2

A SIMPLE PROOF THAT CONNECTIVITY OF FINITE GRAPHS IS NOT FIRST-ORDER DEFINABLE

Haim Gaifman

Institute of Mathematics and Computer Science

Hebrew University of Jerusalem

Jerusalem, Israel

Moshe Y. Vardi†

Center for Study of Language and Information

Stanford University, California

Bull. EATCS 26 (1985), 43-45

† Research supported by a Weizmann Post-Doctoral Fellowship and AFOSR Grant 80-0212.

1. Introduction

First-order properties of finite relational structures are of considerable interest both in mathematical logic (e.g., [BH, F1, F2, Ga, Ha]) and in computer science (e.g., AU, CH, Im]). Perhaps one of the simplest properties that is not first-order is connectivity of graphs. Several proofs to that effect appeared in the literature. The earliest proof appeared in [F1][1], and it uses Erenfeucht-Fraisse games, as do the proofs in [CH, Im]. In [BH] it is claimed that they can show that connectivity is not first-order by an ultraproduct argument. The proof in [Ha] uses the method of semisets, and the proof in [AU] uses quantifier elimination. Finally, the claim follows immediately from a general characterization of first-order properties in [Ga]; a characterization that is proved by quantifier elimination. (Observe that a straightforward compactness argument shows that connectivity for arbitrary graphs is not first-order, and it is only the finite case that poses some difficulty.)

We feel that connectivity is an important enough property to deserve having a simple and direct proof to the effect that it is not first-order. The only tools we use are the compactness and Lowenheim-Skolem theorems. We shall prove the claim for directed graphs, and it can be easily modified to the case of undirected graphs.

2. The Main Result.

The language L we use has one extralogical symbol R of arity two. A structure A for L is a directed graph $\langle D,R\rangle$, where D is a nonempty set of nodes, and $R \subseteq D^2$ is a set of edges. A property π of graphs is *definable* if there is a first-order sentences σ such that for all graphs $A = \langle D,R\rangle$, A has the property π if and only if $A \models \sigma$. π is *finitely definable* if there is a first-order sentences σ such that for all finite graphs $A = \langle D,R\rangle$, A has the property π if and only if $A \models \sigma$.

---

[1] In fact, it is shown there that connectivity is not expressible even in monadic second-order existential logic. See also [dR]).

A graph $A=\langle A,R\rangle$ is *connected* if for all nodes x,y in D there is a sequences of nodes $x_1,\ldots,x_n$, $n>1$, such that $x=x_1$, $y=x_n$, and $(x_i,x_{i+1})\in R$ for $0<i<n$.

Theorem 1. Connectivity is not definable.

Proof. Let $\varphi_n(x,y)$ be a formula saying that there is a path of length n from x to y. We define the φ_n's by induction: $\varphi_0(x,y)$ is the formula $x=y$, and $\varphi_{n+1}(x,y)$ is the formula $\exists z(R(x,z)\&\varphi_n(z,y))$.

Assume now that connectivity is definable by a sentences σ. We now expand the language L by two constants c_1 and c_2. Let

$$S=\{\sigma\}\cup\{\neg\varphi_k(c_1,c_2):0\leq k<\omega\}.$$

It is easy to see that every finite subset of S is satisfiable, but S is not satisfiable - contradiction. □

The above proof uses the Compactness Theorem, and therefore does not carry over to finite definability.

Theorem 2. Connectivity is not finitely definable.

Proof. Let A_n be the finite structure with nodes $\{1,\ldots,n\}$ and edges $\{\langle i,i+1\rangle:1\leq i\leq n-1\}\cup\{\langle n,1\rangle\}$. Let B_n be the finite structure with nodes $\{1,\ldots,2n\}$ and edges $\{\langle i,i+1\rangle:1\leq i\leq 2n-1$ and $i\neq n\}\cup\{\langle n,1\rangle,\langle 2n,n\rangle\}$. That is, A_n is a cycle of length n, so it is connected, and B_n is two cycles of length n, so it is disconnected. Let S_1 be the theory

$$\{\sigma:\sigma \text{ holds in all but finitely many } A_n\text{'s}\},$$

and let S_2 be the theory

$$\{\sigma:\sigma \text{ holds in all but finitely many } B_n\text{'s}\}.$$

We now expand the language L with countably many constants $c_0,\ldots,c_n,\ldots$. Let T be the theory $\{\neg\varphi_k(c_i,c_j):0\leq i,j,k\leq\omega$ and $i\neq j\}$ (the φ_k's are defined in the proof of Theorem 1). Now take T_1 to be $S_1\cup T$, and we take T_2 to be $S_2\cup T$.

We argue by compactness that T_1 is satisfiable. Let Σ be a finite subset of S_1, and let Δ be a finite subset of T. Each sentence in Σ holds in all but finitely many A_n's. It follows that there is some n_0 so that for all $n \geq n_0$, $A_n \models \Sigma$. Let now k be the number of constants occurring in sentences of Δ, and let m be the maximal one such that $\varphi_m(c_i,c_j) \in \Delta$ for some constants c_i and c_j. Let $n > \max(n_0, k(m+2))$. It is easy to see that we can interpret the k constants in Δ by elements from $\{1, \ldots, n\}$, so that for any pair of constants c_i and c_j that occur in sentences of Δ, the shortest path in A_n between the elements that interpret these constants is of length $n+1$. It follows that $A_n \models \Sigma \cup \Delta$. Thus, T_1 is satisfiable. By an identical argument T_2 is also satisfiable.

Let A and B be countable models of T_1 and T_2, respectively. These models exist by Lowenheim-Skolem Theorem, since the expanded language $L(c_0, \ldots)$ is countable. Consider the model A. Clearly, in A every element has a unique successor and a unique predecessor, because this is true in all the A_n's. Thus, a connected component of A is either a cycle or a line[2]. The formula $\tau_n(x)$: $\exists y(x \neq y \& \varphi_{n-1}(x,y) \& R(y,x))$ says that there is a cycle of length n going through x. For all $n>0$, $\forall x(\neg\tau_n(x))$ is in S_1, so A can not have cycles but only lines. Finally, because for each pair of constants c_i and c_j and for all n, we have $A \models \neg\varphi_n(c_i,c_j)$, no two constants can be interpreted as elements on the same line in A. It follows that A has countably many lines. The same argument applies to the model B. Thus, the reductions of A and B to the language L (i.e., ignoring the constants) are isomorphic and elementarily equivalent.

Suppose now that connectivity for finite directed graphs is a finitely definable. Then, there is a sentence σ such that for all n, $A_n \models \sigma$ and $B_n \models \neg\sigma$. It follows that $\sigma \in T_1$ and $\neg\sigma \in T_2$, so $A \models \sigma$ and $B \models \neg\sigma$ - a contradiction. □

---

[2] A line is a directed graph isomorphic to $\{\langle i, i+1 \rangle : -\omega < i < \omega\}$.

References

[AU] Aho, A.V., Ullman, J.D.: Universality of data retrieval languages. Proc. 6th ACM Symp. on Principles of Programming Languages, 1979, 110-117.

[BH] Blass, A., Harary, F.: Properties of almost all graphs and complexes. J. Graph Theory 3(1979), 225-240.

[CH] Chandra, A.K., Harel, D.: Structure and complexity of relational queries. Proc. 21st IEEE Symp. on Foundations of Computer Science, 1980, 333-347.

[dR] de Rougemont, M.: Uniform definability on finite structures with successor. Proc. 16th ACM Symp. on Theory of Computing, Washington DC, April 1984, pp. 409-417. To appear in J. of Computer and System Sciences.

[F1] Fagin, R.: Monadic generalized spectra. Zeitschr. f. math. Logik und Grundlagen d. Math. 21(1975), 89-96.

[F2] Fagin, R.: Probabilities on finite models. J. Symbolic Logic 41(1976), 50-58.

[Ga] Gaifman, H.: On local and non-local properties. Proc. Herbrand Logic Colloq., Marseille, 1981.

[Ha] Hajek, P.: On logics of discovery. Proc. Symp. on Mathematical Foundations of Computer Science, Lecture Notes in Computer Science - Vol. 32, Springer-Verlag, 1975, pp. 15-30.

[Im] Immerman, N.: Number of quantifiers is better than number of tape cells. J. Computer and Systems Sciences 22(1981), 384-406.