ARBORESCENȚE INDEXATE

DE

C. JALOBEANU

Se stabilește că orice arborescență (X, Γ) este izomorfă cu o N-mulțime canonică, abstracție făcînd de o permutare. Se introduce suma a două arborescențe și se dă o teoremă de descompunere a unei arborescențe în arborescențe simple, bazată pe observația că o arborescență este determinată pînă la o permutare de perechile strict terminale.

INTRODUCERE

În lucrare se întreprinde un studiu al arborescențelor, util în cercetările privind generarea cuvintelor într-un limbaj formal. S. Ginsburg [1] asociază un arbore de generare fiecărui cuvînt dintr-un limbaj acontextual generat conform unei gramatici date. Vom folosi în cele ce urmează notația din [1].

Se dă o nouă definiție arborescenței, noțiune utilizată de C. Pair [2]. Rezultatele centrale ale lucrării sînt : teorema de reprezentare 3.2 și teorema de descompunere 3.4. Pentru a ajunge cît mai firesc la aceste teoreme am studiat în paragraful 1 N-mulțimile, structuri algebrice despre care am demonstrat că sînt arborescențe. Rezultatul este conținut în teorema 3.2. Aplicația de indexare, utilizată și de T. Rus în [3], permite obținerea unor metode efective de calcul a sumei a două arborescențe și de determinare a componentelor simple ale unei arborescențe.

Spre deosebire de operațiile booleene tratate de F. Harary și G. W. Wilcox în [4], operația introdusă este internă în mulțimea arborescențelor.

§ 1. N-MULȚIMI

Considerăm alfabetul N al numerelor naturale și operația de juxtapunere a simbolurilor din N, notată cu $*$. Mulțimea cuvintelor formate cu elementele alfabetului N, prin operația de juxtapunere este un semigrup liber în raport cu această operație. Această mulțime va fi notată cu N^*.

ST. CERC. MAT., TOM. 24, NR. 10, P. 1561–1566, BUCUREȘTI, 1972

DEFINIȚIA 1.1. *Dacă $s, t \in N^*$ sînt cuvinte peste alfabetul N, atunci $s \subset t$ dacă există $q \in N^*$ astfel ca $t = s*q$.*

Ordonarea naturală în N induce o ordonare în N^* prin aplicația $h: N^* \to N$ care ataşază fiecărui cuvînt din N^* lungimea sa.

DEFINIȚIA 1.2. *Cuvintele $p_1, p_2 \in N^*$ sînt în relația $p_1 \alpha p_2$ dacă şi numai dacă $h(p_1) < h(p_2)$ sau $h(p_1) = h(p_2)$ şi există $m, q_1, q_2 \in N^*$, eventual vide, astfel ca $p_1 = m*i_1*q_1$, $p_2 = m*i_2*q_2$, cu $i_1, i_2 \in N$ şi $i_1 < i_2$.*

Relația α introdusă este o ordonare totală în N^* şi operația $*$ păstrează această ordonare.

DEFINIȚIA 1.3. *O submulțime P din N^* se numeşte N-mulțime, dacă are următoarele proprietăți:*

i) $P \cap N = \{1\}$.

ii) Dacă $m \in P$, $q \subset m$ atunci $q \in P$; dacă $m = q*i$, $i \in N$, $i > 1$, atunci $q*(i-1) \in P$.

DEFINIȚIA 1.4. *N-mulțimea P se zice canonică dacă oricare ar fi $q \in P$ şi $p_1, p_2 \in N$ are loc proprietatea următoare:*

— dacă $q*p_1 \in P$ şi $q*p_2 \in P$, $q*p_1*1 \in P$ şi $q*p_2*1 \notin P$,

atunci $p_2 < p_1$.

TEOREMA 1.1. *Dacă P şi Q sînt N-mulțimi, atunci $P \cap Q$ şi $P \cup Q$ sînt N-mulțimi.*

Demonstrație. Cum singurul cuvînt din P ca şi din Q, de lungime unu este *1*, rezultă axioma i) pentru $P \cap Q$. Fie $m = q*p$ în $Q \cap P$, rezultă atunci imediat $q \in P$ şi $q \in Q$, iar dacă $p \in N$ şi $p > 1$ atunci $q*(p-1)$ trebuie să aparțină şi lui P şi lui Q din moment ce acestea sînt N-mulțimi, prin urmare $q*(p-1) \in P \cap Q$.
Cu totul analog se arată că în $P \cup Q$ sînt verificate condițiile i) şi ii).

§ 2. ARBORESCENȚE

Fie X o mulțime finită, $X \neq \emptyset$ şi $\Gamma \subset X \times X$, o relație binară în X. Notăm cu $\Gamma(x) = \{y \in X \mid (x, y) \in \Gamma\}$, $\Gamma^{-1}(x) = \{y \in X \mid (y, x) \in \Gamma\}$ şi cu $|Y|$ cardinalul submulțimii $Y \subset X$.

DEFINIȚIA 2.1. *Numim arborescență o pereche (X, Γ) cu $\Gamma \in X \times X$, verificînd condițiile:*

(1) $(a, a) \notin \Gamma$, oricare ar fi $a \in X$,

(2) dacă $a, b, c \in X$ şi $(a, c) \in \Gamma$, $(b, c) \in \Gamma$, *atunci* $a = b$,

(3) *există un singur element* $0 \in X$ *cu proprietatea* $\Gamma^{-1}(0) = \emptyset$.

Observație. Condiția (3), care se mai poate scrie $X - \bigcup_{x \in X} \Gamma(x) = \{0\}$ (unde $\{0\}$ este mulțimea cu singurul element 0), ne asigură că graful (X, Γ) este conex.

Un element $a \in X$ se zice terminal, dacă $\Gamma(a) = \emptyset$. Un şir de elemente din X, $a_1, a_2, \ldots, a_n$ se numeşte drum de lungime $n-1$ de la a_1 la a_n dacă $(a_i, a_{i+1}) \in \Gamma$, oricare ar fi $i = 1, 2, \ldots, n-1$. Notăm atunci $a_n \in \Gamma^{n-1}(a_1)$.

Observație. Pentru orice $x \in X$ există un drum de lungime $l(x)$ unic, de la 0 la x.

Un număr n din N se zice nivelul elementului $x \in X$, dacă $n = l(x)$. l fiind o aplicație din X în N, vom nota cu $l^{-1}(n)$ mulțimea elementelor de pe nivelul n.

DEFINIȚIA 2.2. *O arborescență (X, Γ) se numeşte simplă dacă oricare ar fi un număr natural n există cel mult un element y în X astfel ca $l(y) = n$ şi $\Gamma(y) \neq \emptyset$.*

TEOREMA 2.1. *Arborescența (X, Γ) este simplă dacă şi numai dacă oricare ar fi $n \in N$ există $y \in X$ astfel încît $l^{-1}(n) = \Gamma(y)$.*

Demonstrație. Fie (X, Γ) o arborescență simplă şi fie $n \in N$. Dacă $l^{-1}(n) = \emptyset$, atunci orice element terminal cu proprietatea $\Gamma(y) = \emptyset$, verifică condiția.

Dacă $l^{-1}(n) \neq \emptyset$, fie $x \in l^{-1}(n)$, atunci prin definiție există cel mult un $y \in X$ pe nivelul $n-1$, astfel ca $\Gamma(y) \neq \emptyset$ şi deci $x \in \Gamma(y)$. Dar, $l^{-1}(n) \subseteq \Gamma(y)$ implică $l^{-1}(n) = \Gamma(y)$.
Invers, fie (X, Γ) o arborescență astfel ca pentru orice $n \in N$ să existe $y \in X$ cu $l^{-1}(n) = \Gamma(y)$. Evident $l^{-1}(n) \neq \emptyset$ implică $l(y) = n-1$ şi $\Gamma(y) \neq \emptyset$. Dacă presupunem că $z \in X$ este un alt element de pe nivelul $n-1$ pentru care $\Gamma(z) \neq \emptyset$, atunci $\Gamma(z) \subseteq l^{-1}(n)$ şi deci $\Gamma(z) \subseteq \Gamma(y)$. Dar, $x \in \Gamma(z)$ şi $x \in \Gamma(y)$ pe baza condiției (2) din definiția 2.1 conduce la $y = z$.

TEOREMA 2.2. *Orice N-mulțime este o arborescență.*

Demonstrație. Aleg $\Gamma = \{(p, q) \mid p, q \in P, q = p*i, i \in N\}$, unde am notat cu P, N-mulțimea considerată. Pentru Γ astfel ales se verifică condițiile din definiția arborescenței:

1. Dacă $(p, q) \in \Gamma$, atunci prin construcție $p \subset q$ şi deci $p \neq q$.

2. Fie p, q şi r din P astfel ca $(p, r) \in \Gamma$ şi $(q, r) \in \Gamma$. Atunci $r = p*i$ şi $r = q*j$ cu $i, j \in N$, r, p, q, fiind cuvinte în N^*, rezultă $p = q$ şi $i = j$.

3. Arătăm că *1* este singurul element ce verifică condiția $\Gamma^{-1}(1) = \emptyset$. Să presupunem că există $p \in \Gamma^{-1}(1)$. Atunci $1 \in \Gamma(p)$ şi deci $1 = p*i$ cu $i \in N$ ceea ce nu se poate. Să presupunem acum că există $m \neq 1$ astfel ca $\Gamma^{-1}(m) = \emptyset$, deci că $m \in X - \Gamma(X)$. Dar $m \neq 1$ înseamnă $h(m) > 1$ şi deci $m = r*j$ cu $r \in N^*$ şi $j \in N$, adică $m \in \Gamma(r)$, contradicție ce completează demonstrația teoremei.

§ 3. APLICAȚII DE ARBORESCENȚE

DEFINIȚIA 3.1. *$f: X \to Y$ se numeşte aplicație a arborescenței (X, Γ) în arborescența (Y, Γ') dacă pentru orice $x \in X$, $f(\Gamma(x)) \subset \Gamma'(f(x))$. Dacă aplicația $f: X \to Y$ este bijectivă o vom numi izomorfism de arborescențe.*

Evident, dacă $f: (X, \Gamma) \to (Y, \Gamma')$ este izomorfism de arborescențe, atunci $f(0) = 0'$ şi $f(\Gamma(x)) = \Gamma'(f(x))$ pentru orice $x \in X$ (unde $0 = X - \Gamma(X)$ şi $0' = Y - \Gamma'(Y)$).

DEFINIȚIA 3.2. *O arborescență* (X', Γ') *se numește subarborescență a arborescenței* (X, Γ), *dacă* $X' \subset X$ *și* Γ' *este restricția lui* Γ *la mulțimea* X'. *Dacă* (X', Γ') *este o subarborescență a lui* (X, Γ) *astfel încît și* $(X - X', \Gamma'')$ *este subarborescență a lui* (X, Γ), Γ'' *fiind restricția lui* Γ *la* $X - X'$, *atunci* (X', Γ') *se numește subarborescență plină a lui* (X, Γ).

TEOREMA 3.1. Arborescența (X', Γ') este subarborescență plină a arborescenței (X, Γ) dacă și numai dacă există $x \in X$ astfel încît $X' = \bigcup_{n=0}^{m} \Gamma^n(x)$, unde $m = \max l(X)$.

Demonstrație. Fie $x \in X$, verificînd condițiile teoremei, $\bigcup_{n=0}^{m} \Gamma^n(x)$ este o subarborescență și $l(x) = p$. Atunci

$$X - X' = \bigcup_{n=0}^{m} \Gamma^n(0) - \bigcup_{n=0}^{m} \Gamma^n(x) = \bigcup_{n=0}^{p-1} \Gamma^n(0) \cup \bigcup_{n=0}^{m} \Gamma^n(\Gamma^p(0) - (x)).$$

Rezultă că $(X - X', \Gamma'')$ este subarborescență și deci (X', Γ') este subarborescență plină.

Invers, dacă (X', Γ') este o subarborescență plină, $\Gamma' = \Gamma \cap (X' \times X')$ și $X' = \bigcup_{n=0}^{m'} \Gamma'(0')$, unde $0'$ este elementul definit în 2.1 (3). Cum, $0' \in X$, rezultă $\bigcup_{n=0}^{m'} \Gamma'^n(0') \subset \bigcup_{n=0}^{m} \Gamma^n(0')$. Are loc chiar egalitatea deoarece dacă am presupune că $y \in \bigcup_{n=0}^{m} \Gamma^n(0') - \bigcup_{n=0}^{m'} \Gamma'^n(0')$ ar rezulta pe de o parte că $y \in X - X'$ și pe de alta că $y \in \Gamma^j(0')$, în contradicție cu faptul că $(X - X', \Gamma'')$ este subarborescență. Deci $\bigcup_{n=0}^{m'} \Gamma'^n(0') = \bigcup_{n=0}^{m} \Gamma^n(0')$.

DEFINIȚIA 3.3. *Tripletul* $(X, \Gamma, <)$ *se numește arborescență parțial ordonată dacă* (X, Γ) *este o arborescență, iar* $(X, <)$ *este o mulțime parțial ordonată astfel încît pentru orice* $x \in X, \Gamma(x)$ *este o mulțime total ordonată în* $(X, <)$ *și* $x < y$ *implică* $x' < y'$, *pentru orice* $x' \in \Gamma(x)$ *și orice* $y' \in \Gamma(y)$.

DEFINIȚIA 3.4. *Dacă* $(X, \Gamma, <)$ *și* (X', Γ', α) *sînt două arborescențe parțial ordonate, aplicația* $f: (X, \Gamma) \to (X', \Gamma')$ *se numește izotonă dacă este aplicație de arborescențe și pentru orice* $x, y \in X$, $x < y$ *implică* $f(x) \alpha f(y)$.

DEFINIȚIA 3.5. *Dacă* $(X, \Gamma, <)$ *și* (X, Γ, α) *sînt două ordonări parțiale ale aceleiași arborescențe, un izomorfism izoton* $f: (X, \Gamma, <) \to (X, \Gamma, \alpha)$ *se numește permutare.*

Aplicația identică $i: (X, \Gamma) \to (X, \Gamma)$ este o permutare a lui $(X, \Gamma, <)$ în (X, Γ, α).

TEOREMA 3.2. *Orice arborescență* (X, Γ) *este izomorfă cu o* N*-mulțime canonică, abstracție făcînd de o permutare a lui* (X, Γ).

Demonstrație. Fie $(X, \Gamma, <)$ o ordonare în arborescența (X, Γ) astfel încît dacă $x, y \in \Gamma(z)$ și $\Gamma(x) \neq \emptyset$, iar $\Gamma(y) = \emptyset$, atunci $y < x$.

Fie $f: X \to N^*$ o aplicație definită de $f(0) = 1$ și dacă $f(x) = m$, atunci $f(\Gamma(x)) = (m*1, m*2, \ldots, m* |\Gamma(x)|)$, astfel ca pentru $x_1, x_2 \in \Gamma(x)$ și $x_1 < x_2$ să aibă loc $f(x_1) < f(x_2)$. Vom numi indexare această aplicație.

Se verifică imediat că $f(X)$ este o N-mulțime canonică. Într-adevăr, cum $\Gamma^{-1}(x) \neq \emptyset$, oricare ar fi $x \neq 0$, rezultă că pentru orice $m \in f(X)$, avem $h(m) > 1$, deci $f(X) \cap N = \{1\}$. Apoi, dacă $m \in f(X)$, atunci există $x \in X$ astfel că $f(x) = m$; pentru $h(m) > 1$, fie $y \in \Gamma^{-1}(x)$, atunci $m = f(y)*j$ cu $j \in N$ și $1 \leqslant j \leqslant |\Gamma(y)|$; deci $j > 1$ implică $f(y)*(j-1) \in f(X)$.

Fie acum $x_1, x_2 \in X$, dacă $f(x_1) = q*i$, $f(x_2) = q*j$ și $q*i*1 \notin f(X)$, $q*j*1 \in f(X)$, rezultă $\Gamma(x_1) = \emptyset$, $\Gamma(x_2) \neq \emptyset$ și $x_1 < x_2$; prin urmare $f(x_1) < f(x_2)$, adică $f(X)$ este N-mulțime canonică.

Notînd acum N-mulțimea canonică $f(X)$ cu P și folosind teorema 2.3 obținem că $f: X \to P$ este aplicație surjectivă și izotonă de arborescențe (orice N-mulțime este o arborescență parțial ordonată — în sensul definiției 3.2 — prin ordonarea din N^*). Pentru a arăta că f este izomorfism, este suficient deci să arătăm că oricare ar fi x, y din X, $x \neq y$ implică $f(x) \neq f(y)$. Sînt de deosebit două cazuri: 1) $l(x) \neq l(y)$ și atunci prin construcție $h(f(x)) \neq h(f(y))$ și cu atît mai mult $f(x) \neq f(y)$;

2) $l(x) = l(y)$ și atunci sau $x < y$ sau $y < x$, de unde prin construcție $f(x) < f(y)$ sau $f(y) < f(x)$ și deci $f(x) \neq f(y)$. Ordonarea introdusă în arborescența (X, Γ) face ca izomorfismul f să se compună eventual cu o permutare.

În cele ce urmează, pentru orice arborescență (X, Γ) vom nota cu f aplicația de indexare definită în demonstrația teoremei precedente.

DEFINIȚIA 3.6. *În arborescența* (X, Γ) *perechea* $(u, f(u))$ *cu* u *element terminal pentru care* $f(u) = f(\Gamma^{-1}(u)) * |\Gamma(\Gamma^{-1}(u))|$ *se numește pereche strict terminală.*

COROLAR 3.2.1. O arborescență (X, Γ) este determinată de perechile strict terminale. Arborescența (X, Γ) admite o singură pereche strict terminală dacă și numai dacă este arborescență simplă.

Demonstrație. Fie, mai întîi, arborescența (X, Γ) cu unica pereche strict terminală (u, m) ($u \in X$ și $m \in N^*$). Atunci $l^{-1}(l(u)) = \Gamma(\Gamma^{-1}(u))$ și cu atît mai mult $l^{-1}(l(u) - 1) = \Gamma(\Gamma^{-2}(u))$, sau punînd $l(u) = n$, $l^{-1}(n-1) = \Gamma(\Gamma^{-2}(u))$ și pas cu pas $l^{-1}(n-i) = \Gamma(\Gamma^{-i-1}(u))$ cu $i = 2, 3, \ldots, n-1$, pînă la $l^{-1}(1) = \Gamma(\Gamma^{-n}(u)) = \Gamma(0)$. Teorema 2.2 ne asigură că (X, Γ) este în acest caz simplă. Evident $h(m) = n$.

Fie acum (u, m) și (v, p) două perechi strict terminale în (X, Γ) și $n = \min(h(m), h(p))$; fie spre exemplu $n = h(m)$. Dacă (X, Γ) este simplă, atunci există $x \in X$ astfel ca $l^{-1}(n) = \Gamma(x)$. Dar $\Gamma^{-h(p)-n}(v)$ aparține lui $l^{-1}(n)$ și deci lui $\Gamma(x)$. Iar (u, m) fiind pereche strict terminală $m = f(x) * |\Gamma(x)|$ (căci $x = \Gamma^{-1}(u)$). Deoarece $y = \Gamma^{-h(p)-n}(v) \in \Gamma(x)$, rezultă $f(y) = f(x) * j$ cu $1 \leqslant j \leqslant |\Gamma(x)|$, deși $f(x)*j*1 \in f(X)$ și $f(x) * |\Gamma(x)| * 1 \notin f(X)$ (odată ce $p \in f(X)$ și u-terminal). Ultimele două relații contrazic faptul că $f(X)$ este N-mulțime canonică. Prin urmare ipoteza că o arborescență simplă ar putea conține mai mult decît o pereche strict terminală este falsă.

Prima afirmație a corolarului 3.2.1 rezultă din teorema 3.2 și din definiția N-mulțimii canonice.

TEOREMA 3.3. *Dacă* (X', Γ') *și* (X'', Γ'') *sînt două arborescențe, atunci există o arborescență* $(X, \Gamma) = (X', \Gamma') \oplus (X'', \Gamma'')$, *cea mai mică arborescență ce conține arborescențele date ca subarborescențe.*

Demonstrație. În baza teoremei 1.1 reuniunea N-mulțimilor canonice $f(X')$ și $f(X'')$ este tot o N-mulțime, dar nu neapărat canonică. Fie $f(X' \cup X'')$ N-mulțimea canonică ce se obține prin aplicarea teoremei 3.2 la arborescența $f(X') \cup f(X'')$, și fie $(X' \cup X'', \Gamma)$ arborescență izomorfă cu aceasta. Prin construcție rezultă imediat că (X', Γ') și (X'', Γ'') sînt subarborescențe ale arborescenței obținute, iar condiția de minimalitate rezultă din faptul că $X = X' \cup X''$ și X', X'' sînt mulțimi finite.

TEOREMA 3.4. *Orice arborescență* (X, Γ) *admite o descompunere în arborescențe simple.*

$$(X, \Gamma) = \overset{n}{\underset{i=1}{\oplus}} (X_i, \Gamma_i),$$ *n fiind numărul perechilor strict terminale.*

Demonstrație. Corolarul 3.2.1 ne asigură că fiecare pereche strict terminală determină o arborescență simplă, evident subarborescență a lui (X, Γ). Apoi, ordonarea din $f(X)$ fiind totală, perechile strict terminale se pot aranja într-un șir și nu rămîne decît să aplicăm teorema 3.3.

Primită la redacție la 19 iulie 1971

Institutul de calcul
Academia Republicii Socialiste România,
Filiala din Cluj

INDEXED ROOTED TREES

(ABSTRACT)

The paper states that any rooted tree (X, Γ) is isomorphic with a canonical N-set disregarding a permutation.

The sum of two rooted trees is introduced and a theorem is given of the decomposition of an rooted tree into simple rooted tree based on the remark that any rooted tree is determined up to a permutation, by the strictly terminal pairs.

BIBLIOGRAFIE

1. S. GINSBURG, *The Mathematical Theory of Context free Languages*, McGrow-Hill New York, 1966.
2. C. PAIR, *Etude de la notation de pile application à l'analyse syntaxique*, Public. de l'Univ. de Nancy, Collection de thèses, 1966.
3. T. RUS, *On the matricial representation of the trees*, Mathematica, **6**, *2*, 327—334 (1964).
4. F. HARARY, G. W. WILCOX, *Boolean Operations on Graphs*, Math. Scand., **20**, *1*, 41—52 (1967).