INDEPENDENCE FRIENDLY LOGIC AS A STRATEGIC GAME

THEO M.V. JANSSEN

ILLC
University of Amsterdam
theo@science.uva.nl

Abstract *In this contribution investigate an alternative interpretation for Hintikka's Independence Friendly logic. IF logic is not seen as an extensive game, but as a strategic game. We base our semantics on one assumption: the players are rational: they do not play a strategy if there is a better one. In this semantics signalling is not possible. The semantics gives more adequate results for certain linguistic applications.*

1. Introduction

Hintikka has raised attention for the role of independence in the semantics of natural languages (see Hintikka and Sandu 1997, Hintikka 1996). His most well known examples concern branching quantifier sentences like

(1) *Some friend of each townsman and some neighbour of each villager hate each other*

In the intended meaning the friend of the townsman can be chosen independently of the neighbour of the villager. This meaning is often discussed, but we will accept that those sentences have such a meaning. But I think that the intended meaning is not captured by the traditional semantics for Hintikka's Independence Friendly logic (henceforth IF logic).

Hintikka argues that independence is a widespread phenomenon and suggests many other applications: in linguistics e.g. the *de dicto - de re* distinction as in *John seeks a unicorn*, but also applications in other fields, e.g. in quantum mechanics.

The the meaning of (1), using some suggestive abbreviations, is represented by:

(2) $\forall x_1 \exists x_2 \forall x_3 \exists x_{4/x_1,x_2} [T(x_1) \wedge V(x_3) \rightarrow F(x_2, x_1) \wedge N(x_4, x_3) \wedge H(x_2, x_4)]$

The subscript $_{/x_1,x_2}$ indicates that the choice of x_4 has to be made independent of x_1 and x_2; it may depend, however, on x_3. In this paper we will see IF logic just as first order logic extended with quantifiers like $\exists x_{4/x_1,x_2}$ and $\forall x_{3/x_2}$, in the literature one finds extensions.

The interpretation of IF logic originally is defined by means of a game between two players, the one aiming at confirming the formula, the other at refuting it. Independence is captured by the number of arguments that the strategies have. Some authors prefer to interpret it as a game between two teams. In later publications the game has been analyzed as an extensive game. Then independence is captured by indiscernible information sets. It has been argued that the rules for playing are not clear and that one has to make rather unnatural assumptions about the properties of the players (Sevenster 2005, van Benthem 2004). For instance, they have to forget decisions made before (even their own), and then to remember them again. Furthermore the obtained semantics has been criticized (e.g. Janssen 2002).

In this paper we will describe the game as an extensive game. The inspiration came from a paper by Sevenster 2005. He considers a spectrum of possible strategies, en several criteria for making a select from that spectrum. A player may choose only strategies of a certain type, other players may know this information and use it to eliminate strategies from their spectrum, etc. In this contribution we investigate properties of one natural assumption from this spectrum: rationality (weakly dominance in his paper). The resulting semantics coincides for many examples with Hintikka's semantics, but is not equivalent. It will be shown that our semantics is more useful for one of the applications in linguistics.

2. New semantics

For the ease of discussion, we assume that the formula consists of a quantifier free part, preceded by a list of quantifiers which may be slashed for some variables. We assume that each quantifier has its own indexed variable, and that they appear in order, first the quantifier for x_1, then for x_2, etc. These restrictions are not essential for the approach, it can be extended to arbitrary IF-formulas.

With a variable x_i is associated a player p_i who determines the value for that variable. Players who determine the values the existential variables aim at truth value **1**, and players for the universal quantifiers aim at truth value **0**. The strategy of the player for $\exists x_3$ may have as arguments all variables with wider scope; for $\exists x_{4/x_2,x_3}$ the variables x_2, x_3 are not allowed as argument. A side remark for experts: we do *not* have implicit slashing.

We base our semantics on one assumption: the players use only *rational* strategies. It probably is the easiest to understand if we start with the opposite. A strategy for player i is *irrational* if there is a strategy for her which is better against at least one combination of strategies of the other players and gives the same result in other cases. A strategy is *rational* if it is not irrational. The case that a player has no rational strategies will be discussed separately. Note that a player may have several rational strategies. The notion of rationality is well known in game theory, Apt 2004 gives an overview with many results.

At the start of the game each player determines which her rational strategies are. We define a sentence to be 'true', if any combination of rational strategies guarantees the outcome **1**.

Some examples are the given below; they are played on the domain **N**.

Consider $\exists x_1 \exists x_2 [x_1 = x_2]$. Player p_1 has two strategies: choose $x_1 := 0$ and $x_1 := 1$. It depends on the choice by p_2 whether the sentence is evaluated **1** or not. None of x_1's two strategies is better, so both strategies are rational. For player 2 there are 4 strategies: $x_2 := 0, x_2 := 1, x_2 := x_1, x_2 := (\textit{if } x_1 = 1 \textit{ then } 0 \textit{ else } 1)$. Of these is $x_2 := x_1$ is the only which against all other strategies yields **1**, so it is (the only) rational strategy. Since it guarantees that the result will be **1**, the sentence is 'true'.

In $\exists x_1 \exists x_{2/x_1} [x_1 = x_2]$. Both players p_1 and p_2 have two strategies: choose **0** and **1**. For none of the players there is a strategy that improves all others, both play arbitrarily. For some combinations of strategies the sentence yields **1**, for others it yields **0**. So the sentence is not true.

This last example illustrates that our semantics is not equivalent with Hintikka's semantics. In his approach there is only one player $p_\exists$ who plays $x_1 := 0$ for $x_2 := 0$. The strategy $x_2 := 0$ is a constant strategy, i.e. does not have x_1 as argument, and therefore is allowed (in his semantics). One might say that the player has to forget her choice for $\exists x_1$, but remembers her strategy. In approaches with more players, one has to allow that they communicate on their strategies although their choices must be independent.

In my opinion it is counterintuitive that the player(s) have to make independent choices, but never the less have a method to guarantee two identical choices. In Janssen 2002 it is argued that there are many other examples where Hintikka's results are against intuitions on independence.

A problem for our approach is the situation that a player has no rational strategy. Consider $\exists x_1 \exists x_2 [x_2 < x_1]$. The only rational choice for x_2 is $x_2 := 0$ (because it is better than e.g. $x_2 := 1$). But x_1 has no rational strategies: $x_1 := 1$ is better than $x_1 := 0$, but $x_1 := 2$ is even better $x_1 := 1$ etc. For some choices (viz. $x_1 := 0$) the sentence evaluates to **0**. So the definition of truth has to be adapted somehow for this case. We will not pursue this issue because we are here interested in presenting properties of rational strategies.

3. Properties of the semantics

Inn this section an important property of the logic will proven: the semantics can be defined bottom up, i.e. compositionally.

Definition 3..1 *Two strategies f and g are called equivalent if f in any combination of strategies for the other players yields the same value as g.*

Theorem 3..2 *For any rational strategy of a player for φ there is an equivalent strategy that depends only on variables that occur in φ.*

Proof. Without loss of generality, we may restrict our considerations to the situation that the player is $n + 1$, and that x_1 is the variable that does not occur in φ. We will show that a strategy $f_{n+1}(x_1, \ldots, x_n)$ can be replaced by an equivalent strategy $f^*_{n+1}(x_2, \ldots .x_n)$.

As first step we consider the situation that the values of $x_2, \ldots, x_n$ are fixed, say $a_2, \ldots a_n$ respectively. We call such situations A-situations; they differ only with respect of the value of x_1. Let b a value for x_1 such that the interpretation of φ is as maximal as is possible in A-situations (i.e.for b φ is **1** if that is possible at all in A-situations), and otherwise b is arbitrary.

Define now

$$f'_{n+1} = \begin{cases} f_{n+1}(b, x_2, \ldots x_n) & \text{in A-situations,} \\ f_{n+1}(x_1, x_2, \ldots x_n) & \text{otherwise} \end{cases}$$

So f'_{n+1} gives the same value for x_{n+1} in all A-situations. Since x_1 does not occur in φ, application of strategy f'_{n+1} yields the same value of φ in all A-situations as f_{n+1} does, no matter what the value of x_1 is. So after application of f'_{n+1} the value of φ is at least as good as the value obtained by following f_{n+1} (because of the choice of b). So f'_{n+1} is at least as good as f_{n+1}. Since f_{n+1} was rational, f'_{n+1} cannot be better, hence f'_{n+1} is equivalent with f_{n+1}.

Next we apply the same procedure to all other situations (values for $x_2, \ldots x_n$), thus obtaining a strategy f^* that is equivalent with f, but does not have x_1 as argument.

Summarizing. Define first $g(y_2, \ldots, y_n)$ by:

$$g(a_2, \ldots, a_n) = \begin{cases} b, \text{ where } b \in A \text{ such that} \\ \qquad [\![\varphi]\!]^{\{x_1:b, x_2:a_2, \ldots x_n:a_n, x_{n+1}:f(b, a_2 \ldots, a_n)\}} = 1, \\ \qquad \text{if such an element exists,} \\ \text{arbitrary otherwise} \end{cases}$$

Define $f^* = f(g(x_2, \ldots, x_n), x_2, \ldots, x_n)$.

Then f^* is equivalent with f, but does not have as argument the variable x_1 that does not occur in φ .

End of proof.

A consequence of this theorem is that empty quantifiers can make no difference. The sentence $\forall x \exists x_{/y}[y = x]$ gets the same truth value as $\forall x \exists z \exists x_{/y} y = x$ (both are *not* true). So our semantics is not equivalent with the semantics of Hodges 1997 where the second sentence gets the value 'true'. More generally: signalling is not possible in our approach.

Another example where information from context is used is the following. First note that $\forall x[\exists y_{/x}[x \neq y]$ is not true (in all approaches): it is impossible to make a choice independent of x for y such that the two are unequal. But consider:

(3) $\forall x[\exists y_{/x}[x \neq y] \vee \exists y_{/x}[x \neq y]]$

This sentence is true in the semantics of Hintikka and Hodges. In the left disjunct the player chooses $y := 0$ and in the right $y := 1$, and for the disjunction she chooses

L if $x = 1$. So surprisingly $\varphi \vee \varphi$ is not always equivalent with φ. In our approach the formula is not true. In the left disjunct there are several rational strategies, e.g $y := 0$ and $y := 1$, and the same in the right disjunct. Whatever strategy is followed for the $\vee$, no guarantee is obtained that all combinations of rational strategies yields **1**.

A second consequence of the theorem is that an inductive definition of satisfaction seems possible. We give here a inductive definition of the case that the slashed quantifier is the last one.

Definition 3..3 *Let $\varphi(\bar{x})$ be a formula with possibly $\bar{x}$ as free variables. Then by $\mathcal{A} \models_G^+ \varphi(\bar{x})[v]$ is understood that for any combination of rational strategies φ get value* **1** *if for $\bar{x}$ we take $v(\bar{x})$ as value.*

Theorem 3..4 $\mathcal{A} \models_G^+ \exists y_{/\bar{x}}\, \varphi(\bar{x}, y)[v] \Leftrightarrow$ *there is a function $f\colon (Fr(\varphi) \setminus \{y\}) \to A$ such that $(A) \models \varphi(\bar{x}, f(\bar{x})) \wedge \forall z[\exists u[\varphi(\bar{x}, u)] \to \varphi(z, f(z))][v]$*

Sketch of proof. ($\Leftarrow$) The first conjunct says that f yields **1** for the choices made earlier for $\bar{x}$ (given by v). The second conjunct says that if other values for x would have been chosen (captured by $\forall z$) and if there was then a choice that would make the formula **1**, then also strategy f would yield 1. This means that that f cannot be improved by any other strategy, i.e. it is a rational strategy.
($\Rightarrow$) The first conjunct is obvious, the second conjunct says that for other values f cannot be improved, which is a property of rational strategies
End of proof.

As a matter of fact, the truth definition mentioned in the last theorem resembles the one put forward by Janssen 2002, one which is based upon investigations of many examples of independence and in an attempt to formalize the notion 'independent choice' upon intuitions on independence'. The present work can be seen as an argument for such a semantics that is now based upon game theoretical arguments.

4. A linguistic application

We return to the linguistic application mentioned in the beginnings. We repeat the sentence, and its representation in the logic:

(4) *Some friend of each townsman and some neighbour of each villager hate each other*

(5) $\forall x_1 \exists x_2 \forall x_3 \exists x_{4/x_1,x_2}[T(x_1) \wedge V(x_3) \to F(x_2, x_1) \wedge N(x_4, x_3) \wedge H(x_2, x_4)]$

Consider now the following situation. Among the friends of the townsmen and two groups are distinguished, viz. male and female ones, and the same among the neighbors of the villagers. Assume now that hating is a relation between all pairs of male friends and male neighbors, and also between female friends and female neighbors, but not between friends and neighbors of different sexes. In this situation the choices of friends for townsmen and neighbors for villagers have to correspond: in both cases male ones, or female ones. So intuitively the choice for $\exists x_{4/x_1,x_2}$ cannot be made independently of the choice for $\exists x_2$. So in this model sentence (4) should not be true. However, (5) comes out true in Hintikka semantics: coordinate the strategies such such that both yield male friends and neighbors respectively. This shows that the required independence is not captured by his interpretation.

Let us now consider our analysis of (5). Suppose now that in the model under discussion a male friend has been chosen. Then a male neighbor must be chosen as well, say Jacob. And if a female friend had been chosen only the choice for a female neighbor would be winning. But the condition of rationality requires that for female friend the original choice Jacob would be winning as well. That is not the case in the given model, so the formula is not true. For this example our semantics gives the desired result, whereas that is not the case for game theoretical semantics. I that also for other applications (e.g. quantum mechanics) that will be the case.

5. Conclusions

We have proposed an alternative interpretation for Hintikka's game theoretical semantics for IF logic. It is game theoretically very natural and it yields better results for the investigated linguistic application.

Acknowledgements

I thank Merlijn Sevenster for several discussions on IF; without his inspiring paper Sevenster 2005 this could not have been done.

Bibliography

Apt: 2004, The many faces of rationalizabilty, unpublished paper

van Benthem, J.: 2004, *The epistemic logic behind IF games*, in R. Auxier editor Hintikka Volume, Library of Living Philosophers.

Hintikka, J.: 1996, *The Principles of Mathematics Revisited*, Cambridge University Press

Hintikka, J. and Sandu, G.: 1997, Game-theoretical semantics, in J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic & Language*, pp 361–410, Elsevier Science

Hodges, W.: 1997, Compositional semantics for a language of imperfect information, *Bulletin of the IGPL* 5 (4), 539–563

Janssen, T. M. V.: 2002, Independent choices and the interpretation of IF-logic, *JoLLI* 11, 367–387

Sevenster, M.: 2005, *A player perspective on IF semantic games*, to appear