

## PROBABILITIES ON MODELS OF UNIVERSAL SENTENCES<sup>1</sup>

DOUGLAS N. HOOVER<sup>2</sup>

ABSTRACT. We show that the asymptotics of conditional probabilities of first order universal sentences on finite models are the same as those of general first order sentences. This answers a question of R. Fagin.

By an *L-theory* we mean a consistent set of sentences in a first order language  $L$ ; by an *L-model*, a model for  $L$ ; and by an *L-n-model*, an  $L$ -model whose universe is  $\{0, \dots, n-1\}$ .  $\beta_n^L$  is the number of  $L$ - $n$ -models, and if  $T$  is an  $L$ -theory then  $\beta_n^L(T)$  is the number of  $L$ - $n$ -models satisfying  $T$ . We shall always assume that our languages have a vocabulary consisting of finitely many relation symbols (no function symbols); hence  $\beta_n^L$  is finite.  $\mu_n(T)$ , the probability that a model of size  $n$  satisfies  $T$ , is defined by

$$\mu_n(T) = \beta_n^L(T) / \beta_n^L,$$

where  $L$  is any finite language containing all predicates occurring in  $T$ . It is easily seen that the ratio does not depend on  $L$ , because the addition of new predicates multiplies numerator and denominator by the same factor. We define also the conditional probability

$$\mu_n(T_1 | T_2) = \mu_n(T_1 \cup T_2) / \mu_n(T_2),$$

provided that  $\mu_n(T_2) \neq 0$ . If  $\phi$  and  $\psi$  are sentences, we put

$$\mu_n(\phi) = \mu_n(\{\phi\}), \quad \mu_n(\phi | \psi) = \mu_n(\{\phi\} | \{\psi\}), \quad \mu_n(\phi | T) = \mu_n(\{\phi\} | T),$$

etc. We also use  $T, \phi$  for  $T \cup \{\phi\}$ , i.e.,  $\mu_n(T, \phi) = \mu_n(T \cup \{\phi\})$ .

Evidently, for  $\phi, \psi$  and  $T$  in  $L$ ,  $\mu_n(\phi | T, \psi) = \beta_n^L(T, \phi, \psi) / \beta_n^L(T, \psi)$ .

Glebskii et al. [3] and Fagin [1] independently showed that for  $\phi$  a first order sentence of a finite relational language  $L$ ,  $\mu_n(\phi)$  converges to either 0 or 1 as  $n$  tends to infinity. Fagin observed that a similar convergence cannot be expected for the conditional probability  $\mu_n(\phi | \psi)$ , even when  $\mu_n(\psi) > 0$  for all  $n$ . In fact, Corollary 2 shows that conditional probabilities can approximate the characteristic function of the finite spectrum of any first order sentence, and such spectra are

Received by the editors April 21, 1983 and, in revised form, August 24, 1984.

1980 *Mathematics Subject Classification*. Primary 03C13; Secondary 03B48, 03D15.

*Key words and phrases*. Probabilities on finite models, conditional probabilities, finite spectra.

<sup>1</sup>This research has been supported by a grant-in-aid from the Institute for Advanced Study, and by an operating grant from NSERC Canada. The author is an NSERC University Research Fellow.

<sup>2</sup>The author wishes to thank Haim Gaifman for suggestions which have greatly improved the exposition of these results, and for suggesting Lemma 1, which led to an improvement of the main result.

known to be quite complex sets (Jones and Selman [4], Fagin [1]). If  $\psi$  is restricted to be existential, though, then if  $\psi$  has finite models,  $\mu_n(\psi) \rightarrow 1$ , so that  $\mu_n(\phi|\psi)$  is asymptotic to  $\mu_n(\phi)$  in this case. Fagin asked what happens in the dual case when  $\psi$  is universal. The answer is that the asymptotics are the same as in the general case.

**MAIN THEOREM.** *Given  $\phi, \psi$  such that  $\mu_n(\psi) \geq 0$  for infinitely many  $n$ 's there exist universal sentences  $\phi^*, \psi^*$  such that  $\mu_n(\psi^*) > 0$  for all  $n$  and  $(\mu_n(\phi|\psi) - \mu_n(\phi^*|\psi^*)) \rightarrow 0$ , as  $n$  runs through numbers for which  $\mu_n(\psi) > 0$ .  $\square$*

First we show that we can assume, with no loss of generality, that our given  $\psi$  is such that  $\mu_n(\psi) > 0$  for all  $n$ . This lemma is due to Haim Gaifman.

**LEMMA 1.** *Given  $\phi, \psi$  such that  $\mu_n(\psi) > 0$  for infinitely many  $n$ 's, there exist sentences  $\hat{\phi}, \tilde{\psi}$  such that  $\mu_n(\tilde{\psi}) > 0$  for all  $n$  and  $(\mu_n(\phi|\psi) - \mu_n(\hat{\phi}|\tilde{\psi})) \rightarrow 0$ , as  $n$  runs through the numbers for which  $\mu_n(\psi) > 0$ .*

**PROOF.** Let  $\phi, \psi$  be in the language  $L_0$ . Let  $L_1$  be in the language obtained by adjoining new unary predicate,  $P$ . Put  $\gamma = \wedge(\forall \bar{x}_i R_i(\bar{x}_i))$ , where  $R_i$  ranges in the conjunction over all predicates of  $L_1$ . Evidently for every  $n$ ,  $\beta_n^{L_1}(\gamma) = 1$ . If  $\sigma$  is any sentence of  $L_0$ , we have  $\beta_n^{L_1}(\sigma) = 2^n \cdot \beta_n^{L_0}(\sigma)$ , because  $P$  has  $2^n$  different interpretations in models of size  $n$ . Consequently for  $\sigma$  in  $L_0$ :

$$\begin{aligned} 1 &\leq \beta_n^{L_1}(\sigma \vee \gamma), \\ 2^n \cdot \beta_n^{L_0}(\sigma) &\leq \beta_n^{L_1}(\sigma \vee \gamma) \leq 2^n \cdot \beta_n^{L_0}(\sigma) + 1, \\ 2^n \cdot \beta_n^{L_0}(\sigma) - 1 &\leq \beta_n^{L_1}(\sigma \wedge \neg\gamma). \end{aligned}$$

Now put  $\hat{\phi} = \phi \wedge \neg\gamma$ ,  $\tilde{\psi} = \psi \vee \gamma$ . Then  $\hat{\phi} \wedge \tilde{\psi}$  is equivalent to  $\phi \wedge \psi \wedge \neg\gamma$ . The first two inequalities with  $\sigma = \psi$  and the last inequality with  $\sigma = \phi \wedge \psi$  imply the desired properties of  $\hat{\phi}$  and  $\tilde{\psi}$ .  $\square$

By noting in the proof of Lemma 1 that  $\mu_n(\phi) = 0$  implies  $\mu_n(\hat{\phi}) = 0$ , we get the following result.

**COROLLARY 2.** *For any first order sentence  $\phi$  there are sentences  $\hat{\phi}$  and  $\tilde{\phi}$  such that  $\mu_n(\hat{\phi}|\tilde{\phi}) \rightarrow 1$ , as  $n$  runs through the numbers for which  $\mu_n(\phi) > 0$ ;  $\mu_n(\hat{\phi}|\tilde{\phi}) \rightarrow 0$ , as  $n$  runs through the numbers for which  $\mu_n(\phi) = 0$ .  $\square$*

**DEFINITION 3.** Let  $T$  be a theory in the language  $L$ ,  $T'$ —a theory in  $L'$ , where  $L \subseteq L'$ . Say that  $T'$  satisfies (\*) with respect to  $T$  if the following holds:

(\*) There are sequences  $k_n, l_n, n = 1, 2, \dots$ , such that  $k_n > 0, l_n \cdot \beta_n^L/k_n \rightarrow 0$  and, for every  $L$ -model,  $M$ , of size  $n$ : If  $M \models T$  then  $M$  has  $k_n, L'$ -expansions satisfying  $T'$ , and if  $M \not\models T$  then  $M$  has no more than  $l_n L'$ -expansions satisfying  $T'$ .  $\square$

**LEMMA 4.** *Assume that  $T'$  satisfies (\*) with respect to  $T$ . Let  $\phi, \psi \in L$  and let  $\mu_n(T, \psi) > 0$  for all  $n$ . Then*

$$(\mu_n(\phi|T, \psi) - \mu_n(\phi|T', \psi)) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

PROOF. By (\*) we have

$$\begin{aligned}\beta_n^{L'}(T', \phi, \psi) &= k_n \cdot \beta_n^L(T, \phi, \psi) + r_n, \\ \beta_n^{L'}(T', \psi) &= k_n \cdot \beta_n^L(T, \psi) + s_n\end{aligned}$$

where  $r_n$  and  $s_n$  are the numbers of  $L'$ -models not satisfying  $T$ , which satisfy, respectively,  $T' \cup \{\phi, \psi\}$  and  $T' \cup \{\psi\}$ . Evidently  $r_n, s_n \leq l_n \cdot \beta_n^L$ . Hence  $r_n/k_n$  and  $s_n/k_n$  tend to 0.  $\beta_n^L(T, \psi) \geq 1$ , hence

$$(\beta_n^{L'}(T', \phi, \psi)/\beta_n^{L'}(T', \psi) - \beta_n^L(T, \phi, \psi)/\beta_n^L(T, \psi)) \rightarrow 0. \quad \square$$

We are ready to begin the proof of the theorem. Let  $\phi, \psi$  be sentences in  $L_0$ . Assume with no loss of generality that  $\mu_n(\psi) > 0$  for all  $n$ . We will first introduce Skolem functions for  $\phi$  and  $\psi$  with Skolem theory  $T$  in such a way that

(1) The Skolem functions are uniquely defined, so that  $\mu_n(\phi|T, \psi) = \mu_n(\phi|\psi)$ .

(2) The only existential quantifiers in  $T$  occur in the axioms that say that the Skolem functions are always defined. (The Skolem functions are, of course, represented by relations, since we are not allowing function symbols.)

Then we will find a finite universal theory  $T'$  which has property (\*) with respect to  $T$ . If  $\phi^u, \psi^u$  are the universal equivalents of  $\phi, \psi$  in the Skolem theory  $T$ , then  $\phi^* = \phi^u$  and  $\psi^* = \psi^u \wedge \wedge T'$  will satisfy the conclusion of the theorem.

Extend  $L_0$  to  $L_1$  by adding a binary relation  $<$  and let  $T_1$  consist of the universal sentences asserting that  $<$  is a total ordering. Each  $L_0$ -model has exactly  $n!$   $L_1$ -expansions satisfying  $T_1$ . Hence  $\mu_n(\phi|\psi) = \mu_n(\phi|T_1, \psi)$ .

Rewrite  $\phi$  and  $\psi$  in prenex normal form and consider Skolem functions for the subformulas such that, for every  $\sigma(\vec{x}, y)$ , the Skolem function picks the  $<$ -first  $y$  such that  $\sigma(\vec{x}, y)$ , or, if no such  $y$  exists, the  $<$ -first  $y$ . Add new predicates to denote such functions. The sentences asserting that  $R$  is a graph of a function are

$$\forall \vec{x} \forall y_1, y_2 [R(\vec{x}, y_1) \wedge R(\vec{x}, y_2) \rightarrow y_1 = y_2], \quad \forall \vec{x} \exists y R(\vec{x}, y).$$

The axioms expressing the Skolem-function properties are

$$\begin{aligned}\forall \vec{x}, y [R(\vec{x}, y) \rightarrow \forall v [v < y \rightarrow \neg \sigma(\vec{x}, v)]], \\ \forall \vec{x}, y [R(\vec{x}, y) \rightarrow (\sigma(\vec{x}, y) \vee \forall v (\neg \sigma(\vec{x}, v) \wedge y \leq v))].\end{aligned}$$

If in our theory  $\sigma$  is equivalent both to universal and existential sentences, these last two axioms can be put in universal form. Moreover the four axioms imply

$$\exists y \sigma(\vec{x}, y) \leftrightarrow \forall y [R(\vec{x}, y) \rightarrow \sigma(\vec{x}, y)].$$

Hence if  $\sigma$  is equivalent to both forms, the same is true, given the axioms for  $\exists y \sigma(\vec{x}, y)$ . Since  $\forall y \sigma(\vec{x}, y)$  is equivalent to  $\neg \exists y \neg \sigma(\vec{x}, y)$ , we can, by adding also Skolem functions for  $\neg \sigma$ , express  $\forall y \sigma(\vec{x}, y)$  in both forms.

Working from the inmost quantifiers in  $\phi$  and  $\psi$  outwards, and adding at each stage new predicates for Skolem functions along with their corresponding axioms, we get an extension  $L \supseteq L_1$  and a theory  $T \supseteq T_1$  such that all axioms of  $T$  are either universal or of the form  $\forall \vec{x} \exists y R(\vec{x}, y)$ , and, in  $T$ ,  $\phi$  and  $\psi$  are equivalent to universal formulas, say  $\phi^u$  and  $\psi^u$ . Moreover, every  $L_1$ -model satisfying  $T_1$  has a

unique  $L$ -expansion satisfying  $T$ , because the Skolem functions are uniquely determined by the ordering. Hence

$$\mu_n(\phi|T_1, \psi) = \mu_n(\phi|T, \psi) = \mu_n(\phi^u|T, \psi^u).$$

Note also that  $\mu_n(T, \psi^u) > 0$ .

The last step is to find a universal theory  $T'$  which has property (\*) with respect to  $T$ . Then we shall have

$$\mu_n(\phi^u|T, \psi^u) - \mu_n(\phi^u|T', \psi^u) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

By putting  $\phi^* = \phi^u$  and taking  $\psi^*$  as the universal sentence equivalent to the conjunction of the axioms of  $T'$  with  $\psi^u$  we shall get the desired result.

Let  $R_1, \dots, R_m$  be all the Skolem-predicates that we have added, let  $R_i(\bar{x}_i, y_i)$  be their corresponding atomic formulas, where the strings  $\bar{x}_i, y_i, i = 1, \dots, m$ , consist of distinct variables and have no variables in common. Put  $\bar{x} = \bar{x}_1, \dots, \bar{x}_m, \bar{y} = y_1, \dots, y_m$  and let  $p$  be the length of  $\bar{x}$ . Fix some sufficiently large  $r$ , to be presently specified, and let  $S$  be a new  $(p + m + r)$ -ary predicate. Get  $T'$  from  $T$  by replacing all the axioms  $\forall \bar{x}_i \exists y_i R_i(\bar{x}_i, y_i)$  with the single axiom

$$\forall \bar{x} \forall \bar{y} \forall \bar{w} \left[ S(\bar{x}, \bar{y}, \bar{w}) \rightarrow \bigwedge_i R_i(\bar{x}_i, y_i) \right].$$

Put  $\sigma(\bar{x}, \bar{y}) = \bigwedge_i R_i(\bar{x}_i, y_i)$ . If  $M$  is an  $L$ -model which has an  $L'$ -expansion satisfying  $T'$ , then, given  $\bar{x}$ , there is at most one  $\bar{y}$  such that  $\sigma(\bar{x}, \bar{y})$ . If  $M \models T$  then there is always such a  $\bar{y}$ ; if  $M \not\models T$  there is at least one  $\bar{x}$  without such a  $\bar{y}$ . For fixed  $\bar{x}, \bar{y}$  if  $M \models \sigma(\bar{x}, \bar{y})$  then we have  $2^r$  ways of defining  $S(\bar{x}, \bar{y}, \dots)$ ; otherwise we have only one way (i.e.,  $S(\bar{x}, \bar{y}, \bar{w})$  must fail for all  $\bar{w}$ ). Put  $r_1 = n^r, p_1 = n^p$ . Then the number of  $L'$ -expansions satisfying  $T'$  is  $(2^{r_1})^{p_1}$  if  $M \models T$ ; otherwise it is no more than  $(2^{r_1})^{(p_1-1)}$ .

The first number is  $2^{(n^p \cdot n^r)}$ . The second is  $2^{(n^p-1)n^r}$ , and its ratio to the first is  $2^{-n^r}$ . The number of  $L$ -models with domain  $\{0, \dots, n-1\}$  is bounded by  $2^{n^c}$  for some constant  $c$ . Choosing  $r > c$  we get the property (\*) for  $T'$ .

### REFERENCES

1. Ronald Fagin, *Generalized first order spectra and polynomial-time recognizable sets*, Complexity of Computation (R. M. Karp, ed.), SIAM-AMS Proc., vol. 7, Amer. Math. Soc., Providence R. I., 1974.
2. \_\_\_\_\_, *Probabilities on finite models*, J. Symbolic Logic **41** (1976), 50-58.
3. Y. V. Glebskii, D. I. Kogan, M. I. Liogon'kii, and V. A. Talanov, *Range and degree of realizability of formulas in the restricted predicate calculus*, Kybernetika (Prague) **5** (1969), 142-154.
4. N. D. Jones and A. L. Selman, *Turing machines and the spectra of first order formulas with equality*, Proc. 4th ACM Sympos. on Theory of Computing, 1972, pp. 157-167.

DEPARTMENT OF MATHEMATICS AND STATISTICS, QUEEN'S UNIVERSITY, KINGSTON, CANADA K7L 3N6