A Subjective Game-Theoretic Semantics for First Order Logic

B. Sidney Smith

27 March 2008

ABSTRACT: A subjective variation of game-theoretic semantics is introduced in which the choice trees represent the state of knowledge of the reasoner. The semantics is a generalization of the standard semantics of first-order quantification to situations in which the reasoner's state of knowledge about the domain is incomplete or contingent, or in which the domain of discourse is imperfectly specified or not set-like.

1 Introduction and Motivation

Except over finite domains, the standard Tarski semantics for first-order predicate logic (FOL) takes the quantifiers "for all" and "there exists" to be primitive logical notions. Over a finite domain they may be taken to correspond to finite sentences (conjunctions and disjunctions respectively), but this interpretation is not available over infinite domains, to say nothing of uncountable ones. It may be objected that in mature treatments of FOL the semantics of the quantifiers is defined model-theoretically, i.e., as statements about sets. However, this can only follow some development of set theory in which quantification must be used—for example for extensionality—so on our present understanding the first use of quantification is always primitive, and all subsequent refinements of the semantics trace to such usage.

Prior to Tarki's treatment, in 1923 David Hilbert introduced the ε-operator in an effort to prove the consistency of mathematical analysis[1]. This was a logical operator that replaced bound variables with "indeterminate constants" having the property that they may be interpreted as witnesses (or falsifiers) of the formulas in which they occur. Hilbert was motivated by a desire to replace the intrinsically non-finitistic quantifiers with a finitistic statement about logical choice. The precise relationship between Hilbert's operator and Skolem functions, introduced at about the same time, is rather difficult to pin down. Both represent a means of quantifier elimination in favor of introducing new constant or function symbols to represent witnesses or falsifiers. However, a Skolem function is explicitly a function, and picks out a distinguished representative. By contrast, the ε-operator was not treated as though it distinguished or determined its representative, so the "witnessing" that occurs is not particular (except over well-founded domains, where one may assume one has the least such witness)[3]. In this respect, the ε-operator preserves the non-specificity of "primitive" quantification.

As a "logical choice" function, the ε-operator is often taken to be tantamount to a statement about set existence, effectively subsuming it under the Tarskian conception of quantification. (On this view, it is equivalent to the Axiom of Choice.) An alternative would be to understand "logical choice" as a statement about certain capacities of the reasoner, specifically, about the (possibly contingent) capability of the reasoner to *select* witnesses or falsifiers for a predicate. There is no reason to believe that a subjective interpretation such as this motivated Hilbert, but it is the possibility of such an interpretation that motivates the development we undertake here.

Understanding logical choice subjectively has immediate consequences. First, sentences can fail to have truth values. There may be a predicate P over some domain and an element a of that domain such that Pa is in fact true (or false), but it may not be within the capacity of the reasoner to select a or any other element to witness (or falsify) P. In this event, under a subjective semantics, $\exists x Px$ will not have a determinate truth value.[1] Of necessity then, under a subjective semantics FOL will be a tri-valent logic.

Another consequence is that for n-ary predicates with $n \geq 2$, the order in which elements are chosen to fill the predicate may affect the choices available to the chooser for those elements, irrespective of the quantifiers involved. Moreover, earlier choices may affect later ones.

To illustrate these consequences, consider a large room divided by a transparent partition. There are one-way doors that permit entry (only) into each side of the room. There is also a door in the partition, programmed to open only in response to the sound of a car horn. On one side of the partition are several late-model automobiles, all locked. On the other side, neatly laid out on a table, are some car keys.[2] In this tableau, let the domain of discourse consist of the keys and cars, let Pxy represent the open formula "x unlocks y", and consider the truth of the sentence $\exists x \exists y Pxy$. This sentence is true, *á la* Tarski, so long as one of the keys fits one of the cars. But for someone acting as an agent in the tableau of this domain the truth of this sentence is contingent. For suppose Sam enters the side of the room containing the cars, and Sue enters the side of the room containing the keys. Sue has the capacity to witness or falsify the sentence, for she has only to pick up each key in turn and press the "honk" button. If one causes a car on the other side of the partition to honk, the partition door will unlock and she can go through to identify the corresponding automobile, and doing so has witnessed the sentence. If none of them does, she has falsified the sentence. Either way, the sentence has a determinate truth value for Sue. On the other hand, Sam has no way of determining whether any of the cars on his side have a corresponding key on the other side, so he cannot witness the sentence. Neither can he determine that there are *no* keys that will fit, so he cannot falsify the sentence. The subjective truth of $\exists x \exists y Pxy$ is indeterminate for Sam.

Some further observations about this example are pertinent. If we take the order of the quantifiers in $\exists x \exists y Pxy$ to indicate the order in which objects are chosen to fill the corresponding variables, then this order matters (subjectively). Anyone understanding the situation described above can confirm the truth of $\exists x \exists y Pxy$ by walking first through the one-way door that leads to the keys, finding a key

---

[1] It is tempting to treat a sentence that cannot be witnessed as false. However, in the event that it's negation also can't be witnessed, we would have a false sentence whose negation is also false—a sentential contradiction.

[2] These keys are of the recent electronic variety, with buttons to lock and unlock the car doors automatically, and to make it honk.

for x, and then *with this key in hand* selecting the corresponding car to substitute for y. However, none can first choose a car and *then* find a key that fits it, because once the room has been entered on the side of the cars no key can be chosen. Consequently, there is a definable sense in which $\exists y \exists x Pxy$ is not equivalent to $\exists x \exists y Pxy$. In this instance the latter is determinable and the former is not, but we will show below that under a formalized subjective semantics sentences of this form can both be determinable, yet have opposite truth values (for the same reasoner).

We note also that although Sam's disadvantage in this case was owing to circumstances that pertained in the situation, we do not suppose "subjectivity" of logical choice always to arise from external constraints. It is possible, for instance, that we lied to Sam and that in fact the partition door is not locked. His failure to choose a key then arises from a constraint that is internal to Sam—his erroneous belief that he lacks the means. It is also possible that Sam has a defective concept of "key" that prevents him recognizing the car keys for what they are, and so on. It is even possible that we have underestimated Sam; perhaps he kicks the tires of each car in turn until the anti-theft alarm of one is triggered, honking and flashing and opening the locked partition door for him, so he can determine the truth of the sentence after all.

By the "subjectivity" of logical choice, then, we mean that the capacity to choose a witness or falsifier is specific to the agent and not determined *a priori*. We do not, however, imply any subjectivity about *propositional truth*: we treat the truth of any variable-free sentence as objectively determined. If Sam is able to choose both keys and cars, then he will fix the same truth value for $\exists x \exists y Pxy$ as any other agent able to choose exactly the same keys and cars.

2 Subjective Interpretations

2.1 Preliminaries

Our language is that of standard first-order predicate logic with quantifiers (FOL), with the usual symbols and syntax. We will consider only sentences; any open formulas will be taken to be existentially generalized over the free variables. We will assume that the domain of discourse is a set and that any given n-ary predicate symbol is associated with a corresponding n-ary relation over the domain. An n-ary predicate P followed by n constant symbols will be called an **atomic proposition**, and will be called a true atomic proposition if the n-tuple of constants is an element of the corresponding relation, and false otherwise. In general, we use the term **proposition** for any well-formed formula involving only atomic propositions and logical connectives.

2.2 Knowledge Functions

To determine a given reasoner's subjectively available choices of individual witness or falsifier to replace a variable in a given sentence, at a given point in the process of attempting to witness or falsify the sentence, several pieces of information must be available. These include the predicates involved, the places of the predicates for which choices have previously been made, what those choices were, and the places of the predicates for which the current choice is to be made. We begin by defining a class of functions, called **knowledge functions**, for each predicate in the language.

Given an n-ary predicate, we require one function for each combination of (1) a possible subset of argument places for which prior choices will have been made and (2) a non-empty subset of the remaining argument places to fill.

To visualize this, let $P\circ\circ\circ\circ$ represent a 4-place predicate in which none of the argument places has been filled. We'll represent filled places by $\bullet$ and places for which a choice is presently to be made by $\square$. Thus, $P\bullet\square\bullet\circ$ represents the situation in which places 1 and 3 have been filled by prior choices (or possibly by the same choice), and a choice is now to be made for place 2. On the other hand, $P\bullet\square\bullet\square$ represents the situation in which a choice or choices have been made for places 1 and 3, and a single choice is now to be made that will fill both places 2 and 4. The currently available choices must be represented by a function that maps tuples of elements of the domain to subsets of the domain. In either one of these example cases the tuple would be an ordered pair, representing the prior choices, and the subset to which it is mapped would be the subset of elements from which the current choice may be drawn, given those prior choices. We now make this precise:

> **Definition 2.2.1** Knowledge Functions
>
> Let P be an n-ary predicate over a domain D, let α denote a proper subset of $\{1, 2, \ldots, n\}$, and let β denote a non-empty subset of $\{1, 2, \ldots, n\} \backslash \alpha$. We denote by $|\alpha|$ the number of elements in α. A function $P_{(\alpha,\beta)} : D^{|\alpha|} \longrightarrow \mathcal{P}(D)$ mapping $|\alpha|$-tuples of elements of D to subsets of D is called a **knowledge function** for P.

We observe that α may be empty, in which case β may (and may not) be the entire set of indices of argument places for P. If α is empty we take $D^{|\alpha|}$ to be the 1-element set $\{\emptyset\}$, and the knowledge function then points to the subset of D from which a choice may be drawn to fill the argument places whose indices are in β when no prior choices have been made for any argument places of P. We observe also that the empty set may be in the range of $P_{(\alpha,\beta)}$; when the empty set is returned it means that no choice is possible.

Given a knowledge function $P_{(\alpha,\beta)}$ and an $|\alpha|$-tuple $\overline{a} \in D^{|\alpha|}$, we may refer to $P_{(\alpha,\beta)}(\overline{a})$ as a **knowledge set**.

The knowledge functions of a predicate form the subjective basis of a semantics for that predicate. Taken together, they give a complete picture of how, for a given reasoner, the predicate applies to objects in the world. In short, they are the elements of comprehension.

> **Definition 2.2.2** Comprehensions
>
> Let P be an n-ary predicate over a domain D, and let Γ_P be a set of knowledge functions for P consisting of one knowledge function for each distinct pair (α, β), where α and β are defined as above. Then Γ_P is called a **comprehension** of P.

The comprehension of an n-ary predicate always consists of $3^n - 2^n$ functions. This is because the number of distinct subsets α of $\{1, 2, \ldots, n\}$ of size i is $\binom{n}{i}$, and for each of these there are

$2^{n-i}-1$ choices for β. We have

$$\begin{aligned}
|\Gamma_P| &= \sum_{i=0}^{n}\binom{n}{i}\left(2^{n-i}-1\right)\\
&= \sum_{i=0}^{n}\binom{n}{i}2^{n-i}-\sum_{i=0}^{n}\binom{n}{i}\\
&= (1+2)^n-(1+1)^n \qquad \text{(Binomial Theorem)}\\
&= 3^n-2^n
\end{aligned}$$

So a unary predicate has one knowledge function with domain $\{\emptyset\}$, corresponding to the situation $P\Box$, which may be identified with the corresponding knowledge set, i.e., the set of choices for P. (We emphasize that the knowledge set for a unary predicate P is not the set of objects that witness P, but the set of objects that—for the reasoner—may be chosen to witness *or* falsify P.) The comprehension of a binary predicate has 5 knowledge functions, corresponding to these situations:

$$P\circ\Box \quad P\Box\circ \quad P\bullet\Box \quad P\Box\bullet \quad P\Box\Box$$

The 19 knowledge functions for a 3-ary predicate correspond to the following situations:

$$\begin{array}{lllllll}
P\circ\circ\Box & P\circ\bullet\Box & P\bullet\circ\Box & P\bullet\bullet\Box & P\circ\Box\Box & P\bullet\Box\Box & P\Box\Box\Box\\
P\circ\Box\circ & P\circ\Box\bullet & P\bullet\Box\circ & P\bullet\Box\bullet & P\Box\circ\Box & P\Box\bullet\Box & \\
P\Box\circ\circ & P\Box\circ\bullet & P\Box\bullet\circ & P\Box\bullet\bullet & P\Box\Box\circ & P\Box\Box\bullet &
\end{array}$$

2.3 Percepts

Comprehensions force upon us a new kind of logical expression. To introduce these expressions we return to our example with keys, cars, and the divided room, where the predicate P was defined by $Pxy \Leftrightarrow$ "x unlocks y." Suppose that Sue, having entered on the side of the room with the keys, finds that key a causes car b to honk, unlocking the door in the center partition, and that she is then able to witness the sentence $\exists x\exists yPxy$ by substituting a for x and b for y. We could say that Sue has *perceived* that with key a in hand she can then choose car b to witness the sentence $\exists x\exists yPxy$ by means of the true proposition Pab. However, regarding Pab itself as the witnessing proposition neglects the subjectively essential priority of the key. To indicate the priority of a in the witnessing process we shall write Pab^a, where the superscript indicates that the choice of car b depends upon the prior choice of key a, and we will call this expression a **percept**. Thus, a percept is a proposition in which the constant symbols are ordered.

Definition 2.3.3 Percepts

Let P be an n-ary predicate and let Γ_P be a comprehension for P. A linearly ordered set of constant symbols $\{a_1,\ldots,a_k\}_{k\leq n}$, together with an assignment of those symbols to the argument places of P, is called a **percept for P under Γ_P** if and only if

1. every constant in the set is assigned to one or more argument places of P, and

2. for every constant a_j in the set, if α is the set of indices of argument places to which constants superior to a_j in the order are assigned and β is the set of indices of argument places to which a_j itself is assigned, then a_j is an element of the corresponding knowledge set under Γ_P.

Thus a comprehension of a predicate determines a set of percepts for that predicate.

As indicated, percepts will be denoted using standard literal notation plus superscripts. Thus, $Pab^a c^{ab} b^a$ denotes a percept for a 4-ary predicate in which $a > b > c$ and where a is assigned to the first argument place, b to the second and fourth argument places, and c to the third argument place. This percept asserts that an object a may be chosen, that an object b may subsequently be chosen, and that with a and b in hand a third object c may be chosen, such that $Pabcb$ is a subjectively meaningful proposition, i.e., either true or false.

In the definition just given we specified a linear order on the constant symbols because the subjective semantics we are developing will be defined as a game on sentences in which the variable symbols in literals are linearly ordered by the order of their respective quantifiers. However, it would be more natural to demand only a partial order of the constants, allowing for a capacity of the reasoner to pick witnesses or falsifiers for distinct variables independently of one another. We will introduce this change in Chapter ***, and the resulting logic will be “independence friendly” in the sense of Hintikka [2].

A percept is characterized, up to uniform substitutions of the constant symbols, by the ordering among the symbols assigned to argument places, and we will refer to this as the **pattern** of the percept. Let $S(n,k)$ represent the Stirling number for (n,k), the number of ways of partitioning an n-element set into k non-empty subsets. Then the number $\mathbf{p}(n)$ of distinct linearly ordered patterns for an n-ary predicate is given by $\mathbf{p}(n) = \sum_{k=1}^{n} k!\, S(n,k)$.

2.4 Strategies

In the next section we will define the truth value of a sentence as the outcome of a game between a witness-player and a falsifier-player. In this section we define the game.

Definition 2.4.4 Semantic Game for Atomic Sentences

Let φ be an atomic sentence of FOL, i.e., φ is a predicate symbol P followed by one or more variables and/or constants and preceded by quantifiers such that each variable symbol falls within the scope of a quantifier. Let Γ_P be a comprehension for P, and let $\mathcal{P}_\varphi$ be the set of percepts for P under Γ_P in which any constants already assigned to argument places of P by φ are superior in the pattern of the percept to any constants yet to be chosen, and which otherwise match the pattern determined by the order of the quantifiers that scope the variables occuring in φ. If $\mathcal{P}_\varphi$ is empty then φ is undetermined. Otherwise, the players proceed by selecting constant symbols to replace the variable symbols, if any, in such a way that:

1. The witness-player selects for existentially quantified variables, and the falsifier-player selects for universally quantified variables,

2. The constants are chosen in the order corresponding to the pattern of φ, and
3. Every choice by either player is consistent with some sequence of possible choices for any remaining variable symbols such that the resulting percept is in $\mathcal{P}_\varphi$.

The witness-player wins the game if the atomic proposition that results is true, and the falsifer-player wins otherwise. Any choice made by a player that is part of a winning strategy is said to be a *strategic choice* for that player. We may refer to a possible choice that is not part of a winning strategy as being "strategic" for the opposing player.

Since this is a finite game, either the witness-player or the falsifier-player has a strategy, but not both, so long as $\mathcal{P}_\varphi$ is not empty. That is, an atomic sentence φ is determined just in case there are any percepts of the correct pattern for its predicate. For example, suppose that P is a 3-ary predicate, and that under a given comprehension the set of percepts for P is

$$\{Pab^a c^{ab}, Pa^b a^b b\}$$

Then the sentences $Pabc$, $Paab$, $\exists x \forall y Paxy$, and $\forall y \forall x Pxxy$ are determined, but the sentences $Pbac$, $\exists x Paax$, and $\forall x \forall y Pxxy$ are not determined.

Definition 2.4.5 Compatible Strategies

Let φ and ψ be sentences of FOL in prenex form, such that if x and y are any two variable symbols that have occurrences in both φ and ψ then they are bound by like quantifiers in φ and ψ, and if the quantifier binding x precedes the quantifier binding y in φ, then the quantifier binding x likewise precedes the quantifier binding y in ψ. Let S_φ be a witness-player (or falsifier-player) strategy for φ, and let S_ψ be a witness-player (resp. falsifier-player) strategy for ψ. Then S_φ and S_ψ are **compatible strategies** if the same choices are made in each strategy for any variable symbols that have occurrences in both φ and ψ.

This notion of compatible strategies permits us to extend the semantic game to non-atomic sentences of FOL by recursion on logical connectives. We may assume without loss of generality that any sentence is in prenex form, i.e., that all quantifiers have been moved to the front. This ensures that the conditions of the definition for compatible strategies will have been met for any subformulas.

Definition 2.4.6 Semantic Game for Non-Atomic Sentences

Let χ be a sentence of FOL in prenex form.

1. If χ is of the form $\neg\varphi$, then the witness-player has a strategy for χ if and only if the falsifier-player has a strategy for φ, and the falsifier-player has a strategy for χ if and only if the witness-player has a strategy for φ.
2. If χ is of the form $\varphi \vee \psi$, then the witness-player has a strategy for χ if she has a strategy for either φ or ψ, and the falsifier-player has a strategy for χ if she has compatible strategies for φ and ψ.

3. If χ is of the form $\varphi \wedge \psi$, then the witness-player has a strategy for χ if she has compatible strategies for φ and ψ, and the falsifier-player has a strategy for χ if she has a strategy for either φ or ψ.
4. If χ is of the form $\varphi \rightarrow \psi$, then a player has a strategy for χ if she has a strategy for $\neg\varphi \vee \psi$.

It is evidently impossible for both players to have a strategy for a sentence, but both can fail to have a strategy. If neither player has a strategy for χ then χ is undetermined.

2.5 Truth and Consequences

We are now in a position to present the complete semantics.

Definition 2.5.7 Interpretations

A **subjective, game-theoretic interpretation** I of FOL consists of:

1. A domain D.
2. An assignment of each constant symbol to some element of D.
3. An assignment of a comprehension Γ_P to each predicate symbol P.
4. An assignment of each n-ary predicate symbol to an n-ary relation on D.
5. A truth-valued function V_I on sentences, given by

$$V_I(\varphi) = \begin{cases} T & \text{if the witness-player has a strategy for } \varphi \\ F & \text{if the falsifier-player has a strategy for } \varphi \\ Null & \text{if } \varphi \text{ is undetermined.} \end{cases}$$

We follow the convention of saying that an interpretation **satisfies** a sentence φ if $V_I(\varphi) = T$, and that a sentence is **satisfiable** if there is some interpretation that satisfies it. However, the presence of comprehensions mediating the truth values of sentences alters the nature of the semantics for FOL considerably. In particular, there are no *logically valid* sentences under a subjective semantics, since under any given interpretation any sentence may evaluate to $Null$. Thus, while we may write "$\varphi \models \psi$" to indicate that every interpretation that satisfies φ satisfies ψ, we will never write "$\models \varphi$" for any φ because for some interpretation I we will have $V_I(\varphi) = Null$.

Thus for instance, the sentence $\exists x P x \wedge \neg\exists x P x$ is not satisfiable, since the witness-player has a strategy for $\exists x P x$ if and only if the falsifier-player has a strategy for $\neg\exists x P x$, and *vice-versa*. However, the negation of this sentence, $\neg\exists x P x \vee \exists x P x$, is not logically valid, even though it is a sentential tautology, since for any interpretation I in which the knowledge set for P is empty the sentence is not determined, and so not satisfied.

3 Subjective Proofs

3.1 Tautological Inference

It follows from the definitions in the previous chapter that for any sentences φ and ψ and any subjective interpretation I the valuation function V_I obeys Kleene's rules for (strong) ternary logic:

$V_I(\varphi)$	$V_I(\psi)$	$V_I(\neg\varphi)$	$V_I(\varphi \vee \psi)$	$V_I(\varphi \wedge \psi)$	$V_I(\varphi \rightarrow \psi)$
T	T	F	T	T	T
T	F	F	T	F	F
T	$Null$	F	T	$Null$	$Null$
F	T	T	T	F	T
F	F	T	F	F	T
F	$Null$	T	$Null$	F	T
$Null$	T	$Null$	T	$Null$	T
$Null$	F	$Null$	$Null$	F	$Null$
$Null$	$Null$	$Null$	$Null$	$Null$	$Null$

As a consequence, although there are no tautologies, we may nonetheless make valid tautological inferences from a non-empty set of premises provided the inference is a tautological consequence in strong tri-valent propositional logic. That is, if $\varphi_1, \ldots, \varphi_k$ are sentences of FOL and if $(\varphi_1 \wedge \cdots \wedge \varphi_k) \models_{\text{Kleene}} \psi$, then ψ may be inferred as a tautological consequence of $\varphi_1, \ldots, \varphi_k$. For an example we demonstrate the soundness of the cut rule.

Theorem 3.1.1 Soundness of the Cut Rule

Let δ, λ, and φ be sentences of FOL. Then for any subjective interpretation I, if $V_I(\delta \vee \varphi) = T$ and $V_I(\lambda \vee \neg\varphi) = T$, then $V_I(\delta \vee \lambda) = T$.

Proof: If $V_I(\varphi) = Null$ then $V_I(\delta \vee \varphi) = T \Rightarrow V_I(\delta) = T$, so $V_I(\delta \vee \lambda) = T$. If $V_I(\varphi) = T$ then $V_I(\neg\varphi) = F$, so $V_I(\lambda \vee \neg\varphi) = T \Rightarrow V_I(\lambda) = T$ and hence $V_I(\delta \vee \lambda) = T$. Finally, if $V_I(\varphi) = F$ then $V_I(\delta \vee \varphi) = T \Rightarrow V_I(\delta) = T$, so again $V_I(\delta \vee \lambda) = T$.

3.2 Strategic Inference

Under the usual semantics the non-tautological inference rules for FOL depend upon a set-theoretic interpretation of the quantifiers. Under a subjective semantics, in which the quantifiers represent the reasoner's comprehension of the predicates with respect to the domain, we find that few of the usual inference rules are generally valid. Moreover, we are led to new distinctions and characterizations concerning comprehensions themselves that affect how large a subset of the usual inference rules remain valid for a given interpretation.

Theorem 3.2.2 Strategic Inference Rules (General)

The following rules of inference are sound for all subjective, game-theoretic interpretations of FOL and all formulas φ and ψ with one free variable x.

1. $\dfrac{\forall x \varphi}{\exists x \varphi}$
2. $\dfrac{\forall x \neg \varphi}{\neg \exists x \varphi} \qquad \dfrac{\exists x \neg \varphi}{\neg \forall x \varphi}$

Proof:

1. If $\forall x \varphi$ holds then a choice exists for x for φ, and since it is not a strategic choice for the falsifier-player it must be a strategic choice for the witness player, hence $\exists x \varphi$ also holds.
2. This is an immediate consequence of Definition 2.4.6 (1).

. . . in progress as of December 2008 . . .

BIBLIOGRAPHY

[1] Hilbert, D., *Die logischen Grundlagen der Mathematik*, Math. Annalen 88, 151-165, 1923.

[2] Hintikka, J., *The Principles of Mathematics Revisited*, New York: Cambridge UP, 1996.

[3] Leisenring, A. C., *Mathematical Logic and Hilbert's Epsilon-Symbol*, London: Macdonald, 1969.

[4] Mycielski, J., P. Pudlak, and A.S. Stern, *A Lattice of Chapters of Mathematics*, Memoirs of the AMS, Vol. 84, No. 426, 1990.