344
Notre Dame Journal of Formal Logic
Volume XVII, Number 3, July 1976
NDJFAM

A LOGIC OF BELIEF

ALEX BLUM

Our object in this paper* is to construct a purely extensional first-order system **S** adequate for the systematization of first-order belief sentences.

1 Any satisfactory systematization of belief sentences would have to fulfill, it would appear, the following conditions:

One, that if

(1) Ralph believes of Ortcutt, that he is a spy

is true, so is

(2) $(\exists x)$ Ralph believes of x, that x is a spy

and hence, (1) and

(3) Ortcutt = the mayor of Hanoi

entail

(4) Ralph believes of the mayor of Hanoi that he is a spy.

Two, that even if (3) and

(5) Ralph believes that Ortcutt is a spy

are true,

(6) Ralph believes that the mayor of Hanoi is a spy need not be true.

And *three*, that (2) entails

(7) Ralph believes that $(\exists x)$ x is a spy.

2 To facilitate understanding, we begin with the semantic motivation for **S**. We view the universe as a set of domains (not all distinct) of individuals

*I am deeply indebted to Professor Raziel Abelson and to my students, Judith Rosenberg and David Widerker, for very helpful discussions on the logic of belief.

Received May 9, 1973

[Di], such that an individual α is an element of Di, if and only if, i believes that α exists. We have as a distinct domain Dg, where g is an individual whose set of beliefs are identical to the set of truths. Dg thus consists of the set of all existent individuals. On our intended interpretation, each predicate of S is a belief-predicate. Thus, no sentence in S will be read as 'Ralph is a spy'. Instead, we have 'G believes that Ralph is a spy' or more briefly 'Ralph is a spy for g'.

3 We now move to S. S is an aleph$_0$-sorted first-order system whose language is built up from the following elements:

(i) An infinite list of individual constants of each sort, a^a, b^a, c^a, . . ., a^b, b^b, c^b, . . . (to be read: 'the entity believed by a to be a, by a to be b, by a to be c, . . ., by b to be a, by b to be b, by b to be c, . . .' or more briefly: 'a for a, b for a, c for a, . . . a for b, b for b, c for b, . . .');
(ii) An infinite list of individual variables of each sort, x^a, y^a, z^a, . . ., x^b, y^b, z^b, . . .;
(iii) An infinite list of n-place predicates for each n:$n \geq 1$, Fn, Gn, Hn, . . ., $F'n$, $G'n$, $H'n$, . . . ('Fx' and 'Fxy' are to be read as 'x believes that F' and 'y believes that x is F' or more briefly 'F for x' and 'x is F for y');
(iv) The three place identity predicate I ('Ixyz' is read 'z believes that x and y are identical', or more briefly 'x is identical to y for z');
(v) The logical constants, $\sim$, $\supset$, $\vee$, ., $\equiv$, and for each individual variable x^i, (x^i) and $(\exists x^i)$;
(vi) The standard punctuation marks.

The formation rules for sentencehood in S are the standard ones for first-order languages. And, the deductive apparatus of S is the same as those of standard many-sorted systems without identity, with the addition of the following three axiom schemata:

[To simplify matters, we shall be guided by the following convention: An individual symbol without a superscript is to be understood as having the superscript of its quantifier, if any, otherwise it has 'g' for its superscript.]

(I1) $(x^i)(\exists y^g)\mathrm{I}xyi$

i.e., each x for i, is identical to some g for i;

(I2) $(x^i)\mathrm{I}xxg$

i.e., each x for i is self-identical for g;

(I3) $(x^i)(y^j)(Fx \,.\, \mathrm{I}xyg \,.\supset Fy)$

i.e., if it is F for x, and x and y are identical for g, then it is F for y.

Hence, while

(T1) $(x^i)(y^j)(z^k)(\mathrm{I}xyg \,.\, \mathrm{I}yzg \,.\supset \mathrm{I}xzg)$

and

(T2) $(x^i)(y^j)(z^k)(\mathsf{I}xyg \supset \mathsf{I}yxg)$

are theorems of **S**, the following are not:

(F1) $(x^i)\mathsf{I}xxi$
(F2) $(x^i)(y^j)(z^k)(u^e)(\mathsf{I}xyu\ .\ \mathsf{I}yzu\ .\supset \mathsf{I}xzu)$
(F3) $(x^i)(y^j)(z^k)(\mathsf{I}xyz \supset \mathsf{I}yxz)$
(F4) $(x^i)(\exists y^g)\mathsf{I}xyg$

(We argue for these assertions in the appendix.)

4 Semantically we view the predicate of an expression as consisting of the predicate letter and its last argument. Thus, for example

(i) $\mathsf{I}xyj$

is to be viewed as

(i′) $\langle x, y\rangle \in \mathsf{I}_j$

while

(ii) $(z)\mathsf{I}xyz$

and

(iii) $(\exists z)\mathsf{I}xyz$

are to be thought of as second-order statements, entailing, and being entailed, respectively, by any (and all) of the following:

(iv) $\langle x, y\rangle \in \mathsf{I}_a, \langle x, y\rangle \in \mathsf{I}_b, \langle x, y\rangle \in \mathsf{I}_c, \ldots$

Our domain $\mathbf{D}g$ corresponds to the universal domain of classical quantification theory ('Q' for short) and our g-subscript predicates correspond to their unsubscribed counterparts in Q. The difference between a g-subscribed predicate and its Q counterpart is that the range of a g-subscribed predicate is not limited to $\mathbf{D}g$ while that of Q is.

5 Let us now see how sentences (1) to (7), our original motivation for **S**, fare in **S**. In **S**, (1) to (7) become, respectively,

(1′) $Sor\ .\ (\exists y^r)\mathsf{I}oyg$
(2′) $(\exists x^g)(\exists y^r)(\mathsf{I}xyg\ .\ Sxr)$
(3′) $\mathsf{I}o, (\iota x^g)Mxg, g$
(4′) $S(\iota x^g)Mxg, r$
(5′) $(\exists x^r)(Sxr\ .\ \mathsf{I}xor)$
(6′) $(\exists x^r)(Sxr\ .\ \mathsf{I}(\iota x^g)Mxg, x, r)$
(7′) $(\exists x^g)Sxr$

Our rendition of (1) to (6) into (1′) to (6′) is fairly straightforward. (7) into (7′), however, calls for explanation. It would appear that: (i) while the truth of (7′) commits us to the existence of some entity which need not be Ralph, (7) does not; and (ii) while (7′) says that Ralph believes of someone (or other) that he is a spy, (7) does not. With regard to (i), not only does (7′) commit us to the nonemptiness of **D**g, but so does (7) in Q, given the validity of '$(\exists x)(\exists y)(Fx \supset Fy)$'. The point raised in (ii) brings into focus the rationale of **S**. (7′) is clearly consistent with the claim that Ralph has absolutely no identificatory beliefs about the object x of which he thinks is a spy. But then (7′) surely says no more than (7). For believing absolutely nothing about a being other than that it exists and that it is a spy is no more than believing that there is a being who is a spy. Our motivating desiderata are now met. For clearly (1′) entails (2′), (2′) entails (7′), (1′), and (3′) entail (4′), while (5′) and (3′) fail to entail (6′).

But is **S**, expressively complete? How would we express complex belief sentences such as:

(8) Ralph believes that if O lies then O is a spy.
(9) Ralph believes that J believes that O is a spy.
(10) Ralph believes that Ralph believes that O is a spy.

and

(11) God believes that if O lies then O is a spy?

We express them as follows:

(8′) Kor

where 'K' is 'if ① lies then ① is a spy for ②';

(9′) $Mojr$

where 'M' is '① is a spy for ②, for ③';

(10′) $Morr$

and

(11′) Kog

or, more perspicuously, as

(11″) $Log \supset Sog$

where 'L' is '① lies for ②' and 'S' is '① is a spy for ②'.

(11″) is also a formalization of:

(11^x) If O lies then O is a spy.

That is, **S**'s expressive power is adequate to exhibit all the logical structure that is needed for the logic of first-order belief sentences. For the only postulate regarding the logical acumen of ordinary individuals is given by the closed-ended postulate (I1). As far as g is concerned, the logical acumen of g is given by the rules of Q, expressed in **S** by the ordinary axiom schemata and (I2) and (I3).

APPENDIX

Theorem 1 (T1) *is a theorem of* S.

Proof: Let '*F*', '*x*', and '*y*' in (I3) be replaced respectively by '$\mathsf{I}x②g$', '*y*', and '*z*'. We then have

$$\mathsf{I}y^{x②g} \,.\, \mathsf{I}yzg. \supset \mathsf{I}z^{x②g}$$

and by UG, we get (T1).

Theorem 2 (T2) *is a theorem of* S.

Proof: Let '*F*' in (I3) be replaced as before and let x and y remain as they are. We then have

$$\mathsf{I}x^{x②g} \,.\, \mathsf{I}xyg. \supset \mathsf{I}y^{x②g}$$

and by (I2) and UG get (T2).

The invalidity of (F1), (F2), and (F3) follow from the following consideration: The only assumption we make about the logical acumen of ordinary individuals is given by (I1). (F1), (F2), and (F3), however, are not logical consequences of (I1).

If (F4) were a theorem, then for each i, Di would be a subset of Dg. But the only link between Dg and Di are given in (I1), and the relation there is I_i and not I_g.

Bar-Ilan University
Ramat-Gan, Israel

and

The Hebrew University
Jerusalem, Israel