

A törzs és az ágak 1.

(Horváth Tibor kandidátusi értekezésének bemutatása)

Darányi Sándor

1. Bevezetés

Az alábbiakban néhány technikai megjegyzéssel szeretném segíteni az olvasót, mielőtt kísérelőjeül - talán kalauzául is - elfogad. Dolgozatomba a fejezetcímeket gyakran az értekezés szerzőjétől vettem át, elhagytam viszont a hivatkozásait. A téma iránt érdeklődő ezeket az eredeti munkában úgyis megtalálja.* Ami azonban a szerkesztéssel járó kisebb stilizálásnak indult, az gyakran *Horváth Tibor* gondolatainak erős tömörítésévé alakult át. Ezenközben gondolatmenetét igyekeztem nem csorbítani. Ha saját szubjektív széljegyzeteimet (a szövegben kurzívval szedve) mondandójától nem mindig sikerült elválasztanom, azért elsősorban tőle kérek elnézést.

2. Az értelem mint függvény

A tökéletes rend álma bizonyára az írásbeliséggel egyidős. A könyvtár mindennapos tevékenysége, az osztályozás megvalósulása révén kerül hol távolabb ettől az eszménytől, hol közelebb hozzá. Ezért a tökélyre való törekvés leginkább az osztályozás fejlődésében érhető tetten.

Horváth Tibor műve méltón képviseli ezt a hagyományt. Mint értekezése alcímében közli, ismeretszervezésről és szintaxisról ír, gondolatmenete azonban ennél jóval többet tartalmaz és ígér. Az ismeretszervezésbe ugyanis belehallszik nemcsak az, hogy ma még az ismeret-, azaz az információfeldolgozás korszakát éljük, hanem az is, hogy erről a tudásfeldolgozásra fogunk előbb-utóbb áttérni – márpedig tudáson éppen az igaznak tudott állítások, ismeretek strukturált összességét értjük. Elegendő a „szervezés” szót a strukturálással helyettesítenünk, máris megsejtjük, mekkora, az osztályozástól egész a szakértői rendszerekig nyíló távlatokat hordoz ez a szó.

Mi a szintaxis szerepe az ismeretszervezésben? Egyelőre elégedjünk meg azzal, hogy bármilyen mű kulturális termék, így a nyelv is közlésre szolgál, mégpedig értelmes kijelentések közlésére. Az értelem, me-

* Lásd: Horváth Tibor: „Egy törzsnek az ága.” Ismeretszervezés és szintaxis. Kandidátusi értekezés. 1987. 164 p. (Megtalálható a KMK Könyvtártudományi Szakkönyvtár állományában.)

lyet másképp az adott kód szemantikájának hívunk, nem ugyanolyan független komponense a közlésnek, mint annak szókészlete (lexikája) vagy mondattana (szintaxisa). Utóbbi, mint a nyelvtan (grammatika) megtestesítője felel azért, hogy a szókészlet elemeit különféle viszonyegyettesekbe rendezve, a közlés értelme kibontakozzék. A szemantika tehát a természetes nyelvben e viszonyegyettesek, tágabban a szintaxis függvénye, ezért a mesterséges nyelvek, ezeken belül az indexelés elméletében is kitüntetett hely illeti meg. Az így alkotott indextétel „analógiával élve: „szavak, halmaza, összessége helyett” mondat, szintűvé válik” (Horváth 1978b:10). Az értekezés ezt a gondolatot kíséri következetesen végig, az elméleti alapvetéstől a gyakorlati alkalmazásig.

Így a mű lényegében egy osztályozástörténeti vonulatra, és ennek a vonulatnak a logikáját megindokló problémafelvetésekre, illetve -megoldásokra tagolható. Ezekhez csatlakozik a PRECIS mélyindexelő nyelv szintaxisának megfeleltetése a magyar nyelvtan követelményeinek.

3. Az értekezés célja

„Értekezésem célja az osztályozáselmélet egy fontos gondolatának kifejtése és igazolása” – Horváth Tibor ezekkel a szavakkal kezdi gondolatmenetét. – „Az indextételnek nemcsak az a feladata, hogy az osztályozandó objektumok változatos ismérveit megadja, hanem az is, hogy az indexelő kifejezéseket olyan magasabb egységben szintetizálja, amely az eredeti közlemény gondolatának, mondandójának tömör reprodukálását is eléri. Oly módon, ahogyan a természetes nyelv szavait a grammatika mondatokká kapcsolja össze. Az osztályozó/indexelő eljárások legtöbbször az általánosan elterjedt és elfogadott modell értelmében az objektumokhoz a kifejezések egy részalmazát rendeli. Ebből az eljárásból hiányzik a szintaxis, ugyanis a bevezetett relációk érvénye a kifejezéseket tartalmazó szótárra korlátozódik, a relációk ennek folytán nem képesek olyan szintaktikai funkciók betöltésére, amelynek révén a szerző által az eredeti szövegben létrehozott fogalmi kapcsolatok az indextételekben megőrizhetők.”

4. Az osztályozás feladata

„Két feladatban szokás megjelölni a könyvtári (beleértve a dokumentációs, információs vagy informatikai) osztályozás célját” – folytatja a szerző. – Az első az egészként vagy egységként tekintett tudománynak, tágabban az emberi ismereteknek részeire vagy specicsereire való bontása, felosztása. A második az individuumok, tények, eszmék, elméletek besorolása egy átfogóbb fogalmi osztályba, amelybe valamelyik jellemző tartalmi ismérvük alapján tartoznak. Szigorúan az utóbbit lehet osztályozásnak nevezni, amennyiben az osztályozás nem más, mint osztályba sorolás. Az első feladat szabatosan a felosztás logikai terminusával fejezhető ki. Ennek alapján az információnak (vagy a szakirodalomnak, mint rögzített információnak) osztályozási elmélete a taxonómia általános elméletéből vezethető le. Az utóbbinak egy, a könyvtári-informatikai területre szülő alkalmazásaként határozható meg.

Ez az általánosan elterjedt álláspont nem elegendő a teljes probléma felvetéséhez, mert éppen az utolsó két-három évtized eredményei, elméleti megfontolásai és megállapításai esnek ki abból a keretből, amelyet a fenti megközelítés enged meg. Létezik ugyanis egy harmadik feladat, amely problémaként mindig is fennállt, de az adatbázisok kapcsán különös élességgel bukkant felszínre. Ez pedig az, hogy nemcsak a tudás felosztásáról van szó, nemcsak objektumok, individuumok osztályozásáról, hanem minden olyan megállapítás, ismérv, tulajdonság meghatározásáról és leírásáról, amely valamely objektumról, individuumról, elméletről, eszméről állítható. Ezek az ismérvek, tulajdonságok nevezhetők predikátumoknak.

A klasszikus logikán nyugvó osztályozási elméletektől elszakadni igyekvő osztályozási koncepciók megjelenését – és nyomukban a gyakorlatét – az ötvenes évekre tehetjük. Kibontakozásuk azonban a hatvanas évek eredménye. Az individualizáló „osztályozás” vagy indexelés megjelenése – amely a predikátumok kezelését biztosította – nem tette feleslegessé a klasszikus, felosztás elvű osztályozást sem. Az első két feladat is megmaradt. Nem arról van szó tehát, hogy az indexelő eljárások vagy az információkereső nyelvek felváltották a klasszikus, hagyományos rendszereket, hanem arról, hogy a módo-

sult, vagy új indexelési, osztályozási feladatok új elméletet és gyakorlatot követeltek és hívtak életre, megteremtve az eljárások koegzisztenciáját.

A mai teljes feladatrendszer szolgálatában álló törekvések és elméletek a következő utakat járták, vagy inkább járják.

1. Megmaradtak a hagyományos tudományfelosztáson nyugvó rendszerek, de újak konstruálódtak a régi elveken, keresvén a kor tudományosságát jobban érvényesítő felosztásokat (kidolgozásra került a BSO, *Dahlberg* ontikus rendszere, a legmagasabb szintű rubrikátor). Ezzel párhuzamosan a gyakorlatban korábban bevezetett hagyományos rendszereket minden eddiginél radikálisabb reformmal igyekeztek korszerűsíteni (ETO).

2. A klasszikus logika alapján álló, de a felosztás-osztályozás merevségét oldani kívánó eljárások születtek azzal, hogy a kategóriaelmélet hagyományait elevenítették fel. Eredményként üdvözölhetjük a facettás vagy fazettás osztályozásokat, vagy a facetták beépítését különböző más osztályozások eszköztárába. Ebben az esetben is még mindig osztályozásról van szó a kifejezés szigorú értelmében.

3. A statisztikai elmélet és indexelő eljárások abban kerestek megoldást, hogy az objektumok és a *term*-ek halmazán értelmezett leképezéssel létesítettek kapcsolatot az objektumok és a *term*-ek között. Objektumon bármely fizikai létezőt, élőlényt, dokumentumot, szöveget lehet érteni, a *term*-ek pedig a predikátumok halmazát jelentik (*terms, properties, features*). Ezt az iskolát G. Salton tökéletesítette a klaszterálásnak (*clustering*) az információkeresésben megvalósuló bevezetésével.

4. A nyelvészeti iskolák a predikátumot úgy fogták fel, mint olyan állítmányt, amely valamely objektumról elmondható. Egy indextétel valójában mondat, amelyben a *term*-ek (tárgyszó, vagy bármilyen osztályozási kifejezés) szerepét a nyelvtani esetrendszerrel analóg viszonyokat tükröző ún. operátorok tüntetik fel. Az ebben az esetben létrehozott minden rendszert egy szintaxis definiál, amely az osztályozási kifejezések bármely gyűjteményére alkalmazható (PRECIS). Ezek a mondatok azonban nemcsak a kifejezések halmazai, dokumentumokhoz rendelt csoportjai, hanem új minőségűek, amennyiben az osztályozási kifejezések szavait olyan mondatokká fogják össze, amelynek jelentése más, mint amit a szavak összessége hordozni képes.”

5. Honnan hová

5.1 A tudáspiramis

A következőkben egy igen fontos megkülönböztetésre kerül sor. Előbb a klasszikus rendszerekkel foglalkozván, az osztályozás két lehetséges vezérfonalát határozza meg az értekezés. Eszerint a filozófiai tudományrendszerek az arisztotelészi klasszikus logika felosztáselvén működnek: felosztani annyit tesz, mint az egészet valamely - lényeges - elv alapján részre bontani. De mi olyan lényeges, hogy a felosztás alapja lehessen? Az arisztotelészi módszernek ez a gyöngye, vitákat provokáló pontja hamar ismeretessé vált.

„Az egyik póluson tehát megjelenik a tárgy felosztás arisztotelészi elve. A másik pólust a lehetséges felosztási alapok megválasztásához – teljes kifejllettségében – a kibernetika nyújtja, azaz a nevezetes tételével, hogy igen nagy rendszer vagy bonyolult tárgy vizsgálatának eredményei nem függetlenek a megfigyelőtől.

A megfigyelő álláspontja különböző szempontokban jelenik meg felosztási alapként; feltűnik, mint megismerési fokozat, vizsgálati módszer, uralkodó tudományos paradigma, ideológiai nézőpont.

A kibernetika és a rendszerelmélet csupán egzakt szaktudományi igazolását adják annak, hogy a megfigyelő nézőpontja is lehet felosztási alap; azzal nyújtják ezt, hogy cáfolják azt a sokáig uralkodó kartézianus dichotómiát, amely az anyagi világot, mint kívülről függetlenítette a megismerő tudattól. Ennek ellenére a megfigyelő nézőpontja jóval korábban megjelenik felosztási alapként. Valószínű, hogy *Francis Bacon* volt az első, aki ezt érvényesítette. Tudományrendszerét három intellektuális képesség, az emlékezés, képzelet és értelem alakítja.”

Bacon egyébként a tudományrendszertannak azt az igényét, hogy a világ a művön keresztül egységes egésznek lássék, így fejezte ki - s innen a *disszertáció címe is* - : „minden tudomány ága egyazon fa törzsében találkozik” (*Horváth 1987:10*). A fa törzse tehát, s így a piramis csúcsa a filozófia lett.

Az újkantiánus Rickert óta a tudományokat tárgyuk és módszerük szerint, kétdimenziósan volt szokás felosztani, majd az osztályozás finomítását a felosztási alapok számának növelésével igyekeztek elérni. Mind-

ez már az összetett problémák sok, nem pedig kevés ismérv szerint való kezelése felé mutatott.

5.2 „Nincs világgép”

Jogosult-e azonban a tudáspiramis koncepciója, és ha igen, mennyiben?

„Aki a tudományokat hiererchiába rendezi, eleve feltételezi, hogy a valóság is ilyen szerkezetű.” Erre aligha található igazolás.

Lehetséges, hogy mégis a filozófus feladata nehezebb. Az ő vezérelve az, hogy egységes elméletbe foglalja a világ totalitását, elméletébe építve a tudományokat is. Világgépet kíván alkotni. És ez az, ami miatt a filozófia mindig is meg fogja kísérelni a tudományokat oly módon rendszerezni, hogy azok világgépébe illeszkedjenek.

Az pedig egészen más kérdés, hogy a filozófiának ez az ambíciója hogyan valósul meg, vagy megvalósulhat-e egyáltalán. Pillanatnyilag nincs egységes elmélet, világgép, még egy tudományon belül sincs. Nincs a világról szintézise a fizikának, az élettudományoknak. A kutatás a részeknél tart. A filozófus számára a világgép híján annál inkább szükség van a tudományrendszer-tanra, mert ezek szerint ez az egyetlen eszköz, ami az emberi tudás egészét keretekbe fogja, megteremtve a részeknek és az egésznek összefüggéseit valamely elv alapján. Talán *Jaspersnek* van igaza. „Nincs világgép, csak a tudományok rendszere van.”

6. Új struktúrák

A második világháborút követően nagy gyorsasággal szaporodtak azok a tudományos eredmények, amelyek nem voltak beilleszthetők a kialakult tudományrendszerek egyikébe sem. Sőt egész diszciplínák születtek, amelyek szintén nem voltak elhelyezhetők. A jelenséget általában az interdiszciplináris területek megjelenéseként, illetve a tudományok integrálódásaként szokták magyarázni. Okának azt tartották, hogy a tudomány a tárgyaként adódó valóságot egyre részletesebben vizsgálja, a konkrét kutatás egyre kisebb, szűkebb témákat elemez, ezeket azonban teljes komplexitásukban.

A jelenségnek mélyebb okai vannak. Valóban a határterületi diszciplínák sokasodása jelzi ezt a folyamatot. Velük párhuzamosan azonban olyan új tudományok is kialakulnak, amelyek nemcsak azzal minősíthetők, hogy a közös tárgyat, jelenséget több tudomány közelíti, és ebből előbb-utóbb önálló arculatú, új, a tudományhatárok metszetében álló diszciplína keletkezik. Hanem azzal is, hogy ezek az új „tudományközi” diszciplínák a tudományos modellek létrehozásában a kereszteződő tudományok lehetőségeit meghaladó, eltérő, azokéból közvetlenül nem származtatható új eszközökkel mélyítik el a megismerést.

Az a tény, hogy a tudományok az antik kortól a 20. századig a tudományrendszerek tanulsága szerint tárgyukból határozták meg magukat, fel-fedi e harmadfélezer éves tudomány természetét. A „tárgy” ugyanis az érzékelhető valóság volt akkor is, ha az érzékelhetőség többszörösére tágult a műszerek segítségével. A tudományoknak ez a felállása lehetővé tett egyfajta absztrakciót, amely egyben a megismerés mélységét határozta meg. A lényeg megismerésében addig lehetett elérni, amennyire a tudományos absztrahálás lehetővé tette.

A második világháború után kialakult tudományok jelentős részének – nem előzmények nélkül – olyan tárgya, vizsgálódási területe van, amely nem az érzékelhető valóság feldarabolásából keletkezett. Tárgyuk olyan entitás, amely már emberi absztrakció eredménye. Ez az entitás nem köthető az érzékelhető világ egyik részéhez sem, hanem jelen van a valóság teljességében.

Több olyan új tárgya van a tudománynak, melyekre egyaránt igaz, hogy mintegy „felülről” fedik le az eddigi diszciplínákat, lévén azoknál elvontabbak. Ezek közül is első a rendszerelmélet.

A rendszer olyan, sok komponensű entitás, melyben szükségszerű kölcsönhatások zajlanak le. Vizsgálata új lehetőségeket teremt, mindenekelőtt a tudományok ún. egzaktági dilemmája tekintetében. Általánosságban az a tudomány egzakt, amely minél több változó viselkedését jósolja meg pontosan. Az eddig megismert természeti törvények két-három komponensűek, ez alól Einstein híres $E=mc^2$ képlete sem kivétel; ha pedig ennél bonyolultabb, több összetevőjű problémát vizsgáltak, le kellett mondaniuk a pontosságról. „Tehát vagy egzaktság, de kevés változójú törvények, vagy sok változó, de egzaktság nélkül”. Bertalanffy ezért tartotta sokváltozós jelenségek egzakt vizsgálatát a rendszerelmélet egyik fő feladatá-

nak. Emellett az foglalkozik még a különböző tudományok törvényeinek izomorfiájával is, keresvén bennük azt, ami közös. A szintézishez nem vezethet más út.

Hasonlóan átszervező, másként láttató szerepe van a kibernetikának, a kommunikációnak, a szemiotikának és az információelméletnek is. E diszciplínák tárgya nem a tapasztalható valóság, hanem annak tudati leképezése; karöltve a tudományok fokozódó formalizálásával, rendkívül elvontak, és modellezésre egyaránt alkalmasak. Ami tehát a részterületek integrációjának tűnik, az a harmadik síkba helyezett tudományok visszatüremlése a klasszikus területekbe.

„A fentebb elmondottakat – nem igazolás-ként, csak példa gyanánt – anekdotikusan is alá lehet támasztani. A *Tudomány* c. folyóirat érdekes interjúja közölte a hazánkban járt B.B. Mandelbrottal, akiről az igen érdekes tulajdonságú Mandelbrot halmazokat elnevezték. Az egyik kérdés úgy szólt, hogyan lehetséges, hogy életének különböző periódusaiban más-más egyetemeken tanított, először nyelvész volt, majd közgazdaságtant oktatott, fizikusként a turbulencia jelenségét kutatta, foglalkozott a Nílus és más folyók viselkedésének szabályszerűségeivel, nevét jegyzik a biológiában, eredményei vannak a matematikában (eredetileg matematikus), jelenleg az IBM-nél kutató. Mandelbrot válaszában lényege az volt, hogy egész életében ugyanazt csinálta. Rátalált egy összefüggésre vagy törvényre, felfedezte a fraktálokat. Nem tehet arról, hogy ezzel a legtávolabbi tudományok problémái magyarázhatók. A válasz annyira frappáns, hogy az itt tárgyalt jelenséget akár Mandelbrot-effektusnak is lehetne hívni.”

7. A megfigyelő nézőpontja és a facettás rendszerek

Az osztályozási eljárások másik hagyományos vonulatának a facettás rendszerek tekinthetők. Ezek egyfelől Arisztotelész kategóriaelméletéből, másrészt

– Ranganathan kettőspontos osztályozása révén – a hindu filozófiából erednek.

A fogalmi kategóriák olyan állítások, amelyek a mondattól függetlenül mondhatók a szavakról: a lét struktúrában a létezés legáltalánosabb módjait adják. Céljuk az, hogy a létezőket a legáltalánosabb genusok alá lehessen sorolni. „Ebben mutatkozik meg kapcsolatuk a felosztással. A különbség az, hogy a

kategóriák nem eredményeznek a felosztáshoz hasonló hierarchiát, hanem a fogalmak természetét mutatják, hogy mi módon vannak önmagukban” (Horváth 1987:28). Az osztályozáselmélet a létezés legáltalánosabb kategóriáit nevezi vetületeknek (angolul face). Innen az elnevezés.

„Az ötvenes és hatvanas években úgy tűnt, hogy a facettás osztályozás megoldást nyújt számos elméleti kérdésre is. Ennek a két évtizednek erőfeszítései ma már lanyhultak, mert új paradigmák, a statisztikai és nyelvészeti alapozású eljárások tűntek fel. Ezek visszaszorították a facetták gondolatát. Nem mindörökre, mert ebben az eljárásban maradandó elvek is vannak. Maga az alapgondolat, a kategóriák elmélete harmadfélezer éves. A facetták örökös társul szegődtek az osztályozáshoz. Fel-feltűnnek mai rendszerekben is. Sikerüket és maradandóságukat annak köszönhetik, hogy összekötik az eltérő elvű rendszereket. A facetták néha alig különböznek a felosztástól.

- A kettőspontos osztályozásban – a kettőspont egyik elválasztó jel a facetták közt, innen az elnevezés – már megfigyelhető a facettás osztályozások néhány jellegzetessége.
- Feltételezik, hogy valamennyi osztályozandó téma azonos szerkezetben írható le.
- Sorrendet írnak elő; e sorrendben a facetta funkciót határoz meg, tehát egyfajta szintaktikai feladatot lát el. Az osztályozási kifejezések különböző facettákban jelenhetnek meg. Pl. az arany lehet M (*matter*=anyag), de lehet P (*personality*=egyediség, karakter) is, a kettőspontos osztályozó szerint.
- Feltételezik az osztályozási kifejezések valamilyen gyűjteményét (tárgyszavak, deskriptorok jegyzéke).
- Az osztályozót rákényszerítik a téma elemzésére, mintegy kikérdezik, facettáról facettára haladva.
- Pontosítják az osztályozási kifejezések jelentését.”

Most tehát az eddigiekben többször előbukkant vezérgondolat, a jelenségek sokdimenziós, sok változó szerinti osztályozása mellett megjelenik a másik. Egy állandó struktúrán, a facetták változatlan sorrendjén belül a facetták egymáshoz való viszonya testesíti meg a szintaxist, a tetszőleges facettába „tölthető” osztályozási kifejezések felelnek a szemantikáért. A facetták tehát az arisztotelészi forma, a kifejezések a

szubsztancia szerepét játsszák. Horváth Tibor értékelése az, hogy a facettás rendszerek egyfelől a hagyomány rugalmas összegzését eredményezték, másfelől a nyelvészetben gyökerező elméletek kiindulási alapja is lettek. Ennek bizonyítására két rendszert ismertet.

Farradane 1950 és 1955 között pszichológiai alapokon két fogalomnak az ideiglenestől a megszilárdult együttes előfordulásig tartó kapcsolata, valamint e kapcsolat érzékelésének elmélyülése szerint három-három szakaszt, összesen kilenc fokozatot különböztet meg. Így olyan, nyelvtől független operátorokhoz jutott, amelyek alkalmasnak mutatkoztak nemzetközi indexelési eljárások kimunkálására. Ezek az operátorok az értekezésben még visszatérnek.

Jean-Claude Gardin régészeti leletanyag számítógépes feldolgozása során a facettákból indult ki, de a szintaxis következetes alkalmazásával elődeinél jóval tovább jutott. A leletek morfológiai leírása után ugyanis az ornamentikát elemezte, s ekkor döbbsent rá, hogy nem az individuális díszítő elemet, hanem a kompozíció egészét kell látni. „Ahogyan a nyelvben az egyes szavak fokozatosan alakulnak át szintagmákká és mondatokká, ugyanúgy alakulnak át az ornamentika elemei a teljes jelentést hordozó kompozícióvá. Ez azonban már nem morfológiai elemzés, hanem szintaktikai” (Horváth 1987a:37). Ezt egészítette ki az ikonográfiai leírás, az alany és a tárgy megkülönböztetésének problémájával. Ha ugyanis egy vázán oroszlánt és embert látunk, egyaránt állítható, hogy „az oroszlán megeszi az embert”, illetve hogy „az ember megeszi az oroszlánt”, nem mindegy azonban, a „megesz” reláció kire irányul. Gardin a probléma megoldására öt esetet ismerő ragozást vezetett be, végigvívva a nyelvi analógiát.

„Valójában akkor derülhet ki igazán, hogy felosztással vagy fogalmi kategóriákkal van dolgunk, ha a részletek is kibomlanak. A facettás rendszerek nem hoznak létre hierarchiát. Akkor sem, ha egy-két szinten strukturáltak. Arisztotelész gondolata, hogy általános genusok alá rendelje a létezőket, a felosztás gondolatát sugallja.

Másik oldalon a facetták kapcsolatot teremtettek a fogalmi relációkkal. Az ő kibontásuk relációkat – közöttük hierarchikus relációkat is – eredményez.

Azok a nehézségek, amelyekkel a klasszikus tudományrendszerek küszködnek, itt nem merülnek fel. Mert a valóságot nem akarják feldarabolni és sémákba erőltetni, hanem a valóság egy-egy arculatát kívánják felmutatni. Ennyi is elég.

Gardin gondolata pedig – szintaxisára gondolva, tehát arra, hogy az osztályozásban, leírásban szerepet kapott tárgyilag összetartozó elemek egy teljes egészzé álljanak össze, már egy nyelvhez vezet el, amelyet információs nyelvnek szokás nevezni.

A facetták joggal tűnnek fel tehát azokban az új próbálkozásokban is, amelyeket a felosztások utolsó mohikánjainak hívhatunk.”

8. A jelen alapproblémái

Az osztályozási rendszerek történetében itt rövid kitérőt teszünk. Ennek során az információkereső nyelvek egységeinek megválasztása lesz az egyik fő kérdés, e kifejezések szaknyelvi pontossága, specifikussága a másik. A szintaktikus „alkatrészek” formai és tartalmi meghatározása után kerülhet sor a közöttük lehetséges kapcsolatok, mindazon relációk tárgyalására, melyek a közlés formai-tartalmi oldalát mondat-szintű, nagyobb szemantikai egységgé szervezik.

8.1 Kvantálás

Amit sommásan a tudomány fejlődésének nevezünk, az a valami a formalizálhatóság függvénye. A fejlődés mint művelet sor kitarotán keres egy univerzális függési rendszert, függő és független változókat, a formalizálás pedig nem más, mint ehhez a kiindulási változók meghatározása. Ilyen változókat könnyebb találni a természettudományokban, mint a bölcsészetben, ezért haladnak az előbbieket látványosabban ugyanazon eszköz révén. Ám ettől még a formalizálás első számú problémája egy és ugyanaz marad: hogyan szegmentáljunk? Melyek lesznek azok az egységek, melyekre vonatkozóan aztán a „mi függ mitől” kérdését feltehetjük?

Mielőtt továbbmennénk, ezen a ponton rögzítsünk néhány fogalmat. Mint arra Horváth Tibor is utal, az információ kvantitatív egysége a bit. Ezzel az egységgel azonban a dokumentumok tartalmának leírásához nem jutunk közelebb. Mivel a tartalom egységeivel szeretnénk műveleteket végezni - Edmund Leach szavaival: holmi szemantikai algebra ismeretleneit keressük -, a tartalmat nem bitekben, hanem nagyobb,

ám gondolkodásunk számára még mindig elemien kicsi struktúrákban szeretnénk megtalálni.

A tartalmi információ e kvalitatív egysége lehet pl. néhány, erősen kapcsolódó képzetársítás. Ekkor a neve: kognitívum, és a határát az asszociációs mező, az asszociatív kapcsolatok ereje szabja meg. A kognitívumot legegyszerűbben egy terebélyes, hol ritkább, hol sűrűbb hálóbogaként képzelhetjük el. Ha a bog, vagyis egy sűrűbb hálódarab szögpontjaiban indexkifejezések állnak, melyeket „és”-kapcsolat köt össze, akkor magunk előtt látjuk az osztályozás kognitívumát, az osztályozási kifejezések együttes előfordulását. A kifejezések mátrixából ezeket a bogokat a mátrix parcializálásával, feldarabolásával keressük ki, a gyöngye kapcsolatokat elválasztva az erősektől, az anyagot mintegy „kicsontozva”.

Az információ kvantálásának jelentősége tehát abban áll, hogy a tárolandó anyag természetes egységeit lehet vele meghatározni. Így szakíthatunk az eddigi gyakorlattal, mely egy tárolási egységnek egy formális egységet, azaz egy dokumentumot feleltet meg.

8.2 Heurisztikus információ

A kutató ma már örök tanuló. Ennyiben nem osztom Horváth Tibor optimizmusát, aki szerint „a kutató azért kutató, mert témájáról minden eddigi eredményt ismer. Ha nem így lenne, akkor nem kutatási, hanem tanulási feladatai lennének” (Horváth 1987a:61).

Ám kétségtelen, hogy a kutatói intuíció feltétlenül különbözik abban - például - az anyai ösztöntől, hogy utóbbi tárgya fogható, az előbbié viszont teljességgel foghatatlan. Ráadásul a kutató e foghatatlan felé megtett minden lépését abból a szempontból vizsgálja, vajon érveléséhez a hiányzó láncszemet sikerült-e megtalálnia. Megkaphatná, amit keres, ha tudná, heurisztikusan hol fordult elő hasonló probléma. Vagyis itt viszont a problémák izomorfiájáról van szó. Az információkeresés feladata nem a meglévő ismeretek szaporítása, hanem a meghaladásuk. Vagyis a kutató az információt heurisztikusan keresi, szakmánk viszont a tematikus keresés technikáihoz szokott.

„Egy szakmai példa. Derek de Solla Price törvényszerűséget keresett az információ (dokumentumok) mennyiségi növekedésében. A növekedés tapasztalati tény volt. A probléma tehát úgy merült fel, hogy ez a növekedés mutat-e törvényszerűséget. Van-e valahol már felismert törvény? Valóban, ilyenek adódnak a demográfiából

(népességnövekedés), biológiából (sejtszaporodás), közgazdaságtanból (tőkenövekedés) stb. Valamennyi statisztikai összefüggést jelentett, így el lehetett jutni a matematikához, Euler általános növekedési formulájához. A probléma ezek után két részletre volt korlátozható: az adatok illeszkednek-e Euler exponenciális görbéjéhez, illetve meg kellett határozni a görbe meredekségét meghatározó együtthatót.

Számos példa hozható arra is, hogy az egyik tudomány heurisztikusan hogyan termékenyíti meg a másikat. Az előző fejezetek néhány esete is idézhető. *Farradane*-t a lélektan inspirálja. *Gardin* a diszciplínák egész sorából építkezik, míg esetrendszerét nyelvészeti megfontolások kísérik. A logika, ismeretelmélet pedig az osztályozási rendszerek fundamentumának bizonyultak.”

Mindezt érdemes egybevetnünk a Mandelbrot-hatásról, vagyis a „tudás fölé emelkedésről” mondottakkal. Valószínűleg kibontakozóban van egy olyan irányzat, mely nem elégszik meg a törvények Bertalanffy-féle izomorfiájával, hanem a problémák izomorfiája mentén igyekszik előbbre jutni. Ennek egyik előfordulása lehet az az érdekes cikk, mely kimondja: az emlősök élete, a cickánytól az elefántig, lényegében azonos hosszúságú, noha a legnagyobb és a legkisebb ezt az azonos tartamot különböző idő alatt élék át (Gould 1979.). Biológiai és fizikai időt egy koordináta-rendszerben ábrázolva, a speciális relativitáselmélet illusztrációjára bukkanunk – az élővilágban.

8.3 A „követő” és a precedens nélküli osztályozás

Hogyan hozzuk le a már ismert, a meglévő tudás szintjére az újat, a még alig-ismertet? Hiányoznak hozzá a megfelelő szavak. A nyelvet tanuló gyerek, a teológia és az osztályozás dilemmája ezen a ponton közös: az „alakját (formáját) kereső tartalom” tipikus esete.

„Az uralkodó osztályozási gyakorlat követő osztályozást produkál. Ezért ellentmondás áll fenn a tudomány folytonos, terminológiájában is megvalósuló változása és az osztályozási rendszerek lehetőségei között. Minél nagyobb az időbeli eltolódás, annál inkább a tudományos újdonságok esnek ennek áldozatul. Ha az osztályozó rendszer nagyon régi, elkendőzi az újdonságot ahelyett, hogy felszínre hozná: az osztályozás saját feladata ellen fordul.

A kutatói információs rendszereknek a név-
mók megtalálása fontos, nem traktálhatnak sen-
kit tudományos közhelyekkel. Ami közismert, azt
a kutató már tudja. Az újdonság tulajdonsága pe-
dig az, hogy precedens nélküli. Akkor osztályo-
zásában is érvényesíteni kell a precedens nélküli
osztályozás elvét. Azt kell kifejezni, besorolni,
ami még nem volt.”

*Ha végignézzünk az eddigieken, mindhárom prob-
lémára – a kvantálásra, a heurisztikus információra és
a precedens nélküli osztályozásra – az áll, hogy egye-
dűli orvosságuk az együttes előfordulás, a Luhn szor-
galmazta statisztikai kulcsszóelemzés. Ez választja el
az összetartozó fogalmakat a nem egymáshoz tartozó-
aktól, ez bukkan fel az izomorfia mint eszköz haszná-
latakor, és ebbe, a fogalmi kapcsolatrendszerbe
akaszkodva húzhatjuk le a tegnap még magas isme-
reteket a magunk szintjére – vagy kapaszkodhatunk
fel hozzájuk.*

8.4 Specifikusság

*Az indexkifejezések specifikusságának elve azt
mondja ki, hogy az osztályozásra használt szavak ne
legyenek se általánosabbak, se különösebbek a köz-
nyelvben használnál. Ezt a szegmentálási szintet
Luhn alapállása - hogy a keresett ismérvek a doku-
mentumban vannak, nem kívülről -, valamint az általa
ajánlott eljárások jóvoltából képesek vagyunk megta-
lálni: a keresett szövegszavak és szintagmák számító-
géppel meghatározhatók.*

„Már Luhn és Meadow is tudták, hogy a szig-
nifikáns kifejezések egy meghatározott gyakori-
sági intervallumban találhatók, amelyet diszcipli-
nákra jellemzően – tudományonként ugyanis bi-
zonyos eltérések mutatkoznak – tapasztalatilag
lehet meghatározni. A generikus fogalmak gy-
akorúsága magasabb, a csökkenő gyakoriság – a
szignifikáns kifejezések intervallumán belül –
összefügg a specifikusság növekedésével. A
specifikus szavak gyakorisága általában ala-
acsony. Másfelől a generikus fogalmak eloszlása
egyenletesebb, a specifikusaké általában nem
egyenletes; eloszlásuk rokonságot mutat a ritka
fémek (pl. wolfram) eloszlásával. Ennek hatása
van az elemzésre kiválasztott szövegminták
meghatározásában.

A specifikusság növekedik akkor is, ha a szö-
vegből egységyszavak (*uniterm*) helyett szerkeze-
teket, összetett kifejezéseket határozunk meg.”

Horváth Tibor ezután röviden kitér az automatikus
indexelés és az automatikus osztályozás viszonyára.
Az előbbi az osztályozási kifejezések általánosabb
agglomerációit, együttes előfordulásait vizsgálja, és
ezzel a szakterületi keresőnyelvek lexikájának, szó-
készletének meghatározása felé mutat. A másik irány
emerre - mondhatni - merőleges: automatikus osztá-
lyozáskor a dokumentumhoz a tulajdon szövegéből
géppel meghatározott indexértékű kifejezéseket ren-
deljük hozzá.

Az eddigieket összegezve: mind a jelenlegi
gyakorlat, a követő osztályozás, mind pedig az
indexkifejezések specifikussága mint követel-
mény arra ösztönöz, hogy géppel meghatározott
szövegszavakat használjunk osztályozásra. Ezzel
azonban csak a szókészletet, vagyis a lexikát le-
het meghatározni, a szövegszavak közül ugyanis
hiányzik a szintaxis.

8.5 Relációmegőrző transzformáció

*Az indextétel viszonyrendszere legyen a doku-
mentum relációinak tükré - erről szól a következő trak-
tátus.*

„Az információkereső nyelveknek eddig két
tartópillérére derült fény. Az egyiknek az osztá-
lyozandó egységek, kvantumok megválasztása,
a másikkal a specifikus, az élő szaknyelvvvel egy-
behangzó kulcsszavak bizonyultak. Megvannak
tehát az elemek. Bizonyára hiányérzet támad
bárkiben, aki látta, hogy ezeket az elemeket ér-
telmes közléssé, összképpé, – *Gardint* idézve –
ikonná, kompozícióvá, mondatá kell összekap-
csolni.

A szövegszavas elemzés elemi hiányossága,
hogy nem eléggé érzékeny a relációkra. Ez
azonban újabb alapprobléma.

Gardin következetesen végiggondolt mód-
szerének ismeretében látható volt, hogy az elem-
zés harmadik fázisában azt kellett megoldania,
hogy az individuális elemek hogyan szerveződ-
nek kompozícióvá. Ez már szintaktikai termé-
szetű elemzést kívánt meg, s *Gardin* nem is ha-
bozott nyelvteni esetrendszereket átvenni ahhoz,
hogy az individuális elemekből mondatokat ge-
neráljon. Azaz: szavakból mondatok állnak elő.

Nyelvet konstruált tehát, és a nyelvnek – formális definíciója értelmében – bizonyos elemeket (szavakat) és viszonyokat kell tartalmaznia. (A formális nyelvek elméletében a jelentés, mint harmadik komponens nem is definíció értékű.) Ha *Gardin* problémáját az osztályozandó szövegekre, kognitívumokra, dokumentumokra átvisszük, akkor az osztályozásmélet legmélyebb, legszigorúbb gondolatáig jutunk el. A szövegekben ugyanis a fogalmak bizonyos elrendezésben állnak, rendezettségüket az a viszonyrendszer határozza meg, amelybe a szerző állítja a fogalmakat.

Landry indexelési elméletében az egyik posztulátum erre vonatkozik. A posztulátum azt mondja ki alapelveként, hogy a transzformáció, amely az eredeti szöveget átviszi reprezentációjába (indextételébe), őrizze meg a fogalmak szövegen belüli rendezettségét. Más szóval az a kép, amelyet egy tétel osztályozása nyújt egy szövegről, őrizze meg a fogalmaknak a szövegben egzisztáló viszonyrendszerét. Ez pedig ismét csak szintaktikai probléma.”

Mindez persze felveti a maga érdekes kérdéseit. Elsősorban azt, hogy a jelentés függő, a szemantika a szintaxis függvénye. Ennek a függésnek különösen jó

példáját ismeri a magyar az ún. mozgó jelentés esetében. A „Jössz-e hozzám feleségül?” és a „Feleségül hozzám jössz-e?” kérdések más értelmi árnyalatait jól halljuk, noha a két mondat egyedül a szavak sorrendjében tér el egymástól. A szavak viszonyrendszerének változása tehát a jelentés változását okozza. Ez még nehezebb feladat elé állítja a fogalomhű, szemantikailag hű szurrogátumok készítőit.

Irodalom:

Gould, S.J. (1979.): One standard lifespan, *New Scientist* [Feb. 8.]. vol.81, No.1841, 388-389.pp.

Horváth T. (1987a): „Egy törzsnek az ága...” Ismeretszerzés és szintaxis. Kandidátusi értekezés. Budapest.

Horváth Tibor (1987b): „Egy törzsnek az ága...” Ismeretszerzés és szintaxis. Kandidátusi tézisek (OPKM) Budapest.

A cikk második részét lapunk 3. számában publikáljuk. (A szerk.)

ETO ALAPÍTVÁNYNAK adja át a FID 1991 végén az ETO gondozását. Az „UDC Consortium” néven megalakuló új szervezetnek a FID-en kívül egyelőre az angol szabványügyi intézet, a spanyol szabványügyi társaság, a japán tájékoztatástudományi társaság, a belga-vallon közművelődési könyvtári központ és egy holland cég a tagja. További szervezetek csatlakozását is várják. Az alapítvány öfenntartó vállalkozásként folytatja tovább azt a munkát, amelyet eddig a FID-nek az ETO-val foglalkozó szervei végeztek. Elsődleges feladatának tekinti az anyag teljes számítógépesítését, amely alapul szolgálhat speciális, különféle nyelvű és különféle szakterületi (kézi vagy gépi használatú) kiadványok előállításához.
(FID News Release, 1991. nov.)