


# Biológiai rendszerek modellellenőrzése bayesi megközelítésben

Gál Tamás Zoltán

Szoftver verifikáció és validáció  
kiselőadás, 2013. ősz

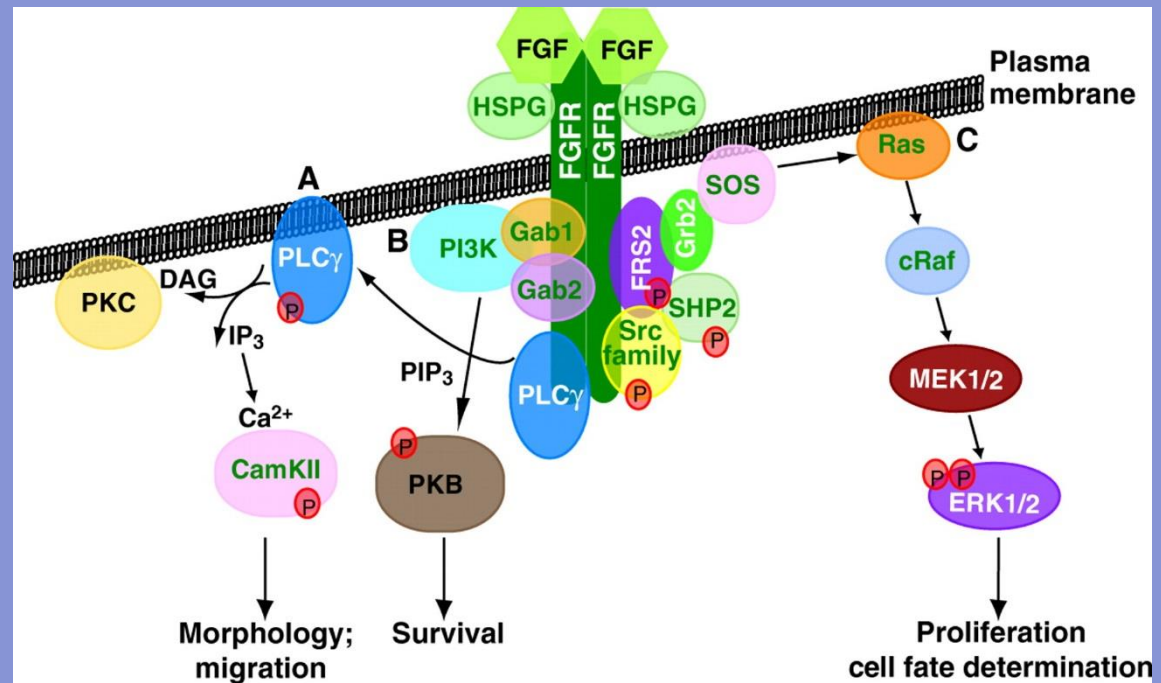
# Biológiai folyamatok



- Sejtalkotók közötti kölcsönhatások
    - Molekulák, fehérjék, gének, ...
  - Tudományterület: rendszerbiológia
  - Példa: (nem csak) **fibroblaszt növekedési faktorok** ( $F_{\text{ibroblast}}$   $G_{\text{rowth}}$   $F_{\text{actor}}$ )
    - Fibroblaszt: kollagén kibocsátás, kötőszövet alkotója (extracelluláris mátrix)
    - Oszródás (+differentiálódás, túlélés, migráció, stb.) elősegítése
- 

# FGF


- Embrionális fejlődés
- Érképződés
- Csont-, és porcképződés
- Sebgyógyulás



Forrás: <http://thenode.biologists.com/tag/fgf/>

# Modellezés



- Szeretnénk...
    - Okokat feltárni, ellenőrizni, jósolni
  - Kérdések:
    - (Mennyi idő alatt gyógyul be a seb...?)
    - *„Nagyobb lesz-e 0,5-nél a valószínűsége annak, hogy Grb2 20 időegységen belül hozzákapcsolódik FRS2-höz?”*
- 

# FGF PRISM modellje

```
formula Frs = relocFrs2=0 ∧ degFrs2=0; // FRS2 not relocated or degraded

module FRS2

  FrsUbi   : [0..1] init 0; // ubiquitin modification of FRS2
  relocFrs2 : [0..1] init 0; // FRS2 relocated
  degFrs2   : [0..1] init 0; // FRS2 degraded
  Y196P     : [0..1] init 0; // receptor Y196 phosphorylated
  Y306P     : [0..1] init 0; // receptor Y306 phosphorylated
  Y471P     : [0..1] init 0; // receptor Y471 phosphorylated
  // compounds bound to FRS2
  FrsFgfr   : [0..1] init 0; // 0: FGFR not bound, 1: FGFR bound
  FrsGrb     : [0..2] init 0; // 0: Grb2 not bound, 1: Grb2 bound, 2: Grb2:Sos bound
  FrsShp     : [0..1] init 0; // 0: Shp2 not bound, 1: Shp2 bound
  FrsSrc     : [0..8] init 0; // 0: Src not bound, 1: Src bound, 2: Src:Spry, 3: Src:SpryP,
  // 4: Src:SpryP:Cbl, 5: Src:SpryP:Grb, 6: Src:SpryP:Grb:Cbl
  // 7: Src:SpryP:Grb:Sos, 8: Src:SpryP:Grb:Sos:Cbl

  // FGFR+FRS2 ↔ FGFR:FRS2 (4)
  [fgfr_bind] Frs ∧ FrsFgfr=0 → 10 : (FrsFgfr'=1);
  [fgfr_rel]  Frs ∧ FrsFgfr=1 → 0.001 : (FrsFgfr'=0);

  // phosphorylation of receptors (5)
  [] Frs ∧ Y653P=1 ∧ Y654P=1 ∧ FrsFgfr=1 ∧ Y196P=0 → 0.2 : (Y196P'=1); // Y196
  [] Frs ∧ Y653P=1 ∧ Y654P=1 ∧ FrsFgfr=1 ∧ Y306P=0 → 0.2 : (Y306P'=1); // Y306
  [] Frs ∧ Y653P=1 ∧ Y654P=1 ∧ FrsFgfr=1 ∧ Y471P=0 → 0.2 : (Y471P'=1); // Y471

  // dephosphorylation of Y196 (6) - remove Src if bound
  [] Frs ∧ FrsShp=1 ∧ Y196P=1 ∧ FrsSrc=0 → 12 : (Y196P'=0);
  [src_rel] Frs ∧ FrsShp=1 ∧ Y196P=1 ∧ FrsSrc>0 → 12 : (Y196P'=0) ∧ (FrsSrc'=0);

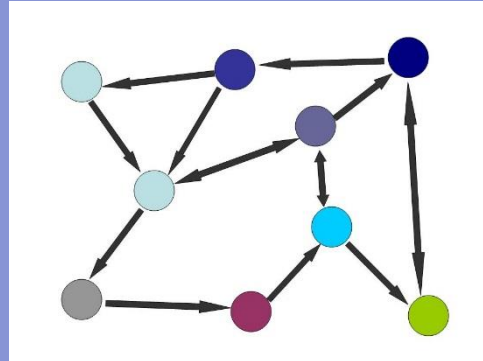
  // dephosphorylation of Y306 (6) - remove Grb2 if bound
  [] Frs ∧ FrsShp=1 ∧ Y306P=1 ∧ FrsGrb=0 → 12 : (Y306P'=0);
  [grb_rel] Frs ∧ FrsShp=1 ∧ Y306P=1 ∧ FrsGrb>0 → 12 : (Y306P'=0) ∧ (FrsGrb'=0);

  // dephosphorylation of Y471 (6) - remove Grb2 if bound
  [] Frs ∧ FrsShp=1 ∧ Y471P=1 ∧ FrsGrb=0 → 12 : (Y471P'=0);
  [grb_rel] Frs ∧ FrsShp=1 ∧ Y471P=1 ∧ FrsGrb>0 → 12 : (Y471P'=0) ∧ (FrsGrb'=0);

endmodule
```

Forrás: J. Heath et al.: Probabilistic model checking of complex biological pathways, Theoretical Computer Science 391 (2008) 239–257

# Milyen modell?



Determinisztikus:  
biológiai folyamatok  
nem ilyenek



Sztochasztikus:  
**Probabilistic Model  
Checking (PMC)**

# PMC probléma



- $M$ : sztochasztikus modell  $\rightarrow$  adott
- $S$ : állapotok halmaza  $\rightarrow$  adott
- $s_0$ : kezdőállapot  $\rightarrow$  adott
- $\Phi$ : temporális logikai kifejezés  $\rightarrow$  vizsgált tulajdonság
- $\Theta \in [0,1]$ : valószínűségi határérték  $\rightarrow$  elfogadás e fölött

$$M, s_0 \models P_{\geq \Theta}(\Phi) ?$$


# Temporális logika

$M, s_0 \models P_{\geq \theta}(\Phi) ?$











- *BLTL: Bounded Linear Temporal Logic*
  - *SV: modell állapotváltozóinak halmaza*
  - *Relációkifejezés:  $x \sim v: \sim \in \{\geq, \leq, =\}; x \in SV; v \in R$*
  - *R.kif.-ek között:  $\vee, \wedge, \neg$*
  - *Temporális operátorok:  $U^t, F^t, G^t$*
- $\Phi$ : *BLTL-lel megfogalmazott formula*
- **Bejárás** (*M egy futtatása*):  $\sigma = (s_0, t_0), (s_1, t_1), \dots$ 
  - *Erre értelmezhető:*

$\sigma \models \Phi ?$

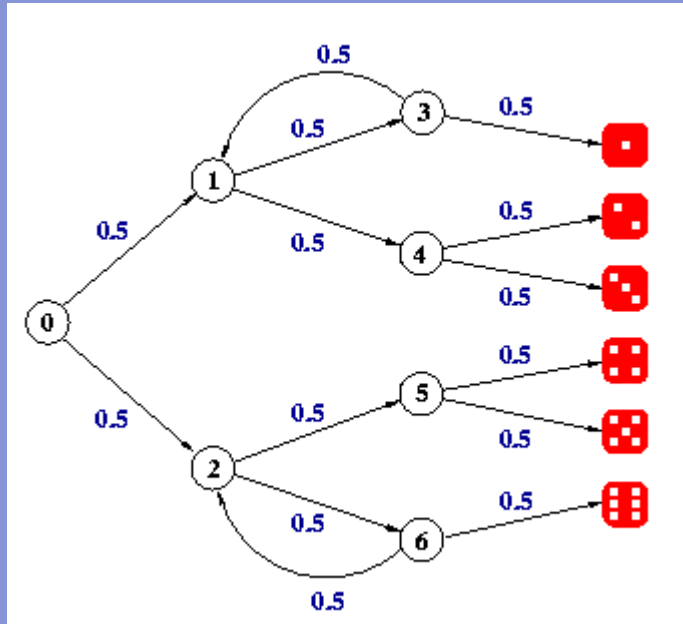


# PBLTL

$$M, s_0 \models P_{\geq \Theta}(\Phi) ?$$


- *Ha veszünk egy bejárást, erre már meg lehet kérdezni, hogy  $\sigma \models \Phi$  ?*
  - *PBLTL: egy  $P_{\geq \Theta}(\Phi)$  formula, amely akkor igaz, ha a fenti kérdés bekövetkezésének a valószínűsége nagyobb, vagy egyenlő, mint  $\Theta$*
  - *Innentől vizsgálható, hogy  $M$  kielégíti-e ezt a formulát*
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 

# Milyen modell?



PMC

Numerikus, egzakt: kis  
rendszerekre alkalmas csak

**Közelítő**

Valószínűségek  
becslése

**Hipotézis  
tesztelés**

# Hipotézis tesztelés



- *Kérdés:  $M, s_0 \models P_{\geq \Theta}(\Phi)$  ?*


- *Hipotézisek:*

$$H_0: P_{\geq \Theta}(\Phi)$$

$$H_1: P_{< \Theta}(\Phi)$$

- *Itt csak azt kell eldönteni, hogy a két hipotézis közül melyik valószínűbb*

- *Hatékonyabb a becslésnél, ha  $\Theta$  lényegesen különbözik a valódi (ismeretlen) küszöbértéktől ( $p$ )*



# Hipotézis tesztelős PMC-k



1. *Wald SPRT módszere – frekventista szemlélet*

$$\Theta_0 < \Theta_1 \quad H_0 \leftarrow \Theta_0 \geq \prod_{i=1}^k \frac{p_1(X_i)}{p_0(X_i)} \geq \Theta_1 \Rightarrow H_1$$

2. *PRISM: fix mintaszám  $\#\sigma_{SAT} / \#\sigma$*

3. *bayesi*

- *kevesebb mintaadatot igényel*
  - *„a priori” tudással feljavítható*
- 

# Bayes-féle megközelítés



- Legyen  $p$  " $\sigma \models \Phi$ ?" valószínűsége. (Ismeretlen!) Ekkor a PMC probléma:

$$H_0: p \geq \Phi$$

$$H_1: p \leq \Phi$$

- $\sigma_i$ : *Több független, azonos eloszlású bejárás*
- $X_i = 1$ , ha  $\sigma_i \models \Phi$ , amúgy pedig  $X_i = 0$
- $p$ -ről lehet „*a priori*” tudásunk, amelyet  $g(\cdot)$  prior eloszlás határoz meg



# Bayes-féle megközelítés



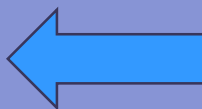
- Adottak  $X_1, \dots, X_n$  bejárási kimenetek, ezek halmaza legyen  $d$ .

$$P(H_0|d) = \frac{P(d|H_0)P(H_0)}{P(d)}$$

*posterior* *prior 1*

*prior 2*

Modell  
„ok”



Futtatások  
„okozatok”



# Bayes-féle megközelítés



- *Ily módon kifejezzük  $H_0$  és  $H_1$  posterior értékeit, elosztva őket egymással kapjuk:*

$$\frac{P(H_0|d)}{P(H_1|d)} = \frac{P(d|H_0)P(H_0)}{P(d|H_1)P(H_1)} = B \frac{P(H_0)}{P(H_1)}$$

- ***B: Bayes-faktor*** ->  $H_0$  és  $H_1$  posteriorok közötti relatív konfidencia mértéke (fix priorok mellett)



# Bayes-féle megközelítés



- *T: szabadon választott küszöb*  
(Jeffreys:  $T := 100$ , mi:  $T := 10000$ )
- *Döntés:*  $B > T \rightarrow H_0$  ,  $B < 1/T \rightarrow H_1$
- *Bayes-faktor a mintaadatokból (d) kiszámítható, a hipotézis modellek prior értékének felhasználásával*



# Bayes-faktor



- $d = (X_1, \dots, X_n)$  bejárás-megfigyelések
- Együttes eloszlás sűrűségfüggvénye:  $f(d | \cdot)$
- *i.i.d.*  $\Rightarrow f(d | \cdot) = f(X_1 | \cdot) \dots f(X_n | \cdot)$
- Prior eloszlás sűrűségfüggvénye:  $g(\cdot)$

$$B = \frac{P(X_1, \dots, X_n | H_0)}{P(X_1, \dots, X_n | H_1)} = \frac{\int_{\Theta} f(X_1 | u) \dots f(X_n | u) g(u) du}{\int_0^{\Theta} f(X_1 | u) \dots f(X_n | u) g(u) du}$$



# Algoritmus



$n := 0$  //eddigyi bejárások száma

$x := 0$  // $\Phi$ -t kielégítő eddigyi bejárások

repeat

$\sigma := \text{futtass\_egy\_bejarast}(M)$

$n++$

$\sigma.\text{satisfies}(\Phi) ? x++ : \text{pass}$

$B := \text{Bayes\_faktor}(n,x)$


until ( $B > T$  or  $B < 1/T$ )

$B > T ? \text{return } H0\_accepted : \text{return } H1\_accepted$



# Kiegészítések



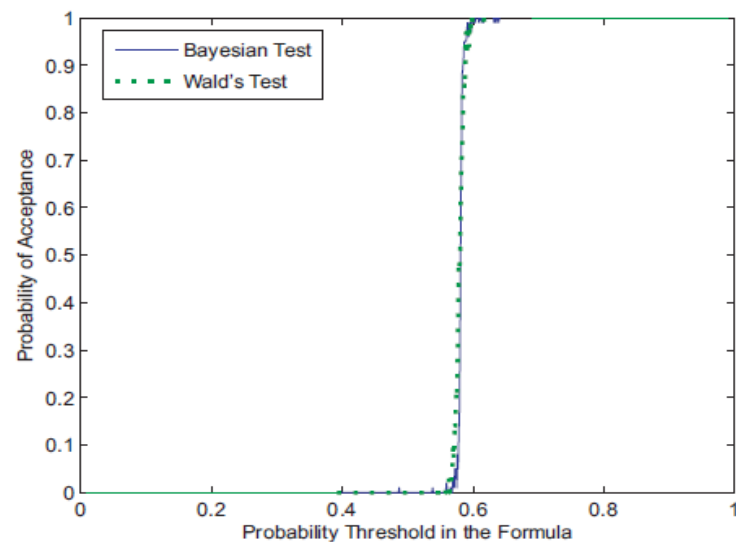
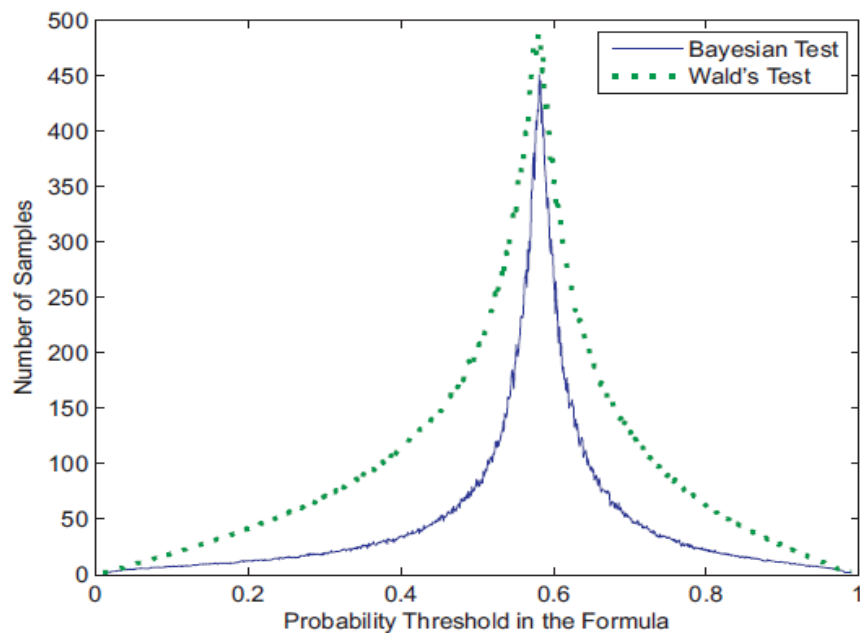
- Tétel 1: Az algoritmus 1 valószínűséggel véget ér ún. Béta-priorokkal és i.i.d. bejárásokkal
  - Tétel 2: A hibázás valószínűsége „elég kicsi” (nem rosszabb korábbi módszerekénél)
  - Bármilyen prior közelíthető Béta-priorokkal
  - Nem nagy baj, ha „rossz” (nem informatív, félrevezető) a prior, az algoritmus akkor is lefut, csak több lépésben, és kevésbé „éles”
- 

# Példa: FGF modell

- „Nagyobb lesz-e 0,5-nél a valószínűsége annak, hogy Grb2 20 időegységen belül hozzákapcsolódik FRS2-höz?”

$$H_0: M \models P_{\geq 0.5} [F^{20} (FRS2_{GRB2} > 0)]$$

- $T = 10000$ ; egzakt módszer nem működik





Thomas Bayes  
(1701-1761)

**Köszönöm a figyelmet!**