

Automatikus következtetés

1. Rezolúció

Feladat:

A_1 : Ha süt a nap, akkor Péter strandra megy.

A_2 : Ha Péter strandra megy, akkor úszik.

A_3 : Péternek nincs lehetősége otthon úszni.

Lássuk be, hogy ezekből következik:

B : Ha süt a nap, akkor Péter nem marad otthon.

Formalizálás:

- | | | | |
|------------------------|-----|---------|------------------------|
| – süt a nap: | p | A_1 : | $p \rightarrow q$ |
| – Péter strandra megy: | q | A_2 : | $q \rightarrow r$ |
| – Péter úszik: | r | A_3 : | $\neg(s \wedge r)$ |
| – Péter otthon marad: | s | B : | $p \rightarrow \neg s$ |

Átalakítás

logikai következmény

- Kell: $p \rightarrow q, q \rightarrow r, \neg(s \wedge r) \Rightarrow p \rightarrow \neg s$
- Minden olyan interpretáció (igazságértékelés), amely kielégíti a feltételeket, az kielégíti a következményt is.
 - vagy: nincs olyan interpretáció (igazságértékelés), amely a feltételeket is, és következmény negáltját is kielégítené.
 - azaz: $(p \rightarrow q) \wedge (q \rightarrow r) \wedge \neg(s \wedge r) \wedge \neg(p \rightarrow \neg s)$ kielégíthetetlen
vagy: $(\neg p \vee q) \wedge (\neg q \vee r) \wedge (\neg s \vee \neg r) \wedge p \wedge s$ kielégíthetetlen

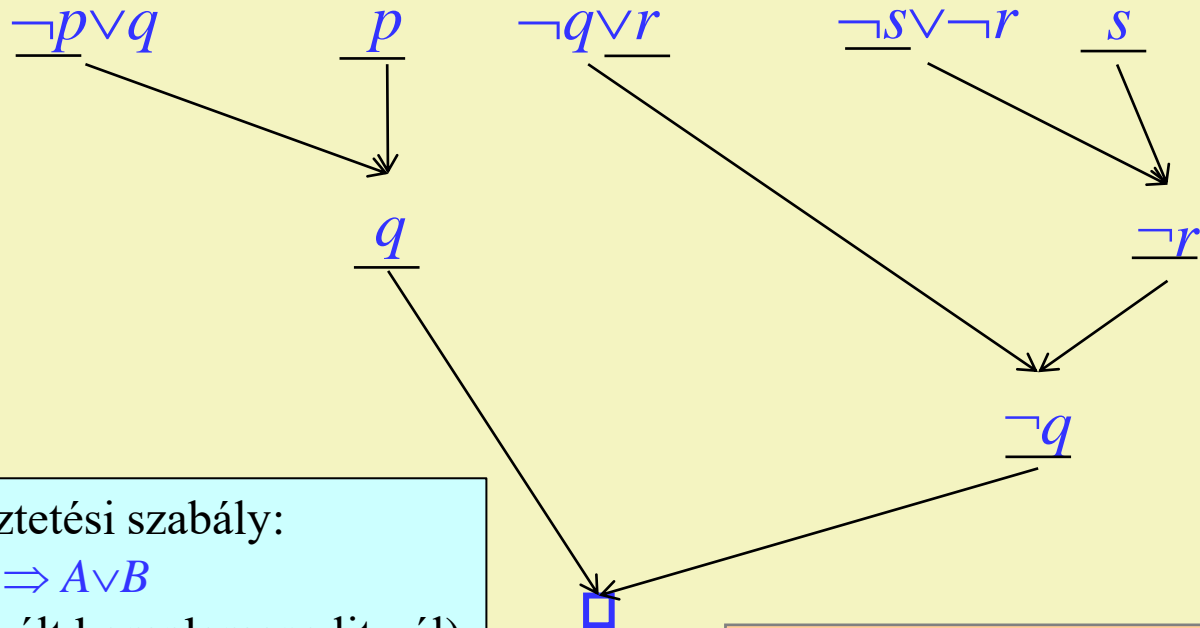
KNF: klózek között 'és' művelet
klóz: literálok között 'vagy' művelet
literál: ítéletváltozó vagy annak negáltja

- Tehát meg kell mutatnunk, hogy bármelyik interpretációval (igazságértékeléssel) **legalább az egyik klóz hamis** lesz.

Rezolúció = indirekt bizonyítás

- ❑ Tekintsük a klózok halmazát és tegyük fel indirekt módon, hogy van olyan interpretáció, amikor **mindegyik klóz igaz**.
- ❑ Ekkor például p is, és $\neg p \vee q$ is *igaz*. Ha azonban p *igaz*, akkor $\neg p$ *hamis*, és ekkor a $\neg p \vee q$ csak úgy lehet *igaz*, ha a q is *igaz*.
- ❑ A q – amely ugyancsak egy klóz – tehát akárcsak a többi klóz *igaz* az indirekt feltevés szerinti interpretációban. Vegyük hát hozzá az eddigi klózhalmazhoz.
- ❑ Az előbbihez hasonló módon bővítsük tovább a klózhalmazt addig, amíg ellentmondáshoz nem jutunk. (Például egyszerre megjelenik a klózhalmazban a q és $\neg q$.)

Rezolúciós eljárás

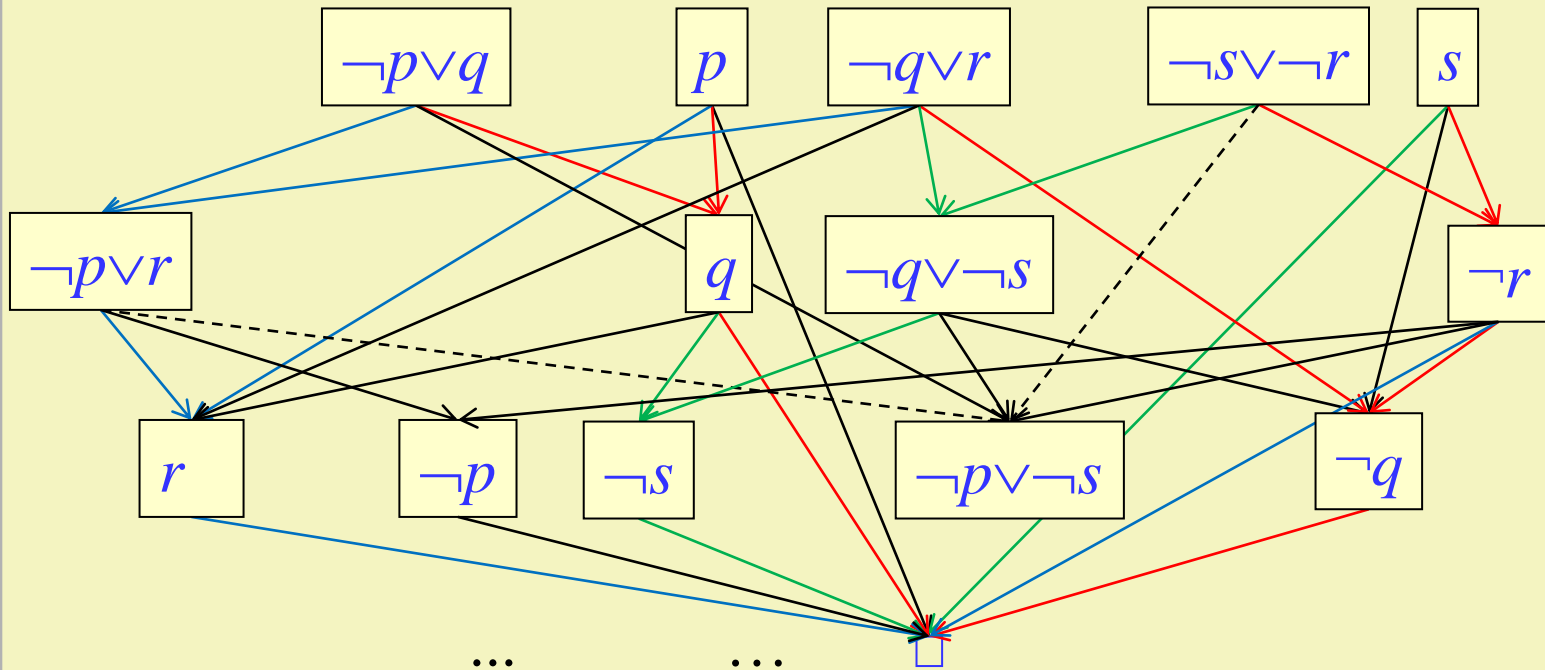


A következtetési szabály:
 $\underline{l \vee A}, \underline{\neg l \vee B} \Rightarrow \underline{A \vee B}$
(l az eliminált komplement literál)

Ellentmondást jelző üres klóz:
minden interpretációban hamis

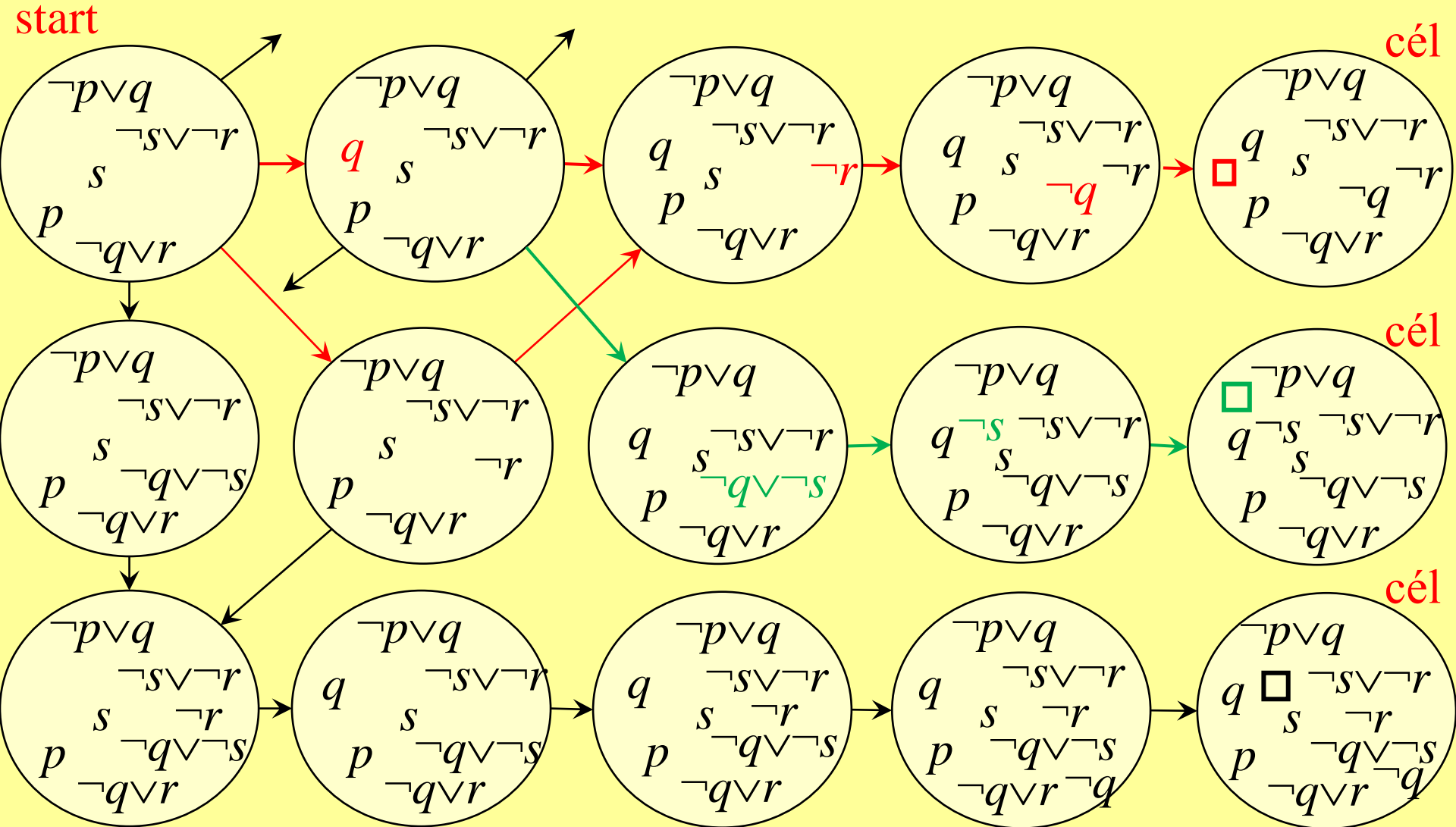
Tehát ha süt a nap, akkor Péter nem marad otthon.

Cáfolati-, rezolúciós gráf



- Cáfolati gráf: az üres klózt előállítását bemutató gráf
- Rezolúciós gráf: az összes klóz előállítását mutató gráf

Reprezentációs gráf



A reprezentációs gráf tulajdonságai

- Egy csúcs egyetlen klózzal tartalmaz többet a szülőcsúcsánál: olyan, amelyik a szülő csúcs két klózából vezethető le.
 - Mindegyik csúcs tartalmazza a kiinduló klózatokat.
 - Nincsenek körök.
 - Az a rezolúciós lépés, amely egy csúcsban elvégezhető, az annak azon gyerek csúcsában is elvégezhető, amelyhez egy másik rezolúciós lépés árán jutottunk el.
- Ha van cáfolat, akkor minden csúcsból el lehet jutni egy üres klózt is tartalmazó célcsúcsba.
 - Nincs rossz döntés, legfeljebb csak felesleges.
- Ítéletkalkulusban a gráfnak csak véges sok különböző csúcsa lehet, predikátumkalkulusban lehet végtelen sok is.

Példa: Kuruzslók-e a doktorok?

A_1 : Van olyan páciens, aki minden doktorban megbízik.

A_2 : A kuruzslókban egyetlen páciens sem bíz meg.

Lássuk be, hogy

B : Egyetlen doktor sem kuruzsló.

Formalizálás:

$P(x)$: x egy páciens

$A_1 : \exists x\{P(x) \wedge \forall y[D(y) \rightarrow M(x,y)]\}$

$D(y)$: y egy doktor

$A_2 : \forall x\{P(x) \rightarrow \forall y[K(y) \rightarrow \neg M(x,y)]\}$

$K(y)$: y egy kuruzsló

$B : \forall x[D(x) \rightarrow \neg K(x)]$

$M(x,y)$: x megbíz az y -ban

Kell: $\exists x\{P(x) \wedge \forall y[D(y) \rightarrow M(x,y)]\} \wedge \forall x\{P(x) \rightarrow \forall y[K(y) \rightarrow \neg M(x,y)]\}$

$\wedge \neg \forall x[D(x) \rightarrow \neg K(x)]$

kielégíthetetlen

Formulák klóz-formára (SKNF) hozása

1. Kiküszöböljük az \leftrightarrow és a \rightarrow műveleti jeleket (logikai törvények).
2. Redukáljuk a negációk hatáskörét (DeMorgan azonosságok).
3. Standardizáljuk a változókat (kvantonkénti átnevezés).
4. Egzisztenciális kvantorok kiküszöbölése. (Skolemizálás:
 $\forall x_1 \dots \forall x_n \exists z F(\dots, z, \dots)$ helyett $\forall x_1 \dots \forall x_n F(\dots, f(x_1, \dots, x_n), \dots)$
– nem ekvivalens átalakítás, de kielégíthetőség tartó)
5. Univerzális kvantorok kiemelése a formula elejére a sorrendjük megtartásával. (prenex normál forma)
6. A formula többi részét konjunktív normálformára alakítjuk (kommutatív, asszociatív, disztributív törvények).
7. Kialakítjuk a klózokat (a kvantorokat, és a konjunkciós műveleti jeleket elhagyjuk, a változókat klózonként egyedivé nevezzük át.)

Skolemizált konjunktív normálforma (SKNF)

$a \rightarrow b$ helyett $\neg a \vee b$

$$A_1: \exists x\{P(x) \wedge \forall y[D(y) \rightarrow T(x,y)]\} = \exists x\{P(x) \wedge \forall y[\neg D(y) \vee T(x,y)]\} \approx$$

$$\approx P(\mathbf{a}) \wedge \forall y[\neg D(y) \vee T(\mathbf{a},y)] = \forall y\{P(\mathbf{a}) \wedge [\neg D(y) \vee T(\mathbf{a},y)]\}$$

$$P(\mathbf{a}), \neg D(y) \vee T(\mathbf{a},y)$$

Skolemizálás
 \mathbf{a} a Skolem konstans

$$A_2: \forall x\{P(x) \rightarrow \forall y[Q(y) \rightarrow \neg T(x,y)]\} =$$

$$= \forall x\{\neg P(x) \vee \forall y[\neg Q(y) \vee \neg T(x,y)]\} =$$

$$= \forall x \forall u\{\neg P(x) \vee \neg Q(y) \vee \neg T(x,y)\}$$

$$\neg P(x) \vee \neg Q(u) \vee \neg T(x,u)$$

$a \rightarrow b$ helyett $\neg a \vee b$

változó átnevezés

$a \rightarrow b$ helyett $\neg a \vee b$

$$B: \neg \forall x[D(x) \rightarrow \neg Q(x)] = \neg \forall x[\neg D(x) \vee \neg Q(x)] =$$

$$= \exists x[D(x) \wedge Q(x)] \approx D(\mathbf{b}) \wedge Q(\mathbf{b})$$

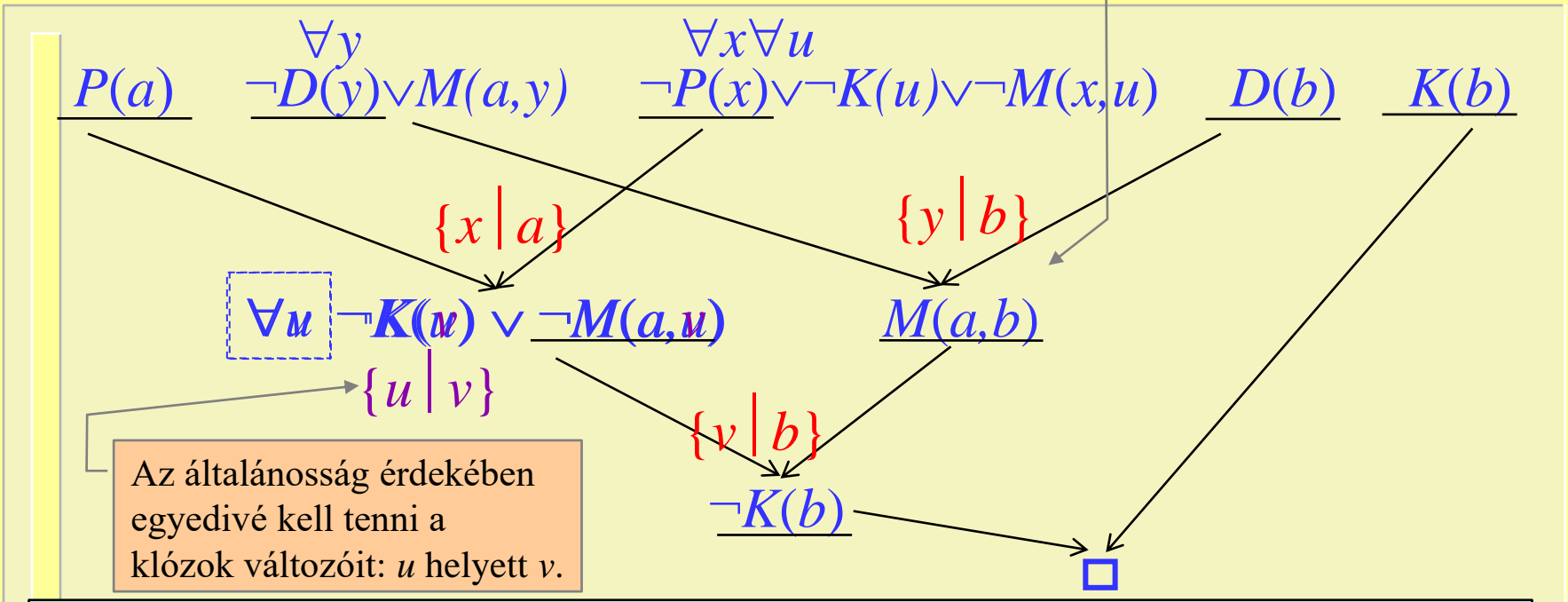
$$D(\mathbf{b}), Q(\mathbf{b})$$

De Morgan's törvény

\mathbf{b} a Skolem konstans

Egy klózpár rezolválásához olyan változó helyettesítésére van szükség, amellyel az elhagyásra kiszemelt komplementer literálok azonos alakra hozhatók. (Az **egyesítő algoritmus**sal egy ilyen helyettesítés található meg.) Ezt követően a klózpár ezen helyettesítéssel kapott példányait rezolváljuk.

Rezolúciós eljárás



Általános rezolúciós szabály:

$C_1 = P(t_{11}, \dots, t_{1n}) \vee \dots \vee P(t_{r1}, \dots, t_{rn}) \vee C_1'$; $C_2 = \neg P(u_{11}, \dots, u_{1n}) \vee \dots \vee \neg P(u_{s1}, \dots, u_{sn}) \vee C_2'$
 C_1' vagy C_2' lehet üres, de tartalmazhatják $P(\dots)$ illetve $\neg P(\dots)$ további előfordulásait.
 ha $P(t_{11}, \dots, t_{1n}), \dots, P(t_{r1}, \dots, t_{rn}), P(u_{11}, \dots, u_{1n}), \dots, P(u_{s1}, \dots, u_{sn})$ egyesíthetők a δ változó-helyettesítéssel, akkor C_1 és C_2 rezolvense: $R(C_1, C_2) = C_1' \delta \vee C_2' \delta$

Rezolúció = lokális keresés

- globális munkaterület: aktuális klózalmaz
- kiindulási érték: az „axiómák \Rightarrow célállítás” klózzai
- terminálási feltétel:
 - sikeres: üres klóz
 - sikertelen: nincs újabb rezolvens klóz
- kereső szabály: rezolvens képzés
- vezérlési stratégia: nem-módosítható
- heurisztika: jó lenne a hatékonyság miatt, de sajnos nincs

ADAT := kezdeti érték

while \neg terminálási feltétel(ADAT) **loop**

 SELECT SZ FROM alkalmazható szabályok

 ADAT := SZ(ADAT)

endloop

Rezolúció algoritmus

$A_1, A_2, \dots, A_n \Rightarrow B$ helyett azt vizsgáljuk, kielégíthetetlen-e az
 $A_1, A_2, \dots, A_n, \neg B$ formulák klóz formája

1. **KLÓZOK** := A_1, A_2, \dots, A_n és $\neg B$ formulák klózzai
2. **loop**
3. **if** $\square \in \text{KLÓZOK}$ **then return** *kielégíthetetlen*
4. **if** nincs olyan $C_1, C_2 \in \text{KLÓZOK}$, amelyre $R(C_1, C_2)$
 még nem ismert (nincs a *KLÓZOK* közt)
 then return *nem kielégíthetetlen*
5. **select** $C_1, C_2 \in \text{KLÓZOK}$, ahol $R(C_1, C_2)$ nem ismert
6. **KLÓZOK** := $\text{KLÓZOK} \cup R(C_1, C_2)$
7. **endloop**

Rezolúció tulajdonságai

- ❑ **Helyes** (eljárás): ha terminál, akkor helyes eredményt ad. (Üres klóz megtalálásakor a kiinduló klóz halmaz kielégíthetetlen, ha nem tud újabb klózt előállítani, akkor a kiinduló klóz halmaz kielégíthető.) Ugyanakkor elsőrendű logikában nem biztosan terminál.
- ❑ **Teljes** (eljárás): egy kielégíthetetlen klóz halmazból véges lépésben levezethető az üres klóz.
- ❑ Elsőrendű logikában a kielégíthetlenség csak **parciálisan dönthető el**, mert nem terminál garantáltan a módszer:
 $\{ \neg P(x), P(y) \vee \neg P(f(y)), P(a) \}$

Válaszadás rezolúcióval

*“Ha Fifi mindenhová követi Jánost, és
János most az iskolában van,
akkor hol van most Fifi?”*

Formalizáció:

$H(y,x) \sim y$ dolog az x helyen van

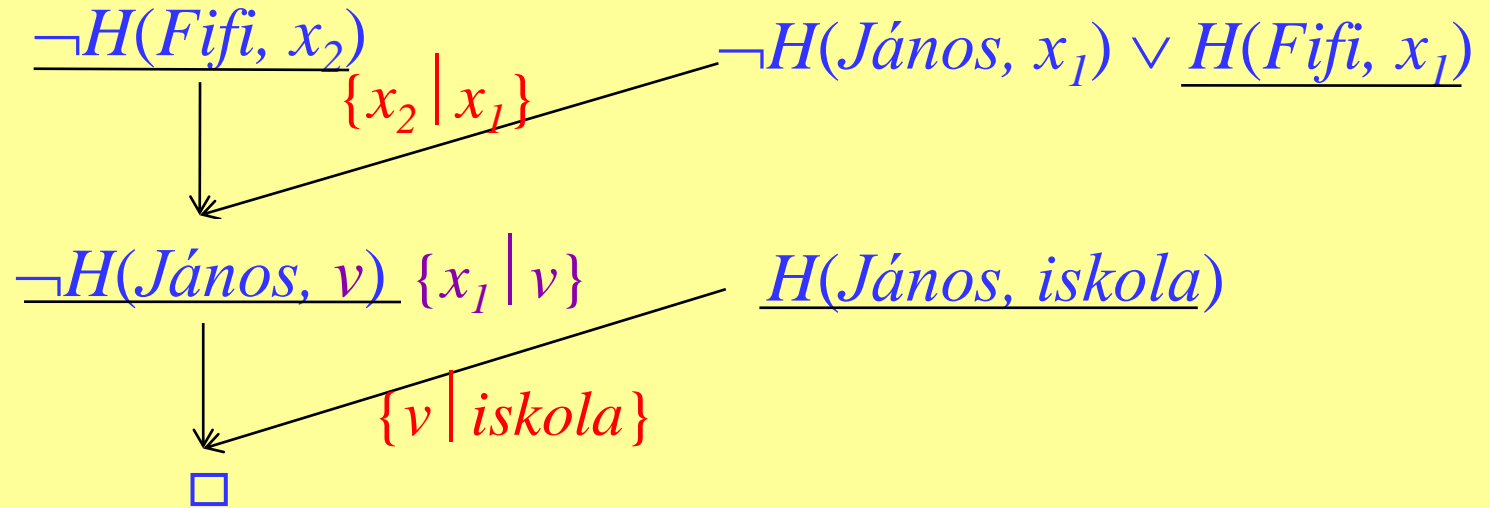
$\forall x[H(\text{János},x) \rightarrow H(\text{Fifi},x)]$

$H(\text{János}, \text{iskola})$

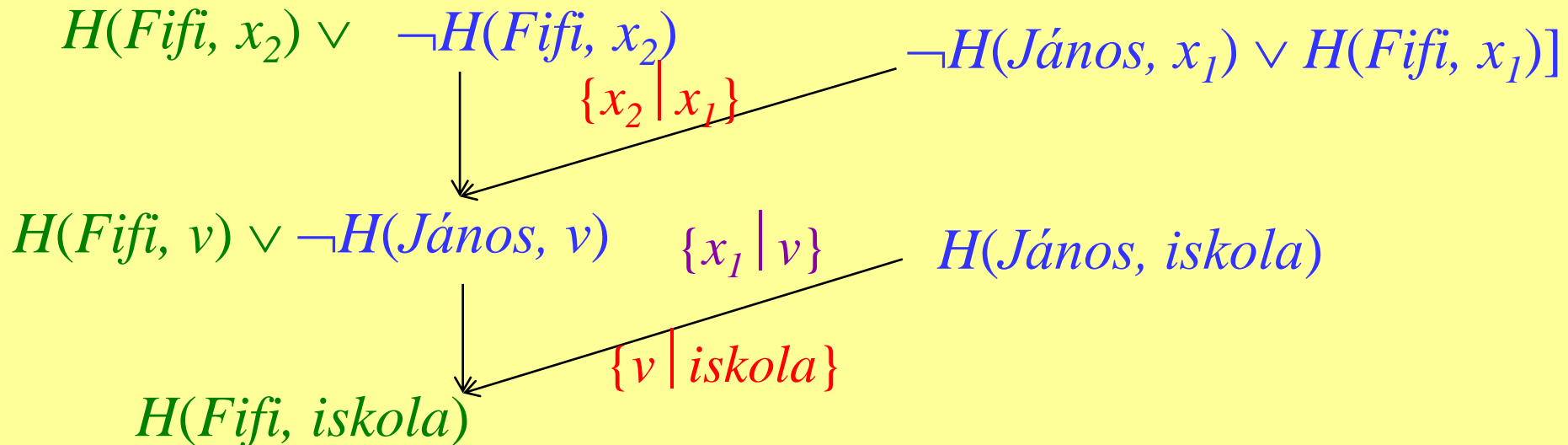
Először lássuk be, hogy létezik-e Fifi számára hely a világban?

$\exists xH(\text{Fifi},x)$

Cáfolati gráf



Válaszadási gráf



Válaszadási eljárás

1. A kérdést (ki, mit, hol, mikor, mennyiért) egy „van-e válasz a kérdésre” célállítással helyettesítjük.
2. Rezolúcióval belátjuk, hogy a célállítás következik az axiómákból.
3. A célállítás negáltjából származó klózokat negáltjaik hozzáfűzésével érvényes formulákká egészítjük ki.
4. A cáfolati gráf által meghatározott rezolúciót követve létrehozuk a hasonló szerkezetű válaszadási gráfot, amelynek gyökere tartalmazza az egyik választ.

Rezolúciós stratégiák

- A rezolúció **nem-determinisztikus**. Egy lépésben
 - egyszerre több rezolválható klóz pár lehet
 - egy klóz párban több komplementer literál pár lehet
 - ugyanannak a literálnak több előfordulása lehet

$$\{P(x,f(a)) \vee P(x,f(y)) \vee Q(y), \quad \neg P(z,f(a)) \vee \neg Q(z), \quad P(u,f(a)) \vee \neg Q(a)\}$$

modellfüggő vezérlési stratégiák: csak klóz alapú reprezentáció esetén értelmezhetőek.

- Egy rezolúciós stratégia a rezolúció alapalgorithmusát kiegészítő olyan előírás, amely
 - **sorrendet ad** a rezolvens képzésekre (sorrendi stratégia)
 - **korlátozza** egy adott pillanatban előállítható rezolvensek körét (vágó strat.)

sérülhet a módszer teljessége

Rezolúció kritikája

- A rezolúció nem jó MI módszer:
 - A számos modellfüggő vezérlési stratégia ellenére **sem hatékony**, sok felesleges rezolúciós lépést végez.
 - **Nem építhető heurisztika** a vezérlési stratégiába, mert az állítások a klóz-formára hozás után már nem emlékeztetnek a feladatban betöltött szerepükre, ezért nehéz „súgni”, hogy mely klózokkal érdemes próbálkozni.

A formulák alakja segítheti a következtetést

Ha például be kell látnunk azt, hogy

$$A, C \rightarrow \neg A, A \rightarrow B, \neg B \rightarrow D \Rightarrow B$$

akkor könnyű kitalálni, hogy mely feltételekre van szükség a bizonyításnál: $A, A \rightarrow B \Rightarrow B$

De a rezolúció nem képes felhasználni ezt a segítséget, hiszen eliminálja az implikációt a formulákból:

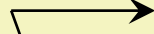
$$A, \neg C \vee \neg A, \neg A \vee B, B \vee D, \neg B$$

Olyan következtetési eljárás kell, ahol az állítások megőrzik eredeti alakjukat, különösen az implikációt.

2. Szabályalapú logikai következtetés

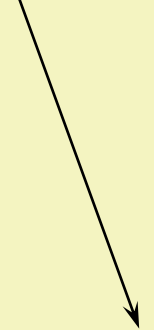
- Egy $A_1, \dots, A_n \Rightarrow C$ probléma esetén az axiómákat két csoportba soroljuk: szabályokra és tényekre.

axiómák



tények

konkrét ismeret
implikáció nélküli
formulákban



szabályok

általános ismeret
implikációs formulákban
 $A \rightarrow B$

Az előre- illetve a hátrafelé láncolás

- **Előrefelé láncolás:** egy alkalmas (illeszthető) szabály segítségével egy állításból új állítást vezet le.

tény: $Kutya(Fifi) \wedge Postás(Jani)$
szabály: $\forall x \forall y Kutya(x) \wedge Postás(y) \rightarrow Harap(x, y)$ \Rightarrow $Harap(Fifi, Jani)$

Nehezebb lenne, ha itt nem alakí, hanem logikai ekvivalenciát kellene igazolni. Pl. a tény: $\neg(\neg Kutya(Fifi) \vee \neg Postás(Jani))$

Illeszthetőség vizsgálat: a tény ekvivalens a szabály előfeltételével az $\{x | Fifi, y | Jani\}$ helyettesítés mellett (amit az egyesítő algoritmus számol ki).

- **Hátrafelé láncolás:** egy állítás bizonyítását visszavezeti egy alkalmas (illeszthető) szabály előfeltételének igazolására.

cél: $Kutya(Fifi)$
szabály: $\forall x Ugat(x) \rightarrow Kutya(x)$

Szerencsénk van, hogy az illesztésnél literált literállal kellett összevetni. Ekkor alakí azonosság = logikai ekvivalencia

\sim elég belátni: $Ugat(Fifi)$

Illeszthetőség vizsgálat: a cél ekvivalens a szabály következményével az $\{x | Fifi\}$ helyettesítés mellett (amit az egyesítő algoritmus számol ki).

Szabályalapú következtetés irányai

- Egy tényekkel, szabályokkal és célállítással megadott probléma bizonyítható
 - *előre haladva*: a tényekből indulva előrefelé láncolással új állításokat vezetünk le, majd azokból még újabbakat, amíg a célállítást meg nem kapjuk
 - *visszafelé haladva*: a célt hátrafelé láncolással részcélokra cseréljük le, a részcélokat további részcélokkal váltjuk fel, amíg tények által igazolható részcélokhoz nem jutunk.

Ezek a módszerek nem teljesek:

Például a $P \rightarrow Q$, $\neg P \rightarrow Q$ szabályokból a fenti módszerek egyike sem vezeti le a Q célállítást, pedig $P \rightarrow Q$, $\neg P \rightarrow Q \Rightarrow Q$

Előre haladó szabályalapú reprezentáció

célja, hogy a hátrafelé láncolásnál literált literállal kelljen összevetni

ÉS/VAGY formájú (ÉVF) kifejezések:

- literálok
- $A \wedge B$, $A \vee B$ alakú formulák, ahol az A és B is ÉVF kifejezés.

□ **Tény:**

- univerzálisan kötött **tetszőleges** ÉVF kifejezés

□ **Szabályok:**

- $L \rightarrow W$ alakú univerzálisan kötött kifejezések, ahol L egy literál, a W pedig ÉVF kifejezés

□ **Cél:**

- $L_1 \vee \dots \vee L_n$ alakú egzisztenciálisan kötött kifejezés, ahol L_1, \dots, L_n literálok.

Példa előre haladó szabályalapú következtetésre

Tény:

$$(A \vee \neg B) \wedge C$$

Szabályok:

$$A \rightarrow D \wedge E$$

$$\neg B \rightarrow D \vee \neg H$$

Cél:

$$D \vee G \vee \neg H$$

elég az egyik részből
levezetni a célt

mindkét részből le
kell vezetni a célt

szabály illesztés

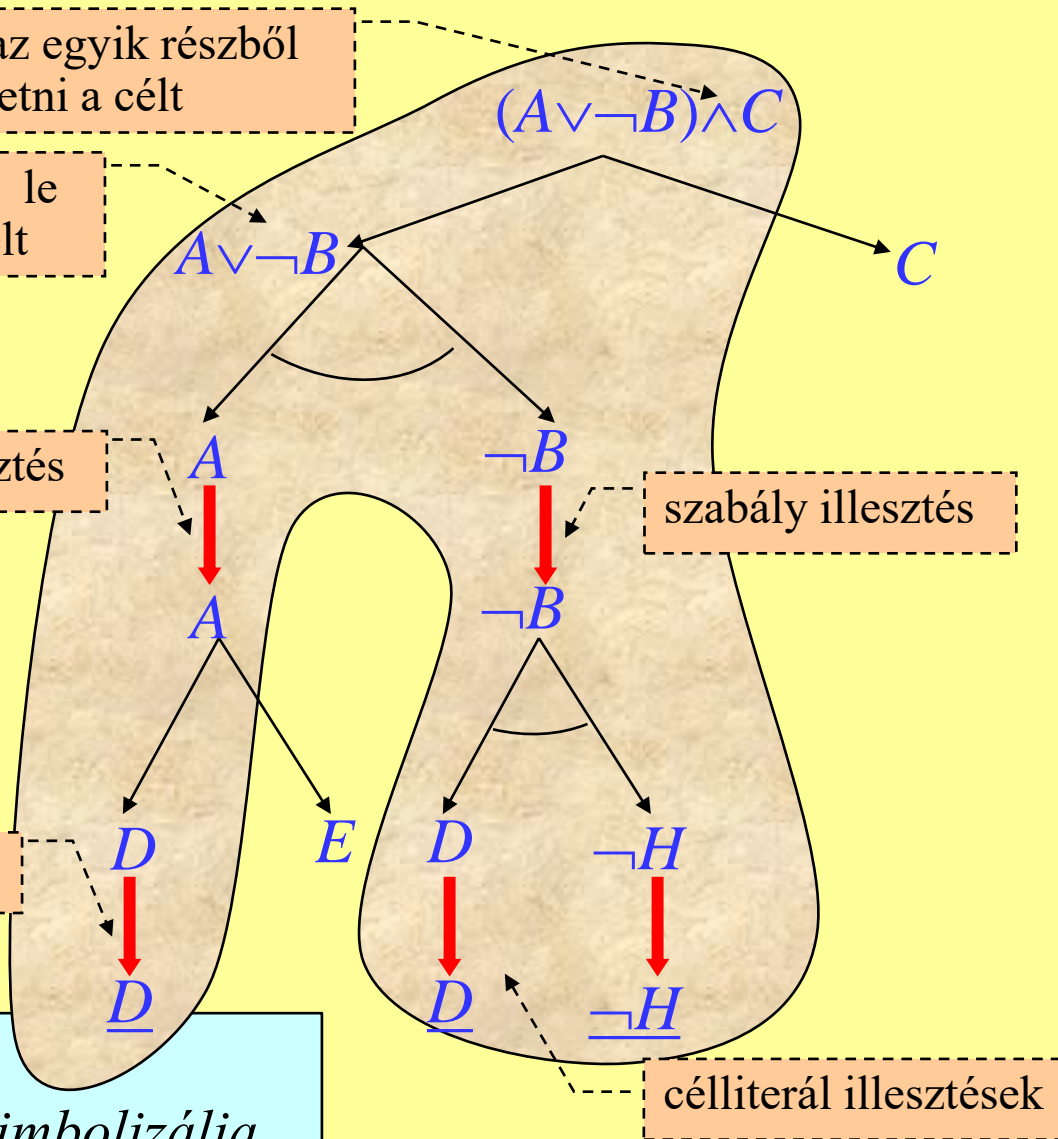
szabály illesztés

célliterál illesztés

célliterál illesztések

*A bizonyítást a tényből
célliterálokba vezető hiperút szimbolizálja*

Mesterséges intelligencia



Visszafelé haladó szabályalapú reprezentáció

célja, hogy a hátrafelé láncolásnál literált literállal kelljen összevetni

□ Tény:

- $L_1 \wedge \dots \wedge L_n$ alakú univerzálisan kötött kifejezés, ahol L_1, \dots, L_n literálok.

□ Szabályok:

- $W \rightarrow L$ alakú univerzálisan kötött kifejezések, ahol L egy literál, a W pedig ÉVF kifejezés

□ Cél:

- egzisztenciálisan kötött **tetszőleges** ÉVF kifejezés

ÉS/VAGY formájú (ÉVF) kifejezések:

- literálok
- $A \wedge B, A \vee B$ alakú formulák, ahol az A és B is ÉVF kifejezés.

Példa visszafelé haladó szabályalapú következtetésre

Tény:

$$A \wedge C \wedge \neg D$$

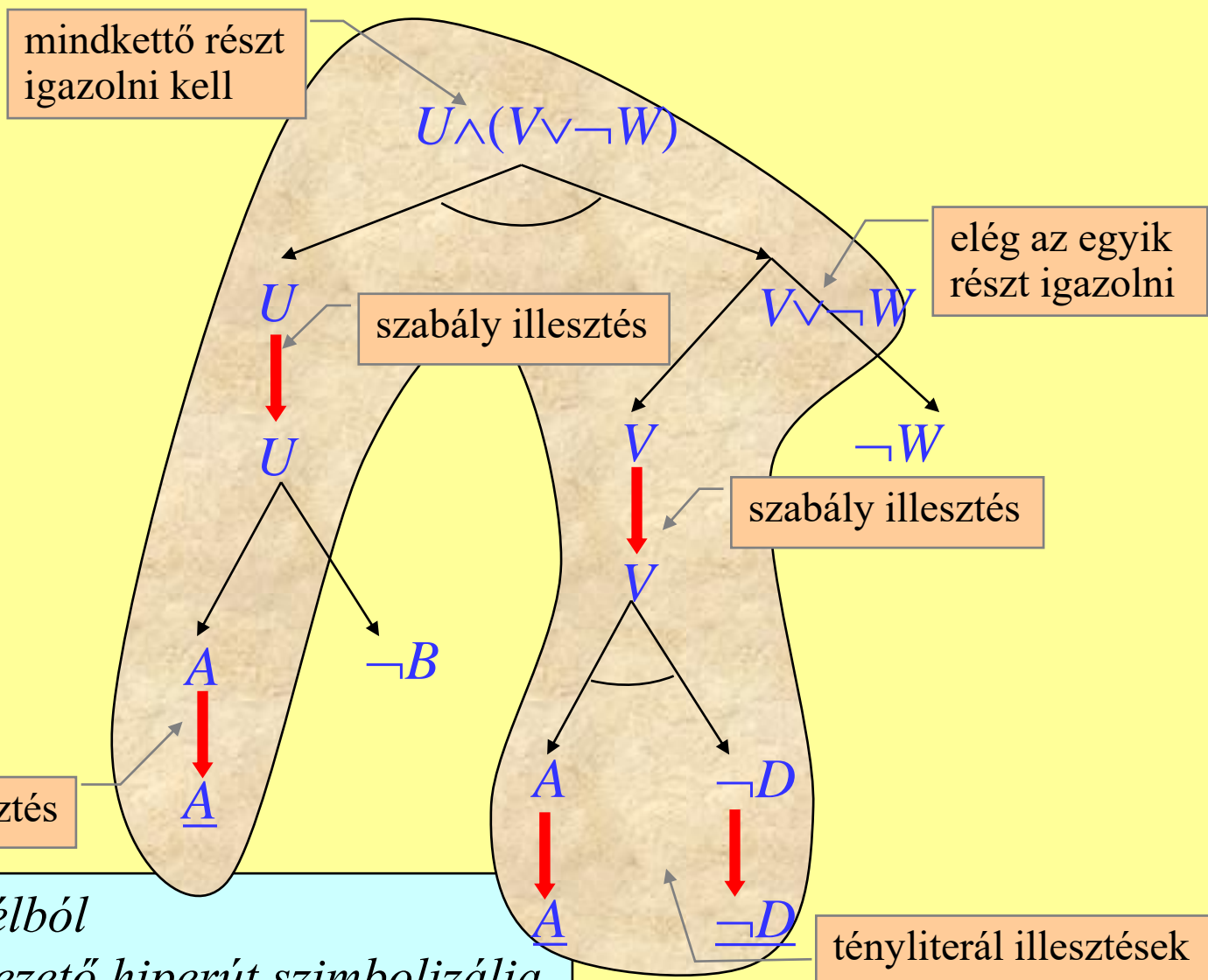
Szabályok:

$$A \vee \neg B \rightarrow U$$

$$A \wedge \neg D \rightarrow V$$

Cél:

$$U \wedge (V \vee \neg W)$$



A bizonyítást a célból tényliterálokba vezető hiperút szimbolizálja

Példa

Fifi és Gyilkos kutyák, Fifi csóválja a farkát, Cili nyávog. A nyávogó állatok a macskák. Az a kutya, amelyik csóválja a farkát, barátságos. A macskák nem félnek a barátságos kutyáktól. Nevezzünk meg olyan kutya-macska párt, ahol a macska nem fél a kutyától!

Formalizálás:

$K(x) \sim x$ kutya

$M(x) \sim x$ macska

$Cs(x) \sim x$ csóvál,

$Ny(x) \sim x$ nyávog

$B(x) \sim x$ barátságos

$F(x,y) \sim x$ fél y -tól

Tény: $K(Fifi) \wedge K(Gyilkos) \wedge Cs(Fifi) \wedge Ny(Cili)$

Szabályok:

$\forall x (Ny(x) \rightarrow M(x))$

$\forall x (K(x) \wedge Cs(x) \rightarrow B(x))$

$\forall x \forall y (K(x) \wedge B(x) \wedge M(y) \rightarrow \neg F(y,x))$

Cél: $\exists x \exists y (M(x) \wedge \neg F(x,y) \wedge K(y))$

válaszadáshoz:
van-e keresett kutya-macska pár

Bizonyítás

Tény: $K(Fifi), K(Gyilkos), Cs(Fifi), Ny(Cili)$

Szabályok: $Ny(x_1) \rightarrow M(x_1)$

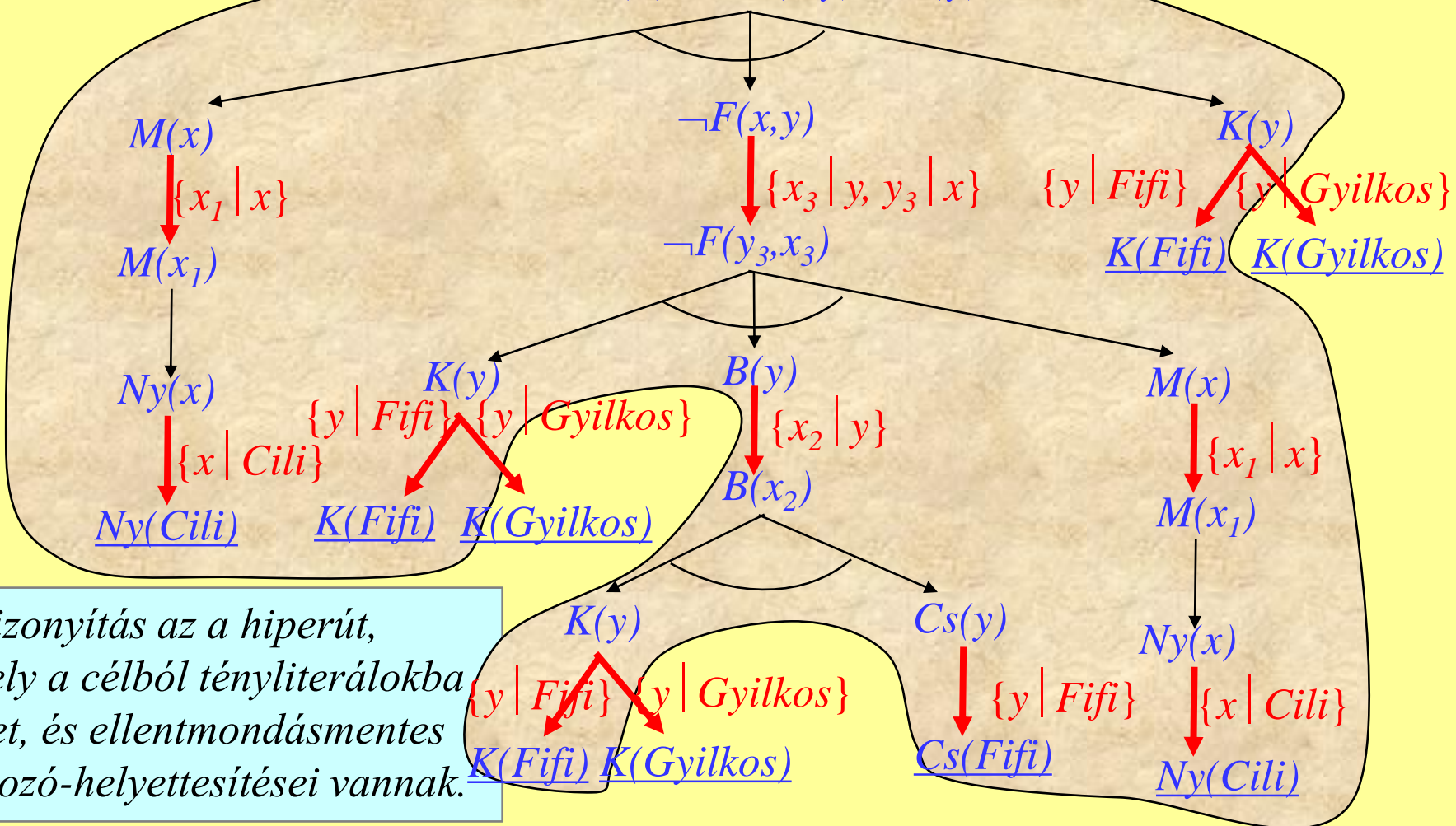
$K(x_2) \wedge Cs(x_2) \rightarrow B(x_2)$

$K(x_3) \wedge B(x_3) \wedge M(y_3) \rightarrow \neg F(y_3, x_3)$

A hiperút változó-helyettesítéseiből olvasható ki a válasz: $\{x \mid Cili, y \mid Fifi\}$

Cél:

$M(x) \wedge \neg F(x, y) \wedge K(y)$



A bizonyítás az a hiperút, amely a célból tényliterálokba vezet, és ellentmondásmentes változó-helyettesítései vannak.

Szabályalapú következtetés = visszalépéses keresés

- ❑ Célja egy bizonyítás keresése, amit egy ÉS/VAGY gráfbeli ellentmondásmentes megoldás gráf reprezentál.
- ❑ Kereső rendszer
 - **Globális munkaterület:** megkezdett bizonyítás (hiperút)
 - **Kereső rendszer szabályai:** láncolások illetve visszalépés
 - **Vezérlési stratégia** (elsődleges): visszalépéses stratégia
 - **Modellfüggő stratégiák**
 - Formulák alakjának kihasználása
 - A tény (cél) illesztése előzze meg a szabály-illesztést.
 - **Heurisztikák:** az adott feladat speciális ismeretei
 - Metaszabályok, kiértékelő függvény