

Csuhaj Varjú Erzsébet

Multi-ágens rendszerek – formális nyelvi alapon

Csuhaj Varjú Erzsébet a KLTE matematikus szakán végzett 1977-ben. 1993-tól a matematikatudomány kandidátusa, 2003-tól az MTA doktora. 1979 óta dolgozik a SZTAKI-ban. Jelenleg az intézet tudományos tanácsadója, az Elméleti Számítástudományi Kutatócsoport vezetője. Érdeklődési köre szerteágazó, a formális nyelvektől a természet által motivált kiszámítási modellekig az elméleti számítástudomány több határterületére kiterjed. Ezek némelyike a mesterséges intelligencia kutatást is érinti.



Matematikusként kezdte. Mennyire határozza meg a matematikusszemlélet az MI területén folytatott kutatásait?

Engedje meg, hogy először pontosítsak. Nem az MI területén végeztem és végzek jelenleg is kutatásokat, hanem többek között az elméleti számítástudomány olyan részterületein, amelyek indíttatásul az MI egyes, az érdeklődés középpontjában álló problémái állnak. Így például foglalkozom multiágens rendszerek formális modellezésével, az osztott kooperatív rendszerek és viselkedésük szintaktikai jellemzőinek vizsgálatával. Mint matematikus törekszem arra, hogy minél egzaktabb megközelítéseket adjak, illetve hogy minél pontosabban és meggyőzőbben írjam le absztrakt eszközökkel ezen rendszerek lényegét, általános és/vagy közös jellemzőit. Így nemcsak matematikusként kezdtem, hanem az is vagyok, olyan matematikus, akinek szakterülete az elméleti számítástudomány és annak egyes határterületei.

Érdeklődési körként a formális nyelvek elméletét, a multiágens, illetve az osztott és kooperatív rendszereket, a természetmotivált számítástudományt, valamint a természetes nyelvek gépi feldolgozását jelölte meg. Milyen pontokon kapcsolódnak ezek egymáshoz?

A formális nyelvek elméletén keresztül, amelynek eszköztárát használom multiágens rendszerek, valamint biológiai indíttatású rendszerek és viselkedésük leírására, illetve ezen eszközök használhatóságát szeretném tanulmányozni a természetes nyelvek matematikai (formális) modellezésében.

A formális nyelvek elmélete a számítástudomány egyik legrégebbi, tradicionális ága. Talán megengedik nekem azt az egyszerűsítést, hogy azt mondjam, ez az elmélet szimbólumsorozatok – azaz szavak –, a belőlük alkotott halmazok – azaz nyelvek –, valamint nyelvosztályok leírásával és tulajdonságaival

foglalkozik. A nyelvek meghatározása többféle módon történhet. Például Turing-gépek segítségével, ez esetben a Turing-gépek által elfogadott nyelvekről beszélünk, de történhet ún. grammatikák segítségével is, amikor egy véges elem-ből álló szabályrendszert adunk meg, amely szabályok megfelelő alkalmazásával egy adott szóból kiindulva a nyelv minden egyes szavát előállíthatjuk, de más, a nyelvhez nem tartozó szavakat nem.

Ezek a grammatikák vagy szabályrendszerek a nyelv meghatározásának egyik természetes módját jelentik. Gondoljunk csak arra, hogy a természetes nyelvek sajátosságaiból kiindulva vonatkoztatta el Noam Chomsky a generatív grammatika fogalmát, amely mint felismerés és fogalomalkotás a formális nyelvek tudományágának egyik kiindulópontja lett. Könnyen mondhatja valaki erre, hogy valóban, a természetes nyelvek modellezésében elképzelhető valamiféle létjogosultsága ennek a megközelítésnek. De miféle alapja lehet annak a gondolatnak, hogy multiágens rendszerek leírására használjunk nyelveket és grammatikákat? A válasz erre a következőkben foglalható össze.

Az ágensek, amelyek lehetnek tudásforrások, biológiai egyedek vagy egyéb más, viselkedést mutató entitások, leírhatók, modellezhetők az őket működtető szabályok együttesével.

(Természetesen ezek a szabályrendszerek sohasem a teljesség igényével adóttak, hanem valamilyen szempontból elvonatkoztatott, egyszerűsített szabályokról van szó). A szabályhalmazok, azaz a grammatikák működésük során szavakat határoznak meg, amely szavakat úgy tekintjük, mint az ágensek és adott környezetük viselkedésének leírását. Ily módon szintaktikai eszközökkel jellemezni tudjuk egy ágensközösség, egy multiágens rendszer viselkedését, le tudjuk írni sajátosságait.

Meg kell jegyezmem, hogy a szabályokat a legáltalánosabb értelemben vesszük: a programnyelvek utasításaitól kezdve a biológiai egyedek fejlődését, illetve viselkedését leíró (egyszerűsített és formalizált) szabályokat is beleértve.

Ez a gondolat, vagyis hogy a multiágens rendszereket, illetve viselkedésüket leírhatjuk együttműködő grammatikák segítségével, vezetett a formális nyelvek „grammatikarendszerek” elnevezésű elismert tárgykörének kiépítéséhez. E terület egyik elindítója és meghatározó művelője voltam, és jelenleg is vagyok. Azt mondhatjuk, hogy a grammatikarendszerek elmélete nem egyéb, mint a multiágens rendszerek elméletének formális nyelvi, szintaktikai megfelelője.

Milyen kutatásokat végzett a formális nyelvek, valamint a természetesnyelv-feldolgozás terén?

A természetes nyelvek gépi feldolgozása területén konkrét kutatásokat nem végeztem, bár tervezek kutatásokat a jövőben. Néhány dolgozatot írtam a természetes nyelvek matematikai modellezése tárgykörében. Szerzőtársaimmal

együtt azt vizsgáltam, hogyan alkalmazhatók az előbb említett „grammatikarendszerek” tárgykör eszközei a pragmatikában, a dialógusok, illetve az együtt élő kulturális, nyelvi közösségek viselkedésének modellezésében.

A formális nyelvek elméletében kiterjedt kutatásokat végeztem.

Mint már az előbb említettem, egyik elindítója voltam a grammatikarendszerek elméletének, amelyet számos modellel és meghatározó eredménnyel gazdagítottam szerzőtársaimmal együtt.

Az elmélet kiindulópontjául az általam és Jürgen Dassow szerzőtársam által 1988-ban bevezetett kooperatív osztott grammatikarendszer fogalmát tekinthetjük, amely nem egyéb, mint az MI-kutatásból jól ismert tábla-architektúra egyik formális modellje. A tábla típusú problémamegoldó rendszerek esetében független, egymással közvetlenül információt nem cserélő ágensek úgy oldanak meg egy kitérőt problémát, hogy egy együttműködési protokoll előírásait követve felváltva módosítják egy közös adatbázis, az ún. tábla tartalmát. Ez az adatbázis tartalmazza minden időpillanatban a problémamegoldás folyamatára, illetve a megoldásra vonatkozó információkat. Az általunk bevezetett modellben az ágenseket grammatikákkal, a tábla tartalmát egy, a grammatikák által az együttműködési protokoll szerint időről időre módosított szóval jellemezzük.

A kooperatív osztott grammatikarendszer fogalma inspirálólágg hatott más kutatókra is, ami új modellek, kutatási irányok megjelenéséhez vezetett. Az eszköztár használhatónak bizonyult tudáshálózatok szintaktikai jellemzésére, de általa egyszerű reaktív rendszerek, vagy a mesterséges élet különböző paradigmái is modellezhetők voltak formális nyelvi eszközökkel. Jürgen Dassow mellett meg kívánom említeni Gheorghe Paun, Jozef Kelemen, illetve Alica Kelemenová kollégáinkat, akiknek tőlünk függetlenül vagy velünk együtt kidolgozott alapvető modellek létrehozását köszönhetjük, és akik számos fontos koncepcionális eredménnyel járultak hozzá a terület fejlesztéséhez. Ugyancsak meg szeretném említeni Vaszil György kollégámat, korábbi PhD-tanítványomat, aki jelentős munkásságot fejtett ki a grammatikarendszerek hálózatokat modellező részterületén.

A grammatikarendszerek elméletét egy nagyobb nemzetközi közösségen belül fejlesztettük ki elismert tárgykörre. Az érdeklődő olvasó csoportunk honlapján talál hasznos információkat (<http://www.sztaki.hu/tcs/>).

A történeti hűség kedvéért meg kell jegyezmem, hogy a kooperáló grammatikát mint terminológiát 1978-ban Grzegorz Rozenberg és Robert Meersman használta először, többszintű grammatikák, illetve konkurens operációs rendszerek modellezésére egy nemzetközi konferencián publikált dolgozatukban. A kutatásokat azonban nem folytatták. 1988-ban észrevettem, hogy az általuk bevezetett fogalom általánosítható, alkalmazható az osztott kooperatív rendszerek modellezésére. Ez a felismerés vezetett a kooperatív osztott grammatikarendszer fogalmának bevezetéséhez.

Az utóbbi időben a molekuláris számítástudomány modelljeivel is foglalkozom. Így például a DNS rekombinációs viselkedését utánozó műveletekre alapozott osztott struktúrák, a kémcsőrendszerek elméletével és a membránkiszámítások elméletével. Ez utóbbi egy rendkívül dinamikus fejlődő, nagyon pozitív nemzetközi visszhanggal rendelkező terület, amelyet Gheorghe Paun indított el 1998-ban. E kutatásokat vele és más külföldi társainkkal műveljük Vaszil György kollégámmal együtt.

Nagy megtiszteltetés számomra, hogy mind a DNS-kiszámításhoz kapcsolódóan, mind pedig a grammatikarendszerek egyes hálózat típusú modelljeinek területén együtt dolgozhatom Arto Salomaa professzorral, aki az elméleti számítástudomány, a formális nyelvek elméletének kimagasló, meghatározó alakja – többek között a Magyar Tudományos Akadémiának is tiszteletbeli tagja.

Munkásságom a formális nyelvek elméletén belül még egy nagyobb tárgykörhöz kapcsolódik: az úgynevezett leírási bonyolultság (*descriptive complexity*) elméletéhez, amely azt vizsgálja, hogy a nyelvosztályokat meghatározó eszközök – grammatikák vagy automaták – mennyire tömör leírását tudják adni a vizsgált nyelvosztály elemeinek. Megpróbálom a nyelvosztály elemeit előállító grammatikák paramétereit minimalizálni – ha ez lehetséges egyáltalán –, meghatározni, mire van szüksége a megfelelő grammatika osztálynak az adott kiszámítási erő eléréséhez. Ezek a kérdések azért érdekesek a multiágens rendszerek elmélete szempontjából, mert tulajdonképpen azt vizsgáljuk, hogy nagyon egyszerű, korlátozott méretű ágensek együttese mennyire bonyolult nyelvet, viselkedést tudnak meghatározni. Mennyire egyszerűsíthető le az ágens formális modellje ahhoz, hogy még alkalmas legyen bonyolult viselkedésű ágensközösségek leírására?

Mennyire?

Attól függ. A grammatikarendszerek elméletében számos nagyon érdekes eredmény található. Ilyen például Lucian Ilie és Arto Salomaa eredménye, amely kimondja, hogy a véges automatáknak megfelelő reguláris grammatikák bizonyos típusú – utasításra szűrőkön át információt cserélő komponensekkel rendelkező – hálózataik esetében már a három együttműködő grammatikából álló rendszerek is alkalmasak a Turing-gépek kiváltására. A reguláris grammatikák pedig köztudottan rendkívül egyszerűek. Ha az információcseré nem utasításra, hanem igény szerint történik, akkor is hasonló a helyzet. Vaszil György kollégámmal megmutattuk, hogy a párhuzamosan dolgozó és igény szerint kommunikáló környezetfüggetlen grammatikák hálózatai a Turing-gépekkel egyenlő erejű kiszámítási eszközök, sőt, még olyan grammatikákból álló csoportok is képesek ezt az erőt felmutatni, ahol a grammatikák legfeljebb hét szabállyal rendelkeznek. Ha a szabályok számát nem korlátozzuk, akkor a hálózat elemeinek számát korlátozhatjuk egy kis konstanssal. A környezetfüggetlen grammatikák szintén viszonylag kiszámítási erőt képviselnek.

Ön vezeti a az Elméleti Számítástudományi Kutatócsoportot, korábbi nevén Multiágens Rendszerek Modellézése Kutatócsoportot.

1997-ben adódott az önálló csoport létrehozásának lehetősége. Volt tanítványommal, a kiváló kutató Vaszil Györggyel jelenleg ketten vagyunk törzstagok, de a csoport létszáma időről időre változik, hiszen hazai és külföldi PhD-hallgatók csatlakoznak hozzánk.

Így kitűnő tanítványom és munkatársunk volt Csima Judit, aki ma a Műegyetemen adjunktus, María Dolores Jiménez-López, aki jelenleg a tarragonai Rovira I Virgili Egyetem (Spanyolország) tanára, valamint Maurice H. ter Beek Hollandiából, aki először mint diplomamunkáját író végzős egyetemi vendéghallgató, majd mint ERCIM ösztöndíjas posztdoktor dolgozott csoportunkban. Jelenleg a pisai CNR kutatóintézetben tudományos munkatárs.

Csoportunk kiterjedt nemzetközi kapcsolatrendszerrel rendelkezik, külföldi kollégáink rendszeresen ellátogatnak hozzánk hosszabb-rövidebb időre közös kutatás céljából. Nagyon fontosnak tartom, hogy ne csak a hazai kutatók utazzanak külföldi tanulmányutakra, hanem fordított irányú áramlás is legyen, azaz építsünk ki olyan kutatóhelyeket, teremtsünk olyan feltételeket, hogy a külföldi kutatók is örömmel látogassanak el hozzánk. Különösen fontos ez most, hogy hazánk az Európai Unió tagja lett. Természetesen intenzív kapcsolatban állunk a hazai kutatóhelyekkel, egyetemekkel is. Időről időre tanítunk, jómagam pedig az ELTE Informatikai Doktori Iskolájában is aktívan tevékenykedem, tagja vagyok a doktori tanácsnak, és egyike vagyok az iskola külső alapító tagjainak.

Mely eddigi projektjeit tartja a legjelentősebbeknek?

Minthogy alapkutatással foglalkozunk, talán nem használnám a projekt szót, beszéljünk inkább támogatott kutatási programokról. Munkáinkat támogatta, és jelenleg is támogatja az OTKA, tagjai vagyunk egy ún. EU-támogatott tematikus hálózatnak a molekuláris számítástudomány tárgykörében, és több kétoldalú kormányközi egyezmény alapján meghirdetett kutatási támogatásban is részesültünk. Intézetünk 2000-ben megkapta az EU Centre of Excellence címet, és ebben a kutatási programcsomagban is részt veszünk egy résztéma erejéig.

Legjelentősebbnek azt tartom, hogy a grammatikarendszerek tárgykörét sikerült a formális nyelvek elméletének elismert területévé fejlesztenünk, és hozzájárulnunk a formális nyelvek elméletének fejlődéséhez számos értékes és érdekes modellel, eredménnyel, olyan eredményekkel, amelyek kapcsolódási pontot jelentenek és jelenthetnek más tudományágakhoz, így a mesterséges intelligencia kutatásához is. Ugyancsak fontosnak tartom a molekuláris számítástudomány területén kifejtett munkásságunkat.

Hogyan látja az ágens kutatás jelenlegi helyzetét, illetve az önök szerepét ebben a helyzetben?

Az ágens kutatás egészére nincs teljes rálátásom. Úgy gondolom viszont, a mi

kutatásainkban még számos lehetőség rejlik. Rendkívül egyszerű – és ezáltal akár egyszerűen megtanulható, felfogható – gépek vagy szoftverek alapjait igyekszünk meghatározni, az esetek egy részében mégis rendkívül bonyolult feladatokat lehet segítségükkel megoldani. Így kutatásainknak nemcsak az elméleti vonalában látok perspektívákat, hanem az alkalmazásaiban is. Szorosabb együttműködés más tudományterületek, például az evolúciós számítástudomány, a robotika művelőivel pontosabb modellek megalkotásához vezetne, és megnyitná az utat a számukra is közvetlenül használható eszköztár kiépítéséhez. Ugyanez vonatkozik a természetes nyelvek gépi feldolgozására, például a többnyelvű környezetben való kommunikációt támogató szoftverek kiépítésére.

Vizsgálataink új típusú kiszámítási modellek létrehozásához járulnak hozzá, ami inspirálólag hathat a más tudományágakkal foglalkozó kutatókra.

Hadd említsem meg, hogy a formális nyelvek elméletében egészen a nyolcvanas évekig általában, sőt majdnem kizárólagosan arra törekedtek, hogy egy nyelvet egy eszközzel – grammatikával vagy automatával – határozzanak meg. Ez a felfogás megfelelt annak az elképzelésnek, hogy a kiszámításokat egy nagyon hatékony processzor, kiszámító eszköz segítségével vigyék végbe. Így látható, hogy a formális nyelvek elméletében az a gondolat, hogy esetleg több grammatika együttműködve és egymással információt cserélve határozzon meg egy nyelvet, koncepcionálisan újszerű volt. A fogalom kialakítását és elfogadtatását azonban nagyban inspirálta és segítette az a tény, hogy megjelentek a számítógép-hálózatok, realitássá vált a számítógépen végzett, osztott, kooperatív munka.

Jelenlegi projektjeikben – MolCoNet, biológiailag inspirált számítások, molekuláris számítások elosztott modelljei – fontos szerepet játszik a biológia.

A biológiai indíttatású kiszámítási modellek a jelenkori számítástudomány érdeklődésének homlokterében állnak. Céljuk az, hogy az élő szervezetek vagy azok alkotórészeinek működési elveit, viselkedését utánozva nagy hatékonyságú, a hagyományostól eltérő elveken működő kiszámítási eszközöket hozzanak létre, illetve új megvilágításba helyezve azt, a kiszámítás fogalmának pontosabb megértéséhez jussanak.

Két fő témakörrel foglalkozunk ezen a területen. Az első terület az úgynevezett membránkiszámítások elmélete, amely az élő sejt felépítését és működését utánozni próbáló kiszámítási modellek elmélete. A kiszámítási eszköz, az ún. membránrendszer egy osztott architektúra, amelyben az egyes komponensek és szabályaik formális nyelvi műveletekkel vannak reprezentálva. Az architektúra felépítésében megfelel a sejt membránstruktúrájának, az egyes komponensek pedig szabályaik segítségével szavakból álló multihalmazokon végeznek műveleteket. (A multihalmaz esetében egy szó több példányban is előfordulhat a szavakból álló együttesben). A szavak (bio-)kémiai alkotórészeknek, a műveletek

kémiai reakcióknak felelnek meg. A szavak közlekednek az egyes komponensek között, hasonlóan a sejtekben lezajló folyamatokhoz.

A DNS-kiszámítás területén ugyancsak végzünk kutatásokat. A DNS szét-törésének és újrakombinálásának elveit imitáló formális nyelvi műveletekre alapozott kiszámítási eszközöket hozunk létre, és vizsgáljuk tulajdonságait. A DNS-szálak szavakkal is reprezentálhatók, így a formális nyelvi megközelítések megalapozottak. Elsősorban úgynevezett kémcsőrendszereket, azaz ezen eszközökből épített osztott rendszereket tanulmányozunk, amelyek egy része a Turing-gépekkel egyenlő kiszámítási erejű eszközt képvisel.

Bár kiindulásukban különböznek, sem a membránrendszerek, sem a kémcsőrendszerek nem különböznek alapvető konstrukciós elveikben a grammatikarendszerektől – osztott architektúrák, amelyek komponensei formális nyelvi műveletekre alapozott kiszámítási eszközök.

Miben látja a kutatás-fejlesztési projektek sikerének, eredményességének a titkát?

Csak az alap kutatás területén végzett kutatásokról tudok nyilatkozni, de megvallo-
lom, nem szívesen teszem. Már az is kérdéses, mikor mondhatjuk, hogy egy alap-
kutatás sikeres. A sikeresség feltételei között feltétlen ott van az újszerű gondolat
és/vagy látásmód, az elmélyült gondolkodás, a kitartó munka, és a szerencse.
Természetesen ebben az esetben a sikeresség rövidtávon nem ítélni meg.

*Ha visszamehetnének az időben, és most lenne egyetemista, mivel foglalkozna
legszívesebben? Milyen témakörben kutatási területben látna komoly perspek-
tívát?*

Hogy mivel foglalkoznék legszívesebben, nem tudom, mivel ehhez bele kellene
illeszkednem egy mai fiatal gondolkodás- és érzésvilágába. Viszont minden
egyetemistának szívesen ajánlom figyelmébe szakterületemet, az elméleti
számítástudományt és határterületeit, mert olyan érdekesek. Az utóbbi évtized
számos izgalmas új terület megjelenését hozta magával, de nem szeretnék az e
területek hosszú távú jövőjét illetően jóslásokba bocsátkozni.

Csuhaj Varjú Erzsébet

MTA SZTAKI Elméleti Számítástudományi Kutatócsoport

1111 Budapest, Kende u. 13-17.

<http://www.sztaki.hu/reszleg/TCS>