

MISKOLCI EGYETEM DOKTORI (PHD) TÉZISFÜZETEI
HATVANY JÓZSEF INFORMATIKAI TUDOMÁNYOK DOKTORI ISKOLA

**FOGALOMHÁLÓK ALKALMAZÁSA
OSZTÁLYFELBONTÁSI PROBLÉMÁKRA**

KÉSZÍTETTE:

KÖREI ATTILA

OKLEVELES MATEMATIKUS

AKI DOKTORI (PHD) FOKOZAT ELNYERÉSÉRE PÁLYÁZIK

TUDOMÁNYOS VEZETŐ:

DR. RADELECZKI SÁNDOR

A MATEMATIKAI TUDOMÁNY KANDIDÁTUSA

DOKTORI ISKOLA VEZETŐ:

PROF. DR. TÓTH TIBOR

A MŰSZAKI TUDOMÁNY DOKTORA

MISKOLC, 2008

Bíráló bizottság

Elnök:

Dr. Szigeti Jenő CSc, dr. habil., egyetemi tanár (ME)

Titkár:

Dr. Vadász Dénes CSc, egyetemi docens (ME)

Tagok:

Dr. Koltai Tamás CSc, dr. habil., egyetemi tanár (BME)

Dr. Erdélyi Ferenc CSc, ny. egyetemi docens

Dr. Horváth Mátyás DSc, professor emeritus (BME)

Dr. Mátyási Gyula CSc, egyetemi docens (BME)

Dr. Váncza József CSc, laborvezető-helyettes (MTA-SZTAKI)

Hivatalos bírálók:

Dr. Kovács László PhD, egyetemi docens (ME)

Dr. Szegh Imre PhD, egyetemi docens (BME)

Tartalomjegyzék

1. Bevezetés	4
2. Osztályokra bontás fogalomháló és dobozháló felhasználásával	4
3. A kutatás célja	8
4. Irodalmi áttekintés	8
5. Új tudományos eredmények	10
5.1. Dobozhálók elemeinek meghatározása	10
5.2. Adott kontextushoz tartozó fogalomháló és dobozháló egyenlősége	11
5.3. Részkontextus dobozelemei és a kontextus elemi bővítése	11
5.4. Az elméleti eredmények alkalmazása	13
6. További kutatási feladatok	14
7. Köszönetnyilvánítás	15
8. Az értekezés témakörében készített saját publikációk	16
Hivatkozott irodalom	16
Summary	19

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.

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.

2. Osztályokra bontás fogalomháló és dobozháló felhasználásával

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 adatbányászat-

beli 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 fogalomháló 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 egy (G, M, I) hármas, ahol G objektumok, M attribútumok (tulajdonságok) egy halmaza, I pedig egy reláció a $G \times M$ halmazon, a következő előírással: $g \in G, m \in M$ esetén $(g, m) \in I$ pontosan akkor teljesül, ha a g objektum rendelkezik az m tulajdonsággal. A kontextus alapján G és M bizonyos részhalmazaiból összetartozó párokat állíthatunk elő. Ha $A \subseteq G, B \subseteq M$ akkor az (A, B) párt **fogalomnak** nevezzük, ha az A -beli objektumok közös tulajdonságai alkotják a B halmazt, és ugyanakkor a B -beli tulajdonságok mindegyikével csak az A halmaz objektumai rendelkeznek. Azaz $A' = B$ és $B' = A$, ahol

$$\begin{aligned} A' &= \{m \in M \mid (g, m) \in I \quad \forall g \in G\} \\ B' &= \{g \in G \mid (g, m) \in I \quad \forall m \in M\}. \end{aligned}$$

Az (A, B) fogalom objektumrészét a fogalom **extenziójának**, míg az attribútumokból álló részt a fogalom **intenziójának** nevezzük. A fogalmak extenziói zárt halmazok a közös tulajdonságaikra nézve, azaz csak az extenzióba eső objektumok rendelkeznek ezekkel a tulajdonságokkal.

A fogalmak között az

$$(A_1, B_1) \leq (A_2, B_2) \iff A_1 \subseteq A_2$$

előírással egy részbenrendezési relációt definiálunk. Ezzel a relációval a (G, M, I) kontextusból származtatható fogalmak halmaza egy teljes hálót alkot, amelyet a kontextushoz tartozó **fogalomhálónak** nevezünk, és $\mathcal{F}(G, M, I)$ -vel jelölünk. (Az L részbenrendezett halmaz akkor teljes háló, ha tetszőleges $H \subseteq L$ esetén $\inf H$ és $\sup H$ létezik.)

A G halmaz

$$\{G_j \mid j \in J, G = \bigcup_{j \in J} G_j, G_j \cap G_k = \emptyset, \text{ ha } j \neq k\}$$

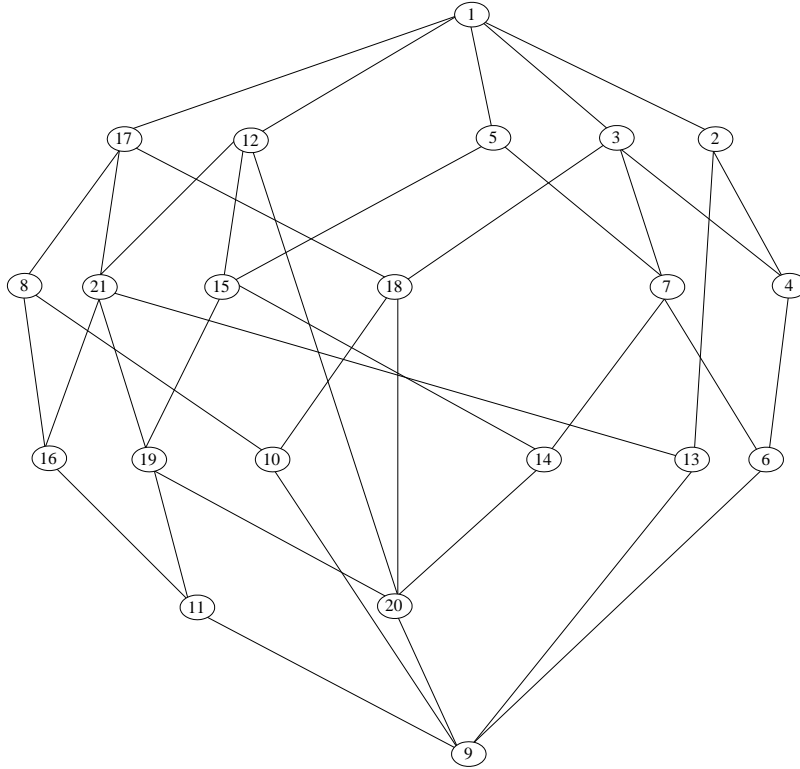
partíciói közül azokat, amelyeknél a G_j blokkok mindegyike valamely $\mathcal{F}(G, M, I)$ -beli fogalom extenziója, a G halmaz **extenzió-partícióinak** hívjuk; egy ilyen partíció blokkjait objektumok olyan halmazai alkotják, melyek zártak az objektumok közös tulajdonságaira nézve.

A $\{g\}' = \{m \in M \mid (g, m) \in I\}$ halmazt a $g \in G$ objektumhoz tartozó **objektum-intenzió**nak nevezzük, és röviden g' -vel jelöljük. Azt mondjuk, hogy a (G, M, I) kontextus **sor-redukált**, ha G nem tartalmaz olyan g objektumokat, amelyekre a g' halmaz előáll más objektum-intenziók metszeteként. Ha a (G, M, I) kontextus sor-redukált, akkor a G halmaz tetszőleges $\{A_j \mid j \in J\}$ extenzió-partíciójának természetes módon megfelelő $\{(A_j, A'_j) \mid j \in J\}$ rendszer a $\mathcal{F}(G, M, I)$ fogalomháló egy ún. osztályozási rendszere. Ez a fogalom általánosabb struktúrákban is értelmezhető.

Az L teljes háló nemüres részhalmazából 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. $x = \bigvee_{j \in J} (x \wedge a_j) \quad \forall x \in L$.

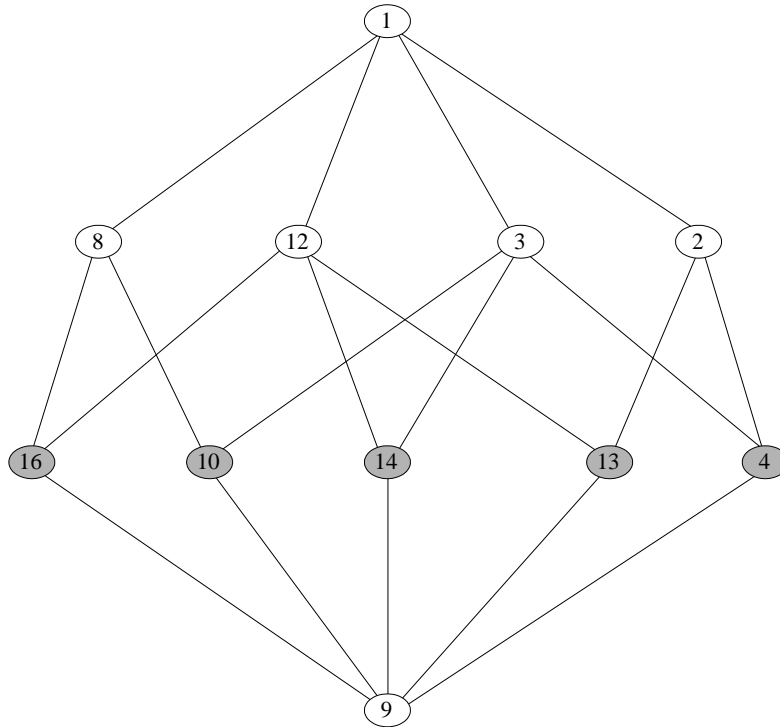
(A definícióban szereplő \wedge jel két elem infimumát, a \bigvee szimbólum egy halmaz elemeinek szuprémumát jelöli, $\mathbf{0}$ pedig a háló legkisebb eleme.)



1. ábra. Fogalomháló diagramja

Azokat a fogalmakat, amelyek elemei lehetnek a fogalomháló valamely osztályozási rendszerének, **dobozelemeknek** nevezzük. Jelölje $\mathcal{D}(G, M, I)$ a dobozelemek halmazát, melyet kiegészítettünk a fogalomháló legkisebb elemével.

$\mathcal{D}(G, M, I)$ a fogalomhálón adott rendezési reláció leszűkítésével szintén egy teljes hálót alkot, melyet a fogalomháléhoz (vagy a kontextushoz) tartozó **dobozhálónak** nevezünk. A dobozháló atomisztikus háló, azaz bármely eleme előáll atomok szuprémumaként, ahol atomok alatt a háló legkisebb elemének rákövetkezőit értjük.



2. ábra. Dobozháló diagramja. Az atomokat sötétebb ellipszisek jelölik, a nagyobb elemek előállnak atomok szuprémumaként.

Az 1. ábra egy fogalomháló, a 2. ábra egy dobozháló diagramját mutatja be. A csomópontok hálóelemeket reprezentálnak, úgy, hogy $x < y$ pontosan akkor igaz, ha az x -hez tartozó csomópontból vezet az y -hoz tartozó csomópontig végig „felfelé” haladó töröttvonal. Esetünkben mindkét háló ugyanabból a kontextusból származik, így az azonos számmal jelölt csomópontok megegyező fogalmakat jelölnek.

Bebizonyítható, hogy egy sor-redukált kontextushoz tartozó fogalomháló és dobozháló osztályozási rendszerei megegyeznek, így az objektumhalmaz extenzió-partícióit a kevesebb elemet tartalmazó, szabályosabb szerkezetű dobozháló segítségével is előállíthatjuk. Ez az eredmény jelenti a dolgozatban bemutatott osztályfelbontási módszer alapját.

3. 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.

4. Irodalmi áttekintés

A kutatási téma adatbányászati alapjait, az osztályozás és klaszterezés viszonyát illetve módszereit Han és Kamber bevezető jellegű művének egy-egy fejezete tárgyalja [8]. A legfontosabb klaszterező eljárások leírása az Iványi Antal által szerkesztett *Informatikai algoritmusok* c. könyvben is megtalálható [10]. Mivel az objektumok osztályozása az adatbányászat kezdetektől vizsgált problémája, természetesen önálló munkák is rendelkezésre állnak a témában való elmélyüléshez [11], [15].

A részbenrendezett halmazok és hálók elméletének tanulmányozásához Szász Gábor műve szolgálhat alapul [21], egy újabb mértékadó forrás pedig Czédli Gábor jegyzete, mely már a fogalomhálókkal kapcsolatos alapvető összefüggéseket is tartalmazza [3]. Természetesen a hálóelmélet témakörében számos kiváló, idegen nyelvű monográfia is rendelkezésre áll, például Grätzer György *General Lattice Theory* c. könyve [7], vagy a közérthető formában megírt *Introduction to Lattices and Order* c. munka Davey és Priestley tollából [4].

A fogalomanalízis területén Bernhard Ganter és Rudolph Wille könyve megkerülhetetlen forrásnak számít [6], 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* [13] alkalmazása. Ez a kutatási terület adja a háttérét Ho T. B. cikkének, aki három klaszterezési feladatra ad modellt és megoldó algoritmust [9]. 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 [26]. 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 [27]. 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 [16], [17], [18].

Az értekezésben bemutatott gyakorlati feladatok háttérében csoporttechnológiai elvek alkalmazása áll. A csoporttechnológia fogalmával, kialakulásával és korai fejlődési szakaszaival foglalkozik a [5] 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 [22]. A csoporttechnológiai elvek korszerű alkalmazási területeiről nyújtanak áttekintést a [12] és [20] 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 [19], az egyes technikákat pedig a [1], [2], [14], [24], [25] publikációk alapján tanulmányoztam.

5. Új tudományos eredmények

5.1. 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 szuprimumaként. 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. tézis:

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 az értekezés 3. fejezetének első három alfejezete tartalmazza, illetve a [K6], [K10], [K11], [K12] publikációk is ezekre az eredményekre épülnek.

5.2. 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 fogalomhálója az eredeti dobozhálóval 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 fogomanalízisbeli 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 dobozháléhoz, 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.

5.3. Részkontextus dobozelemei és a kontextus elemi bővítése

A K^* kontextus konstruálásánál az objektumhalmaz nem változott az eredeti kontextushoz képest. Érdeemes 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-partíciójában az elemek 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. Kידolgoztam 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 dobozhálójában csak a triviális elemeket tartalmazza. Ezután egyesével bővítve az objektumhalmazt, 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 értekezés 5. fejezete és a [K1] publikáció tartalmazza.

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

6. 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 függetlenül 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-extendziókat 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 dobozelemeken 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. [23]). A projektek és a végrehajtásukhoz szükséges operációk kontextusából kiindulva, a dobozháló atomjainak meghatározásával választjuk ki a tervezendő alprojekt-hez 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 árára (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.

7. 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.

8. 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.**: *Fogalomháló 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] Burbidge J. L.: Production Flow Analysis for Planning Group Technology. Clarendon Press, Oxford, 1996.
- [2] Chan H. M., Milner D. A.: *Direct clustering algorithm for group formation in cellular manufacturing*. Journal of Manufacturing Systems 1(1), 1982, 65-75.
- [3] Czédli G.: Hálóelmélet. Egyetemi jegyzet, JATE, Szeged, 1996.
- [4] Davey B. A., Priestley H. A.: Introduction to Lattices and Order. Cambridge University Press, Cambridge, 2002.
- [5] Detzky I., Erdélyi F., Fridrik L., Tóth T. (szerk): *A csoporttechnológia számítógépes tervezése*. Elemző tanulmány. NME Gépészmérnöki Kar, Miskolc, 1987.
- [6] Ganter B., Wille R.: Formal Concept Analysis, Mathematical Foundations. Springer Verlag, Berlin-Heidelberg-New York, 1999.
- [7] Grätzer G.: General Lattice Theory. Birkhäuser Verlag, Basel-Boston-Berlin, 2003.
- [8] Han J., Kamber M.: Adatbányászat, Konceptiók és technikák. Panem Könyvkiadó, Budapest, 2004.
- [9] 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.
- [10] Iványi A.: Informatikai algoritmusok 2. ELTE Eötvös Kiadó, Budapest, 2005.
- [11] Jajuga K., Sokolowski A., Bock, H.: Classification, Clustering, and Data Analysis: Recent Advances and Applications. Springer, 2002.
- [12] Kamrani A. K., Logendram R. (editors): Group Technology and Cellular Manufacturing, Methodologies and Applications. Gordon and Breach Science Publishers, 1998.
- [13] Kent R. E.: Rough Sets, Fuzzy Sets and Knowledge Discovery. Springer Verlag, 1994.
- [14] Kusiak A.: *The Generalized Group Technology Concept*. International Journal of Production Research, 25, 1987, 561-569.
- [15] Mirkin B.: Mathematical Classification and Clustering. Kluwer Academic Publishers, Dordrecht, 1996.

- [16] Radeleczki S.: *Fogalomhálóok és alkalmazásuk a csoporttechnológiában*. Proceedings of International Computer Science, microCAD, Miskolc, 1999, 3-8.
- [17] Radeleczki S.: *Classification systems and the decomposition of a lattice into direct products*. Mathematical Notes, Vol.1. No.1, 2000, 145-156.
- [18] Radeleczki S.: *Classification systems and their lattice*. Discussiones Mathematicae, General Algebra and Applications 22, 2002, 167-181.
- [19] 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.
- [20] 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.
- [21] Szász G.: *Hálóelmélet*. Tankönyvkiadó, Budapest, 1975.
- [22] Tóth T.: *Termelési rendszerek és folyamatok*. Miskolci Egyetemi Kiadó, Miskolc, 2004.
- [23] Tóth T., Radeleczki 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.
- [24] 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.
- [25] 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.
- [26] Ventos V., Soldano H.: *Alpha Galois Lattices: An Overview*. ICFCA, 2005, Lecture Notes in Computer Science 3403, Springer, 2005, 299-314.
- [27] Wille R.: *Subdirect decomposition of concept lattices*. Algebra Universalis 17, 1983, 275-287.

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.