KYBERNETIKA — VOLUME 8 (1972), NUMBER 4

A Planar Test of Linguistic Projectivity

LADISLAV NEBESKÝ

In a mathematical model of the dependency structure of the sentence, the condition of projectivity occupies a significant place. In the present paper it is shown that the question whether this model is projective can be reshaped as the question whether a certain finite graph is planar.

One of the important notions of mathematical models of language is what we call here an L-tree.

Definition 1. An ordered quadruple $L = (V_0, E_0, r, \leqq_L)$ will be called an L-tree, if

(i) the ordered couple (V_0, E_0) is a tree,
(ii) $r \in V_0$,
(iii) $\leqq_L$ is a complete ordering of the set V_0.

An L-tree is non-trivial if it has at least one edge. (For the notions of graph theory see, for example, Busacker and Saaty [1].)

An L-tree $(V_0, E_0, r, \leqq_L)$ models the dependency structure of a sentence in natural languages. The tree (V_0, E_0) together with the vertex r (called the root) describes its syntactic relations, the ordering $\leqq_L$ corresponds to its word order (for the position of trees in mathematical models of the dependency structure of sentences, see Novák [5]). L-trees of sentences often fulfil the condition of projectivity:

Definition 2. An L-tree $L = (V_0, E_0, r, \leqq_L)$ will be called projective if for any $u, v, w \in V_0$ such that $(u \,\&\, w) \in E_0$ and either $u <_L v <_L w$ or $w <_L v <_L u$, it holds that if the vertex u lies on the chain joining the vertices r and w, then the vertex u lies on the chain joining the vertices r and v.

L-trees and projective L-trees corresponds to the simple strings and simple projective strings in Chapter VI of Marcus's book [2] (which also contains a bibliography of the concept of projectivity). For some other properties of linguistic projectivity see also [3], Chapter IV; cf. [4].

There is a geometrical criterion for an L-tree $L = (V_0, E_0, r, \leqq_L)$ to be projective. We map L in a plane with orthogonal coordinates x, y such that there is a point $P_v = (x_v, y_v)$, where $y_v > 0$, to correspond to each vertex $v \in V_0$, and an abscissa P_uP_w corresponds to each edge $(u \,\&\, w) \in E_0$, such that for any vertices $s, t \in V_0$ it holds that

(i) $x_s < x_t$ if and only if $s <_L t$, and

(ii) $y_s > y_t$ if and only if $d(r, s) < d(r, t)$, where $d(p, q)$ is the distance between vertices p and q.

Let us denote $Q_v = (x_v, 0)$, for any $v \in V_0$. L is projective if and only if for any edges $(s \,\&\, t)$, $(u \,\&\, w) \in E_0$ and any vertex $v \in V_0$ it holds that neither abscissae P_sP_t and P_uP_w nor abscissae P_uP_w and P_vQ_v cut across each other. Cf. Marcus [3], pp. 237–240.

The criterion of projectivity mentioned here is of great practical use, but from the point of view of graph theory it is not very transparent. Nevertheless it does indicate the possibility that there is some relationship between linguistic projectivity and planarity in graph theory. It is our aim to find this relationship.

Definition 3. Let $L = (V_0, E_0, r, \leqq_L)$ be a non-trivial L-tree such that

$$V_0 = \{v_1, \ldots, v_n\}, \quad v_1 <_L v_2 <_L \ldots <_L v_n,$$

where $n \geqq 2$. Let us assume that C be a simple circuit of length $n + 1$ with vertices different from the vertices of V_0 and with the edges $(w_0 \,\&\, w_1), \ldots, (w_n \,\&\, w_0)$. Then by G_L we shall denote the undirected graph (V, E) such that

$$V = V_0 \cup \{w_0, \ldots, w_n\},$$

$$E = E_0 \cup \{(w_0 \,\&\, w_1), (w_1 \,\&\, w_2), \ldots, (w_n \,\&\, w_0)\} \cup \{(r \,\&\, w_0)\} \cup$$

$$\cup \{(v_1 \,\&\, w_1), \ldots, (v_n \,\&\, w_n)\}.$$

The graph G_L has $2n + 1$ vertices and $3n + 1$ edges. The following theorem gives a planar test of linguistic projectivity:

Theorem. *Let L be a non-trivial L-tree. A necessary and sufficient condition for L to be projective is that the graph G_L be planar.*

Proof. We shall utilize the notions, symbols and assumptions from Definition 3. We put $r = v_i$, $1 \leqq i \leqq n$.

Necessity. First we shall give the mentioned mapping of L in a plane. By G_0 we shall denote the graph which we obtain from the tree (V_0, E_0) by adding the vertices (points) $Q_{v_1}, \ldots, Q_{v_n}$ and the edges (abscissae) $Q_{v_1}Q_{v_2}, \ldots, Q_{v_{n-1}}Q_{v_n}$, $P_{v_1}Q_{v_1}, \ldots$ $\ldots, P_{v_n}Q_{v_n}$. As L is projective, then G_0 is planar. The planarity will be preserved when we complete G_0 on the graph G_L by the suitable addition of a vertex (point) Q and edges (simple open curves) joining Q with Q_{v_1}, Q with Q_{v_n} and Q with P_{v_i}.

Sufficiency. Let us assume that G_L is planar, but L is not projective. This means that there exist $s, t, u \in V_0$ such that (a) $(s \& t) \in E_0$, (b) either $s <_L u <_L t$ or $t <_L u <_L s$, (c) s lies, in the tree (V_0, E_0), on the simple chain joining r and t, and (d) s does not lie, in (V_0, E_0), on the simple chain joining r and u. Without loss of generality let us assume that $s <_L u <_L t$. There are j, k such that $s = v_j$ and $t = v_k$.

(A) Let $u = r$. By D we shall denote the simple chain joining r and t in (V_0, E_0). We shall consider the graph G_r created by the simple circuit C, the simple chain D and the set of edges

$$\{(r \& w_0), (s \& w_j), (r \& w_i), (t \& w_k)\}.$$

In the graph G_r exactly six vertices have degree 3, namely r, s, w_0, w_j, w_i and w_k. It is easy to see that the graph G_r is isomorphic to within vertices of degree 2 with a type 2 Kuratowski graph (see [1], Chapter 4). Thus G_L is not planar, which is a contradiction.

(B) Let $u \neq r$. From (A) it follows that either $r <_L s$ or $t <_L r$. This means that there exist $p, q \in V_0$ such that (a) $(p \& q) \in E_0$, (b) $s <_L q <_L t$, (c) either $p <_L s$ or $t <_L p$, (d) s does not lie, in (V_0, E_0), on the simple chain joining r and q, and (e) p lies on this chain. There are g, h such that $p = v_g$, $q = v_h$, Evidently either $1 \leqq g < j < h < k \leqq n$ or $1 \leqq j < h < k < g \leqq n$. By z and $\bar{z}$ we shall denote such vertices that $\{z, \bar{z}\} = \{p, q\}$ and $d(t, z) = d(t, \bar{z}) + 1$, where d is the distance in (V_0, E_0). It is obvious that the vertices s and $\bar{z}$ lie on the simple chain F joining the vertices t and z in the tree $(V_0. E_0)$. We shall consider the graph $\bar{G}$ created by the simple circuit C, the simple chain F and the set of edges

$$\{(p \& w_g), (s \& w_j), (q \& w_h), (t \& w_k)\}.$$

In the graph $\bar{G}$ exactly six vertices have degree 3, namely $s, \bar{z}, w_g, w_j, w_h$ and w_k. It is easy to see that $\bar{G}$ is isomorphic to within vertices of degree 2 with a type 2 Kuratowski graph, which is a contradiction. The proof is completed.

(Received November 8, 1971.)

REFERENCES

[1] R. G. Busacker and T. L. Saaty: Finite Graphs and Networks: An Introduction with Applications. McGraw-Hill Book Company, New York 1965.
[2] S. Marcus: Algebraic Linguistics; Analytical Models. Academic Press, New York 1967.
[3] L. Nebeský: Algebraic Properties of Trees. Karlova universita, Praha 1969.
[4] L. Nebeský: Left-right double trees. Discrete Mathematics *1* (1971), 73–81.
[5] P. Novák: Postscript. [3], 83–95.

Rovinný test lingvistické projektivity

LADISLAV NEBESKÝ

V matematickém modelu závislostní struktury věty zaujímá významné místo podmínka projektivity. V článku je ukázáno, že otázka, zda tento model je projektivní, může být převedena na otázku, zda jistý konečný graf je rovinný.

Ladislav Nebeský, prom. mat., filosofická fakulta Karlovy university (Faculty of Philosophy, Charles university), Krasnoarmějců 2, Praha 1.