth of an IF sentence ϕ on M in terms of the outcome of playing the game $G^{\mathcal{P}}(\phi, M)$ by players that are empowered to reason according to the procedure $\mathcal{P}$. For instance, it is the case that $(\{a, b\}, =) \models^{\mathcal{P}}_{\text{WDS}} \tau$.

First of all, note that, epistemically, player n needs to know nothing about the other players in order to pick a weakly dominated strategy, i.e., to act in accordance with $\mathcal{P}$. Now, player $n - 1$ needs to know that player n is indeed rational in order for it to be rational to consider game $G^n(\phi, M)$. In general, to explain why the players would execute $\mathcal{P}$, one has to assume that every player i is rational and i knows that $i+1$ knows that ... knows that n is rational. Now, this is quite strong an assumption to make. Much stronger in any case than WDS' mere requirements that all the players are rational.

Secondly, we observe that for $\phi \in$ IF

$$M \models_{\text{WDS}} \phi \text{ implies } M \models^{\mathcal{P}}_{\text{WDS}} \phi \quad \text{and} \quad M \models^{\mathcal{P}}_{\text{WDS}} \phi \text{ implies } M \models_{\text{GTS}} \phi, \qquad (10)$$

but the converses do not hold, witnessing τ and θ on $M = (\{a, b\}, =)$, respectively.

Thirdly, in Theorem 14 we observe that the expressive power increases when switching from $\models_{\text{WDS}}$ to $\models^{\mathcal{P}}_{\text{WDS}}$ with respect to FOL. Also, we draw the conclusion from this theorem that every FOL formula behaves under WDS plus $\mathcal{P}$ as it does under Tarski semantics. What is the expressive power of IF logic under WDS plus $\mathcal{P}$ is left open.

Theorem 14 *Let $\phi \in$ FOL and let M be a suitable model. Then,*

$$M \models_{\text{Tarski}} \phi \quad \textit{iff} \quad M \models^{\mathcal{P}}_{\text{WDS}} \phi.$$

The procedure $\mathcal{P}$ turns out to be the strategic counterpart of the *backwards induction algorithm* as applied to the extensive game tree of an FOL game. The proof of Theorem 14 boils down to showing that a tuple of Skolem functions $\overline{f}$ is a witness of $M \models_{\text{Tarski}} \phi$ iff it is contained in $S^{\mathcal{P}}_1 \times \ldots \times S^{\mathcal{P}}_n$.

6 Conclusion

In this paper, we set up a strategic framework for IF semantic games, which are traditionally studied extensively. Naturally, by giving up the extensive structure that is traditionally given to IF games, we avoid conceptual issues that arise with the playability of IF games (i.e., lack of perfect recall). We observed that truth of IF logic under GTS can be characterized by the solution concept of Nash equilibrium. We saw that other issues arise in the strategic framework: how are players supposed to coordinate or, more eloquently, how are Nash equilibria supposed to arise?

We used the strategic framework to define to semantic interpretations for IF logic inspired by solution concepts related to weakly dominant strategies: $\models_{\text{WDS}}$ and $\models^{\mathcal{P}}_{\text{WDS}}$. The former does not require any of the involved players to know anything about the other players. We showed that under $\models_{\text{WDS}}$ the expressive power of IF logic collapses to that of a fragment of first-order logic (under Tarski semantics). The epistemic demands of $\models^{\mathcal{P}}_{\text{WDS}}$ were seen to be higher than that of $\models_{\text{WDS}}$. We showed that the expressive power of FOL (under Tarski semantics) is left intact when evaluated under $\models^{\mathcal{P}}_{\text{WDS}}$. Thus, all of IF logic (under $\models^{\mathcal{P}}_{\text{WDS}}$) has expressive power of at least FOL (under Tarski semantics). Our findings can be summarized in the following table:

Solution concept $\mathcal{S}$	Expressive power $\models_{\mathcal{S}}$
Nash equilibrium	High ($= \Sigma^1_1$)
WDS + $\mathcal{P}$	Medium-high ($\geq$ FOL)
WDS	Low ($<$ FOL)

Further research will have to flesh out this table and determine what are the dependencies between solution concepts and the expressive power of IF logic evaluated under the associated solution concept. This enterprise would explore correlations between notions of agency and semantic interpretations of logical languages.

References

[1] P. J. Cameron and W. Hodges. Some combinatorics of imperfect information. *Journal of Symbolic Logic*, 66(2):673–684, 2001.

[2] Boudewijn de Bruin. *Explaining Games, On the Logic of Game Theoretic Explanations*. PhD thesis, ILLC, Universiteit van Amsterdam, 2004.

[3] Francien Dechesne. *Game, Set, Maths: Formal investigations into logic with imperfect information*. PhD thesis, SOBU, Tilburg university and Technische Universiteit Eindhoven, 2005.

[4] Leon Henkin. Some remarks on infinitely long formulas. *Infinitistic Methods*, pages 167–183, 1961.

[5] Jaakko Hintikka. *Principles of mathematics revisited*. Cambridge University Press, Cambridge, 1996.

[6] Jaakko Hintikka and Gabriel Sandu. Game-theoretical semantics. In Johan F. A. K. van Benthem and Alice ter Meulen, editors, *Handbook of logic and language*, pages 361–481. North Holland, Amsterdam, 1997.

[7] T. M. V. Janssen. Independent choices and the interpretation of IF logic. *Journal of Logic, Language and Information*, 11:367–387, 2002.

[8] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[9] A. Pietarinen and Tero Tulenheimo. An introduction to IF logic. Lecture notes for the 16th European Summerschool in Logic, Language and Information, 2004.

[10] Gabriel Sandu and A. Pietarinen. Informationally independent connectives. In G. Mints and R. Muskens, editors, *Logic, Language and Computation*, volume 9, pages 23–41. CSLI Publications, Stanford, 2003.

[11] Merlijn Sevenster. *Branches of imperfect information: logic, games, and computation*. PhD thesis, ILLC, University of Amsterdam, 2006.

[12] J.F.A.K. van Benthem. Logic and games, lecture notes. Draft version, 2000.

[13] W. Walkoe. Finite partially-ordered quantification. *Journal of Symbolic Logic*, 35:535–555, 1970.