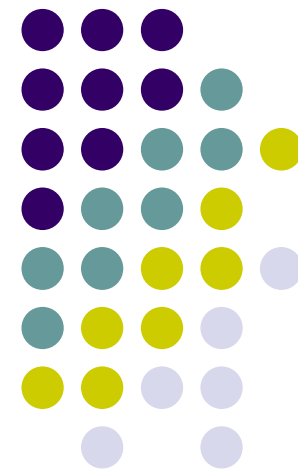


Heuristiline meetod tõenäosuslike boole'i võrkude genereerimiseks kasutades siirde tõenäosusmaatriksit



Mairo Leier
2010



Sisukord

- Boole'i võrkude ülevaade (BN)
- Geeni siirded
- Tõenäosuslikud boole'i võrgud (PBN)
- Algoritmid PBN arvutamiseks
- Algoritmide efektiivsus



Sissejuhatus

- Tõenäosuslik Boole'i Võrk (PBN) on Boole'i Võrgu(BN) laiendus
- BN: geeni olemasolu 0 või 1
- PBN: geeni olemasolu 0...1
- PBN: võetakse arvesse vigade esinemist
- PBN dünaamikat saab õppida Markov'i ahela teooria alusel
 - Siirde tõenäosus
 - Stabiilne olek

Võimalikud kasutusala

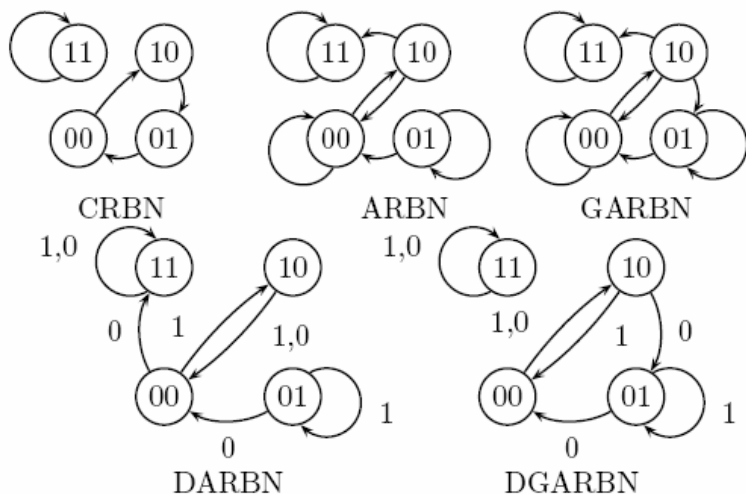
- Haiguskuulude jälgimine
- Organismi füsioloogia
- Diagnooside panemine
- Ravimite arendus





Boole'i võrgustiku klassifikatsioon

- DDN – diskreetsed dünaamilised võrgud
- CRBN – klassikalised võrgud, sünkroonne uuendus
 - kõikide geenide olekute uuendamine igal ajahetkel ($t+1$)
- ARBN – asünkroonsed võrgud
 - geenide olekute uuendamine juhuslik
- DGARBN – deterministlikud (ennustavad) üldised asünkroonsed võrgud
- DARBN – deterministlikud asünkroonsed võrgud



Geeni uuendustsüklid, millest tulenevad tõeväärtustabelid

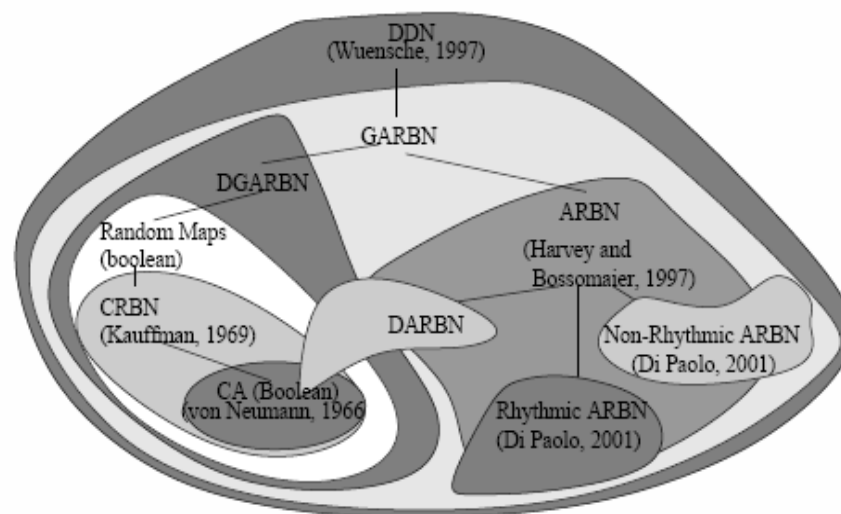


Figure 1: Classification of Random Boolean Networks



BN ülevaade

- BN $G(V, F)$ koosneb tippudest (näit. geenid) ($V = \{v_1, v_2, \dots, v_n\}$) ja boole'i funtsioonidest (F)
- Tipu v_i olek mingil ajahetkel t võib olla 0 või 1
- $v(t) = (v_1(t), v_2(t), \dots, v_n(t))^T$ kutsutakse Geeni Aktiivsuse Profiiliks (GAP)
- Geeni aktiivsust kirjeldab geeni siirdetabel



Geeni siirdetabel

abc(t) abc(t+1)

000 011

001 101

010 111

011 011

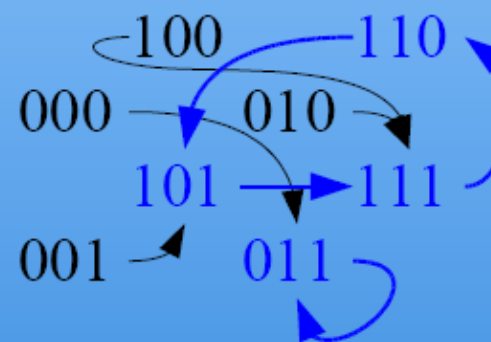
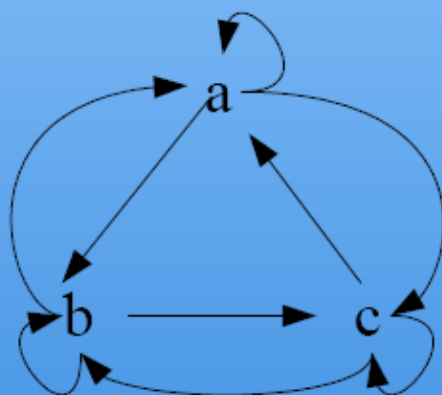
100 111

101 111

110 101

111 110

$N=3, K=3$



- Lõplikud olekud (2^N) (N – geenide arv võrgustikus)
 - Punktilised olekud - peale uuendamist jääb olek samaks
 - Tsüklilised olekud – geeni olek muutub teatud tsükli alusel

Geenide siirde tõenäosus



- P - siirde tõenäosus
- 000 või 111 olekust protsess enam ei välju kuni ei teki mutatsiooni

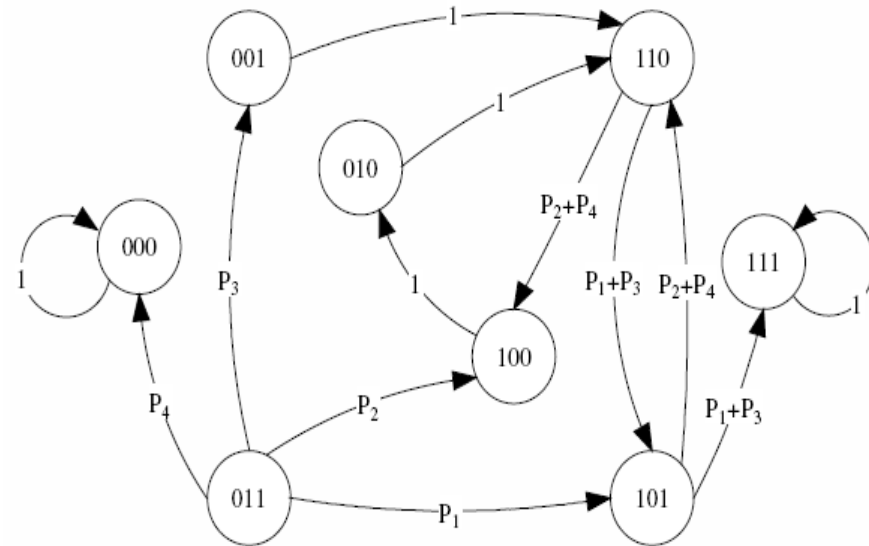


Fig. 5. State transition diagram from Example 1.

PBN



- Geenivõrgustiku juhtimiseks on võimalik muuta osasid geene kuid mitte kõiki
- Geenide jälgimisel välistegurid:
 - Geeni avaldumine stohhastiline
 - Müra keerulise mõõtmisprotsessi tõttu
 - Sõltuvus välisteguritest, mida ei mõõdeta (proteiinid, välistingimused jm)

Geeni transformatsiooni tõenäosus



- A on kõikide boole'i võrkude maatriksite summa A_i

$$A = \sum_{i=1}^N q_i A_i$$

- Kus q_i on BN maatriksi A_i valimise tõenäosus.
- Leiame q_i ehk mingi geeni võrgustiku esinemise kui on antud A ja A_i



Algoritm 1

- Boole'i võrgu A lineaarse kihtlahutuse leidmine

$$A = \begin{pmatrix} 0,1 & 0,7 \\ 0,9 & 0,3 \end{pmatrix} = 0,1 \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0,0 & 0,6 \\ 0,9 & 0,3 \end{pmatrix} \equiv 0,1A_1 + R_2$$

$$R_2 = \begin{pmatrix} 0,0 & 0,6 \\ 0,9 & 0,3 \end{pmatrix} = 0,3 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0,0 & 0,6 \\ 0,6 & 0,0 \end{pmatrix} \equiv 0,3A_2 + R_3$$

$$R_3 = \begin{pmatrix} 0,0 & 0,6 \\ 0,6 & 0,0 \end{pmatrix} = 0,6 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0,0 & 0,0 \\ 0,0 & 0,0 \end{pmatrix} \equiv 0,6A_3 + R_4$$

$$R_4 = 0,0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv 0,0A_4$$

$$A = 0,1A_1 + 0,3A_2 + 0,6A_3$$

Sammud:

- 0. $R_1 = A; k=0$
- 1. $k=k+1$
- 2. Valime väikseima R_k 'st nullist erineva arvu q_k . Leiame kõikidest teistest veergudest suurima numbri.
- 3. $R_{k+1} = R_k - q_k A_k$
- 4. Kui $R_{k+1} \neq 0$, siis samm 1
- 5.

$$A = \sum_{i=1}^M q_k A_k$$

- Summaarne tõenäosus on 1

Algoritm 2

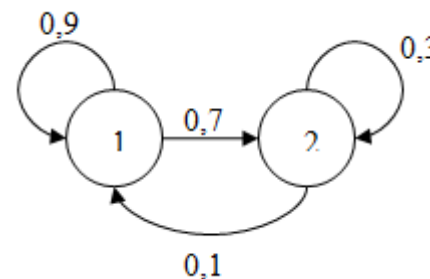


Näide:

$$A = \begin{pmatrix} 0,1 & 0,7 \\ 0,9 & 0,3 \end{pmatrix} = 0,7 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0,1 & 0,0 \\ 0,2 & 0,3 \end{pmatrix} \equiv 0,7A_1 + R_2$$
$$R_2 = \begin{pmatrix} 0,1 & 0,0 \\ 0,2 & 0,3 \end{pmatrix} = 0,2 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0,1 & 0,0 \\ 0,0 & 0,1 \end{pmatrix} \equiv 0,2A_2 + R_3$$
$$R_3 = \begin{pmatrix} 0,1 & 0,0 \\ 0,0 & 0,1 \end{pmatrix} = 0,1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0,0 & 0,0 \\ 0,0 & 0,0 \end{pmatrix} \equiv 0,1A_3 + R_4$$
$$R_4 = 0,0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv 0,0A_4$$
$$A = 0,7A_1 + 0,2A_2 + 0,1A_3$$

Sammud:

- Kõik sammud samad v.a. 2. samm kus valime mõlema veeru suurimatest arvudest väiksema



Kahe geeni siirde näide

Entroopia definitsioon



- Entroopia aitab mõõta jaotusfunktsiooni mitteühtlasust
- Entroopia definitsioon:
$$H_{bin}(p) \equiv -p \log p - (1-p) \log(1-p)$$
- Entroopia $H_{bin}(p)$ on suurim $p = 1/2$ puhul
- p - kahe teineteisest sõltuva tulemuse tõenäosus, mis summaarselt on 1

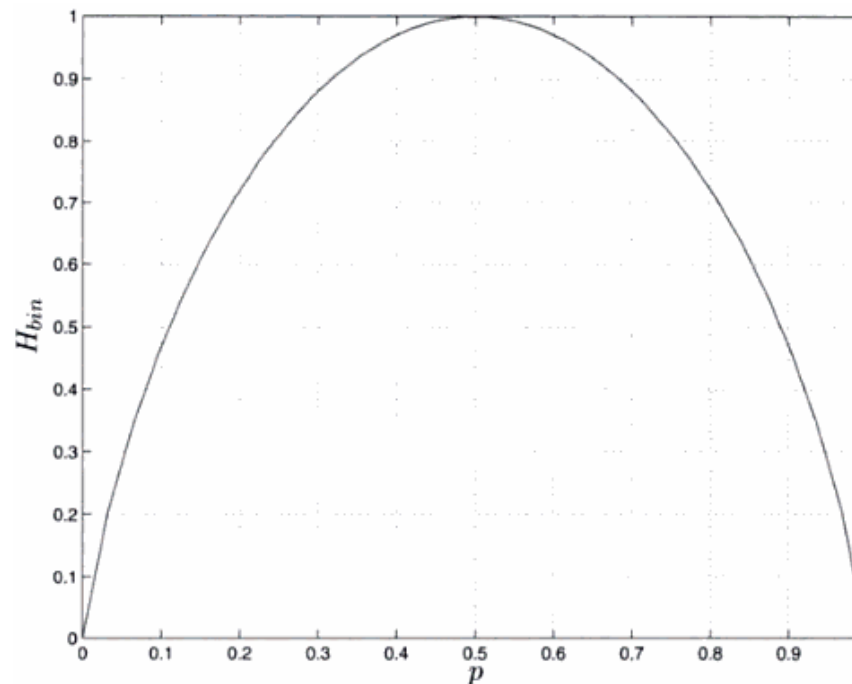


Figure 11.1. Binary entropy function $H(p)$.

Algoritmide tõhusus



- Entroopia leidmine:

$$-\sum_{i=1}^M q_i \log q_i$$

- Algoritm 1:

$$-\sum_{i=1}^M q_i \log q_i = -0,1 \cdot \log 0,1 - 0,3 \cdot \log 0,3 - 0,6 \cdot \log 0,6 \approx 0,38$$

- Algoritm 2:

$$-\sum_{i=1}^M q_i \log q_i = -0,7 \cdot \log 0,7 - 0,2 \cdot \log 0,2 - 0,1 \cdot \log 0,1 \approx 0,347$$

- Väiksema entroopiaga algoritmi puhul peab tegema vähem arv kihtlahutuse tehteid ehk jõuab kiiremini lõpptulemuseni

Kasutatud kirjandus



- A Heuristic Method for Generating Probabilistic Boolean Networks from a Prescribed Transition Probability Matrix, W. Ching, etc..., 2008, IEEE
- From Boolean to Probabilistic Boolean Networks as Models of Genetic Regulatory Networks, Ilya Shmulevich, etc..., 2002, IEEE
- Generating Probabilistic Boolean Networks from a Prescribed Transition Probability Matrix, W.-K. Ching, X. Chen, etc..., 2009, IET
- Probabilistic Boolean Networks: A Rule-Based Uncertainty Model for Gene Regulatory Networks, I. Shcmulevich, etc..., 2002, IEEE
- Introduction to Random Boolean Networks, Carlos Gershenson, Univ. Brussel
- Increasing Detection Performance of Surveillance Sensor Networks, Litvak, Altaf, Barbu, etc
- Classification of Random Boolean Networks, C. Gershenson, 2002, Univ. Brussel
- Quantum Computation and Quantum Information, Michael A., Nielsen, etc