

Jogszabályszövegek gépi elemzésének tanulságai

A normák természetére és változására fókuszáló kutatásunk⁴ keretében jogszabályok gépi feldolgozásával kezdtünk el foglalkozni. A munkának ebben a fázisában elemeznünk és értékelnünk kellett a jogszabályszövegekben gyakran felbukkanó speciális listákat számítógép segítségével. A gépi feldolgozás rámutatott ezeknek a listáknak néhány sajátosságára, és arra, hogy értelmezésük, különösképp automatizált értelmezésük⁵ során milyen nehézségek merülhetnek fel.

A jogszabályok szerkesztésével kapcsolatos elvárásokat a 61/2009. (XII. 14.) IRM-rendelet rögzíti. A szerkesztésre vonatkozó előírások részben a jogszabályok szerkezetére vonatkoznak. Ilyen például az – a jogászok számára triviális, a cikk szempontjából viszont kiemelő – előírás, mely szerint a jogszabályszövegeket szerkezeti egységekre kell tagolni úgy, hogy a jogszabályok „szerkezeti egységeit folyamatos sorszámozással vagy a latin ábécé betűivel meg kell jelölni. Jogszabály tervezete jelöletlen szerkezeti egységet nem tartalmazhat.”⁶

Ezek a szövegek tehát *jelölt szerkezeti egységekből* állnak (csak azokból állhatnak). További fontos elvárás, hogy minden egyes szerkezeti egység a rendeletben meghatározott típusok valamelyikébe kell tartozzon. A rendelet „a jogszabály tervezetének összetettségétől függően, a szerkezeti egységek növekvő szintjének sorrendjében:

- a) az alpont,
- b) a pont,
- c) a bekezdés,
- d) a szakasz,
- e) az alcím,
- f) a fejezet,
- g) a rész és
- h) a könyv...”

típusokat különíti el.⁷ Alább látni fogjuk, hogy a jogszabályszerkesztési gyakorlatban milyen problémák merülnek fel a taxonómiával kapcsolatban.

Szintén a rendeletben rögzített, hogy a bekezdések és a bekezdéseknél magasabb szintű szerkezeti egységek esetében nem megengedett a bekezdésen túlnyúló hatókör.⁸ Ez utóbbiból következik, hogy lehetnek olyan szerkezeti egységek – a *pontok* és az *alpontok* –, amelyek esetében a mondatok túlnyúlhatnak a szerkezeti egységek határain. Az ilyen szerkezeti egységekben szereplő szövegrészek önmagukban nem teljesek, így csak más szerkezeti egységekben található szövegrészekkel kiegészítve, azokkal együtt válnak teljessé, szemantikailag érvényessé. Nézzük meg a következő példát:

„(1) Az alapszabály az alaptőke húsz százalékát meg nem haladó mértékben rendelkezhet olyan részvény kibocsátásáról, amely alapján a részvényre vonatkozóan

¹ Doktori hallgató, ELTE Filozófiatudományi Doktori Iskola, Logika Tanszék; egyetemi tanársegéd, BME Üzleti Jogi Tanszék

² Egyetemi docens, BME Szociológia és Kommunikáció Tanszék

³ Egyetemi docens, BME Szociológia és Kommunikáció Tanszék

⁴ OTKA 83887

⁵ A szemantikus web Berners-Lee-féle programja (Berners-Lee, 2010) úgy szól, hogy a web első évtizedében megtanítottuk a gépeket a szöveget olvasni, a most következő évtizedben – mondja 2000-ben – meg kell tanítanunk arra őket, hogy értsék is azt, amit olvasnak.

⁶ A jogszabályszerkesztésről szóló 61/2009. (XII. 14.) IRM-rendelet (továbbiakban 61/2009 IRM r.) 37. § (1)

⁷ 61/2009 IRM r. 36. § (2)

⁸ „Bekezdések alkalmazása esetén egy szakaszon belül valamennyi mondatnak teljes terjedelmében valamely bekezdésbe kell tartoznia.” (61/2009. IRM. r. 44. §. (5))

- a) a részvénytársaságot vételi jog;
- b) a részvényest eladási jog; vagy
- c) a részvénytársaságot vételi jog és a részvényest eladási jog illeti meg.”⁹

Az idézett rész a bekezdés alá sorolt pontokban három állítást tartalmaz, és meg is jelöli mindegyik szerkezeti egységet: az (1)-es jelű bekezdést, illetve az a), b) és c) jelű pontokat. Ha valamely más jogszabály hivatkozik az idézett szövegrészen belül, például, a b) pontra, akkor pusztán a b)-vel jelölt szöveg alapján nem rekonstruálható, hogy mit akart kifejezni a jogalkotó ebben a szerkezeti egységben. Természetesen a jogszabályszöveg olvasásában gyakorlott olvasó tudja, hogy a b) pont tartalmát ki kell egészíteni az (1)-gyel jelölt nyitó mondatrészszel, és tudja azt is, hogy hozzá kell még olvasnia az utolsó – itt a c) – pont után következő, nem jelölt szövegrészt is. Az értelmes mondatot tehát három mondatrész (vagy három részmondat) összeolvasásával kapjuk meg. A kutatás keretében elvégzett fejlesztési munka egyik célja az volt, hogy kidolgozzuk azt az eljárást, amellyel az összeolvasás eredményeként előálló szöveg a gépi feldolgozás során rekonstruálható.

I. Ellipszisek

Jelen cikkben nem az eljárás technikai részleteit mutatjuk be,¹⁰ inkább azokról a tapasztalatokról szeretnénk számot adni, amelyeket e munka során szereztünk, és amelyek a jogszabályszerkesztési gyakorlat számára is hasznosíthatók.

Érdemes mindenekelőtt tisztázni, hogy miféle nyelvi vagy nyelvi reprezentációs jelenséggel van dolgunk a vizsgált hiányos szerkezet esetében A nyelvészetben a hiányos mondatok alkalmazását *ellipszisteknikának* nevezik. Az ellipszisteknikák alkalmazásának több módja lehetséges, és az ilyen rövidítésekkel a természetes nyelvben nagyon gyakran élünk. A beszédben (írásban) sokszor elhagyjuk a fölöslegesnek vélt ismétléseket, hogy rövidebbé, tömörebbé tegyük mondandónkat. Tegyük fel, hogy a következő információkat akarjuk közölni valakivel:

Réka olvas. Gábor olvas. Syi olvas.

Ezt a három mondatot az esetek többségében lerövidítjük, és csak annyit mondunk (írunk), hogy:

Réka, Gábor és Syi olvas.

Ezzel ugyanazt a tartalmat közöljük, csak rövidebb formában, és ezt az elliptikus mondatot is jól formálnak és érvényesnek tartjuk szintaktikai, illetve szemantikai értelemben. Fontos tény azonban, hogy az érvényesség megállapításakor (nyugtázásához) a háttérben elvégzünk egyfajta érvényességvizsgálatot a mondaton. Nem mindig jutunk arra az eredményre, hogy helyes mondatot kaptunk. Addig, amíg nem tudjuk, hogy a ‘Syi’ kifejezéssel ugyanúgy egy személyt jelölünk, mint ahogy ezt tesszük a ‘Réka’ és a ‘Gábor’ terminusokkal, addig bizonytalanok lehetünk abban, vajon a harmadik mondat(rész) szemantikailag érvényes tartalmat fejez-e ki, de ha rendelkezésünkre áll az az információ, hogy a ‘Syi’ kifejezéssel is egy személyt jelölünk, a bizonytalanságtól megszabadulhatunk. Azonban a következő mondat feldolgozásánál komoly interpretációs munkát kell végeznünk, hogy – pl. metaforikusan – értelmesnek tekintsük a mondatot, ha éppen nem utasítjuk el mint értelmetlent (szemantikailag helytelen).

⁹ A Polgári Törvénykönyvről szóló 2013. évi V. törvény (továbbiakban Ptk.)

¹⁰ Erről bővebben lásd: Syi – Hamp Gábor – Markovich Réka: Goody-listák jogszabályszövegekben. (kutatási műhelytanulmány) 2014, és Hamp Gábor – Syi – Markovich Réka: Elliptikus listák jogszabályszövegekben. In Tanács A. – Varga V. – Vincze V. (szerk.) XI. Magyar Számítógépes Nyelvészeti Konferencia. Szeged, 2015.

Az elutasítás vagy az interpretációs többletfeladat rávilágít arra a „háttérmunkára”, amit a mondat olvasása közben végzünk el. Ez a háttérmunka egyfajta osztálybasorolás. A világról való tudásunk alapján ugyanis kijelenthetjük, hogy a szék nem tud olvasni, ezért nem tartjuk érvényesnek (igaznak) a róla szóló kijelentést, és ebből következően nem fogadjuk el helyesnek a mondat egészét sem.¹¹

II. Goody-listák és Goody-mondatok

A jogszabályokban másfajta ellipszistechnikákat is találhatunk.¹² Mégpedig azokban a szerkezeti egységekben (pontokban, alpontokban), amelyekről a bevezetőnkben tettünk említést. A pont és alpont szintű szerkezeti egységekben gyakran találhatunk felsorolásokat, listákat. Jack Goody írásában¹³ az írásbeliség specifikumait keresve említi, hogy az írásban alkalmazunk olyan technikákat, amelyek a szóbeli kommunikáció során nem állnak (mert nem állhatnak) rendelkezésünkre. Példaként a listákra, a táblázatokra vagy a keresztretjvények különös térbeli elrendezéseire utal. Ezek mindegyikének közös jellemzője az a mód, ahogy az – egyébként szükségszerűen – lineáris jellegű szöveget (pontosabban annak egy részét) kétdimenziós formában rendezzük el. Utalhatunk itt például jelen írás elején található listára, amely felsorolja a jogszabályok lehetséges szerkezeti egységeit. A felsorolás elvégezhető úgy is, hogy egyetlen sorba rendezzük el a lista elemeit. Ha így tettünk volna, akkor illeszkedtünk volna a szöveg (írás) szokásos, egydimenziós, lineáris formátumához. Azzal azonban, hogy más reprezentációs formát választottunk, lehetőségünk nyílt arra, hogy kifejezhessünk valami többletet a lineárisan rendezettségű írásmóddhoz képest. Tartalmilag ugyan nem keletkezett többletinformáció, a formai megjelenítés megváltoztatásával – egy másik dimenzióban – könnyebben befogadhatóvá tettünk egy olyan háttértudást, amely a fejünkben megvolt addig is, csak az írásban hagyományos módon használt egyetlen dimenzióban nem tudtunk kifejezni. A listatételek egymás alá rendezése egy másik (a függőleges) dimenzióba láthatóvá tette, hogy a felsorolt elemek mind ugyanabba az osztályba, itt: a 'pont' típusú szerkezeti egységek osztályába tartoznak. Továbbá ezzel a megoldással arra is mód nyílt, hogy a tételeket egyenként megjelölhessük, ami pedig azért jelenthet előnyt, mert a rájuk való hivatkozás lehetősége így jóval könnyebbé és egyértelműbbé válik. Az ilyen jelölt listákra a továbbiakban Goody-listákként hivatkozunk.

Hat törvény teljes szövegét töltöttük be egy adatbázisba úgy, hogy minden szerkezeti egység az adatbázis egy-egy sorába (rekordjába) került. Összesen 10 500 sort (szerkezeti egységet) kaptunk. Ezek egy adott része valamilyen magasabb szintű szerkezeti egység címét vagy jelölését tartalmazta. Mivel ezek a sorok tartalmilag kevésbé voltak érdekesek az elemzés számára, ezért ezeket elkülönítettünk a többi rekordtól (az ilyen „tartalmatlan” sorok teszik ki a korpusz negyedét). Első lépésben megtanítottuk a számítógépet arra, hogy ismerje fel a szerkezeti egységek típusát, ami nem volt túl nehéz feladat, mivel a jogszabály-szerkesztési gyakorlat ezeket – régtől fogva – egységesen, megbízható módon és magas szinten formalizálja, és a formalizálás nyomán létrejövő mintázatok gépi felismertetése viszonylag egyszerű feladat. A következő feladat volt a Goody-listák gépi felismertetése, és a listák belső struktúrájának gépi beazonosítása (hogy ez mit jelent pontosan, arra később még kitérünk bővebben). A Goody-listák megtalálása után kiderült, hogy az összes – tízezer sornyi – szerkezeti egység közül háromezer Goody-listába

¹¹ Bővebben lásd: Syi et al. 2014 i. m.

¹² A szintaktikai és a pragma-szemantikai ellipszisekről jó áttekintést nyújt pl. Bánréti Zoltán: Az ellipszis mondatnana és a lexikai szelekció. In *Nyelvtudományi Közlemények*, 1998, 98, 7–70., illetve Stainton, R. J.: *Words and Thoughts. Substances, Ellipsis and the Philosophy of Language*. OUP, 2006, pp. 80–144.

¹³ Goody, J.: *Nyelv és írás*. In Nyíri K. – Szécsi G. (szerk.): *Szóbeliség és írásbeliség*. Áron Kiadó, Budapest, 1988, 189–221. o.

tartozó rekord van. Ez azt jelenti, hogy az elemzésbe vont teljes jogszabályi szövegtörzset közel harmada Goody-listás szövegnek minősíthető. Ez az egyetlen arányszám önmagában jelzi, hogy fontos jelenségről van szó.

A továbblépéshez szükséges az ilyen felsorolások belső logikai szerkezetének feltárása. A jogszabályszerkesztésről szóló rendelet meghatározza, hogy – a pontokba vagy alpontokba szedett – lista milyen szerkezeti összetevőkből áll. A cikk elején bemutatott példákra (Ptk. 3:239.§) újra rátekintve könnyen felismerhetők a felsorolások bevezető és lezáró részei, valamint az ezek között található listaelemek. Az IRM-rendelet értelmében a bevezető szövegrész kötelező eleme a listának, a záró rész opcionális: „[a]nnak érdekében, hogy a pontokat tartalmazó szerkezeti egység ne váljon hiányos mondatná, a bekezdés szövege keretbe is foglalhatja a pontokat. Ez esetben a pontokat követő szövegrészt (a továbbiakban: záró szövegrész) sortöréssel kell a pontoktól elválasztani.”¹⁴

A nyelvészeti hagyományokhoz igazodva a felsorolások felvezető részét a lista *nyitófejtételének* nevezzük,¹⁵ a záró szövegrészt pedig a lista *zárófejtételének*,¹⁶ míg a felsorolt elemeket *listatételeknek*. A listaösszetevőkre vonatkozóan annyit rögzíthetünk még, hogy a nyitófejtételből csak egy lehet egy listán belül, és ennek kötelezően léteznie kell, a listatételekből legalább kettő elemnek léteznie kell (különben értelmetlen lenne felsorolásról beszélni),¹⁷ végül a zárófejtétel jelenléte opcionális, tehát nem szükségszerű része a listáknak.

Kidolgoztuk azokat az eljárásokat, amelyekkel a listaösszetevőket automatikusan felismertetjük, valamint a Goody-listák elemei összeilleszthetők. Az összeillesztés nyomán előállt nem elliptikus szerkezeteket Goody-mondatoknak nevezzük. A Goody-mondatok azok a gép által előállított teljes mondatok, amelyek a nyitófejtétel bevezető részéből, az adott listatételből és az opcionális zárófejtételből állnak össze. Ezek az ellipszist feloldó, szemantikailag teljes mondatoknak tekinthetők. A tízezres korpuszban 540 listát azonosítottunk be. Ez tehát 540 nyitófejtételt jelent, 130 a zárófejtételek száma, 2300 sorban szerepeltek listaelemek, tehát ennyi Goody-mondatot. Ennyi Goody-mondatot lehet a gép feldolgozással képezni.

III. Goody-listák problémái

A jogszabályok gépi elemzése során a számítógéptől arra tanítottuk (tanítjuk), hogy az előzetes elvárásoknak megfelelő szabályszerűségeket keressen (és találjon) a korpusz egészén belül. Ez a folyamat azonban alkalmas volt arra is, hogy a szabálytól való eltéréseket, jogszabály- szerkesztési „hibákat” megmutassa, és felhívja a figyelmünket olyan – teljesen szabályszerű – helyzetekre, ahol a gépi elemzés további feladatok megoldásával válik csak alkalmassá a kívánt eredmény nyújtására. A továbbiakban ezekről adunk számot.

III.1. A taxonómiában nem szereplő szerkezeti egység-típus

Korábban láttuk, hogy a listákat összefogó (nyitó) fejtételek csak bekezdés vagy pont szintű szerkezeti egységekben fordulhatnak elő, továbbá, hogy Goody-listák listaelemei csak pont vagy alpont szintű szerkezeti egységben lehetnek. Ebből a hierarchiából az is következik, hogy ha egy lista alpontokat is tartalmaz, akkor ezek nyilván a lista pontjai alá lehetnek csak besorolva. Ilyen esetekben az alpontok egy beágyazott Goody-lista listatételei, amelyeknek a fejtétele már a

¹⁴ 61/2009. IRM. 47. § (3).

¹⁵ A rendeletben „nyitó szövegrész” a megnevezése [61/2009. IRM. r. 47. § (1) bekezdés].

¹⁶ A rendeletben „záró szövegrész” a megnevezése [61/2009. IRM. r. 47. § (3) bekezdés].

¹⁷ Ebből következően az egyelemű Goody-lista a józan észnek és a jogszabályalkotásról szóló rendelet előírásainak ellentmond, de jogszabálmódosítások következményeként előállhat egyetlen listatételből álló alakzat. Ilyen például a nemzeti felsőoktatásról szóló törvény 31. § (1) bekezdése.

„külső” lista listatétele is egyben. A pontok tehát kettős szerepet is betölthetnek: lehetnek egyfelől a fölöttük levő bekezdés mint listafejtétel alá sorolt listatételek, másfelől funkcionálhatnak fejtételként az alájuk besorolt alpontok tekintetében. Nézzünk meg egy példát a kétszeres, beágyazott Goody-listákra (vagy másként: az alpontokat tartalmazó felsorolásokra).

- „(4) A felsőoktatási információs rendszer működéséért felelős szerv
- a) a működési engedéllyel rendelkező felsőoktatási intézménynek intézményi azonosító számot (a továbbiakban: intézményi azonosító szám);
 - b) annak, akit oktatói, kutatói, illetve tanári munkakörben közalkalmazotti vagy más, munkavégzésre irányuló jogviszonyban foglalkoztatnak
 - ba) oktatói azonosítót,
 - bb) kérelemre oktatói, kutatói, alkalmazotti igazolványt;
 - c) a felsőoktatási felvételi eljárásra jelentkezőknek, a hallgatóknak hallgatói azonosítót;
 - d) a hallgatóknak kérelemre diákigazolványt ad ki”¹⁸

Ebben az esetben a ba) és bb) alpontok csak akkor értelmezhetők, ha kétszeresen egészítjük ki a hiányos mondataikat: először a bekezdésben szereplő fejtétel, majd a közvetlenül följük rendelt b) pont tartalmával. Az ilyen típusú Goody-listák gépi keresése meglepő – terminológiai jellegű – inkonzisztenciára mutatott rá néhány jogszabályban. Előfordultak ugyanis olyan felsorolások, amelyeket nem lehetett konzisztens módon leírni a jogszabályszerkesztésről szóló rendeletben megadott szerkezetegység-típusok segítségével. A rendelet előírja, hogy a pontokat kisbetűkkel vagy arab számokkal kell jelölni a felsorolás elemszámától függően. Az a tény, hogy egy pontot vagy számmal, vagy betűvel egyaránt lehet jelölni, még nem jelent önmagában problémát. Az viszont már igen, hogy olykor előfordulnak olyan számozott pontok, amelyek alá kétszeresen beágyazott felsorolások vannak rendelve. Nézzünk meg egy példát erre.

- „13. kar: egy vagy több képzési területen, tudományterületen több, a képzési programban rögzített szakmailag összetartozó képzés oktatási és tudományos kutatási, illetve alkotó művészeti tevékenység feladatait ellátó szervezeti egység;
- a) egyetemi kar az a szervezeti egység, amelyen
 - aa) a teljes munkaidőben foglalkoztatott oktatók létszáma eléri a 40 főt
 - ...
 - b) főiskolai kar az a szervezeti egység, amelyen
 - ba) a teljes munkaidőben foglalkoztatott oktatók létszáma eléri a 35 főt
 - bb) a teljes munkaidőben foglalkoztatott oktatók legalább harmada rendelkezik tudományos fokozattal;”¹⁹

Ezzel a megoldással az a gond, hogy nem lehet egyértelműen megnevezni a 13. pont alá tartozó szerkezeti egységeket. Az a) vagy a b) jelű szerkezeti egységet a jelölési konvenció alapján mondhatnánk pontnak, de a hierarchikus építkezés szabályai miatt a 13. pont alá nem rendelhetünk ‘pont’ típusú szerkezeti egységet, így a terminológiai tisztaság azt követelné meg, hogy alpontokként kezeljük őket. Ez viszont a jelölési-megnevezési konvenciónak mondana ellent. Ugyanilyen problémát okoz az aa), ba), bb) jelű egységek minősítése is. A jelölési konvenció alapján alpontoknak kéne definiálnunk ezeket, de a hierarchikus pozíciójuk szerint voltaképp „al-alpontok”-nak kéne tartanunk őket. Az ilyen típusú szerkezeti egység létezését (megnevezését) viszont nem engedi a vonatkozó rendelet. Ez nem pusztán elnevezési kérdés. Ez a többértelműség nem teszi lehetővé azt, hogy pontosan – a rendeletben rögzített hierarchiát és terminológiai szabályokat betartva – lehessen hivatkozni ezekre a szerkezeti egységekre. Ha

¹⁸ A nemzeti felsőoktatásról szóló 2011. évi CCIV. törvény (továbbiakban 2011. évi CCIV. tv.)

¹⁹ 2011. évi CCIV. tv.

ragaszkodunk az elnevezési konvenciókhoz, akkor egy másik jogszabályban hivatkozhatunk az „adott törvény adott bekezdés 13. pont a) pont ab) alpontjára”, ami szerencsétlen megoldás lenne, hiszen a ‘pont’-ot kétszer használnánk. Terminológiai szempontból jobb lenne az a megoldás, ha az utalás második elemére ‘alpont’-ként, az utolsó elemre pedig ‘al-alpont’-ként utalnánk, ami viszont a hatályos szabályozás és a tényleges joggyakorlat számára ismeretlen kifejezés lenne. Az ilyen háromszintű egymás alá rendelés ritkán fordul elő, eddig csak a törvények értelmezési rendelkezései között találtunk rá példát, de elvileg elképzelhetőnek tartjuk, hogy máshol is felbukkanjanak.

Az ‘al-alpont’-ra mint szerkezeti egységre tehát olykor szükség lenne, de a bevezetésével kapcsolatban mi sem tudnánk egyértelmű álláspontot képviselni. A terminológiai tisztaság, az egyértelmű utalás biztosítása végett érdemes lenne felvenni a szerkezeti egységek közé még akkor is, ha csak ritkán kellene igénybe venni a gyakorlatban. Az új szerkezeti egység bevezetése azonban azt is jelentené, hogy egy újabb szint jelenne meg a szerkezeti egységek hierarchiájában, illetve a felsorolási listák háromszintűvé válnának. Ez nem lenne szerencsés, mert a befogadhatóságot és az egyértelmű értelmezést megnehezítené. A Goody-listák elemzéséből kiderül, hogy a jogszabályszerkesztés során a jogalkotó igyekszik elkerülni az ilyen többszörös egymásba ágyazásokat, elég nyilvánvalóan azt az elvet követve (mindegy, hogy ösztönösen vagy tudatosan), hogy a felsorolásokban található ellipszisek feloldása ne legyen túl nagy feladat a listák „összeolvasásakor”.

III.2. Kiszervezett listatétel

A többszörösen egymásba ágyazott Goody-listák elkerülésére olykor azt a technikát alkalmazza a jogalkotó, hogy egy szabályos listatételen belül hivatkozik egy következő felsorolási elemre.

- „(1) A mezőgazdasági igazgatási szerv - a 23. § (1) és (2) bekezdésében meghatározott eseteken túl - az adás-vételi szerződés jóváhagyását megtagadja
- a) ha a helyi földbizottság
 - aa) állásfoglalása alapján egyik elővásárlásra jogosulttal és az adás-vételi szerződés szerinti vevővel sem támogatja az adás-vételi szerződés jóváhagyását,
 - ab) a 26. §-ban foglalt felhívás ellenére nem adja ki az állásfoglalását;
 - b) a (2) bekezdésben meghatározott esetekben.”²⁰

A b) pontba felvett listatételben tartalmilag nincs állítás kifejtve, csak annyi információt kapunk, hogy egy másik Goody-lista fejtételére hivatkoznak. Ha a (2) bekezdés tartalmát az itteni b) pontban sorolták volna fel, akkor már túl bonyolult lett volna az így kialakuló szerkezet, amit a passzus olvasója nehezebben tudott volna feloldani. Hogy a többszörös egymásba ágyazások elkerülésének szándékáról lehet itt szó, mutatja az is, hogy az idézett törvényhelyen a következő (3) bekezdés is még ugyanazt a kérdést szabályozza egy újabb Goody-listával. A túl mély hierarchia elkerülésére alkalmaz a jogalkotó olyan megoldásokat is, amikor visszafelé hivatkozik az adott jogszabályon belül.

- „(4) Az (1) bekezdés b)-d) pontjában, valamint a (2)-(3) bekezdésben meghatározott földműves jogosulti csoportokon belül az előhaszonbérletre jogosultak sorrendje a következő
- a) a családi gazdálkodó, illetve a gazdálkodó család tagja,
 - b) fiatal földműves,
 - c) pályakezdő gazdálkodó.”²¹

²⁰ A nemzeti vagyronról szóló 2011. évi CXCVI. törvény (továbbiakban 2011. évi CXCVI. tv.) 27. §

Ezek a megoldások úgy értelmezhetők, hogy ilyen esetekben a jogalkotó több listát is összekapcsol szemantikai értelemben, de ezt az összekapcsolást nem támogatja szintaktikai eszközökkel, ezáltal a hierarchia mélyítését – legalábbis a felszínen – elkerüli.

III.3. *Bújtatott listaelem*

A beágyazási technikáknak egy furcsa, kevert típusát láthatjuk a következő idézetben.

„A jogszabály tervezetében a valamely

a) nemzetközi szerződésre,

b) az Európai Unió intézményei által az Európai Unió működéséről szóló szerződés vagy az Európai Atomenergia-közösség létrehozásáról szóló szerződés alapján elfogadott jogi aktusra (a továbbiakban együtt: uniós jogi aktus)

vagy ezek szerkezeti egységére való hivatkozásra e fejezet rendelkezéseit – az ebben az alcímben foglalt eltérésekkel – kell alkalmazni.”²²

Ebben a passzusban a zárófejtételben szerepel egy olyan beágyazott diszjunkció („vagy ezek szerkezeti egysége”), amelyet az a) és b) pontok feloldásakor (kiegészítések) mindegyik listatételhez illeszteni kellene, így voltaképpen a zárófejtételben szereplő kiegészítés egy (pontosabban kettő) „bújtatott” listaelemnek is tekinthető. Ez nem igazán szerencsés megoldás, mert itt a zárófejtétel lezáró, kiteljesítő funkciója keveredik az egyes listatételekben kifejezett lehetséges alternatívák felsorolásának funkciójával.

III.4. *Elcsúszott zárófejtétel*

A zárófejtételek kezelésében találhatunk egy másfajta inkonzisztenciát, amely a gépi feldolgozást mindenképpen és jelentősen akadályozza, de bizonyosan az egységes kezelésmód ellenében hat, és vélhetőleg az emberi befogadást is inkább nehezíti, mint megkönnyíti. A záró szövegrészekben írt mondat(ok)nak ugyanis az az alapvető funkciója (ahogy ezt idéztük korábban), hogy lezárja (kiegészítse) a listaelemek hiányos mondatait. Ezt a feladatát akkor tudja jól betölteni, ha valóban önálló egységként jelenik meg a felsorolás legvégén (sortöréssel elválasztva, elkülönítve az összes listaelemtől – ahogy ezt egyébként a jogszabályszerkesztésről szóló rendelet is előírja). Előfordul azonban (meglepően sokszor), hogy sérül ez a szabály. Íme egy példa.

„48/Q. § (1) A volt magyar állami (rész)ösztöndíjas hallgató kérelmére a magyar állami ösztöndíj feltételei teljesítésének nyilvántartásáért felelős szerv

a) ötmillió forint alatti összegű állami ösztöndíj visszatérítési kötelezettsége esetén legfeljebb tíz évre szóló,

b) ötmillió forint feletti összegű állami ösztöndíj visszatérítési kötelezettsége esetén legfeljebb tizenöt évre szóló részletfizetést engedélyez.”²³

Ennek a listának formailag nincsen elkülönült záró szövegrésze, mégis látjuk, hiszen tartalmilag le tudjuk választani a b) pontban olvasható mondat utolsó két szavát, amikre a lista egészének értelmezésekor úgy kell gondolnunk, mint ami az a) pont lezárásaként is funkcionál.

²¹ A mező- és erdőgazdasági földek forgalmáról 2013. évi CXXII. törvény (továbbiakban: 2013. évi CXXII. tv.) 46.§

²² 61/2009. IRM. r. 26. §

²³ 2011. évi CCIV. tv.

Ezt a formai inkonzisztenciát az emberi olvasó „korigálni”, kezelni tudja, de ettől még formai hiba marad. A gépi elemzés praktikus eredményének tekinthető, hogy bár a gép az ilyen eseteket egyelőre nem tudja jól értelmezni, azzal, hogy az ilyen helyeken leáll, felhívja a figyelmet arra, hogy azokon a pontokon a jogszabályokban formai hibák vannak. És ahogy jeleztük, meglepően sok ilyen esetet találtunk.

III.5. A listatípusok logikai problémái

Kutatásunk hosszú távú célja megérteni, megkeresni, kitalálni azt, hogy hogyan lehet megtanítani a számítógépet arra, hogy értelmezni tudja a jogszabálysövegeket. A kutatásban résztvevők számára ez a kérdés jelenti az igazi kihívást. Tudva, hogy e cél elérésétől még nagyon távol vagyunk, a kutatásban, ahol csak lehet, megpróbáljuk a kezdetektől fogva alkalmazni a logika eszköztárát azt remélve, hogy az apró sikerek bizonyos részproblémák megoldásával idővel majd lehetővé teszik a jelentősebb áttörést is ezen a fronton. A Goody-listákkal kapcsolatos problémák egyike szorosan kapcsolódik a logikához, így értelemszerűen sokat foglalkoztunk vele. Az IRM- rendelet felsorolja, hogy a felsorolásokban milyen kapcsolat lehetséges a lista elemei között. Négy típust különít el, amelyeknek nem ad külön neveket, mi azonban a kijelentéslogika kategóriái segítségével meg fogjuk nevezni őket. A rendelet így fogalmaz:

- „Felsorolás alkalmazása esetén egyértelművé kell tenni, hogy a felsorolás elemei közül
- a) valamennyinek teljesülnie kell,
 - b) egyik sem teljesülhet,
 - c) pontosan egynek kell teljesülnie vagy
 - d) legalább egynek teljesülnie kell... a joghatás kiváltásához.”²⁴

A kutatás részeként az elemzési korpuszunkban található összes Goody-listát elemeztük, és besoroltuk a rendeletben elkülönített típusok valamelyikébe. Anélkül, hogy bővebben kifejténénk, itt csak jelezzük, hogy a logika nyelvén ezeket a mondatok közti összekapcsolást biztosító műveletekre az alábbi módon hivatkoznak. Az a) pont esetében logikai *konjunkción* („és”-kapcsolatról) van szó, a d) pont esetében logikai *diszjunkción* („megengedő vagy”-kapcsolat) köti össze az elemeket, a c) pont a *bisubtrakción* (a „kizáró vagy”-kapcsolat) esetére utal, végül a b) pontban leírtaknak a *konnegáción* művelete felel meg.²⁵ Érdeemesnek tartjuk jelezni, hogy az ilyen listák elemzésekor nagyon érdekes logikai problémák merülnek fel. Első lépésként el kell különítenünk a Goody-listákon belül lehetséges mondatösszekapcsolások két típusát egymástól. Az idézett rendelet a listatételek közti kapcsolat minőségét tipizálja, amit mi a lista *belső kapcsolatának* nevezünk. A belső kapcsolat két listatétel között állhat csak fent, és ezekre a kapcsolatokra nézve helytálló, elfogadható az IRM-rendelet taxatív felsorolása. A listák egészén belül azonban létezik egy másik kapcsolódástípus, amit mi *külső kapcsolatnak* hívunk. Ez a kapcsolat köti össze a lista fejtételét és a listatételek együttesét, egészét. Az egyik nagyon gyakori külső kapcsolódási forma az ún. retrokondicionális művelet, ami a kondicionális kapcsolódás megfordítottja. Ennek legfontosabb formai jegye abban ragadható meg, hogy előbb megfogalmaz egy kijelentést, majd utána elmondja (felsorolja) azokat a feltételeket, amelyek teljesülése szükséges az első tagmondatban kifejtett állításhoz.²⁶ Szintén gyakran fordul elő ún. bikondicionális viszony, amikor definiálás-jellegű deklarációval adja meg a jogalkotó valaminek a szükséges-elégséges feltételeit. A külső kapcsolatok tulajdonképpen a fejtétel és a listatételek között érvényesítendő normatív viszonyt fejezik ki. Például a kötelezést, a tiltást vagy a

²⁴ 61/2009. IRM r. 7. §

²⁵ Syi: syi.hu/cse. L'Harmattan – Könyvpont Kiadó, Budapest, 2014

²⁶ A kijelentések közti logikai kapcsolatok tipizálásáról lásd: Syi 2014 i. m.

megengedést kifejező formai jegyeket a fejtételben (logikai szempontból: deontikus operátort helyeznek a fejtételben) helyezi el a jogalkotó, és a listatételekben sorolja fel azokat a cselekvési lehetőségeket, amelyekre vonatkoztatni akarja a fejtételben kifejezett normatív szándékot (logikailag: a fejtétel deontikus operátorának hatóköre alá sorolja be a listatételek összességét). Olyan esetekben, amikor a fejtételben megengedést fejeznek ki, akkor gyakran jelentkezik a szabad választású engedély paradoxona, ami első látásra meglepő következménnyel jár.²⁷ Amikor ugyanis megengedő vagy-gyal (diszjunkcióval) kötünk össze jogi feltételeket (például: „a gondnok vagy a tulajdonos használhatja”), akkor ha az engedélyt külön-külön karjuk kifejezni a két szereplőre, az így létrejött két összetevőt már és-kapcsolattal (konjunkcióval) kell összefognunk, hogy ugyanazt a tartalmat fejezzük ki továbbra is („a gondnok használhatja és a tulajdonos használhatja”). Ezt a tételt fontos figyelembe venni akkor, amikor a Goody-listák hiányos mondataiból teljes Goody-mondatokat akarunk előállítani és jelezni akarjuk a köztük levő logikai kapcsolatokat is. A vonatkozó jogszabályi rendelet a belső kapcsolatokra vonatkozóan négyféle típus előfordulását tételezi, amelyek mindegyikére találtunk is számos példát a gépi feldolgozás során.²⁸

A formai hibák további típusaként említhetjük még azt is, amikor az IRM-rendelet szerint a felsorolásokban a listák utolsó előtti elemének végére ki kell tenni a listaelemek közti kapcsolatot jelző kötőszót, de ezt elég gyakran elmulasztják, vagy ami talán még zavaróbb, sok esetben logikai értelemben másként kell minősítenünk a fennálló tényleges belső kapcsolatot a lista összetevői között ahhoz képest, mint amit a kirakott kötőszó hétköznapi értelmezése sugall: a vagyos kapcsolatok sok esetben konjunkcióval összekötött mondatokban bukkannak fel, ami tulajdonképpen szándékolatlan megmutatkozása annak, ahogyan a fentebb említett szabad választású paradoxon mögötti megfeleltetés hogyan jelenik meg a természetes nyelvben. Ez a jelenség előfordul másfajta természetes nyelvű szövegben is, ebben a tekintetben a nyelvhasználatunk pongyola, vagy ha úgy tetszik, esendő. Azonban – mint minden pongyolaság – lehetőség szerint ez is kerülendő a jogszabályszövegezés során; s egyúttal tovább keresendő az, hogy az elkerülhetetlen nyelvi esendőségnél a gépet miképpen tanítsuk meg a helyes értelmezésekre.

Hivatkozások

- Bánréti Z.: Az ellipszis mondattana és a lexikai szelekció. In *Nyelvtudományi Közlemények*, 1998, 98, 7–70. o.
- Tim Berners-Lee, James Hendler, and Ora Lassila: The Semantic Web. *Scientific American*, May 2001
- Goody, J.: Nyelv és írás. In Nyíri K. – Szécsi G. (szerk.): Szóbeliség és írásbeliség. Áron Kiadó, Budapest, 1988, 189–221. o.
- Markovich R. – Hamp G. – Syi: A kondicionálisok problémája jogszabályszövegekben. In Tanács A. – Varga V. – Vincze V. (szerk.): X. Magyar Számítógépes Nyelvészeti Konferencia. Szeged, 2014, 295–302. o.

²⁷ Ezt a kérdéskört alaposabban kifejtjük egy másik cikkünkben: Hamp et al. 2015. i. m.

²⁸ Logikailag érdekes problémát jelent az egyes listákban előírt sorrendiség kérdése. Vannak ugyanis olyan felsorolások (pl. Ptk. 7:84. § (1) bekezdés), amelyekben a felsorolt listatételekben rögzített jogi feltételek teljesülésére valamilyen sorrendiséget írnak elő, és kérdés, hogy hogyan lehet mindezt megragadni a logika eszköztárával. Amikor egy sorrendiséget írnak elő valamilyen feltételek (dolgok, ágensek stb.) között (mondjuk öröklési sorrend esetén), akkor az logikailag együtt jár azzal, hogy kizáró vagy-os kapcsolatot feltételezhetünk az elemek között, hiszen akármelyik listatételben rögzített feltétel esetén lép életbe a joghatás, ez együtt jár azzal, hogy az összes többi feltétel már nem teljesülhet, és ez a logikai mintázat a ‘kizáró vagy’-os logikai műveletre (vagy másként: a biszubtrakcióra) jellemző. Azzal a minősítéssel azonban, hogy azt mondjuk, hogy egy felsorolásban biszubtrakciós művelet van az összetevők között, még nem tudjuk megragadni a jogszabálynak rögzített sorrendiséget. Ennek az adekvát logikai leírását és gépi megoldását még keresnünk kell.

- Hamp G. – Syi – Markovich R.: Elliptikus listák jogszabálysövegekben. In Tanács A. – Varga V. – Vincze V. (szerk.) XI. Magyar Számítógépes Nyelvészeti Konferencia. Szeged, 2015, 273–281. o.
- Syi – Hamp G. – Markovich R.: Goody-listák jogszabálysövegekben. (kutatói műhelytanulmány) 2014
- Stainton, R. J.: Words and Thoughts. Substances, Ellipsis and the Philosophy of Language. OUP, 2006
- Syi: syi.hu/cse. L'Harmattan – Könyvpont Kiadó, Budapest, 2014