Fabejáró és kavics fabejáró automaták

Absztrakt

Muzamel Loránd

A *fabejáró automata* fogalmát elsőként 1971-ben A. V. Aho és J. D. Ullman definiálta, hogy segítségével modellezni tudják a szintaxis-vezérelt fordításokat. Mára, az XML nyelv elterjedése miatt, a fabejáró automata az XML lekérdezések modellezésére alkalmas eszközzé nőtte ki magát.

Fabejáró automatán egy olyan A, véges állapotú, szekvenciálisan működő felismerőt értünk, amely rendelkezik egy, az input fa valamely csúcsára mutató pointerrel. Ez a pointer az input fa élein „sétál”, lépésről lépésre változtatva az állapotait. Minden lépés előtt A „leteszteli” az aktuális u csúcsot (é.i: a pointer által mutatott csúcsot), melynek az eredménye az u csúcs címkéje (az input ábécé egy eleme) és egy nemnegatív egész, az u csúcs ú.n. „gyermek-száma” (child number: megmutatja, hogy az u csúcs hányadik fia az apjának, a gyökér gyermek-száma mindig 0). Ezektől a teszt eredményektől, továbbá az aktuális állapottól függ a következő lépés. Az aktuális csúcsok tesztelése miatt A szisztematikus fabejárásokra is képes lehet, mint például a preorder, vagy postorder keresés. Kezdetben A mindig kezdőállapotban van, a pointer pedig az input fa gyökerére mutat, ez a „kezdő konfiguráció”, továbbá A az input fát akkor fogadja el, ha létezik egy olyan kezdő konfigurációból induló lépéssorozat, hogy A végállapotba jut.

A fabejáró automaták által felismert nyelvek regulárisak, de napjainkig megválaszolatlan volt az a kérdés, hogy vajon fabejáró automatával fel lehet-e ismertetni az összes reguláris nyelvet. M. Bojańczik és T. Colcombet 2005-ben válaszolták meg, hogy fabejáró automaták nem ismerhetnek fel minden reguláris nyelvet. Felmerült tehát a kérdés, hogy lehet-e olyan – továbbra is szekvenciálisan működő – faautomatát definiálni, amelynek felismerő kapacitása erősebb, mint a fabejáró automatáké, de nem haladja meg a reguláris nyelvek osztályát.

Ilyen célból J. Engelfriet és H. J. Hoogeboom 1999-ben megjelenő cikkükben a fabejáró automata modellt kiterjesztették úgy, hogy az addigi modellhez hozzátettek egy véges sok, n kavicsból álló vermet. A kavicsok 1-től n-ig számozottak és a verem tetejéről lefelé haladva emelkedő rendezésben vannak tárolva. Ez az eszköz az *n-kavics fabejáró automata*, amely annyival tud többet a fabejáró automatánál, hogy a kavicsokat leteheti az input fa csúcsira, majd felveheti onnan, a következő szabályokat betartva: Az aktuális csúcsra csak a verem tetejéről lehet kavicsot tenni; Az aktuális csúcsról csak akkor lehet az l kavicsot visszatenni a verembe, ha l a legnagyobb kavics az input fán;

A letett kavicsok rendre egy-egy bites megjegyzendő információkat jelentenek, amelyek módosíthatják a feldolgozást. Például, egy csúcsra letett kavics jelentheti azt, hogy az automata már járt azon a csúcson. A kavics fabejáró automaták által felismert nyelvek szintén regulárisak, de az továbbra is nyitott kérdés, hogy vajon kavics fabejáró automatával fel lehet-e ismertetni minden reguláris nyelvet.

A tanszéki szeminárium-sorozat két alkalmával a (kavics) fabejáró automaták, és azok módosított változatainak felismerési kapacitásáról fogok beszélni. Sok szeretettel várok mindenkit.