

# BayesCube felhasználói kézikönyv

## Tartalomjegyzék

<b>1. Bayes-háló modellek szerkesztése</b>	<b>2</b>
1.1. Új modell létrehozása	2
1.2. Létező modell megnyitása	2
1.3. Modellstruktúra szerkesztése	2
1.4. Csomópontok típusainak szerkesztése	2
1.5. Csomópontok csoportjainak szerkesztése	3
1.6. Csomópontok szerkesztése	4
1.6.1. Valószínűségi csomópont - feltételes valószínűségi tábla	4
1.6.2. Valószínűségi csomópont - döntési fás	5
1.6.3. Döntési csomópont	6
1.6.4. Hasznossági csomópont	6
1.6.5. Hasznossági csomópont - döntési fás	7
1.7. Dinamikus Bayes-hálók	7
1.7.1. Homogén lánc	7
1.7.2. Összetett dinamikus Bayes-háló	7
1.8. Adatfájlok formátuma	8
1.9. Paraméterek beállítása adatfájl alapján	8
<b>2. Strukturális információk megjelenítése</b>	<b>10</b>
2.1. Élek megjelenítése	10
2.2. Húrgráf megjelenítése	11
2.3. Klikkfa megjelenítése	11
<b>3. Következtetés Bayes-hálóban</b>	<b>12</b>
3.1. Evidenciák/döntések bevitele	12
3.2. Marginális eloszlások számítása	12
3.2.1. A mintavételi nézet (Sampling view)	12
3.3. Következtetés érzékenységének vizsgálata	13
3.3.1. A „Sensitivity of Inference Diagram” értelmezése	14
3.4. Adatfájl alapján történő score-számítás	15
<b>4. Annotáció és magyarázatkeresés</b>	<b>16</b>
<b>5. Hatáserősség</b>	<b>18</b>
<b>6. Becslések torzítása</b>	<b>19</b>

## 1. Bayes-háló modellek szerkesztése

### 1.1. Új modell létrehozása

Új modellt a FILE > NEW ... menü vagy az erre szolgáló ikon kiválasztásával hozhatunk létre. Az ekkor felugró fájlkiválasztó dialógusablakban kell meghatározni az új modellt tartalmazó fájl mentési helyét, majd ha ez megtörtént, kezdhetjük a modell szerkesztését.

### 1.2. Létező modell megnyitása

Egy már korábban fájlba mentett modellt a FILE > OPEN ... menü, vagy a megfelelő ikon kiválasztásával nyithatunk meg: a felbukkanó fájlkiválasztó dialógusablakban kell kiválasztani a modell leírását tartalmazó fájl elérési útját.

### 1.3. Modellstruktúra szerkesztése

A megnyitott modellhez a szerkesztési paletta segítségével adhatunk új csomópontokat: a főablak jobb oldalán elhelyezkedő palettát kinyitva kiválaszthatjuk a hozzáadni kívánt csomópont típusát, majd ezután az egérrel a modell területére klikkelve ott megjelenik az új csomópont.

Az élek modellhez adásához hasonlóan történik: a palettáról kiválasztjuk a megfelelő ikont, majd a modellben először a forrás, majd a célcsomópontra klikkelünk.

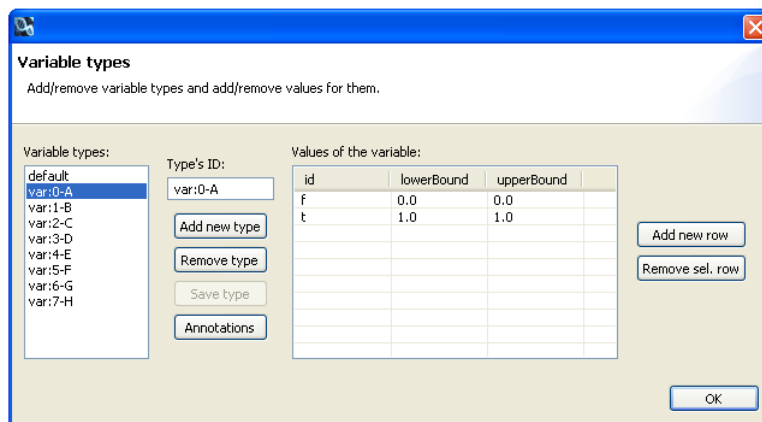
A palettán a kiválasztott (a modellhez adandó) elemet reprezentáló ikonok „beragadnak”, azaz a kiválasztás utáni első elemhozzáadás után még tetszőleges számú ugyanolyan elemet adhatunk a modellhez, anélkül, hogy a palettán való kijelölést újra meg kellene tennünk. Az alap ú.n. kiválasztási (select) üzemmódban az ESCAPE billentyű megnyomásával térhetünk vissza.

A modell egy tetszőleges elemét (csomópontot, vagy élet) törölni kattintással kijelölés után a DELETE billentyű megnyomásával lehet.

### 1.4. Csomópontok típusainak szerkesztése

Minden valószínűségi és döntési csomópont egy-egy típushoz tartozik, ez a típus határozza meg hogy hány és milyen értéket vehet fel az adott csomópont. A csoportokhoz hasonlóan a típusok listája is a változóktól függetlenül szerkeszthető, a főablak VARIABLE TYPES ... környezeti menüjén keresztül.

A menü kiválasztására felbukkanó párbeszédablakban szerkeszthetjük (1) a típusok listáját, (2) az egyes típusok által felvehető értékek listáját és az azokhoz rendelt valós intervallumokat, illetve (3) a típushoz rendelt szöveges annotációkat.

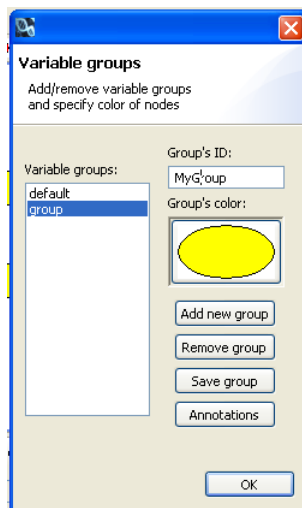


1. ábra. A változók lehetséges típusainak szerkesztésére szolgáló dialógusablak.

Alapértelmezésben minden új változó a DEFAULT típusba tartozik, amely nem szerkeszthető. A csomópontokhoz rendelt típust a csomópont tulajdonságai között szerkeszthetjük.

## 1.5. Csomópontok csoportjainak szerkesztése

A modellben található csomópontok csoportokba szervezhetők. E csoportok listája a főablak VARIABLE GROUPS ... környezeti menüjéből érhető el. A menüre kattintva a felugró dialógusablakban szerkeszthetők, a változók csoportjainak (1) nevei, (2) a csoportokhoz rendelt megjelenítési szín, és (3) a csoportokhoz tartozó szöveges annotációk listája.



2. ábra. A változói csoportok szerkesztésére szolgáló dialógusablak.

Alapértelmezésben minden változó a DEFAULT csoporthoz tartozik, ez a cso-

port nem szerkeszthető. Egy-egy változó csoporthoz rendelése a változó tulajdonságai között szerkeszthető.

## 1.6. Csomópontok szerkesztése

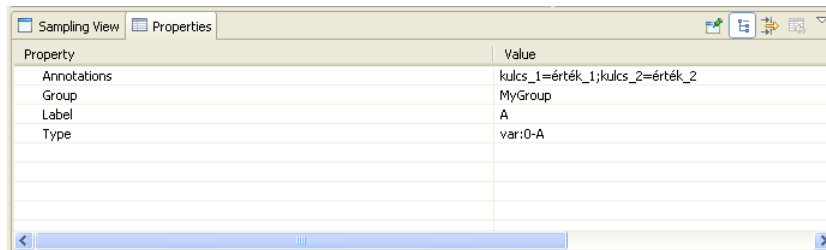
A főablak alatt elhelyezkedő PROPERTIES nézet szolgál a csomópontok alapvető információinak megjelenítésére és szerkesztésére. Bármely csomópontra ráklickeelve, azt kiválasztva a nézetben táblázatos formában megjelennek a következő információk:

**Annotations.** Szöveges annotációk listája, egyetlen szövegfűzérbe fűzve. Az annotációk (kulcs, érték) alakú párosok, szerkesztésük egy külön dialógusablakban lehetséges, amely a nézet ANNOTATIONS sorát kiválasztva annak ... gombjára kattintva érhető el.

**Group.** A változó csoportja. A már definiált csoportok közül választható ki, hogy az adott változó melyikbe tartozzon. Egy változó pontosan egy csoportnak a tagja.

**Label.** A változó neve, ez a tulajdonság magában a főablakban is szerkeszthető, a változót kiválasztva, majd annak nevére még egyszer klikkelve.

**Type.** A változó típusa. Hasonlóan a csoporthoz, a már definiált típusok közül választható pontosan egy. Ez a tulajdonság csak a valószínűségi és döntési csomópontok esetén érhető el, hasznosságcsomópontok esetén nem.



3. ábra. A csomópontok alapvető tulajdonságait megjelenítő/szerkesztő nézet.

A csomópontok által leírt valószínűségi/hasznossági függvények a csomópontra duplán klikkelve elérhető dialógusablakokban szerkeszthetők. Ezek leírása a következő alfejezetekben található.

### 1.6.1. Valószínűségi csomópont - feltételes valószínűségi tábla

A valószínűségi csomópontok egy-egy diszkrét, véges értékészletű valószínűségi változót reprezentálnak. Ezek legfontosabb jellemzője az általuk reprezentált feltételes valószínűségi eloszlás, azaz, hogy szülei adott értékkonfigurációi mellett milyen eloszlást vesznek fel lehetséges értékeik felett.

A feltételes valószínűségi táblás (conditional probability table – CPT) csomópontok ezt a lehető legegyszerűbb módon, a valószínűségi értékeket egy táblázatban tárolva valósítják meg.

(G, E)	f	t
(f, f)	0.05	0.95
(f, t)	0.95	0.05
(t, f)	0.95	0.05
(t, t)	0.95	0.05

vvs/p  fixed row height

OK Cancel

4. ábra. Valószínűségi csomópont feltételes valószínűségi táblájának megjelenítése.

A táblázat sorai a változó egy-egy szülői konfigurációjának felelnek meg, az ezen belül található celláknak kell teljesíteniük a valószínűségi eloszlásokra vonatkozó feltételt.

### 1.6.2. Valószínűségi csomópont - döntési fás

Egy valószínűségi csomópontnak a szüleitől való feltételes függését a táblázatos forma mellett meg lehet adni egy döntési fával is. Ebben a döntési fában a belső csomópontok (azaz az elágazások) a szülőkkel lesznek felcímkézve, míg a levelek egy-egy eloszlást fognak tartalmazni.

Egy újonnan felvett csomópont alapértelmezés szerint a szülei bármely konfigurációja esetén egyenletes eloszlással rendelkezik, azaz döntési fája egyetlen (egyenletes eloszlású) levélből áll.

A döntési fát a következő műveletekkel építhetjük fel.

**Új elágazás hozzáadása.** A szerkesztő jobb oldalán található palettából egy kattintással kiválasztjuk az egyik szülő csomópontot, majd egy második kattintással kiválasztjuk azt a levelet, amely elé be szeretnénk illeszteni a kiválasztott szülő szerinti elágazást (A rendszer ellenőrzi, hogy az adott szülő elhelyezhető-e a fa kiválasztott részén, és ha ez nem lehetséges, nem ajánlja fel a hozzáadás lehetőségét).

Ekkor a levélcsoomópont a fában lecserélődik egy részfára, amely az adott szülő szerinti döntési csomópontot, és az ezen „lógó” annyi levél-eloszlást tartalmaz, ahány értéke a szülőnek van.

**Részfa törlése.** Egy döntési csomópontot kijelölve, majd ezután a DELETE billentyűt megnyomva, az adott csomópont, és az alatta levő teljes részfa törlődik, a helyére pedig egyetlen, egyenletes eloszlású levél kerül.

**Levél-eloszlás szerkesztése.** Egy levélcsoomópontra duplán kattintva a CPT-szerkesztőhöz hasonló dialógusablak jelenik meg, amelyben a már ismert módon szerkeszthető az adott levélhez tartozó eloszlás.

**Részfa áthelyezése.** A döntési fa egy részfája áthelyezhető egy másik levél helyére, a kérdéses csomópontot az egérrel a levél fölé húzva.

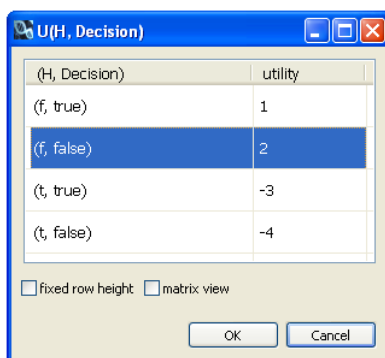
A fentiekén túl lehetőség van a fa grafikus megjelenítésének (a csomópontok elhelyezkedésének) meghatározására is: Egy vagy több kijelölt elemet az egér segítségével „arrébb húzhatunk” a szerkesztővászonon, illetve a TREE LAYOUT környezeti menü kiválasztására, a program automatikusan elrendezi a fa csomópontjait.

### 1.6.3. Döntési csomópont

A döntési csomópontok a felhasználó által (a rendszer szempontjából kívülről) hozott döntéseket reprezentálják, azaz ezek nem függhetnek a modell más csomópontjaitól, nem lehetnek szüleik. Ennek megfelelően a döntési csomópontok nem reprezentálnak semmilyen függést, így hozzájuk nem tartozik ilyen jellegű szerkesztő.

### 1.6.4. Hasznossági csomópont

A hasznossági csomópontok a valószínűségeiktől eltérően nem eloszlásokat, hanem csak hasznossági értékeket tartalmaznak, ennek megfelelően a hozzájuk tartozó, szintén táblázatos függvénytárolási forma annyiban egyszerűsödik, hogy soronként nem egy teljes eloszlást, hanem csak egyetlen valós számértéket kell megadnunk (ez természetesen lehet 1.0-nél nagyobb vagy negatív érték is).



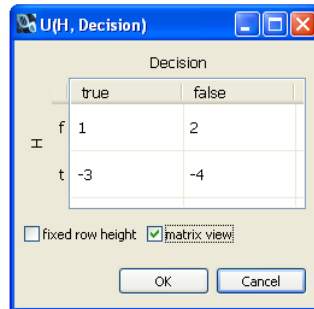
(H, Decision)	utility
(f, true)	1
(f, false)	2
(t, true)	-3
(t, false)	-4

fixed row height    matrix view

OK   Cancel

5. ábra. Hasznossági függvény táblázatos megjelenítése/szerkesztése.

A táblázat jobb áttekinthetősége érdekében 2 szülő esetén lehetőség van a táblázat mátrixos formában való megjelenítésére is: ekkor az oszlopok az egyik, a sorok a másik szülő lehetséges értékeivel vannak felcímkézve.



6. ábra. Hasznossági függvény mátrixos megjelenítése/szerkesztése.

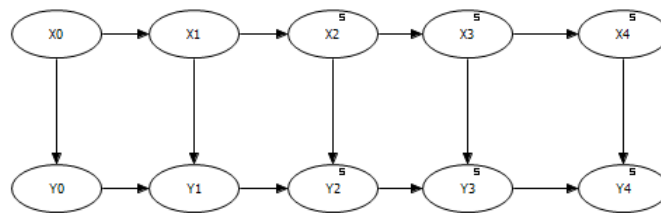
### 1.6.5. Hasznossági csomópont - döntési fás

Hasonlóképpen a valószínűségi csomópontokhoz, a hasznossági csomópontoknak is létezik döntési fás változata. Ezek esetén a döntési fa szerkesztése hasonlóképpen történik, mint a valószínűségi csomópontoknál, annyi eltéréssel, hogy a levélcsomópontok ebben az esetben egyetlen hasznosságértéket tartalmaznak, nem egy eloszlást.

## 1.7. Dinamikus Bayes-hálók

### 1.7.1. Homogén lánc

Feltételes valószínűség táblát használó valószínűségi csomópontok segítségével létre lehet hozni homogén láncokat. Először készítsük el az első tálcát, majd jelöljük ki a csomópontokat. Kijelölés után, a főablak CREATE NEXT TEMPORAL GROUP FROM THE PREVIOUS környezeti menüjének segítségével létre lehet hozni tetszőleges számú újabb tálcát. A második tálcá csoomópontjainak eloszlásait öröklük az ezt követő tálcák, így elegendő csak a második tálcá csoomópontjainak eloszlásait szerkeszteni a korábban bemutatott módon, azok megjelennek a későbbi tálcák csoomópontjainál.

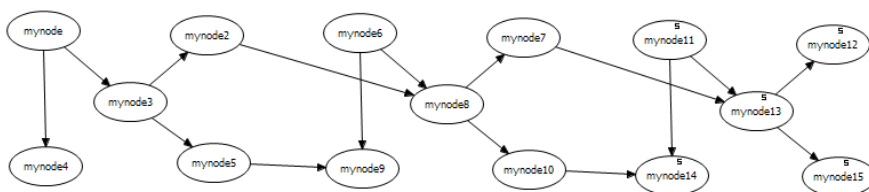


7. ábra. Homogén lánc.

### 1.7.2. Összetett dinamikus Bayes-háló

Feltételes valószínűség táblát használó valószínűségi csomópontok segítségével létre lehet hozni összetett dinamikus Bayes-hálókat. Először készítsük el az első tálcát, majd jelöljük ki a csomópontokat és jobb gomb segítségével a környezeti

menüből válasszuk a COPY SELECTED SUBGRAPH majd a PASTE SUBGRAPH menüpontokat vagy pedig a DUPLICATE SELECTED SUBGRAPH menüpontot. A két tábla között hozzuk létre a kívánt kapcsolatokat. A korábban bemutatott módon hozzunk létre kettő temporális csoportot és az első tábla csomópontjait kijelölve adjuk hozzá az első temporális csoporthoz, a másodikat a másodikhoz. Ezek után a főablak CREATE NEXT HOMOGENEOUS TEMPORAL GROUP FROM THE PREVIOUS környezeti menüjének segítségével létrehozható tetszőleges számú újabb tábla. A második tábla csomópontjainak eloszlásait öröklik az ezt követő táblák, így elegendő csak a második tábla csomópontjainak eloszlásait szerkeszteni a korábban bemutatott módon, azok megjelennek a későbbi táblák csomópontjainál.



8. ábra. Összetett dinamikus Bayes-háló.

## 1.8. Adatfájlok formátuma

A Bayes-hálókön végrehajtható számos művelet igényli egy adatfájl meglétét, amely a háló által reprezentált csomópontokra vonatkozó megfigyeléseket tartalmaz. A szoftver a vesszővel tagolt szövegfájlokat (comma-separated vectors – csv) támogatja, ezeknek a következőknek kell megfelelniük.

- A fájl egy-egy során belül a cellákat vesszőkkel kell egymástól elszeparálni.
- A fájl első sora az ún. fejléc, ebben a sorban a modellben található változók neveinek kell szerepelnie. A változók sorrendje tetszőleges lehet (nem kell megegyeznie a modellen belüli sorrenddel), de minden változónak szerepelnie kell az adatfájlból. Esetlegesen további oszlopok is lehetnek az adatfájlból, ez nem jelent problémát.
- A további sorok tartalmazzák a megfigyelési értékeket, ezeknek az egyes változók lehetséges értékei definíciójához megadott intervallumok valamelyikébe eső számoknak kell lenniük, nem egész szám esetén tizedespontot alkalmazva (nem vesszőt). A cellák az ismeretlen megfigyelési értékek jelölésére hagyhatók üresen is.

## 1.9. Paraméterek beállítása adatfájl alapján

A modellhez a fentiek szerint illeszkedő adatfájl felhasználható arra, hogy a modell numerikus paramétereit (azaz a valószínűségi csomópontok feltételes valószínűségi tábláinak bejegyzéseit) az alapján beállítsuk.

Az OPERATIONS > LEARNING PARAMETERS ... menüpontot kiválasztva, meg kell adnunk a felhasználni kívánt adatfájlt, és ez alapján lezajlik a paraméterek beállítása. Ami valójában ekkor történik, az, hogy minden egyes változóra



a szoftver megszámolja az adatfájlban a változó családjához<sup>1</sup> tartozó teljes konfigurációk előfordulási gyakoriságait és ezekkel a számokkal tölti ki a csomópont CPT-jét, majd a kérdéses CPT-t normalizálja.

---

<sup>1</sup>Egy változó családja magát a változót és annak szüleit jelenti.

## 2. Strukturális információk megjelenítése

A diszkrét csomópontokat tartalmazó Bayes-hálókból való következtetésre a szoftver a PPTC (*propagation of probabilities in trees of cliques*) algoritmust alkalmazza (részletesen lásd [?]). Az eljárás folyamán az eredeti gráfstruktúrából több lépésen keresztül egy másodlagos struktúra (*klikkfa*) jön létre, amely a csomópontok meghatározott részhalmazait tartalmazza csomópontokként.

A PPTC algoritmus során végrehajtott teljes strukturális átalakítás a következő lépésekből áll (az alábbiakban nem részletezett műveletek pontos leírása a [?] cikkben található).

**Morális gráf.** A morális gráfot az eredetiből úgy kapjuk, hogy az egyes csomópontok szüleiből képezhető összes párt (irányítatlan éllel) összekötjük, majd az eredeti él irányítását is töröljük (vagyis ez már egy irányítatlan struktúra lesz).

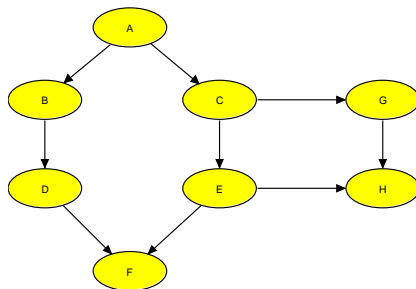
**Háromszögesített (húr-) gráf.** A következő lépésben fel kell számolni a morális gráfban lévő minden olyan kört, amely 3-nál több csomópontot tartalmaz, és nem tartalmaz egy kisebb kört is (vagyis a 3-nál nagyobb köröket „háromszögekre kell bontani”).

**Klikkfa.** Az immár háromszögesített gráfból már megkonstruálható a klikkfa: a háromszögesített gráf minden egyes maximális klikkje a klikkfa egy csomópontja lesz, ezeket pedig úgy kell egy minimális (tehát fa) gráfstruktúrába szervezni, hogy az számításigény szempontjából lehetőleg optimális legyen.

A fentiekkel összhangban meghatározható, hogy a Bayes-hálóra jellemző strukturális információk közül melyek legyenek megjelenítve. Ezek felsorolása a következő alfejezetekben található.

### 2.1. Élek megjelenítése

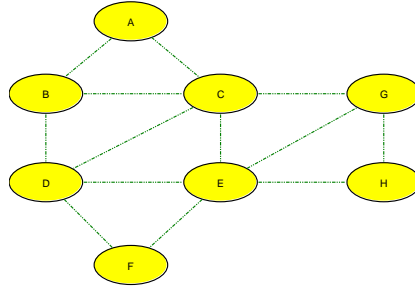
Az OPERATIONS > SHOW CONNECTIONS menüpontra kattintva kapcsolható ki, illetve be az irányított körmentes gráf éleinek (vagyis a Bayes-háló definiáló éleinek) a megjelenítése. Alapértelmezésben ez az egyetlen opció van bekapcsolva, a következők kikapcsolt állapotban indulnak.



9. ábra. Bayes-háló modell az élek megjelenítésével.

## 2.2. Húrgráf megjelenítése

Az OPERATIONS > SHOW CHORDAL GRAPH menüpontra kattintva kapcsolható be, illetve ki a húrgráf éleinek megjelenítése.



10. ábra. Bayes-háló modell az háromszögesített (húr-) gráf éleinek megjelenítésével.

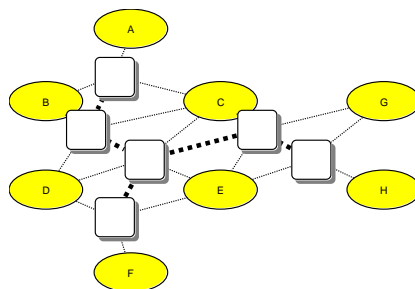
## 2.3. Klikkfa megjelenítése

Az OPERATIONS > SHOW CLIQUE TREE menüpontra kattintva kapcsolható be, illetve ki a klikkfa megjelenítése. Az itt megjelenített elemek a következők:

**Klikkfa csomópontjai.** A klikkfa csomópontjait lekerekített sarkú fehér négyzetek jelképezik. Az egyes klikkfa-csomópontok által tartalmazott csomópontok listája a kurzort a klikkfa-csomópont fölé véve jeleníthető meg: ezt a felbukkanó tooltip tartalmazza.

**Klikkfa „élei”.** A klikkfában lévő klikkek egymáshoz ú.n. *határoló halmazokon* (sepset) keresztül kapcsolódnak. Bár maguk a határoló halmazok nem jelennek meg, azok között a klikkek között, amelyek ily módon össze vannak kötve, vastag szaggatott élek futnak.

**A klikkfa és az eredeti csomópontok közötti összeköttetések.** Az egyes klikkek által tartalmazott csomópontokat vékony szaggatott vonal köti az őket tartalmazó klikk(ek)hez.



11. ábra. Bayes-háló modell a hozzá tartozó klikkfa klikkjeinek és azok kapcsolatainak megjelenítésével.

## 3. Következtetés Bayes-hálókbán

Az OPERATIONS > INFERENCE MODE menüre kattintva, azt aktívává téve (ezt a mellette megjelenő pipa jelzi) válthatunk következtetési üzemmódba. Ekkor a főablak jobb oldalán megjelenik a csomópontok listáját tartalmazó ún. *következtetési nézet* (Inference view), amelynek használatával az alapvető következtetési feladatokat elvégezhetjük.

### 3.1. Evidenciák/döntések bevitele

A szoftver jelenleg a *biztos evidenciák* (hard evidence) kezelését támogatja, azaz ennek megfelelően egy adott csomópontnak a következtetés során vagy teljes biztonsággal ismert az értéke (van biztos evidencia) vagy semmilyen információnk nincsen róla (nincs evidencia).

Evidencia beállítása a következő módon történhet: az adott csomópontoz tartozó értéklístát megjelenítjük (a csomópont neve melletti +/- szimbólum segítségével), majd a kiválasztott értéken jobb-klikkelve a felbukkanó menüben kiválasztjuk a SET AS EVIDENCE pontot.

Evidenciát törölni hasonlóképpen, a változónévre jobb-klikkelve, a CLEAR EVIDENCE menüvel lehetséges, ha csak a kérdéses változóra vonatkozóan akarunk törölni, az összes evidencia törlésére pedig a CLEAR ALL EVIDENCES menü szolgál.

Döntési csomópontok esetén a kiválasztott érték bevitele hasonlóképpen történik, a SET AS DECISION menü segítségével. Döntési csomópont esetén törlésre nincs lehetőség, mivel a döntési csomópontoknak minden pillanatban rendelkezniük kell egy kiválasztott értékkel.

**Evidenciák/döntések megjelenítése** Az evidenciával rendelkező csomópontok nevei a főablakban félkövérrel szedve láthatók, a név alatt a felvett értékkel; a viselkedésükkel összhangban, a döntési csomópontok mindig így jelennek meg következtetési módban. A következtetési nézetben a csomópontok nevei mellett szintén megjelennek az általuk felvett evidenciák/döntések értékei.

### 3.2. Marginális eloszlások számítása

A prediktív következtetési műveletek közül a legalapvetőbb az aktuális evidenciák mellett az egyes csomópontok által felvett marginális eloszlások kiszámítása. Ezt a műveletet a szoftver automatikusan elvégzi minden evidencia/döntés módosítás után. Az egyes változók marginálisai a következtetési nézetben láthatók: a változót kattintással kiválasztva, annak eloszlása a nézet alsó részén jelenik meg, piros oszlopokkal ábrázolva. Az oszlopok által jelképezett valószínűség-értékek (3 tizedesjegy pontossággal) az egérkurzort az oszlop fölé víve jeleníthetők meg.

#### 3.2.1. A mintavételi nézet (Sampling view)

A marginális eloszlások evidenciák függvényében történő változásának pontosabb nyomonkövetése érdekében lehetőség van az eredmények táblázatba szervezésére: erre szolgál a főablak alatt található ún. *Mintavételi nézet* (Sampling view). Ennek működése a következő.

	1.	2.	3.
A/f	0.5	0.57009	1.0
A/t	0.5	0.42991	0.0
B/f	0.45	0.45701	0.5
B/t	0.55	0.54299	0.5
C/f	0.45	0.58879	0.80328
C/t	0.55	0.41122	0.19672

12. ábra. „Mintavételi” nézet a következtetési esetek során előállt marginálisok számon tartására.

A következtetési nézetben a csomópontok, illetve azok értékeinek környezeti menüjében a WATCH menüpontot kiválasztva az adott érték(ek) sorcímkeként a mintavételi nézet táblázatához adhatók. A mintavételi nézet SAMPLE gombját megnyomva a táblázathoz egy új oszlop adódik, amely tartalmazott értékek aktuális konfiguráció melletti értékeit tartalmazza.

A táblázat a fenti alapfunkciókon kívül még a következő funkcionalitásokat támogatja.

- Egy adott sor bármely pontjára klikkelve a sor kiválasztódik. Ez az áttekinthetőséget javító szerepén túl még arra használható fel, hogy az ily módon kiválasztott sor a REMOVE SELECTED környezeti menü kiválasztásával törölhető a táblázatból.
- A COPY TABLE TO CLIPBOARD AS TEXT környezeti menü kiválasztására a táblázat tartalma (tabulátorokkal szeparált) szöveges formátumban a vágólapra kerül; innen pedig már könnyen beilleszthető valamilyen külső táblázatkezelő programba.
- A REM. COLUMNS gomb megnyomására a táblázat által tartalmazott minta-értékek törlődnek, de a sorok címkéi megmaradnak.
- A REMOVE ALL gomb törli a táblázat teljes tartalmát.

### 3.3. Következtetés érzékenységének vizsgálata

Az érzékenységi vizsgálat (sensitivity of inference – SOI) során arra keressük a kérdést, hogy egy adott célkonfiguráció valószínűségét hogyan határozzák meg egy másik, az ú.n. feltételcsomópontok halmazába tartozó változók lehetséges értékei, illetve, hogy a feltételcsomópontoknak egymás után értéket adva, ezek az értékadások hogyan módosítják a célkonfiguráció feltételes valószínűségét annak korábbi értékéhez képest.

A tipikus érzékenységvizsgálat a következő lépésekből áll.

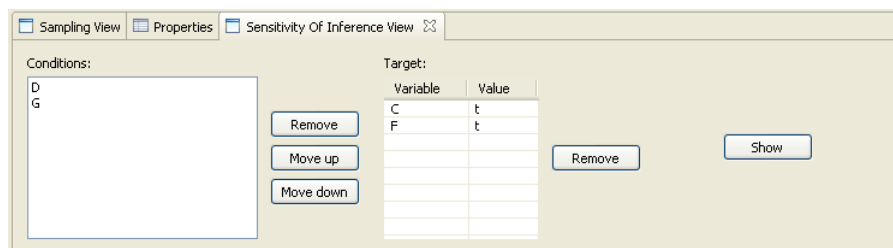
**Evidenciák bevitele.** A szokásos módon (3.1) bevisszük a rendszerbe a fixen tartani kívánt evidenciákat és döntéseket.

**Feltételek és vizsgálati sorrendjük meghatározása.** A feltételváltozókat az INFERENCE VIEW nézetben a változók ADD TO SOI CONDITIONS környezeti menüjére kattintva jelölhetjük ki, majd ez után, ha szükséges, sorrendjükön a SENSITIVITY OF INFERENCE VIEW nézet MOVE UP és MOVE DOWN gombjaival módosíthatunk.

**Célkonfiguráció megadása.** A célkonfigurációba tartozó értékeket az INFERENCE VIEW nézetben, a változóértékek SET AS SOI TARGET környezeti menüjének segítségével vehetjük fel.

**Számítás végrehajtása.** Ha összeállítottuk a kívánt vizsgálati konfigurációt, a SENSITIVITY OF INFERENCE VIEW nézet SHOW gombjával hajthatjuk végre a szükséges számításokat.

A számítások lefutása után egy új szerkesztőablak nyílik meg, amely az érzékenységvizsgálat eredményeit ábrázoló diagramot tartalmazza.

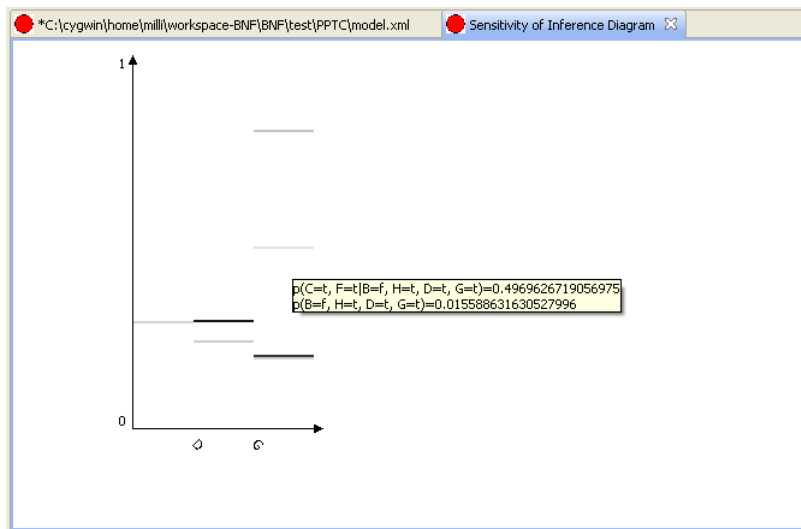


13. ábra. Az érzékenységvizsgálat paraméterezésének összeállítására szolgáló nézet.

### 3.3.1. A „Sensitivity of Inference Diagram” értelmezése

Az érzékenységvizsgálat során a fix evidenciákat (azok meghatározott sorrendjében) iteratíván bővítjük ki a feltételváltozók lehetséges értékeivel. Minden egyes ilyen lépésben kiszámítjuk (1) az aktuális evidenciahalmaz valószínűségét és (2) a célkonfiguráció valószínűségét, az evidenciák halmaza, mint feltétel mellett.

Összefoglalva, a kapott valószínűség párok egy fába rendezhetők, melynek gyökerében a  $(P(\underline{T} = \underline{t} | E_0 = e_0))$  található, ennek gyermekei pedig az első feltétel lehetséges értékeivel kibővített evidenciát tartalmazzák:  $(P(\underline{T} = \underline{t} | E_0 \cap \{C_0\} = e_0 \cap \{c_0\}))$ , így folytatva az evidenciák halmazának bővítését a második, harmadik stb. feltételváltozókkal a fában lefelé haladva.

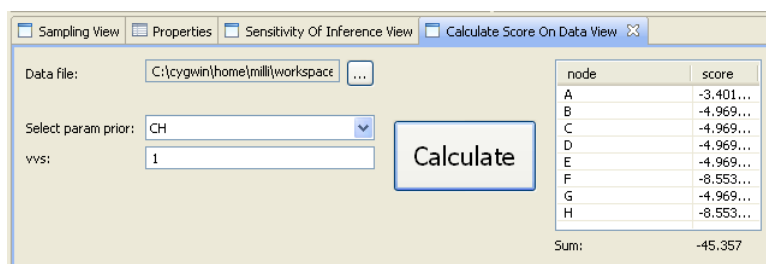


14. ábra. Az érzékenységvizsgálat eredményét megjelenítő diagram.

A diagram vízszintesen felosztva oszlopokra van tagolva, melyek a fent leírt fa egy-egy szintjének felelnek meg, egy-egy oszlopon belül pedig vízszintes vonalak jelképezik az ábrázolt  $(P(E), P(T|E))$  valószínűség-párokat: a vonal annál sötétebb, minél nagyobb a  $P(E)$  valószínűség, a függőleges elhelyezkedése pedig a  $P(T|E)$  mennyiséggel arányos.

### 3.4. Adatfájl alapján történő score-számítás

Az aktuális modellstruktúra jósága minősítésének eszköze az adatfájl alapján számított score-érték. A vonatkozó nézet az OPERATIONS > SHOW CALCULATE SCORE ON DATA VIEW menü segítségével hívható elő, és az alábbi működési elemeket tartalmazza.



15. ábra. Az adatok alapján történő score-számítás paraméterezésének összeállítására és az eredmények megjelenítésére szolgáló nézet.

#### Bemeneti mezők

A DATA FILE mezőben választható ki az adatokat tartalmazó fájl.

A SELECT PARAM. PRIOR mezőben választható ki, hogy a *CH* vagy a *BD<sub>eu</sub>* mutató legyen kiválasztva.

Az kétféle számítás pontos leírását a [], illetve a [] cikk tartalmazza.

A VSS mezőben adható meg a számítás során alkalmazandó virtuális mintaméret.

**CALCULATE gomb** Ennek megnyomására zajlik le a számítás maga.

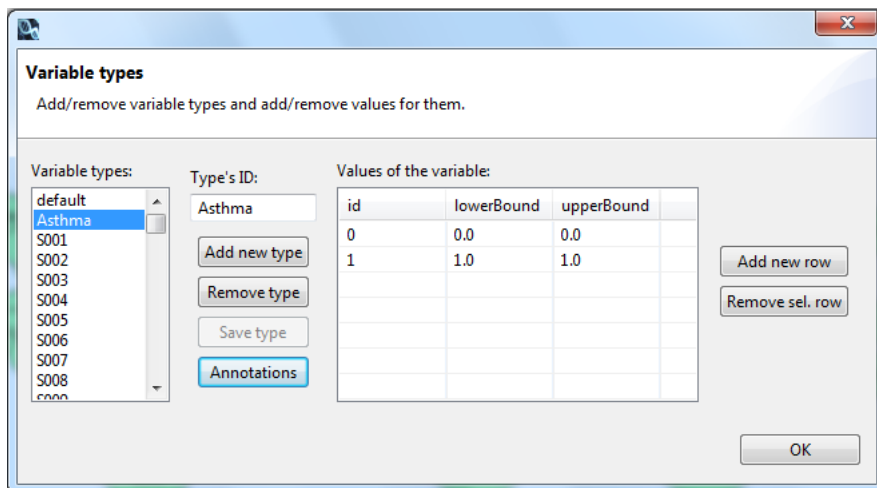
### Kimeneti mezők

A score-számítás csomópontonkénti részeredményeit a CALCULATE gombtól jobbra eső táblázat tartalmazza.

A teljes score (vagyis a fenti táblázat elemeinek összege) a SUM címke mellett, a táblázat alatt található.

## 4. Annotáció és magyarázatkeresés

Az annotáció lényege, hogy egy modell különféle absztrakciós szintjein lévő elemeit kiegészítő információkkal gazdagítsuk. Az változókhöz, illetve változó csoportokhoz publikációs hivatkozást vagy akár adatbázis referenciát rendelhetünk, egyes kiemelten fontos elemekhez további részleteket fűzhetünk. De lehetővé válik korábbi eredmények adott elemhez való linkelése is. A cél az, hogy az elemzést elősegítse a rendszerezett addicionális információ.

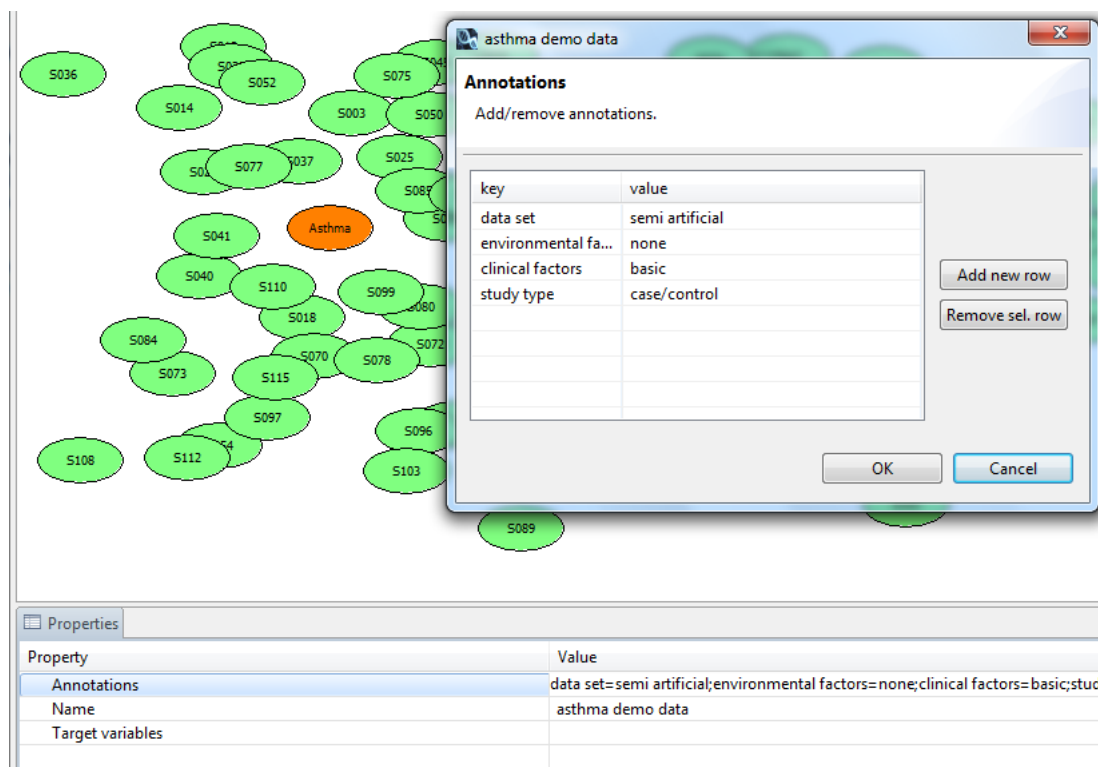


16. ábra. Változó típus elemek annotálása.

Az annotációs mezők elérhetőek a változók, változó típusok (Variable Types), változó csoportok (Variable groups) és a teljes modell szintjén is. Az egyes változók és a modell esetében az adott entitás PROPERTIES mezőjén belül található ANNOTATIONS nyomógomb segítségével. A változó csoportok és változó típusok

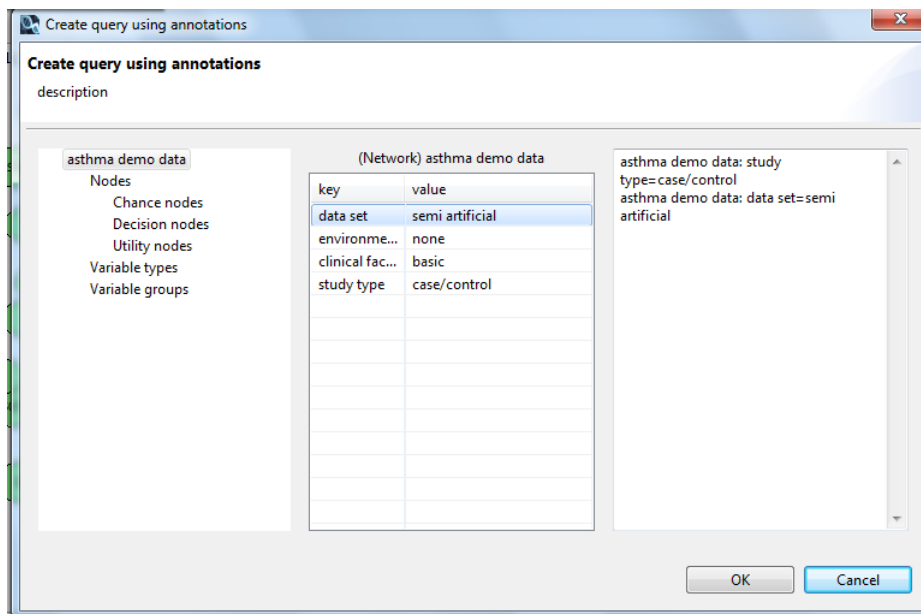


esetén pedig a tulajdonságok beállítására szolgáló ablakon belül lévő ANNOTATIONS nyomógomb által. Az annotációs mezőkbe kulcs - érték (key - value) párok vihetők be. Ez azt a cél szolgálja, hogy az annotációk rendszerezetten álljanak rendelkezésre. Lehetséges egy elemet többféle annotációval ellátni, továbbá több elemet hasonló típusú (key) annotációval gazdagítani.



17. ábra. Modell elemek annotálása. Az annotáció strukturáltan hozzárendelhető az egyes elemekhez key - value párok formájában.

Az annotációk egy központi menüpontból is elérhetők az elemek hierarchiája szerint tagoltan. Ugyanerről a felületről nyílik lehetőség lekérdezések összeállítására. Ez az OPERATIONS menü CREATE QUERY USING ANNOTATIONS parancsával érhető el. Az lekérdezéshez használni kívánt annotációk kiválasztása dupla kattintással történik. Ennek következtében a jobb oldalon lévő mezőben gyűlnek a kiválasztott annotációs kifejezések. Az OK gomb megnyomásával indítható egy általános keresés a Google keresőjében.



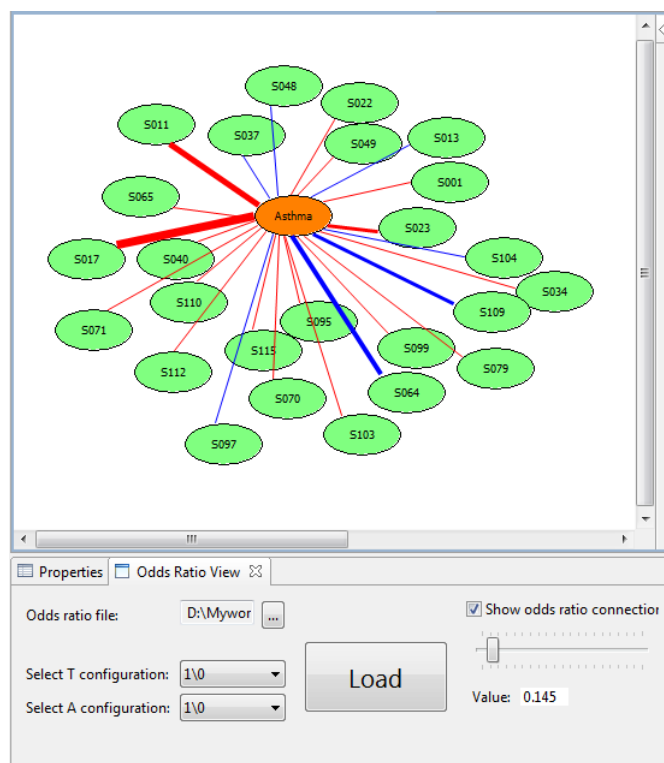
18. ábra. Query létrehozása annotált elemekből.

E funkció segítségével kapcsolódó dokumentumokat, illetve háttéranyagokat kereshetünk, amelynek célja az eredmények értelmezésének elősegítése, a magyarázatkeresés támogatása.

## 5. Hatáserősség

A Bayes-háló alapú bayesi elmezés főként strukturális jellemzők feltárását teszi lehetővé. A strukturális relevancia mellett azonban nem elhanyagolható a parametrikus relevancia, azaz a hatáserősségek kérdése. Tehát amellett, hogy az egy-egy asszociációs elemzés célja annak feltárása, hogy a célváltozóval mely változók vannak függésben, a függőségek erősségének a vizsgálata is lényeges információt hordoz.

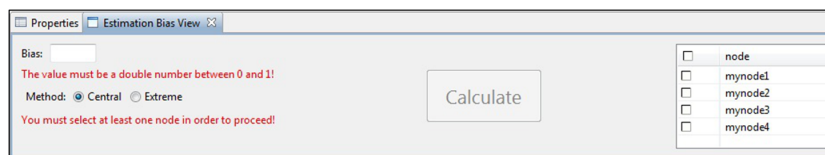
A felület lehetővé teszi a bayesi relevanciaelemzés eredményeinek hatáserősség információval történő gazdagítását. Ez a funkció az OPERATIONS menű SHOW ODDS RATIO VIEW menüpontjára kattintva érhető el. Az ekkor megnyíló nézetben ki kell választani a kérdéses információkat tartalmazó adatfájlt (Odds ratio file) és az értékek megjelenítéséhez meg kell nyomni a LOAD gombot. A változók értékészletétől függően többféle hatáserősség mutató állhat rendelkezésre. A célváltozó értékei közötti választásra a SELECT T CONFIGURATION, a magyarázó változók értékei közötti választásra a SELECT A CONFIGURATION legördülő menü szolgál.



19. ábra. Hatáserősség (odds ratio megjelenítése).

## 6. Becslések torzítása

A valószínűségi következtetés és modellezés szempontjából kétféle következménye lehet az embereknél megfigyelhető valószínűségi becslési heurisztikáknak a paraméterek meghatározását tekintve: *túlzott magabiztosság* (overconfidence) vagy a *magabiztosság hiánya* (underconfidence). A *túlzott magabiztosság* a valószínűségi skála szélsőséges értékei felé való eltolódást eredményez ( $p < 0.5 : p \rightarrow 0, p > 0.5 : p \rightarrow 1$ ). Ezzel szemben a magabiztosság hiánya az előzővel ellentétes, centrális tendenciát mutat, azaz minden esetben a maximális bizonytalanságot jelentő  $p=0.5$  felé való torzítást jelent.



20. ábra. A becsléstorzítás nézet funkciói.

A becsléstorzítási funkciók az ESTIMATION BIAS menüpontból érhetőek el az OPERATIONS menüin belül. A 20 ábrán látható nézetben megadható a torzítás mértéke a BIAS pontban egy [0, 1] közötti skálán (1 jelenti a legerősebb értéket.)

A METHOD pontban pedig kiválasztható az alkalmazandó torzítás típusa: EXTREME (túlzott magabiztosság), vagy CENTRAL (magabiztosság hiánya). A jobboldali listában kiválasztható, hogy mely csomópontoknál, vagyis azok feltételes valószínűségi tábláiban megadott értékeinél, kerüljön sor a torzítás alkalmazására. A beállításokat követően a CALCULATE gomb megnyomásával indítható a számítás.