

Miskolci Egyetem

GÉPÉSZMÉRNÖKI ÉS INFORMATIKAI KAR

**FOGALOMHÁLÓK ALKALMAZÁSA
OSZTÁLYFELBONTÁSI PROBLÉMÁKRA**

PhD értekezés

KÉSZÍTETTE:

KÖREI ATTILA

OKLEVELES MATEMATIKUS

HATVANY JÓZSEF INFORMATIKAI TUDOMÁNYOK DOKTORI ISKOLA

DOKTORI ISKOLA VEZETŐ:

Prof. Dr. Tóth Tibor

A MŰSZAKI TUDOMÁNY DOKTORA

TUDOMÁNYOS VEZETŐ:

Dr. Radeleczki Sándor

A MATEMATIKAI TUDOMÁNY KANDIDÁTUSA

Miskolc, 2008

A témavezető ajánlása
Körei Attila: „Fogalomhálók alkalmazása osztályfelbontási problémákra” című
PhD értekezéséhez

Körei Attila jelölt szakmai munkáját már 14 éve ismerem. Kezdetben fraktálokkal és valós függvényekkel foglalkozott és sikeresen bekapcsolódott a Miskolci Egyetem Analízis Tanszékének, később pedig az Alkalmazott Informatikai Intézetének a munkájába. Oktatói tevékenységét mindig maximális lelkiismeretességgel végezte. 2003 szeptemberében megkezdte tanulmányait a Hatvany József Informatikai Tudományok Doktori iskola keretében, ahol érdeklődése középpontjába diszkrét matematikai struktúrák műszaki-informatikai alkalmazásai kerültek, így lettem az adott témában a választott témavezetője. Az elmúlt évek eredményes és kiváló színvonalú munkájának köszönhetően sikeres abszolutóriumot szerzett és elindította doktori eljárását.

A jelölt doktori disszertációjában fogalomhálóknak és ezekhez kapcsolódó ún. dobozhálóknak műszaki osztályozási (klaszterezési) feladatokban való alkalmazásával foglalkozik. A dolgozatban tovább tökéletesítette azt a modellt, ami az osztályozási feladatot egy adott kontextushoz tartozó dobozháló bizonyos elemrendszereinek a kiválasztására vezeti vissza. Ezt a modellt ki is bővítette az ún. kitüntetett tulajdonságok (attribútumok) fogalmának a bevezetésével, amelyek segítségével számos, a csoporttechnológiában felmerülő osztályozási probléma sokkal jobban megragadható. (Ez utóbbi módszernek a teljes kidolgozása a jelölt jelenlegi kutatásaiban is központi szerepet kap.) Algoritmust fogalmazott meg és implementált annak a klasszikus problémának a megoldására is, hogy egy új munkadarab miként illeszthető be meglévő osztályozási rendszerekbe, a dobozháló aktualizálása révén. Mind a meglévő, mind a kifejlesztett új algoritmusokat leleményes és újtó módon implementálta, MATLAB környezetben elvégezhető egyszerű mátrix műveletekre visszavezetve a feladatokat. Az ún. „atomszerkezet” fogalmának bevezetésével és felhasználásával lényegesen egyszerűsítette egy kontextus dobozhálójának a megkonstruálását. Mindezek lehetővé tették a felhasznált módszerek valós gyakorlati problémákra való alkalmazását is, ahol a kiinduló kontextusok mérete meglehetősen nagy is lehet. Jó példa erre az egyik vizsgált probléma: gyártócellák kialakítása, a munkadarabok és a megmunkálásukra felhasznált gépek attribútumai alapján. Elmondható, hogy a dolgozatban a jelölt által javasolt és implementált algoritmus versenyképes a szakirodalomban ismert módszerekkel.

Az elmúlt évek sikeres munkája és termékeny publikációs tevékenysége igazolta, hogy a jelölt képes színvonalas, önálló kutatómunkára. A disszertáció téziseivel kapcsolatos tudományos eredményeit mind magyar, mind angol nyelven rangos nemzetközi konferenciákon illetve folyóiratokban publikálta. Publikációi között egyaránt megtalálhatók elméleti jellegű és műszaki-informatikai alkalmazásokat tartalmazó publikációk, így megállapítható hogy azok teljes mértékben eleget tesznek a Hatvany József Informatikai Tudományok Doktori Iskola publikációs követelményeinek. Közös OTKA keretében is együtt dolgozhattam a jelölttel, aki itt is igazolta, hogy képes kreatívan bekapcsolódni a csoportos kutatómunkába. Maga az értekezés nagyon aprólékos és gondos munkát tükröz, felépítése, nyelvezete kiváló és esztétikus. A jelöltnek a tézisekhez kapcsolódó eredményei saját, önálló és helyes eredmények. Mindezekre tekintettel, a jelölt számára határozottan javaslom a PhD cím odaítélését.

Miskolc, 2008. április 3.

Dr. Radeleczi Sándor
tudományos vezető

Köszönetnyilvánítás

Az értekezés a Hatvany József Informatikai Tudományok Doktori Iskola keretein belül 2003-ban kezdett kutatómunkám eredményeit foglalja össze. Köszönettel tartozom tudományos vezetőmnek, **Dr. Radeleczki Sándornak**, aki már korábban megismertetett a fogalomhálók elméletével és kutatómunkám irányítását is készséggel vállalta. Együttműködésünk során szakmai és baráti tanácsokkal látott el, értelmes és elérhető célokat tűzött ki elém, a dolgozat matematikai eredményeinek kidolgozásában pedig pótolhatatlan segítséget nyújtott.

Köszönettel tartozom a doktori iskola vezetőjének, **Prof. Dr. Tóth Tibornak**, aki az Alkalmazott Informatikai Tanszékre kerülésem óta ösztönzött a tudományos kutatómunkára és ahhoz minden támogatást megadott. Lehetővé tette, hogy a „Csoporttechnológia-alapú tervezés és ütemezés támogatása diszkrét matematikai modellekkel és módszerekkel” c. OTKA pályázat munkatársaként hazai és külföldi konferenciákon vehessek részt.

Köszönetet mondok **Bernhard Ganter** professzornak, akinek javaslatai nagyban hozzájárultak a kontextus bővítését tárgyaló fejezetben a jelölések egyszerűsítéséhez és így az eredmények érthetőbb közléséhez.

Megköszönöm **Dr. Erdélyi Ferenc** értékes javaslatait, melyek a dolgozat eredményeinek gyakorlati hasznosíthatóságára vonatkoztak.

Köszönet illeti kollégáimat a Miskolci Egyetem Alkalmazott Informatikai Tanszékén és Matematikai Intézetében, akik szakmailag vagy egyéb módon segítettek munkámat.

Hálás vagyok családom tagjainak, különösen feleségemnek, biztatásukért és türelmükért.

Tartalomjegyzék

Jelölések jegyzéke	4
1. Bevezetés	5
1.1. Osztályokra bontás fogalomháló segítségével	6
1.2. A dolgozat felépítése	7
1.3. Irodalmi áttekintés	7
1.4. A kutatás célja	9
2. Háló, fogalomháló, dobozháló	10
2.1. Hálók	10
2.2. Fogalomhálók	13
2.3. Dobozhálók	19
3. Az osztályozási rendszerek meghatározása	24
3.1. A dobozháló atomjainak meghatározása	24
3.2. A dobozelemek meghatározása	27
3.3. Az osztályozási rendszerek előállítása	30
3.4. Az adatok reprezentációja	32
3.5. Az algoritmusok egy lehetséges megvalósítása	35
3.6. Az algoritmusok hatékonysága	42
4. Fogalomháló és dobozháló egyenlősége	45
5. Részkontextus és bővített kontextus	50
5.1. Részkontextushoz tartozó extenzió-partíciók	50
5.2. A kontextus elemi bővítése	52
6. Alkalmazások	58
6.1. Alkatrészek osztályozása	58
6.2. Gyártócellák kialakítása	63
6.2.1. A feladat modellezése gép-alkatrész incidencia mátrixszal	63
6.2.2. A gyártócellák meghatározásának különböző módszerei	66
6.2.3. Gyártócellák meghatározása dobozelemek segítségével .	68
7. Új tudományos eredmények	74
8. További kutatási feladatok	79
9. Összefoglalás	80
Az értekezés témakörében készített saját publikációk	82

TARTALOMJEGYZÉK

Irodalomjegyzék	83
Ábrák, táblázatok, algoritmusok jegyzéke	88

Jelölések jegyzéke

(G, M, I)	formális kontextus
$\mathcal{F}(G, M, I)$	a (G, M, I) kontextushoz tartozó fogalomháló
$\mathcal{D}(G, M, I)$	a (G, M, I) kontextushoz tartozó dobozháló
$J(L)$	az L háló egyesítés-irreducibilis elemeinek halmaza
$\mathcal{C}(L)$	az L háló osztályozási rendszereinek halmaza
$\mathcal{D}(L)$	az L háló dobozelemeinek halmaza
$\mathcal{A}$	a dobozháló atomjainak halmaza
$a \prec b$	a b elem rákövetkezője a -nak
0_P vagy $\mathbf{0}$	a $(P, \leq)$ részbenrendezett halmaz legkisebb eleme
1_P vagy $\mathbf{1}$	a $(P, \leq)$ részbenrendezett halmaz legnagyobb eleme
$\wedge A, \vee A$	az A halmaz infimuma, szuprémuma
$L_1 \cong L_2$	az L_1 struktúra izomorf L_2 -vel
A'	az A halmaz objektumainak közös tulajdonságaiból álló halmaz
B'	a B halmaz attribútumainak mindegyikével rendelkező objektumok halmaza
$Extent(G, M, I)$	a (G, M, I) kontextusból származó extenziók halmaza
$Intent(G, M, I)$	a (G, M, I) kontextusból származó intenziók halmaza
$Ext(a)$	az a fogalom extenziója
Π_0	az objektumhalmaz legfinomabb extenzió-partíciója
$a(H)$	a H objektumhalmaz atomszerkezete
A_g vagy $A_g(G)$	a G halmaz legfinomabb extenzió-partíciójában a g elemet tartalmazó blokk
n_1	a műveletek száma a gép-alkatrész incidencia mátrixban
n_0	a kihasználatlan gépek száma
n_k	kivételes alkatrészek száma
Ω	kivételes alkatrészek aránya
Λ	gépek kihasználtságának mértéke
Γ	csoportképzési határfok
$h_j(a)$	az a alkatrészt megmunkáló gépek aránya a j -edik gyártócellán belül

Az algoritmusok pszeudo-kódjainak leírásában alkalmazott további jelöléseket a 3.4. alfejezet bevezetője tartalmazza.

1. Bevezetés

A műszaki tervezés és irányítás, a termelés menedzsment területén gyakran jelentkező probléma a valós objektumok hasonlóság alapján történő csoportosítása. Ezt a feladatot kell például megoldani, ha különböző alkatrészeket osztályozunk geometriai-fizikai jellemzőik, megmunkálási követelményeik figyelembevételével, ha megmunkáló gépeket kell gyártócellákba sorolni, vagy adott új megrendeléshez (projekthez) valamilyen értelemben hasonló, már megvalósított projekteket keresünk. Az ilyen feladatok a gyakorlati élet más területein is megjelennek: ha vevőket osztályozunk vásárlói szokásaik alapján, ha betegségek besorolását végezzük az észlelt tünetekből kiindulva, vagy dokumentumokat kell csoportosítanunk információtartalmuk szerint. A nagyméretű adathalmazok elemeinek (objektumainak) ismert tulajdonságaik (attribútumaik) alapján való csoportosítása az adatbányászat néven ismert információ-feldolgozási technológiának is az egyik legfontosabb feladata. A csoportok kialakítása osztályozással vagy klaszterezéssel valósítható meg. A két művelet között az az alapvető különbség, hogy az osztályozásnál már van valamilyen előzetes elképzelésünk a választható osztályokról, amelyekbe az elemeket besoroljuk, míg a klaszterezés esetén nem állnak rendelkezésre az osztályok címkéi, az adatok maguk alakítják ki a klaszterek határait. Másképp kifejezve, az osztályozás felügyelt, a klaszterezés pedig felügyelet nélküli csoportosítást jelent. A cél azonban mindkét esetben azonos: a csoportosítandó objektumok halmazának egy olyan partícióját keressük, amelyben a közös osztályba kerülő elempárok lényegesen hasonlóbbak egymáshoz, mint azok, amelyeket eltérő osztályba soroltunk. Az objektumok hasonlósági viszonyait attribútumaik alapján térképezzük fel. Ez a folyamat leggyakrabban az adatmátrixból indul ki, melyben az objektumok mindegyikéhez tartozik egy, az adott objektum tulajdonságait leíró vektor.

A gyártástudomány területén az osztályozással kapcsolatos problémákat csoporttechnológiai szemlélettel érdemes megközelíteni. A csoporttechnológia az alkatrészgyártás átfogó szervezési elve, mely a különféle entitások (alkatrészek, gépek, konstrukciós tervek, technológiai folyamatok) közötti hasonlóságok kihasználásán alapul. A hasonló entitások osztályainak kialakítása elősegíti a gyártási funkciók egységesítését és a szabványosítást, ami az átfutási idők és a költségek csökkenéséhez vezet. A csoporttechnológia koncepciója a gyártási folyamat szinte valamennyi fázisában alkalmazható és meghatározó szerepet játszik a számítógéppel segített konstrukciós tervezés és a számítógéppel támogatott gyártás integrációjában is.

1.1. Osztályokra bontás fogalomháló segítségével

Értekezésemben a hálóelméletből ismert egyik struktúrának, a fogalomhálónak az alkalmazhatóságát vizsgálom meg egy adott objektumhalmaz osztályokra bontásának feladatában. A fő cél az, hogy az azonos osztályba sorolt elemek közötti hasonlóság nagy legyen, a különböző osztályokba tartozó elemek pedig lényegesen különbözzenek egymástól. A feladat tehát megegyezik az említett adatbányászati problémával, az általunk alkalmazott eljárás azonban mind az osztályozástól, mind a klaszterezéstől eltér. Látni fogjuk, hogy a kialakítandó osztályokról nem rendelkezünk előzetes információval, azaz felügyelet nélküli csoportosítást végzünk (az osztályozással ellentétben). Eljárásunk mégsem nevezhető klaszterezésnek, mert az elemek közötti hasonlóság kifejezésére nem használunk mértéket (az általános klaszterező módszerektől eltérően).

A fogalomhálók elméletét a fogalomanalízis tárgyalja. Ez a fiatal tudományterület 1980 körül, egy darmstadti kutatócsoport munkájának eredményeként született és fejlődése azóta is töretlen. A fogalomháló elemei a fogalmak, melyek alatt objektumok és tulajdonságok szorosan összetartozó egységeit értjük, és amelyek az emberi gondolkodásban megjelenő fogalmak absztrakt megfelelőinek tekinthetők. A fogalmak generálása a formális kontextusból kiindulva végezhető, mely az adatmátrixhoz hasonlóan objektumok egy halmazának, az objektumokat jellemző tulajdonságoknak és e két halmaz kapcsolatának leírását adja. Minden fogalom egy párból áll, melynek egyik eleme egy objektumhalmaz, a másik pedig egy szorosan az objektumokhoz köthető tulajdonsághalmaz. A fogalmak hálóba rendezésével fogalmak, fogalmi rendszerek hierarchiáját építhetjük fel, és az adathalmaz elemei közötti rejtett összefüggésekre deríthetünk fényt.

Dolgozatomban a fogalomhálók elméletének egy kevésbé vizsgált alkalmazásával, az objektumhalmaz elemeinek osztályokra bontásával foglalkozom. A fogalomháló és az azt alkotó fogalmak ismerete azért használható az ilyen jellegű feladatok megoldásában, mert a fogalmak objektumrészei már kijelölnek egyfajta csoportokat, és e csoportok elemeit éppen közös tulajdonságaik alapján válogattuk össze. A kutatómunka egyik legfontosabb eredményeként megmutatjuk, hogy az osztályok meghatározásához nem szükséges az összes fogalom ismerete. Ez azért nagy előny, mert egy fogalomháló a kiindulásul szolgáló kontextus méretétől függően kezelhetetlenül sok elemet is tartalmazhat. Az objektumhalmaz felbontását egy kevesebb elemből álló, szabályosabb struktúra, a dobozháló segítségével végezzük. A dobozháló kedvező tulajdonságait kihasználva a keresett osztályok kialakítását célzó, hatékony algoritmusok fejleszthetők.

1.2. A dolgozat felépítése

Az értekezés e bevezető szakaszában a téma irodalmi áttekintése után megfogalmazom a kutatás célkitűzéseit. A második fejezetben összefoglalom a kutatómunka elméleti háttéréül szolgáló matematikai ismereteket, kitérve a hálóelmélet és fogalomanalízis legalapvetőbb összefüggéseire, valamint a dobozhálókkal kapcsolatos eredményekre. A harmadik fejezetben részletesen ismertetem az objektumhalmaz felbontására kifejlesztett új módszer lépéseit és azok matematikai vonatkozásait. Ebben a részben a módszer megvalósítására alkalmas algoritmusok mellett egy olyan adatrepresentációs technikát is bemutatok, amely a feladat programozása során igen hasznosnak bizonyult. Az új osztályfelbontási technika alapját a dobozháló jelenti, amelynek elemei a fogalomhálóból kerülnek ki. A kutatás során felmerült a kérdés, hogy van-e olyan fogalomháló, melyben minden fogalom lényeges az objektumhalmaz felbontása szempontjából, azaz amelynek minden eleme egyúttal a dobozhálónak is eleme. Az adott kontextusból származtatható fogalomháló és dobozháló egyenlőségére fogalmazok meg szükséges és elégséges feltételt a negyedik fejezetben. Az ötödik fejezetben a kontextus szűkítésével és bővítésével kapcsolatos kérdésekkel foglalkozom. Megvizsgálom, hogyan módosul a dobozháló a kontextus objektumhalmazának megváltoztatásakor, és hogyan generálhatóak az osztályfelbontásokat meghatározó új dobozelemek a korábbi állapotnak megfelelő elemek ismeretében. A hatodik fejezet az elért eredmények gyakorlati hasznosíthatóságát mutatja be a termelési rendszerek és folyamatok tervezése és irányítása területén. Megmutatom, hogyan alkalmazható az új osztályfelbontási módszer csoporttechnológiai problémákban, az alkatrészek osztályozása illetve a gyártócellák kialakításának feladatában. A dolgozat két másik, fontos eredménye is rendelkezik gyakorlati vonatkozással: a negyedik fejezetben bevezetett kitüntetett tulajdonság koncepció az osztályfelbontások szempontjából lényeges attribútumok meghatározását teszi lehetővé, az ötödik fejezet kontextusbővítéssel kapcsolatos eredményei pedig megoldják egy új alkatrész megfelelő családba sorolásának problémáját. Az említett feladatok a Miskolci Egyetem Alkalmazott Informatikai Tanszékének kutatási témái között szerepelnek és azokkal a „Csoporttechnológia-alapú tervezés és ütemezés támogatása diszkrét matematikai modellekkel és módszerekkel” c. OTKA-pályázat keretein belül foglalkoztam. A hetedik fejezetben összefoglalom a kutatás legfontosabb eredményeit és megfogalmazom az értekezés téziseit.

1.3. Irodalmi áttekintés

A kutatási téma adatbányászati alapjait, az osztályozás és klaszterezés viszonyát illetve módszereit Han és Kamber bevezető jellegű művének egy-egy

fejezete tárgyalja [22]. A legfontosabb klaszterező eljárások leírása az Iványi Antal által szerkesztett Informatikai algoritmusok c. könyvben is megtalálható [24]. Mivel az objektumok osztályozása az adatbányászat kezdetektől vizsgált problémája, természetesen önálló munkák is rendelkezésre állnak a témában való elmélyüléshez [25], [38].

A részbenrendezett halmazok és hálók elméletének tanulmányozásához Szász Gábor műve szolgálhat alapul [53], egy újabb mértékadó forrás pedig Czédli Gábor jegyzete, mely már a fogalomhálókkal kapcsolatos alapvető összefüggéseket is tartalmazza [14]. Természetesen a hálóelmélet témakörében számos kiváló, idegen nyelvű monográfia is rendelkezésre áll, például Grätzer György *General Lattice Theory* c. könyve [21], vagy a közérthető formában megírt *Introduction to Lattices and Order* c. munka Davey és Priestley tollából [15].

A fogalomanalízis területén Bernhard Ganter és Rudolph Wille könyve megkerülhetetlen forrásnak számít [19], mely a fogalomhálók elméletének matematikai megalapozása mellett a tématerület kutatásának legfontosabb irányait is kijelöli. A fogalomanalízissel és alkalmazásaival kapcsolatos legújabb eredményekről a téma kutatói által üzemeltetett honlapról kiindulva tájékozódhatunk (<http://fcahome.org.uk>).

A fogalomhálók és az osztályozási, klaszterezési feladatok összekapcsolásának ötlete már korábban is felmerült az irodalomban. Egy lehetséges megközelítés a fogalomanalízis és a durva (közelítő) halmazok elméletének házasságából született *rough concept analysis* [29] alkalmazása. Ez a kutatási terület adja a háttérét Ho T. B. cikkének, aki három klaszterezési feladatra ad modellt és megoldó algoritmust [23]. Ventos és Soldano az extenzió fogalmának tágabb értelmezést adva bevezetnek egy új struktúrát, az alfa-fogalomhálót (*Alpha Galois Lattice*), abból a célból, hogy az eredeti fogalomháló struktúráját megőrző, de kisebb méretű hálót hozzanak létre [60]. Az extenzió fogalmának kiterjesztése a kontextus objektumhalmazának egy partícióján alapszik, tehát itt előzetesen kell az alaphalmazt osztályokra bontanunk.

Az objektumhalmaz felbontásának duális problémáját (azaz az attribútumok halmazának particionálását) Rudolf Wille vetette fel egy 1983-as cikkében [63]. A kontextus attribútumhalmazából kiinduló osztályozási módszerek fejlesztésével sokan foglalkoznak; a Miskolci Egyetemen Kovács László és csoportja végez ilyen jellegű kutatásokat, szöveg-osztályozó algoritmusok létrehozására.

Az említett Wille-cikk volt a kiindulópontja témavezetőm, Radeleczki Sándor kutatásának, aki a dobozháló fogalmának bevezetésével egy alapvető eszközt definiált az osztályfelbontási probléma megoldására. A kutatás eredményeként született publikációk szolgálnak a dobozhálók és a kapcsolódó struktúrák elméletéhez referenciául [40], [41], [42].

Az értekezésben bemutatott gyakorlati feladatok háttérében csoporttech-

nológiai elvek alkalmazása áll. A csoporttechnológia fogalmával, kialakulásával és korai fejlődési szakaszaival foglalkozik a [17] tanulmány. Az utóbbi húsz évben a gyártási technológiák és a számítógépes módszerek fejlődésének köszönhetően a csoporttechnológia fogalomköre jelentősen kibővült, a tervezés és gyártás integrációjának fontos tényezője lett [54]. A csoporttechnológiai elvek korszerű alkalmazási területeiről nyújtanak áttekintést a [28] és [52] kiadványok. A dolgozatban részletesen vizsgált témakörben, a gyártócellák és alkatrészcsaládok meghatározásának területén az elmúlt harminc évben számtalan publikáció született, elsősorban a feladat újabb és újabb megoldási módszereit ismertetve. E módszerek csoportosítására a Shafer által javasolt kategóriákat tekintettem mérvadónak [50], az egyes technikákat pedig a [5], [6], [34], [56], [59] publikációk alapján tanulmányoztam.

1.4. A kutatás célja

A tudományos kutatómunka célkitűzéseit három fő feladat köré szervezve fogalmaztam meg.

1. Az objektumok hasonlóság alapján történő csoportosítását megvalósító, fogalomanalízisen alapuló módszer elméleti háttérének kidolgozása. Olyan matematikai eredmények bizonyítása, amelyek alapján az osztályfelbontás a fogalomháló megkerülésével, a dobozháló elemeinek ismeretében valósítható meg. Az objektumokat és tulajdonságaikat leíró kontextus vizsgálata és manipulálása annak érdekében, hogy a leendő osztályok kialakításában döntő szerephez jutó tulajdonságokat előre meghatározhassuk. Az objektumhalmaz bővítésének következményeit leíró elmélet kidolgozása.
2. A matematikai háttér kutatásával párhuzamosan az új osztályfelbontási módszer feladatmegoldó részének algoritmizálása. Számítógépes program készítése, mely egy adott kontextusból kiindulva meghatározza az objektumhalmaz keresett partícióit. A program legyen képes a kontextusbővítés feladatának kezelésére, azaz új objektum hozzáadása esetén végezze el annak besorolását vagy létesítsen új osztályt számára.
3. Az elméleti eredmények hasznosítása a termelési rendszerek és folyamatok tervezése és irányítása területén. A Miskolci Egyetem Alkalmazott Informatikai Tanszékének termelésinformatikai kutatási projektjeiben megjelenő osztályozási feladatok elemzése, azok modellezése. A csoporttechnológiai témájú kutatási projektben megfogalmazott feladatok részletes analízise, megoldó algoritmus és számítógépes alkalmazás kidolgozása.

2. Háló, fogalomháló, dobozháló

A részbenrendezett halmazok és a hálók egyszerűen definiálható, mégis általánosan használható algebrai eszközök, amelyek segítségével egy adott halmazba tartozó elemek közötti hierarchikus viszonyok szemléletesen megjeleníthetők. Ebben a fejezetben bemutatjuk kutatásunk elméleti hátterét: a részbenrendezett halmaz általános fogalmából kiindulva megadjuk a hálók, fogalomhálók és dobozhálók definícióját és tulajdonságait. A rövidség kedvéért csak a későbbi fejezetek megértése szempontjából fontos fogalmakra és tulajdonságokra térünk ki, az eredményeket is bizonyítás nélkül közöljük. A hálókkal kapcsolatban magyar nyelven Szász Gábor [53] vagy Czédli Gábor [14] könyve szolgálhat referenciául, de számos idegennyelvű forrás is rendelkezésre áll [15], [21]. A fogalomhálók tulajdonságait vizsgáló fogalomanalízis alapvető műve Bernhard Ganter és Rudolph Wille könyve [19]. A fejezet utolsó részében tárgyalt osztályozási hálókra és dobozhálókra vonatkozó eredmények igazolása Radeleczki Sándor cikkeiben található [41], [42].

2.1. Hálók

2.1. Definíció. Egy P nemüres halmaz elemei között értelmezett és " $\leq$ "-vel jelölt bináris relációt **részbenrendezésnek** nevezünk, ha minden $x, y, z \in P$ -re teljesülnek az alábbi feltételek:

1. $x \leq x$ (reflexivitás),
2. $x \leq y$ és $y \leq x$ esetén $x = y$ (antiszimmetria),
3. $x \leq y$ és $y \leq z$ esetén $x \leq z$ (transzitivitás).

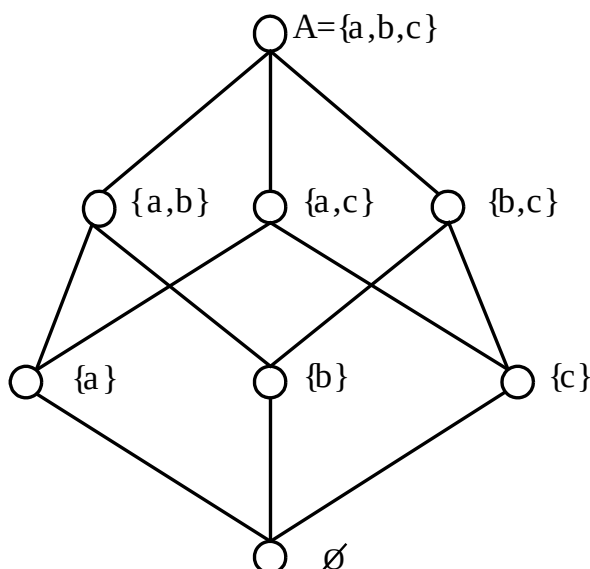
Ha a nemüres P halmazon adott egy $\leq$ részbenrendezés, akkor P -t (pontosabban a $(P, \leq)$ párt) **részbenrendezett halmaznak** nevezzük. A $(P, \leq)$ részbenrendezett halmaz a, b elemére az $a < b$ jelölést használjuk, ha $a \leq b$, de $a \neq b$. Az $a \prec b$ jelölés pedig azt fejezi ki, hogy b **rákövetkezője** a -nak, azaz $a < b$ és nincs olyan $x \in P$, amelyre $a < x < b$ teljesül.

Példa: Jelölje $P(A)$ az A halmaz hatványhalmazát, azaz A részhalmazainak összességét. A részhalmazok közötti $\subseteq$ tartalmazási reláció ekkor részbenrendezése $P(A)$ -nak, $(P(A), \subseteq)$ pedig részbenrendezett halmaz. Ha $B, C \in P(A)$, és $B \subseteq C$, akkor C rákövetkezője B -nek, ha a $C \setminus B$ halmaz egyelemű.

2.2. Definíció. Legyen $(P, \leq)$ egy részbenrendezett halmaz. Azt mondjuk, hogy $a \in P$ a P **legkisebb eleme**, ha minden $x \in P$ -re $a \leq x$. Hasonlóan, $b \in P$ -t a P **legnagyobb elemének** nevezzük, ha minden $x \in P$ -re $x \leq b$. A $(P, \leq)$ részbenrendezett halmaz legkisebb elemét 0_P -vel (vagy egyszerűen **0**-val), a legnagyobbat 1_P -vel (vagy **1**-el) jelöljük.

Példa: $(P(A), \subseteq)$ legkisebb eleme az üres halmaz, legnagyobb eleme maga az A halmaz.

2.3. *Megjegyzés.* Minden véges $(P, \leq)$ részbenrendezett halmaz a rákövetkezés fogalmát felhasználva szemléltethető az ún. Hasse diagrammal, a következő módon: A P halmaz elemeinek kis köröket feleltetünk meg a síkon. Ha $a < b$, akkor a b -nek megfelelő kört az a -t jelölő kör fölé rajzoljuk és összekötjük őket egy egyenes szakasszal. Egy ilyen ábrából az $x < y$ reláció teljesülése könnyen leolvasható: $x < y$ pontosan akkor igaz, ha az x -hez tartozó körből vezet az y -hoz tartozó körbe végig „felfelé” haladó töröttvonal. Az 1. ábrán az $A = \{a, b, c\}$ háromelemű halmaz esetén ábrázoltuk a $(P(A), \subseteq)$ részbenrendezett halmazt.



1. ábra. Háromelemű halmaz részhalmazhálójának Hasse diagramja

2.4. Definíció. Legyen $(P, \leq)$ egy részbenrendezett halmaz, és A a P egy nemüres részhalmaza. A $p \in P$ elemet az A halmaz **alsó korlátjának** nevezzük, ha $p \leq a$ minden $a \in A$ -ra, és hasonlóan, a $p \in P$ elem az A halmaz **felső korlátja**, ha $p \geq a$ minden $a \in A$ -ra. A p elemet az A halmaz legnagyobb alsó korlátjának (**infimumának**) nevezzük, ha p alsó korlátja A -nak, és A minden egyéb k alsó korlátja esetén $p \geq k$. Az A halmaz infimumát gyakran elemei metszetének nevezzük, és $\inf A$ -val vagy $\wedge A$ -val jelöljük. A p elemet az A halmaz legkisebb felső korlátjának (**szuprémumának**) nevezzük, ha p felső korlátja A -nak, és A minden egyéb k felső korlátjára $p \leq k$ teljesül. Az A halmaz szuprémumát gyakran elemei egyesítésének is nevezzük és $\sup A$ -val vagy $\vee A$ -val jelöljük.

2.5. Definíció. Az $(L, \leq)$ részbenrendezett halmazt **hálónak** nevezzük, ha bármely $a, b \in L$ -re $\inf \{a, b\}$ és $\sup \{a, b\}$ létezik. Ha $\inf H$ és $\sup H$ minden $H \subseteq L$ nemüres részhalmazra létezik, akkor $(L, \leq)$ -t **teljes hálónak** nevezzük.

2.6. Megjegyzés. (i) Ha $(L, \leq)$ egy véges háló, akkor teljes is, hiszen bármely $H = \{h_1, h_2, \dots, h_n\} \subseteq L$ halmazra n szerinti teljes indukcióval beláthatjuk $\inf H$ és $\sup H$ létezését.

(ii) Minden $(L, \leq)$ teljes hálónak van legkisebb és legnagyobb eleme, méghozzá $\mathbf{0} = \inf L$ és $\mathbf{1} = \sup L$.

A hálók kétműveletes algebrai struktúráként is definiálhatóak.

2.7. Definíció. Az $(L, \wedge, \vee)$ algebrai struktúrát hálónak nevezzük, ha $\wedge$ és $\vee$ bináris műveleteket jelölnek az L nemüres halmaz elemei között, úgy, hogy $\forall x, y, z \in L$ -re az alábbi tulajdonságok teljesülnek:

1. $x \wedge y = y \wedge x$ és $x \vee y = y \vee x$ (kommutativitás),
2. $x \wedge (y \wedge z) = (x \wedge y) \wedge z$ és $x \vee (y \vee z) = (x \vee y) \vee z$ (asszociativitás),
3. $x \wedge x = x$ és $x \vee x = x$ (idempotencia),
4. $x \vee (x \wedge y) = x$ és $x \wedge (x \vee y) = x$ (elnyelési tulajdonságok).

Ha egy $(L, \leq)$ részbenrendezett halmaz háló a 2.5. Definíció értelmében, akkor az $a \wedge b = \inf \{a, b\}$ és $a \vee b = \sup \{a, b\}$ összefüggéssel definiált műveletek teljesítik az előző definícióban felsorolt tulajdonságokat, így $(L, \wedge, \vee)$ háló a 2.7. Definíció értelmében is. Megfordítva, ha adott az $(L, \wedge, \vee)$ algebrai struktúra, ami háló a 2.7. Definíció értelmében, akkor az $a \leq b \Leftrightarrow a = a \wedge b$ összefüggés egy részbenrendezést definiál, és igazolható, hogy az $(L, \leq)$ részbenrendezett halmaz háló a 2.5. Definíció értelmében. Ezen megfeleltetés értelmében például a $(P(A), \subseteq)$ hálózhoz a $(P(A), \cap, \cup)$ algebrai struktúra rendelhető hozzá.

A továbbiakban a hálókra csak az alaphalmaz megadásával hivatkozunk, ha ez nem okozhat félreértést, így $(L, \leq)$ illetve $(L, \wedge, \vee)$ helyett is csak L -et írunk. Láthattuk, hogy a hálók kétféle módon való definíciója ugyanahhoz a struktúrához vezetett. Két háló azonban elemeik különbözősége esetén is lehet algebrai szempontból teljesen azonos, azaz izomorf.

2.8. Definíció. Az L_1 és L_2 hálókat **izomorfaknak** hívjuk, ha van olyan $\Phi : L_1 \rightarrow L_2$ bijektív leképezés, hogy $a \leq b$ L_1 -ben pontosan akkor teljesül, ha $\Phi(a) \leq \Phi(b)$ L_2 -ben. Jelölés: $L_1 \cong L_2$.

Ha az L_1 és L_2 hálók a 2.7. Definícióval adottak, akkor izomorfiájuk azt jelenti, hogy van olyan $\Phi: L_1 \rightarrow L_2$ bijektív leképezés amelyre $\Phi(a \wedge b) = \Phi(a) \wedge \Phi(b)$ és $\Phi(a \vee b) = \Phi(a) \vee \Phi(b)$ teljesül $\forall a, b \in L_1$ esetén.

A definícióból következik, hogy az izomorf hálókat ugyanazzal a Hasse diagrammal szemléltethetjük.

A következő két definícióval speciális tulajdonságú hálóelemeket, illetve speciális hálókat határozunk meg.

2.9. Definíció. Legyen p az L háló egy $\mathbf{0}$ -tól különböző eleme. Ha a $p = \bigwedge \{x_j \mid x_j \in L, j \in J\}$ egyenlőségből $p = x_{j_0}$ következik valamely $j_0 \in J$ -re, akkor p -t (teljesen) **metszet-irreducibilis** elemnek nevezzük. Hasonlóképpen, ha $p = \bigvee \{x_j \mid x_j \in L, j \in J\}$ csak akkor lehetséges, ha $p = x_{j_0}$ valamely $j_0 \in J$ -re, akkor p -t (teljesen) **egyesítés-irreducibilis** elemnek nevezzük. Másképp fogalmazva, egy metszet-irreducibilis elem nem állhat elő más elemek infimumaként, egy egyesítés-irreducibilis elem pedig más elemek szuprémumaként. Az L hálót **CJ-generált hálónak** nevezzük, ha minden, $\mathbf{0}$ -tól különböző eleme előáll egyesítés-irreducibilis elemek szuprémumaként.

A CJ rövidítés az angol *completely join-irreducible* szavakból ered, a CJ-generált hálókkal még találkozunk a későbbi fejezetekben.

2.10. Definíció. Ha az L hálónak van legkisebb eleme, akkor annak rákövetkezőit a háló **atomjainak** nevezzük. Tehát $a \in L$ atom, ha $\mathbf{0} \prec a$. Az L hálót **atomisztikusnak** nevezzük, ha minden eleme előáll atomok szuprémumaként, azaz $\forall x \in L$ esetén $x = \bigvee \{a \mid a \leq x, \mathbf{0} \prec a\}$.

Az előző két definícióból adódik, hogy egy háló atomjai (ha léteznek) mindig egyesítés-irreducibilis elemek.

2.2. Fogalomhálók

Rudolf Wille 1982-ben megjelent cikkével fogalomanalízis néven új irányzat jelent meg a hálóelméletben, azzal a céllal, hogy a hálóelmélet módszereinek különböző alkalmazási lehetőségeit vizsgálja [62]. A fogalomanalízisben definiált fogalmak adott objektumok és az őket jellemző tulajdonságok egységeként értelmezhetőek. A fogalmak hálót alkotnak, így a közöttük fennálló hierarchia szemléletes módon megragadható, a fogalomhálókat pedig adatok, ismeretek tömör tárolására és rejtett összefüggések feltárására is használhatjuk. Ennek köszönhető, hogy az elmúlt három évtizedben a fogalomanalízis az alkalmazott hálóelmélet egyik leggyorsabban fejlődő területévé vált, és a matematikusokon kívül a legkülönbözőbb tudományterületek kutatói (nyelvészek, biológusok, fizikusok) fedezték fel a benne rejlő lehetőségeket.

2.11. Definíció. Legyen adott objektumoknak egy G halmaza, M pedig olyan tulajdonságoknak (attribútumoknak) a halmaza, melyekkel a G -beli objektumok rendelkezhetnek. Legyen továbbá $I \subseteq G \times M$ egy reláció, melynél $(g, m) \in I$ azt jelenti, hogy a g objektum rendelkezik az m tulajdonsággal. A (G, M, I) hármast (formális) **kontextusnak** nevezzük. Ha G és M egyaránt véges halmazok, akkor azt mondjuk, hogy a (G, M, I) kontextus **véges**.

A kisebb méretű kontextusok táblázatban szemléltethetők, melynek i -edik sorának j -edik oszlopába akkor teszünk x-et, ha az i -edik objektum rendelkezik a j -edik tulajdonsággal.

1. táblázat. Példa kontextusra

K	Attribútumok (M)						
Objektumok (G)		m ₁	m ₂	m ₃	m ₄	m ₅	m ₆
	g ₁	x	x				
	g ₂				x	x	x
	g ₃	x				x	x
	g ₄		x	x		x	
	g ₅		x	x		x	x
	g ₆	x	x	x			
	g ₇			x	x	x	x
	g ₈		x		x		x

2.12. Definíció. Vezessük be a következő jelöléseket: $A \subseteq G$ esetén legyen

$$A' = \{m \in M \mid (g, m) \in I, \quad \forall g \in A\}$$

az A -beli objektumok közös tulajdonságainak halmaza, és $B \subseteq M$ esetén

$$B' = \{g \in G \mid (g, m) \in I, \quad \forall m \in B\}$$

azoknak az objektumoknak a halmaza, melyek minden B -beli tulajdonsággal rendelkeznek. Az (A, B) párt a (G, M, I) kontextushoz tartozó **fogalomnak** nevezzük, ha

$$A \subseteq G, \quad B \subseteq M, \quad A' = B, \quad B' = A.$$

Az A halmazt az (A, B) fogalom objektumrészének vagy **extenziójának**, míg a B halmazt az (A, B) fogalom tulajdonságrészének vagy **intenziójának** nevezzük.

Az 1. táblázat kontextusának egy fogalma például a $(\{g_4, g_5, g_6\}, \{m_2, m_3\})$ halmazpár. Mivel egy (A, B) fogalom A extenziója és B intenziója egymást kölcsönösen meghatározza, egy fogalom ismeretéhez elegendő a halmazpár „felének” ismerete.

A továbbiakban $Extent(G, M, I)$ -vel, illetve $Intent(G, M, I)$ -vel jelöljük a (G, M, I) kontextusból származtatható extenziók illetve intenziók halmazát. Az előző definícióban szereplő $()'$ operátorral kapcsolatos, legfontosabb tulajdonságokat foglalja össze a következő állítás.

2.13. Állítás. *Legyen (G, M, I) egy kontextus, A, A_1, A_2 objektumok tetszőleges halmazai, B, B_1, B_2 pedig attribútumokból álló halmazok. Ekkor*

1. $A_1 \subseteq A_2 \Rightarrow A'_2 \subseteq A'_1$ és $B_1 \subseteq B_2 \Rightarrow B'_2 \subseteq B'_1$.
2. $A \subseteq A''$ és $B \subseteq B''$.
3. $A' = A'''$ és $B' = B'''$.
4. (A'', A') és (B', B'') fogalmak.
5. A'' a legszűkebb A -t tartalmazó extenzió és B'' a legszűkebb B -t tartalmazó intenzió.
6. $A \in Extent(G, M, I) \Leftrightarrow A = A''$ és $B \in Intent(G, M, I) \Leftrightarrow B = B''$.

A következő állítás szerint tetszőleges számú extenzió metszete is egy extenzió és ugyanez az intenziókra is fennáll.

2.14. Állítás. *Legyen J egy tetszőleges indexhalmaz és (A_j, B_j) a (G, M, I) kontextushoz tartozó fogalmak $\forall j \in J$ -re. Ekkor $\bigcap_{j \in J} A_j \in Extent(G, M, I)$ és $\bigcap_{j \in J} B_j \in Intent(G, M, I)$.*

A következő definíció lehetővé teszi a fogalmak összehasonlítását és egy algebrai struktúrába rendezi azokat.

2.15. Definíció. Legyenek (A_1, B_1) és (A_2, B_2) a $K = (G, M, I)$ kontextushoz tartozó fogalmak. Azt mondjuk, hogy

$$(A_1, B_1) \leq (A_2, B_2), \text{ ha } A_1 \subseteq A_2 \text{ (ekkor } B_2 \subseteq B_1 \text{ is teljesül).}$$

Az így definiált reláció egy részbenrendezés a kontextus fogalmai között. A (G, M, I) kontextushoz tartozó fogalmak összességét a fenti rendezéssel ellátva egy hálót kapunk, melyet a $K = (G, M, I)$ kontextushoz tartozó **fogalomhálónak** nevezünk és $\mathcal{F}(G, M, I)$ -vel vagy $\mathcal{F}(K)$ -val jelölünk.

A következő fontos tétel a fogalomhálóok teljességét állítja.

2.16. Tétel (A fogalomhálóok alaptétele). *Az $\mathcal{F}(G, M, I)$ fogalomháló egy teljes háló, melyben tetszőleges halmazok infimumát illetve szuprémumát a következő összefüggésekkel határozhatjuk meg:*

$$\bigwedge_{j \in J} (A_j, B_j) = \left(\bigcap_{j \in J} A_j, \left(\bigcup_{j \in J} B_j \right)'' \right)$$

$$\bigvee_{j \in J} (A_j, B_j) = \left(\left(\bigcup_{j \in J} A_j \right)'', \bigcap_{j \in J} B_j \right).$$

2.17. Definíció. Legyen (G, M, I) egy kontextus. A $g \in G$ objektumhoz tartozó **objektum-intenzió** a $\{g\}' = \{m \in M \mid (g, m) \in I\}$ halmazt értjük, és röviden g' -vel jelöljük. Hasonlóképpen az $m \in M$ tulajdonsághoz tartozó **attribútum-extenzió** alatt az $\{m\}' = \{g \in G \mid (g, m) \in I\}$ halmazt értjük, melyet m' -vel jelölünk. A (g'', g') alakú fogalmakat **objektum-fogalmaknak**, míg az (m', m'') alakúakat **attribútum-fogalmaknak** nevezzük. (Itt g'' a $\{g'\}'$ halmazt, m'' pedig az $\{m'\}'$ halmazt jelöli.)

2.18. Állítás. *Az $\mathcal{F}(G, M, I)$ fogalomháló minden eleme előáll objektum-fogalmak szuprémumaként illetve attribútum-fogalmak infimumaként a következő formában: ha $(A, B) \in \mathcal{F}(G, M, I)$, akkor*

$$\bigvee_{g \in A} (g'', g') = (A, B) = \bigwedge_{m \in B} (m', m'').$$

Ezt az eredményt összevetve a szuprémum és infimum fogalomhálóokban érvényes kiszámítási módjával (2.16. Tétel), az alábbi következményre jutunk.

2.19. Következmény. *Az $\mathcal{F}(G, M, I)$ fogalomháló minden fogalmának objektumrésze előáll attribútum-extenziók metszeteként, és minden tulajdonság-rész előállítható objektum-intenziók metszeteként.*

A 2.19. Következmény alapján egyszerű eljárást adhatunk egy adott kontextusból származó összes fogalom előállítására. A következő algoritmus egy L listára gyűjti az összes olyan halmazt, mely attribútum-extenziók metszeteként előáll, azaz a folyamat végén L az $\mathcal{F}(G, M, I)$ háló fogalmainak extenzióit tartalmazza. Ezt a módszert gyakran nevezik a fogalomháló-építés „naív” algoritmusának.

2.20. Algoritmus (Fogalom-extenziók meghatározása).

Az L lista első eleme G halmaz (a (G, G') fogalom extenziója)

minden $m \in M$ attribútum esetén

minden $A \in L$ halmazra

ha $A \cap m' \notin L$

akkor az L listát bővítsd az $A \cap m'$ halmazzal.

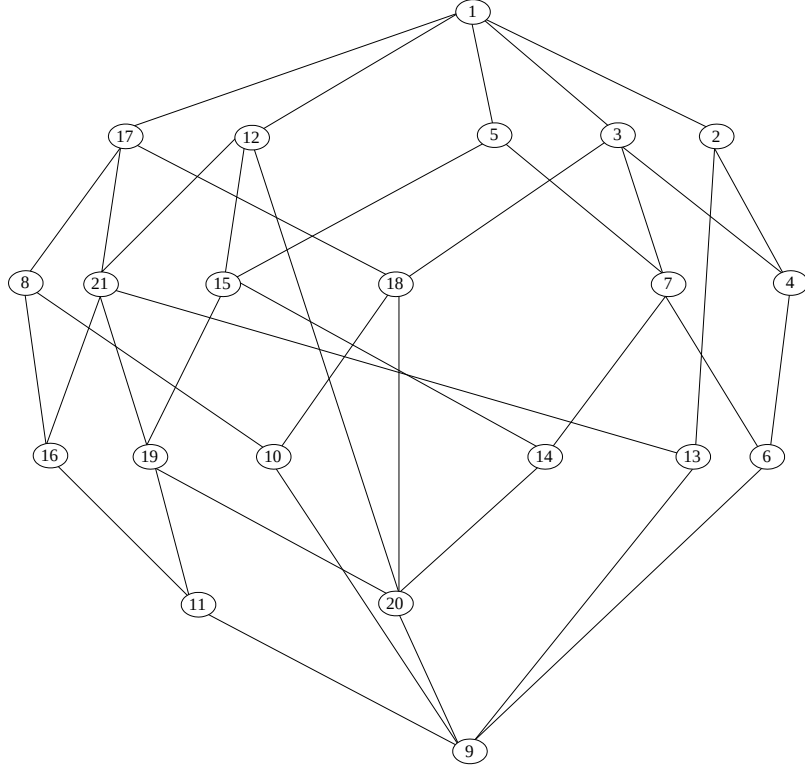
A következő táblázat az algoritmus által előállított L lista elemeinek képzését mutatja be az 1. táblázatban szereplő kontextus esetén. A táblázat középső oszlopában azt az attribútumot jelöltük, mely a 2.20. Algoritmus külső ciklusának változója az adott extenzió előállításakor.

2. táblázat. Fogalom-extenziók meghatározásának algoritmus

Sorszám	Attribútum	Extenzió
1.		$\{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\}$
2.	m_1	$\{g_1, g_3, g_6\}$
3.	m_2	$\{g_1, g_4, g_5, g_6, g_8\}$
4.	m_2	$\{g_1, g_6\}$
5.	m_3	$\{g_4, g_5, g_6, g_7\}$
6.	m_3	$\{g_6\}$
7.	m_3	$\{g_4, g_5, g_6\}$
8.	m_4	$\{g_2, g_7, g_8\}$
9.	m_4	$\emptyset$
10.	m_4	$\{g_8\}$
11.	m_4	$\{g_7\}$
12.	m_5	$\{g_2, g_3, g_4, g_5, g_7\}$
13.	m_5	$\{g_3\}$
14.	m_5	$\{g_4, g_5\}$
15.	m_5	$\{g_4, g_5, g_7\}$
16.	m_5	$\{g_2, g_7\}$
17.	m_6	$\{g_2, g_3, g_5, g_7, g_8\}$
18.	m_6	$\{g_5, g_8\}$
19.	m_6	$\{g_5, g_7\}$
20.	m_6	$\{g_5\}$
21.	m_6	$\{g_2, g_3, g_5, g_7\}$

A listán szereplő extenziók és a kontextus alapján a fogalmak intenzióit is előállíthatjuk. Az így kapott fogalomháló Hasse diagramja a 2. ábrán látható. Minden csomópont egy fogalmat jelöl, a csomópontot jelölő szám pedig

megfelel az adott fogalom extenziójához tartozó, a 2. táblázatban szereplő sorszámoknak.



2. ábra. Fogalomháló Hasse diagramja

A 2.20. Algoritmus a legegyszerűbb módszer a fogalom-extenziók meghatározására, nagyméretű kontextusok esetén azonban érdemes hatékonyabb eljárást alkalmazni. Ilyen például a G részhalmazainak lexikografikus rendezésén alapuló Ganter-féle *next closure* algoritmus [19], vagy Godin inkrementális módszere, mely a kontextus fokozatos bővítésével állítja elő a fogalmak halmazát [20]. Az inkrementális módszer továbbfejlesztésével Kovács László az eddigieknél még hatékonyabb hálóépítő algoritmust javasolt [31].

A továbbiakban a kontextus olyan átalakításairól lesz szó, melyek nem befolyásolják a kontextushoz tartozó fogalomháló szerkezetét, azaz a módosított kontextushoz tartozó fogalomháló izomorf az eredeti kontextus fogalomhálójával. Az egyik ilyen művelet a kontextusban szereplő azonos sorok illetve oszlopok „összefésülése”, ilyenkor a megegyező sorokból illetve oszlopokból elegendő csak egyet megtartani a kontextusban.

2.21. Definíció. Egy (G, M, I) kontextust **tisztított**nek nevezünk, ha minden $g, h \in G$ objektumra $g' = h'$ csak $g = h$ esetén lehetséges, és minden

$m, n \in M$ attribútumra $m' = n'$ csak $m = n$ esetén következhet be.

A kontextus redukciója pedig azt jelenti, hogy töröljük azokat a g objektumokat, amelyek esetén a g' halmaz előáll más objektum-intenziók metszeteként, illetve azokat az m tulajdonságokat, amelyekre m' előáll attribútum-extenziók metszeteként.

2.22. Definíció. Egy tisztított (G, M, I) kontextust **sor-redukálnak** nevezünk, ha minden $g \in G$ objektum esetén a (g'', g') fogalom egyesítés-irreducibilis. A kontextust **oszlop-redukálnak** hívjuk, ha minden $m \in M$ -re az (m', m'') fogalom metszet-irreducibilis. Ha egy kontextus sor és oszlop-redukált is, akkor **redukálnak** nevezzük.

A következő állítás szerint minden véges háléhoz megadható egy olyan redukált kontextus, melyből a hálóval izomorf fogalomháló származtatható.

2.23. Állítás. Minden véges V háló esetén van egy olyan redukált $K(V)$ kontextus, amelyre $V \cong \mathcal{F}(K(V))$ teljesül. A kontextus objektumhalmaza megegyezik a V háló egyesítés-irreducibilis elemeinek halmazával, attribútumhalmaza pedig a metszet-irreducibilis elemek halmazával, a reláció pedig a V háló $\leq$ részbenrendezése.

2.3. Dobozhálók

Ebben a fejezetben bemutatjuk azokat az eszközöket, amelyek segítségével a fogalomhálókat osztályfelbontási feladatok megoldására is alkalmazhatjuk. Az objektumok csoportosítását azok közös tulajdonságai alapján végezzük, úgy, hogy az egy osztályba sorolt objektumok rendelkeznek egy adott tulajdonságcsoporthal, de az osztályon kívüli objektumok közül már egyik sem rendelkezhet e tulajdonságok mindegyikével. Ez a fogalomanalízis terminológiájára lefordítva azt jelenti, hogy az objektumok osztályai és az osztályokat jellemző tulajdonságcsoporthok fogalmakat alkotnak.

2.24. Definíció. Legyen G egy halmaz és $P = \{G_j \mid j \in J\}$ a G halmaz nemüres részhalmazáiból álló rendszer. A P halmazrendszert G egy **partíciójának** nevezzük, ha $\bigcup_{j \in J} G_j = G$ és $G_j \cap G_k = \emptyset$ minden $j \neq k$ esetén. A G_j részhalmazokat a P partíció **blokkjainak** nevezzük.

Ugyanazon halmaz különböző partícióit a következő definícióban adott módon hasonlíthatjuk össze.

2.25. Definíció. Legyen $P_1 = \{A_i \mid i \in I\}$ és $P_2 = \{B_j \mid j \in J\}$ a G halmaz két partíciója. Azt mondjuk, hogy P_1 **finomabb**, mint P_2 (jelölés: $P_1 \leq P_2$), ha P_2 minden egyes blokkja előáll P_1 -beli blokkok egyesítéseként.

Egy kontextus objektumhalmazának partíciói közül számunkra azok fontosak, amelyeknél a blokkok jellemezhetőek elemeik közös tulajdonságaival.

2.26. Definíció. Legyen (G, M, I) egy kontextus és $P = \{G_j \mid j \in J\}$ a G egy partíciója. Azt mondjuk, hogy a P partíció a G halmaz egy **extenzió-partíciója**, ha $G_j'' = G_j$ teljesül minden $j \in J$ esetén (azaz a G_j blokkok zártak közös tulajdonságaikra nézve).

Az elnevezést az indokolja, hogy a G halmaz egy $P = \{G_j \mid j \in J\}$ extenzió-partíciója esetén, minden egyes G_j blokk egyben az $\mathcal{F}(G, M, I)$ fogalomháló (G_j, G_j') fogalmának az extenziója (az, hogy a (G_j, G_j') pár egy fogalom, a $G_j'' = G_j$ egyenlőség következménye).

2.27. Megjegyzés. A G halmaz extenzió-partícióit Radeleczki Sándor a kutatás alapjául szolgáló cikkekben a G halmaz *osztályozásainak* nevezi. Hogy elkerüljük az adatbányászati terminológiából adódó esetleges félreértéseket, az osztályozás, mint halmazrendszer helyett az extenzió-partíció, az osztályozás, mint csoportképzési művelet helyett az osztályfelbontás szavakat használjuk. Az extenzió-partíciók blokkjait azonban (különösen a műszaki problémákban) gyakran nevezzük osztályoknak.

Az extenzió-partícióknak megfeleltethető fogalmak rendszere a fogalomhálóban is kitüntetett szereppel bír. Az ilyen rendszerek általános tulajdonságait írja le a következő definíció.

2.28. Definíció. Legyen L egy teljes háló. Az L nemüres részhalmazaiából álló $S = \{a_j \mid j \in J\}$ rendszert a háló egy **osztályozási rendszerének** nevezzük, ha

1. $a_j \wedge a_k = \mathbf{0} \quad \forall j \neq k$ esetén,
2. $x = \bigvee_{j \in J} (x \wedge a_j) \quad \forall x \in L.$

A G halmaz extenzió-partíciói és az $\mathcal{F}(G, M, I)$ háló osztályozási rendszerei között az alábbi tétel teremt kapcsolatot [41].

2.29. Tétel. Ha (G, M, I) egy tetszőleges kontextus, és $\{G_j \mid j \in J\}$ a G halmaz egy extenzió-partíciója, akkor a $\{(G_j, G_j') \mid j \in J\}$ rendszer az $\mathcal{F}(G, M, I)$ fogalomháló egy osztályozási rendszere. Megfordítva, ha (G, M, I) egy sor-redukált kontextus és $\{(A_j, B_j) \mid j \in J\}$ az $\mathcal{F}(G, M, I)$ fogalomháló egy osztályozási rendszere, akkor az $\{A_j \mid j \in J\}$ halmazok G egy extenzió-partícióját alkotják.

Ha L egy sor-redukált (G, M, I) kontextusból származó fogalomháló, akkor az előző tétel értelmében L minden S osztályozási rendszerének megfelel a G halmaznak egy P_S extenzió-partíciója. Ezt a kapcsolatot használjuk fel a fogalomháló osztályozási rendszereinek összehasonlítására: azt mondjuk, hogy az S_1 osztályozási rendszer finomabb S_2 -nél, ha az általuk indukált extenzió-partíciókra $P_{S_1} \leq P_{S_2}$ teljesül. Ahhoz, hogy az osztályozási rendszerek közötti rendezési relációt ettől általánosabb formában definiálhassuk, szükségünk van a CJ-generált hálók osztályozási rendszereinek egy tulajdonságára.

Legyen L egy CJ-generált teljes háló és jelölje $J(L)$ a háló egyesítés-irreducibilis elemeinek összességét. Egy $a \in L$ elemre legyen $J(a) = \{p \in J(L) \mid p \leq a\}$, azaz az a -nál kisebb vagy egyenlő egyesítés-irreducibilis elemek halmaza. Igazolható, hogy az L háló egy $S = \{a_k \mid k \in K\}$ osztályozási rendszere a $J(L)$ halmazon egy partíciót indukál, $\Pi_S = \{J(a_k) \mid k \in K\}$ formában, így az osztályozási rendszerek rendezése a $J(L)$ halmaz partíciói segítségével is lehetséges.

2.30. Definíció. Legyen L egy CJ-generált teljes háló, és S_1, S_2 a háló két osztályozási rendszere. Azt mondjuk, hogy S_1 **finomabb** S_2 -nél ($S_1 \leq S_2$), ha az S_1 által a $J(L)$ halmazon indukált Π_{S_1} partíció finomabb S_2 által indukált Π_{S_2} partíciónál.

Megjegyezzük, hogy fogalomhálók esetében az osztályozási rendszerek közötti rendezést végezhetjük az objektumhalmazon indukált partíciók illetve az egyesítés-irreducibilis elemek halmazán indukált partíciók szerint is, mindkét esetben ugyanazt a relációt kapjuk.

Jelölje $\mathcal{C}(L)$ az L teljes háló osztályozási rendszereinek halmazát. Könnyen igazolható, hogy az előzőekben megadott reláció részbenrendezés az osztályozási rendszerek halmazán, tehát $\mathcal{C}(L)$ egy háló. Ettől azonban több is igaz.

2.31. Állítás. Ha L egy CJ-generált teljes háló, akkor a $\mathcal{C}(L)$ háló teljes.

2.32. Definíció. A $\mathcal{C}(L)$ hálót az L háló **osztályozási hálójának** nevezzük, melynek **0**-eleme, azaz a legfinomabb osztályozási rendszer az összes osztályozási rendszer metszeteként áll elő: $S_0 = \bigwedge \{S \mid S \in \mathcal{C}(L)\}$.

Az osztályozási rendszerek meghatározásának megkönnyítésére bevezetünk egy újabb struktúrát.

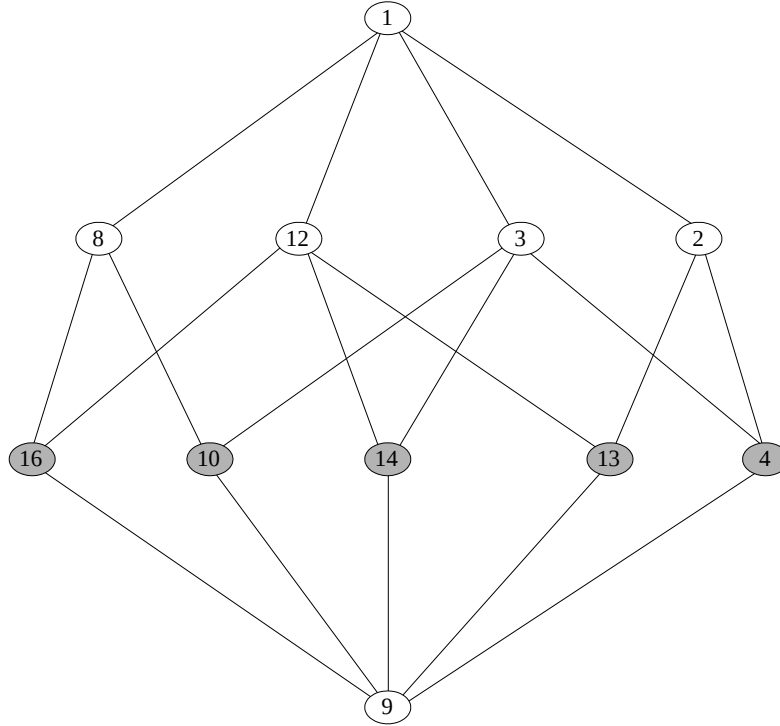
2.33. Definíció. Az L teljes háló egy d elemét **dobozelemnek** nevezzük, ha vagy megegyezik a háló **0** elemével, vagy a háló valamely osztályozási rendszerének egy eleme, azaz $d \in S$, ahol $S \in \mathcal{C}(L)$. Az L háló dobozelemeinek összességét $\mathcal{D}(L)$ -el jelöljük. Az L hálón adott részbenrendezés $\mathcal{D}(L)$ -re való

leszűkítésével a dobozelemek hálót alkotnak, melyet az L hálóhoz tartozó **dobozhálónak** neveziünk. Ha az L háló a $K = (G, M, I)$ kontextushoz tartozó fogalomháló, akkor dobozhálójára a $\mathcal{D}(\mathcal{F}(G, M, I))$ jelölés helyett a rövidebb $\mathcal{D}(G, M, I)$ jelölést használjuk.

Az alábbi tétel a dobozháló jellemzését adja.

2.34. Tétel. *Ha L egy CJ-generált teljes háló, akkor $\mathcal{D}(L)$ egy teljes, atomisztikus háló, melynek atomjai a legfinomabb osztályozási rendszer elemei.*

Az 1. táblázat kontextusából származtatható dobozháló Hasse diagramja a 3. ábrán látható, ahol az atomoknak megfelelő csomópontokat sötétebb ellipszisek jelölik. A dobozháló a 2. ábra fogalomhálójának elemeiből áll, a két diagram azonos számmal jelölt csomópontjai megegyező fogalmakat jelölnek.



3. ábra. Dobozháló Hasse diagramja

A dobozhálók jelentőségére a következő eredmény világít rá.

2.35. Tétel. *Ha L egy CJ-generált teljes háló, akkor $\mathcal{C}(L) = \mathcal{C}(\mathcal{D}(L))$.*

Vizsgáljuk meg az utóbbi két tétel következményét arra az esetre, ha $L = \mathcal{F}(G, M, I)$, ahol a (G, M, I) kontextus sor-redukált. Ez utóbbi tulajdonság azért lényeges, mert sor-redukált kontextus esetén minden $g \in G$ -re a (g'', g') fogalmak egyesítés-irreducibilisek, és a 2.18. Állítás szerint a fogalomháló minden eleme előáll ilyen alakú fogalmak szuprémumaként. Ebből az következik, hogy L egy CJ-generált háló, így a két utóbbi tétel L -re is érvényes, azaz egy sorredukált kontextushoz tartozó fogalomháló dobozhálója teljes, atomisztikus háló, melynek osztályozási rendszerei megegyeznek a fogalomháló osztályozási rendszereivel. Ezek az eredmények lehetővé teszik, hogy az $\mathcal{F}(G, M, I)$ fogalomháló osztályozási rendszereinek (vagy a G halmaz extenzió-partícióinak) meghatározásakor a feladatot a fogalomhálónál jóval kevesebb elemet tartalmazó, atomisztikus $\mathcal{D}(G, M, I)$ dobozháló osztályozási rendszereinek megadására vezessük vissza. (Az, hogy a dobozháló a fogalomhálónál könnyebben kezelhető struktúra, a 2. és 3. ábra összehasonlításából is érzékelhető.) A dobozháló elemeinek és osztályozási rendszereinek meghatározását a következő fejezetben tárgyaljuk.

3. Az osztályozási rendszerek meghatározása

Az előző fejezetben megmutattuk, hogy egy sor-redukált kontextushoz tartozó fogalomháló osztályozási rendszerei és a kontextus objektumhalmazának extenzió-partíciói megfeleltethetők egymásnak, hiszen az osztályozási rendszerek bizonyos fogalmakból, az extenzió-partíciók pedig e fogalmak extenzióiból állnak. A megfelelő fogalmak kiválogatása a fogalomhálóból nehezen kivitelezhető feladat, főleg nagyméretű kontextusok esetén, hiszen a fogalomháló elemeinek száma a kontextus méretének exponenciális függvénye is lehet. A dobozhálót éppen ennek a nehézségnek a kiküszöbölésére használhatjuk fel, hiszen a dobozháló osztályozási rendszerei azonosak a fogalomháló osztályozási rendszereivel, úgy, hogy a dobozháló atomisztikus szerkezetű és a fogalomhálónál jóval kevesebb elemet tartalmaz. A dobozháló osztályozási rendszereinek meghatározását három, egymástól jól elkülöníthető lépésben végezzük:

1. Meghatározzuk a dobozháló atomjait (a dobozatomokat).
2. Az atomok ismeretében előállítjuk a további dobozelemeket.
3. A dobozelemek felhasználásával meghatározzuk az osztályozási rendszereket.

A továbbiakban e három lépés kivitelezésére mutatunk be egy módszert. Az egyes lépéseket megvalósító algoritmusok mellett megadjuk azok egy lehetséges számítógépes megvalósítását, és megvizsgáljuk az algoritmusok hatékonyságát is.

3.1. A dobozháló atomjainak meghatározása

A dobozháló atomjai a fogalomháló legfinomabb osztályozási rendszerét alkotják, így kézenfekvőnek tűnik az a gondolat, hogy a fogalomháló legkisebb fogalmaiból kiindulva keressük ezt a minimális osztályozási rendszert. Ha (G, M, I) egy sor-redukált kontextus, akkor az $\mathcal{F}(G, M, I)$ fogalomháló egy osztályozási rendszere olyan fogalmakból áll, melyek extenziói az objektumhalmaz egy partícióját adják, tehát a legfinomabb osztályozási rendszer elemeinek extenziói is le kell fedjék a G halmazt. Ha $g \in G$, akkor (g'', g') a legkisebb olyan fogalom, melynek extenziója g -t tartalmazza, így az $\mathcal{A}^0 = \{(g'', g') \mid g \in G\}$ halmaz fogalmainak extenziói lefedik a G halmazt. Az $\mathcal{A}^0$ halmazból kiindulva, fogalmak fokozatos egyesítésével egy osztályozási rendszert állítunk elő a következő algoritmussal: tegyük fel, hogy $\mathcal{A}^0$ -nak van két olyan különböző, c és d eleme, amelyekre $c \wedge d \neq \mathbf{0}$ (ahol $\mathbf{0}$ a fogalomháló legkisebb elemét jelöli), ekkor a c és d fogalmakat töröljük $\mathcal{A}^0$ -ból és vegyük

fel helyettük a $c \vee d$ fogalmat. Jelölje $\mathcal{A}^1$ ezt az új, eggyel kevesebb elemből álló halmazt, és ismételjük meg az $\mathcal{A}^0$ -ra alkalmazott műveletet, ha lehetséges. Általában, legyen $\mathcal{A}^i = (\mathcal{A}^{i-1} \setminus \{c, d\}) \cup (c \vee d)$ ahol c és d ($c \neq d$) az $\mathcal{A}^{i-1}$ halmaz egy olyan fogalompárja, melyre $c \wedge d \neq \mathbf{0}$. Ha a G objektumhalmaz n elemből áll, akkor a folyamat legfeljebb $n - 1$ fogalom-összevonás után megáll, egy $\mathcal{A}^l$ ($0 \leq l \leq n - 1$) halmazt eredményezve. Jelölje $\mathcal{A}$ az eredményül kapott $\mathcal{A}^l$ halmazt, világos, hogy $\mathcal{A}$ bármely két elemének az infimuma a háló legkisebb elemével egyezik meg, vagy pedig $\mathcal{A}$ egyetlen fogalomból, a háló maximális eleméből áll.

Azt állítjuk, hogy az $\mathcal{A} = \{a_j = (A_j, B_j) \mid 1 \leq j \leq k\}$ rendszer a fogalomháló minimális osztályozási rendszere. A 2.28. Definíció alapján $\mathcal{A}$ akkor osztályozási rendszere az $\mathcal{F}(G, M, I)$ fogalomhálónak, ha

1. $a_i \wedge a_j = \mathbf{0} \quad \forall i \neq j$ esetén,
2. $x = \bigvee_{1 \leq j \leq k} (x \wedge a_j) \quad \forall x \in \mathcal{F}(G, M, I).$

Az első feltétel $\mathcal{A}$ képzési szabálya miatt automatikusan teljesül. A második feltétel igazolásához először vegyük észre, hogy $\bigcup_{1 \leq j \leq k} A_j = G$. Ez azért teljesül,

mert az $\mathcal{A}^0$ rendszert alkotó fogalmak extenziói, a g'' halmazok lefedik G -t, és ez a tulajdonság az $\mathcal{A}^i$ ($1 \leq i \leq l$) rendszerek fogalmaira is érvényes. Valóban, ha az a_i és a_j fogalmakat kicseréljük az $a_i \vee a_j$ fogalomra, ez utóbbi extenziója tartalmazza az előzőek extenzióját:

$$\text{Ext}(a_i) \cup \text{Ext}(a_j) \subseteq (\text{Ext}(a_i) \cup \text{Ext}(a_j))'' = \text{Ext}(a_i \vee a_j).$$

($\text{Ext}(a)$ -val jelöltük az a fogalom extenzióját.)

Legyen $x = (A, B)$ a fogalomháló egy tetszőleges eleme. A fogalmak metszetére és egyesítésére vonatkozó 2.16. Tételt alkalmazva vizsgáljuk meg az $\bigvee_{1 \leq j \leq k} (x \wedge a_j)$ fogalom extenzióját:

$$\begin{aligned} \text{Ext}\left(\bigvee_{1 \leq j \leq k} (x \wedge a_j)\right) &= \text{Ext}\left(\bigvee_{1 \leq j \leq k} (A \cap A_j, (B \cup B_j)'')\right) = \\ &= \left(\bigcup_{1 \leq j \leq k} (A \cap A_j)\right)'' = (A \cap \bigcup_{1 \leq j \leq k} A_j)'' = (A \cap G)'' = A'' = A = \text{Ext}(x). \end{aligned}$$

Mivel a fogalomháló két fogalma megegyezik, ha extenzióik azonosak, $x = \bigvee_{1 \leq j \leq k} (x \wedge a_j)$, azaz $\mathcal{A}$ egy osztályozási rendszer.

Be kell még látnunk, hogy az $\mathcal{A}$ osztályozási rendszer minimális. Legyen $\mathcal{P} = \{p_i = (P_i, P'_i) \mid 1 \leq i \leq s\}$ a fogalomháló egy tetszőleges osztályozási rendszere. Minden $q = (Q, Q') \in \mathcal{P}$ fogalom esetén legyen

$\mathcal{A}_q^0 = \{(g'', g') \mid g \in Q\}$ és képezzük az $\mathcal{A}_q^1, \mathcal{A}_q^2, \dots, \mathcal{A}_q^{l_q} = \mathcal{A}_q$ sorozatot az $\mathcal{A}$ rendszer előállításának mintájára. Vegyük észre, hogy tetszőleges $0 \leq i \leq l_q$ indexre, $x \in \mathcal{A}_q^i$ -ből $x \leq q$ következik, hiszen kezdetben minden $x \in \mathcal{A}_q^0$ esetén $x \leq q$, ha pedig $x \leq q, y \leq q$, akkor $x \vee y \leq q$, tehát bármely, egyesítéssel kapott fogalom is kisebb vagy egyenlő q -nál. Ebből az is következik, hogy ha p és q a $\mathcal{P}$ rendszer két különböző fogalma, $\mathcal{A}_p^i$ az $\mathcal{A}_p$, $\mathcal{A}_q^j$ pedig az $\mathcal{A}_q$ rendszert előállító sorozat tetszőleges tagja, akkor $x \in \mathcal{A}_p^i, y \in \mathcal{A}_q^j$ esetén $x \wedge y = \mathbf{0}$, (hiszen $x \leq p, y \leq q$ és $p \wedge q = \mathbf{0}$). Azt kaptuk, hogy az $\mathcal{A}$ rendszer elemeit egymástól független halmazsorozatokkal is előállíthatjuk, ahol a sorozatok kezdőelemét a $\mathcal{P}$ osztályozási rendszer által indukált partíció alapján határozzuk meg, azaz $\mathcal{A} = \bigcup_{q \in \mathcal{P}} \mathcal{A}_q$. Legyen $q = (Q, Q')$ a $\mathcal{P}$ osztályozási rendszer egy tetszőleges fogalma, ekkor a Q blokk előáll $Q = \bigcup_{q \in \mathcal{A}_q} \text{Ext}(q)$ alakban (ezt ugyanúgy láthatjuk be, mint korábban a $G = \bigcup_{a \in \mathcal{A}} \text{Ext}(a)$ egyenlőséget). Mivel az $\mathcal{A}_q$ -beli fogalmak egyben $\mathcal{A}$ -beliek is, kaptuk, hogy az $\mathcal{A}$ által indukált extenzió-partíció finomítja a $\mathcal{P}$ által indukált extenzió-partíciót, azaz az $\mathcal{A}$ osztályozási rendszer finomabb $\mathcal{P}$ -nél.

Bebizonyítottuk, hogy az $\mathcal{A}$ rendszer a fogalomháló minimális osztályozási rendszere. Mivel egy fogalom intenzióját extenziója ismeretében könnyen megadhatjuk, az $\mathcal{A}$ rendszer elemeinek előállításához elegendő az $\mathcal{A}$ -beli fogalmak extenzióit meghatározni, melyek együttesen a G halmaz legfinomabb extenzió-partícióját alkotják. Jelöljük Π_0 -nal ezt a minimális partíciót. A következő algoritmus a Π_0 rendszer elemeit határozza meg az előzőekben ismertett módszer alapján, kihasználva, hogy ha a c és d fogalmakra $c \wedge d \neq \mathbf{0}$, akkor $\text{Ext}(c) \cap \text{Ext}(d) \neq \emptyset$, illetve $\text{Ext}(c \vee d) = (\text{Ext}(c) \cup \text{Ext}(d))''$.

3.1. Algoritmus (Atom-extenziók meghatározása).

Kezdetben legyen $\Pi_0 = \{g'' \mid g \in G\}$

amíg van olyan $X \neq Y \in \Pi_0$, amelyre $X \cap Y \neq \emptyset$

addig Π_0 -t bővítsd az $(X \cup Y)''$ halmazzal

és Π_0 -ből töröld az X és Y halmazt.

Az 1. táblázatban adott példakontextusra az algoritmus az alábbi öt fogalom-extenziót adja meg a legfinomabb extenzió-partíció elemeiként:

$$\Pi_0 = \{\{g_1, g_6\}, \{g_2, g_7\}, \{g_3\}, \{g_4, g_5\}, \{g_8\}\}.$$

3.2. A dobozelemek meghatározása

A dobozháló atomjainak ismeretében a további dobozelemek meghatározására ad egy módszert Bukszár Zoltán [4]-ben. Eljárása a dobozháló atomisztikus tulajdonságát használja ki: mivel minden dobozelem előáll atomok szuprémumaként, a fogalomháló elemei közül kiválogatva azokat a fogalmakat, amelyek dobozatomból egyesítései, megkapjuk a dobozháló további elemeit. Nagy elemszámú fogalomháló esetén az eljárás időigénye is megnő, ezért mindegyik hasznos olyan módszerek kidolgozása, melyek a dobozelemeket anélkül határozzák meg, hogy szükség lenne a fogalomháló összes elemének kiszámítására. A továbbiakban egy ilyen módszert mutatunk be.

Legyen (G, M, I) egy véges, sor-redukált kontextus és legyen $\mathcal{A}$ - az előző alfejezet jelölésével összhangban - a $\mathcal{D}(G, M, I)$ dobozháló atomjainak halmaza. Bevezetünk egy fogalmat, melyet a dobozelemek meghatározásának algoritmusában használunk majd.

3.2. Definíció. Ha $H \subseteq G$ egy tetszőleges objektumhalmaz, akkor H **atomszerkezetén** azoknak az atomoknak az összességét értjük, melyek extenziói H részhalmazai. Ezt a halmazt $a(H)$ -val jelölve:

$$a(H) = \{x \in \mathcal{A} \mid Ext(x) \subseteq H\}.$$

A következő két eredmény az atomszerkezet fogalmával kapcsolatos.

3.3. Lemma. Minden $H, F \subseteq G$ esetén $a(H \cap F) = a(H) \cap a(F)$.

Bizonyítás. $x \in a(H \cap F) \Leftrightarrow Ext(x) \subseteq H \cap F \Leftrightarrow Ext(x) \subseteq H$ és $Ext(x) \subseteq F \Leftrightarrow x \in a(H)$ és $x \in a(F) \Leftrightarrow x \in a(H) \cap a(F)$. $\square$

3.4. Tétel. Az (E, E') fogalom pontosan akkor eleme a $\mathcal{D}(G, M, I)$ dobozhálónak, ha

$$E = \bigcup_{x \in a(E)} Ext(x).$$

Bizonyítás. Legyen (E, E') egy dobozelem. Ha $x \in a(E)$, akkor $Ext(x) \subseteq E$, tehát $\bigcup_{x \in a(E)} Ext(x) \subseteq E$. Legyen $g \in E$ és jelölje x_g azt az atomot, melynek extenziója g -t tartalmazza. Mivel $Ext(x_g)$ a G halmaz legfinomabb extenzió-partíciójának egy blokkja, E pedig egy másik osztályozásnak a g elemet tartalmazó blokkja, $Ext(x_g) \subseteq E$. Ebből $x_g \in a(E)$ következik, tehát $g \in \bigcup_{x \in a(E)} Ext(x)$. Ezzel beláttuk a fordított irányú, $E \subseteq \bigcup_{x \in a(E)} Ext(x)$ tartalmazást is.

Megfordítva, tegyük fel, hogy az (E, E') fogalomra $E = \bigcup_{x \in a(E)} Ext(x)$ teljesül.

Ha $\mathcal{A}$ jelöli a dobozháló atomjainak halmazát, az $\{Ext(x) \mid x \in \mathcal{A}\}$ rendszer

G legfinomabb extenzió-partíciója. Mivel $E = \bigcup_{x \in a(E)} Ext(x)$ egy fogalom extenziója, az $\{E\} \cup \{Ext(x) \mid x \in \mathcal{A} \setminus a(E)\}$ rendszer is egy extenzió-partíciója a G halmaznak. Ebből következik, hogy az $\{(E, E')\} \cup \{x \mid x \in \mathcal{A} \setminus a(E)\}$ rendszer a fogalomháló egy osztályozási rendszere, tehát (E, E') a $\mathcal{D}(G, M, I)$ dobozháló egy eleme.

□

Bukszár Z. bevezetőben említett eljárása a fogalmak generálására a 2.20. Algoritmust használja. Ez az algoritmus a fogalmak extenzióit egy listára gyűjti, úgy, hogy előállítja az attribútum-extenziók összes lehetséges metszeteit. A dobozelem-extenziók meghatározásához azonban nincs szükség az összes ilyen metszetre, megmutatjuk, hogy az algoritmus módosításával hogyan kerülhető el a felesleges fogalom-extenziók előállítása. A 3.4. Tételből következik, hogy különböző dobozelemeknek az atomszerkezete is különböző, és az atomoktól, valamint a háló $\mathbf{0}$ elemétől különböző dobozelemek atomszerkezete legalább kételemű halmaz. Ezt az észrevételt kihasználva a 2.20. Algoritmust a következőképpen módosítjuk: két, kezdetben üres listával dolgozunk. Az első listára extenziókat gyűjtünk, a másodikra rendre kiszámítjuk az első lista megfelelő elemének atomszerkezetét. A listák első eleme G , illetve G atomszerkezete, az $\mathcal{A}$ halmaz. Ezután sorra vesszük az m attribútumokat, és képezzük minden, az első listán szereplő A halmaz esetén az $F = A \cap m'$ halmazt. Az eredeti algoritmustól eltérően az F extenziót csak akkor vesszük fel az első listára, ha eleget tesz a következő két feltételnek:

- (i) $a(F)$ legalább két elemet tartalmaz,
- (ii) $a(F)$ nem szerepel a második listán, azaz nincs olyan H halmaz az első listán, amelyre $a(F) = a(H)$ teljesülne.

Az (i) feltétel kizárja az első listáról azokat a fogalom-extenziókat, melyek 0 vagy 1 atom-extenziót tartalmaznak. Az ilyen tulajdonságú extenziók törölhetők, hiszen sem ezek, sem további attribútum-extenziókkal való metszeteik nem eredményezhetnek atomoknál nagyobb fogalmakat, az atomokat pedig már ismerjük. Az (ii) feltétel azt biztosítja, hogy az első listára csak különböző atomszerkezetű extenziók kerülhessenek. Ezekből, az atomszerkezetük szerint is különböző extenziókból fogjuk a keresett dobozelem-extenziókat meghatározni. Felmerülhet azonban a kérdés, hogy a második feltétel alapján kizárt extenziók a később sorra kerülő attribútum-extenziókkal metszve nem eredményezhetnek-e olyan extenziókat, amelyekből származó dobozelemek az első lista alapján nem generálhatóak és így elvesznek. A következő gondolatmenet szerint ez nem fordulhat elő. Tegyük fel, hogy a H halmaz már rajta van az első listán, és az algoritmus következő lépésében egy olyan F halmaz felvételéről kell dönten, amelyre $a(F) = a(H)$ teljesül. Az (ii) feltétel miatt

az F halmazt nem vesszük fel az első listára és vele együtt az $F \cap J$ halmazok is hiányozni fognak onnan, minden olyan J extenzióra, melyet az algoritmus az F -et követően vizsgál. A kérdés az, hogy az $F \cap J$ halmazok atomszerkezete lehet-e a listán már szereplő halmazokétól különböző, mert ha igen, akkor az $F \cap J$ extenzióból olyan dobozelem generálható, amely a listán szereplő extenziók egyikéből sem hozható létre. A válasz azonban nemleges, hiszen a 3.3. Lemma felhasználásával $a(F \cap J) = a(F) \cap a(J) = a(H) \cap a(J) = a(H \cap J)$ adódik, ami azt jelenti, hogy az $F \cap J$ extenzióval azonos atomszerkezetű extenzió már van a listán a $H \cap J$ halmaz formájában. Tehát az $F \cap J$ halmazok egyetlen F -et követő J halmaz esetén sem kerülnének az első listára, így arra sincs szükség, hogy F -et felvegyük arra.

A következő algoritmus az atomoktól és a háló $\mathbf{0}$ elemétől különböző dobozelemek extenzióit határozza meg. A dobozelemek extenzióit a D halmaz, azok atomszerkezeteit az S halmaz tárolja.

3.5. Algoritmus (Dobozelem-extenziók meghatározása).

Kezdetben legyen $D = \{G\}$ és $S = \{a(G)\}$

minden $m \in M$ attribútumra

minden $E \in D$ -re

legyen $F = E \cap m'$

ha $a(F)$ legalább kételemű és $a(F) \notin S$

akkor D -t bővítsd F -el és S -et bővítsd $a(F)$ -el

minden $E \in D$ -re

legyen $C = \bigcup_{x \in a(E)} \text{Ext}(x)$

ha $E \neq C$ akkor D -ből töröld az E -t és S -ből töröld $a(E)$ -t

ha $C = C''$ akkor D -t bővítsd C -vel és S -et bővítsd $a(C)$ -vel.

Az utolsó négy sor az első listára került extenziók ellenőrzéséről szól, ugyanis nem biztos, hogy mindegyik dobozelem-extenzió is egyben. Előfordulhat, hogy egy (E, E') dobozelem extenziója helyett, egy nála bővebb, de megegyező atomszerkezetű fogalom-extenzió hamarabb kerül az első listára, s így az (ii) feltétel miatt E hiányozni fog onnan. Ezért az első listán szereplő E extenziók esetén a második lista $a(E)$ atomszerkezete felhasználásával előállítjuk a $C = \bigcup_{x \in a(E)} \text{Ext}(x)$ halmazt. Ha $E = C$, akkor E dobozelem-extenzió.

Ha $E \neq C$, akkor E helyett a C halmaz lehet dobozelem-ekstenzió, amennyiben teljesíti a $C = C''$ egyenlőséget, azaz C egy fogalom-ekstenzió.

A 3. táblázat a példakontextusból származtatható, nem-atomai dobozelemek meghatározásának folyamatát szemlélteti. A középső oszlopban szereplő halmazok elemei az előző alfejezetben meghatározott $a_i = (A_i, A'_i)$ ($i = 1, 2, \dots, 5$) atomok, a következő ekstenziókkal:

$$A_1 = \{g_1, g_6\}, \quad A_2 = \{g_2, g_7\}, \quad A_3 = \{g_3\}, \quad A_4 = \{g_4, g_5\} \quad A_5 = \{g_8\}.$$

3. táblázat. Dobozelem-ekstenziók meghatározása

Fogalom-ekstenzió (első lista)	Atomszerkezet (második lista)	Dobozelem-ekstenzió
$\{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\}$	$\{a_1, a_2, a_3, a_4, a_5\}$	$\{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\}$
$\{g_1, g_3, g_6\}$	$\{a_1, a_3\}$	$\{g_1, g_3, g_6\}$
$\{g_1, g_4, g_5, g_6, g_8\}$	$\{a_1, a_4, a_5\}$	$\{g_1, g_4, g_5, g_6, g_8\}$
$\{g_4, g_5, g_6, g_7\}$	$\{a_4, a_5\}$	
$\{g_2, g_7, g_8\}$	$\{a_2, a_5\}$	$\{g_2, g_7, g_8\}$
$\{g_2, g_3, g_4, g_5, g_7\}$	$\{a_2, a_3, a_4\}$	$\{g_2, g_2, g_4, g_5, g_7\}$
$\{g_2, g_3, g_5, g_7, g_8\}$	$\{a_2, a_3, a_5\}$	
$\{g_2, g_3, g_5, g_7\}$	$\{a_2, a_3\}$	

Az attribútum-ekstenziók metszetével adódó halmazok közül csak azok kerülhettek az első listára, amelyek atomszerkezete legalább két atomból áll és a korábbi atomszerkezetektől különbözik. Ezeknek a feltételeknek nyolc ekstenzió felelt meg, melyek közül öt bizonyult végül dobozelem-ekstenziónak is. A fogalmak illetve a dobozelemek előállítását szemléltető 2. és 3. táblázat összehasonlításából kiderül, hogy a dobozelem-ekstenziók meghatározásához a 21 lehetséges fogalom-ekstenzióból csak 8-at használtunk fel.

3.3. Az osztályozási rendszerek előállítása

Az osztályozási rendszerek meghatározásához a következő eredményt használjuk.

3.6. Állítás. Legyen $D = \{(D_j, D'_j) \mid j \in J\}$ a $\mathcal{D}(G, M, I)$ dobozháló elemeinek egy rendszere. D pontosan akkor osztályozási rendszere a dobozhálónak, ha a megfelelő atomszerkezetekből álló $\{a(D_j) \mid j \in J\}$ rendszer az atomok halmazának partíciója.

Bizonyítás. Tegyük fel, hogy $D = \{(D_j, D'_j) \mid j \in J\}$ a dobozháló egy osztályozási rendszere, és képezzük az $\{a(D_j) \mid j \in J\}$ rendszert. A 3.3. Lemma

miatt $i \neq j$ esetén $a(D_i) \cap a(D_j) = a(D_i \cap D_j) = a(\emptyset) = \emptyset$. Másrészt, tegyük fel, hogy van olyan x atom, melyre $x \in \mathcal{A} \setminus \cup\{a(D_j) \mid j \in J\}$. Ha x egy atom, extenziója nem üres, legyen $g \in \text{Ext}(x)$. Mivel $G = \cup\{D_j \mid j \in J\}$, van olyan i index, melyre $g \in D_i$. Az atomok extenzióiból álló osztályozás finomabb a $\{D_j \mid j \in J\}$ partíciónál, ezért $\text{Ext}(x) \subseteq D_i$, de akkor $x \in a(D_i)$, tehát $x \in \cup\{a(D_j) \mid j \in J\}$, ami ellentmondás. Ezzel belátuk, hogy $\cup\{a(D_j) \mid j \in J\} = \mathcal{A}$, tehát az $\{a(D_j) \mid j \in J\}$ rendszer valóban partíciója az atomok halmazának.

Megfordítva, tegyük fel, hogy a $D = \{(D_j, D'_j) \mid j \in J\}$ dobozelemekből álló rendszer esetén $\{a(D_j) \mid j \in J\}$ az $\mathcal{A}$ halmaz egy partíciója. A 3.4. Tétel alkalmazásával $i \neq j$ esetén

$$D_i \cap D_j = \bigcup_{x \in a(D_i)} \text{Ext}(x) \cap \bigcup_{x \in a(D_j)} \text{Ext}(x) = \bigcup_{x \in a(D_i) \cap a(D_j)} \text{Ext}(x) = \emptyset.$$

Másrészt,

$$\bigcup_{j \in J} D_j = \bigcup_{j \in J} \left(\bigcup_{x \in a(D_j)} \text{Ext}(x) \right) = \bigcup_{x \in \bigcup_{j \in J} a(D_j)} \text{Ext}(x) = \bigcup_{x \in \mathcal{A}} \text{Ext}(x) = G.$$

□

Az előző állítás szerint az osztályozási rendszerek előállításához a dobozelemeket úgy kell csoportosítani, hogy az egy rendszerbe tartozó elemek atomszerkezetei $\mathcal{A}$ egy partícióját alkossák. Minden atom atomszerkezete egyetlen elemből, magából az atomból áll, a többi, $\mathbf{0}$ -tól különböző dobozelem atomszerkezetét pedig a 3.5. Algoritmus eredményeként kapott S lista tárolja. Az $\mathcal{A}$ halmaz egy atomszerkezetekből álló partícióját úgy kapjuk, hogy S részhalmazai közül páronként diszjunkt halmazokat választunk, majd a kiválasztott atomszerkezetek egyikében sem szereplő atomokkal kiegészítjük a rendszert. A következő algoritmus először egy $\mathcal{H}$ listára felveszi minden egyes dobozelem atomszerkezetét, majd ezt a listát bővíti S azon részhalmazai-val, melyek páronként diszjunkt atomszerkezeteket tartalmaznak. Mivel a $\mathcal{H}$ lista elemei atomokkal kiegészíthetőek $\mathcal{A}$ partícióivá, minden $H \in \mathcal{H}$ halmazhoz hozzárendeljük atomok és nem-atomi dobozelemek egy halmazát, melyek megfelelnek a dobozháló egy osztályozási rendszerének. Az eredményt az $\mathcal{L}$ listán tároljuk, melynek első eleme az $\mathcal{A}$ halmaz, azaz a legfinomabb osztályozási rendszer.

3.7. Algoritmus (Osztályozási rendszerek előállítás).

Kezdetben a $\mathcal{H}$ lista a nem-atomi dobozelemek atomszerkezeteiből áll,

az $\mathcal{L}$ lista egyetlen eleme az $\mathcal{A}$ halmaz

amíg van olyan H részhalmaza S -nek, hogy $\forall X, Y \in H$ -ra $X \cap Y = \emptyset$

addig $\mathcal{H}$ -t bővítsd a H halmazzal

minden $H \in \mathcal{H}$ esetén

legyen $A = \mathcal{A} \setminus \bigcup \{x \mid x \in X, X \in H\}$ és legyen

$D = \bigcup \{(B, B') \mid (B, B') \text{ dobozelem}, \exists X \in H, a(B) = X\}$

és az L listára vedd fel az $A \cup D$ halmazt.

A korábban vizsgált példában az S halmazt a 3. táblázat második listájából megmaradó öt atomszerkezet alkotja. Ezekből még képezhetjük az $\{\{a_1, a_3\}, \{a_2, a_5\}\}$ diszjunkt atomszerkezetekből álló halmazt, melyhez szintén tartozik egy osztályozási rendszer. A következő táblázatban felsoroltuk az atomok halmazának azon partícióit, melyeknek osztályozási rendszerek felelnek meg, illetve megadtuk a G halmazon indukált extenzió-partíciókat is.

4. táblázat. Az extenzió-partíciók meghatározása

Az atomok halmazának partíciói	Az objektumhalmaz extenzió-partíciói
$\{a_1\}\{a_2\}\{a_3\}\{a_4\}\{a_5\}$	$\{g_1, g_6\}\{g_2, g_7\}\{g_3\}\{g_4, g_5\}\{g_8\}$
$\{a_1, a_2, a_3, a_4, a_5\}$	$\{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\}$
$\{a_1, a_3\}\{a_2\}\{a_4\}\{a_5\}$	$\{g_1, g_3, g_6\}\{g_2, g_7\}\{g_4, g_5\}\{g_8\}$
$\{a_1, a_4, a_5\}\{a_2\}\{a_3\}$	$\{g_1, g_4, g_5, g_6, g_8\}\{g_2, g_7\}\{g_3\}$
$\{a_2, a_5\}\{a_1\}\{a_3\}\{a_4\}$	$\{g_2, g_7, g_8\}\{g_1, g_6\}\{g_3\}\{g_4, g_5\}$
$\{a_2, a_3, a_4\}\{a_1\}\{a_5\}$	$\{g_2, g_3, g_4, g_5, g_7\}\{g_1, g_6\}\{g_8\}$
$\{a_1, a_3\}\{a_2, a_5\}\{a_4\}$	$\{g_1, g_3, g_6\}\{g_2, g_7, g_8\}\{g_4, g_5\}$

A kapott extenzió-partíciók alapján az osztályozási rendszerek elemeit úgy kapjuk meg, hogy az egyes partíciók blokkjaihoz, mint fogalom-extenziókhöz meghatározzuk a fogalmak intenzióját is.

3.4. Az adatok reprezentációja

Az osztályozási rendszerek előállításának három fázisát az előző alfejezetekben közölt 3.1., 3.5. és 3.7. Algoritmusok írják le. Az említett algoritmusok felhasználásával készítettem egy programot, mely egy adott $K = (G, M, I)$

kontextusból kiindulva előállítja az $\mathcal{F}(G, M, I)$ háló osztályozási rendszereit és a G halmaz extenzió-partícióit. A kontextus tárolására egy mátrixot használtam, az objektumok és attribútumok különböző halmazait pedig vektorokkal azonosítottam. Ez a reprezentáció a feladat megoldása során a következő előnyökkel járt:

- Az extenzió-partíciók meghatározásához szükséges halmazműveletek mátrix-aritmetikai műveletek sorozataként értelmezhetők.
- A legtöbb művelet ugyanazzal a módszerrel végezhető el egyetlen objektumra (attribútumra) mint objektumok (attribútumok) egy halmazára (például, ha $g \in G$ vagy $A \subseteq G$, akkor a $\{g\}'$ illetve A' halmazokat ugyanaz a függvény állítja elő, az első esetben egy vektor, a másodikban egy mátrix bemenettel).
- Az adatok vektoros ábrázolásának köszönhető néhány olyan ötlet, amelyeknek szerepe volt a dolgozat új, tudományos eredményeinek megszületésében (például az atomszerkezet fogalma, vagy a következő fejezetben ismertetendő kitüntetett tulajdonság koncepció.)

A választott adatszerkezet kezelését és a szükséges műveletek elvégzését a Matlab programkörnyezet messzemenően támogatja, ezért az extenzió-partíciók meghatározásához szükséges függvények definícióját Matlab nyelven írtam meg.

Az algoritmusok pszeudokódjainak leírásakor az Algoritmusok c. könyvben rögzített megállapodásokat követtem (Cormen és társai, [12]). Ezek szerint a blokkszerkezetet bekezdések jelzik, a $\leftarrow$ értékadást jelent, a tömbelemek indexeit szögletes zárójelbe tesszük. A $\triangleright$ jel után megjegyzések állnak.

A fejezet további részében a következő jelöléseket alkalmazzuk:

$\mathfrak{M}_{n \times k}$: az $n \times k$ típusú, 0-1 elemeket tartalmazó mátrixok halmaza,

$\mathfrak{M}_{1 \times k}, \mathfrak{M}_{n \times 1}$: 0-1 elemekből álló k -dimenziós sorvektorok illetve n -dimenziós oszlopvektorok halmaza,

$\mathbf{A}, \mathbf{B}, \mathbf{C}$ mátrixokat, $\mathbf{x}, \mathbf{y}, \mathbf{v}$ vektorokat jelöl,

$\mathbf{E}_{n \times n}$ az egységmátrix, $\mathbf{0}_{n \times k}$ a zérusmátrix az indexben jelölt típussal,

$\mathbf{1}_{n \times 1}, \mathbf{1}_{1 \times k}$ csupa egyesekből álló vektorok,

$\mathbf{a}_i$ az $\mathbf{A}$ mátrix i -edik oszlopa, $\mathbf{a}_i^*$ az $\mathbf{A}$ mátrix i -edik sora,

M_{ij} az $\mathbf{M}$ mátrix i, j indexű eleme (ugyanaz a pszeudokódokban $M[i][j]$),

v_j a $\mathbf{v}$ vektor j -edik koordinátája (pszeudokódokban $v[j]$),

$\mathbf{v}^T, \mathbf{M}^T$ a $\mathbf{v}$ vektor, illetve az $\mathbf{M}$ mátrix transzponáltja,

$[\mathbf{A}, \mathbf{a}]$ az a mátrix, amely az $\mathbf{A}$ mátrixból az $\mathbf{a}$ oszlopvektor hozzáírásával keletkezik,

$[\mathbf{S}; \mathbf{s}^*]$ az a mátrix, amely az $\mathbf{S}$ mátrixból az $\mathbf{s}^*$ sorvektor hozzáírásával keletkezik,

$[\mathbf{A} \setminus \mathbf{a}]$ az a mátrix, amely az $\mathbf{A}$ mátrixból az $\mathbf{a}$ oszlopvektor törlésével keletkezik,

$[\mathbf{S} \setminus \mathbf{s}^*]$ az a mátrix, amely az $\mathbf{S}$ mátrixból az $\mathbf{s}^*$ sorvektor törlésével keletkezik,

$\mathbf{a} \in \mathbf{A}$, illetve $\mathbf{s}^* \in \mathbf{S}$ azt jelzi, hogy az $\mathbf{a}$ vektor az $\mathbf{A}$ mátrix egyik oszlopvektora, illetve $\mathbf{s}^*$ az $\mathbf{S}$ mátrix egyik sorvektora.

Az $\mathcal{F}(G, M, I)$ fogalomháló osztályozási rendszereinek, illetve a G halmaz extenzió-partícióinak meghatározásához egyetlen bemenő adatra van szükségünk, a (G, M, I) kontextusra. Tegyük fel, hogy az objektumok halmaza n elemből áll, az attribútumok halmaza pedig k elemű, azaz $G = \{g_j \mid j = 1, 2, \dots, n\}$ és $M = \{m_j \mid j = 1, 2, \dots, k\}$. A (G, M, I) kontextus táblázata egy $\mathbf{K} \in \mathfrak{M}_{n \times k}$ mátrix segítségével jeleníthető meg a legkézenfekvőbb módon:

$$K_{ij} = \begin{cases} 1, & \text{ha } (g_i, m_j) \in I \\ 0 & \text{egyébként.} \end{cases}$$

Az $\mathcal{F}(G, M, I)$ háló fogalmai (A, B) alakú párok, ahol A objektumok, B pedig attribútumok egy halmaza. Az objektumhalmazoknak oszlopvektorokat, az attribútumhalmazoknak sorvektorokat feleltetünk meg a következő definíció értelmében.

3.8. Definíció. Legyen $A \subseteq G$ egy tetszőleges halmaz. Azt mondjuk, hogy az A halmazt az $\mathbf{a} \in \mathfrak{M}_{n \times 1}$ vektor **reprezentálja**, ha $j = 1, 2, \dots, n$ esetén

$$a_j = \begin{cases} 1, & \text{ha } g_j \in A \\ 0 & \text{egyébként.} \end{cases}$$

Hasonlóképpen, a $B \subseteq M$ attribútumhalmazt a $\mathbf{b} \in \mathfrak{M}_{1 \times k}$ vektor reprezentálja, ha $j = 1, 2, \dots, k$ esetén

$$b_j = \begin{cases} 1, & \text{ha } m_j \in B \\ 0 & \text{egyébként.} \end{cases}$$

3.9. *Megjegyzés.* Ha χ_A jelöli az $A \subseteq G$, χ_B pedig a $B \subseteq M$ halmaz karakterisztikus függvényét, akkor az A -t reprezentáló $\mathbf{a}$ vektort és a B -t reprezentáló $\mathbf{b}$ vektort az alábbi képletekkel is megadhatjuk:

$$\mathbf{a} = (\chi_A(g_1), \chi_A(g_2), \dots, \chi_A(g_n))^T$$

$$\mathbf{b} = (\chi_B(m_1), \chi_B(m_2), \dots, \chi_B(m_k)).$$

3.5. Az algoritmusok egy lehetséges megvalósítása

A kitűzött feladat megoldása során az egyik legalapvetőbb művelet egy objektumhalmaz közös tulajdonságainak előállítására, azaz $A \subseteq G$ esetén az A' halmaz meghatározása. Hasonlóképpen gyakran van szükség a B' halmaz előállítására, ha $B \subseteq M$. Legyen $A \subseteq G$, jelölje $\mathbf{a}$ az A halmazt reprezentáló vektort, és képezzük az $\mathbf{c} = \mathbf{a}^T \mathbf{K}$ szorzatot. Vegyük észre, hogy a $\mathbf{c}$ vektor j -edik koordinátája azt mutatja meg, hogy az $m_j \in M$ tulajdonsággal hány A -beli objektum rendelkezik. Ha m_j egy olyan tulajdonság, amellyel A minden objektuma rendelkezik, akkor a $\mathbf{c}$ vektor j -edik koordinátája megegyezik az A halmaz elemszámával. Ennek igazolására tegyük fel, hogy $A = \{g_l \mid l \in L\}$ és jelölje $|A|$ az A halmaz elemeinek számát. Ekkor

$$m_j \in A' \Leftrightarrow (g_l, m_j) \in I \quad \forall l \in L \Leftrightarrow K_{lj} = 1 \quad \forall l \in L \Leftrightarrow$$

$$\Leftrightarrow c_j = \sum_{p=1}^k a_p K_{pj} = \sum_{l \in L} a_l K_{lj} = |L| = |A|.$$

Ezzel igazoltuk, hogy az A' attribútumhalmazt reprezentáló $\mathbf{b}$ vektort a következő képlettel írhatjuk le: $j = 1, 2, \dots, k$ -ra

$$b_j = \begin{cases} 1, & \text{ha } c_j = |A| \\ 0 & \text{egyébként.} \end{cases}$$

Hasonló módon adhatjuk meg egy $B \subseteq M$ halmaz esetén a B -beli attribútumok mindegyikével rendelkező objektumok halmazát. Jelölje $\mathbf{b}$ a B halmazt reprezentáló sorvektort, ekkor a $\mathbf{d} = \mathbf{Kb}^T$ szorzatvektor j -edik koordinátája azt adja meg, hogy a g_j objektum hány B -beli tulajdonsággal rendelkezik. Ha ez a szám megegyezik B elemszámával, akkor g_j minden egyes B -beli

tulajdonsággal rendelkezik, tehát $g_j \in B'$. Ha a B' halmazt az $\mathbf{a}$ vektor reprezentálja, akkor $j = 1, 2, \dots, n$ -re

$$a_j = \begin{cases} 1, & \text{ha } \mathbf{d}_j = |B| \\ 0 & \text{egyébként.} \end{cases}$$

A fenti összefüggéseket és a mátrix-szorzás sajátosságait kihasználva van arra lehetőség, hogy az $A_j \subseteq G, j = 1, 2, \dots, m$ adott halmazok esetén a közös tulajdonságok A'_j halmazait *egyszerre* meghatározzuk. Ezt írja le a KTUL nevű függvény pszeudokódja (a KTUL elnevezés a közös tulajdonságokra utal). A függvény egyik bemeneti változója egy $\mathbf{A} \in \mathfrak{M}_{n \times m}$ mátrix, mely az A_j halmazokat reprezentáló oszlopvektorokból áll, a másik pedig a kontextusnak megfelelő $\mathbf{K} \in \mathfrak{M}_{n \times k}$ mátrix; visszatérési értéke az a $\mathbf{B} \in \mathfrak{M}_{m \times k}$ mátrix, melynek j -edik sora az A'_j halmazt reprezentálja.

KTUL ($\mathbf{A}, \mathbf{K}$)

```

1   $\mathbf{B} \leftarrow \mathbf{0}_{m \times k}$ 
2   $\mathbf{C} \leftarrow \mathbf{A}^T \mathbf{K}$ 
3  for  $i \leftarrow 1$  to  $m$ 
4    for  $j \leftarrow 1$  to  $k$ 
5      if  $C[i][j] = \sum_{l=1}^n A[l][i]$ 
6        then  $B[i][j] \leftarrow 1$ 
7  return  $\mathbf{B}$ .
```

A KTUL függvényhez teljesen hasonló adható meg a KOBJ függvény, amely a $B_j \subseteq M, j = 1, 2, \dots, m$ halmazokhoz meghatározza a megfelelő B'_j objektumhalmazokat. A bemeneti $\mathbf{B}$ mátrix j -edik sora a B_j -t reprezentáló sorvektor, az eredmény pedig az az $\mathbf{A}$ mátrix, melynek oszlopvektorai a B'_j halmazokat reprezentálják.

Vegyük észre, hogy ha a közös tulajdonságokat illetve közös objektumokat meghatározó függvényeket egymás után alkalmazzuk az A objektumhalmazra, eredményül az A'' halmazt kapjuk. Ez a művelet a KTUL és KOBJ függvényekhez hasonlóan egyszerre több objektumhalmazra is végrehajtható. Ha az $\mathbf{A}$ mátrix az $A_1, A_2, \dots, A_m$ objektumhalmazokat reprezentáló oszlopvektorokból áll, akkor a KOBJ (KTUL ($\mathbf{A}, \mathbf{K}$), $\mathbf{K}$) függvényhívás eredménye egy olyan mátrix, melynek oszlopvektorai az $A''_1, A''_2, \dots, A''_m$ halmazokat reprezentálják.

Az atom-extenziók előállítására szolgáló 3.1. Algoritmusban az A'' halmazok meghatározása mellett két halmaz egyesítésének, illetve metszetének kiszámítása is szerepel. Az, hogy két, objektumokból vagy attribútumokból álló halmaznak van-e közös eleme, rendszerünkben egyszerűen eldönthető: pontosan akkor diszjunktak, ha a halmazokat reprezentáló vektorok skaláris szorzata 0. A halmazok metszetét, illetve unióját reprezentáló vektorokat az eredeti vektorok összegéből állítjuk elő:

$$\mathbf{v} = \text{METSZET}(\mathbf{x}, \mathbf{y}), \text{ ahol}$$

$$v_j = \begin{cases} 1, & \text{ha } x_j + y_j = 2 \\ 0 & \text{egyébként.} \end{cases}$$

$$\mathbf{w} = \text{UNIO}(\mathbf{x}, \mathbf{y}), \text{ ahol}$$

$$w_j = \begin{cases} 0, & \text{ha } x_j + y_j = 0 \\ 1 & \text{egyébként.} \end{cases}$$

A felsorolt függvényeket felhasználva a 3.1. Algoritmus felhasználásával megadható az **ATOM** nevű függvény, amely a $\mathbf{K}$ kontextusból kiindulva egy olyan $\mathbf{A}$ mátrixot ad eredményül, melynek oszlopvektorai a dobozatomok extenzióit reprezentálják. Az algoritmus első lépésében a $\{g''\}$ halmazokat állítjuk elő. Ehhez az egységmátrixot használjuk fel, hiszen a $\{\{g_1\}, \{g_2\}, \dots, \{g_n\}\}$ rendszert $\mathbf{E}_{n \times n}$ reprezentálja.

ATOM($\mathbf{K}$)

- 1 $\mathbf{A} \leftarrow \text{KOBJ}(\text{KTUL}(\mathbf{E}_{n \times n}, \mathbf{K}), \mathbf{K})$
- 2 **while** $\exists \mathbf{x}, \mathbf{y} \in \mathbf{A}, \mathbf{x} \neq \mathbf{y}, \text{METSZET}(\mathbf{x}, \mathbf{y}) \neq \mathbf{0}_{n \times 1}$
- 3 $\mathbf{A} \leftarrow [\mathbf{A}, \text{KOBJ}(\text{KTUL}(\text{UNIO}(\mathbf{x}, \mathbf{y}), \mathbf{K}), \mathbf{K})],$
- 4 $\mathbf{A} \leftarrow [\mathbf{A} \setminus \mathbf{x}], \mathbf{A} \leftarrow [\mathbf{A} \setminus \mathbf{y}]$
- 5 **return** $\mathbf{A}$.

Az **ATOM**($\mathbf{K}$) függvény alkalmazásával adódó $\mathbf{A} \in \mathfrak{M}_{n \times m}$ mátrix oszlopai az A_j extenziókat reprezentálják, ahol $\{a_j = (A_j, A'_j) \mid j = 1, 2, \dots, m\}$ a dobozháló atomjainak rendszere. Az $\mathbf{A}$ mátrixot használjuk az objektumhalmazok atomszerkezetének meghatározására is. A $H \subseteq G$ objektumhalmaz $a(H)$ atomszerkezetét azzal az $\mathbf{s}$ sorvektorral reprezentáljuk, melynek j -edik koordinátája pontosan akkor 1, ha $a_j \in a(H)$. Ennek a vektornak az előállításához vegyük észre, hogy ha $\mathbf{h}$ reprezentálja a H halmazt, akkor a $\mathbf{h}^T \mathbf{A}$

vektor j -edik koordinátája azt mutatja meg, hogy H hány objektumot tartalmaz az a_j atom extenziójából; ha ez a szám megegyezik A_j elemszámával, akkor $a_j \in a(H)$. Ennek alapján az $\mathbf{A}$ mátrixból és a $\mathbf{h}$ vektorból az $\mathbf{s}$ sorvektort előállító függvény pszeudokódja a következő:

ATOMSZERKEZET ($\mathbf{A}, \mathbf{h}$)

```

1   $\mathbf{s} \leftarrow \mathbf{h}^T \mathbf{A}$ 
2  for  $j \leftarrow 1$  to  $m$ 
3      if  $s[j] = \sum_{k=1}^n A[k][j]$ 
4          then  $s[j] \leftarrow 1$ 
5          else  $s[j] \leftarrow 0$ 
6  return  $\mathbf{s}$ 
```

Az atomoktól és a háló $\mathbf{0}$ elemétől különböző dobozelemek extenzióit a 3.5. Algoritmus alapján a DOBOZ nevű függvény határozza meg. A függvény két mátrixot használ: a $\mathbf{D}$ oszlopvektorai olyan halmazokat reprezentálnak, amelyek esélyesek arra, hogy különböző dobozelem-extenziók legyenek, e halmazok atomszerkezetei pedig az $\mathbf{S}$ mátrix sorvektorai. Kezdetben mindkét mátrix egyetlen vektorból áll, a csupa egyest tartalmazó, G halmaznak megfelelő vektorokból. A függvény bemenő paraméterei a $\mathbf{K} \in \mathfrak{M}_{n \times k}$ kontextus és az $\mathbf{A} = \text{ATOM}(\mathbf{K}) \in \mathfrak{M}_{n \times m}$ mátrix, melynek oszlopai az atomok extenzióit reprezentálják. Az eredményül kapott $\mathbf{D}$ mátrixból a keresett dobozelem-extenziók, az $\mathbf{S}$ mátrixból pedig azok atomszerkezetei határozhatóak meg.

A 9. sortól kezdve a 3.5. Algoritmus tárgyalásánál említett ellenőrző lépéseket végezzük, de a könnyebb megvalósítás érdekében kissé eltértünk az ott javasolt módszertől. A $\mathbf{D} \leftarrow \mathbf{A}\mathbf{S}^T$ értékadással azt érjük el, hogy a $\mathbf{D}$ mátrix oszlopai az első listára került E extenziók helyett a $C = \bigcup_{x \in a(E)} \text{Ext}(x)$ halmazokat

reprezentálják, ugyanis a dobozelem-extenziók csak ilyen alakúak lehetnek. Ezek után a elegendő a $C = C''$ feltétel ellenőrzése annak eldöntésére, hogy a C halmaz dobozelem-extenzió vagy sem. Ha a feltétel nem teljesül, a C halmazt reprezentáló oszlopot töröljük a $\mathbf{D}$ mátrixból, az atomszerkezetét reprezentáló vektort pedig $\mathbf{S}$ -ből.

DOBOZ($\mathbf{K}, \mathbf{A}$)

```

1   $\mathbf{D} \leftarrow \mathbf{1}_{n \times 1}$ 
2   $\mathbf{S} \leftarrow \mathbf{1}_{1 \times m}$ 
3  for  $\mathbf{k} \in \mathbf{K}$ 
4    for  $\mathbf{d} \in \mathbf{D}$ 
5       $\mathbf{v} \leftarrow \text{METSZET}(\mathbf{k}, \mathbf{d})$ 
6       $\mathbf{a} \leftarrow \text{ATOMSZERKEZET}(\mathbf{A}, \mathbf{v})$ 
7      if  $\sum_{l=1}^m a[l] > 1$  és  $\mathbf{a} \notin \mathbf{S}$ 
8        then  $\mathbf{D} \leftarrow [\mathbf{D}, \mathbf{v}], \mathbf{S} \leftarrow [\mathbf{S}; \mathbf{a}]$ 
9   $\mathbf{D} \leftarrow \mathbf{A}\mathbf{S}^T$ 
10 for  $\mathbf{d} \in \mathbf{D}$ 
11   if  $\mathbf{d} \neq \text{KOBJ}(\text{KTUL}(\mathbf{d}, \mathbf{K}))$ 
12     then  $\mathbf{D} \leftarrow [\mathbf{D} \setminus \mathbf{d}], \mathbf{S} \leftarrow [\mathbf{S} \setminus \text{ATOMSZERKEZET}(\mathbf{A}, \mathbf{d})]$ 
13 return  $\mathbf{D}, \mathbf{S}$ .
```

Az osztályozási rendszerek meghatározásához a DOBOZ függvény eredményeként kapott $\mathbf{S}$ mátrixot használjuk, melynek sorai - az atomoktól és a háló $\mathbf{0}$ elemétől különböző - dobozelemek extenzióinak atomszerkezetét reprezentálják. Ennek a mátrixnak az ij indexű eleme pontosan akkor 1, ha az i -edik dobozelem extenziója tartalmazza a j -edik atom extenzióját. Tegyük fel, hogy a dobozháló m számú atomot tartalmaz, s a hálónak, a $\mathbf{0}$ elemet nem számítva, még l darab dobozeleme van, tehát az $\mathbf{S}$ mátrix $l \times m$ típusú. Az osztályozási rendszereket az általuk tartalmazott nem-atomi dobozelemek felsorolásával adjuk meg (a hiányzó atomokat az $\mathbf{S}$ mátrix alapján fogjuk beazonosítani). E lista elkészítéséhez hozzunk létre egy $\mathbf{T}$ -vel jelölt, l oszlopot tartalmazó mátrixot, melynek ij indexű eleme pontosan akkor 1, ha az i -edik osztályozási rendszer tartalmazza a j -edik nem-atomi dobozelemet. Mivel minden egyes dobozelem kiegészíthető atomokkal egy osztályozási rendszerré, a $\mathbf{T}$ mátrix kezdeti értéke egy $l \times l$ típusú egységmátrix. A 3.7. Algoritmus alapján az osztályozási rendszerek előállításához az atomszerkezetek halmazának minden olyan részhalmazát meg kell keresni, melyek páronként diszjunkt halmazokból állnak. Ez a folyamat az $\mathbf{S}$ mátrixszal végzett

műveletre lefordítva azt jelenti, hogy bármely két olyan sor összege a mátrixhoz írandó, melyek skaláris szorzata 0. (Diszjunkt halmazok unióját az azokat reprezentáló vektorok összege jeleníti meg.) Az $\mathbf{S}$ mátrix bővítésével párhuzamosan a $\mathbf{T}$ mátrixhoz is hozzáírunk egy sort, ami az $\mathbf{S}$ új sorának megfelelő dobozelem-halmazt reprezentálja, és korábbi $\mathbf{T}$ -beli soroknak megfelelő dobozelem-halmazok egyesítését jelenti. Végül a legfinomabb osztályozási rendszert reprezentáló vektorokat is hozzáírjuk az $\mathbf{S}$ és $\mathbf{T}$ mátrixokhoz zérusvektorok formájában.

OSZTALY(S)

```

1   $\mathbf{T} \leftarrow \mathbf{E}_{l \times l}$ 
2  while  $\exists \mathbf{s}_i^* \neq \mathbf{s}_j^* \in \mathbf{S}, \text{METSZET}(\mathbf{s}_i^*, \mathbf{s}_j^*) = \mathbf{0}$ 
3       $\mathbf{S} \leftarrow [\mathbf{S}; \mathbf{s}_i^* + \mathbf{s}_j^*], \mathbf{T} \leftarrow [\mathbf{T}; \mathbf{t}_i^* + \mathbf{t}_j^*]$ 
4   $\mathbf{S} \leftarrow [\mathbf{S}; \mathbf{0}_{1 \times m}], \mathbf{T} \leftarrow [\mathbf{T}; \mathbf{0}_{1 \times l}]$ 
5  return  $\mathbf{S}, \mathbf{T}$ .
```

Az algoritmus eredményeként kapott $\mathbf{S}$ és $\mathbf{T}$ mátrixok kódolják az egyes osztályozási rendszerekbe tartozó atomok és nem-atomi dobozelemek halmazát. Az O_k osztályozási rendszer elemei a mátrixok k -adik sora alapján határozhatóak meg az

$$a_i \in O_k \Leftrightarrow S_{ki} = 0, i = 1, 2, \dots, m,$$

$$d_j \in O_k \Leftrightarrow T_{kj} = 1, j = 1, 2, \dots, l$$

összefüggések alapján, ahol $\{a_1, a_2, \dots, a_m\}$ az atomok, $\{d_1, d_2, \dots, d_l\}$ a $\mathbf{0}$ -tól különböző, nem-atomi dobozelemek halmaza.

A fenti eljárás szemléltetése érdekében leírjuk az osztályozási rendszerek meghatározásának menetét a korábban vizsgált példára. Az ATOM és a DOBOZ függvény alkalmazásával már előállítottuk az atomok és a további dobozelemek extenzióit. Az $a_i = (A_i, A'_i)$ ($i = 1, 2, \dots, 5$) atomok extenziói:

$$A_1 = \{g_1, g_6\}, \quad A_2 = \{g_2, g_7\}, \quad A_3 = \{g_3\}, \quad A_4 = \{g_4, g_5\} \quad A_5 = \{g_8\}.$$

A $\mathbf{0}$ -tól különböző további dobozelemek: $d_i = (D_i, D'_i)$ ($i = 1, 2, \dots, 5$), extenzióik pedig:

$$D_1 = \{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\}, \quad D_2 = \{g_1, g_3, g_6\},$$

$$D_3 = \{g_1, g_4, g_5, g_6, g_8\}, \quad D_4 = \{g_2, g_7, g_8\} \quad D_5 = \{g_2, g_3, g_4, g_5, g_7\}.$$

Az alábbi táblázat az OSZTALY függvény eredményét tartalmazza, a $\mathbf{T}$ mátrix a táblázat baloldali részéből, az $\mathbf{S}$ mátrix jobboldalról olvasható le. $\mathbf{T}$ kezdetben egy 5×5 -ös egységmátrix, az $\mathbf{S}$ mátrix i -dik sora pedig a D_i extenzió atomszerkezetét kódolja. Az $\mathbf{S}$ mátrixban olyan sor-párokat keresünk, melyek diszjunkt atomszerkezettel rendelkező extenziókat kódolnak (azaz a sorvektorok skaláris szorzata 0). A második és negyedik sor ilyen, hiszen D_2 és D_4 atomszerkezetének nincs közös eleme, tehát van olyan osztályozási rendszer, mely a d_2 és d_4 dobozelemeket is tartalmazza. Ezt az osztályozási rendszert kódolja a mátrixok hatodik sora, melyet a második és negyedik sorok összeadásával kaptunk. Most már a bővebb $\mathbf{S}$ mátrixban folytatjuk a „diszjunkt” sorpárok keresését; esetünkben nincs több ilyen, ezért a folyamat leáll.

d_1	d_2	d_3	d_4	d_5	a_1	a_2	a_3	a_4	a_5
1	0	0	0	0	1	1	1	1	1
0	1	0	0	0	1	0	1	0	0
0	0	1	0	0	1	0	0	1	1
0	0	0	1	0	0	1	0	0	1
0	0	0	0	1	0	1	1	1	0
0	1	0	1	0	1	1	1	0	1

Az $\mathbf{S}$ mátrix i -dik sorában 0 jelöli az i -dik osztályozási rendszer atomjait, a további dobozelemeket pedig 1 jelöli a $\mathbf{T}$ mátrix i -dik sorában. A keresett osztályozási rendszerek tehát a következők:

$$O_1 = \{d_1\},$$

$$O_2 = \{d_2, a_2, a_4, a_5\},$$

$$O_3 = \{d_3, a_2, a_3\},$$

$$O_4 = \{d_4, a_1, a_3, a_4\},$$

$$O_5 = \{d_5, a_1, a_5\},$$

$$O_6 = \{d_2, d_4, a_4\}.$$

3.6. Az algoritmusok hatékonysága

A fogalomháló osztályozási rendszereinek előállítására - tudomásunk szerint - eddig még nem született megoldó algoritmus, ezért a bemutatott eljárást nem tudjuk más módszerekkel összevetni hatékonyság szempontjából. A dobozelemek meghatározására Bukszár Z. [4]-ben ugyan javasol egy módszert (mely az összes fogalom előállítása után ellenőrzi, előállnak-e azok atomok szuprérumaként), ennek az algoritmusnak sem készült azonban működő implementációja. Mindenesetre ez a módszer legalább akkora költségű, mint az atomszerkezetek vizsgálatán alapuló 3.5. Algoritmus, hiszen ez utóbbiban nincs szükség a fogalomháló összes elemére. A továbbiakban az osztályozási rendszerek előállításának három lépését megvalósító ATOM, DOBOZ és OSZTALY függvények kiszámításának időigényét szeretnénk megbecsülni. Nyilvánvaló, hogy a megadott algoritmusok javításával, finomhangolásával (vagy más elven működő módszerek kidolgozásával) kedvezőbb költségértékek is elérhetőek. Jelen dolgozat egyik fő célja egy működőképes számítógépes program készítése volt az osztályozási rendszerek meghatározására; a megoldó módszer fejlesztése, optimalizálása a további kutatási feladatok közé tartozik.

Az ATOM függvény a kontextusból kiindulva meghatározza a dobozháló atomjainak extenzióit, és az alábbi pszeudo-kóddal adtuk meg:

```

ATOM(K)
1  A ← KOBJ (KTUL (En×n, K), K)
2  while ∃ x, y ∈ A, x ≠ y, METSZET(x, y) ≠ 0n×1
3    A ← [A, KOBJ (KTUL (UNIO(x, y), K), K)],
4    A ← [A \ x], A ← [A \ y]
5  return A.
```

Tegyük fel, hogy a kontextust egy $n \times k$ típusú mátrix írja le (tehát az objektumok száma n , az attribútumok száma k) és legyen $N = \max\{n, k\}$. Az 1. sorban minden g objektum esetén előállítjuk a g'' halmazt a KTUL és KOBJ függvények egymás utáni alkalmazásával. A KTUL (**E**_{*n*×*n*}, **K**) művelet egy n^2k műveletigényű mátrixszorzásból, és a kapott szorzat elemeinek ellenőrzéséből áll, majd következik a KTUL függvény ugyanennyi művelettel. A 2. sorban szereplő ciklus feltételében az előzőleg kapott $n \times n$ típusú mátrix mátrix oszlopvektorai között keresünk olyan párokat, melyek összemetsző halmazokat reprezentálnak. Eredménytelen keresés esetén $(n-1) + (n-2) + \dots + 2 + 1 = \frac{n(n-1)}{2}$ alkalommal végezzük el az n műveletigényű metszetképzést. Ha találunk két összemetsző halmazt reprezentáló vektort, akkor a 3. sor szerint a halmazok

egyesítését reprezentáló vektorra kell a KTUL és KOBJ kompozíciót képezni. Az unióképzés n műveletet igényel, ehhez adódik a függvények alkalmazása, ami $O(nk)$ költségű egy vektor esetén. Mivel az összemetsző halmazokat reprezentáló vektort töröljük a mátrixból, legfeljebb $n - 1$ -szer kerülhet sor a 3. sorbeli műveletek végrehajtására. Összefoglalva elmondhatjuk, hogy az atomok halmazának számosságától függetlenül az atom-extenziók meghatározása $O(N^3)$ költséggel megoldható.

Az DOBOZ függvény a fogalmak extenzióinak kiszámítására szolgáló 2.20. Algoritmus módosítása, ezért előbb ennek az algoritmusnak a költségét vizsgáljuk meg. Vegyük észre, hogy a fogalomháló elemeinek száma a bemeneti változóknak exponenciális függvénye is lehet. Ha például a G objektumhalmaz és az M attribútumhalmaz egyaránt n elemű, és $I = \{(g_i, m_j) \mid g_i \in G, m_j \in M, i \neq j\}$, akkor az $\mathcal{F}(G, M, I)$ fogalomháló 2^n elemből áll. Éppen ezért a fogalomgeneráló algoritmusok költségbecslésében az input mérete mellett a létrehozott fogalmak számát is figyelembe kell venni. Ebből a szempontból a 2.20. Algoritmus egy elég költséges eljárás, hiszen egy adott halmaz (attribútum-extenziók egy metszete) csak akkor kerül fel a fogalom-extenziókat gyűjtő listára, ha megbizonyosodtunk róla, hogy még nem szerepel ott. Így ha összesen c számú fogalmat generáltunk, a folyamat során legalább $\frac{c(c-1)}{2}$ alkalommal kell a lista korábbi elemeivel való összehasonlításokat elvégezni, azaz az algoritmus költsége a fogalmak számának kvadrátikus függvénye. Ezért van költségcsökkentő hatása a listára kerülő extenziók előzetes szűrésének, melyet az atomszerkezetek vizsgálatával valósítunk meg. Ez esetben is meg kell azonban azt vizsgálni, hogy a soronkövetkező halmaz atomszerkezetével találkozunk-e már korábban, így a DOBOZ függvény is legalább $O(d^2)$ költségű, ahol d az atomoktól és a háló $\mathbf{0}$ elemétől különböző dobozelemek száma. Ezt a számot nehéz előre megbecsülni; lehetséges, hogy a dobozháló csupán a fogalomháló minimális és maximális eleméből áll, de az is előfordulhat, hogy a fogalomháló minden eleme egyben dobozelem is (ennek feltételeit a következő fejezetben tárgyaljuk).

Ha a dobozelemek száma túl nagy, akkor algoritmusunk nagyméretű kontextuson lassan fog működni. Ilyenkor érdemes megfontolni a dobozelemek meghatározásának Bukszár Z. dolgozatában javasolt módját, csak éppen a „naív” algoritmus helyett hatékonyabb fogalomgeneráló módszer alkalmazása után kell az összes fogalom közül a dobozelemeket kiválogatni. Ilyen módszer például a 2. fejezetben már említett *next closure* algoritmus, mely szintén a fogalmak extenzióit állítja elő egy előzőleg G részhalmazai között definiált lineáris rendezés szerinti növekvő sorrendben. Ennek köszönhető, hogy ugyanazt az extenziót nem generálhatjuk még egyszer, így nincs szükség arra, hogy minden új extenzió esetén megvizsgáljuk a listán már szereplő elemeket. Az algoritmus az összes fogalom-extenziót $O(n^2kc)$ időben állítja elő, ahol n az objektumok, k az attribútumok, c a fogalmak száma [36]. A fogalom-

extenziók közül a dobozelem-extenziókat a 3.4. Tétel alapján választhatjuk ki: ha az E extenzióra

$$E = \bigcup_{x \in a(E)} Ext(x)$$

teljesül, akkor E egyúttal egy dobozelem extenziója is. Ezt a vizsgálatot az ATOMSZERKEZET függvény segítségével $O(nm)$ költséggel elvégezhetjük, ahol m a dobozháló atomjainak számát jelöli. A további kutatási feladatok közé tartozik annak vizsgálata, hogy a *next closure* algoritmust lehet-e a 3.5 Algoritmus mintájára úgy módosítani, hogy a fogalom-extenziók helyett eleve dobozelem-extenziókat állítson elő.

Végül az OSZTALY függvény feladata a dobozháló atomjainak és dobozelemeinek ismeretében a lehetséges osztályozási rendszerek meghatározása. A 3.6. Állítás értelmében az atomok halmazának partícióit kell előállítani dobozelemek atomszerkezeteiből és atomokból. Ez a feladat tekinthető egy nevezetes NP-teljes nyelv, a halmaz-partíció probléma kereső feladatának. A halmaz-partíció problémában adott egy H halmaz és részhalmazainak egy rendszere, el kell döntenünk, hogy a részhalmazokból kiválaszthatóak-e olyanok, melyek H egy partícióját adják. A kérdés eldöntésére valószínűleg nincs polinomiális idejű algoritmus [37]. Esetünkben azonban biztos, hogy létezik a feladatnak megoldása, hiszen minden egyes atom ott van a felhasználható részhalmazok között. A kereső feladat megoldásához, azaz az atomhalmaz partícióinak előállításához az OSZTALY függvény egy listára gyűjti az összes olyan rendszert, melynek elemei diszjunkt atomszerkezetű nem-atomi dobozelemek. A listára kerülő új rendszerek esetén meg kell vizsgálni, hogy a listára majd később kerülő rendszerekkel kombinálva alkothatnak-e újabb megoldást. A listát kezdetben azok a rendszerek alkotják, melyek az egyes nem-atomi dobozelemek atomszerkezetéből állnak, majd a megfelelő párok keresése során a külső ciklus $i = 1$ -től, a belső $j = i + 1$ -től fut a lista aktuális elemszámáig. Ezért az algoritmus költsége legrosszab esetben $O(s^2)$, ahol s az osztályozási rendszerek száma.

Fontos hangsúlyozni, hogy a műszaki gyakorlatban felmerülő osztályozási problémák esetén a dobozháló elemszáma lényegesen kisebb a fogalmak számánál, így a vizsgált függvények rövid idő alatt eredményt szolgáltatnak (ezt a 6. fejezetben fogjuk látni). Túl sok dobozelem esetén olyan nagyszámú is lehet az osztályozási rendszerek halmaza, hogy az osztályfelbontási feladat egyszerűen értelmét veszti. Az algoritmusok elemzésekor említett legrosszab eset ugyanis azt a szituációt írja le, amikor az objektumok tetszős szerint csoportosítása is megoldást jelent, ami a valós feladatokban nyilván nem fordulhat elő.

4. Fogalomháló és dobozháló egyenlősége

A dobozelemek definíciója szerint a dobozhálót a fogalomháló azon elemei alkotják, amelyek a fogalomháló valamely osztályozási rendszerének is elemei. A dobozháló tehát azokból a fogalmakból áll, amelyek a lehetséges osztályozások előállítása szempontjából lényegesek. Felmerül a kérdés, hogy lehetséges-e olyan eset, amikor a fogalomháló minden eleme lényeges, azaz minden fogalom egyben dobozelem is. A következő, [42]-ben szereplő állítás erre ad választ.

4.1. Állítás. *Ha L egy teljes, atomisztikus háló, akkor az L háló és a $\mathcal{D}(L)$ dobozháló megegyezik.*

Bizonyítás. A bizonyításhoz azt kell észrevennünk, hogy $\forall x \in L \setminus \{\mathbf{0}\}$ esetén az $S_x = \{x\} \cup \{a \text{ atom } L\text{-ben} \mid a \wedge x = \mathbf{0}\}$ rendszer az L háló egy osztályozási rendszere, így minden $x \in L$ elem egyúttal dobozelem is. $\square$

Az állítás szerint, ha az $\mathcal{F}(G, M, I)$ fogalomháló atomisztikus, akkor megegyezik a $\mathcal{D}(G, M, I)$ dobozhálóval. A továbbiakban a két háló egyenlőségére egy olyan szükséges és elégséges feltételt adunk, amely fogalomháló helyett a kontextus sajátosságait használja. Az eredmény a következő kérdés vizsgálatából született: ha adott egy (G, M, I) kontextus, az ahhoz tartozó $\mathcal{F}(G, M, I)$ fogalomháló és a $\mathcal{D}(G, M, I)$ dobozháló, akkor lehet-e olyan (G, M^*, I^*) kontextust megadni, melyhez tartozó fogalomháló izomorf a $\mathcal{D}(G, M, I)$ dobozhálóval.

4.2. Megjegyzés. A 2.23. Állítás szerint tetszőleges véges V háléhoz van olyan kontextus, amelyhez tartozó fogalomháló izomorf V -vel, tehát $\mathcal{D}(G, M, I)$ -hez is létezik ilyen kontextus. Ez a kontextus azonban nagyon általános, objektumhalmaza a háló egyesítés-irreducibilis elemeinek összessége. A fejezet fő eredményeként megmutatjuk, hogy az objektumhalmaz megváltoztatása nélkül is lehet olyan kontextust konstruálni, melynek fogalomhálójá $\mathcal{D}(G, M, I)$ -vel izomorf.

Legyen (G, M, I) egy kontextus és jelölje Π_0 a G halmaz legfinomabb extenzió-partícióját. A 2.34. Tétel értelmében Π_0 a $\mathcal{D}(G, M, I)$ dobozháló atomjainak extenzióiból áll. Jelölje A_g a Π_0 extenzió-partíciónak azt a blokkját, amely a $g \in G$ elemet tartalmazza. A következő állításban az A_g halmazok segítségével jellemezzük a dobozelemek extenzióit.

4.3. Állítás. *Legyen (G, M, I) egy kontextus és $E \subseteq G$ egy extenzió. A következő három állítás ekvivalens:*

- (i) *E a $\mathcal{D}(G, M, I)$ dobozháló valamely elemének az extenziója,*
- (ii) *$\forall g \in E$ esetén $A_g \subseteq E$,*
- (iii) *$\forall g \notin E$ esetén $A_g \cap E = \emptyset$.*

Bizonyítás. (i) $\implies$ (ii) Tegyük fel, hogy $(E, E') \in \mathcal{D}(G, M, I)$ és $g \in E$. Az A_g halmaz G legfinomabb osztályozásának g -t tartalmazó blokkja, ezért a durvább partícióban szereplő E halmaz szükségképpen tartalmazza A_g -t. (ii) $\implies$ (iii) Tegyük fel, hogy $g \notin E$ és $A_g \cap E \neq \emptyset$. Ha $y \in A_g \cap E$, akkor $y \in A_g$ és ekkor $g \in A_y$, másrészt $y \in E$, de akkor (ii) miatt $A_y \subseteq E$. Ebből azonban $g \in E$ következik, ami ellentmondás. (iii) $\implies$ (i) Ha minden $g \notin E$ esetén $A_g \cap E = \emptyset$ teljesül, akkor az $E \cup \{A_g \mid g \notin E\}$ rendszer G egy extenzió-partíciója, ami azt jelenti, hogy E valamely dobozelem extenziója a $\mathcal{D}(G, M, I)$ hálóban. $\square$

4.4. Állítás. *Dobozelemek extenzióinak metszete is dobozelem-extenzió.*

Bizonyítás. Legyen $\{E_j \mid j \in J\}$ dobozelemek extenzióinak egy rendszere. A 2.14. Állítás szerint a $\bigcap_{j \in J} E_j$ halmaz is egy extenzió, és tegyük fel, hogy egy g objektumra $g \notin \bigcap_{j \in J} E_j$. Akkor $g \notin E_j$ valamely $j \in J$ esetén, így $A_g \cap E_j = \emptyset$ az előző állítás értelmében. Ekkor $A_g \cap \bigcap_{j \in J} E_j = \emptyset$, amiből, szintén a 4.3. Állítás alapján az következik, hogy a $\bigcap_{j \in J} E_j$ halmaz egy dobozelem-extenzió. $\square$

Rátérünk a 4.2. Megjegyzésben említett konstrukcióra.

4.5. Állítás. *Legyen (G, M, I) egy tetszőleges kontextus, $\mathcal{F}(G, M, I)$ a hozzá tartozó fogalomháló, $\mathcal{D}(G, M, I)$ pedig a fogalomháló dobozhálója. Ekkor*
(i) *van olyan (G, M^*, I^*) kontextus, amelyre $\mathcal{F}(G, M^*, I^*) \cong \mathcal{D}(G, M, I)$,*
(ii) *ha a $p \in M$ attribútumra $(p', p'') \in \mathcal{D}(G, M, I)$, akkor van olyan $q \in M^*$, amelyre $q' = p'$.*

Bizonyítás. A (G, M^*, I^*) kontextus konstrukciója a következő: minden $\mathbf{0}$ -tól és $\mathbf{1}$ -től különböző $(D, D') \in \mathcal{D}(G, M, I)$ dobozelem esetén képzünk egy $m(D)$ -vel jelölt attribútumot, amelyekkel pontosan azok a $g \in G$ objektumok rendelkeznek, amelyekre $g \in D$ is teljesül. Ezzel az előírással kapjuk a (G, M^*, I^*) kontextust, ahol $M^* = \{m(D) \mid (D, D') \in \mathcal{D}(G, M, I)\}$, a reláció pedig $I^* = \{(g, m(D)) \mid g \in G, m(D) \in M^*, g \in D\}$. Az $\mathcal{F}(G, M^*, I^*)$ és a $\mathcal{D}(G, M, I)$ háló izomorfiaja abból következik, hogy a két háló fogalmainak extenziói megegyeznek. Ezt a következőképpen láthatjuk be: legyen E az $\mathcal{F}(G, M^*, I^*)$ fogalomháló egy tetszőleges elemének extenziója. Mivel minden fogalom-extenzió előáll attribútum-extenziók metszeteként, van olyan $\{m(D_j) \mid j \in J\} \subseteq M^*$ attribútum-halmaz, hogy $E = \bigcap_{j \in J} (m(D_j))'$.

Az $m(D_j)$ attribútumok definíciója alapján $(m(D_j))' = D_j$, így $E = \bigcap_{j \in J} D_j$.

Ebből a 4.4. Állítás alkalmazásával már következik, hogy E a $\mathcal{D}(G, M, I)$

5. táblázat. A K^* kontextus konstrukciója és a K^{**} kontextus

K^*	Attribútumok									
Objektumok		$m(D_1)$	$m(D_2)$	$m(D_3)$	$m(D_4)$	$m(D_5)$	$m(D_6)$	$m(D_7)$	$m(D_8)$	$m(D_9)$
	g_1	x	x					x		
	g_2			x	x				x	
	g_3	x			x	x				
	g_4		x		x		x			
	g_5		x		x		x			
	g_6	x	x					x		
	g_7			x	x				x	
	g_8		x	x						x

valamely elemének extenziója.

Megfordítva, ha D a $\mathcal{D}(G, M, I)$ háló egy elemének extenziója, akkor az M^* halmaz $m(D)$ tulajdonságával az $(m(D))' = D$ azonosságból az következik, hogy D egy attribútum-extenzió az $\mathcal{F}(G, M^*, I^*)$ fogalomhálóban, tehát az $\mathcal{F}(G, M^*, I^*)$ háló egy fogalmának extenziója.

Az (ii) állítás bizonyításához legyen $(p', p'') \in \mathcal{D}(G, M, I)$, akkor a $q = m(p') \in M^*$ attribútumra $q' = (m(p'))' = p'$ teljesül. $\square$

4.6. *Megjegyzés.* a) Ha a $\mathcal{D}(G, M, I)$ dobozháló n elemet tartalmaz, akkor a (G, M^*, I^*) kontextus attribútumainak száma $n - 2$ (a legkisebb és legnagyobb dobozelemhez nem készítünk új attribútumot). A kontextus méretének csökkentése érdekében érdemes elvégezni a kontextus oszlopredukcióját, az így kapott kontextusra is igaz az, hogy a hozzá tartozó fogalomháló izomorf $\mathcal{D}(G, M, I)$ -vel. A korábbi fejezetekben szereplő $K = (G, M, I)$ kontextushoz a $\mathcal{D}(G, M, I)$ háló elemeinek ismeretében előállítottuk a $K^* = (G, M^*, I^*)$ kontextust, majd az ezen elvégzett oszlopredukció a K^{**} kontextust eredményezte. Az 5. táblázat a K^* és K^{**} kontextust mutatja be, ez utóbbit sötétebb oszlopok jelzik. A K^{**} kontextushoz tartozó fogalomháló megegyezik a $\mathcal{D}(G, M, I)$ dobozhálóval.

b) Az olyan $p \in M$ tulajdonságokat, amelyekre p' valamely dobozelemnek az extenziója, **kitüntetett tulajdonságoknak** hívjuk. Az (ii) állításból az is következik, hogy az új, M^* attribútumhalmaz az összes kitüntetett tulajdonságot tartalmazza, hiszen ha valamely $p \in M$ -re $(p', p) \in \mathcal{D}(G, M, I)$, akkor van olyan $q \in M^*$, amelyre $q' = p'$.

c) A dobozelemeken alapuló osztályfelbontási technika továbbfejlesztésében a kitüntetett tulajdonságok vizsgálata meghatározó jelentőségű. Egy osztályozási feladat végrehajtásakor legtöbbször vannak olyan jellegű tulajdonságai az objektumoknak, amelyeket hangsúlyosabban figyelembe kell venni az osztályok kialakításakor, míg más jellemzők kevésbé fontosak. Például az alkatrészgyártás területén a gyártás minél hatékonyabb szervezése érdekében az alkatrészeket elsődlegesen a megmunkálásukhoz szükséges gépek szerint érdemes osztályozni. A csoportosítás lényeges szempontjait már a kontextus megadásakor is kiemelhetjük a következő meghatározással: azt mondjuk, hogy a (G, M, P, I) négyes egy *kitüntetett tulajdonságokkal jellemzett kontextus*, ha (G, M, I) egy kontextus, $P \subseteq M$ attribútumok egy nemüres halmaza úgy, hogy

$$g \in p' \Rightarrow A_g \subseteq p' \quad \forall p \in P$$

(A_g a G halmaz legfinomabb extenzió-partíciójában a g elemet tartalmazó blokkot jelöli).

A (G, M^*, I^*) kontextust a következő állítás bizonyításában is felhasználjuk.

4.7. Állítás. *Legyen (G, M, I) egy tetszőleges kontextus. $\mathcal{F}(G, M, I) = \mathcal{D}(G, M, I)$ akkor és csak akkor teljesül, ha minden $p \in M$ -re $(p', p'') \in \mathcal{D}(G, M, I)$.*

Bizonyítás. Tegyük fel, hogy $\mathcal{F}(G, M, I) = \mathcal{D}(G, M, I)$, és $p \in M$ egy tetszőleges attribútum. Ekkor p' egy attribútum-extenzió, azaz (p', p'') az $\mathcal{F}(G, M, I)$ fogalomháló egy eleme. Mivel $\mathcal{F}(G, M, I) = \mathcal{D}(G, M, I)$, a (p', p'') fogalom a dobozhálónak is eleme.

Megfordítva, tegyük fel, hogy minden $p \in M$ -re $(p', p'') \in \mathcal{D}(G, M, I)$ és legyen (A, B) az $\mathcal{F}(G, M, I)$ fogalomháló egy eleme. Az A extenzió előáll attribútum-extenziók metszeteként, tehát $A = \bigcap_{j \in J} p'_j$, ahol $p_j \in M, j \in J$.

Legyen (G, M^*, I^*) az 4.5. Állításban szereplő konstrukció szerint létrehozott kontextus, akkor az állítás második része alapján $M \subseteq M^*$. Ebből az következik, hogy p'_j minden $j \in J$ esetén attribútum-extenzió az $\mathcal{F}(G, M^*, I^*)$ hálóban, így $A = \bigcap_{j \in J} p'_j$ egy fogalom-extenzió az $\mathcal{F}(G, M^*, I^*)$ hálóban, és

a 4.5. Állítás szerint a $\mathcal{D}(G, M, I)$ hálóban is. Tehát $(A, B) \in \mathcal{D}(G, M, I)$, amiből $\mathcal{F}(G, M, I) \subseteq \mathcal{D}(G, M, I)$ következik. Mivel a fordított tartalmazás triviális, beláttuk a fogalomháló és a dobozháló egyenlőségét. $\square$

A fejezet eredményeit foglalja össze az alábbi következmény, mely a 4.1. és 4.7. Állítás, valamint a kitüntetett tulajdonságok fogalmának felhasználásával könnyen belátható.

4.8. Következmény. *Legyen (G, M, I) egy tetszőleges kontextus. Az alábbi állítások ekvivalensek:*

- (i) Az $\mathcal{F}(G, M, I)$ fogalomháló és a $\mathcal{D}(G, M, I)$ dobozháló megegyezik.*
- (ii) Minden $p \in M$ tulajdonság kitüntetett.*
- (iii) Az $\mathcal{F}(G, M, I)$ fogalomháló atomisztikus.*

5. Részkontextus és bővített kontextus

A továbbiakban arra keresünk választ, hogy egy (G, M, I) kontextus objektumhalmazának szűkítése illetve bővítése esetén az új kontextushoz tartozó dobozelemek milyen kapcsolatban állnak az eredeti kontextusból származtatható dobozháló elemeivel. Az objektumhalmaz módosításakor az attribútumok halmazát nem változtatjuk, így a kontextus szűkítése bizonyos sorok eltörlését jelenti a kontextust reprezentáló táblázatból, míg a kontextus bővítése esetén új sorokat írunk a már meglévőkhöz. A kontextus bővítése esetén csak az elemi bővítésekkel foglalkozunk, ebben az esetben egyetlen objektummal egészítjük ki a G halmazt. Ha a kontextust több objektummal bővítjük, egymás utáni elemi bővítések elvégzése után határozhatjuk meg az új kontextushoz tartozó dobozelemeket.

Fontos megjegyezni, hogy a felvetett kérdéseket a dobozhálótól, osztályozási hálótól függetlenül is vizsgálhatjuk. Ez a szemlélet jellemzi a [K1] publikációt, melyben Bernhard Ganter javaslatára az extenzió-partíciókra koncentrálnak, így leírhatjuk a kontextus szűkítésének és bővítésének következményeit. Bizonyítható, hogy tetszőleges (G, M, I) kontextus esetén a G halmaz extenzió-partíciói teljes hálót alkotnak, így az is igaz, hogy G legfinomabb osztályozása, Π_0 mindig létezik. Ha a kontextus sor-redukált, akkor Π_0 blokkjai megegyeznek a dobozháló legfinomabb osztályozási rendszerét alkotó fogalmak, az atomok extenzióival. Annak érdekében, hogy a korábbi, hálókra megfogalmazott eredményeket alkalmazhassuk a bizonyításokban, a fejezet további részében kontextus alatt mindig sor-redukált kontextust értünk.

5.1. Részkontextushoz tartozó extenzió-partíciók

Legyen $K = (G, M, I)$ egy kontextus és $H \subseteq G$ objektumok egy halmaza. Vezessük be az $I_H = I \cap (H \times M)$ jelölést, és képezzük a $K_H = (H, M, I_H)$ részkontextust. A félreértések elkerülése végett a $()'$ operátor helyett a továbbiakban az $()^I$ illetve $()^{I_H}$ jelölést alkalmazzuk, attól függően, hogy az eredeti vagy a részkontextus alapján keressük egy objektumhalmaz közös tulajdonságait, vagy egy attribútumhalmaz mindegyik tulajdonságával rendelkező objektumok halmazát. Mivel a K_H részkontextust reprezentáló táblázat úgy kapjuk, hogy töröljük a K kontextus táblázatából a $G \setminus H$ halmaz objektumaihoz tartozó sorokat, világos, hogy egy $A \subseteq H$ halmaz esetén $A^{I_H} = A^I$, és $B \subseteq M$ esetén $B^{I_H} = B^I \cap H$. Ezeket az összefüggéseket a következő lemma bizonyításában is felhasználjuk. Azt, hogy $E \subseteq G$ az $\mathcal{F}(G, M, I)$ fogalomháló egy elemének extenziója, az egyszerűbb fogalmazás érdekében úgy fogjuk mondani, hogy E egy extenzió a (G, M, I) kontextusban.

5.1. Lemma. *Ha E egy extenzió a K kontextusban, akkor $E \cap H$ egy extenzió a K_H részkontextusban.*

Bizonyítás. Azt kell belátnunk, hogy $(E \cap H)^{I_H I_H} = E \cap H$. Az $E \cap H \subseteq (E \cap H)^{I_H I_H}$ reláció automatikusan teljesül. A fordított tartalmazás igazolásához induljunk ki az $E^I \subseteq (E \cap H)^I = (E \cap H)^{I_H}$ összefüggésből, amiből $(E \cap H)^{I_H I} \subseteq E^{II}$ következik. Mivel E egy extenzió a K kontextusban, $E^{II} = E$, tehát $(E \cap H)^{I_H I} \subseteq E$. Az $(E \cap H)^{I_H I_H}$ halmazt felírhatjuk $\left((E \cap H)^{I_H}\right)^{I_H} = \left((E \cap H)^{I_H}\right)^I \cap H$ alakban, amiből az előzőek alapján $(E \cap H)^{I_H I_H} = \left((E \cap H)^{I_H}\right)^I \cap H \subseteq E \cap H$ következik. $\square$

5.2. Állítás. Ha $\Pi = \{E_j \mid j \in J\}$ a G halmaz egy extenzió-partíciója, akkor

$$\Pi_H = \{E_j \cap H \mid E_j \cap H \neq \emptyset, j \in J\}$$

a H halmaz egy extenzió-partíciója.

Bizonyítás. Mivel az $\{E_j \mid j \in J\}$ halmazrendszer G egy partíciója, az $\{E_j \cap H \mid E_j \cap H \neq \emptyset, j \in J\}$ rendszer a H halmaz egy partíciója. Az E_j halmazok mindegyike egy extenzió a K kontextusban, így az előző lemma alapján az $E_j \cap H, j \in J$ halmazok extenziók a K_H kontextusban, tehát Π_H valóban extenzió-partíciója a H halmaznak. $\square$

5.3. Megjegyzés. Ha $\Pi = \Pi_0$, tehát a G halmaz legfinomabb extenzió-partíciójából indulunk ki, akkor a leszűkítéssel kapott Π_H rendszer a H halmaz egy extenzió-partíciója lesz, de nem feltétlenül a legfinomabb; a H halmaz legfinomabb extenzió-partíciója, $\Pi_0(H)$ finomítja a Π_H partíciót.

5.4. Következmény. Ha E a $\mathcal{D}(G, M, I)$ dobozháló valamely elemének az extenziója, akkor $E \cap H$ a $\mathcal{D}(H, M, I_H)$ dobozháló valamely elemének az extenziója.

Bizonyítás. Ha $E \cap H = \emptyset$, akkor $E \cap H$ a $\mathcal{D}(H, M, I_H)$ dobozháló $\mathbf{0}$ elemének, a $(\emptyset, M)$ fogalomnak az extenziója. Ha $E \cap H \neq \emptyset$, akkor E sem lehet üres, így egy blokkja a G halmaz valamely Π extenzió-partíciójának. Az 5.2. Állítás alapján az $E \cap H$ halmaz szerepel a H egy extenzió-partíciójában, tehát $E \cap H$ a $\mathcal{D}(H, M, I_H)$ dobozháló valamely elemének az extenziója. $\square$

Ha $H \subseteq G$, jelöljük $A_x(H)$ -val, illetve $A_x(G)$ -vel a H , illetve a G halmaz legfinomabb extenzió-partíciójában az x elemet tartalmazó blokkot. A következő állítás a kitüntetett tulajdonságokkal jellemzett kontextus részkontextusát írja le.

5.5. Állítás. Legyen (G, M, P, I) egy kitüntetett tulajdonságokkal jellemzett kontextus, és tegyük fel, hogy (H, M, I_H) a (G, M, I) kontextus részkontextusa. Akkor a (H, M, P, I_H) négyes szintén egy kitüntetett tulajdonságokkal jellemzett kontextust ad meg.

Bizonyítás. Azt kell igazolnunk, hogy bármely $p \in P$ attribútum esetén $x \in p^{I_H}$ -ból $A_x(H) \subseteq p^{I_H}$ következik. Legyen $p \in P$, $x \in H$ és tegyük fel, hogy $x \in p^{I_H}$. Akkor $x \in p^I$ is teljesül, és így $A_x(G) \subseteq p^I$, hiszen (G, M, P, I) egy kitüntetett tulajdonságokkal jellemzett kontextus. Ebből $A_x(G) \cap H \subseteq p^I \cap H = p^{I_H}$ következik. Az 5.3. Megjegyzés értelmében a $\Pi_0(H)$ partíció finomabb, mint a Π_0 partíció H -ra való leszűkítése, így $A_x(H) \subseteq A_x(G) \cap H$. Tehát $A_x(H) \subseteq p^{I_H}$, ami azt jelenti, hogy (H, M, P, I_H) egy kitüntetett tulajdonságokkal jellemzett kontextus. $\square$

5.2. A kontextus elemi bővítése

Az előző alfejezetben azt vizsgáltuk, hogy a kontextus objektumrészének leszűkítésével kapott részkontextus esetén az eredeti kontextusból származó információk hogyan használhatóak fel az új kontextushoz tartozó struktúrák meghatározásakor. Általános tapasztalatként az adódott, hogy az eredeti kontextus alapján meghatározott extenziók, dobozelem-extenziók, extenzió-partíciók öröklődnek, és az örökítést egyszerűen a megfelelő objektumhalmazok leszűkítésével valósíthatjuk meg. Bonyolultabb, és a gyakorlat szempontjából fontosabb kérdés, hogy a kontextus bővítésével adódó új helyzetben a szűkebb kontextus esetén már meghatározott ismereteket mennyire hatékonyan tudjuk felhasználni. A továbbiakban a kontextus elemi bővítésével foglalkozunk, ezen az objektumhalmaz egy elemmel való bővítését értjük, mely új elem ugyanabból az attribútumhalmazból rendelkezhet bizonyos tulajdonságokkal, mint az eredeti objektumok.

5.6. Állítás. *Legyen $H = G \setminus \{z\}$ és E a $\mathcal{D}(G, M, I)$ dobozháló valamely elemének az extenziója. Ekkor a következő két állítás egyike biztosan teljesül:*
 (i) *E a $\mathcal{D}(H, M, I_H)$ dobozháló egy elemének az extenziója és $E \cap A_z(G) = \emptyset$.*
 (ii) *$E \setminus \{z\}$ a $\mathcal{D}(H, M, I_H)$ dobozháló egy elemének az extenziója.*

Bizonyítás. Az 5.4. Következmény szerint az $E \cap H$ halmaz a $\mathcal{D}(H, M, I_H)$ dobozháló egy elemének az extenziója, így elegendő bizonyítani, hogy az első esetben E , a másodikban pedig $E \setminus \{z\}$ megegyezik az $E \cap H$ halmazzal.
 (i) Ha $E \cap A_z(G) = \emptyset$, akkor $E \cap \{z\} = \emptyset$ és így $E \subseteq H$, tehát $E = E \cap H$.
 (ii) Ha $E \cap A_z(G) \neq \emptyset$, akkor $E \setminus \{z\} = E \cap H$. $\square$

A következő állítás a z elemmel való bővítés után írja le az objektumhalmaz legfinomabb extenzió-partíciójának szerkezetét.

5.7. Állítás. *Ha $H = G \setminus \{z\}$, akkor G legfinomabb extenzió-partíciója az $A_z(G)$ halmazból és H legfinomabb extenzió-partíciójának $A_z(G)$ -től diszjunkt blokkjaiból áll, azaz*

$$\Pi_0 = A_z(G) \cup \{A_h(H) \mid h \in H \setminus A_z(G)\}.$$

Bizonyítás. A Π_0 halmazrendszer lefedi a G halmazt, hiszen z -t (és esetleg a H egy részét) lefedi az $A_z(G)$ halmaz, a $H \setminus A_z(G)$ maradék részt pedig lefedik az $A_h(H)$ halmazok. Az $A_h(H)$ halmazok páronként diszjunktak is definíciójukból adódóan, azt kell még látnunk, hogy bármely $h \in H \setminus A_z(G)$ esetén $A_z(G) \cap A_h(H) = \emptyset$ is teljesül. Tegyük fel, hogy $h \in H \setminus A_z(G)$ és $y \in A_z(G) \cap A_h(H)$. Ekkor $y \in A_z(G)$, amiből $A_z(G) = A_y(G)$ következik, és $y \in A_h(H)$, ami azt jelenti, hogy $h \in A_y(H)$. A 5.3. Megjegyzés alapján $A_y(H) \subseteq A_y(G)$, így $h \in A_y(H) \subseteq A_y(G) = A_z(G)$, ami ellentmond a $h \in H \setminus A_z(G)$ feltételnek. Π_0 tehát egy partíciója a G halmaznak, meg kell mutatnunk, hogy minden egyes blokk egy extenzió is a (G, M, I) -hez tartozó fogalomhálóban. Ez az $A_z(G)$ halmaz esetén definíciójából eredően teljesül, a többi blokk esetében az $(A_h(H))^{II} = A_h(H)$ egyenlőséget kell igazolnunk minden $h \in H \setminus A_z(G)$ esetén. Legyen $h \in H \setminus A_z(G)$ egy tetszőleges objektum. Ekkor $h \notin A_z(G)$, és a 4.3. Állítás szerint $A_z(G) \cap A_h(G) = \emptyset$, hiszen $A_z(G)$ egy dobozelem extenziója a $\mathcal{D}(G, M, I)$ hálóban. Ebből $A_h(G) \subseteq H$ következik. Másrészt $A_h(H) \subseteq A_h(G)$, így $(A_h(H))^{II} \subseteq (A_h(G))^{II} = A_h(G)$, azaz $(A_h(H))^{II} \subseteq H$ is igaz. Ebből pedig $(A_h(H))^{II} = A_h(H)$ következik az alábbi gondolatmenet alapján:

$$\begin{aligned} (A_h(H))^{II} &= \{x \in G \mid (x, m) \in I \ \forall m \in (A_h(H))'\} = \\ &= \{x \in H \mid (x, m) \in I_H \ \forall m \in (A_h(H))'\} = A_h(H). \end{aligned}$$

Az utolsó egyenlőség azért áll fenn, mert $A_h(H)$ egy extenzió a $\mathcal{F}(H, M, I_H)$ fogalomhálóban.

Azt kell még belátnunk, hogy Π_0 a G halmaz legfinomabb extenzió-partíciója. Képezzük a

$$\Pi = \{A_g(G) \mid g \in G\} = A_z(G) \cup \{A_h(G) \mid h \in H \setminus A_z(G)\}$$

partíciót, mely az $A_g(G)$ halmazok definíciója miatt a legfinomabb extenzió-partíciója a G halmaznak. Ha $h \in H \setminus A_z(G)$, akkor $A_h(H)$ a (G, M, I) kontextus szerint is egy h -t tartalmazó extenzió, ezért $A_h(G) \subseteq A_h(H)$. Másrészt, mivel H a G halmaz leszűkítése, $A_h(H) \subseteq A_h(G)$. Beláttuk, hogy $A_h(H) = A_h(G)$ minden $h \in H \setminus A_z(G)$ esetén, így a Π_0 és Π partíciók megegyeznek, tehát valóban Π_0 a G halmaz legfinomabb extenzió-partíciója. $\square$

A Π_0 extenzió-partíció előállításához a 3.1. Algoritmust használjuk, de nincs arra szükség, hogy az eljárást az objektumok szintjéről kezdjük. Az algoritmus első lépésében kihasználjuk, hogy a H halmaz legfinomabb extenzió-partíciójának elemeit már ismerjük, ezek közül kell valamelyikbe az új elemet beillesztenünk, vagy ha ez nem lehetséges, új osztályt kell létrehozni számára.

5.8. Algoritmus (Atom-extenziók meghatározása elemi bővítés után).

Legyen $\Pi_0 = \Pi_0(H) \cup \{z\}^{II}$

amíg van olyan $X \neq Y \in \Pi_0$, amelyre $X \cap Y \neq \emptyset$

addig Π_0 -t bővítsd az $(X \cup Y)^{II}$ halmazzal

és Π_0 -ból töröld az X és Y halmazt.

A következő állítás szerint a Π_0 partícióban a $\{z\}$ halmaz pontosan akkor alkot önálló blokkot, ha a $(\{z\}, \{z\})'$ fogalom a dobozháló egy eleme.

5.9. Állítás. $(\{z\}, \{z\})' \in \mathcal{D}(G, M, I) \Leftrightarrow A_z(G) = \{z\}$.

Bizonyítás. Tegyük fel, hogy $(\{z\}, \{z\})'$ a dobozháló egy eleme. A 4.3. Állítás miatt $A_z(G) \subseteq \{z\}$. A fordított tartalmazás $A_z(G)$ definíciójából következik, így $A_z(G) = \{z\}$. Megfordítva, ha $A_z(G) = \{z\}$, akkor $(\{z\}, \{z\})' = (\{A_z(G)\}, \{A_z(G)\})'$, ami a dobozháló egy eleme, hiszen $A_z(G)$ a legfinomabb extenzió-partíció egy blokkja. $\square$

Az állításban szereplő kérdésre 5.13.-ban még visszatérünk, a gyakorlatban jobban alkalmazható feltételelt megfogalmazva az $A_z(G) = \{z\}$ egyenlőségre. Ennek előkészítésére bevezetünk egy új fogalmat.

5.10. Definíció. Azt mondjuk, hogy a $P \subseteq H = G \setminus \{z\}$ objektumhalmaz a $z \in G$ objektum egy **premisszája**, ha $z \in P^{II}$. Vezessük be a $P_z^\cup$ jelölést a $\Pi_0(H)$ partíció azon blokkjainak az uniójára, amelyek z -nek premisszái, azaz legyen

$$P_z^\cup := \cup \{B \in \Pi_0(H) \mid z \in B^{II}\}.$$

Vegyük észre, hogy ha E a (H, M, I_H) kontextus egy fogalmának az extenziója, akkor vagy premisszája z -nek, vagy egy olyan fogalomnak az extenziója, ami a (G, M, I) kontextushoz tartozik, hiszen ha $z \notin E^{II}$, akkor $E = E^{I_H I_H} = E^{II} \cap H = E^{II}$.

5.11. Tétel. *A z elemet tartalmazó $E \subseteq G$ halmaz pontosan akkor extenziója a $\mathcal{D}(G, M, I)$ dobozháló egy elemének, ha teljesülnek az alábbi feltételek:*

- (1) $E^{II} = E$,
- (2) $E \setminus \{z\}$ a (H, M, I_H) kontextus egy dobozelemének az extenziója,
- (3) E a $\Pi_0(H)$ partíció minden olyan blokkját tartalmazza, ami z -nek premisszája.

Bizonyítás. Legyen E a (G, M, I) kontextus egy dobozelemének az extenziója. Akkor (1) nyilván teljesül és alkalmazva az 5.6. Állítást a (2) feltételt is rögtön megkapjuk. Tegyük fel most, hogy valamely $B \in \Pi_0(H)$ -ra $z \in B^{II}$ teljesül. Mivel $\Pi_0(H) \leq \Pi_0$, ezért B részhalmaza a Π_0 partíció valamely $A \in \Pi_0$ blokkjának. De akkor $z \in B^{II} \subseteq A^{II} = A$, így $A = A_z(G)$. Mivel az E halmaz előáll úgy, mint Π_0 valamely blokkjainak az uniója, és mivel $z \in E$, ezért $A_z(G) \subseteq E$, amiből $B^{II} \subseteq E$ is következik. Így a (3) feltétel is teljesül. Fordítva, tegyük fel hogy az E halmazra teljesülnek az (1), (2) és (3) feltételek. Akkor $G \setminus E \subseteq H$ és így $G \setminus E = H \setminus (E \setminus \{z\})$. Mivel a (2) feltétel értelmében az $E \setminus \{z\}$ halmaz a $\Pi_0(H)$ partíció valamely blokkjainak az uniója, ezért a $G \setminus E = H \setminus (E \setminus \{z\})$ halmaz is valamely $B_k \in \Pi_0(H)$, $k \in K$ ($K \neq \emptyset$) blokkoknak az uniója. A (3) feltétel miatt $z \notin B_k^{II}$, $k \in K$, így $B_k^{II} \subseteq H$ minden $k \in K$ -ra teljesül. Mivel minden B_k , $k \in K$ halmaz extenzió a (H, M, I_H) kontextusban, ezért $B_k = B_k^{I_H I_H} = B_k^{II} \cap H = B_k^{II}$, $k \in K$, vagyis minden B_k , $k \in K$ a (G, M, I) kontextusban is extenzió. Így a $G = E \cup (G \setminus E) = E \cup \{B_k \mid k \in K\}$ és az $E^{II} = E$ összefüggések alapján $\Pi = \{E\} \cup \{B_k \mid k \in K\}$ a (G, M, I) kontextus egy extenzió-partíciója. De akkor az (E, E^I) fogalom a $\mathcal{D}(G, M, I)$ dobozháló egy eleme, E pedig ennek a dobozelemnek az extenziója. $\square$

5.12. Következmény. Az $E = A_z(G)$ halmaz az a legkisebb extenzió a (G, M, I) kontextusban, amelyre egyszerre teljesül a $\{z\} \cup P_z^\cup \subseteq E$ összefüggés és az, hogy $E \setminus \{z\}$ a $\mathcal{D}(H, M, I_H)$ dobozháló egy elemének az extenziója.

5.13. Következmény. A $\{z\}$ halmaz pontosan akkor extenziója a $\mathcal{D}(G, M, I)$ dobozháló egy elemének, ha $\{z\}^{II} = \{z\}$ és $P_z^\cup = \emptyset$.

Bizonyítás. Ha $E = \{z\}$ a $\mathcal{D}(G, M, I)$ dobozháló egy elemének az extenziója, akkor $\{z\}^{II} = \{z\}$, $\{z\} = A_z(G)$ és $P_z^\cup \subseteq H$ miatt 5.12.-ből $P_z^\cup \subseteq E \setminus \{z\} = \emptyset$, vagyis $P_z^\cup = \emptyset$ következik. Fordítva, ha $E = \{z\} = \{z\}^{II}$, akkor az $E \setminus \{z\} = \emptyset$ halmaz a $\mathcal{D}(H, M, I_H)$ dobozháló $(\emptyset, M)$ dobozelemének az extenziója, és mivel $P_z^\cup = \emptyset$, az 5.11. Tétel mindhárom feltétele teljesül. Így $E = \{z\}$ a $\mathcal{D}(G, M, I)$ dobozháló egy elemének az extenziója. $\square$

A következő tétel azt írja le, hogy egy $\mathcal{D}(H, M, I_H)$ -beli dobozelem extenziója (vagy annak z -vel való kiegészítése) milyen feltételek mellett lesz dobozelem-extenzió a $\mathcal{D}(G, M, I)$ hálóban.

5.14. Tétel. Legyen $H = G \setminus \{z\}$ és A a $\mathcal{D}(H, M, I_H)$ dobozháló valamely elemének extenziója. Akkor

- (i) A a $\mathcal{D}(G, M, I)$ dobozháló egy elemének extenziója $\Leftrightarrow A_z(G) \cap A^{II} = \emptyset$.
- (ii) $A \cup \{z\}$ a $\mathcal{D}(G, M, I)$ dobozháló egy elemének extenziója $\Leftrightarrow A_z(G) \cap H \subseteq A$ és $(A \cup \{z\})^{II} = A \cup \{z\}$.

Bizonyítás. (i) Tegyük fel, hogy $(A, A^I) \in \mathcal{D}(G, M, I)$. Ekkor $A^I = A$, és $A \subseteq H$ miatt az A halmaz a Π_0 partíció $A_z(G)$ -től diszjunkt blokkjainak uniójaként írható fel, tehát $A_z(G) \cap A^I = A_z(G) \cap A = \emptyset$. Megfordítva, ha $A_z(G) \cap A^I = \emptyset$, akkor $A^I \subseteq H$, amiből az 5.7. Állítás bizonyításában szereplő gondolatmenet alapján $A^I = A$ következik. Tehát A egy extenzió a (G, M, I) kontextus alapján is. Legyen $x \in A$ egy tetszőleges objektum. Mivel $A_z(G) \cap A = \emptyset$, $x \in H \setminus A_z(G)$, így ismét az említett bizonyításra hivatkozva $A_x(G) = A_x(H)$. Az $A_x(H)$ a legszűkebb x -et tartalmazó dobozelem-extenzió a (H, M, I_H) kontextusban, így $A_x(H) \subseteq A$, de akkor $A_x(G) \subseteq A$ is teljesül, ami a 4.3. Állítás szerint azt jelenti, hogy A a $\mathcal{D}(G, M, I)$ dobozháló valamely elemének az extenziója.

(ii) Tegyük fel, hogy $A \cup \{z\}$ a $\mathcal{D}(G, M, I)$ dobozháló egy elemének extenziója. Akkor $(A \cup \{z\})^I = A \cup \{z\}$ automatikusan teljesül, és az $A \cup \{z\}$ extenzió tartalmazza a minimális $A_z(G)$ extenziót. Így $A_z(G) \cap H \subseteq (A \cup \{z\}) \cap H = A \cap H = A$. A fordított állítás igazolásához először vegyük észre, hogy minden $x \in A$ esetén $A_x(H) \subseteq A$, mert A egy dobozelem-extenzió a $\mathcal{D}(H, M, I_H)$ hálóban. Ezután végezzük el az $A \cup \{z\}$ halmaz következő átalakításait:

$$\begin{aligned} A \cup \{z\} &= (A \setminus A_z(G)) \cup A_z(G) \subseteq \cup \{A_x(H) \mid x \in A \setminus A_z(G)\} \cup A_z(G) \subseteq \\ &\subseteq A \cup A_z(G) = A \cup \{z\} \cup (A_z(G) \cap H) \subseteq A \cup \{z\} \cup A = A \cup \{z\}. \end{aligned}$$

Azt kaptuk, hogy $A \cup \{z\} = \cup \{A_x(H) \mid x \in A \setminus A_z(G)\} \cup A_z(G)$, ahol a jobb-oldali unióban szereplő halmazok mindegyike a Π_0 partíció valamely blokkja az 5.7. Állítás szerint. A 3.2. Definícióban bevezetett atomszerkezet fogalmát használva ezt úgy is fogalmazhatjuk, hogy az $A \cup \{z\}$ halmaz megegyezik az atomszerkezetébe tartozó atomok extenziójának uniójával. Mivel $A \cup \{z\}$ valamely fogalom extenziója is az $(A \cup \{z\})^I = A \cup \{z\}$ feltételből következően, így a 3.4. Tételt alkalmazva kapjuk, hogy $A \cup \{z\}$ egyben dobozelem-extenzió is. $\square$

Felmerül a kérdés, hogy a $\mathcal{D}(G, M, I)$ dobozháló összes elemének extenzióját megkaphatjuk-e a $\mathcal{D}(H, M, I_H)$ háló elemeinek ismeretében az előző tételben megadott lehetőségek szerint. A kérdésre az 5.6. Állításban már pozitív választ adtunk, megmutatva, hogy ha E egy dobozelem extenziója $\mathcal{D}(G, M, I)$ hálóban, akkor vagy a szűkebb kontextus alapján is extenziója valamelyik dobozelemnek, vagy $E = B \cup \{z\}$ alakban kapható meg a $\mathcal{D}(H, M, I_H)$ háló egy (B, B^{I_H}) dobozeleméből.

A fenti eredmények lehetővé teszik, hogy a kontextus elemi bővítése során az eredeti kontextusból származó dobozelemek ismeretében meghatározzuk a bővített kontextushoz tartozó összes dobozelemet. (Pontosabban a dobozelemek extenzióit, de azokból a dobozelemek egyértelműen származtathatóak.) Legyen $H = G \setminus \{z\}$ és jelölje DH azt a halmazt, amely a $\mathcal{D}(H, M, I_H)$

dobozháló elemeinek extenzióból áll, DG pedig legyen ugyanez a halmaz a $\mathcal{D}(G, M, I)$ dobozháló esetén. A bővített kontextushoz tartozó dobozelemek extenzióit meghatározó algoritmus a DH halmaz elemein kívül az $A = A_z(G)$ extenziót használja fel.

5.15. Algoritmus (Dobozelem-extenziók meghatározása elemi bővítés után).

Legyen $DG = \{A\}$

minden $E \in DH$ esetén

ha $E^{II} \cap A = \emptyset$ akkor $DG \leftarrow DG \cup \{E\}$

ha $A \cap H \subseteq E$ akkor $B \leftarrow E \cup \{z\}$

ha $B^{II} = B$ akkor $DG \leftarrow DG \cup \{B\}$.

Vegyük észre, hogy az algoritmus többszöri alkalmazásával egy kezdeti, kisebb objektumszámú kontextusból kiindulva nagyméretű kontextusok dobozhálójának elemeit is előállíthatjuk, úgy, hogy a kontextus objektumhalmazát egyesével bővítjük. Ez a módszer egy hasznos alternatíváját jelenti a 3. fejezetben bemutatott, dobozelemeket meghatározó algoritmusnak, melyben a kontextussal együtt az objektumok száma is rögzített.

6. Alkalmazások

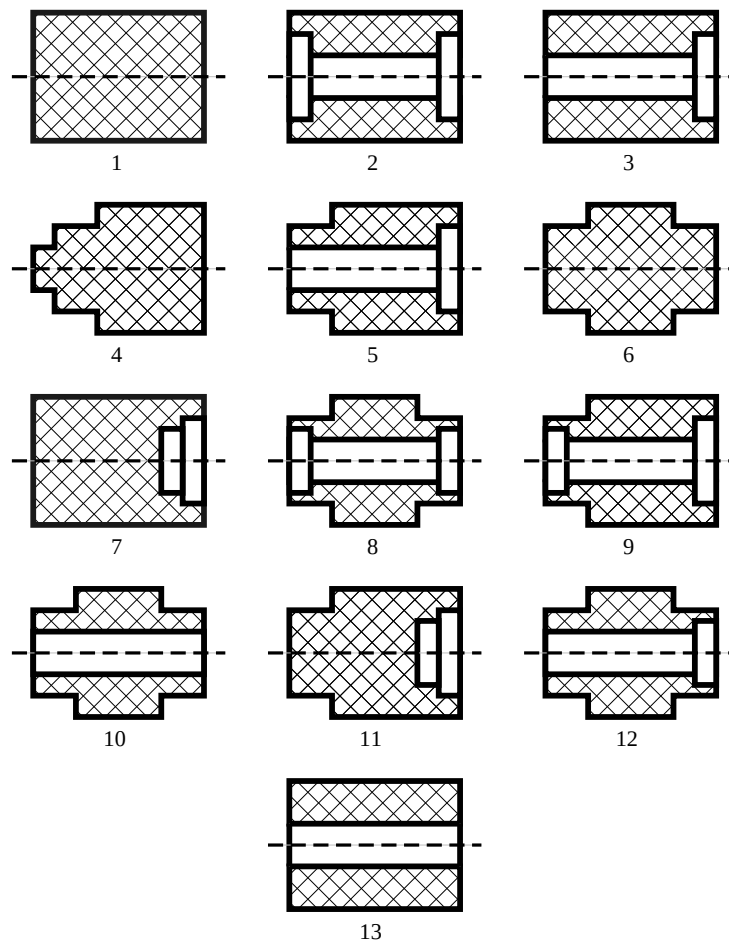
A dolgozat eredményei a Miskolci Egyetem Alkalmazott Informatikai Tanszékének termelésinformatikai témájú projektjeiben közvetlenül hasznosíthatóak. Ebben a fejezetben az előzőekben ismertetett elméleti eredmények alkalmazását mutatjuk be a tanszék kutatási profiljába illeszkedő, csoporttechnológiai jellegű feladatokon keresztül. A csoporttechnológia a diszkrét termelési folyamatok során alkalmazott klasszikus tervezési módszer, melynek fő célja a különböző alkatrészek, gépek, gyártási folyamatok közötti hasonlóságok felismerése, és az így szerzett tudás hasznosítása. A csoporttechnológia elveinek jól átgondolt alkalmazása számos előnnyel jár: a termelés irányítása egyszerűbbé válik, a specializálódott munkahelyek révén magasabb termelékenységi színvonal érhető el, a szerszámozás és készülékezés költségei csökkennek, az anyagmozgatási távolságok rövidülnek. Közelebb jutunk a konstrukciós tervezés és gyártás integrációjához, a szabványosítás egy magasabb formája érhető el, aminek hatására az átfutási idők lerövidíthetőek [54].

A korábbi fejezetekben ismertetett elméleti eredmények hasznosíthatóságát két jellegzetes, egymással is szoros kapcsolatban álló csoporttechnológiai feladat esetén vizsgáljuk meg.

6.1. Alkatrészek osztályozása

Az alkatrészek, munkadarabok osztályozása során nemcsak egyszerűen partitionáljuk az alkatrészhalmazt, hanem egyúttal az is a célunk, hogy alkatrészcsaládok hierarchikus rendszerét hozzuk létre. Az egyes családokba tartozó alkatrészeket úgy válogatjuk össze, hogy azok hasonló geometriai-fizikai jellemzőkkel rendelkezzenek és megmunkálási igényük is közel azonos legyen. Az alkatrészcsaládok strukturálásakor először csak durva kategorizálást végzünk, kevés közös tulajdonsággal is egy családba sorolva alkatrészeket, majd ezt a felbontást finomítjuk, mindaddig, míg olyan csoportokat nem nyerünk, melyeknek elemei ugyanazon berendezésekkel és közelítőleg azonos felszerszámozással gyárthatóak. A legszűkebb osztályok elemei között csak csekély különbségek vannak, melyek a megmunkálás során csak bizonyos paraméterek kismértékű megváltoztatását igénylik (pl. forgácsolási hossz növelése, kisebb átmérőjű homlokmaró alkalmazása).

A következő példában az alkatrészcsaládok kialakítását a 3. fejezetben ismertetett, dobozhálón alapuló osztályfelbontási technikával végezzük. Adottak a 4. ábrán látható, forgástest jellegű alkatrészek, melyek azonos anyagból készülnek és befoglaló méreteik is megegyeznek, megmunkálási igényüket tekintve azonban eltérnek egymástól. (A 4. és a későbbi, 6. ábrán látható alkatrészek rajzai Vadász Dénes kandidátusi értekezéséből származnak [57].)

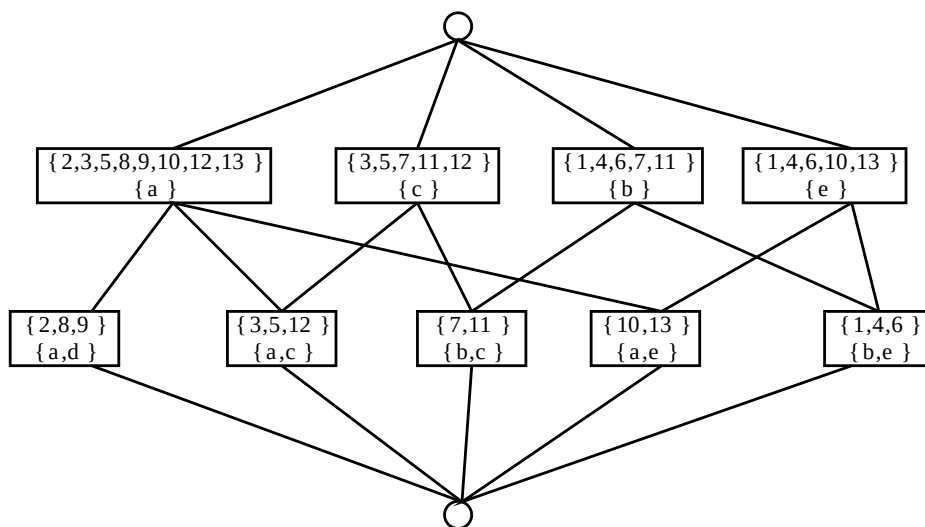


4. ábra. Osztályozandó alkatrészek

Az alkatrészek jellemzésére az alábbi attribútumokat használjuk:

- a) átfúrt
- b) nem átfúrt
- c) belülről lépcsőzött egy irányból
- d) belülről lépcsőzött mindkét irányból
- e) belülről nem lépcsőzött
- f) kívülről lépcsőzött egy irányból
- g) kívülről lépcsőzött mindkét irányból.

Megadva az egyes alkatrészek tulajdonságait, az eredményt egy formális kontextusban ábrázolhatjuk (6. táblázat).



5. ábra. Dobozháló alkatrészek osztályozásához

6. táblázat. Kontextus alkatrészek osztályozásához

	a	b	c	d	e	f	g
1		×			×		
2	×			×			
3	×		×				
4		×			×	×	
5	×		×			×	
6		×			×		×
7		×	×				
8	×			×			×
9	×			×		×	
10	×				×		×
11		×	×			×	
12	×		×				×
13	×				×		

A 3. fejezetben ismertetett módon kiszámítva a kontextushoz tartozó dobozelemeket, az 5. ábrán látható dobozhálót kapjuk. Minden egyes téglalap egy-egy nem-triviális dobozelemet jelöl, a téglalapon belül az első sorban a dobozelem extenzióját, a második sorban az intenzióját tüntettük fel.

A dobozelemek megfelelő csoportosításával nyerjük az osztályozási rendszereket, minden egyes osztályozási rendszer egy extenzió-partíciót indukál az alkatrészek halmazán. A legfinomabb extenzió-partíció a dobozatok

extenzióiból áll:

$$\Pi_0 = \{\{1, 4, 6\}, \{2, 8, 9\}, \{7, 11\}, \{3, 5, 12\}, \{10, 13\}\}.$$

A következő extenzió-partíciók olyan osztályozási rendszerekhez tartoznak, melyek egy kivételével atomokból állnak:

$$\Pi_2 = \{\{2, 3, 5, 8, 9, 10, 12, 13\}, \{1, 4, 6\}, \{7, 11\}\}$$

$$\Pi_3 = \{\{1, 4, 6, 10, 13\}, \{2, 8, 9\}, \{7, 11\}, \{3, 5, 12\}\}$$

$$\Pi_4 = \{\{1, 4, 6, 7, 11\}, \{2, 8, 9\}, \{3, 5, 12\}, \{10, 13\}\}$$

$$\Pi_5 = \{\{3, 5, 7, 11, 12\}, \{1, 4, 6\}, \{2, 8, 9\}, \{10, 13\}\}.$$

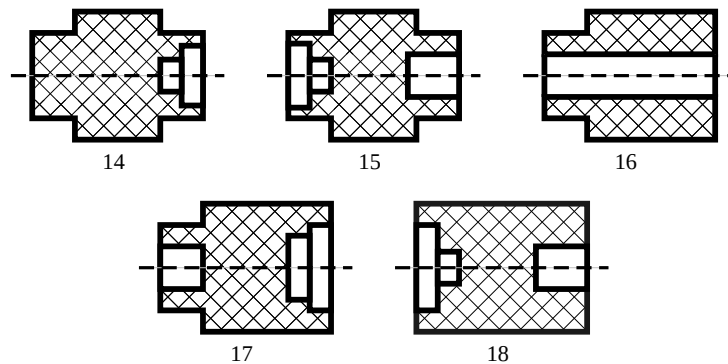
A fentiekén kívül még két felbontást határoz meg a dobozháló, melyek technológiai szempontból nagyon értékes információt hordoznak:

$$\Pi_6 = \{\{(2, 3, 5, 8, 9, 10, 12, 13), \{1, 4, 6, 7, 11\}\}$$

$$\Pi_7 = \{\{3, 5, 7, 11, 12\}, \{2, 8, 9\}, \{1, 4, 6, 10, 13\}\}.$$

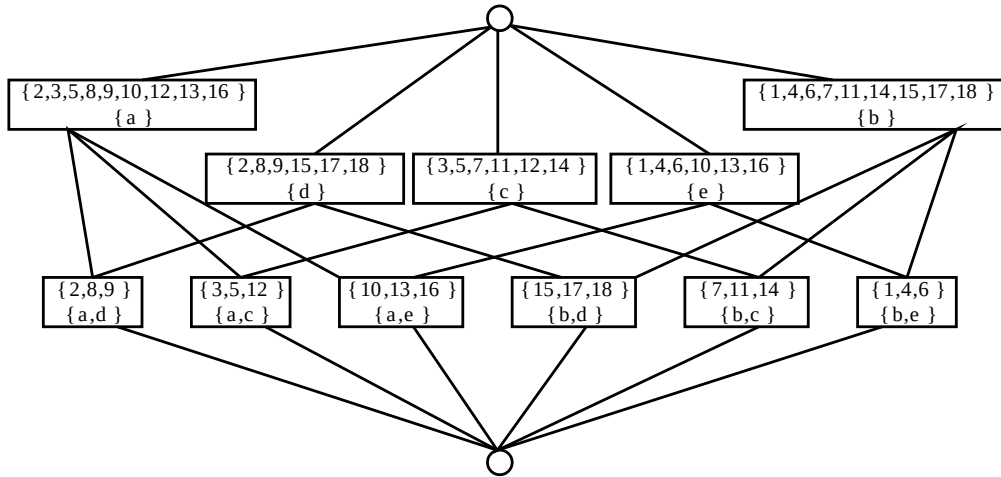
A Π_6 extenzió-partíció aszerint választja szét az alkatrészeket, hogy át vannak-e fúrva teljesen, vagy sem (a vagy b tulajdonság). Π_7 esetében a három csoport a belső lépcsőzöttség alapján alakult ki (c , d vagy e tulajdonság). Megfigyelhetjük, hogy a dobozháló elemeinek intenziói az f és g attribútumokat nem tartalmazzák, ami azt jelenti, hogy az alkatrészek külső lépcsőzöttsége nem lényeges szempont az osztályok kialakítása során.

A továbbiakban az 5. fejezet eredményeit használva megvizsgáljuk a kontextus elemi bővítésének hatását a dobozháló szerkezetére, úgy, hogy a 6. ábrán látható alkatrészeket illesztjük egymás után a kontextus objektumhalmazába.



6. ábra. Új, besorolandó alkatrészek

Minden új objektum esetén először azt kell meghatározni, hogy a bővítéssel kapott dobozháló mely atomjának extenziójába illeszthető az adott alkatrész (ha van ilyen). Például a 14-es jelű alkatrész besorolható egy már létező atom extenziójába (hiszen a 7-es és 11-es alkatrészhöz hasonlóan ez az új elem sincs átfúrva és belülről egy irányból lépcsőzött), míg a 15-ös jelű alkatrész egyik atom extenziójába sem illeszthető be, ezért számára új blokkot kell létrehozni a legfinomabb extenzió-partícióban. A következő lépés a bővítéssel kapott dobozháló elemeinek előállítása. Ehhez az 5.15. Algoritmust használjuk, mely az új alkatrészt tartalmazó atom-extenzió és a korábbi dobozelemek kapcsolata alapján megadja az új dobozelemeket. Ha a kontextust egymás után mind az öt alkatrésszel bővítjük, végül a 7. ábrán látható dobozhálót kapjuk. A háló szembetűnő szabályossága annak köszönhető, hogy a vizsgált 18 alkatrésszel minden olyan objektumot megadtunk, mely az adott attribútumok mellett a kontextusban szerepelhet.



7. ábra. Dobozháló a bővített kontextushoz

Végül megvizsgáljuk a 4. fejezetben bevezetett kitüntetett tulajdonságok szerepét az osztályfelbontások létrehozásában. A 4.6. Megjegyzés *b)* pontja értelmében azok a p tulajdonságok kitüntetettek, melyekre p' valamely dobozelemnek az extenziója. A 7. ábrán látható dobozelemek közül 5 olyan van, melynek egyetlen attribútumból áll az intenziója, tehát az a, b, c, d, e attribútumok mindegyike egy-egy kitüntetett tulajdonság. Az osztályfelbontások szempontjából csak ezek a tulajdonságok lényegesek, ahogy azt már a szűkebb, 13 alkatrészből álló halmaz extenzió-partícióinak megadásakor megfigyeltük. Ha a teljes, 18×7 -es méretű K kontextusra végrehajtjuk a 4. fejezetben leírt konstrukciót, eredményül egy olyan K^{**} kontextust kapunk, melyben a 18 alkatrészt csak az említett 5 tulajdonság jellemzi. A K^{**} kon-

textushoz ugyanaz a dobozháló tartozik, mint az eredetihez, tehát az alkatrészhalmoz keresett felbontásai egy attribútum-oldalról redukált kontextusból kiindulva is meghatározhatóak. Az osztályfelbontás szempontjából lényeges tulajdonságok ismerete és a kontextus redukciójának lehetősége különösen fontos lehet az olyan esetekben, amikor az objektumok jellemzésére nagyszámú attribútumot használunk.

6.2. Gyártócellák kialakítása

Az alkatrészcsaládok meghatározása mellett a csoporttechnológia másik alapproblémája az egyes családok elemeit megmunkáló gépek csoportjainak kijelölése, azaz a gyártócellák megtervezésének, kialakításának feladata. Ez a feladat végezhető az alkatrészcsaládok beazonosítása után, de gyakran egyidejűleg határozzuk meg a hasonló technológiával készülő alkatrészek osztályait és az egyes osztályok elemeit megmunkáló gépek csoportjait. A fő cél a gépek és az alkatrészek halmazának olyan felbontása, hogy az alkatrészek cellák közötti mozgatása minimális mértékű legyen és az egyes cellákon belül a gépeket a lehető legjobban kihasználjuk.

6.2.1. A feladat modellezése gép-alkatrész incidencia mátrixszal

A gyártócellák kialakításának folyamata a legtöbb megoldási módszer esetében a gép-alkatrész incidencia mátrixból indul ki. Ha a vizsgált gyártórendszerben n darab gépen k darab alkatrész megmunkálása folyik, akkor viszonyukat az $n \times k$ típusú, $A = [a_{ij}]$ incidencia mátrix segítségével írjuk le, ahol

$$a_{ij} = \begin{cases} 1, & \text{ha az } i\text{-dik gép megmunkálja a } j\text{-dik alkatrészt,} \\ 0 & \text{egyébként.} \end{cases}$$

A következő példában szereplő incidencia mátrix 7 gép és 8 alkatrész viszonyát szemlélteti. A jelölés egyszerűsítése érdekében a gépek azonosítására számokat használunk, az alkatrészeket pedig kisbetűkkel jelöljük.

		Alkatrészek							
		a	b	c	d	e	f	g	h
G	1	0	1	1	0	1	0	0	0
	2	1	0	0	0	0	1	0	0
	3	0	0	0	1	0	0	1	0
	4	1	0	0	0	0	1	0	0
	5	0	0	1	0	1	0	0	1
	6	0	0	0	1	0	0	0	0
	7	0	1	1	0	1	0	0	1

A sorok és oszlopok alkalmas átrendezésével egy olyan mátrixot nyerhetünk melyben három, egymástól jól elkülöníthető részmátrix jelöli ki a gépek és alkatrészek keresett csoportjait:

		Alkatrészek							
		<i>b</i>	<i>c</i>	<i>e</i>	<i>h</i>	<i>a</i>	<i>f</i>	<i>d</i>	<i>g</i>
G	1	1	1	1	0	0	0	0	0
	5	0	1	1	1	0	0	0	0
	7	1	1	1	1	0	0	0	0
	2	0	0	0	0	1	1	0	0
	4	0	0	0	0	1	1	0	0
	3	0	0	0	0	0	0	1	1
	6	0	0	0	0	0	0	1	0

Az átrendezés után kapott mátrixból a következő megoldást olvashatjuk le: az $\{1, 5, 7\}$ jelű gépekből álló gyártócella munkálja meg a $\{b, c, e, h\}$ alkatrészekből álló családot, a $\{2, 4\}$ gyártócella gépei dolgoznak az $\{a, f\}$ alkatrészcsoport elemein, míg a $\{3, 6\}$ gyártócellához a $\{d, g\}$ alkatrészhalmaz tartozik. A valóságban általában nem ilyen egyszerű a feladat megoldása, hiszen jóval nagyobb méretű mátrixok reprezentálják a gyártási rendszereket, így nem könnyű hatékony algoritmust találni a sorok és oszlopok alkalmas átrendezésére. Másrészt a diszjunkt részmátrixokra való szeparálhatóság is csak ideális esetben fordul elő, a gyakorlati példák majd mindegyikében találkozunk a *szűk keresztmetszetek* problémájával. Szűk keresztmetszet kétféleképpen is adódhat: az egyik esetben van olyan alkatrész, melyet különböző gyártócellákba tartozó gépeknek is meg kell munkálniuk (bottleneck part), vagy van olyan gép, mely több alkatrészcsoport elemein is dolgozik (bottleneck machine). Az első esetben a szűk keresztmetszet feloldása úgy történik, hogy a kérdéses alkatrészt átszállítjuk a további megmunkálását végző gyártócellába, vagy módosíthatjuk úgy az adott alkatrész terveit, hogy funkcióját megtartva gyártása beilleszthető legyen egy már létező cellába. Ha ez nem lehetséges, az alkatrész mozgatása pedig költséges, akkor végső esetben gyártását alvállalkozóra is bízhatjuk. Ha valamelyik gép miatt lép fel szűk keresztmetszet, meg kell fontolni egy másik, ugyanolyan gép beszerzését, hogy elérhessük a cellák függetlenségét, vagy ki kell alakítanunk egy olyan optimális elrendezést, melyben az érintett alkatrészcsoportok egyaránt használhatják a kérdéses gépet.

A gyártócellák és alkatrészcsoportok meghatározására alkalmas módszerek gyakran egymástól különböző megoldásokra vezetnek. A feladat optimális megoldása csak egyes speciális esetekben értelmezhető egyértelműen, például, ha minden egyes alkatrészt csak a családjához rendelt cella gépei munkálnak meg, de azok közül mindegyik részt vesz a megmunkálásban. Általában a

megoldások minőségét egy vagy több mutató értéke alapján értékeljük, így az adott mutatókban figyelembe vett szempontok alapján beszélhetünk csak a megoldás optimalitásáról. Ha a gépek és alkatrészek összetartozó csoportjait meghatároztuk, a megoldás minősítésére a legtöbb mutató az alábbi paramétereket veszi figyelembe:

- n_1 : a műveletek száma (azaz az 1-esek száma az incidencia mátrixban),
- n_0 : kihasználatlan gépek száma (azaz a 0-k száma a kialakított csoportoknak megfelelő részmátrixokon belül),
- n_k : kivételes alkatrészek száma (azaz az 1-esek száma a kialakított csoportoknak megfelelő részmátrixokon kívül).

A megoldások minőségének legegyszerűbb jellemzését a kivételes alkatrészek aránya adja:

$$\Omega = \frac{n_k}{n_1}.$$

A gépek kihasználtságának mértékét fejezi ki az

$$\Lambda = \frac{n_1 - n_k}{\sum m_k p_k}$$

tört, ahol m_k a gépek, p_k az alkatrészek száma a k -dik csoportban, és az összegzést az összes csoportra el kell végezni [9]. Az irodalomban további mérőszámok alkalmazására is találunk példát: grouping efficiency [7], grouping index [48], grouping capability index [47]. A leggyakrabban és legáltalánosabban használt mutató a megoldások értékelésére a *csoportképzési hatásfok* (grouping efficacy), melyet Kumar és Chandrasekharan vezetett be [32]-ben a következő képlettel:

$$\Gamma = 1 - \frac{n_0 + n_k}{n_0 + n_1} = \frac{n_1 - n_k}{n_1 + n_0}.$$

A továbbiakban a Γ változó értéke alapján választjuk ki a feladat megoldásai közül a „legjobbat”. A fejezet elején szereplő példában $n_1 = 17$, $n_0 = 3$, $n_k = 0$, így a csoportképzési hatásfok: $\Gamma = \frac{17}{20}$. Ideális esetben kihasználatlan gépek és kivételes alkatrészek sincsenek, így a Γ mutató maximális értéke 1. Általában elmondhatjuk, hogy minél nagyobb a csoportképzési hatásfok, annál jobb az adott gyártócella és alkatrészcsalád-kiosztást megvalósító megoldás. A hatásfok szerepét azonban nem szabad túlértékelnünk, mindig a konkrét feladat dönti el, hogy mely tényezőket kell hangsúlyosabban figyelembe vennünk [46].

6.2.2. A gyártócellák meghatározásának különböző módszerei

Az elmúlt évtizedekben a gyártócellák meghatározására az elméleti szakemberek számos megoldási technikát dolgoztak ki, sőt, több publikációnak is témája a különböző módszerek rendszerbe foglalása és e módszerek összehasonlító elemzése [11], [51], [58]. S. M. Shafer szerint a megoldási technikákat úgy célszerű csoportosítani, hogy az alapjukat jelentő elméleti háttérrel vesszük figyelembe. Javaslatára alapján a feladat megoldására használható módszereket öt nagy csoportba sorolhatjuk [50]. Ezek a kategóriák a következők:

manuális módszerek,

statisztikai klaszter-analízis,

szortírozó algoritmusok a gép-alkatrész mátrixra,

mesterséges intelligencia módszerek,

matematikai módszerek.

A *manuális módszerek* csoportjába azok az eljárások tartoznak, amelyekben az alkatrészcsaládok és gyártócellák kialakítása egy vagy több szakértő sorozatos döntéseinek eredményeképpen valósul meg. A legismertebb ilyen módszert J. L. Burbidge dolgozta ki az 1970-es években, Production Flow Analysis néven [5]. A szinte folyamatos szakértői jelenlét következtében a módszer rendkívül rugalmas, hátránya azonban, hogy számítógépes támogatása legtöbbször nehézségekbe ütközik. Másrészt a szakértő tudása és tapasztalata is elégtelennek bizonyulhat, ha egy üzemben több ezer, bonyolult felületű alkatrész megmunkálását végzik.

A *statisztikai klaszter-analízis* alkalmazásához először egy mérőszámot kell definiálnunk, mely kifejezi két alkatrész vagy két gép hasonlóságát (vagy éppen az eltérőségét, távolságát). Az egyik legáltalánosabban használt ilyen mérőszám a Jaccard-féle hasonlósági együttható, mely a következő képlettel adható meg:

$$J_{ik} = \frac{N_{ik}}{N_i + N_k - N_{ik}},$$

ahol J_{ik} az i és k alkatrész között definiált Jaccard-féle hasonlósági együttható, N_{ik} azon gépek száma, melyek az i és k alkatrészt egyaránt megmunkálják, N_i illetve N_k pedig az i , illetve k alkatrész megmunkálását végző gépek száma. A hasonlósági együtthatót minden egyes objektumpárra kiszámítjuk és az eredményeket egy mátrixban, a hasonlósági mátrixban tároljuk. A hasonlósági mátrix értékeiből kiindulva határozzuk meg egy klaszterező algoritmussal a

keresett alkatrészcsaládok illetve gyártócellák elemeit. A hasonlósági együtt-hatók definiálására számtalan lehetőség kínálkozik, attól függően, hogy a csoportok képzésével egyidejűleg milyen feltételeknek kell teljesülnie. A téma alapos áttekintését nyújtja Yin és Yasuda [56] cikke, mely bemutatja, rendszerbe foglalja és értékeli az irodalomban javasolt hasonlósági és különbözőségi együtt-hatókat. Érdeemes megemlíteni, hogy a hasonlósági együtt-hatók alkalmazására más, főleg mesterséges intelligenciabeli módszerekkel kombinálva is találunk példát: Jeon és Leep genetikusan [27], Adenso-Díaz pedig tabu keresést alkalmazva oldotta meg a feladatot [1].

A *gép-alkatrész incidencia mátrix szortírozó algoritmusain* olyan eljárásokat értünk, amelyek a mátrix sorainak és oszlopainak alkalmas átrendezésével az alkatrészcsaládok és gyártócellák meghatározását egyidejűleg végzik. A folyamat végén a mátrix főátlója mentén elkülönülő blokkokból olvasható le a megoldás, ezért az ilyen módszereket blokk-diagonális módszereknek is nevezik. A két legnevezetesebb blokk-diagonális módszer a King által kifejlesztett ROC (rank order clustering) algoritmus [30] és a Chan és Milner nevéhez fűződő DCA (direct clustering algorithm) [6].

A számítástechnikai kapacitás és teljesítmény növekedésével párhuzamosan a kutatók egyre gyakrabban alkalmazzák a *mesterséges intelligencia* körébe tartozó módszereket a csoporttechnológiai problémák megoldásában. A gyártócellák és alkatrészcsaládok meghatározásának feladata nagyméretű incidencia mátrix esetén kombinatorikus robbanást eredményezhet, mellyel heurisztikus algoritmusok alkalmazásával könnyebben megbirkózhatunk, mint egzakt megoldási módszerekkel. A két leggyakoribb megoldási technika a neurális hálók és a genetikusan [30] és a Chan és Milner nevéhez fűződő DCA (direct clustering algorithm) [6]. A számítástechnikai kapacitás és teljesítmény növekedésével párhuzamosan a kutatók egyre gyakrabban alkalmazzák a *mesterséges intelligencia* körébe tartozó módszereket a csoporttechnológiai problémák megoldásában. A gyártócellák és alkatrészcsaládok meghatározásának feladata nagyméretű incidencia mátrix esetén kombinatorikus robbanást eredményezhet, mellyel heurisztikus algoritmusok alkalmazásával könnyebben megbirkózhatunk, mint egzakt megoldási módszerekkel. A két leggyakoribb megoldási technika a neurális hálók és a genetikusan [30] és a Chan és Milner nevéhez fűződő DCA (direct clustering algorithm) [6]. A számítástechnikai kapacitás és teljesítmény növekedésével párhuzamosan a kutatók egyre gyakrabban alkalmazzák a *mesterséges intelligencia* körébe tartozó módszereket a csoporttechnológiai problémák megoldásában. A gyártócellák és alkatrészcsaládok meghatározásának feladata nagyméretű incidencia mátrix esetén kombinatorikus robbanást eredményezhet, mellyel heurisztikus algoritmusok alkalmazásával könnyebben megbirkózhatunk, mint egzakt megoldási módszerekkel. A két leggyakoribb megoldási technika a neurális hálók és a genetikusan [30] és a Chan és Milner nevéhez fűződő DCA (direct clustering algorithm) [6].

A gyártócellák és alkatrészcsaládok meghatározására természetesen *matematikai módszerek* is alkalmazhatóak. A leggyakrabban alkalmazott módszerek elméleti hátterében a gráfelmélet, kombinatorikus analízis és a lineáris

programozás áll. Ez utóbbi területen Kusiak p -medián modellje volt az első olyan módszer, amely alkatrészcsaládok meghatározására szolgált [34]. A p érték a létrehozandó alkatrészcsaládok számára utal, a célfüggvény pedig az egy családon belüli alkatrészek maximális hasonlóságára irányul, ahol a hasonlóságot két alkatrész között azon gépek számával adjuk meg, amelyek mindkét alkatrészt megmunkálják. A módszer újdonsága abban rejlett, hogy az alkatrészek közötti hasonlóság alapján létrehozott hasonlósági mátrix egy matematikai programozási feladat bemenő adataként szolgált. Az elmúlt húsz évben a kutatók a kezdeti modell számos olyan variánsát is kifejlesztették, melyek gyártócellák meghatározására is alkalmasak. A lineáris programozáson alapuló módszerek alapjaikban kevésbé, a célfüggvényeket illetően azonban jelentősen eltérhetnek egymástól. A célfüggvényekben leggyakrabban megfogalmazott követelmények az egy csoportba kerülő alkatrészek illetve gépek maximális hasonlósága mellett a gyártócellák kihasználásának maximalizálása, illetve a gépbeállítási költségek, a cellaközi szállítások, a szűk keresztmetszetek megmunkálási költségének minimalizálása. A lineáris programozás alkalmazása a változatos célfüggvények ellenére sem terjedt el a gyakorlatban, hiszen az eljárás bonyolult matematikai apparátust igényel és a nagyméretű feladatok megoldása még a mai számítógépes kapacitás mellett is nehézségekbe ütközik.

6.2.3. Gyártócellák meghatározása dobozelemek segítségével

A 3. fejezetben bemutatott, dobozhálón alapuló osztályfelbontási módszer segítségével is meghatározhatjuk egy adott gép-alkatrész incidencia mátrixból kiindulva az alkatrészek és gépek egy csoportba tartozó elemeit. Ez a módszer, absztrakt algebrai háttérét figyelembe véve, egy újabb matematikai technikának tekinthető. A gép-alkatrész incidencia mátrix megfeleltethető a kontextust definiáló táblázatnak, objektumoknak pedig válasszuk a gépeket, ha a gyártócellák meghatározása az elsődleges célunk, ha pedig az alkatrészcsaládok meghatározása a fő feladat, akkor az alkatrészek legyenek az objektumok, és az őket megmunkáló gépek jelentsék a tulajdonságokat. Példaként tekintsük a 6.2.1. alfejezetben szereplő 7×8 -as méretű incidencia mátrixot, amelynek megfelelő K kontextust a 7. táblázat mutatja be.

A K kontextushoz tartozó extenzió-partíciók meghatározásához először a dobozháló atomjait keressük meg. Algoritmusunk első lépésével az atomok extenzióit állítjuk elő, melyek a következők: $\{1, 5, 7\}$, $\{2, 4\}$, $\{3, 6\}$ (az objektumokra sorszámaikkal hivatkoztunk). Ez a legfinomabb extenzió-partíciója az objektumhalmaznak, és nincs is több nem-triviális osztályfelbontás, mert az atom-extenziók uniói már nem adnak újabb extenziókat. Ez az extenzió-partíció határozza meg a gyártócellákat, melyek megegyeznek a fejezet elején megadott megoldással. Az alkatrészcsaládok meghatározása ezek után

7. táblázat. Kontextus létrehozása incidencia mátrixból

K	Attribútumok (M)								
Objektumok (G)		a	b	c	d	e	f	g	h
	1		x	x		x			
	2	x					x		
	3				x			x	
	4	x					x		
	5			x		x			x
	6				x				
	7		x	x		x			x

úgy történik, hogy az azonos gyártócellába kerülő gépek által megmunkált alkatrészeket egy csoportba soroljuk, így kapjuk a $\{b, c, e, h\}$, $\{a, f\}$, $\{d, g\}$ halmazokból álló alkatrészcsaládokat. Ugyanezt a megoldást közvetlenül is megkapjuk, ha a kontextus objektumrészének az alkatrészek halmazát feleltetjük meg, és az így adódó 8×7 -es mátrixból kiindulva határozzuk meg az extenzió-partíciókat.

Az előző példában a dobozelemes módszer alkalmazásával nyert felbontások közül nem volt nehéz kiválasztani a legjobbat, hiszen csak egyetlen, nem-triviális extenzió-partíció adódott. Megmutatjuk, hogy bonyolultabb esetben is elkerülhetjük az az összes felbontás vizsgálatát, gyakran elegendő az extenzió-partíciók egy szűkebb halmazából kiválasztani a megoldást. Ha feltételezzük, hogy a keresett megoldás a gépek G halmazának egy olyan $\{B_j \mid j \in J\}$ partíciója, melynek blokkjaira $B'_i \cap B'_j = \emptyset$ teljesül minden $i \neq j$ esetén, akkor ezzel azt írjuk elő, hogy nincsenek olyan alkatrészek, amelyeket különböző gyártócellák gépeinek mindegyike meg kell, hogy munkáljon. A feltételből azonban a fogalmak szuprémumképzési szabálya alapján levezethető a

$$(B_i, B_i)' \vee (B_j, B_j)' = ((B_i \cup B_i)'', B'_i \cap B'_j) = (G, \emptyset) = \mathbf{1}$$

összefüggés. Ez azt jelenti, hogy elegendő a „maximális” osztályozási rendszerekkel foglalkozni, amelyekben bármely két különböző elem szuprémuma a háló legnagyobb elemével egyezik meg. A maximális osztályozási rendszerek által indukált extenzió-partíciók megadásával kijelöltük a gyártócellákat, azokhoz azonban még hozzá kell rendelnünk az alkatrészek megfelelő csoportjait. Világos, hogy egy adott cella gépeiből álló objektumhalmaznak az intenziója biztosan része a megfelelő alkatrészcsaládnak, hiszen az intenzióba tartozó alkatrészeket a cella minden gépének meg kell munkálnia. A kimaradt alkatrészeket pedig abba a családba soroljuk be, amelyhez tartozó cella

esetén a vizsgált alkatrészt megmunkáló gépek aránya maximális. Ezzel az eljárással az a célunk, hogy egy adott gyártócellakiosztáshoz hozzárendeljük azt az alkatrészcsalád-rendszert, melyre a megoldás csoportképzési hatásfoka a lehető legnagyobb. Jelölje B_j ($j \in J$) a j -dik gyártócella gépeinek halmazát, A_j a B_j cellához tartozó alkatrészcsaládot és legyen A az alkatrészek halmaza. Egy $a \in A$ alkatrész esetén jelölje $h_j(a)$ az a alkatrészt megmunkáló gépek arányát a j -dik gyártócellán belül:

$$h_j(a) = \frac{\text{a } B_j \text{ cella gépei közül az } a \text{ alkatrészt megmunkálók száma}}{\text{a } B_j \text{ cella gépeinek száma}}.$$

A következő algoritmus egy adott gyártócellakiosztás B_j celláihoz meghatározza a cellába tartozó gépek által megmunkált A_j alkatrészcsalád elemeit.

6.1. Algoritmus (Alkatrészcsaládok meghatározása).

Minden B_j esetén legyen $A_j = B'_j$

minden $a \in A \setminus \bigcup_{j \in J} A_j$ esetén

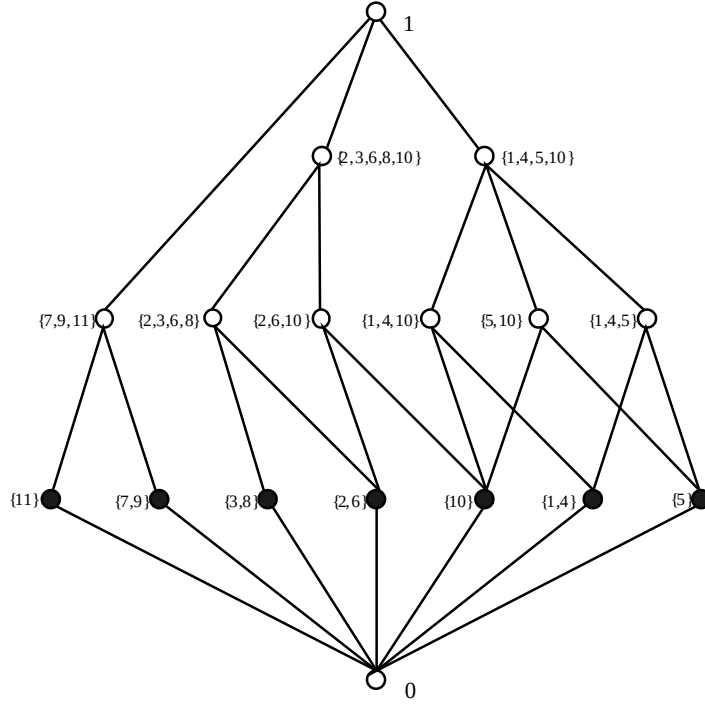
legyen k a legkisebb olyan index, melyre a $h_j(a)$ mennyiség maximális

legyen $A_k = A_k \cup a$.

Az elmondottak illusztrálására tekintsünk egy nagyobb méretű incidenciamátrixot, melyet Cheng a gyártócellák meghatározására irányuló, 12 különböző módszer összehasonlító elemzésére használt fel [10]. A szerző értékelése szerint a két legmegbízhatóbb megoldást Chandrasekharan és Rajagopalan ZODIAC nevű algoritmus, illetve a Seifoddini és Wolf által kidolgozott, hasonlósági együtthatókon alapuló ALC (average linkage clustering) algoritmus szolgáltatja (a módszerek leírása a [8], illetve a [49] cikkekben található).

		Alkatrészek																					
		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>	<i>m</i>	<i>n</i>	<i>o</i>	<i>p</i>	<i>q</i>	<i>r</i>	<i>s</i>	<i>t</i>	<i>u</i>	<i>v</i>
G é p e k	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	1	1
	2	0	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
	3	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0
	4	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	1	1
	5	1	1	1	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0	1	1	1
	6	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0
	7	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0	0
	8	0	0	0	0	1	1	0	0	1	1	0	1	1	1	0	0	1	1	1	0	0	0
	9	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0	0
	10	1	1	1	0	1	0	1	1	0	0	1	1	0	0	1	1	0	1	0	0	0	0
	11	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0

A fenti mátrixnak megfelelő kontextus egy olyan fogalomhálót határoz meg, melyből egy 17 elemű dobozháló származtatható. Ennek a hálónak 7 atomja és 10 további eleme van, diagramja a 8. ábrán látható, melyen a dobozelemeket extenzióik azonosítják, az atomokat sötét karikák jelzik.



8. ábra. Dobozháló diagramja a Cheng tanulmányában szereplő kontextushoz

Az osztályozási rendszerek meghatározására szolgáló algoritmussal 29 különböző megoldást kaptunk. Ezek közül csak kettő bizonyult maximálisnak, a hozzájuk tartozó osztályozásokat az alábbi blokkok alkotják:

$$\Pi_1 = \{\{1, 4, 5, 10\}, \{7, 9, 11\}, \{2, 3, 6, 8\}\},$$

$$\Pi_2 = \{\{1, 4, 5\}, \{7, 9, 11\}, \{2, 3, 6, 8, 10\}\}.$$

Mindkét esetben lefuttatva 6.1. Algoritmust, Π_1 esetére adódott nagyobb csoportképzési hatások, így a módszerünk alkalmazásával kapott megoldás az $\{1,4,5,10\}$, $\{7,9,11\}$, $\{2,3,6,8\}$ gyártócellákból és a hozzájuk rendelt $\{a,b,c,g,k,o,p,t,u,v\}$, $\{d,i,j,n,q,r\}$, $\{e,f,h,l,m,s\}$ alkatrészcsoportokból áll. A kivételes alkatrészek száma: $n_k = 11$, a kihasználatlan gépek száma: $n_0 = 15$. A következő táblázatban a dobozelemek segítségével kapott megoldás mutatóit hasonlítjuk össze a Cheng által legjobbnak ítélt megoldások paramétereivel.

Módszer	n_1	n_0	n_k	Ω	Λ	Γ
Dobozelemek	78	15	11	0.14	0.82	0.72
ALC	78	16	11	0.14	0.81	0.71
ZODIAC	78	15	10	0.13	0.82	0.73

A vizsgált mutatók alapján nincs lényeges minőségi különbség a három megoldás között (Ω esetén a kisebb, Λ és Γ esetén a nagyobb érték számít jobbnak az értékeléskor). A módszerek összehasonlításánál vegyük figyelembe azt is, hogy míg a dobozelemeken alapuló technika egy általános célú osztályozási módszer, addig a Cheng cikkében szereplő algoritmusokat kifejezetten a gyártócellák meghatározására fejlesztették ki.

A 3.6. alfejezetben az osztályozási rendszereket meghatározó algoritmusok hatékonyságát elemezve, arra a következtetésre jutottunk, hogy legrosszabb esetben a fogalomháló elemszáma exponenciális is lehet a kontextus méretének függvényében, sőt elméletileg ez a dobozháló esetén is fennáll. A műszaki gyakorlatban felmerülő osztályozási problémák esetén szerencsére más a helyzet, a kiindulási kontextusból viszonylag kevés dobozelemet lehet generálni. A 8. táblázatban néhány gép-alkatrész incidencia-mátrix esetére meghatároztuk a mátrixhoz, mint kontextushoz tartozó fogalomháló, dobozháló, osztályozási háló elemszámát, illetve az elemek előállításához szükséges futási időt másodpercben mérve (1.73 GHz-es processzorral rendelkező személyi számítógépen futtatva a MATLAB programot). A táblázatban szereplő példamátrixok gyakran idézett, gyártócellák kialakításával foglalkozó publikációkból származnak.

8. táblázat. Hálók elemszáma és az elemek előállításához szükséges idő (s)

A mátrix mérete, forrása	fogalmak száma/futási idő	dobozelemek száma/futási idő	oszt. rendszerek száma/futási idő
11×22 , [10]	28/0.42	17/0.3	29/0.59
20×20 , [39]	34/0.37	11/0.12	5/0.22
40×24 , [7]	39/0.66	13/0.22	3/0.29
27×27 , [13]	336/45.80	35/2.31	188/49.48

A vizsgált mátrixok esetén a dobozháló elemeinek előállítása a fogalmak generálásánál is kevesebb időt vett igénybe. (A fogalmakat csak emiatt az összehasonlítás miatt szerepeltetjük a táblázatban, hiszen a dobozelemek előállításához nincs szükség rájuk.) A legköltségesebb művelet az osztályozási rendszerek meghatározása, ami nagy elemszámú dobozháló esetén időigényes feladat. Programunk az összes lehetséges osztályozási rendszert előállítja, világos, hogy ezek között sok olyan megoldás is van, ami a műszaki probléma szempontjából érdektelen. A csoporttechnológiához köthető osztályozási

feladatokban a hasznos megoldások egy osztályozási fának nevezett struktúrába rendezhetők, melyből a megfelelő osztályozási rendszer kiválasztása a műszaki szakember feladata. Az osztályozási fák elméletével kapcsolatban kutatócsoportunk a „Csoporttechnológia-alapú tervezés és ütemezés támogatása diszkrét matematikai modellekkel és módszerekkel” c. OTKA pályázat keretein belül ért el kezdeti eredményeket.

A feladat komplexitása miatt az egzakt megoldási módszerek nagy számításigénnyel járnak, ezért a felsoroltaknál nagyobb méretű mátrixok ritkán szerepelnek az irodalmi hivatkozásokban. Ha egy gyártócellakialakítási problémában 100-nál több gép vagy alkatrész szerepel, akkor általában valamilyen heurisztikus módszerrel keresik a megoldást (jellemzően genetikus algoritmust vagy neurális hálót használnak). Ez azonban nem jelenti azt, hogy az egzakt módszerek (mint például a dobozhálós technika) algoritmusait nem érdemes fejleszteni, hiszen a számítási teljesítmény növekedésével az ilyen módszerek alkalmazásának lehetőségei is bővülnek és egy egzakt módszerrel kapott megoldás minősége legtöbbször felülmúlja a heurisztikus meggondolásokkal nyert megoldásokét.

7. Új tudományos eredmények

Dobozhálók elemeinek meghatározása

A kontextus objektumhalmazának felbontását (az extenzió-partíciók meghatározását) a kontextusból származtatható dobozháló segítségével végezzük. A dobozháló elemei egyúttal fogalmak is, így azok a kontextushoz tartozó fogalomháló elemei közül is kiválogathatóak lennének. A dobozháló atomjainak ismeretében ezt úgy oldhatjuk meg, hogy az atomisztikus tulajdonságot kihasználva megkeressük azokat a fogalmakat, amelyek előállnak atomok szuprémumaként. Nagyméretű kontextus esetén azonban a fogalomháló elemeinek száma is igen nagy lehet, ezért szükség volt egy olyan módszer kidolgozására, mely a dobozelemeket a fogalomháló elemeinek előállítása nélkül határozza meg.

1. tétel:

Megvizsgálva a dobozháló elemei és a kontextus közötti viszonyt, tételeket fogalmaztam meg és igazoltam a dobozelemek megkonstruálásával kapcsolatban. Az eredmények lehetővé teszik a dobozháló, illetve az osztályozási rendszerek elemeinek előállítását közvetlenül a fogalomháló kontextusából.

Az osztályozási rendszerek meghatározása három fázisra bontható. Az első lépésben a dobozháló atomjait (azaz a minimális osztályozási rendszer elemeit) állítjuk elő, ehhez megadtam egy algoritmust, és formálisan levezettem a helyességét. A további dobozelemek meghatározásának érdekében bevezettem az **atomszerkezet** fogalmát: ez egy H objektumhalmaz esetén azoknak az atomoknak a halmaza, melyeknek extenziói teljes egészében benne vannak H -ban. Az atomszerkezet fogalmát felhasználva szükséges és elégséges feltételt adtam arra nézve, hogy egy fogalom-extenzió mikor dobozelem-extenzió is egyben. Ezt követően egy, a fogalom-extenziók előállítására alkalmas ismert módszert úgy módosítottam, hogy az összes fogalom-extenzió helyett csak azokat kapjuk meg, amelyek dobozelem-extenziók lehetnek. A módosított eljárás segítségével a fogalmak egy részének előállítását elkerülhetjük (például elhagyhatóak azok, amelyek atomszerkezete 0 vagy 1 elemből áll, az azonos atomszerkezetű extenziók közül pedig elegendő csak egyet megtartani). Emellett igazoltam, hogy a redukció ellenére az eljárás a dobozelemek mindegyikét előállítja. Végül, az osztályozási rendszerek meghatározásának fázisában is hasznosnak bizonyult az atomszerkezet fogalma: igazoltam, hogy dobozelemek egy rendszere pontosan akkor osztályozási rendszer, ha az elemek extenzióinak atomszerkezetei az atomok halmazát partícionálják.

Az említett tételeket és bizonyításait értékezés 3. fejezetének első három

alfejezete tartalmazza, illetve a [K6], [K10], [K11], [K12] publikációk is ezekre az eredményekre épülnek.

Adott kontextushoz tartozó fogalomháló és dobozháló egyenlősége

A dobozháló a fogalomháló elemeiből álló atomisztikus háló, melynek előállításakor a fogalomháló bizonyos elemeit figyelmen kívül hagyhatjuk. Felmerül a kérdés, van-e olyan eset, amikor az osztályfelbontás szempontjából minden egyes fogalom fontos, azaz a kontextushoz tartozó fogalomháló és dobozháló megegyezik. A válasz pozitív, ha a fogalomháló már eleve atomisztikus háló: ebben az esetben minden fogalom dobozelem is egyben. Ennek ellenőrzése helyett másik feltételt kerestem a kérdés eldöntésére.

2. tézis:

Egy adott kontextushoz tartozó dobozháló elemeinek felhasználásával, a kontextus objektumhalmazát változatlanul hagyva olyan új kontextust konstruáltam, melynek fogalomhálóját az eredeti dobozhálójával izomorf. Ezt a konstrukciót felhasználva, szükséges és elégséges feltételt adtam az attribútumokkal kapcsolatban arra nézve, hogy egy adott kontextus esetén fennálljon a kontextushoz tartozó fogalomháló és dobozháló egyenlősége.

A konstrukció lényege a következő: legyen adott egy $K = (G, M, I)$ kontextus, és tegyük fel, hogy ismerjük a $\mathcal{D}(G, M, I)$ dobozháló elemeit. A $K^* = (G, M^*, I^*)$ kontextust úgy definiáltam, hogy az M^* halmaz attribútumai az egyes dobozelemek extenzióiba való tartozást fejezzék ki. Igazoltam, hogy ekkor az $\mathcal{F}(G, M^*, I^*)$ fogalomháló és a $\mathcal{D}(G, M, I)$ dobozháló elemeinek extenziói azonosak, tehát a két háló izomorf. Egy ismert fogalomanalízisbeli tétel szerint tetszőleges, véges V háló esetén van olyan kontextus, melyhez tartozó fogalomháló izomorf V -vel. Fontos hangsúlyozni azonban, hogy esetünkben nem a tételben szereplő kontextust határoztuk meg a dobozhálózhoz, hanem egy olyan kontextust, amelynek objektumhalmaza megegyezik a dobozháló alapkontextusának objektumhalmazával.

Bebizonyítottam, hogy egy adott (G, M, I) kontextushoz tartozó dobozháló és fogalomháló pontosan akkor esik egybe, ha minden $p \in M$ attribútum esetén a (p', p'') fogalom egyben dobozelem is. A feltételt teljesítő p attribútumokra bevezettem a **kitüntetett tulajdonság** elnevezést, mert ezek azok az attribútumok, amelyek valóban fontosak az objektumok osztályainak kialakítása szempontjából. A fogalomháló és dobozháló egyenlőségével kapcsolatos eredményeket és azok bizonyításait az értekezés 4. fejezete, illetve a [K2] és [K3] publikáció tartalmazza.

Részkontextus dobozelemei és a kontextus elemi bővítése

A K^* kontextus konstruálásánál az objektumhalmaz nem változott az eredeti kontextushoz képest. Érdekes megvizsgálni azt a kérdést is, hogy az objektumok halmazának szűkítése vagy bővítése az osztályfelbontásokat hogyan érinti. Az első esetben a részkontextus (amely az objektumhalmaz bizonyos elemeinek elhagyásával keletkezik) alapján számítható dobozelemek kapcsolatát kell feltárnunk a bővebb kontextus dobozhálójának elemeivel. A második esetben az a kérdés, hogy ha új objektumokkal bővítjük a kontextus objektumhalmazát, a bővebb kontextushoz tartozó dobozháló elemei származtathatóak-e valamilyen módon az eredeti kontextus dobozelemeiből.

Legyen adott egy $K = (G, M, I)$ kontextus, és $H \subseteq G$ esetén képezzük a $K_H = (H, M, I \cap H \times M)$ részkontextust. Igazoltam, hogy ha $E \subseteq G$ az eredeti kontextushoz tartozó dobozhálóban egy elemnek az extenziója, akkor az $E \cap H$ halmaz a szűkített kontextushoz tartozó dobozháló valamely elemének az extenziója. Ezzel összefüggésben azt is beláttam, hogy a G halmaz egy extenzió-partíciójának elemeit H -ra leszűkítve a H halmaz egy extenzió-partícióját nyerjük.

A kontextus bővítése kapcsán az objektumhalmaz elemi (azaz egyetlen objektummal való) bővítését vizsgáltam, hiszen ilyenek sorozataként tetszőleges bővítést előállíthatunk.

3. tézis:

Megvizsgáltam a kontextus szűkítésének illetve elemi bővítésének hatását a dobozháló struktúrájára. Ennek során olyan tételeket fogalmaztam meg és bizonyítottam be, melyek alapján az eredeti dobozelemekből kiindulva a megváltozott dobozháló elemei előállíthatóak. Kidolgoztam és implementáltam a dobozháló inkrementális frissítésének algoritmusát.

Tegyük fel, hogy $G = H \cup z$ és a $K = (G, M, I)$ kontextushoz tartozó dobozelemeket keressük a $K_H = (H, M, I \cap H \times M)$ kontextusból származó dobozelemek ismeretében. Bebizonyítottam, hogy ha E a K kontextus egy dobozelemének az extenziója, akkor vagy E , vagy $E \setminus \{z\}$ a szűkebb kontextus valamely dobozelemének az extenziója. Ezután arra adtam választ, hogy a K_H kontextus dobozelem-extenziói mely esetben maradnak dobozelem-extenziók az új kontextusban, illetve mikor kell bővíteni azokat az új elemmel, hogy dobozelem-extenziók legyenek. Továbbá megadtam egy olyan feltételt az új objektumra, melynek teljesülése esetén az objektum egyetlen, már létező osztályba sem sorolható be, azaz új osztályt kell létesíteni a számára.

Az eredmények felhasználásával lehetőség nyílik arra, hogy a dobozhálót inkrementális módon felépítsük (szemben az 1. tézis alapján adódó batch módszerrel). Induljunk ki a G objektumhalmaz egy tetszőleges g eleméből, a

g -re leszűkített kontextus dobozhálója csak a triviális elemeket tartalmazza. Ezután egyesével bővítve az objektumhalmazzt, minden bővítés után meghatározzuk a dobozelemeket, míg végül a teljes kontextushoz tartozó dobozhálót kapjuk.

A kontextus szűkítésével és bővítésével kapcsolatos állításokat és azok bizonyításait az 5. fejezet és a [K1] publikáció tartalmazza.

Az elméleti eredmények alkalmazása

Az 1. tézisbeli elméleti eredményekre támaszkodva algoritmusokat dolgoztam ki a dobozháló felépítésére egy adott (G, M, I) kontextusból kiindulva. Készítettem egy programot, mely mátrix-aritmetikai eszközökkel operálva meghatározza a $\mathcal{D}(G, M, I)$ dobozháló atomjait és további elemeit, valamint előállítja az $\mathcal{F}(G, M, I)$ fogalomháló osztályozási rendszereit.

Az $\mathcal{F}(G, M, I)$ háló osztályozási rendszereivel egyidejűleg a G objektumhalmaz extenzió-partícióit is előállítjuk, ami lehetőséget teremt osztályozási, csoportképzési feladatok megoldására. Ilyen jellegű problémákkal gyakran találkozunk a csoporttechnológia keretein belül, például az alkatrészek osztályozásának vagy a gyártócellák kialakításának feladatában. A művelettervezési és gyártócella tervezési feladatok modelljében kulcsszerepet játszó **gép-alkatrész incidencia mátrix** a formális kontextussal szoros analógiát mutató fogalom, mely lehetővé teszi a dobozhálón alapuló osztályfelbontási technika alkalmazását. A módszer azon kívül, hogy alkalmas a gyártócellák és alkatrészcsaládok elemeinek meghatározására, képes rugalmasan alkalmazkodni a gyártási körülményekben bekövetkező változásokhoz és lehetővé teszi a teljesítménymutatók széleskörű használatát. Ezek a tulajdonságok az igény szerinti tömeggyártásban jelentkező új, „*agile manufacturing*” paradigmához való illeszkedését jelzik. Az alkalmazás kapcsolódik az Alkalmazott Informatikai Tanszék „Csoporttechnológia-alapú tervezés és ütemezés támogatása diszkrét matematikai modellekkel és módszerekkel” c. OTKA pályázatához és az MTA-SZTAKI vezetésével folyó „Valós idejű, kooperatív vállalatok” (VITAL) c. projekthez.

A gép-alkatrész incidencia mátrixot kontextusnak tekintve, a gépek halmazának extenzió-partíciói előállíthatóak. Minden egyes felbontás kijelöl egy gyártócella-kiosztást. Kidolgoztam egy algoritmust, mellyel minden cellakiosztás esetén meghatározható az az alkatrészcsalád rendszer, melyre a megoldást értékelő csoportképzési hatások a legnagyobb. Módszeremet véletlenszerűen generált, illetve az irodalomban fellelhető mátrixok esetében is teszteltem.

A csoportképzési probléma megoldása mellett azt is vizsgáltam, hogy a további tézisek eredményei hogyan hasznosíthatóak az adott problémákban. A 2. tézisben vázolt kontextus-konstrukcióval lehetőség nyílik az alkat-

részek tulajdonságait leíró kontextus redukciójára, úgy, hogy az attribútum-oldalon csak az osztályfelbontás szempontjából fontos tulajdonságok maradjanak meg. A 3. tézisbeli kontextusbővítéssel kapcsolatos eredmények pedig abban az esetben alkalmazhatóak, ha egy új objektum (alkatrész vagy gép) gyártórendszerbe való illesztését kell megoldanunk.

4. tézis:

Kidolgoztam, implementáltam és hatékonyság szempontjából elemeztem az osztályozási rendszerek meghatározásának algoritmusát. Megmutattam, hogy a dobozhálón alapuló osztályfelbontási módszer csoporttechnológiai alapeladatokban eredményesen alkalmazható, és ezekben a feladatokban megadtam a 2. és 3. tézisbeli eredmények hasznosításának módját is.

A módszer algoritmikus leírását és a megvalósítás módját a 3.4. és 3.5. alfejezet tárgyalja, a 3.6. alfejezet pedig az algoritmusok hatékonyságát elemzi. Az elért eredmények gyakorlati alkalmazásának lehetőségeit az értekezés 6. fejezete, valamint a [K4], [K5], [K7], [K8] és [K9] publikációk ismertetik.

8. További kutatási feladatok

A kutatás továbbfejlesztésére több lehetőség is kínálkozik. Az egyik irány a kitüntetett tulajdonságokra vonatkozó elmélet kidolgozása. Megmutattuk, hogy ezek a tulajdonságok az osztályok kialakításának meghatározó tényezői, de a dobozelemektől függetlenített meghatározásukról, a kontextusbővítésben betöltött szerepükről nem esett szó. Az utóbbi kérdéssel kapcsolatban szeretnénk meghatározni a kontextusbővítésre invariáns kitüntetett tulajdonságok körét, azaz megadni azokat a kitüntetett tulajdonságokat, amelyek a kontextus bővítése után is kitüntetettek maradnak.

Egy másik feladat az új osztályfelbontási módszer megvalósítására javasolt algoritmusok fejlesztése. A dobozelemek meghatározását egy egyszerű, fogalom-extenziókat előállító algoritmus módosításával végeztük, hatékonyabb fogalom-generáló algoritmus (például egy alkalmas inkrementális módszer) adaptálásával módszerünk is jobbá tehető. A dobozelemek meghatározásának érdekében más, új módszerű dobozháló-építő algoritmusok fejlesztése is elképzelhető.

A dobozelemeneken alapuló osztályfelbontási módszer közvetlenül alkalmazható a gyártócellák és alkatrészcsaládok meghatározásának feladatában. A feladat sajátosságait figyelembe véve szeretnénk módszerünk finomításával olyan megoldási technikát kidolgozni, amely még jobb eredményre vezet a vizsgált mutatók (gépek kihasználtsága, csoportképzési hatások) tekintetében.

A dobozháló az egyedi gépgyártásban jelentkező projektek megvalósítási idejének becslésében is jól használható eszköz (a probléma leírását illetően ld. [55]). A projektek és a végrehajtásukhoz szükséges operációk kontextusából kiindulva, a dobozháló atomjainak meghatározásával választjuk ki a tervezendő alprojekthez leginkább hasonló, a gyár korábbi gyakorlatában már előfordult alprojekteket, és ezek adatainak felhasználásával adunk becslést a vizsgált projekt megvalósításához szükséges időre. Ez a feladat tovább bővíthető, hiszen a már befejezett, hasonló projektek adatainak felhasználásával nemcsak az időszükségletre, hanem a tervezendő gép várható ára (költségére), az anyagigény jellegzetességeire, a szükséges beszállítói körre vonatkozóan is adhatunk becsléseket.

A dolgozat eredményeinek további alkalmazási területei az alábbi feladatok:

- Ügyfelek, megrendelők, vevők osztályozása marketing, karbantartási kötelezettségek, újra-hasznosítási feladatok és más ügyfél-relációs feladatok megoldása érdekében.
- Gépsorok üzemeltetése során *uncertainty handling* feladatok megoldása a korábbi szituációk és a hozzárendelt eredményes irányítási aktivitások kontextusának kiértékelése alapján.

9. Összefoglalás

Dolgozatomban a fogalomháló és a hozzá kapcsolódó struktúrák alkalmazhatóságát vizsgáltam olyan feladatokban, ahol egy objektumhalmaz elemeit kell hasonló tulajdonságaik alapján csoportokra bontani. Az objektumok és az őket jellemző attribútumok egy formális kontextus által adottak, melyből kiindulva fogalomháló generálható. Az objektumhalmaz felbontásához nincs szükség az összes fogalomra, elegendő egy szűkebb struktúra, a dobozháló elemeit ismerni. A dobozháló a fogalomhálónál általában kevesebb elemet tartalmazó, atomisztikus háló, melynek elemeit közvetlenül a kontextusból is meghatározhatjuk. A dobozelemek előállítására és a keresett osztályfelbontások megadására algoritmust és számítógépes programot készítettem.

A fogalomháló és a dobozháló nem feltétlenül különbözik. Ha egy kontextusból származó fogalomháló atomisztikus, akkor minden fogalom egyben dobozelem is. A kontextus attribútumhalmaza és a dobozelemek közötti összefüggés vizsgálatával a két struktúra egybeesésére másik feltétel is adható: ugyanazon kontextus fogalomhálója és dobozhálója pontosan akkor egyezik meg, ha az attribútumok mindegyike kitüntetett tulajdonság.

Megvizsgáltam azt a kérdést is, hogy egy új objektum megfelelő osztályba való besorolása hogyan oldható meg az osztályfelbontási folyamat újraindítása nélkül. Olyan tételeket igazoltam, melyek alapján az eredeti dobozelemek ismeretében az új elem beillesztése után adódó új dobozelemek meghatározhatóak. Megadtam a dobozháló inkrementális frissítésének algoritmusát is, ami alapján lehetőség nyílik a dobozháló felépítésére a kontextus elemi bővítéseinek sorozata révén.

A módszer és a kapcsolódó eredmények hasznosíthatóságát csoporttechnológiai jellegű feladatokra mutattam meg. A dobozhálón alapuló eljárás eredményesen alkalmazható az alkatrészek osztályozásának, gyártócellák meghatározásának feladatában. A kontextus bővítésére vonatkozó eredményeknek pedig akkor vesszük hasznát, ha egy létező osztályozás esetén új objektumok megfelelő osztályokba sorolását kell megoldani.

A kutatás elején megfogalmazott hármas célkitűzésnek megfelelően a problémakör matematikai, informatikai és műszaki vonatkozásait egyaránt vizsgáltam. Igyekeztem a gyakorlatban is használható, matematikailag szilárd alapokon nyugvó módszert kidolgozni a kitűzött feladatok megoldására. Az osztályfelbontási feladat újszerű megközelítése kapcsán számos új kérdés is felmerült, ezek megválaszolása a matematikai háttér további kutatását igényli és egyúttal újabb eredményeket is ígér.

Applying Concept Lattices in Classification Problems Summary

Concept lattices and their related structures can successfully be applied in solving classification problems. In these problems given objects are to be grouped based on their similarities. The objects and their attributes are given in a formal context, from which a concept lattice can be generated. There is no need for knowing all of the concepts to form the suitable partitions of the object set, instead we can use the box lattice for this purpose. Box lattice has less elements and simpler structure than the concept lattice and its elements can be determined directly from the context as well. The objective of the research was to develop a method for generating the elements of the box lattice and the classification systems from which the demanded partitions of the object set can be identified. The developed process proved to be very useful for solving problems in the field of Group Technology.

The new scientific results are summarized as follows:

Thesis 1. Determination of the box lattice

Having studied the relationship between the box elements and the context, theorems have been formulated and verified for constructing the box lattice. The results make it possible to generate the elements of the classification systems directly from the context.

Thesis 2. The equality of the concept lattice and the box lattice

Using the box elements of a given context a new context is constructed with the same object set, the concept lattice of which is isomorphic with the original box lattice. Based on this construction a necessary and sufficient condition related to the attributes has been proved for the equality of the concept and the box lattice corresponding to the same context.

Thesis 3. Subcontext and extension of a context

Having analyzed the reducing and extending of the context and its effect on the structure of the box lattice new results have been formulated and verified aiming to generate the elements of the modified box lattice using the former box elements. An algorithm has been elaborated and implemented for building the box lattice in incremental mode.

Thesis 4. Applying theoretical results

Based on the results of Thesis 1, an algorithm has been implemented for determining classification systems. The new method and the developed program can be successfully applied in problems originated from Group Technology such as grouping parts and formation of machine cells. In addition, the usefulness of the results achieved in Thesis 2 and Thesis 3 is also demonstrated in the mentioned practical problems.

Az értekezés témakörében készített saját publikációk

- [K1] Ganter B., **Körei A.**, Radeleczki S.: *Extent partitions and context extensions*. (Manuscript, 2007)
- [K2] **Körei A.**: *Reducing Machine-Part Incidence Matrix for Cell Formation*. IFAC Workshop on Manufacturing Modelling, Management and Control, Budapest, November 2007, 187-191.
- [K3] **Körei A.**: *The Equality of the Concept Lattice and the Box Lattice Corresponding to the Same Context*. microCAD 2007 International Scientific Conference, Vol. G, Miskolc, 2007, 53-58.
- [K4] Tóth T., **Körei A.**: *New Approaches in Solving Machine-Part Grouping Problems*. Production Systems and Information Engineering, Vol.3, 2006, 37-46.
- [K5] **Körei A.**, Tóth T.: *Mathematical Methods in Solving Machine-Part Grouping Problems*. MITIP 2006 Proceedings, Budapest, September 2006, 409-414.
- [K6] **Körei A.**, Radeleczki S.: *Box Elements in a Concept Lattice*. Contributions to ICFCA 2006, Dresden, February 2006, 41-56.
- [K7] **Körei A.**: *Fogalomhálók alkalmazása gyártócellák és alkatrészcsaládok meghatározására*. GÉP, LVII. évfolyam, 2006/10 szám, 2006, 14-20.
- [K8] **Körei A.**: *Solving Cell Formation Problems with Concept Lattices*. Machine Engineering, Vol. 5, 3-4, 2005, 37-46.
- [K9] **Körei A.**: *Different Methods for Solving Cell Formation Problems*. Doktoranduszok Fóruma Szekciókiadvány, Miskolc, 2005, 98-103.
- [K10] **Körei A.**: *Using the Box Lattice Method in Solving Classification Problems*. microCAD 2004 International Scientific Conference, Vol. E, Miskolc, 2004, 59-64.
- [K11] **Körei A.**: *Osztályozási feladatok megoldása dobozhálókkal*. Doktoranduszok Fóruma Szekciókiadvány, Miskolc, 2004, 157-164.
- [K12] **Körei A.**: *Osztályozási problémákban felhasznált algoritmusok*. Doktoranduszok Fóruma Szekciókiadvány, Miskolc, 2003, 137-141.

Hivatkozott irodalom

- [1] Adenso-Díaz B., Lozano S., Eguía I.: *Part-machine grouping using weighted similarity coefficients*. Computers & Industrial Engineering, 48, 2005, 553-570.
- [2] Ballakur A., Steudel H. J.: *A within-cell utilization based heuristic for designing cellular manufacturing systems*. International Journal of Production Research, 25, 1987, 639-655.
- [3] Brown E. C., Sumichrast R. T.: *Evaluating performance advantages of grouping genetic algorithms*. Engineering Applications of Artificial Intelligence, 18, 2005, 1-12.
- [4] Bukszár Z.: Fogalomhálókön alapuló adatbányászási technológiák. Diplomamunka, Miskolci Egyetem, 2001.
- [5] Burbidge J. L.: Production Flow Analysis for Planning Group Technology. Clarendon Press, Oxford, 1996.
- [6] Chan H. M., Milner D. A.: *Direct clustering algorithm for group formation in cellular manufacturing*. Journal of Manufacturing Systems 1(1), 1982, 65-75.
- [7] Chandrasekharan M. P., Rajagopalan R.: *Groupability: an analysis of the properties of binary data matrices for group technology*. International Journal of Production Research, 27(6), 1989, 1035-1052.
- [8] Chandrasekharan M. P., Rajagopalan R.: *ZODIAC - An algorithm for concurrent formation of part-families and machine cells*. International Journal of Production Research, 25(6), 1987, 835-850.
- [9] Chandrasekharan M. P., Rajagopalan R.: *MODROC: an extension of rank order clustering for group technology*. International Journal of Production Research, 24(5), 1986, 1221-1233.
- [10] Cheng C. H.: *Algorithms for Grouping Machine Groups in Group Technology*. Omega, 20(4), 1992, 493-501.
- [11] Chu C. H.: *Clustering analysis in manufacturing cell formation*. Omega, 17, 1989, 289-295.
- [12] Cormen T. H., Leiserson C. E., Rivest R. L.: Algoritmusok. Műszaki Könyvkiadó, Budapest, 1997.

- [13] McCormick W. T., Schweitzer P. J., White T. W.: *Problem decomposition and data reorganization by a clustering technique*. Operation Research, 24, 1972, 993-1009.
- [14] Czédli G.: Hálóelmélet. Egyetemi jegyzet, JATE, Szeged, 1996.
- [15] Davey B. A., Priestley H. A.: Introduction to Lattices and Order. Cambridge University Press, Cambridge, 2002.
- [16] De Lit P., Falkenauer E., Delchambre A.: *Grouping genetic algorithms: an efficient method to solve the cell formation problem*. Mathematics and Computers in Simulation, 51, 2000, 257-271.
- [17] Detzky I., Erdélyi F., Fridrik L., Tóth T. (szerk): *A csoporttechnológia számítógépes tervezése*. Elemző tanulmány. NME Gépészmérnöki Kar, Miskolc, 1987.
- [18] Falkenauer E.: *The grouping genetic algorithm - widening the scope of GAs*. JORBEL - Belgian Journal of Operations Research, Statistic and Computer Science, 33, 1992, 79-102.
- [19] Ganter B., Wille R.: Formal Concept Analysis, Mathematical Foundations. Springer Verlag, Berlin-Heidelberg-New York, 1999.
- [20] Godin R., Missaoui R., Alaoui H.: *Incremental concept formation algorithms based on Galois lattices*. Computational Intelligence, 11(2), 1995, 246-267.
- [21] Grätzer G.: General Lattice Theory. Birkhäuser Verlag, Basel-Boston-Berlin, 2003.
- [22] Han J., Kamber M.: Adatbányászat, Koncepciók és technikák. Panem Könyvkiadó, Budapest, 2004.
- [23] Ho T. B.: *Formal Concept Analysis and Rough Set Theory in Clustering*. In: Do. L. V., Ito M. (editors), The Mathematical Foundation of Informatics, World Scientific, 2005, 43-53.
- [24] Iványi A.: Informatikai algoritmusok 2. ELTE Eötvös Kiadó, Budapest, 2005.
- [25] Jajuga K., Sokolowski A., Bock, H.: Classification, Clustering, and Data Analysis: Recent Advances and Applications. Springer, 2002.
- [26] James T. L., Brown E. C., Keeling K. B.: *A hybrid grouping genetic algorithm for the cell formation problem*. Computers & Operations Research, 2007, 2059-2079.

- [27] Jeon G., Leep H. R.: *Forming part families by using genetic algorithm and designing machine cells under demand changes*. Computers & Operations Research, 33, 2006, 263-283.
- [28] Kamrani A. K., Logendram R. (editors): *Group Technology and Cellular Manufacturing, Methodologies and Applications*. Gordon and Breach Science Publishers, 1998.
- [29] Kent R. E.: *Rough Sets, Fuzzy Sets and Knowledge Discovery*. Springer Verlag, 1994.
- [30] King J. R.: *Machine-component grouping in production flow analysis: an approach using a rank order clustering algorithm*. International Journal of Production Research, 18(2), 1980, 213-232.
- [31] Kovács L.: *Algorithms for Building Concept Set and Concept Lattice*. Production Systems and Information Engineering, 1, 2003, 91-106.
- [32] Kumar C., Chandrasekharan M. P.: *Grouping efficacy: a quantitative criterion for goodness of block diagonal forms of binary matrices in group technology*. International Journal of Production Research, 28, 1990, 233-243.
- [33] Kusiak A.: *Intelligent Design and Manufacturing*. John Wiley & Sons, New York-Chichester-Brisbane-Toronto-Singapore, 1992.
- [34] Kusiak A.: *The Generalized Group Technology Concept*. International Journal of Production Research, 25, 1987, 561-569.
- [35] Kusiak A., Boe J. W., Cheng C. H.: *Designing cellular manufacturing systems. Branch-and-bound and A* approaches*. IEE Transactions, 25(4), 1993, 46-56.
- [36] Kuznetsov S. O., Obiedkov S. A.: *Comparing performance of algorithms for generating concept lattices*. Journal of Experimental and Theoretical Artificial Intelligence, 14(2-3), 2002, 189-216.
- [37] Lovász L.: *Algoritmusok bonyolultsága*. Nemzeti Tankönyvkiadó, Budapest, 1994.
- [38] Mirkin B.: *Mathematical Classification and Clustering*. Kluwer Academic Publishers, Dordrecht, 1996.
- [39] Nagi R., Harhalakis G., Proth J.M.: *Multiple routings and capacity considerations in group technology applications*. International Journal of Production Research, 28, 1990, 2243-57.

- [40] Radeleczki S.: *Fogalomhálók és alkalmazásuk a csoporttechnológiában.* Proceedings of International Computer Science, microCAD, Miskolc, 1999, 3-8.
- [41] Radeleczki S.: *Classification systems and the decomposition of a lattice into direct products.* Mathematical Notes, Vol.1. No.1, 2000, 145-156.
- [42] Radeleczki S.: *Classification systems and their lattice.* Discussiones Mathematicae, General Algebra and Applicationes 22, 2002, 167-181.
- [43] Radeleczki S., Tóth T.: *Fogalomhálók és fuzzy módszerek kapcsolata és alkalmazásuk a csoporttechnológiában.* Kutatási jelentés, Miskolci Egyetem, 1999.
- [44] Radeleczki S., Tóth T.: *Fogalomhálók alkalmazása a csoporttechnológiában. Osztályozórendszerek matematikai alapjainak összehasonlító vizsgálata.* Kutatási jelentés, Miskolci Egyetem, 2001.
- [45] Prezenszky J.: *Logisztika I-II.* Budapest, 2001.
- [46] Sandbothe R. A.: *Two observations on the grouping efficacy measure for goodness of block diagonal forms.* International Journal of Production Research, 36, 1998, 3217-3222.
- [47] Sarker B., Mondal.: *Grouping efficiency measures in cellular manufacturing: a survey and critical review.* International Journal of Production Research, 37, 1999, 285-314.
- [48] Seifoddini H., Djassemi M.: *The treshold value of a quality index for formation of cellular manufacturing systems.* International Journal of Production Research, 34, 1996, 3401-3416.
- [49] Seifoddini H., Wolfe P. M.: *Application of the similarity coefficient method in group technology.* IEE Transactions, 18(3), 1986, 271-277.
- [50] Shafer S. M.: *Part-Machine-Labour Grouping: The Problem and Solution Methods.* In: Suresh N. C., Kay J. M. (editors), *Group Technology and Cellular Manufacturing, State-of-the-Art Synthesis of Research and Practice*, Kluwer Academic Publishers, Boston-Dordrecht-London, 1998, 131-152.
- [51] Singh N.: *Design of cellular manufacturing systems: An invited review.* European Journal of Operational Research, 69, 1993, 284-291.

- [52] Suresh N. C., Kay J. M. (editors), Group Technology and Cellular Manufacturing, State-of-the-Art Synthesis of Research and Practice, Kluwer Academic Publishers, Boston-Dordrecht-London, 1998.
- [53] Szász G.: Hálóelmélet. Tankönyvkiadó, Budapest, 1975.
- [54] Tóth T.: Termelési rendszerek és folyamatok. Miskolci Egyetemi Kiadó, Miskolc, 2004.
- [55] Tóth T., Radelecki S., Erdélyi F.: *Similarity Based Project Planning in the Field of the Production of Individual Machines*. Proceedings of the 37th CIRP International Seminar on Manufacturing Systems, Budapest, 2004, 225-230.
- [56] Yin Y., Yasuda K.: *Similarity coefficient methods applied to the cell formation problem: A taxonomy and review*. International Journal of Production Economics, 101, 2006, 495-510.
- [57] Vadász D.: Tengely- és tárcsaszerű alkatrészek műveleti sorrendtervezésének automatizálása. Kandidátusi értekezés, Budapest, 1982.
- [58] Vakharia A. J.: *Methods of cell formation in group technology: A framework for evaluation*. Journal of Operations Management, 6, 1986, 257-271.
- [59] Venugopal V.: *Artificial Neural Networks and Fuzzy Models: New Tools for Part-Machine Grouping*. In: Suresh N. C., Kay J. M. (editors), Group Technology and Cellular Manufacturing, State-of-the-Art Synthesis of Research and Practice, Kluwer Academic Publishers, 1998, 169-184.
- [60] Ventos V., Soldano H.: *Alpha Galois Lattices: An Overview*. ICFCA, 2005, Lecture Notes in Computer Science 3403, Springer, 2005, 299-314.
- [61] Wettle F., Mayer Gy., Szabó P.: Latex kézikönyv. Panem Könyvkiadó, Budapest, 2004.
- [62] Wille R.: *Restructuring lattice theory: an approach based on hierarchies of concepts*. In: Ivan Rival (editor), Ordered Sets, Reidel, Dordrecht-Boston, 1982, 445-470.
- [63] Wille R.: *Subdirect decomposition of concept lattices*. Algebra Universalis 17, 1983, 275-287.

Ábrák, táblázatok, algoritmusok jegyzéke

Ábrák jegyzéke

1.	Háromelemű halmaz részhalmazhálójának Hasse diagramja . .	11
2.	Fogalomháló Hasse diagramja	18
3.	Dobozháló Hasse diagramja	22
4.	Osztályozandó alkatrészek	59
5.	Dobozháló alkatrészek osztályozásához	60
6.	Új, besorolandó alkatrészek	61
7.	Dobozháló a bővített kontextushoz	62
8.	Dobozháló diagramja a Cheng tanulmányában szereplő kontextushoz	71

Táblázatok jegyzéke

1.	Példa kontextusra	14
2.	Fogalom-extenziók meghatározásának algoritmusai	17
3.	Dobozelem-extenziók meghatározása	30
4.	Az extenzió-partíciók meghatározása	32
5.	A K^* kontextus konstrukciója és a K^{**} kontextus	47
6.	Kontextus alkatrészek osztályozásához	60
7.	Kontextus létrehozása incidencia mátrixból	69
8.	Teszteredmények példamátrixokra	72

Algoritmusok jegyzéke

2.20.	Fogalom-extenziók meghatározása	16
3.1.	Atom-extenziók meghatározása	25
3.5.	Dobozelem-extenziók meghatározása	28
3.7.	Osztályozási rendszerek előállítása	31
5.8.	Atom-extenziók meghatározása elemi bővítés után	53
5.15.	Dobozelem-extenziók meghatározása elemi bővítés után	56
6.1.	Alkatrészcsoportok meghatározása	69