

## **A Mesterséges Intelligencia kihívásai: Fuzzy rendszerek, adatbányászat, és döntések modellezése – Interjú Dombi Józseffel**

**Dombi József a Szegedi Egyetem -matematikus- szakán végzett 1972-ben, és azóta is ott dolgozik. Még a végzés évében Szentpétervárott (akkori nevén Leningrádban) majd Lipcsében folytatta tanulmányait. 1979-ben DAAD ösztöndíjat kapott és a Németországi Aacheni Egyetemen végzett kutatásokat. 1986-ban megkapta a Humboldt-ösztöndíjat és Németországban folytatta megkezdett kutatásait. Több külföldi intézményben töltött el hosszabb rövidebb időt. Így járt Bristolban, Linzben, többször volt Belgiumban és Finnország különböző egyetemein, mint vendégelőadó. Az elméleti kutatások mellett az eredményeket gyakorlatban megvalósító cégeket alapított melyek jelentős nemzetközi elismerésben részesültek. Érdeklődési köre rendkívül szerteágazó spektrumot fed le.**

**- Adatbányászat, adatvizualizáció, döntéstámogatás, fuzzy, neurális hálók, genetikus algoritmusok... Melyiket emelné ki a sok kutatási területe közül?**

- Nehéz lenne bármelyiket kiemelni. Nemcsak azért, mert bármelyik fontosabb lenne a másiknál. Most még ez nem eldönthető. Idő kell, hogy rálássunk melyik is lesz fontos a jövőben. Úgy jellemezhetnénk mindegyiket, hogy szorosan kötődnek a heurisztikához, ami probléma, mert a matematikában és egyáltalán amit tudománynak nevezünk, valami bizonyosat szeretnénk állítani, amit tételek garantálnak A heurisztikához

való fordulás egyfajta kényszer, mert nagyon sok olyan probléma merül fel a mérnöki tudományokban amire az MI-nek kellene választ adni, és mai tudásunk szerint nem tudunk. Ezért a természethez fordulunk és ötleteket lesünk el, amelyek -úgy néz ki - jól működnek és így kapjuk az említett heurisztikus közelítéseket. Első látásra úgy tűnik ezek mind különbözőek, de – véleményem szerint – egy még csak körvonalazódó - új egységes irányzat elemei. Közös bennük, hogy mindegyik a természethez fordul segítségért és próbálja ellesni titkát: a genetikai algoritmus és az evolúciós programozás Darwin világtól, a neurális hálózatok az agykutatástól les el fortélyokat. Az immunrendszer működésének is van megfelelője már az informatikában. A hangya algoritmusok a közösségek optimalizáló tevékenysége alapján működnek, a fuzzy rendszerek pedig az emberi szóhasználat pontatlanságából csinál erényt.

A számítástechnika megköveteli, hogy a feladatokat gyorsan lehetőleg rövid idő alatt oldjuk meg. Ehhez az informatika hardver oldalról mindent megtesz és itt nincs is gond, mert a számítógépek sebessége rohamosan nő. A gép azonban kevés, hatékony algoritmusok is szükségesek. Ha 100 várost kell meglátogatnunk, hogy közben a legrövidebb utat tegyük meg az összes lehetőségek száma  $1*2*3*4*...*100$  (az első száz számot kell összeszorozni, ez matematika nyelvén 100 faktoriálisnak mondjuk) ennek nagysága 10 a 155. hatványánál is nagyobb. Ez olyan nagy szám, ha a ma létező leggyorsabb számítógépet tekintjük, ami 1 másodperc alatt 1 milliárd művelet végez el akkor az összes lehetőség kipróbálása 10 a 147. hatványon évig tartana. Tehát sokkal de sokkal hosszabb ideig, mint amióta a

világmindenség létezik. Becslések szerint 13 milliárd év a világmindenség kora és ez eltörpül e szám mellett.

Az elméleti matematika felvetette a kiszámíthatóság problémáját: azaz milyen algoritmusoknál van remény, hogy reális időn belül eredményt kapjunk. Kiderült, hogy a problémák nagy része nem ebbe az osztályba tartozik. Ellenben a fentihez hasonló feladatokat kellene megoldani a mindennapokban. Például 100 benzinkút ellátása nem is olyan sok és az optimum megtalálása lehetetlen. A természet azonban valahogy megoldotta ezeket az összetett problémákat. Például az emberi szemnek nagyon nagy a komplexitása. Nehezebb egy szem optimális tervezése, mint 100 város közti legrövidebb út megtalálása. Ha megnézzük, hogy véletlenszerűen, vagy próbálkozásos kísérletekkel mennyi idő alatt alakulhat ki egy szem, amely viszonylag minimális energiaellátással kvantumos tartományban képes működni, akkor azt mondhatjuk, hogy ez egy olyan probléma, aminek matematikai, vagy számítógépes megoldására még hosszabb időt venne igénybe. Tehát ki se alakulhatott volna, és mégis itt van. Látom a világot. A nehéz feladatra jó példa a gén szekvencia megfelelő sorrendjének kitalálása. A heurisztikus algoritmusok közelítik az optimumot és kísérletekkel meg lehet vizsgálni jóságukat. Ha 95%-os pontossággal megelégszünk akkor egy PC-n az algoritmus fél perc alatt talál jó megoldást a 100 város meglátogatásának problémájára. Ezek az eljárások az utóbbi évek eredményei.

**- Többi évtizede foglalkozik a fuzzy-val.**

- Nagyon érdekes a Zadeh-i gondolat (Zadeh vezette be ezt a fogalmat): a hétköznapi életben a beszédmód, az utasításaink elég nagy lazaságot mutatnak a matematikai szigorhoz képest, ahol a pontos definíció-meghatározás nélkül nem tudnánk tételeket bizonyítani, mégis a mindennapokban nagyon jól elvagyunk vele. A számítógéppel való kommunikáció a robotok irányítása azért sem jó, mert nem tudjuk ezeket a teljesen laza szavakat használni. A fuzzy-t – talán leginkább elmosódott határú halmaznak lehetne fordítani. Hogy pontosabban érthető legyen mit is takar ez nézzünk néhány példát: valaki magas, de mondhatnám azt is magas hegy vagy magas árfolyam. Ilyen szó a fiatal kolléga, fiatal akadémikus, vagy az a kijelentés, hogy gyakran járok moziba. Mi a pontos definíciója a magasnak, fiatalnak, a gyakrannak? Kérdés: Lehet-e matematikai formalizmussal leírni ezeket a lazaságokat és így segíteni a kommunikációt? Ez eleve ellentmondás. Lehetséges a lazaságok szigorú modellezése? Pedig jó lenne, mert akkor azt kérdezhetném egy adatbázistól, hogy keressen egy közepes árfekvésű diesel autót, aminek nagy a csomagtere és fogyasztása nagyon kicsi. A robotot pedig úgy vezérelhetném, most lassabban kell menni, mert egy kicsit csúszós út következik. A fuzzy-nak a természetes nyelvben meglévő fogalmak matematikai modelljéről kellene szólnia, illetve ennél kicsit többről is, mert nem biztos, hogy nyelvünk képes minden lazaságot kezelni és az se biztos, hogy van mindenre szavunk. Például mindennap lemegyünk különböző lépcsőkön, ha arra kerül sor és a lépcső laza definíciójához alkalmazkodnak izmaink és vezérlik a testet, ellenben szavakkal nehéz leírni a járás utasításait!

Visszatérve a fuzzy-ra, létrejöttét egyrészt az ember és a humán oldal motiválta, másrészt az, hogy bizonyos dolgokat össze kell foglalni, hogy ne vesszünk el a részletekben. A túl sok információ, számadat megnehezíti a dolgunk.

A fuzzy rendszerek szorosan kapcsolódnak még a többértékű, folytonos logikákhoz. Már a harmincas-negyvenes években - elsősorban Lengyelországot említhetném - elkezdődött ennek a kutatása. Az általánosítás szerint nem csak igaz és hamis érték létezik, hanem az igazságértékek a 0 és 1 közötti bármely számot felvehetik. Itt 0 jelenti a hamisat és 1 az igazat. Úgy is mondhatnánk nem csak feketében és fehérben kell látni a világot. Létezik egy spektruma az igazságértékeknek.

Mindez azt jelentette, hogy új tudományterület született, amely különböző válságokkal tűzdelve a mai napig éli az életét.

**- Kandidátusi értekezést is a többtényezős döntések és a fuzzy kapcsolatából írta.**

- Nem egyszerű véleményt alkotni a fuzzy kutatásokról.

Fontosnak tartom hogy a felvetett problémára választ adjunk, de nem értek egyet minden próbálkozással. A publikációs kényszer miatt mindenféle kérész életű próbálkozások is születnek és aztán pillanatok alatt eltűnnek. Néhány év és senki sem emlékszik rájuk. Nem egyértelmű, milyen szerepet fog betölteni a későbbiekben a fuzzy. Az is kérdés melyik irányzat lesz a fuzzy-n belül életképes. A fuzzy rendszerekre vonatkozó kutatás szállka volt az elméleti matematikusok szemében. Az első időszakban az elméleti matematika és az informatika is elég agresszíven lépett fel ellene. Mi köze a tudománynak a pontatlansághoz, a pontatlanság modellezéséhez? Évtizedekkel ezelőtt hasonló volt a

helyzet a mára már teljesen elfogadott valószínűség-számítással is. A kialakuláskor, amikor a fő alkalmazói csak a vásári játékosok, a kártyások és egyéb szerencselovagok lehettek, nem foglalkozhatott a tiszta matematika vele. Ki emlékszik már arra az időkre, amikor paradoxonokat gyártottak, hogy elriasszák a komoly kutatókat ezekről a kétes vizekről. Ma pedig, már nem lehet kísérleti eredményeket elfogadtatni anélkül, hogy statisztikai megbízhatósággal ne támasztanák alá az eredményeket. A fuzzy-elméletben – ellentétben a valószínűség-számítással – nem jelentek meg a paradoxonok, ami nem tett jót, mert jobb lett volna pontosan látni már a kezdeteknél a hiányosságokat. A legtöbb tudós elvetette, és nem vette komolyan. A *Mesterséges intelligencia* folyóirat húsz évig egyetlen cikket se volt hajlandó közölni ebből a témából. A múlt években jelentek meg az csak az első cikkek. De nem kell azt gondolni, hogy fórum nélkül maradt volna ez a tudományterület. Most az Interneten egy kereső 3.5 millió választ ad a fuzzy kulcsszóra, míg a teljes mindent felölelő mesterséges intelligencia témákra csak 3 milliót. Azzal, hogy ilyen sokan foglalkoznak a fuzzy-val, az is egy fajta legitimációt biztosít a tudományterületnek.

De nézzük egy kicsit alaposabban meg ezt az elméletet. Ha a 0 és az 1 értéket felvevő logikát egy spektrummal egészítjük ki, akkor sokan azt gondolják, hogy ez a logika egyfajta általánosítása.

Közel sincs így, mert nem két pontra, hanem egy folytonos függvény összes belső pontjára kell bizonyos tulajdonságoknak teljesülniük. Ez pedig jóval szűkebb halmaz, mint amikor csak a határpontokon követelem meg a tulajdonságok teljesülését. Nincs egyetlen olyan folytonos kiterjesztés sem, ami az összes Boole

azonosságot teljesítené. Tehát azt is mondhatjuk, hogy a fuzzy nem egyfajta általánosítása a logikának, hanem leszűkítése, és nagyon kevés olyan struktúra tesz ennek eleget, amelyekkel jól lehet dolgozni.

Elég egyszerű megmutatni, hogy - nem az összes - hanem már négy-öt azonosságot megadva, egyetlen folytonos logika se konstruálható. Ez egyfajta lehetetlenségi tétel, és ezért lehet azt mondani, hogy a kiválasztott tulajdonságok szerint más-más fuzzy-logikát lehet felépíteni. Az egységes elmélet különböző irányzatokra szakadt szét.

Most már az a kérdés hogy miként lehet egy jó fuzzy-elméletet kiválasztani? A válasz: Lehetőleg sok jó tulajdonsága legyen és/vagy negációműveletnek, az implikációnak. Felsoroljuk a megkövetelt tulajdonságokat, amiket fontosnak tartunk és egy gyakorlatban alkalmazható számítási eljárást konstruálunk.

A fuzzy a humán kommunikációtól akar ellesni trükköket, de nyelvünk elég bonyolult egy egységes, használható modellhez. A szavak nagyon sokfélék. A fiatal, a gyakran, a körülbelül igaz, stb. teljesen más fogalmak – egész más funkcióik vannak a nyelvben, nehéz egységes közös alapra hozni ezeket. Különbözik minden szavunk tartalmaz egyfajta fuzzyságot. Nemcsak az előbb említett tulajdonságjelzők, hanem a fogalmaink is pontatlanok. Mi az, hogy szék? Mi az hogy csomag? Ezek is lehetnek fuzzy-fogalomak, mert homályos elképzelésünk van róla. A definíciós kényszer azonban tévútra is vezethet. A mindennapi életben azonban állandóan áthágjuk a definíciókat. A jogalkotás például kénytelen definíciókkal dolgozni és ezért kivételek sokaságát kell állandóan kezelni. Ha visszatérek a székhez, és úgy definiálom, ahogy szoktuk, akkor például a

favágók az erdőben a kivágott fák azon tönkjeit nevezik székeknek, melyekre ráülnek, és azt nevezik asztalnak amin esznek. De ha másnap az asztalra ülnek az lesz szék és a szék asztal lehet. A világnak a fogalmi meghatározása mellett, a funkcionalitás jellegű meghatározása is létezik. Azt kérdezzük ilyenkor mire jó, mire használható? Ezt egyfajta intenciónak, intencionalitásnak lehet nevezni, amivel Dennettre és Pinkerre szeretnék utalni. Mire lehetne használni egy fuzzy-rendszert? Az Internet révén lekérdezésekként éljük meg a világot. Olyan kérdéseket kell generálnunk, melyekre megfelelő válaszokat kaphatunk. A keresések alapvetően csak szavakra szolgálnak, holott a világban adatbázisok sokasága létezik, amikben ma még nem tudunk turkálni. Jó lenne ezeket megszólítani. A számok helyett a mondatok az emberi kommunikáció elemei. Az időjárás-jelentés jó példa erre. Ahelyett hogy Hektopascal, Celsius, fedettségi százalék és csapadék valószínűségi érték adatok tömegét zúdítanak ránk, azt mondják: kissé borongós idő várható helyenként heves zivatarral és hűvös idő várható. Hogyan fordítjuk le ezeket a számokat mondatokká? Ez a technika nagyon sok helyen jól jönne, például a gazdasági elemzések standardizált elemzésében.

A megoldandó feladat következő: vannak numerikus értékeink, változóink – árfolyam, befektetés mértéke, létszám –, amelyekhez fogalmakat kell rendelnünk. A fogalmak hozzárendelése azt jelenti, hogy megfelelő kategóriákra kell bontani ezt a változót. Mikor mondjuk, hogy jó a kategóriára bontás? Ha a gépkocsikat tekintem, és köbcentijük alapján csoportosítom azokat, akkor az 1300, az 1500, az 1600, a 2000 köbcentik a jó határok. Egy programnak olyannak kell lennie,



hogy éppen ezekre a határokra találjon rá, amihez bizonyos intelligens algoritmusok szükségesek. A diszkretizálást még kontextus-függővé is kell tennünk, mielőtt szavakat rendelnénk a tartományokhoz. Például az elvárható profitnak más a tartományi felosztása a jó a bankok és más a mezőgazdasági tevékenység estén. Itt egyfajta intelligenciát kell bevezetni, és ezekre nagyon jók az elmosódott határú halmazok.

A másik alkalmazás a lekérdezések rugalmassá tétele. Az SQL-kérdésekhez a határokat rögzítve, nagyon gyakran vagy kapunk ötezer választ vagy semmit se mert nincs ilyen alternatíva, ami a feltételeknek eleget tesz. Egyikkel se tudok sokat kezdeni, mert a jó válaszhoz valahogy még fel kell térképeznem az adatokat, és ilyenkor egy iteratív játék szokott kezdődni, amivel elmegy a fél napom... Azt kell elérni, ha lehetetlen kérdést is teszek fel, akkor is kapjak valami választ. Mivel több érték áll a rendelkezésemre, a fuzzy-lekérdezés alapvetően képes a rákérdezett alternatívák sorrendjének meghatározására, aszerint, hogy a feltételeimnek mennyiben tesznek eleget az alternatívák.

A kutatásaimra jellemző, hogy mindig speciális, egyszerű feladatot vizsgálok. Az általánosítást ezekből az egyedi esetekből keresem meg. Gyakran azonban az általánosabb formalizmussal egyszerűbb is belátni a tételt.

### **- Hogyan helyezi el önmagát a fuzzy-kutatásban?**

- Más irányzatot képviselek a fuzzy-n belül, mint a kutatók többsége. Mik a fő elvek?

Ha a klasszikus logikát nem tudjuk leírni folytonos logikákkal legalább készítsünk olyan logikát, hogy egyre jobban megközelítsük azt. Ha nem is igaz az azonosság a folytonos

esetben, akkor egy bizonyos sorozat révén egyre inkább megközelíthető legyen.

Szintén lényeges, hogy a fuzzy logikában nagyon fontos szerepet játszó az úgynevezett halmazhoz tartozási függvény, ami leírja, hogy egy-egy elem mennyire tartozik az adott halmazhoz, legyen jól definiált. Mivel a fuzzy elmélet mind a mai napig nem válaszolta meg szemantikus jelentését ennek a függvénynek, erős támadási felületet hagyott az ellenzőknek. Így olyan az egész, mintha a valószínűség-számítást úgy építenénk, hogy el akarunk feledkezni az eloszlás-függvényekről. Pedig statisztika e nélkül nem lehetséges. Aszerint, hogy milyen halmazhoz tartozási függvényt definiálunk, más-más területre merészkedünk. Az approximációs elvből kiindulva ez a függvény olyan lehet csak, ami egy halmaz kétértékű karakterisztikus függvényét közelíti. A határon jelezve, hogy éppen a határon vagyunk  $\frac{1}{2}$ -et vesz fel. Nem kell olyan nagy problémát csinálni belőle, egyszerűen választani kell egy közelítő-függvényt, ami a logisztikus vagy más néven szigmoid és az ehhez tartozó operátorokat kell jól meghatározni, ami már nehezebb probléma.

Az első jelentősebb cikkem ezen a területen született, amiben az operátoroknak a DeMorgan azonossághoz való viszonyát vizsgáltam. Egy példát is adtam arra, hogyan lehet ilyen operátort konstruálni. Az élet csalafintaságaihoz tartozik, hogy a későbbiekben nem a cikk, hanem a példa bizonyult sokkal érdekesebbnek. Ma Dombi-operátornak hívják, és rengetegen használják azóta. Nagyon sok helyen alkalmazzák ezeket.

Japánban a mosógépek vezérlésében – tudtommal – Dombi-operátor szerepel. A mosógépnek azért van jóval kifinomultabb elektronikája Japánban, mert tilos a magas hőmérsékleten való

mosás, nincsenek olyan gépek, amik hatvan vagy kilencven fokon mosnának. Harminc-negyven fok a maximális hőmérséklet, és ez egészen másfajta technológiát követel meg. Az Aachenben végzett kutatásaim is megmutatták, hogy csak a logikai operátorok keveseknek bizonyultak a fuzzy-ban. Amikor döntéseket hozunk, nem csak és-ben és vagy-ban gondolkozunk, hanem aggregáljuk valahogy az értékeket, ezért az aggregációról írtam egy másik, viszonylag jelentősebb cikket, és aztán végül visszatértem a halmazhoz tartozási függvényre egy harmadik cikkben, ami pontosabban a módosítósók általános leírását adja. A tulajdonságjelzők elé tett „nagyon”, „többé kevésbé” stb. módosító transzformációkat adom meg.

Van itt most három különböző terület: a logikai műveletek világa negációval, az aggregáció világa – amihez mai kutatásaim szerint még egy fogalmat hozzá kell tenni: a többértékű, vagy folytonos preferenciát –, a harmadik pedig az említett a módosítósók modellezése. A cikkek után öt-hat évvel, egy rendszerezési munka után jöttem rá, hogy mind a három tudományos munka – melyeket egymástól függetlennek gondoltam – közös alakra, hozható. Ezt a világot *pliant* más szóval *flexibilis rendszernek* neveztem el. Jól lehet számolni vele, egyfajta kalkulus. Olyan, mint az összeadás, szorzás, osztás a mindennapi életünkben. A koncepció közel áll a fuzzy eredeti intencióihoz is.

**- Az adatbányászatban szintén fontos eredményeket ért el.**

- Amellett, hogy egyetemen dolgozom, cégeket is irányítottam. 1992-ben, kidolgoztam egy vizualizációs rendszert és eljárást, ami a kontextus függő lekérdezést támogatja, amire egy belga kutatótársam 1995-ben azt mondta, hogy *adatbányászat*. Ekkor

kezdett ez a terület lábra kapni. A négyfős Cygron cég (melynek alapítója és többségi tulajdonosa voltam) a kifejlesztett programért '97-ben megkapta az Európai Információtechnológiai Díjat, amelyet Brüsszelben az EU akkori elnöke Jacques Santel adott át. Mivel itthon nem nagyon volt igény ilyen technológiákra, külföldi megrendelésekkel tartottuk fenn a céget. A kilencvenes években, még nagyon nehéz volt fenn maradni, és a cég eladása látszott kiútnak – mert fenn akartuk tartani az eredeti koncepciót. A Cygron amerikai-szingapúri tulajdonba került és továbbra is vezető szerepet játszottam irányításában. A nagy siker ezután jött: 1999-ben, már ebben az új formációban az Év Szoftverdíját kaptuk meg Las Vegasban a COMDEX kiállításon ugyanannak a terméknek a továbbfejlesztett változatáért.

A díjnyertes DataScope koncepció lényege, érdekessége és szépsége abban volt, hogy teljesen újfajta vizualizációt valósított meg. A későbbiekben tovább is fejlesztettük, melynek lényege abban állt, hogy a számítógépen szinkronban és párhuzamosan lehetett látni a különböző folyamatokat. Ha valamelyik grafikonon kijelöltem valamit, az összes grafikon szinkronban kijelölte, mire vonatkozik. A computer – ellentétben az emberrel – gyorsan el tudja végezni ezt.

Majd az amerikai vállalat vezetőségében és az igazgatóságban is részt vettem. Majdnem egy évet dolgoztam még náluk.

Úgy gondoltam, hogy ezeknek a technológiáknak egy másik, már nem féltetlenül az adatbányászathoz kötődő, hanem a koncepciót továbbvivő aspektusával kell foglalkoznom. Ekkor alapítottam az Adixo-t, két évet töltöttem vele. Utána hoztam létre a jelenlegi –

adatelemzésen alapuló optimalizálással foglalkozó társaságot, a Dopti céget. Most már elég jelentősek a sikereink.

Visszatérnék az adatbányászatra. A *data mining* kifejezésnek van egy olyan viszonya a világhoz, mint a fuzzy-nak. Nagyon sokan úgy gondolják ez is egyfajta univerzális eszköz, és fontosnak tartják a mielőbbi bevezetését, de valahogy az alkalmazás során derül ki, hogy nem csodaszer. Jól lehet érzékelni a változást megítélésében: 2000-ben még meghatározó szerepe volt, míg az utóbbi években már csak egy-két cég hirdeti magát data mining-gal. Divatok jönnek, divatok mennek.

A világon rengeteg adatbázis készül. Ha bedugjuk a kártyánkat a bankautomatába, eszméletlenül sok adat generálódik. Pár éve az a mondás járta, hogy nyolc hónapon belül duplázódik meg az adat mennyisége. Lehet, hogy ma már fél év ez az idő. Emberi szem nem látja az adatok 90-95%-át. Nem is olvassuk el, nincs is rá időnk. Ki olvassa át részletesen a számláit? Senki. A rengeteg adat felhalmozásával a tárolás-kapacitás mennyisége, vagy mértéke nem tud lépést tartani, akkor bizony ki kell selejtezni. Ez néha automatikusan is megtörténik, mert a technológiai váltás miatt, nincs idő és energia az új hordozóra való átmentésre, lásd nagy- és kis floppy esete. Mielőtt leselejteznénk a felgyülemlett adatokat végig kell gondolni ki kell-e dobni mindet egyáltalán, értékes része van-e. A felhalmozott hatalmas mennyiségből ki kellene keresni az értékeket. Álljunk neki, hátha találunk benne valamit. Ez az adatbányászat l'art pour l'art megfogalmazása. Igazából nincs sok értelme, mert nagyon nagy az energiaráfordítás. Az adatbányászathoz jó adatok kellene. Adatáruházat kell csinálni ehhez. Egy év biztosan rámegegy, és

aztán kezdődhet az elemzés. A világ pedig gyorsan változik, a korábban értékes adatok esetleg már nem is értékesek.

Az adatbányászattal ellentétben, egy vállalati vezetőnek a folyamatokat kellene látnia. Erre szolgált a DataScope koncepció.

Egy másik újítás lényege, hogy szakítottunk azzal, hogy az időben változó adatokat, a nyomtatás hagyományait követve statikus grafikonokként kell megjeleníteni. A számítógépen vannak játékok, amik mozognak. Miért ne lehetne az adatokat animálni és mozgatni, ha azok időben változnak? Erre dolgoztunk ki eljárást és ilyen jellegű dolgokat is fejlesztünk.

### **- Adatvizualizáció?**

- Mind a cégeknél, mind a kutatásaimban rettentő erős volt az adatbányászat és az adatvizualizáció kapcsolata.

De mi történt az adatbányászattal? Kiderült, léteznek jól meghatározott szegmensei: például a CRM, ami abban áll, hogy vannak ügyfeleim, és őket szeretném elemezni, napra készen akarom tudni, hogyan szólítsam meg őket, milyen ajánlatot tegyek nekik, mert szeretném mindegyiket individuálisan kezelni. Itt a fuzzy-nak megint szerep jut. Az ügyfélről kaphatok egy természetes jellemzést. Tehát, nemcsak a lekérdezés oldalt lehet erősíteni, hanem a mondatgeneráló funkció is munkára fogható. Egy másik hasznos és továbbélő szegmense az adatbányászatnak a Churn-analízis: egy telefontársaságtól átvándorolnak a másik telefontársasághoz az ügyfelek, ami nagy kárt okozhat, különösen akkor, ha új ügyfél már nincs is nagyon a piacon. Tehát erre a negatív folyamatra kell figyelni miután telítődött a piac. Hogyan tudnám az adatokból megjósolni, hogy kik azok, akik a közeljövőben elhagyják a céget? Képes vagyok-e ajánlatokkal,

kedvezményekkel megtartani őket? Úgy tűnik, az adatbányászat – l'art pour l'art jellege helyett annak feladatorientált transzformációja megy végbe napjainkban. Méghozzá azokra a területre kell koncentrálni, amik konkrétan hasznot fognak hozni a következő időszakban.

Az adatbányászat létrehozott és létrehoz eszközöket. Ezek az eszközök – például az osztályozás, a tanulás koncepciója, a vizualizáció koncepciója – olyanok, mint egy nyomozónak az eszközei: a nagyító, a háttér labor stb., és amikor adott egy konkrét feladat, akkor ezek az eszközök bevetethők, és egy szakember segítségével megoldhatóvá válik a feladat.

A statisztika és az adatbányászat egyaránt adatelemzéssel foglalkozik. Az előbbi attól szenved és az a fő problémája, hogy egy szűk minta áll rendelkezésre, és abból kell következtetéseket levonnia. (Lásd a gyógyszer-hatékonysági vizsgálatoknál a humán tesztek esete.) Az adatbányászatnál pont fordított a helyzet, mert például a telefonbeszélgetésekről terrabájtnyi adatunk van. Hogyan kezelhető ez a hatalmas adatmennyiség? Az algoritmus hatékonysága nagyon lényegessé válik. Egészen más technológia szükséges a két területhez.

### **- Hogyan kapcsolódnak a fuzzy és az adatbányászat terén végzett kutatásai az MI-hez?**

- Úgy látom, hogy azok az eljárások, melyekkel az MI foglalkozik – genetikus algoritmusok, hangyaalgoritmusok, a neurális hálózatok legkülönbözőbb paradigmái – heterogén és szerteágazó területek. Mint amikor a fizikában különböző dolgokat fedeztek fel, viszont nem sikerült kialakítani egységes

elméletet. Ez várat még magára. Reménykedem benne, hogy többek között a pliant segít ennek a létrehozásában. Bizonyos dolgokat sikerült felfűznöm rá: klaszterezési algoritmusokat, a fuzzy egyes részeit, neuronmodelleket. A formális hasonlóság mellett nehezebbnek tűnik a közös szemantikus gyökér megtalálása.

Az MI mindig olyan tudományterület volt, hogy amikor valami újba, vagy új megközelítésbe fogott az ember, rögtön azt mondta: ez az MI. Az új diszciplínák mindig belekerülnek ebbe a nagy bugyorba, aztán – miután megerősödtek – leváltak. Kicsit hasonló a helyzet, mint a filozófiában. Az is egy hatalmas olvasztótégely, aminek megvan a saját szerepe. De ebből az is kiderül, hogy állandóan változik mindkét diszciplína. Az egyetemek MI tanszékei nehéz helyzetben vannak, mert minden oktatásnak bizonyos értelemben kanonizálnak kell lennie, ami elég nehéz az állandóan változó tárgyak esetében. A nyolcvanas évek közepe, vége felé az akkori kutatók, oktatók nagy része úgy döntött, meg tudja határozni mi az MI, és azt mondták az MI a szakértői rendszerekkel és a logikai programozással azonos. A rendszer megmerevedése lett a következmény és nem volt hajlandó válaszolni az újabb kihívásokra. Ezért fordulhatott elő, hogy se a fuzzy, se a genetikusan algoritmusok nem kerülhettek be sokáig az MI körébe. De a neurális hálóktól mind a mai napig teljesen elzártak az MI folyóiratok.

Elmondhatom, hogy 2000-ben Jelasity Márk PhD-hallgatómmal sikerült egy genetikusan algoritmusmal foglalkozó cikket elfogadtatni a *Mesterséges intelligencia* folyóirattal és ez volt az első ilyen publikáció, ami itt megjelenhetett.



Mára már a szakértői rendszerek, a Prolog nyelv is leváltak az MI-ről. Külön tárgyak keretében oktatják. Ugyanúgy történt, mint a programozási nyelvek, vagy az alakfelismerés esetében. De a példák sorolhatók tovább. A Szegedi Egyetemen az MI-t két lépcsőben oktatjuk. Az alapozó kurzusban a logika, a játékok és a keresési stratégiák szerepelnek, a második lépcsőben az új diszciplínák jelennek meg: fuzzy, genetikus algoritmusok, neurális hálók, döntési fák világa adatbányászattal és vizualizációs eljárásokkal. De helyet kapnak többtényezős döntések eljárásai is. Ezt az általánosabb koncepciót próbálom továbbadni a hallgatóknak.

**- Említette a neurális hálókat. Hogyan látja ezeknek a kutatásoknak a jelenlegi helyzetét?**

- A genetikus algoritmusokhoz hasonlóan szintén a biológiához fordulás érhető itt tetten. Nem nagyon fogjuk fel mennyivel nagyobb teljesítményű az agyunk, amikor azt látjuk és tapasztaljuk, hogy a számítógépek segítségével Interneten óriási virtuális lexikonokat tudunk kezelni. Erre az agyunk képtelen. Vagy az áruházakban a számítógépek pillanatok alatt összeadják az ötszáztizenkét cikk árát. Bizonyos területeken az emberi agynál gyorsabbak a számítógépek, ám az agy hatékonysága nagyságrendekkel nagyobb. Például a szem esetében az adaptáció: a szem és a digitális kamerák között százezres nagyságrendnyi az eltérés a szem javára. A másik példám a szúnyogról szól: csak mikroszkóp alatt érzékelhető milyen kicsi is a feje. Úgy gondoljuk, a neurális hálózata se túl nagy egy számítógép processzorához képest. Viszont rendkívül jól tud manőverezni, akárhova leszáll, megfordul, képes menekülésre

stb. Van látórendszere, szaglórendszere, speciális fűró rendszere, amit működtet. Ha az általunk létrehozott legmodernebb helikopterhez hasonlítom, azok jelenleg sincsenek hasonló fejlettségi szinten. Az energiafelhasználásról nem is beszélve. Rettentő nehéz feladat lenne egy szűnyog funkcióinak megvalósítása számítógépen futtatott algoritmusokkal. Azaz, a valódi neurális rendszerek nagyon jók, nagyon hatékonyak, energia igényük se nagy és ezért érdemes ellesni a természettől, hogyan is működnek. A kutatásaimhoz ez is kapcsolódik. Itt is digitális jelek, impulzusok vannak, amik igazából ingerületeknek felelnek meg. Csomópontokban kapcsolódnak össze és bizonyos műveletek elvégzése után továbbítódnak. A valódi hálózatok komplexitása óriási. Ez azonban még mindig nem ad magyarázatot arra, hogy miért ilyen hatékony a gondolkodásunk, látásunk, érzékelésünk. Képzeljünk el egy óriási kapcsolási rajzot: ha fel-le kapcsolatok ebben a rendszerben, mindig mechanikus dolgok történnek. Tehát kell lenni valami másnak is, ami ezt a rendkívüli plaszticitást létrehozza.

Astrogliának nevezzük az agynak egy másik struktúráját, ami az idegsejteket szintén összekapcsolja. Több mint négy éve érdekel ez az idegsejt környezet. Itt analóg folyamatok mennek végbe. Koncentrációk változnak, amik visszahatnak az impulzusokra. Az agy és a neuronok hatékonysága lehet, hogy azzal magyarázható, hogy kétfajta – egy analóg és egy digitális – rendszer szimbiózisaként működik. Az analóg jellegű számítások nagyon masszív párhuzamos számítások, melyek a molekulák szintjén történnek, míg a digitális a neuron hálózat impulzusai. Az alacsonyabb rendű élőlényeknél alig fordul elő glia. Itt ösztönszerű „behuzalozott” működésről van szó. A

magasabbrendű élőlényeknél viszont annál nagyobb szerepet játszik a glia sejt, melynek száma a neuronénak többszöröse. Az érdekel mostanában, hogy miképp lehetne ilyen jellegű rendszereket készíteni. Az analóg számításnál nehéz a peremfeltételek megadása. Utána viszont pillanatokon belül végbemennek a számítások. A digitális eljárások lassúak, de az inputot könnyű megadni. Egyfajta komplementeritást tapasztalhatunk, ha összehasonlítjuk az analóg és a digitális világot. A gondolkodásban szintén létezik ez a kétfajta megközelítés. Az egyik esetében hosszan állítgatjuk az inputot aztán elengedve a folyamatot kapunk egy jó eredményt. A másik esetében megadjuk az inputot és hosszan várunk, míg az algoritmus kiszámolja az eredményt. Ma mindenki a digitalizációról beszél, ám a számolások szempontjából úgy látszik, vissza kell térni az analóg világra. Mert ott pillanatokon belül kaphatunk eredményt. A robot nem tudna átkelni az úton, ha minden szituáció elemzésére 1 órát kellene fordítania.

**- Miben látja a kutatásfejlesztési projektek sikerének, eredményességének titkát?**

- Legfontosabb, hogy egyáltalán azok is kapjanak projektet, akik innovatív újszerű megoldásokat képesek létrehozni. A pályázatok elbírálása mindenképpen kritikus pont ebből a szempontból. A tapasztalatok inkább negatívak. A kevés pénz se megfelelően kerül elosztásra.

Valós problémák megoldása a siker másik fontos tényezője. A jelen feszítő kérdéseire, ha nincs tudományos válasz, akkor jó megoldások helyett legjobb esetben kevésbé rosszak születnek.

A baj ezekkel nem is az, hogy messze vannak az ideálistól, hanem hogy a jövőre is rányomják a bélyegüket.

Azonban megoldandó feladat túlságos specializált volta sem kedvez a sikernek. A siker ne legyen kérész életű. Egy olyan kutatási megoldás, amiről látszik, hogy pár év múlva más technológia helyettesíti nem szabad tudományos eredménynek tekinteni.

Végezetül a siker segít, de az, aki sikerorientált elfeledkezhet igazi feladatáról. A siker lehet fontosabb az eredményes kutatásnál? Igen, és ez baj. Tenni a dolgunkat a feladat és a siker, elismerés legyen ráadás.

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

- Egyetemista korom a Z 80-as gépek kora, tehát azok a dolgok, amik ma fontosak nem is jelenhettek még meg. Ha most lennék egyetemista, akkor a mostani kutatási területem nagyon is megfelelné. De azért meg tudnék nevezni más területeket is. A logika numerikus módszereit szívesen vizsgálnám. A kognitív pszichológia eredményeinek matematikai alkalmazása is érdekelne. Az evolúció modelljei, különös tekintettel a komplexitás növekedésének kérdése szempontjából nagyon érdekes terület. Az analóg és digitális folyamatok kölcsönhatása alapján működő rendszerrel be is fejezném a felsorolást.

**- Milyen alapelveket érdemes képviselni ahhoz, hogy a csúcstechnológiai kutatásokban komoly eredményeket érjünk el?**

- Azonos területen dolgozó emberek összefogásának elérése. Mindenki jól tudja ki, illetve kik azok, akik jók az adott területen, ahol dolgozik. Egy közös cél érdekében való összefogás elérése mégis szinte lehetetlen. A projektek kikényszeríthetnék az összefogást, mint ahogy egyes országokban ez működőképes is, de ez nem megy itthon. (Még az EU-ban sem megy ez nagyon.) A pénzek megszerzése a lényeg és kutatási jelentéseken való keresztülvergődés után kezdődik a következő projekt összekalapálása.

Maga a pályáztatás se jó. Például Kolumbusz Kristóf hogyan fedezhette volna fel Amerikát, ha a Király attól teszi függővé támogatását, ha pontosan kitölt egy pályázati csomagot, megfelelő gazdaságossági számításokkal, megtérüléssel stb. Mit kell arra válaszolni, hogy: Mennyi ideig tart az út? Mennyi arany és egyéb érc várható?

Az adminisztráció sok értékes időt vesz el az effektív kutatástól. A jelentések valóságtartalma pedig nehezen ellenőrizhető. Csúcstechnológiában koncentrált erőforrással nagyon jól kiválasztott projektekkel lehet sikert elérni.

Érdemes lenne az utolsó 10 év pályázatainak utóéletét megvizsgálni. A sikeresség szempontjából a bírálókat, az értékelőket minősíteni, nem utolsó sorban pedig magukat a résztvevőket újra értékelni.

<http://www.inf.u-szeged.hu/~dombi> (Dombi József)