



INTELLIGENS RENDSZEREK I. LABORATÓRIUM

4. laborgyakorlat segédlete

Bayes-hálók: a valószínűségektől a függetlenségi és oksági viszonyokig

Készítette:

Antal Péter (antal@mit.bme.hu)

Kovács Dániel László (dkovacs@mit.bme.hu)

Méréstechnika és Információs Rendszerek Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem

2008, április.

Tartalomjegyzék

1. BEVEZETÉS	3
2. BAYES-STATISZTIKAI MÓDSZEREK.....	4
A VALÓSZÍNŰSÉG BAYESI ÉRTELMEZÉSE	4
A BAYESI MODELL.....	4
KÖVETKEZTETÉS.....	4
MONTE CARLO MÓDSZEREK	4
BAYES-STATISZTIKAI MEGKÖZELÍTÉS ELŐNYEI	5
3. BAYES-HÁLÓK.....	6
A VALÓSZÍNŰSÉGI DEFINÍCIÓ: SZINTAXIS ÉS SZEMANTIKA.....	6
KAUZÁLIS DEFINÍCIÓ	7
BAYES-HÁLÓK ÉS A TUDÁSMERNŐKSÉG	7
KÖVETKEZTETÉS BAYES-HÁLÓKBAN	8
BAYES-HÁLÓK TANULÁSA	9
BAYES-HÁLÓK TANULÁSA HIÁNYOS ADATOK ALAPJÁN.....	9
JEGYTANULÁS	9
4. KITEKINTÉS	11
5. BAYES HÁLÓK, DÖNTÉSI HÁLÓK ÁGENSEKBEN.....	12
HASZNOSSÁG-ORIENTÁLT ÁGENSEK.....	12
6. BAYES HÁLÓK KOOPERATÍV ÉS SZEKVENCIÁLIS DÖNTÉSEKBEN.....	14
7. FÜGGELÉK	17
A. A BNET SZOFTVER HASZNÁLATA.....	17
1.1. LÉTEZŐ MODELL MEGNYITÁSA	17
1.2. BAYES-HÁLÓK SZERKESZTÉSE	17
1.3. KÖVETKEZTETÉS (ALAP).....	17
• Állítsuk be az ismert értékeket:.....	18
• Végezzük el a következtetést:.....	18
• Ismert értékek/megfigyelések visszavonása	18
1.4. KÖVETKEZTETÉS (MOST PROBABLE EXPLANATION)	19
1.5. KÖVETKEZTETÉS (ÉRZÉKENYSÉG VIZSGÁLAT)	19
B. BNET MODELLEK HASZNÁLATA A JADE KERETRENDSZERBEN.....	20
8. IRODALOMJEGYZÉK	24

1. Bevezetés

A Bayes-háló alapú alkalmazások térhódítása a 90-es években kezdődött el (Heckerman1995b), kezdetben főleg orvosi diagnosztikai és előrejelző rendszereknél. Az elmúlt néhány évben a felhasználási kör olyan változatos területekkel bővült, mint pénzügyi, telekommunikációs vagy hadszíntéri döntéstámogatás és hírszerzői információk integrálása. A felhasználói viselkedéshez kapcsolódó alkalmazások két fő irányvonal mentén fejlődtek, egyrészt a személyre szabott információszolgáltatás terén, mint például a felhasználót segítő sugórendszerek (Horvitz1998), valamint az információs rendszerek biztonsága területén (Wright2004), ahol a rosszindulatú felhasználók kiszűrése a cél a viselkedésmintázatok vizsgálta alapján. Ehhez hasonló osztályozási feladat a spam levelek kiszűrése, melyre számos Bayes-hálón alapuló megoldás született (Sahami1998). Komplex rendszerek működtetésénél, legyen az mozdonyszerelés (Przytula2004) vagy nyomtatórendszer karbantartás (Skaanning1998), ahol a diagnosztika a bonyolult felépítés és bizonytalanság miatt egyszerű szabályalapú módszerekkel nem követhető, szintén hatékonyan alkalmazható a bayesi megközelítés. Mindemellett különböző döntéstámogatási rendszerekben (Tchangani2002) és ezen belül, a kockázat-előrejelzés (Ramamurthy) terén is jelentős pozíciót töltenek be a Bayes-háló alapú alkalmazások. Egyes területeken, mint a Bayes-háló alapú adatbányászatnál (Myllymaki2001) vagy az előbb említett kockázat-előrejelzésnél a Bayes-háló tanulását maga a felhasználó irányíthatja.

2. Bayes-statisztikai módszerek

A valószínűség bayesi értelmezése

A Bayes-statisztika és a Bayes-háló modellosztály közös alapvető célja, hogy a bizonytalan háttértudáson és megfigyeléseken alapuló következtetések számára axiomatikus alapot és gyakorlati alkalmazhatóságot biztosítson. A fellépő bizonytalanságnak számos oka lehet, például a tudás kinyerése során alkalmazott módszer, az adatgyűjtési eljárás, vagy a tudás hiánya, esetleg figyelmen kívül hagyása.

A Bayes-statisztikai módszertan a bizonytalanság kezelésére a valószínűségi keretrendszert alkalmazza, a valószínűség szubjektivistá interpretációját elfogadva, szemben a mérnöki gyakorlatban elterjedtebb, a relatív gyakoriságok határértékein alapuló, ún. frekventista értelmezéstől. A szubjektivistá értelmezésben a valószínűségeket az események bekövetkeztében való, adott kontextushoz tartozó a priori hiedelemnek, elvárásnak, egyfajta meggyőződési mértéknek tekintjük. Az axiomatikus származtatásnál megmutatható, hogy egy döntési problémában¹ minden eseményhez rendelhető egy pozitív való szám, mely az adott esemény valószínűségéként értelmezhető és egy hasznossági érték, melyekkel a preferenciák egzaktul reprezentálhatók és racionális döntések hozhatók.

A bayesi modell

A bayesi módszertan további axiomatikus alapját a reprezentációs tételek (Bernardo1995) jelentik, ezek megmutatják, hogy egy végtelen felcserélhetőséget teljesítő eloszlás (azaz amelyben bármely ' π ' permutációra $p(x_1, \dots, x_n) = p(x_{\pi(1)}, \dots, x_{\pi(n)})$), reprezentálható egy alkalmas adatgenerálási parametrikus modellosztállyal és egy e feletti eloszlással.

A valószínűség szubjektivistá értelmezésére és a fenti tulajdonságú modellosztályok léteire alapozva javasolható a Bayes-statisztikai keretrendszer, amelyben a megfigyelési adatokat valószínűségi változók által paraméterezett modellegyüttesekből származtatjuk, azaz a megfigyelések és a modelparaméterek ugyanolyan modellezési szinten helyezkednek el.

Gyakorlati megközelítésekben az alkalmazott modellosztály paraméterezését hierarchikusan tagolják, leggyakrabban a következő módon, amit a cikkben is követünk: egyrészt a modelltér diszkrét elemek (a lehetséges modellstruktúrák) halmaza, másrészt hozzájuk numerikus paraméterek tartoznak.

Következtetés

A következtetés során a feladat, hogy megbecsüljük egy adott esemény, vagy egy modell feltételes valószínűségét az alapismereteink és a megfigyelési adatok szerint. Az első esetben *prediktív*, a másodikban *parametrikus* becslésről (a posteriori eloszlás számításáról) beszélünk. Mindkét esetben a Bayes-tételből indulunk ki, amelynek segítségével események feltételes valószínűségét számíthatjuk (a továbbiakban D az adatokat, G egy struktúrát, θ pedig egy paraméterezést jelöl):

$$(1) \quad P(G, \theta | D) = \frac{P(D | G, \theta) P(G, \theta)}{P(D)}$$

Az a posteriori eloszlást (röviden *posterior*) az előzőleg említett modelltér egy struktúrájának paraméterezésére, vagy magára a struktúrák terére is kiszámíthatjuk. A paraméterek esetén a képlet formailag megegyezik (1)-gyel, a struktúrákra vonatkozó pedig a paraméterek kiintegrálása után adódik:

$$(2) \quad P(G | D) = \frac{\int P(D | G, \theta) P(G, \theta) d\theta}{P(D)}$$

Predikció esetén még egy lépést teszünk: a keresett valószínűséget kiszámítjuk minden létező modellre, és ezeknek a modelleknek a posteriori valószínűségével súlyozott átlagát vesszük:

$$(3) \quad p(x | D) = \sum_k p(G_k | D) \int p(x | \theta_k) p(\theta_k | G_k, D) d\theta_k$$

Monte Carlo módszerek

A fenti posteriorok gyakran nem mintavételezhetők, ezért Monte Carlo módszereket kell alkalmaznunk, például a fontossági mintavételezést vagy a Markov-lánc Monte Carlo (MCMC) módszerek egyikét (Gilks1995). A posterior vizsgálata helyett a feladat gyakran egy vagy több $\bar{f} = E_{p(x|D)}[f(x)]$ alakú várható érték becslésére egyszerűsödik. Ez megtehető a következő lépésekkel, melyek helyességét az MCMC módszereknél igazolt nagy számok törvénye (2. pont) és centrális határeloszlás tétel (3. pont) biztosítja [Gilks1995]:

1. a $\{x_i\}_{i=1}^N$ minta vételezése
2. $\bar{f}$ becslése az $\hat{f} = \frac{1}{N} \sum_{i=1}^N f(x_i)$ képlet alapján
3. konfidenciabecslés az $|\bar{f} - \hat{f}|$ eltérésre

¹ Egy döntési probléma egy (E, C, A, <) négyessel definiálható, ahol 'E' az események, 'C' a következmények, 'A' a lehetséges cselekvések halmaza, '<' pedig az 'A' elemei feletti preferenciáinkat tükröző rendezés (Bernardo1995).

A Monte Carlo mintavételezés mellett még gyakori a meghatározott számú, legnagyobb a posteriori valószínűségű modellstruktúra alapján történő kiszámítás, amely legegyszerűbb esetben egyetlen, az ún. MAP (maximum a posteriori) modell használatát jelenti. A legnagyobb valószínűségű modell(ek) meghatározását a tanulásról szóló fejezet tárgyalja.

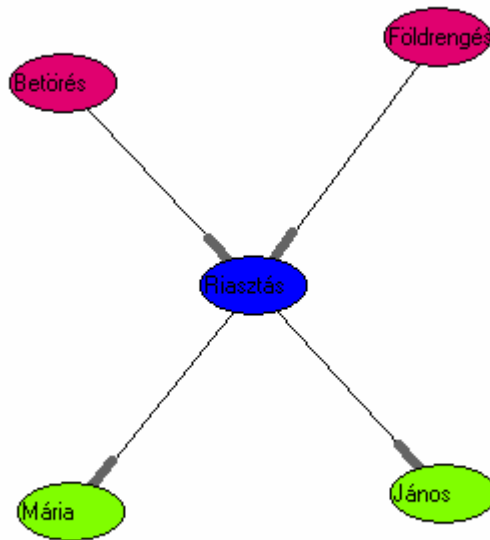
Bayes-statisztikai megközelítés előnyei

A következő listában röviden összefoglaljuk a fentebb ismertetett bayesi módszertannak a klasszikus statisztikával szembeni előnyeit (Roberts2001):

- A paraméterek bizonytalanságát a felettük definiált eloszlással jellemezzük, így minden statisztikai következtetés egy direkt valószínűségi állítás, ami az automatizált többlépéses tanulási rendszereknél és tudásbázisok generálásánál igen előnyös.
- A paraméterbecslés egy inverziós feladatként fogható fel, hisz itt kizárólag az adatból következtetünk arra a paraméterre, amely annak generálását meghatározta. A Bayes-tétel pontosan ezt az inverziót formalizálja, így a következtetést a hipotetikus viselkedés figyelmen kívül hagyásával végzi, szemben a klasszikus statisztika egyes módszereivel.
- Az a priori eloszlások (röviden *priorok*) használata alkalmas az előismeretek összegzésére vagy akár a teljes ismerethiány kifejezésére is.
- A priorok – mivel leggyakrabban korábbi megfigyeléseken vagy vizsgálatokon alapulnak – az ismeretszerzési folyamat egyes fázisainak tekinthetők, hisz új tudásunkat (az a posteriori eloszlást) ez alapján szerezünk.
- A bayesi következtetés a Bayes-tétel segítségével egyenrangú módon, normatívan kombinálja az előismeretekben és az adatokban rejlő információkat. Így a Bayes-tétel használata az adatok és az előismeretek egyfajta súlyozását valósítja meg: az adatok mennyiségének növekedtével azok befolyása is nő az a posteriori eloszlásra.
- Az a posteriori eloszlás használata pontbecslés helyett a predikció során nem csak a legvalószínűbb konfiguráció alapján számol, hanem figyelembe veszi a kevésbé valószínű eseteket is, ami a modell komplexitásához képest kis mennyiségű megfigyelés esetén fontos.

3. Bayes-hálók

A valószínűségi megközelítésben bizonytalan tudásunkat sztochasztikus változók együttes eloszlásával reprezentáljuk. A szisztematikus struktúrával nem rendelkező tárgyerületek esetén (szemben pl. a kép- és hangfeldolgozással) az ilyen eloszlások modellezésére használt elsődleges eszközt ma a Bayes-hálók jelentik. Ezekben egy irányított körmentes gráfban (DAG – directed acyclic graph) reprezentálják a változókat és a köztük lévő összefüggéseket: minden csomópont egy-egy változót jelöl, és minden csomóponthoz tartozik egy lokális feltételes valószínűségi modell, amely leírja a változó függését a szüleitől (a pontos definíciót a következő fejezet tartalmazza).



Mint reprezentációs eszköz, egy Bayes-háló háromféle értelmezést kaphat, ezek a felsorolás sorrendjében egyre erősebb modellezési, értelmezési lehetőséggel bírnak:

- Tekinthető egyszerűen az együttes eloszlás egy hatékony ábrázolásának, hisz a csomópontonkénti feltételes valószínűségi modellekre való faktorizálással a felhasznált paraméterek száma jelentősen csökken.
- Egy adott struktúra meghatározza, hogy az ábrázolt eloszlásban milyen feltételes függések és függetlenségek lehetnek, azaz az élek tekinthetők a közvetlen valószínűségi összefüggések reprezentációjának, míg a teljes gráf a reprezentált eloszlás függési térképének.
- Az előzőnél is erősebb a kauzális értelmezés, amelyben minden élt az érintett két csomópont közötti ok-okozati összefüggésként értelmezzük.

A valószínűségi definíció: szintaxis és szemantika

Egy Bayes-háló struktúrája és a reprezentálni kívánt eloszlás közti kapcsolatot az alábbi négy feltételre alapozhatjuk, melyekről belátható (Cowell1999), hogy ekvivalensek.

- A $P(X_1, \dots, X_n)$ eloszlás *faktorizálható* a G DAG szerint, ha:

$$P(X_1, \dots, X_n) = \prod P(X_i | Pa(X_i)),$$

ahol $Pa(X_i)$ az X_i csomópont szülői halmaza.

- A $P(X_1, \dots, X_n)$ eloszlásra teljesül a *sorrendi Markov-feltétel* G szerint, ha

$$\forall i = 1..n : I(X_{\pi(i)} | Pa(X_{\pi(i)}) | \{X_{\pi(j)} | j < i\} \setminus Pa(X_{\pi(i)}))_P,$$

ahol az $I(X|Y|Z)$ reláció az X feltételes függetlenségét jelenti a Z-től Y feltétellel, π pedig a struktúra egy topologikus rendezése

- A $P(X_1, \dots, X_n)$ eloszlásra teljesül a *lokális (szülői) Markov-feltétel* G szerint, ha bármely változó független nem-leszármazottaitól, feltéve szüleit.
- A $P(X_1, \dots, X_n)$ eloszlásra teljesül a *globális Markov-feltétel* G szerint, ha

$$\forall x, y, z \subseteq \{X_i\} : I(x | z | y)_G \Rightarrow I(x | z | y)_P,$$

vagyis, ha z d -szeparálja² x -et y -től a G gráfban, akkor x független y -től, feltéve z -t.

Egy elfogadott definíció a Markov-feltételek által adott függőségi rendszer tulajdonságaira épít (Pearl1988):

A ' G ' irányított körmentes gráf a ' $P(U)$ ' eloszlás Bayes-hálója (U az összes változó halmaza), akkor és csak akkor, ha minden $u \in U$ változót a gráf egy csomópontja reprezentál, a gráfra teljesül valamelyik (és így az összes) Markov-feltétel, és a gráf minimális (azaz bármely él elhagyásával a Markov-feltétel már nem teljesülne). Míg ez a definíció egyértelműen a valószínűségi függetlenségek rendszerének reprezentációjaként tekint a Bayes-hálóra, addig a mérnöki gyakorlatban közkedvelt az alábbi, praktikus meghatározás:

Az ' U ' valószínűségváltozó-halmaz Bayes-hálója a (G, θ) páros, ha ' G ' irányított körmentes gráf, amelyben a csomópontok jelképezik U elemeit, θ pedig csomópontokhoz tartozó ' $P(X|Pa(X))$ ' feltételes eloszlásokat leíró numerikus paraméterek összessége.

Fontos megjegyezni, hogy a definiált modellosztályban a lehetséges struktúrák száma a csomópontok számában szuperexponenciális, ez pedig pl. a később tárgyalandó tanulás komplexitását is befolyásolja.

Bár egy Bayes-háló egyaránt tartalmazhat diszkrét és folytonos változókat is, mi a továbbiakban kizárólag diszkrét, véges változókkal foglalkozunk, feltéve továbbá, hogy minden lokális feltételes valószínűségi modell a multinominális eloszlásokhoz tartozik, így a paraméterek ún. feltételes valószínűségi táblák (FVT-k) elemei.

Egy adott Bayes-háló struktúrája meghatározza, hogy az milyen függéseket írhat le (pl. külön komponensekben lévő változók közt nem lehet függés), azonban különböző struktúrákhoz is tartozhat azonos implikált függési rendszer. Ha két struktúrából ugyanazok a feltételes függetlenségek olvashatók ki, a két gráfot *megfigyelés-ekvivalensnek* mondjuk. Belátható (Pearl1988), hogy két gráf akkor és csak akkor megfigyelés-ekvivalens, ha irányítás nélküli vázuk, illetve v -struktúráik (az $A \rightarrow B \leftarrow C$ típusú részgráfok, úgy, hogy A és C közt nincs él) megegyeznek.

A megfigyelési ekvivalencia segítségével a struktúrákat diszjunkt osztályokba sorolhatjuk. Minden ilyen ekvivalencia osztályt egy ún. esszenciális PDAG³ gráffal reprezentálhatunk. Az esszenciális gráf váza megegyezik az osztályba tartozó gráfokéval, és csak azok az élei irányítottak, amelyek iránya mindegyik gráfban megegyezik (ezek az ún. kényszerített –compelled– élek).

Kauzális definíció

Az előző, tisztán valószínűségi definíciók bevezetése után formálisan könnyen áttérhetünk a Bayes-hálók kauzális értelmezésére: *egy (G, θ) páros kauzális Bayes-hálója a $P(U)$ eloszlásnak, ha egyrészt a tárgyterület valószínűségi modellje az előző értelmezések szerint, továbbá minden él közvetlen ok-okozati viszonyt jelképez.*

Hasonlóan, itt is létezik egy Markov-feltétel: *egy $P(U)$ eloszlás és egy kauzális relációkat leíró G gráf teljesíti a kauzális Markov-feltételt, ha G és $P(U)$ teljesíti a lokális Markov-feltételt.*

A Markov-feltétel teljesülése biztosítja, hogy minden (kauzális) függés kiolvasható a gráfból, a másik irányhoz, ahhoz tehát, hogy minden a gráfból kiolvasott függés teljesüljön az eloszlásban, annak stabilnak kell lennie. *Egy $P(U)$ eloszlás stabil, ha létezik olyan G gráf, hogy $P(U)$ -ban pontosan a G -ből d -szeparációval kiolvasható függések és függetlenségek teljesülnek benne* (pl. megfelelő paraméterezés mellett előfordulhat, hogy egy $A \rightarrow B \rightarrow C$ struktúrában A és C függetlenek).

A fenti kauzális definíció a modell és a tárgyterület összefüggéseinek értelmezését illetően igen erős, a megfigyelési adatok statisztika elemzésének kereteit meghaladó eszközt szolgáltat. Alkalmazásakor figyelembe kell vennünk, milyen nem kauzális kapcsolatok okozhatnak valószínűségi összefüggést két változó között, azaz milyen korlátai vannak a kauzális értelmezésnek. Ilyenek lehetnek pl.:

- Zavaró változók: a két változó közti függést okozhatja egy közös ősz (az ún. zavaró változó) is.
- Kiválasztási bias: a változók közti függés lehet az adatgyűjtési mód következménye is (pl. ha egy orvosi adatbázisba csak a komolyabb megfázással kezelt betegek kerülnek be, akkor a láz és torokfájás között direkt függést figyelhetünk meg).
- Az ősz-ok, leszármazott-okozat megfeleltetés és a DAG gráfstruktúra kizárja a mechanizmusokban lévő visszacsatolások (ciklikusságok), illetve az oda-vissza ható okozatiság lehetőségét.
- A modellter maga (azaz, hogy milyen változók szerepelnek, illetve azok milyen értékészlettel rendelkeznek) szintén befolyásolja, hogy milyen direkt függések jelennek meg (azaz a gráf struktúráját).

Bayes-hálók és a tudásmérnökség

A fentebb definiált Bayes-háló a tudásmérnökség eszközeként jelent meg a 80-as években, konstruálása jellemzően a szakértőktől származó adatokból történt manuálisan. A kézi konstruálás még napjainkban is jelentős súlyt képvisel a Bayes-hálók alkalmazásában, másrészt ahol az adathoz viszonyítva jelentős a priori tudás áll rendelkezésre, ott a Bayes-hálók tudásmérnöki alkalmazása a bayesi keretrendszer alkalmazásának egy kezdeti fázisát jelenti, nevezetesen a prior konstruálást.

² ' z ' d -szeparálja ' x '-et és ' y '-t a ' G ' gráfban ($x, y, z \subseteq V(G)$), ha minden ' x ' és ' y ' között menő irányítatlan ' p ' utat blokkol, azaz, ha (1) ' p ' tartalmazza ' z ' egy elemét nem összefutó élekkel, vagy (2) ' p ' tartalmaz egy ' n ' csomópontot összefutó élekkel, hogy ' z ' nem tartalmazza sem ' n '-t, sem valamelyik leszármazottját.

³ Egy PDAG (partially directed acyclic graph) gráf vegyesen tartalmaz irányított és irányítatlan éleket.

A tudásmérnökség metodikájára nagy hatással volt a nagy mennyiségű elektronikus tárgyterületi információ megjelenése, a megfelelő mennyiségű statisztikai adat elérhetősége, valamint a Bayes-statisztikai alapú gépi tanulási módszerek elterjedése. A felépített tudásbázissal szemben követelményként jelent meg a bayesi módszerek alkalmazásakor, hogy támogassa a priorok konstruálását, hiszen a valószínűségekkel leírt a priori tudás és a rendelkezésre álló adatok bayesi frissítéssel történő kombinációja szolgáltatja a végső tudásmodellt. Mindemellett fontos, hogy a tudásbázis segítse komplex, akár szabad szöveges háttérismereteket is tartalmazó valószínűségi állítások megfogalmazását, valamint tegye lehetővé a szakértőktől származó szubjektív információ tárolását, mely releváns a bayesi a priori tudásmodell megalkotásánál.

Egy tudásbázis megépítéséhez olyan környezetben, ahol rendelkezésre áll elektronikus tárgyterületi tudás, elegendő statisztikai adat, valamint a megfelelő bayesi módszerek, az alábbi lépések szükségesek (amelyekből a specifikusokat részletezzük):

1. *Célok, alkalmazási terület és modellezési szintek identifikációja*
Terminológia és ontológia elfogadása.

2. *Nem rendszerezett tudás begyűjtése*

Ehhez a lépéshez tartozik az összes releváns elektronikus és egyéb szövegalapú információforrás feldolgozása, ami magába foglalja az a priori információ kinyerését különféle szövegbányászati metódusok alkalmazásával, mint például az általunk kifejlesztett módszer, amit a 4.8. fejezetben mutatunk be.

3. *Struktúra kinyerése*

A G DAG struktúrák feletti $p(G)$ priorok konstruálása, melyek egyesítik a szakértők által megadott információkat az elektronikus forrásokból kinyert információkkal. A $p(G)$ a priori eloszlást többnyire normalizálatlan formában lehet előállítani: pl. egy adott referencia struktúrától való eltérés alapján

$$P(G|\xi) \propto \kappa^\delta; 0 < \kappa < 1,$$

ahol δ a referenciától való tetszőlegesen definiált strukturális tulajdonságokbeli eltéréseknek a száma.

4. *Paraméter és hiperparaméter kinyerése*

A valószínűségi paraméterek számos módon nyerhetők: adatbázisok, szakirodalom vagy szakértők szubjektív véleménye alapján. A $p(\theta|G)$ paraméter prior specifikációja az általunk vizsgált diszkrét, véges esetben egy egyszerű módszerrel megtehető, ha feltehetjük az egyes változókhoz és szülői értékkonfigurációkhoz tartozó paraméterek függetlenségét

$$P(\theta|G_0, \xi) \propto \prod_{i=1 \dots n} \prod_{j=1 \dots q_i} P(\theta_{i,j} | G_0, \xi).$$

Egy szinte kizárólagosan használt eloszláscsalád az adott változó, adott szülői értékkonfigurációjához tartozó $P(\theta_{i,j} | G_0, \xi)$ megadására a Dirichlet eloszlás $\text{Dir}(\theta_{i,j} | \alpha_{i,j}, \xi)$, ahol az $\alpha_{i,j}$ hiperparaméter jelentése a paraméterhez tartozó szülői értékkonfiguráció korábban megfigyelt eseteinek számait jelenti (Cowell1999). Megmutatható, hogy a Dirichlet család az egyetlen lehetséges választás, ha az ugyanazon megfigyelési ekvivalencia osztályba tartozó G struktúrákhoz ekvivalens priorokat szeretnénk megadni, ami kauzális modellezésnél nem szükségszerű (Heckerman1995a). További feltevések mellett az is bizonyítható, hogy az összes struktúrához konzisztens $p(\theta|G)$ definíciója ekvivalens egy teljes modellhez tartozó pontparametrizációnak és egyetlen korábban megfigyelt összesetszámmal jelentő hiperparaméternek a megadásával. E kettő együtt valójában egy a priori adathalmazt definiál, ami korábban megfigyelt eseteket tartalmazza, így az összesetszámmal virtuális vagy a priori mintaszámmak nevezünk.

5. *Érzékenységi analízis, verifikáció és validáció*

A modellek posteriorjának vizsgálata magába foglalja egyrészt az a priori eloszlásokra való érzékenység vizsgálatát (ami különösen fontos a több szakértőt és tudásbázist is magában foglaló automatizáltan származtatott prioroknál), másrészt referencia priorokkal való összehasonlítást. Mindkét esetben gyakran szükséges a modellosztály komplexitása miatt, egyrészt hogy modell jegyeket használjunk, másrészt hogy MAP modellre alapozzuk a vizsgálatot.

Mint ahogy az látható, a tudásbázis építése a bayesi modellkiértékeléssel zárul. A kiértékelés tartalmazza az adat és a modell kompatibilitásának vizsgálatát és az a posteriori valószínűségek vizsgálatát, azaz a tudásmérnöki folyamat lényege az a priori modell konstruálása a későbbi tanulási folyamat számára.

Következtetés Bayes-hálókbán

Egy konkrét Bayes-hálóban való következtetés alapfeladata a $P(X = x | Y = y, G, \theta)$ mennyiség kiszámítása, azaz adott egy struktúra és paraméterezése valamint ismert a bizonyítékváltozók (Y) behelyettesítése, kérdés a lekérdezésváltozók (X) egy adott konfigurációjának valószínűsége.

Könnyen belátható (Glymour1999), hogy a feladat NP-teljes (hiszen pl. visszavezethető a kielégíthetőségi problémára), számításigénye a csomópontok számában exponenciális. Ezért a gyakorlatban vagy szimuláción alapuló, közelítő eredményt adó Monte Carlo módszereket (Gilks1995), vagy a gráfot másodlagos struktúrákba transzformáló ún. junction-tree algoritmusokat (Huang1996) alkalmaznak.

Hogy a $P(X = x | Y = y)$ mennyiséget kiszámíthassuk, azaz valódi bayesi predikciót végezzünk, a (3) képlet szerinti összegzést és integrálást kell elvégezni. Ilyenkor az (2.4) fejezet közelítései alkalmazhatók.

Bayes-hálók tanulása

Mivel a teljes bayesi következtetés annak komplexitása miatt csak különleges esetekben hajtható végre, gyakran a teljes modell helyett csak egyetlen modellt használunk. Ha elegendő statisztikai adat áll rendelkezésre, a fent bemutatott manuális konstruálás mellett szerepet kaphat az optimális modell keresése, a *tanulás*, mely végezhető a szakértői modelltől kiindulva, annak finomításával, vagy tabula rasa alapon is. A tanulás, mint az optimális modell keresése, a parametrikus következtetés alkalmazásának tekinthető, és megmutatható, hogy NP-teljes bonyolultságú (Glymour1999), az adatok szükséges mennyiségére kívánt közelítési hiba mellett (Friedman1996) ad képletet.

A tanulás két szinten lehetséges: kereshetjük adott struktúra mellett az optimális paraméterezést (paramétertanulás), vagy a legjobb struktúrát és annak paraméterezését (struktúratanulás). Az optimalitás valamilyen mérték szerint értendő, ez legegyszerűbb esetben a modell a posteriori valószínűsége.

A MAP modell keresése mellett elképzelhető más kritériumfüggvény is, amely leggyakrabban az a posteriori valószínűség egyenletes priorral, kiegészítve valamilyen, a struktúra bonyolultságát büntető taggal. Az ilyen büntetés alkalmazása felfogható a prior módosításának: minél erősebb a büntetés, annál kisebb a bonyolult struktúrák valószínűsége. A leggyakoribb ilyen minősítési függvény a bayesi információ-kritérium függvény (BIC – Bayesian information criterion), ennek képlete (Friedman1997):

$$(4) \quad BIC(G, D) = \log P(D | G) - \frac{\log N}{2} |\theta|,$$

ahol 'N' a tanító minták, $|\theta|$ pedig a háló paramétereinek száma. A $\log N$ -nel arányos mellett még elképzelhető N-ben lineáris vagy polinomiális büntetés is.

Számítási igényét tekintve a tiszta a posteriori kritériumfüggvény, és a teljes, független, azonos eloszlású minták alapján végzett tanulás a legegyszerűbb. Ekkor, Dirichlet eloszlású paraméterpriorrt feltéve, adott struktúra a posteriori valószínűsége egyszerű, zárt formában számítható (Cooper1992):

$$(5) \quad P(G | D) \propto P(G, D) = P(G) \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{(N'_{ij} + r_i - 1)!}{(N_{ij} + N'_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} \frac{(N_{ijk} + N'_{ijk})!}{N'_{ijk}!},$$

ahol az N_{ijk} az i. változó j. szülői konfigurációjának és k. értékének az előfordulását, q_i az i. változó szülői konfigurációinak a számát és r_i az értékeinek számát jelenti (N_{ij} a megfelelő marginális). Az N'_{ijk} a megfelelő virtuális mintaszámokat jelöli (ezek előismeretek hiányában 1-nek választhatók).

Paramétertanulás esetén az optimális paraméterezés az FVT-k külön-külön, relatív gyakoriságokkal való kitöltésével elérhető, struktúratanulás esetén pedig minden csomóponthoz külön megkereshető az optimális szülői halmaz, feltéve hogy ismert a csomópontok egy kauzális rendezése⁴. Ha ilyen információ nem áll rendelkezésre, ügyelni kell, hogy a DAG tulajdonság ne sérüljön, pl. úgy, hogy minden lehetséges sorrendet külön megvizsgálunk.

Bayes-hálók tanulása hiányos adatok alapján

Amennyiben a tanító adatok hiányosak, azaz bizonyos változók értéke nem minden esetben ismert a tanulás feladata jóval nehezebbé válik. Ilyenkor a paramétertanulásban iteratív eljárások használhatók, a legismertebbek ezek közül a gradiens alapú közelítő eljárások vagy ezek robosztusabb változatai, a konjugált gradiens és a skálázott konjugált gradiens algoritmusok (Bishop1995), vagy az expectation maximization algoritmus (Gelman1995).

Struktúratanulás esetén, mivel a szülői halmazok nem tanulhatók külön még adott sorrendnél sem, a teljes struktúratanulást bejáró keresésre van szükség. Mivel a lehetséges struktúrák száma a csomópontok számával szuperexponenciálisan nő, a gyakorlatban nem teljes keresési eljárásokat kell alkalmazni, pl. mohó keresést vagy szimulált lehűtést (ekkor az elemi lépés pl. egy él törlése, beszúrása, megfordítása lehet).

Ezek az eljárások is csak akkor működnek azonban, ha az adatokra teljesül a véletlenszerű eltűnés (MAR – missing at random) feltétele, azaz ha a bejegyzések eltűnése nem függ az eltűnt értéktől (Gelman1995).

Jegytanulás

A jegytanulás során bizonyos részstruktúrák (ún. jegyek) meglétének valószínűségét keressük. Ilyen jegy lehet a legegyszerűbb esetben egy adott él megléte, vagy Markov-határ keresése. Egy X csomóponthalmaz Markov-takarója egy olyan Y halmaz, melyre igaz, hogy $I(X|Y|U(X \cup Y))$ (azaz Y d-szeparálja X-et a háló többi részétől). Egy csomópont vagy csomóponthalmaz Markov-határa annak minimális Markov-takarója. Ez lehetővé teszi egy szimmetrikus, páronkénti reláció definiálását a Markov-határbeliséget, az egymás Markov-határában való előfordulást (MBM(X,Y) – Markov boundary membership). A jegytanulás alternatív megoldást jelenthet a struktúratanulással szemben, mivel ha segítségével meg tudjuk állapítani a fent említett viszonyok meglétének

⁴Egy kauzális rendezésben a csomópontok szülei csak az őket megelőző változók közül kerülhetnek ki. Egy kauzális rendezés a reprezentáns DAG csúcsainak egy topologikus rendezése.

valószínűségét (azaz, hogy egy csomópont beletartozik-e egy másik Markov-határába), akkor ezzel a MAP modell egy jó közelítését konstruálhatjuk.

A kérdéses valószínűségek számítása, a bayesi következtetés sémáját követi, amiből következően összegeznünk kell azon struktúrák a posteriori valószínűségét, amelyek rendelkeznek a kívánt jeggyel:

$$(6) \quad P(F = f \mid D) = \sum_{k: 1_x[G_k] = f} P(G_k \mid D)$$

Természetesen itt is alkalmazhatók közelítő Monte Carlo módszerek, mivel a struktúrák feletti összegzés túl számításigényes, hacsak nincsenek rendkívül pontos a priori ismereteink a lehetséges struktúrákról.

4. Kitekintés

Az eddigi fejezetek rövid áttekintést adtak a monolitikus Bayes-hálók használatáról. A monolitikus jelző ez esetben arra utal, hogy egy adott problémára konstruált hálóban nincsen hierarchikus vagy moduláris dekomponálás. A következőkben rövid áttekintést adunk a Bayes-hálók kiterjesztésére törekvő irányzatokról.

Az első lépést ebben az irányban az annotált Bayes-hálók vizsgálatával tettük meg, ami lehetőséget adott tetszőleges szemantikai információ bevitelére és automatizált felhasználására. A következő lépés a már említett jegytanulás volt, mivel ennek felhasználásával felfedezhetők reguláris hálórészletek (bizonyos területeken gyakori az ok-okozati mechanizmusokban felfedezhető, ismétlődő mintázat, pl. a biológiában egyes gének aktivációs sémái).

A modularizációs igényre adott formális válasz az objektumorientált Bayes-hálók (OBN) megjelenése volt (Laskey1997). Mint nevük is mutatja, a programozástechnikában ismert objektumorientált paradigmához hasonlóan terjesztik ki a Bayes-hálókat. Egy objektumorientált Bayes-hálózat objektumokból áll, melyek szintén tovább bonthatók objektumokra, vagy egyszerű valószínűségváltozó-csomópontokra. Ezzel a többszintű hierarchiával a teljes rendszer funkcionálisan különálló részei elszigetelhetők egymástól, valamint lehetővé válik előre felépített részhalóknak a teljesbe építése. Hasonló koncepció áll a valószínűségi relációs modellek mögött is (Koller1998).

5. Bayes hálók, döntési hálók ágensekben

Hasznosság-orientált ágensek

A reflex-szerű, memóriával rendelkező, cél-orientált ágens architektúrák közt a legrugalmasabb, ámde nem minden helyzetben optimális a hasznosság-orientált ágens. Elméleti szempontból alapvetően két részre tagolható:

1. Valószínűségi modell
2. Hasznosság modell

Mint ismeretes, a hasznosságfüggvény létezésének feltételei az alternatívák sorrendezhetősége, tranzitivitása, folytonossága, helyettesíthetősége, monotonitása és felbonthatósága [MI].

A bayesi döntéselmélet értelmében ezek felhasználásával az optimális (azaz a várható hasznosságot maximáló) döntés (A) meghozható adott evidenciák (E) mellett.

$$a^* = \operatorname{argmax}_a EU(a|E) = \sum_i P(\text{Eredmény}_i(a)|\text{Tesz}(a), E) U(\text{Eredmény}_i(a))$$

A hasznosság egy általánosabb teljesítmény mérték: a világ állapotainak az összehasonlítása, hogy milyen boldoggá tennék az ágens, ha elérné azokat.

„Boldogság” – tudományosan „előnyösebb” egy másikhoz képest, ha magasabb a **hasznossága (utility)** az ágens számára.

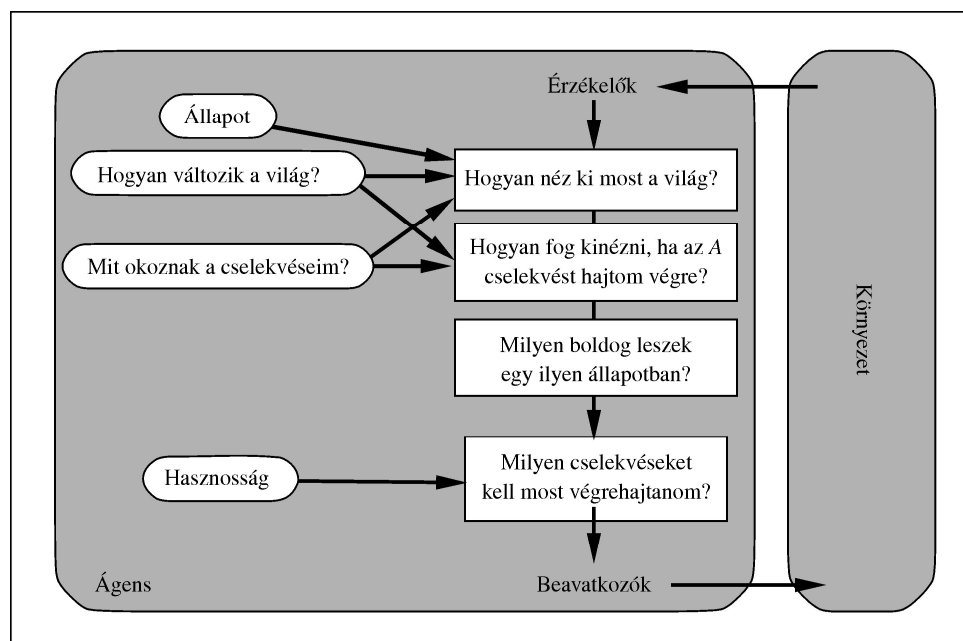
A **hasznosság**: egy függvény, amely egy állapotot egy olyan valós számmá képez le, amelyik a hozzá rendelt boldogság fokát írja le.

A **hasznosság függvény**: racionális döntések, amikor a célok bajban vannak

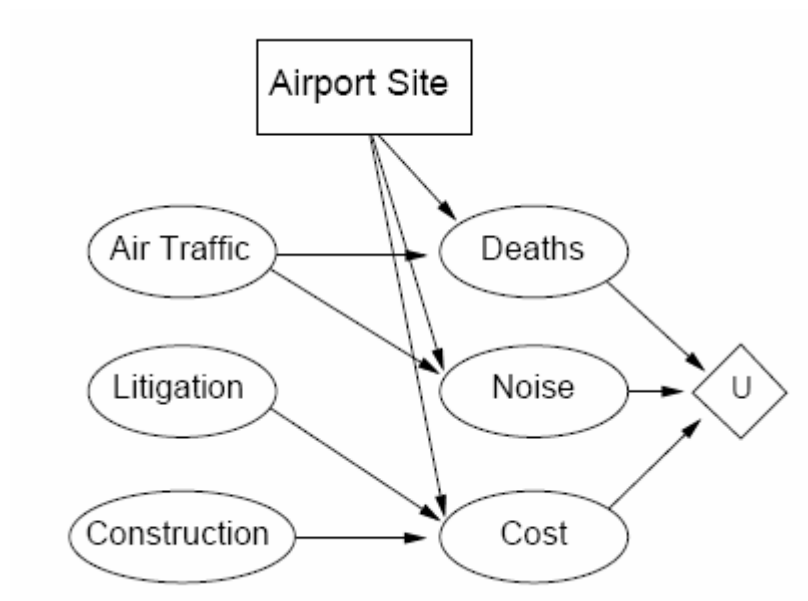
- egymásnak ellentmondó célok (mint pl. a biztonságosság és a sebesség),

1. ha több cél van, amire az ágens törekedhet.

Az általános ágens architektúra a következő.



A világ, a cselekedetek és a hasznosságfüggvény együttes reprezentálására a döntési háló formalizmus használható, ami a Bayes hálók kiegészítése döntési csomópontokkal (négyszögek) és hasznosság csomópontokkal (rombusz).



Megjegyzés: a labor keretei között csak a valószínűségi modellt használjuk.

6. Bayes hálók kooperatív és szekvenciális döntésekben

A Bayes háló reprezentáció ideális eszköz több szakértői modell fuzionálására. A lehetséges megközelítések két szélsőséges példái a következők

- Evidenciák kommunikációja, majd közös, integrált döntés, vagy
- Evidenciák külön kezelése, különálló döntések, majd azok aggregálása.

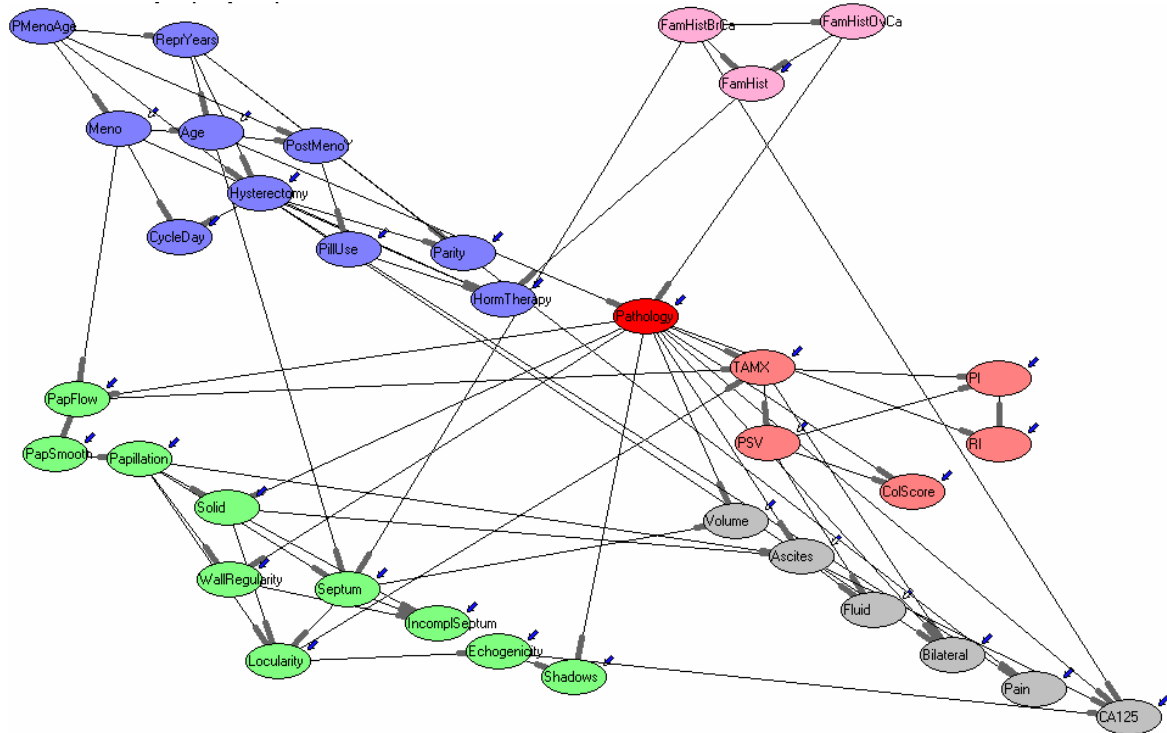


Figure 1 Evidenciák kommunikációja, majd közös, integrált döntés (egyes színek az egyes szakértők).

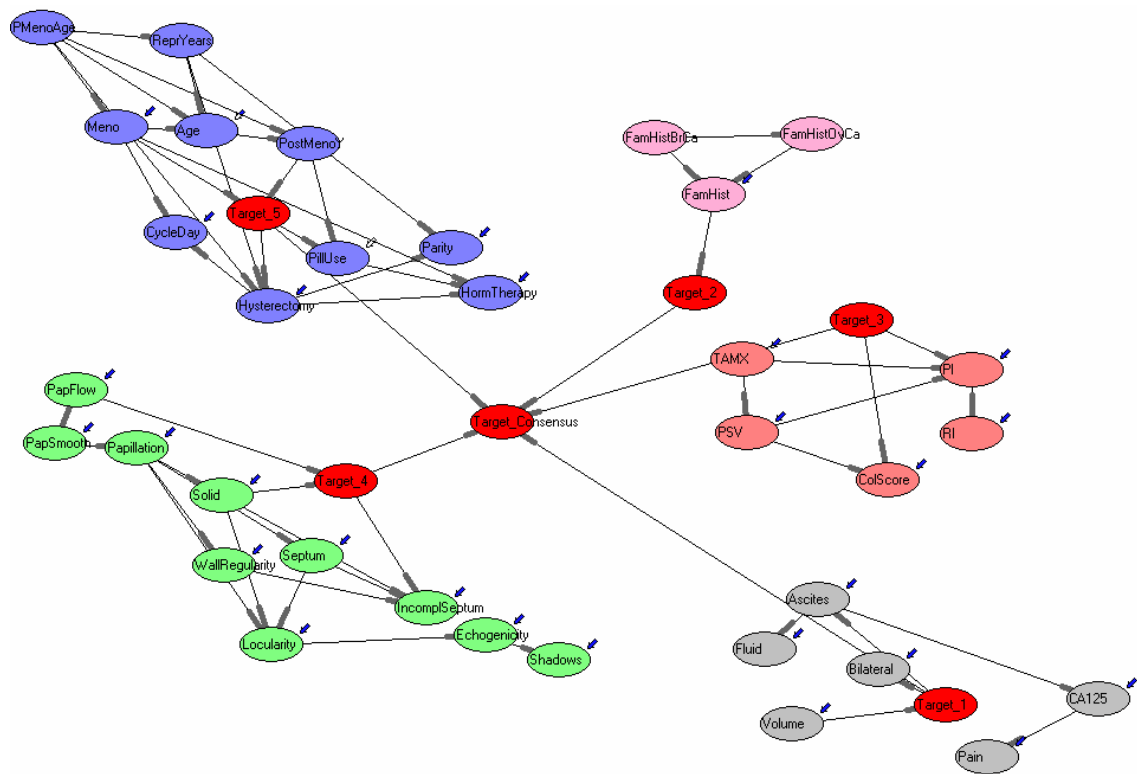


Figure 2 Evidenciák külön kezelése, különálló döntések, majd azok aggregálása.

A laborgyakorlat folyamán az utóbbi esettel foglalkozunk. Feltételezzük, hogy a szakértők döntéseiket szekvenciálisan hozzák meg, és az egyeztetés is külön, külön történik, azaz minden szakértő csupán a többi szakértő által korábban hozott döntéseket ismeri. Feltesszük továbbá, hogy minden szakértőnek csak a legutolsó döntése érhető el. Ezt úgy kell érteni, hogy a döntéshozás körökre van osztva (lásd. alábbi ábra).

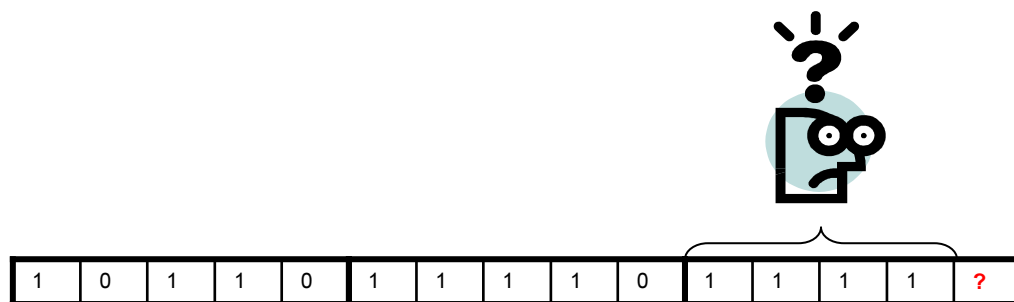


Figure 3. Az 5-ös számú ágens 3-dik körben aggregáláshoz használt bemenete

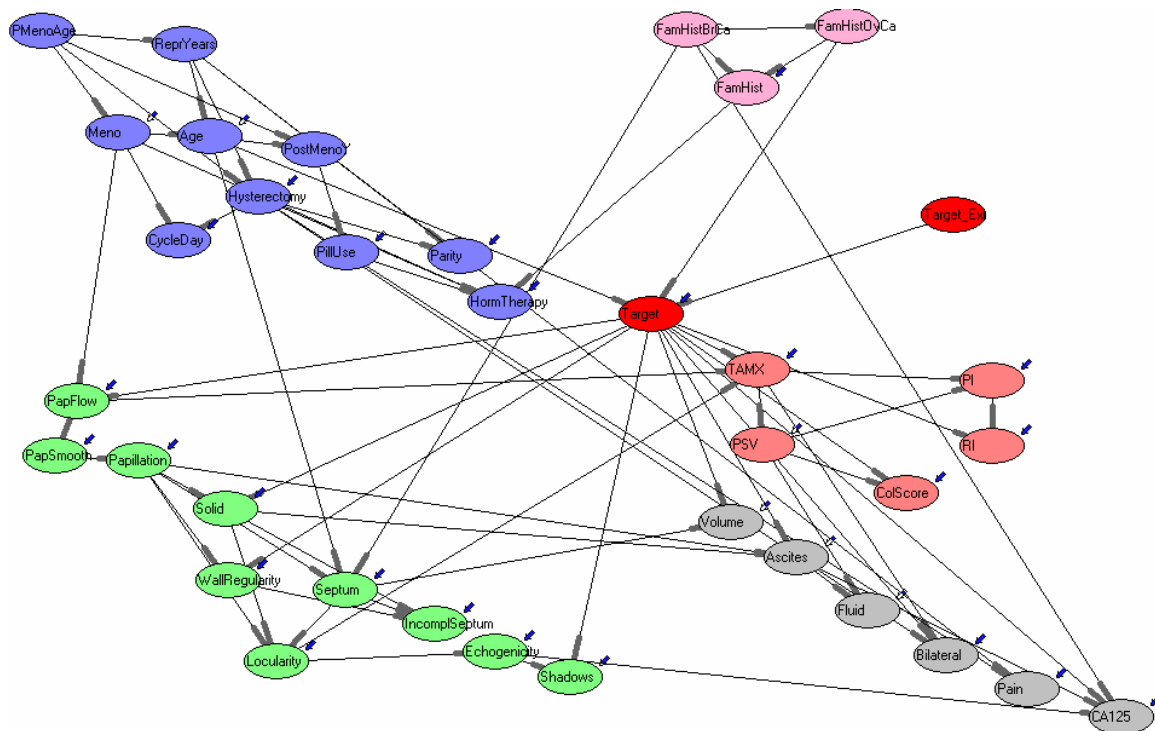


Figure 4 Evidenciák külön kezelése, különálló döntések, majd azok aggregálása és figyelembe vétele.

7. Függelék

a. A BNet szoftver használata

1.1. Létező modell megnyitása

- 'File|Open'
- A felugró fájlkiválasztó menüben válasszuk ki a kívánt (*.bn) fájlt (lásd még a mappa ikonos gombot az alsó menüsorban)

1.2. Bayes-hálók szerkesztése

- Az egyes változókat (változócsoportokat) érintő műveletek dialógusablaka ('Variable Definition' ablak) a változón történő dupla bal kattintással, vagy a szerkesztővásznon történő jobb klikk után a felbukkanó menüben a 'Define|Variables', 'Define|Values' választással hívható elő
- Az ablakban végrehajtott műveletek csak az ablak alján lévő 'OK' gomb megnyomásával véglegesítődnek

Új csoport definiálása

- A kívánt nevet beírjuk a csoport keret 'edit box'-ába, majd az 'Add' gomb megnyomásával véglegesítjük a műveletet

Új változó felvétele

- Mint a csoportdefiníció esetén, csak a 'Variables' keretben
- Ha akarjuk a csoportlistából választva megadhatjuk a változó csoportját

Új értékek felvétele egy változóhoz

- 'Variable Definition' ablak 'Variables' keretében válasszuk ki a változót
- A 'Values' keretben töltsük ki a szövegbeviteli mezőket (balról jobbra a jelentésük: érték neve, érték alsó határa, érték felső határa)
- Az érték listába vételéhez nyomjuk meg az 'Add' gombot (így balról zárt, jobbról nyílt intervallumokat adhatunk meg)

Változó csoportjának megváltoztatása

- Válasszunk egy változót a változók listájából és egy csoportot a csoportlistából

Változó mozgatása a szerkesztővásznon

- A bal gomb nyomva tartása mellett húzzuk a változót a kívánt helyre

Irányított él hozzáadása, elvétele

- A 'CTRL' gomb nyomva tartása mellett a bal egérgombbal húzzunk élt a szülőtől a gyerekig. Az él elvételéhez ugyanezt ismételjük meg még egyszer.

'CPT' (feltételes valószínűségi tábla) módosítása

- A jobb klikkre felugró menüben válasszuk a 'Define|CondProb' pontot
- A 'Probability specification' dialógusablakban a 'Variable' keretben válasszuk ki az adott (gyermek) csomópontot
- A 'Condition (configuration)' keretben állítsuk be a módosítani kívánt táblabejegyzéshez tartozó szülői konfigurációt (a bal oldali listában egy szülőt kijelölve választhatunk annak értéket a jobb listából)
- A 'Point specification of conditional probabilities' keretben válasszuk ki 'TypeScope'-nak először a bal listából 'All'-t, majd a jobb listából 'Table'-t (ezzel a csomópont minden bejegyzésének típusát egyszerű 'pontértékre' állítottuk)
- A keret felső listájában válasszuk ki a gyermek csomópont megfelelő értékét, majd a 'Specification' mezőbe írjuk be a kívánt értéket és nyomjunk 'Enter'-t (ha végeztünk az összes beállítással ne felejtsünk el 'OK'-t nyomni)

1.3. Következtetés (alap)

- Hívjuk fel az 'Inference|Inference' menüpontból elérhető 'Expression specification' dialógusablakot

- **Állítsuk be az ismert értékeket:**

- Válasszuk ki a változót a 'Variables:' keret listájából
- Válasszuk ki a változó értékét a 'Values:' keret listájából
- Nyomjuk meg a 'Hard' gombot a 'Virtual evidence' keretben

- **Végezzük el a következtetést:**

- Nyomjuk meg a 'Values:' keret 'Update' gombját
- Az egyes változók a posteriori eloszlásait megtekinthetjük, a változó kiválasztásával a 'Variables:' listából (az értékek a 'Values:' lista 'P:*' alakú elemei)

- **Ismert értékek/megfigyelések visszavonása**

- Válasszuk ki a változót a 'Variables:' keret listájából
- Válasszuk ki a változó értékét a 'Values:' keret listájából
- A 'Virtual evidence' keret 'Uni' gombjának megnyomásával töröljük az ismeretet

Expression specification

Variables:

Pathology	-	MP:Unknown
WallRegularity	-	MP:Unknown
Meno	-	MP:Unknown
Ascites	-	MP:Unknown
PapSmooth	-	MP:Unknown
Locularity	-	MP:Unknown
CA125	<35(0..35.)	MP:Unknown
ColScore	-	MP:Unknown
PillUse	<Halfyear(0..0.5	MP:Unknown
Bilateral	-	MP:Unknown
Volume	-	MP:Unknown

Max Probability Set

☒ Variable in MP CP(MP config): Update

Values:

Benign(0..0)	E:1.0000 P:0.9152 T:N	Confidence
Malignant(1..1)	E:1.0000 P:0.0848 T:N	

Update

Collective Probability Set

☐ Value in CPS CP(Collective) Update

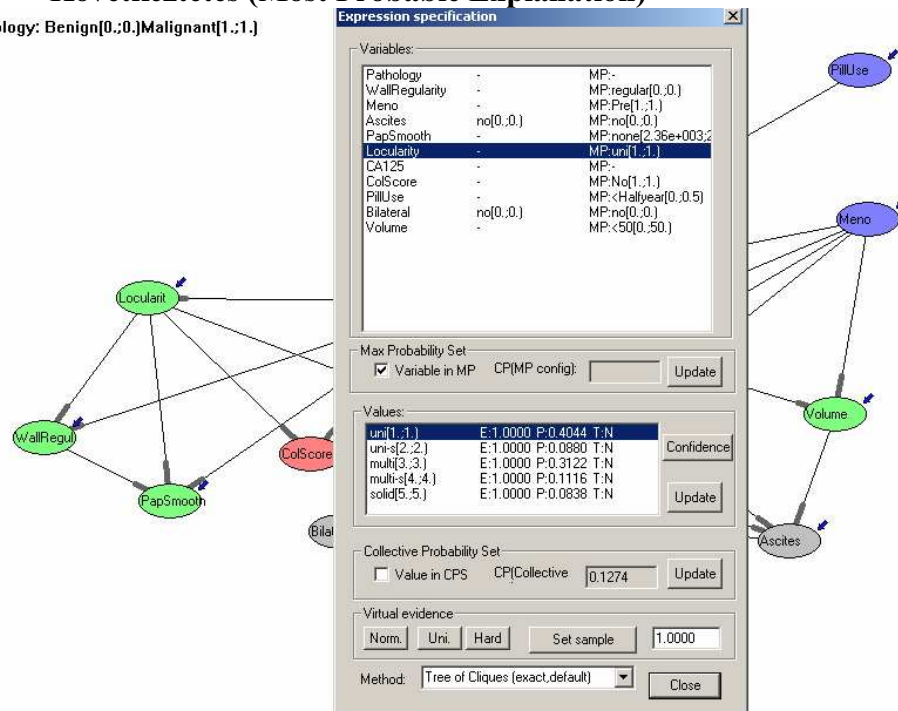
Virtual evidence

Norm. Uni. Hard Set sample 1.0000

Method: Tree of Cliques (exact,default) Close

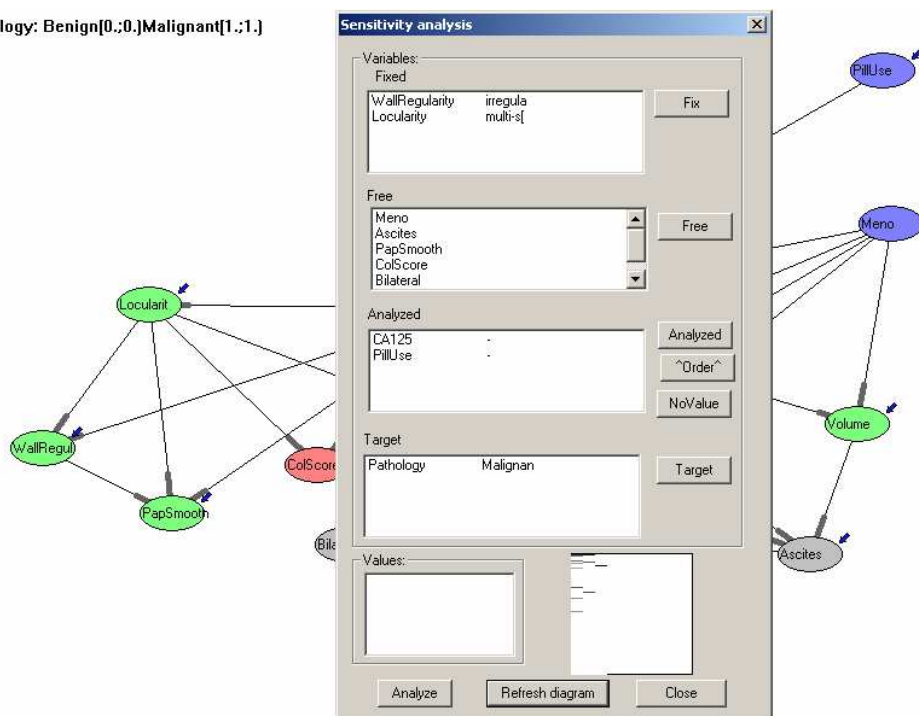
1.4. Következtetés (Most Probable Explanation)

Pathology: Benign[0.;0.]Malignant[1.;1.]



1.5. Következtetés (Érzékenység vizsgálat)

Pathology: Benign[0.;0.]Malignant[1.;1.]



Paramétertanulás

A paramétertanuláshoz szükséges funkciókat a Főmenü 'Learn Parameters' pontjába csoportosítottuk, a tanulás lépései:

- Mintafájl megnyitása: **'Open sample'**
Az állománynak a 'csv' (comma separated vectors) formátumot kell követnie, azaz egy-egy sor felel meg a különböző megfigyeléseknek, melyeken belül az egyes értékeket vesszővel kell egymástól elválasztani. A fájl első sora a fejléc, itt a változók nevei szerepelnek (érdemes azonos neveket használni, mint a modellben, így a program automatikusan el tudja végezni a csomópontok és az oszlopok összerendelését – lásd a következő pontot).
A lépés után a mintafájl betöltődik a program „munkamemóriájába”.
- Modellcsomópontok és mintaoszlopok egymáshoz rendelése: **'Attach fields to nodes'**
Egyszerű „adminisztratív” lépés: a dialógusablakban meg kell adnunk, a modell melyik csomópontja a mintafájl melyik oszlopából kapja az értékeit.
- **'Convert to IIC sample'**: a szöveges adatfájlt egy bináris formára konvertálja.
- A tanulás maga: **'Learn Parameters'**
Ez a következő lépésekből áll:
 - Uniform paraméterprior beállítása: A hálópármeterek felett egyenletes a priori paramétereloszlást állít be az úgynevezett Dirichlet eloszlás felhasználásával (lásd MIHF-sillabusz) 1-es értékűre állítva az összes virtuális mintaszámot, így a tanulás végeredménye ezen pszeudó-mintaszámlálóktól és az adatoktól függ majd.
 - A posteriori paramétereloszlás ($P(\theta|G,D)$) számítása a hiperparaméterek frissítésével.
 - Paraméterek pontbecslése: Minden paraméterre kiszámítja annak várható értékét az a posteriori eloszlásából.

b. BNet modellek használata a JADE keretrendszerben

A **JADE** (**J**ava **A**gent **D**evelopment framework) keretrendszer, mint ismeretes, multi-ágens rendszerek fejlesztésére szolgál. A rendszer résztvevői, az egyes ágensek többféleképpen dönthetnek arról, hogy milyen cselekvéseket hajtsanak végre (pl. miként módosítsák környezetük állapotát, vagy épp milyen típusú és tartalmú üzeneteket küldjenek a többi ágens felé). A döntéshozásban, adott esetben szakértői modulok támogathatják őket (pl. szabályalapú szakértői rendszerek, tervekészítő rendszerek, neurális hálózatok, vagy épp Bayes háló alapú szakértői következtető rendszerek).

Esetükben, a jelen segédlethez kapcsolódó laborgyakorlat során olyan JADE-es ágensekkel fogunk megismerkedni, melyek döntéshozásában a Bayes hálós szakértői következtetés játszik szerepet. Ilyen ágensekről már az 5. fejezetben is olvashattunk. Ebben a szakaszban tehát a JADE-es megvalósítás fontosabb részleteit fogjuk tárgyalni.

A szcenárió röviden a következő: adottak orvosok, és páciensek. A páciensek adott problémákkal (szimptómákkal) fordulnak az orvosokhoz. Az orvosok diagnosztizálják a kapott „leleteket”, és a diagnózist visszaküldik a páciensnek. Mi több, az orvosok képesek adott esetben más orvosok diagnózisát is figyelembe venni. Ehhez a páciens ágensektől kaphatnak információt. A páciens ágensek ugyanis, amennyiben több orvost is ismernek, úgy sorra mindegyiktől kérnek egy-egy diagnózist, és az adott időpontig beérkezett diagnózisokat már a következő orvosnak is továbbítják. Ezt a felállást (az orvos információit) szemlélteti a 6. fejezet 2. ábrája.

Jelen esetben tehát az orvosok Bayes hálók felhasználásával próbálnak meg döntésre jutni a páciensek diagnózisát illetően. Bayes hálójukban a következtetéshez 3 dolgot kell ismerniük: **(1)** a szimptómák értékét, mint evidenciát (igaz, hamis, ismeretlen), **(2)** a célváltozót, aminek adott értékének (jelenleg az IGAZ értéknek) a bekövetkezési valószínűségét kell megjósolni, és esetleg **(3)** az eddigi, célváltozóra vonatkozó becsléseket. Ebből az (1)-es és (3)-as információt a páciensektől kapják. A (2)-es információ pedig, azaz a célváltozó kiemelten szerepel Bayes hálójukban.

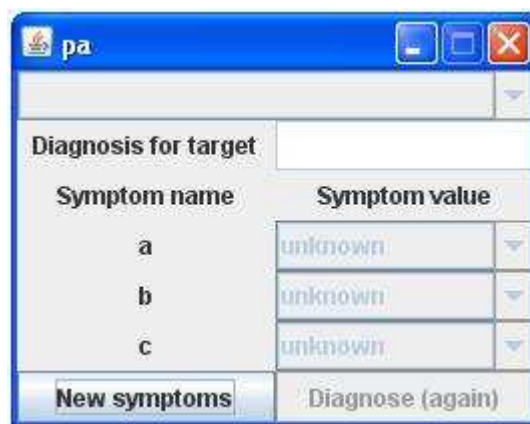
Az elosztott következtetési algoritmus (orvos-páciens dialógusra lebontva) tehát a következő:⁵

1. A páciens ágens felderíti, hogy milyen orvos ágensek vannak a rendszerben.
2. A páciens ágens sorra elküldi nekik a szimptómáinak az értékét (igaz, hamis, ismeretlen), és az eddig kapott, célváltozóra vonatkozó diagnózisokat. Ez utóbbiak kezdetben nyilván még mind „ismeretlen” értékűek. Természetesen a páciens ágens mindig bevárja az adott orvos diagnózisát, mielőtt a következőnek elküldené a „leleteit” (a legfrissebb diagnózisokkal együtt).

⁵ Természetesen tetszőleges számú orvos, és páciens lehet a rendszerben.

3. Az *i.* orvos ágens tehát megkapja a páciens ágenstől (1)-et és (3)-at. (3) az addigi (*i-1*) darab diagnózist, és (*n-i+1*) darab „ismeretlen” tartalmaz. Természetesen a diagnózisok közt „ismeretlen” érték is lehet.
4. Az *i.* orvos ágens aggregálja (3)-at, és ezt – amennyiben az aggregálás nem „ismeretlennek” adódott – evidenciaként, (1)-el együtt beilleszti a Bayes hálójába. A szimptómák közül is nyilván csak azok tekinthetők evidenciának, és azok kerülnek beillesztésre, melyek értéke nem „ismeretlen”.
5. Az *i.* orvos ágens ezek után becslést ad arra, hogy (2) mekkora valószínűséggel IGAZ.
6. A kapott becslést az *i.* orvos ágens visszaküldi a páciens ágensnek.
7. A páciens ágens ekkor vagy új értéket ad a szimptómáinak, és az 1. lépésnél folytatja, vagy rögvest 2. lépésnél folytatja, de ott (3)-ban a legutóbbi körben adódott diagnózisokat küldi el.

Mindez JADE-ben a következőképp néz ki: a páciens ágensek rendelkeznek egy GUI-val (lásd. pl. a következő ábrát). Ezen szerepel a „New Symptoms” gomb. Ezt lenyomva az emberi felhasználó megadhatja a szimptómáinak az értékét, majd a „Diagnose (again)” feliratú gombra kattintva elindíthatja az elosztott következtetést. A következtetés végén, miután mindegyik ágenstől beérkezett a diagnózis, aktiválódik a GUI felső részében található ComboBox, amelyből kiválasztható, hogy mely orvos diagnózisára vagyunk kíváncsiak. Ennek értéke ekkor nyomban megjelenik a ComboBox alatti TextBox-ban.



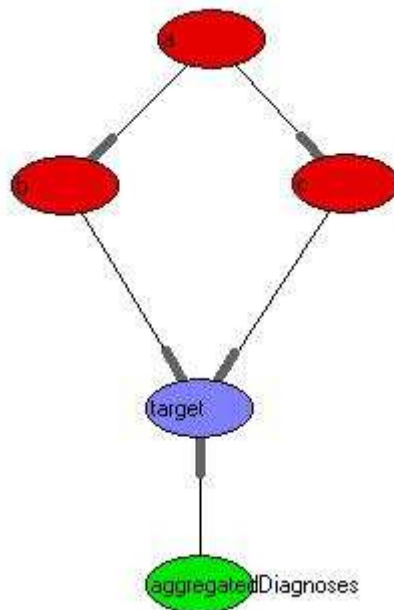
Az ábrán tehát egy „target” nevezetű célváltozó, és három, „a”, „b”, és „c” nevű szimptóma esetén láthatjuk a páciens ágens GUI-ját induláskor.

Miután kigyöngyösködtük magunkat az orvosoktól kapott diagnózisokban, a fentebb leírt algoritmus 7. lépésének megfelelően 2 dolgot tehetünk: **(1)** vagy rákattintunk a „New Symptoms” gombra, és akkor minden kezdődik előlről, vagy **(2)** újra rákattintunk a „Diagnose (again)” gombra, és akkor ennek hatására az eredetileg megadott szimptómákkal, és az utóbbi körben beérkezett diagnózisokkal (azokat lépésről lépésre frissítve) kezdjük el bombázni az orvos ágenseket.

Az orvos ágens is rendelkezik egy GUI-val. Ezen pusztán csak azt lehet beállítani, hogy a beérkező szimptómákból adott esetben (ha nem „ismeretlen” értékűek) melyeket vegyük figyelembe.



Az orvosok tehát rendelkeznek egy Bayes hálóval is. Az előbbi példánál maradva, a Bayes háló például a következő (BNet-ben).



Az ábrán az eddig említett változókon felül még egy „aggregatedDiagnoses” nevezetű változó is látható. Ez szolgál arra, hogy az orvos evidenciaként képes legyen beilleszteni a hálójába az addigi diagnózisokat.

Ebben a formában természetesen egy JADE-es orvos ágens nem tudja felhasználni a Bayes hálót. Ezért először a hálót az orvos ágens által olvasható/értelmezhető formára – jelen esetben XML formára. Ehhez a kattintsunk BNet-ben a „File” menü „Export to Pythia XML” menüpontjára.

A kiexportált hálót érdemes oda menteni, ahol maga az orvos ágens is van (a `\jade\src\milab\lab04\DoctorAgent` könyvtárba). A fentebbi hálót például `example01_bnet.xml` néven, illetve eredetiben, `example01.bn` néven megtalálhatjuk az említett könyvtárban.

Felmerülhet viszont a kérdés, hogy: honnan tudja az orvos ágens, illetve a páciens ágens, hogy mi a szimptómák neve, mi a célváltozó, stb? A válasz igen egyszerű: a rugalmasság kedvéért ezt egy konfigurációs fájl tárolja, amelyet az ágensek újrafordítása nélkül változtathatunk épp úgy, mint ahogyan a Bayes hálót is. Változtatás után viszont az ágenseket újra kell indítani (mivel induláskor olvassák fel a hálót és/vagy a konfigurációs fájlt).

A fenti hálónak megfelelő XML-en túl, amelyet egy JADE-es orvos ágens induláskor felolvas, még szükség van tehát az említett konfigurációs fájlra is. Ez példánkban a `\jade\src\milab\lab04\example01_param.txt` elérésen található. Tartalma a következő:

```
target      a      b      c
```

Látható tehát, hogy egy szöveges fájlról van szó, melyben karakterfüzerek szerepelnek tab-okkal elválasztva. Az első string a célváltozó neve, a többi pedig sorban a szimptómák neve.

A JADE-es páciens ágens (`milab.lab04.PatientAgent.PatientAgent`) indításakor tehát paraméterként meg kell adni azt a konfigurációs (avagy paraméter-leíró) fájlt, amely tartalmazza a páciens lehetséges szimptómáinak felsorolását, és a célváltozó nevét, aminek kapcsán a diagnózisra (hogy mi a valószínűsége annak, hogy IGAZ) kíváncsi. Ezek szerint példánkban a következő paraméterrel kellene indítanunk a `milab.lab04.PatientAgent.PatientAgent` ágenszt:

```
\jade\src\milab\lab04\example01_param.txt
```

A JADE-es orvos ágensnek (`milab.lab04.DoctorAgent.DoctorAgent`) viszont már 2 paraméterre is szüksége van induláskor: **(1)** kiexportált Bayes hálójának elérése, és **(2)** a Bayes hálóban szereplő változók/csomópontok, és a célváltozó nevét tartalmazó konfigurációs fájl elérésére. Ezek szerint példánkban a következő 2 paraméterrel kellene indítanunk a `milab.lab04.DoctorAgent.DoctorAgent` ágenszt:

```
\jade\src\milab\lab04\DoctorAgent\example01_bnet.xml
\jade\src\milab\lab04\example01_param.txt
```

A JADE-es ágensek paramétereit, amennyiben több is van, szóközzel kell elválasztani. Továbbá, részletkérdésnek tűnhet, de mégis sok galibát okozhat az orvos és a páciens konfigurációs fájljainak eltérése. Célszerű tehát, ha a konfigurációs fájl mindkét típusú ágensnél egy és ugyanaz. A laborgyakorlat során is próbáljuk meg ehhez tartani magunkat.

A paraméterezést követően a JADE-es ágensek elindulnak, majd a fentebb leírtaknak megfelelően működésbe lépnek. Vessünk még egy pillantást arra, hogy milyen tartalmú/felépítésű üzenetek cserélnek gazdát közöttük. Példánkban a páciens a következő – \jade\src\milab\lab04\PatientAgent\language01\language01.xsd sémának megfelelő – tartalmú REQUEST üzenetet küldi az orvosnak:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<message>
  <symptoms>1 1 1</symptoms>
  <diagnoses>x</diagnoses>
</message>
```

Az üzenetből látszik, hogy az „a”, „b”, és „c” szimptómák a páciens szerint rendre mind igazak, továbbá egyetlen orvos ágenssel kommunikál(t), és az éppen az, akinek most ezt az üzenetet küldi, és akinek az eddigi diagnózisa „x”, azaz ismeretlen értékű.

Az orvos ágens erre példánkban a következő – \jade\src\milab\lab04\DoctorAgent\language01\language01.xsd sémának megfelelő – tartalmú INFORM üzenettel válaszol a páciensnek:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<diagnose>0.500000</diagnose>
```

Ezek szerint az orvos szerint $\frac{1}{2}$ annal a valószínűsége, hogy a „target” nevű, bináris célváltozó értéke „1”, azaz IGAZ lesz. Ez akkor sem változik, ha a páciens újra diagnózist kér (immár a beérkezett diagnózis fényében), azaz a következő üzenetet küldi az orvosnak:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<message>
  <symptoms>1 1 1</symptoms>
  <diagnoses>0.500000</diagnoses>
</message>
```

Az orvos erre is fentebbi üzenettel válaszol. Miért? Mert hiába használja fel ezt az előbbi diagnózist arra, hogy hálójában az aggregatedDiagnoses nevezetű csomópontnak evidenciaképp értéket adjon, a háló olyan felépítésű, és a benne szereplő valószínűségek olyanok, hogy ez a tény nem változtat a célváltozó eloszlásán.

Elképzelhető azonban, hogy az eddigi diagnózisok másfajta aggregálása mellett, és így másmilyen evidenciával látva el az aggregatedDiagnoses nevezetű csomópontot, más eredmény jönne ki. Példánkban ez éppenséggel nem így van, de általában elképzelhető, sőt, igen gyakori lehet (attól függően, hogy az aggregatedDiagnoses csomópont mennyire meghatározó a célváltozóra nézve).

A páciensről kapott, eddig beérkezett diagnózisok aggregálását a DoctorAgent osztály aggregateDiagnoses metódusa végzi. Tehát ennek átírásával tovább finomíthatjuk JADE-es orvos ágensünk döntéshozását.

8. Irodalomjegyzék

- (Antal2002a) P. Antal, T. Meszaros, D. Timmerman, B. De Moor, T. Dobrowiecki: **Domain knowledge based information retrieval langugae: an application of annotated Bayesian networks**, Fifteenth IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), June 3-7, Maribor, Slovenia, pp 213-218
- (Antal2002b) P. Antal, P. Glenisson, G. Fannes, Y. Moreau, B. De Moor: **On the potential of domain literature for clustering and Bayesian network learning**, The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD), 2002, Edmonton, pp 405-414
- (Antal2003) P. Antal, G. Fannes, D. Timmerman, Y. Moreau, B. De Moor: **Bayesian Applications of Belief Networks and Multilayer Perceptrons for Ovarian Tumor Classification with Rejection**, Artificial Intelligence in Medicine, vol. 29, pp 39-60, 2003
- (Antal2004a) P. Antal, G. Fannes, Y. Moreau, D. Timmerman, B. De Moor: **Using Literature and Data to Learn Bayesian Networks as Clinical Models of Ovarian Tumors**, Artificial Intelligence in Medicine, 2004, vol 30, pp 257-281
- (Antal2004b) P. Antal, A. Millinghoffer: **Learning Causal Bayesian Networks from Literature Data** In Proc. of the 3rd Int. Conf. on Global Research and Education, Inter-Academia'04, Budapest, Hungary, 6-9. September 2004.
- (Bernardo95) J. M. Bernardo: **Bayesian Theory**, Wiley & Sons, Chichester, 1995.
- (Bishop95) C. M. Bishop: **Neural Networks for Pattern Recognition**, Clarendon Press, Oxford, 1995.
- (Cooper92) G. F. Cooper, E. Herskovits: **A Bayesian Method for the Induction of Probabilistic Networks from Data**, Machine Learning, vol 9, pp 309-347, 1992.
- (Cowell1999) R. G. Cowell, A. P. Dawid, S. L. Lauritzen, D. J. Spiegelhalter: **Probabilistic Networks and Expert Systems**, Springer Verlag, 1999.
- (Friedman96) N. Friedman, Z. Yakhini: **On the Sample Complexity of Learning Bayesian Networks**, Proc. of the 12th Conf. on Uncertainty in Artificial Intelligence, pp 274-282, 1996.
- (Friedman97) N. Friedman: **Learning Belief Networks in the Presence of Missing Values and Hidden Variables**, Proc of the 14th international conference on machine learning, pp 125-133, 1997.
- (Friedman2002) N. Friedman and D. Koller: **Being Bayesian about Network Structure**, Journal of Machine Learning Research, vol. 2, pp 1-30, Kluwer Academic Publ., Dordrecht, Netherlands, 2002
- (Gelman1995) A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin: **Bayesian Data Analysis**, Chapman & Hall, London, 1995.
- (Gilks1995) W. R. Gilks, S. Richardson, D. J. Spiegelhalter (editors): **Markov Chain Monte Carlo in Practice**, Chapman & Hall, 1995.
- (Glymour1999) C. Glymour, G. F. Cooper: **Computation, Causation, and Discovery**, AAAI Press, 1999.
- (Heckerman1995a) D. Heckerman, D. Geiger, D. Chickering: **Learning Bayesian networks: The Combination of Knowledge and Statistical Data**, Machine Learning, vol 20, pp 197-243, 1995.
- (Heckerman1995b) D. Heckerman, A. Mamdani, M.P. Wellman: **Real-world applications of Bayesian networks**, Communications of the ACM, Vol. 38, issue 3, p. 24-26, ACM Press, New York, 1995.
- (Horvitz1998) E. Horvitz, J. Breese, D. Heckerman, D. Hovel, K. Rommelse: **The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users**, Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (1998) July. 24-26 1998, Madison, Wisconsin, USA, p. 256-265.

- (Huang1996) C. Huang, A. Darwiche: **Inference in Belief Networks: A procedural guide**, International Journal of Approximate Reasoning, vol 15, pp 225-263, 1996.
- (Hullám2005) Hullám Gábor: **Bayes-hálók strukturális tulajdonságainak tanulása kényszer alapú módszerekkel**, diplomamunka BME-MIT, 2005.
- (Koller1998) D. Koller, A. Pfeffer, **Probabilistic Frame-Based Systems**, In Proc. of the 15th National Conference on Artificial Intelligence (AAAI), Madison, Wisconsin, 1998., pp 580-587
- (Laskey1997) K. Laskey, S. Mahoney: **Network Fragments: Representing Knowledge for Constructing Probabilistic Models**, In Proc. of the 13th Conf. on Uncertainty in Artificial Intelligence (UAI-1997), Morgan Kaufmann, 1997., pp 334-341
- (Myllymaki2001) P. Myllymaki, T. Silander, H. Tirri, P. Uronen: **Bayesian Data Mining on the Web with B-Course**, In Proceedings of The 2001 IEEE International Conference on Data Mining, p.626-629, IEEE Computer Society Press, 2001.
- (Pearl1988) J. Pearl: **Probabilistic Reasoning in Intelligent Systems**, Morgan Kaufmann, San Francisco, 1988.
- (Przytula2004) K. W. Przytula, T.Lu: **Bayesian Networks Based Diagnostic Tools for Locomotives: Model Development and Inference**, The Second Bayesian Modeling Applications Workshop During UAI-04, July 7th, 2004.
- (Ramamurthy) S. Ramamurthy, H. Arora, A. Ghosh : **Operational Risk and Probabilistic Networks – An Application to Corporate Actions Processing**, Banking & Capital Markets Solutions Consulting, Infosys Technologies Limited, <http://www.hugin.com/cases/Finance/Infosys/oprisk.article>
- (Robert2001) C. P. Robert: **The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation**, Springer-Verlag 2001
- (Sahami1998) M. Sahami, S. Dumais, D.Heckermann, E. Horvitz: **A Bayesian Approach to Filtering Junk E-mail**, AAI Workshop on Learning for Text Categorization July 1998, Madison, Wisconsin, AAAI Technical report WS-98-05.
- (Skaanning1998) C. Skaanning, F.V. Jensen, U. Kjærulff, L. Parker, P. Pelletier, L. Rostrup-Jensen: **SACSO - A Bayesian-Network Tool for Automated Diagnosis of Printing Systems**, Machine Intelligence Group, Aalborg University, Technical report, 1998.
- (Tchangani2002) A.P. Tchangani : **Decision Support System with Uncertain Data: Bayesian Networks Approach**, http://www.ici.ro/ici/revista/sic2002_3/art2.htm
- (Timmeman2000) D. Timmerman, L. Valentin, T. H. Bourne, W. P. Collins, H. Verrelst, I. Vergote: **Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) Group**, Ultrasound Obstetrics Gynecology, vol 16, pp500-505, 2000.
- (Wright2004) E.Wright, J. Fitzgerald, D. Barbara, T. Shakelford, K. Laskey, G. Alghamdi, X. Wang: **Detecting insider threats in information systems**, The Second Bayesian Modeling Applications Workshop During UAI-04, July 7th, 2004.