



IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

IKI 30320: Sistem Cerdas

Kuliah 17: Efficient Probabilistic Reasoning

Ruli Manurung

Fakultas Ilmu Komputer
Universitas Indonesia

26 November 2007



Outline

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- 1 Review & kompleksitas
- 2 Independence = efficiency
- 3 Bayes' Rule
- 4 Ringkasan



Outline

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- 1 Review & kompleksitas
- 2 Independence = efficiency
- 3 Bayes' Rule
- 4 Ringkasan



Review: Probabilistic Inference

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

Dengan *joint probability distribution*, probability sembarang proposition bisa dihitung sbg. jumlah probability sample point di mana ia bernilai true.

	<i>toothache</i>		\neg <i>toothache</i>	
	<i>catch</i>	\neg <i>catch</i>	<i>catch</i>	\neg <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
\neg <i>cavity</i>	.016	.064	.144	.576

Mis.: menghitung *conditional probability*:

$$\begin{aligned} P(\neg \text{cavity} | \text{toothache}) &= \frac{P(\neg \text{cavity} \wedge \text{toothache})}{P(\text{toothache})} \\ &= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4 \end{aligned}$$



Normalisasi

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

$$P(\neg cavity | toothache) = P(\neg cavity \wedge toothache) / P(toothache)$$

$$P(cavity | toothache) = P(cavity \wedge toothache) / P(toothache)$$

Perhatikan bahwa pembagiannya sama!

Disebut **konstanta normalisasi** $\alpha \rightarrow$ supaya jumlahnya 1.

$$\mathbf{P}(Cavity | toothache) = \alpha \mathbf{P}(Cavity, toothache)$$

$$= \alpha [\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$$

$$= \alpha [\langle 0.108, 0.016 \rangle + \langle 0.012, 0.064 \rangle]$$

$$= \alpha \langle 0.12, 0.08 \rangle = \langle 0.6, 0.4 \rangle$$

Ide utama:

α dapat diperoleh ketika menghitung **distribusi** utk. *query variable* (mis. *Cavity*) \rightarrow semua kemungkinan nilai.



Beberapa istilah

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

Contoh sebuah "query"

- $P(\neg cavity | toothache)$
- "Berapakah probabilitas gigi Anto tidak berlubang jika diketahui ia menderita sakit gigi?"
- (Bandingkan dengan logical query: "Jika Anto sakit gigi, apakah giginya tidak berlubang?")

- **Query variable**: variable yang ingin dihitung probabilitasnya (*cavity*)
- **Evidence variable**: variable yang kita tahu nilainya (*toothache*)
- **Hidden variable**: variable yang tidak kita tahu nilainya (*catch*)

Inference dengan (full) joint probability distribution:

Tentukan nilai *evidence variable*, jumlahkan probabilitas *query variable* untuk semua kemungkinan nilai *hidden variable*.



Complexity inference dgn joint distribution

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- Jika ada n random variable Boolean, tabel full joint distribution berisi $O(2^n)$ nilai.
- **Space complexity** eksponensial dalam n . Untuk disimpan saja tidak feasible. Apalagi ***menghitung semua nilainya dari data empiris!***
- **Time complexity** juga eksponensial dalam n ($O(2^n)$).

Dalam kenyataan, masalah probabilistic reasoning melibatkan **ribuan** *random variable*. Kita butuh metode yang lebih efisien!



Outline

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- 1 Review & kompleksitas
- 2 Independence = efficiency
- 3 Bayes' Rule
- 4 Ringkasan



Independence

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

Kompleksitas jauh lebih efisien jika diketahui ada random variable yang saling **independent**.

Dua proposition a dan b independent jhj

$$P(a|b)=P(a) \text{ atau } P(b|a)=P(b) \text{ atau } P(a \wedge b)=P(a) \times P(b)$$

Dua variable A dan B independent jhj

$$\mathbf{P(A|B)=P(A)} \text{ atau } \mathbf{P(B|A)=P(B)} \text{ atau } \mathbf{P(A, B)=P(A)P(B)}$$

Contoh: 2 dadu, d_1 dan d_2 , dilempar.

- $P(a \wedge b) = P(a|b) \times P(b) \rightarrow$ (Product rule)
- $P(d_1 = 2 \wedge d_2 = 4) = P(d_1 = 2 | d_2 = 4) \text{ times } P(d_2 = 4)$
- $P(d_1 = 2 \wedge d_2 = 4) = P(d_1 = 2) \times P(d_2 = 4) = 1/6 \times 1/6 = 1/36$

Tabel *full joint distribution* untuk lemparan n dadu = 2^n nilai!



Contoh lain

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

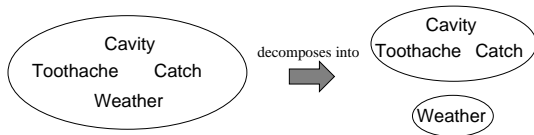
Independence
= efficiency

Bayes' Rule

Ringkasan

Mis. ditambahkan variable *Weather* = $\langle \text{sunny, rain, cloudy, snow} \rangle$ ke masalah dokter gigi (*Cavity*, *Toothache*, *Catch*)

- Tabel full joint distribution-nya ada 32 nilai.
- Tapi cuaca dan gigi tidak berhubungan...
- Untuk menghitung probabilitas *proposition* mengenai cuaca dan gigi, kalikan probabilitasnya, mis.:
 $P(\text{toothache} \wedge \text{Weather} = \text{cloudy}) =$
 $P(\text{toothache}) \times P(\text{Weather} = \text{cloudy})$
- $P(\text{Weather})$ cukup disimpan secara terpisah. $8+4=12$ nilai.





Jadi ingat ini...

IKI30320
Kuliah 17
26 Nov 2007

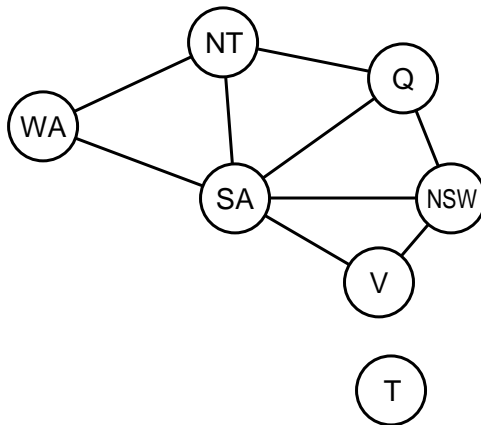
Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan





Conditional Independence

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- Dalam kenyataan, *absolute independence* jarang terjadi.
- Mis. dalam kedokteran gigi, ratusan *variable* biasanya berhubungan.
- Tapi hubungan itu bisa **langsung** atau **tidak langsung**!

Conditional independence

cavity mempengaruhi probabilitas *toothache*

cavity mempengaruhi probabilitas *catch*

Tetapi ini adalah dua efek yang terpisah!

- Nilai *Toothache* memberi kita info tentang nilai *Catch*, tetapi secara tidak langsung (lewat *Cavity*).
- Dkl, kalau kita **sudah tahu** nilai *Cavity*, mengetahui nilai *Toothache* tidak memberikan info baru tentang *Catch*.



Conditional independence = efisien

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- *Catch* dan *Toothache* conditionally independent krn *Cavity*:
 $P(\text{Catch} | \text{Toothache}, \text{Cavity}) = P(\text{Catch} | \text{Cavity})$ (...coba diperiksa!)
- Statement yang ekuivalen:
 $P(\text{Toothache} | \text{Catch}, \text{Cavity}) = P(\text{Toothache} | \text{Cavity})$
 $P(\text{Toothache}, \text{Catch} | \text{Cavity}) = P(\text{Toothache} | \text{Cavity}) P(\text{Catch} | \text{Cavity})$
- Full joint distribution bisa direduksi:
 $P(\text{Toothache}, \text{Catch}, \text{Cavity}) =$
 $P(\text{Toothache}, \text{Catch} | \text{Cavity}) P(\text{Cavity}) =$
 $P(\text{Toothache} | \text{Cavity}) P(\text{Catch} | \text{Cavity}) P(\text{Cavity}) =$
2 nilai + 2 nilai + 1 nilai = 5 nilai

Pengaruh efisiensi

Scr. umum, *conditional independence* mengubah kompleksitas *joint distribution* dari **eksponensial** menjadi **linier** dalam jumlah variable.



Outline

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- 1 Review & kompleksitas
- 2 Independence = efficiency
- 3 Bayes' Rule
- 4 Ringkasan



Dari mana asalnya nilai P ?

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

Jawaban:

Eksperimentasi. Hitung frekuensi.

Mis. lempar koin 100 kali. Muncul *heads* 47x $\rightarrow P(\text{heads}) = 0.47$.

Biasanya, lebih mudah menghitung $P(\text{Akibat}|\text{Sebab})$:

- Ambil sampel 100 orang, mis. penderita *meningitis*
- Mis. ada hipotesa 50 kemungkinan gejala: demam, pusing, leher pegal, mual, dst. \rightarrow lakukan 50 tes pada setiap orang.

Biasanya, lebih sulit menghitung $P(\text{Sebab}|\text{Akibat})$:

- Ambil sampel 50 kelompok @ 100 orang: yang demam, yang pusing, yang leher pegal, yang mual, dst.
- Tes apakah mereka menderita *meningitis*.

Yang **dibutuhkan** dokter (ahli **diagnostik**) adalah $P(\text{Sebab}|\text{Akibat})$!

Adakah cara “menyulap” $P(a|b)$ menjadi $P(b|a)$?



Bayes' Rule

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- **Ada!** Sebuah pendekatan yang praktis (tetapi kontroversial scr. teoritis) adalah **Bayes' Rule**.
- Product rule: $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$
- Bayes' rule: $P(a|b) = \frac{P(b|a)P(a)}{P(b)}$
- Bayes' rule: $P(\text{Sebab}|\text{Akibat}) = \frac{P(\text{Akibat}|\text{Sebab})P(\text{Sebab})}{P(\text{Akibat})}$



Bayes' Rule & Distribution

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- Seringkali ada banyak gejala (akibat), dan satu penyebab. Mis. dokter ingin menghitung:
 $P(\text{meningitis} | \text{leher_pegal} \wedge \text{pusing} \wedge \text{gatal_gatal} \dots)$
- **Masalah:** jika menggunakan Bayes', kita harus hitung *prior probability* kombinasi gejala tsb.
- **Solusi:** yang kita hitung adalah distribusi:
$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} = \alpha P(X|Y)P(Y)$$
- $P(X)$ hanya dibutuhkan untuk normalisasi jumlah probabilitas menjadi 1, jadi bisa diabaikan.



Outline

IKI30320
Kuliah 17
26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- 1 Review & kompleksitas
- 2 Independence = efficiency
- 3 Bayes' Rule
- 4 Ringkasan



Ringkasan

IKI30320

Kuliah 17

26 Nov 2007

Ruli Manurung

Review &
kompleksitas

Independence
= efficiency

Bayes' Rule

Ringkasan

- Inference dengan *full joint distribution* konsepnya sangat mudah dimengerti, tetapi dalam kenyataan tidak feasible (*exponential time & space complexity*)
- Agar inference bisa *tractable*, kita mengambil asumsi **independence**.
- Dalam kenyataan, kita hanya bisa mengambil asumsi **conditional independence**.
- **Bayes' Rule**, ditambah dengan *conditional independence*, adalah mekanisme yang *sangat* berguna.