Modelowanie Niepewności

Na podstawie: AIMA, ch13

Wojciech Jaśkowski

Instytut Informatyki,
Politechnika Poznańska

13 marca 2015

Modelowanie Niepewności

Modelowanie Niepewności
Na podstawie: AIMA, ch13

Wojciech Jaśkowski

Instytut Informatyki,
Politechnika Poznańska

13 marca 2015

Źródła niepewności

Jakie są źródła niepewności?

Źródła niepewności

Jakie są źródła niepewności?

Źródła niepewności

Jakie są źródła niepewności?

- częściowa obserwowalność
- niedeterministyczny

Wynikające często jednak z zawinionej niewiedzy:

- lenistwa i
- ignorancji

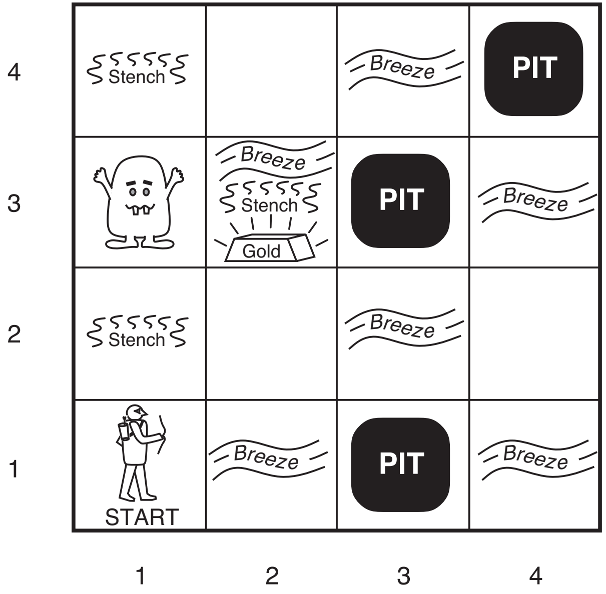
Cel: Racjonalne decyzje w kontekście niepewności.

Źródła niepewności

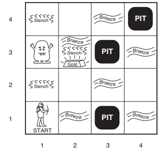
Jakie są źródła niepewności?
- częściowa obserwowalność
- niedeterministyczny
Wynikające często jednak z zawinionej niewiedzy:
- lenistwa i
- ignorancji

Cel: Racjonalne decyzje w kontekście niepewności.

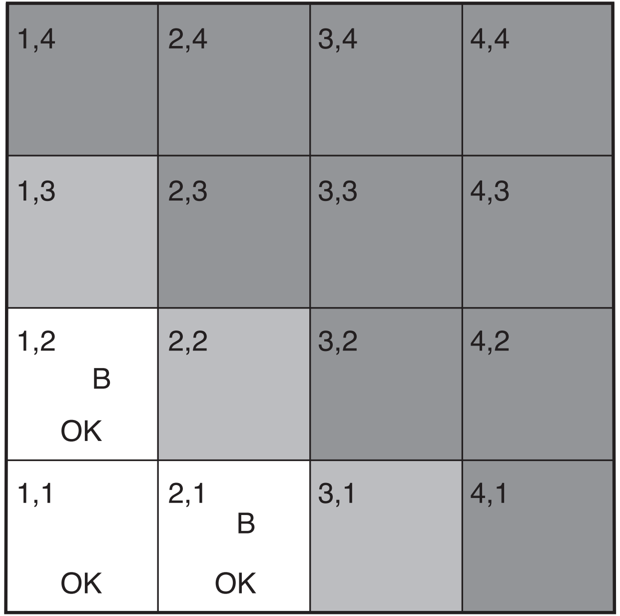
Świat Wumpus'a

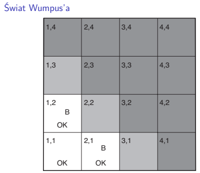


Świat Wumpus'a



Świat Wumpus'a





1. W takiej sytuacji Wumpus nie ma bezpiecznego ruchu, bo bryza w (1,2) i (2,1). Jaki ruch Wumpus powinien więc wykonać? Jakie jest prawdopodobieństwo, że na polu (1,3) jest pułapka? A jakie dla (2,2) i dla (3,1)?

Podstawy: notacja

- **zmienna losowa**

$$Cancer = C = \{\neg c, c\}$$

$$Test = T = \{\neg t, t\}$$

Podstawy: notacja
- zmienna losowa
$Cancer = C = \{\neg c, c\}$
$Test = T = \{\neg t, t\}$

1. rozkład prawd. łącznego: wektor wszystkich kombinacji wartości zmiennych losowych.

Podstawy: notacja

- **zmienna losowa**

$$Cancer = C = \{\neg c, c\}$$

$$Test = T = \{\neg t, t\}$$

- **prawdopodobieństwo zdarzenia**

$$P(c) = 0.3$$

Podstawy: notacja
- zmienna losowa
$Cancer = C = \{\neg c, c\}$
$Test = T = \{\neg t, t\}$
- prawdopodobieństwo zdarzenia
$P(c) = 0.3$

1. rozkład prawd. łącznego: wektor wszystkich kombinacji wartości zmiennych losowych.

Podstawy: notacja

- **zmienna losowa**

$$Cancer = C = \{\neg c, c\}$$

$$Test = T = \{\neg t, t\}$$

- **prawdopodobieństwo zdarzenia**

$$P(c) = 0.3$$

- **rozkład prawdopodobieństwa**

$$\mathbf{P}(Cancer) = \langle P(c), P(\neg c)\rangle = \langle 0.3, 0.7\rangle$$

Podstawy: notacja
- zmienna losowa
$Cancer = C = \{\neg c, c\}$
$Test = T = \{\neg t, t\}$
- prawdopodobieństwo zdarzenia
$P(c) = 0.3$
- rozkład prawdopodobieństwa
$\mathbf{P}(Cancer) = \langle P(c), P(\neg c)\rangle = \langle 0.3, 0.7\rangle$

1. rozkład prawd. łącznego: wektor wszystkich kombinacji wartości zmiennych losowych.

Podstawy: notacja

- **zmienna losowa**

$$Cancer = C = \{\neg c, c\}$$

$$Test = T = \{\neg t, t\}$$

- **prawdopodobieństwo zdarzenia**

$$P(c) = 0.3$$

- **rozkład prawdopodobieństwa**

$$\mathbf{P}(Cancer) = \langle P(c), P(\neg c)\rangle = \langle 0.3, 0.7\rangle$$

- **prawd. łączne**

$$P(c \wedge \neg t) = P(c, \neg t) = 0.4$$

Podstawy: notacja
- zmienna losowa
$Cancer = C = \{\neg c, c\}$
$Test = T = \{\neg t, t\}$
- prawdopodobieństwo zdarzenia
$P(c) = 0.3$
- rozkład prawdopodobieństwa
$\mathbf{P}(Cancer) = \langle P(c), P(\neg c)\rangle = \langle 0.3, 0.7\rangle$
- prawd. łączne
$P(c \wedge \neg t) = P(c, \neg t) = 0.4$

1. rozkład prawd. łącznego: wektor wszystkich kombinacji wartości zmiennych losowych.

Podstawy: notacja

- **zmienna losowa**

$$Cancer = C = \{\neg c, c\}$$

$$Test = T = \{\neg t, t\}$$

- **prawdopodobieństwo zdarzenia**

$$P(c) = 0.3$$

- **rozkład prawdopodobieństwa**

$$\mathbf{P}(Cancer) = \langle P(c), P(\neg c)\rangle = \langle 0.3, 0.7\rangle$$

- **prawd. łączne**

$$P(c \wedge \neg t) = P(c, \neg t) = 0.4$$

- **rozkład prawd. łącznego**

$$\mathbf{P}(Cancer, Test) = \langle P(c \wedge t), P(c \wedge \neg t), P(\neg c, t), P(\neg c, \neg t)\rangle$$

Podstawy: notacja
- zmienna losowa
$Cancer = C = \{\neg c, c\}$
$Test = T = \{\neg t, t\}$
- prawdopodobieństwo zdarzenia
$P(c) = 0.3$
- rozkład prawdopodobieństwa
$\mathbf{P}(Cancer) = \langle P(c), P(\neg c)\rangle = \langle 0.3, 0.7\rangle$
- prawd. łączne
$P(c \wedge \neg t) = P(c, \neg t) = 0.4$
- rozkład prawd. łącznego
$\mathbf{P}(Cancer, Test) = \langle P(c \wedge t), P(c \wedge \neg t), P(\neg c, t), P(\neg c, \neg t)\rangle$

1. rozkład prawd. łącznego: wektor wszystkich kombinacji wartości zmiennych losowych.

Podstawy: prawd. warunkowe

Prawdopodobieństwo warunkowe

$$P(c|t)P(t) = P(c \wedge t)$$

(reguła produkcji)

Podstawy: prawd. warunkowe

Prawdopodobieństwo warunkowe

$P(c|t)P(t) = P(c \wedge t)$

(reguła produkcji)

1. Graficznie: dwa przecinające się zbiory. t zaistniało, więc jakie jest prawd., że zaistnieje c? $P(c|t) = \frac{P(c \wedge t)}{P(t)}$.

(Pełny) rozkład prawd. łącznego

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

(Pełny) rozkład prawd. łącznego

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

Prawd. marginalne (marginalizacja)

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 0]
$P(dziura) = ?$
$\mathbf{P}(Dziura) = ?$

1. $P(dziura) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$
   Marginalizujemy Ból i Test
2. $\mathbf{P}(Dziura) =< 0.2, 0.8 >$
3. $P(dziura \wedge \neg ból) = 0.072 + 0.008 = 0.08$
   Marginalizujemy Test

Prawd. marginalne (marginalizacja)

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 0]
$P(dziura) = ?$
$\mathbf{P}(Dziura) = ?$

Prawd. marginalne (marginalizacja)

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 0]
$P(dziura) = ?$
$\mathbf{P}(Dziura) = ?$
[zadanie 1]
$P(dziura \wedge \neg ból) = ?$

Prawd. marginalne (marginalizacja)

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 0]
$P(dziura) = ?$
$\mathbf{P}(Dziura) = ?$
[zadanie 1]
$P(dziura \wedge \neg ból) = ?$

1. $P(dziura) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$
   Marginalizujemy Ból i Test
2. $\mathbf{P}(Dziura) =< 0.2, 0.8 >$
3. $P(dziura \wedge \neg ból) = 0.072 + 0.008 = 0.08$
   Marginalizujemy Test

Prawd. marginalne (marginalizacja)

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 0]
$P(dziura) = ?$
$\mathbf{P}(Dziura) = ?$
[zadanie 1]
$P(dziura \wedge \neg ból) = ?$

Ogólnie:

$$\mathbf{P}(\mathbf{Y}) = \sum_{x \in \mathbf{X}} \mathbf{P}(\mathbf{Y}, x) = \sum_{x \in \mathbf{X}} \mathbf{P}(\mathbf{Y}|x) P(x),$$

gdzie **Y** i **X** są wektorami zmiennych losowych

1. $P(dziura) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$
   Marginalizujemy Ból i Test
2. $\mathbf{P}(Dziura) = < 0.2, 0.8 >$
3. $P(dziura \wedge \neg ból) = 0.072 + 0.008 = 0.08$
   Marginalizujemy Test

Prawd. całkowite

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 2] Przed obliczeniami: czy to będzie duża wartość?

$$P(dziura|ból) = ?$$

Prawd. całkowite

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 2] Przed obliczeniami: czy to będzie duża wartość?

$P(dziura|ból) = ?$

Prawd. całkowite

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 2] Przed obliczeniami: czy to będzie duża wartość?

$$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

Ogólnie:

$$\mathbf{P}(\mathbf{Y}|\mathbf{Z}) = \sum_{x \in \mathbf{X}} \mathbf{P}(\mathbf{Y}, x|\mathbf{Z}) = \sum_{x \in \mathbf{X}} \frac{\mathbf{P}(\mathbf{Y}, x, \mathbf{Z})}{\mathbf{P}(\mathbf{Z})}$$

Prawd. całkowite

[zadanie 2] Przed obliczeniami: czy to będzie duża wartość?

Ogólnie:

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

[zadanie 3] Policz, rozpisz:

$$P(\neg dziura|ból) = ?$$

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$

[zadanie 3] Policz, rozpisz:

$P(\neg dziura|ból) = ?$

1. Ten sam mianownik! Liczenie mianownika często jest trudne.
2. $0.12 + 0.08 \neq 1.0$, więc normalizujemy je dzieląc przez $0.12 + 0.08 = 0.2$.

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

[zadanie 3] Policz, rozpisz:

$$P(\neg dziura|ból) = \frac{P(\neg dziura \wedge ból)}{P(ból)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064}$$

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

[zadanie 3] Policz, rozpisz:

$$P(\neg dziura|ból) = \frac{P(\neg dziura \wedge ból)}{P(ból)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064}$$

1. Ten sam mianownik! Liczenie mianownika często jest trudne.
2. $0.12 + 0.08 \neq 1.0$, więc normalizujemy je dzieląc przez $0.12 + 0.08 = 0.2$.

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

[zadanie 3] Policz, rozpisz:

$$P(\neg dziura|ból) = \frac{P(\neg dziura \wedge ból)}{P(ból)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064}$$

$$\mathbf{P}(Dziura|ból) =$$

1. Ten sam mianownik! Liczenie mianownika często jest trudne.
2. $0.12 + 0.08 \neq 1.0$, więc normalizujemy je dzieląc przez $0.12 + 0.08 = 0.2$.

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

[zadanie 3] Policz, rozpisz:

$$P(\neg dziura|ból) = \frac{P(\neg dziura \wedge ból)}{P(ból)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064}$$

$$\mathbf{P}(Dziura|ból) = \alpha \mathbf{P}(Dziura, ból) = \alpha \langle 0.108 + 0.012, 0.016 + 0.064 \rangle$$
$$\alpha \langle 0.12, 0.08 \rangle = \langle 0.6, 0.4 \rangle$$

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$P(dziura|ból) = \frac{P(dziura \wedge ból)}{P(ból)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$

[zadanie 3] Policz, rozpisz:

$P(\neg dziura|ból) = \frac{P(\neg dziura \wedge ból)}{P(ból)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064}$

$\mathbf{P}(Dziura|ból) = \alpha \mathbf{P}(Dziura, ból) = \alpha \langle 0.108 + 0.012, 0.016 + 0.064 \rangle$
$\alpha \langle 0.12, 0.08 \rangle = \langle 0.6, 0.4 \rangle$

1. Ten sam mianownik! Liczenie mianownika często jest trudne.
2. $0.12 + 0.08 \neq 1.0$, więc normalizujemy je dzieląc przez $0.12 + 0.08 = 0.2$.

Normalizacja

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

[zadanie 3] Policz, rozpisz:

$$\mathbf{P}(Dziura|ból) = \alpha\mathbf{P}(Dziura, ból) = \alpha\langle 0.108 + 0.012, 0.016 + 0.064\rangle$$
$$\alpha\langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle$$

Ogólnie:

$$\mathbf{P}(\mathbf{X}|\mathbf{e}) = \alpha\mathbf{P}(\mathbf{X}, e),$$

gdzie **X** jest wektorem zmiennych losowych

1. Ten sam mianownik! Liczenie mianownika często jest trudne.
2. $0.12 + 0.08 \neq 1.0$, więc normalizujemy je dzieląc przez $0.12 + 0.08 = 0.2$.

Podstawy: niezależność

Niezależność zdarzeń losowych $c \perp w$ (W – pogoda), jeśli

$$P(c) = P(c|w) \text{ lub } P(w) = P(w|c)$$

lub

$$P(c)P(w) = P(c \wedge w)$$

Podstawy: niezależność
Niezależność zdarzeń losowych $c \perp w$ (W – pogoda), jeśli
$P(c) = P(c|w)$ lub $P(w) = P(w|c)$
lub
$P(c)P(w) = P(c \wedge w)$

1. Graficznie: niezależność zdarzeń losowych, jeśli pole c w stosunku do całości = pole $c \wedge w$ w stosunku do w.

Podstawy: niezależność

Niezależność zdarzeń losowych $c \perp w$ (W – pogoda), jeśli

$$P(c) = P(c|w) \text{ lub } P(w) = P(w|c)$$

lub

$$P(c)P(w) = P(c \wedge w)$$

Niezależność zmiennych losowych $C \perp W$:

$$\mathbf{P}(C)\mathbf{P}(W) = \mathbf{P}(C \wedge W)$$

[zadanie 4] Rozpisać powyższe

Podstawy: niezależność

Niezależność zdarzeń losowych $c \perp w$ (W – pogoda), jeśli

$P(c) = P(c|w)$ lub $P(w) = P(w|c)$

lub

$P(c)P(w) = P(c \wedge w)$

Niezależność zmiennych losowych $C \perp W$:

$\mathbf{P}(C)\mathbf{P}(W) = \mathbf{P}(C \wedge W)$

[zadanie 4] Rozpisać powyższe

1. Graficznie: niezależność zdarzeń losowych, jeśli pole c w stosunku do całości = pole $c \wedge w$ w stosunku do w.

Podstawy: niezależność

Niezależność zdarzeń losowych $c \perp w$ (W – pogoda), jeśli

$$P(c) = P(c|w) \text{ lub } P(w) = P(w|c)$$

lub

$$P(c)P(w) = P(c \wedge w)$$

Niezależność zmiennych losowych $C \perp W$:

$$\mathbf{P}(C)\mathbf{P}(W) = \mathbf{P}(C \wedge W)$$

[zadanie 4] Rozpisać powyższe

$$\begin{aligned} Lewa &= \langle P(c), P(\neg c)\rangle \times \langle P(w), P(\neg w)\rangle \\ &= \langle P(c)P(w), P(c)P(\neg w), P(\neg c)P(w), P(\neg c)P(\neg w)\rangle \\ Prawa &= \langle P(c \wedge w), P(c \wedge \neg w), P(\neg c \wedge w), P(\neg c \wedge \neg w)\rangle \end{aligned}$$

wiedza o niezależności zmiennych jest zwykle wiedzą dziedzinową.

Podstawy: niezależność
Niezależność zdarzeń losowych $c \perp w$ (W – pogoda), jeśli
$P(c) = P(c|w)$ lub $P(w) = P(w|c)$
lub
$P(c)P(w) = P(c \wedge w)$
Niezależność zmiennych losowych $C \perp W$:
$\mathbf{P}(C)\mathbf{P}(W) = \mathbf{P}(C \wedge W)$
[zadanie 4] Rozpisać powyższe
$Lewa = \langle P(c), P(\neg c)\rangle \times \langle P(w), P(\neg w)\rangle = \langle P(c)P(w), P(c)P(\neg w), P(\neg c)P(w), P(\neg c)P(\neg w)\rangle$
$Prawa = \langle P(c \wedge w), P(c \wedge \neg w), P(\neg c \wedge w), P(\neg c \wedge \neg w)\rangle$
wiedza o niezależności zmiennych jest zwykle wiedzą dziedzinową.

1. Graficznie: niezależność zdarzeń losowych, jeśli pole c w stosunku do całości $=$ pole $c \wedge w$ w stosunku do w.

Niezależność

Czwarta zmienna: Pogoda (W)

$$P(W = słoneczna, b, t, d) = P(W = słoneczna|b, t, d)P(b, t, d)$$

Ale przecież (logika!):

$$P(W = słoneczna|b, t, d) = P(W = słoneczna)$$

więc

$$P(W = słoneczna, b, t, d) = P(W = słoneczna)P(b, t, d)$$

Niezależność

Czwarta zmienna: Pogoda (W)

$$P(W = słoneczna, b, t, d) = P(W = słoneczna|b, t, d)P(b, t, d)$$

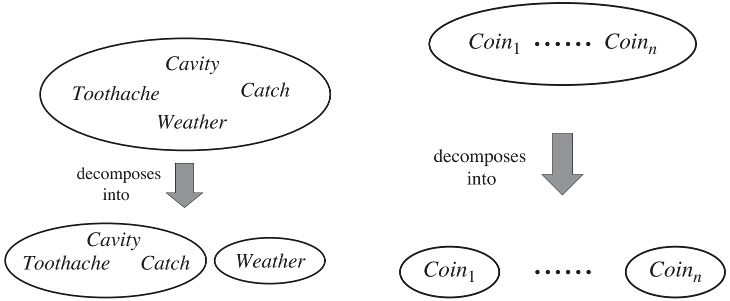
Ale przecież (logika!):

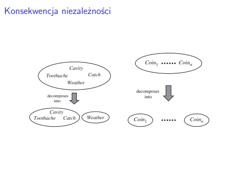
$$P(W = słoneczna|b, t, d) = P(W = słoneczna)$$

więc

$$P(W = słoneczna, b, t, d) = P(W = słoneczna)P(b, t, d)$$

Konsekwencja niezależności





1. Co nam to daje? Zamiast zapisywać rozkład prawd. łącznego za pomocą $2^4 = 16$ liczb, wystarczy nam $2^3 + 2 = 10$ (dekompozycja). Jest lepiej, ale w praktyce to nie wystarcza. Żeby było lepiej trzeba sięgnąć do reguły Bayesa i niezależności warunkowej.

Reguła Bayesa

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

- $P(x)$ często nieznane i rozpisuje się je jako praw. całkowite.

Reguła Bayesa

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

- $P(x)$ często nieznane i rozpisuje się je jako praw. całkowite.

1. Wyprowadzenie korzysta z def. prawd. warunkowego (reguły produkcji)
2. To proste równanie leży u podstawy większości nowoczesnych systemów sztucznej inteligencji opartych na wnioskowaniu probabilistycznym.

Reguła Bayesa

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

- $P(x)$ często nieznane i rozpisuje się je jako praw. całkowite.

Wersja ogólniejsza (zmienne losowe i dodatkowa wiedza $\mathbf{e}$):

$$\mathbf{P}(Y|X,\mathbf{e}) = \frac{\mathbf{P}(X|Y,\mathbf{e})\mathbf{P}(Y|\mathbf{e})}{\mathbf{P}(X|\mathbf{e})}$$

Reguła Bayesa

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

- $P(x)$ często nieznane i rozpisuje się je jako praw. całkowite.

Wersja ogólniejsza (zmienne losowe i dodatkowa wiedza $\mathbf{e}$):

$$\mathbf{P}(Y|X,\mathbf{e}) = \frac{\mathbf{P}(X|Y,\mathbf{e})\mathbf{P}(Y|\mathbf{e})}{\mathbf{P}(X|\mathbf{e})}$$

1. Wyprowadzenie korzysta z def. prawd. warunkowego (reguły produkcji)
2. To proste równanie leży u podstawy większości nowoczesnych systemów sztucznej inteligencji opartych na wnioskowaniu probabilistycznym.

Wnioskowanie

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

Zależność między zmiennymi X i Y. Możemy ją rozważać w dwóch kierunkach:

- **przyczynowym**: $P(efekt|przyczyna)$, np. $P(gorączka|grypa)$
- **diagnostycznym**: $P(przyczyna|efekt)$, np. $P(grypa|gorączka)$

[zadanie 5] Które prawd. łatwiej poznać?

Wnioskowanie

$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$

Zależność między zmiennymi X i Y. Możemy ją rozważać w dwóch kierunkach:

- **przyczynowym**: $P(efekt|przyczyna)$, np. $P(gorączka|grypa)$
- **diagnostycznym**: $P(przyczyna|efekt)$, np. $P(grypa|gorączka)$

[zadanie 5] Które prawd. łatwiej poznać?

1. Łatwiej poznać $P(gorączka|grypa)$, czyli jak często grypa powoduje gorączkę - wystarczy zebrać dane na temat ludzi z grypą. Łatwiej więc pozyskać dane o kierunku przyczynowym. Trudniej o kierunku diagnostycznym.

Przykład

Zapalenie opon mózgowych (M) i sztywność karku (S).

- $P(s|m) = 0.7$ (*kierunek przyczynowy*)
- $P(m) = 1/50000$ (*a priori, że osoba ma zapalenie opon*)
- $P(s) = 0.01$ (*a priori, że osoba ma sztywność kark*)

[zadanie 6] $P(m|s) = ?$

Przykład

Zapalenie opon mózgowych (M) i sztywność karku (S).

- $P(s|m) = 0.7$ (*kierunek przyczynowy*)
- $P(m) = 1/50000$ (*a priori, że osoba ma zapalenie opon*)
- $P(s) = 0.01$ (*a priori, że osoba ma sztywność kark*)

[zadanie 6] $P(m|s) = ?$

1. Jeśli doktor wie, że $P(m|s) = 0.0014$, to nie musi korzystać z reguły Bayes'a, więc po co mu to całe wnioskowanie? Otóż, $P(m|s)$ nie jest stałe! Jeśli tylko wybuchnie epidemia zapalenia opon mózgowych, wtedy $P(m)$ się zwiększy i $P(m|s)$ się zwiększy. $P(m|s)$ jest stałe, bo odzwierciedla zasady fizyki/biologii/etc.

Przykład

Zapalenie opon mózgowych (M) i sztywność karku (S).

- $P(s|m) = 0.7$ (*kierunek przyczynowy*)
- $P(m) = 1/50000$ (*a priori, że osoba ma zapalenie opon*)
- $P(s) = 0.01$ (*a priori, że osoba ma sztywność kark*)

[zadanie 6]

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = 0.0014$$

Przykład

Zapalenie opon mózgowych (M) i sztywność karku (S).

- $P(s|m) = 0.7$ (*kierunek przyczynowy*)
- $P(m) = 1/50000$ (*a priori, że osoba ma zapalenie opon*)
- $P(s) = 0.01$ (*a priori, że osoba ma sztywność kark*)

[zadanie 6]

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = 0.0014$$

1. Jeśli doktor wie, że $P(m|s) = 0.0014$, to nie musi korzystać z reguły Bayes'a, więc po co mu to całe wnioskowanie? Otóż, $P(m|s)$ nie jest stałe! Jeśli tylko wybuchnie epidemia zapalenia opon mózgowych, wtedy $P(m)$ się zwiększy i $P(m|s)$ się zwiększy. $P(m|s)$ jest stałe, bo odzwierciedla zasady fizyki/biologii/etc.

Przykład

Zapalenie opon mózgowych (M) i sztywność karku (S).

- $P(s|m) = 0.7$ (*kierunek przyczynowy*)
- $P(m) = 1/50000$ (*a priori, że osoba ma zapalenie opon*)
- $P(s) = 0.01$ (*a priori, że osoba ma sztywność kark*)

[zadanie 6]

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = 0.0014$$

Prawd. w kierunku przyczynowym jest solidniejsze!

Przykład

Zapalenie opon mózgowych (M) i sztywność karku (S).

- $P(s|m) = 0.7$ (*kierunek przyczynowy*)
- $P(m) = 1/50000$ (*a priori, że osoba ma zapalenie opon*)
- $P(s) = 0.01$ (*a priori, że osoba ma sztywność kark*)

[zadanie 6]

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = 0.0014$$

Prawd. w kierunku przyczynowym jest solidniejsze!

1. Jeśli doktor wie, że $P(m|s) = 0.0014$, to nie musi korzystać z reguły Bayes'a, więc po co mu to całe wnioskowanie? Otóż, $P(m|s)$ nie jest stałe! Jeśli tylko wybuchnie epidemia zapalenia opon mózgowych, wtedy $P(m)$ się zwiększy i $P(m|s)$ się zwiększy. $P(m|s)$ jest stałe, bo odzwierciedla zasady fizyki/biologii/etc.

Warunkowa niezależność zmiennych

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$\mathbf{P}(Dziura|ból \wedge test) = ?$$

Ale to się nie skaluje, gdy mamy wiele zmiennych („wielka tabelka”).

Warunkowa niezależność zmiennych

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$\mathbf{P}(Dziura|ból \wedge test) = ?$

Ale to się nie skaluje, gdy mamy wiele zmiennych („wielka tabelka”).

1. $\mathbf{P}(Dziura|ból \wedge test) = \alpha \langle 0.108, 0.016 \rangle \approx \langle 0.871, 0.129 \rangle$

Warunkowa niezależność zmiennych

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$$\mathbf{P}(Dziura|ból \wedge test) = ?$$

Ale to się nie skaluje, gdy mamy wiele zmiennych („wielka tabelka").

$$\begin{aligned}\mathbf{P}(Dziura|ból \wedge test) &= \alpha\mathbf{P}(ból \wedge test|Dziura)\mathbf{P}(Dziura)\\ &= \alpha\mathbf{P}(ból|Dziura)\mathbf{P}(test|Dziura)\mathbf{P}(Dziura)\end{aligned}$$

Warunkowa niezależność zmiennych

	ból		*¬ból*	
	test	*¬test*	*test*	*¬test*
dziura	0.108	0.012	0.072	0.008
¬dziura	0.016	0.064	0.144	0.576

$\mathbf{P}(Dziura|ból \wedge test) = ?$

Ale to się nie skaluje, gdy mamy wiele zmiennych („wielka tabelka").

$\mathbf{P}(Dziura|ból \wedge test) = \alpha\mathbf{P}(ból \wedge test|Dziura)\mathbf{P}(Dziura) = \alpha\mathbf{P}(ból|Dziura)\mathbf{P}(test|Dziura)\mathbf{P}(Dziura)$

1. $\mathbf{P}(Dziura|ból \wedge test) = \alpha \langle 0.108, 0.016 \rangle \approx \langle 0.871, 0.129 \rangle$

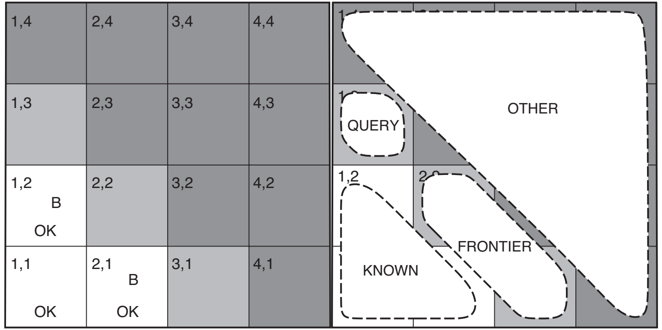
Warunkowa niezależność zmiennych

- C — rak, T_1 — jakiś test na obecność raka, T_2 — jakiś inny test na obecność raka
- C jest zmienną ukrytą. Ale jeśli znalibyśmy C, jakakolwiek wiedza o T_1 nie da nam żadnej dodatkowej wiedzy dot. T_2 , czyli T_1 **i** T_2 **są niezależne warunkowo pod warunkiem** C.
  - $\mathbf{P}(T_2|C, T_1) = \mathbf{P}(T_2|C)$
  - $\mathbf{P}(T_1, T_2|C) = \mathbf{P}(T_1|C)\mathbf{P}(T_2|C)$
- Notacja: $T_1 \perp T_2|C$

Warunkowa niezależność zmiennych
- C — rak, T_1 — jakiś test na obecność raka, T_2 — jakiś inny test na obecność raka
- C jest zmienną ukrytą. Ale jeśli znalibyśmy C, jakakolwiek wiedza o T_1 nie da nam żadnej dodatkowej wiedzy dot. T_2 , czyli T_1 **i** T_2 **są niezależne warunkowo pod warunkiem** C.
  - $\mathbf{P}(T_2|C, T_1) = \mathbf{P}(T_2|C)$
  - $\mathbf{P}(T_1, T_2|C) = \mathbf{P}(T_1|C)\mathbf{P}(T_2|C)$
- Notacja: $T_1 \perp T_2|C$

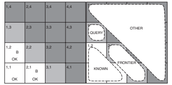
1. Widać to na diagramie. Jeśli znamy C, to on niezależnie wpływa na T1 i T2. W pewnym sensie obcina to co się dzieje w T1 od tego co się dzieje w T2
2. Jeśli zmienne losowe A i B są niezależne pod warunkiem, że X, możemy wnioskować tak: $\mathbf{P}(A, B, X) = \mathbf{P}(A, B|X)\mathbf{P}(X) = \mathbf{P}(A|X)\mathbf{P}(B|X)\mathbf{P}(X)$.

Wumpus



Jakie jest prawd, że w polu (1,3) jest jama jeśli wiatr poczuliśmy w polu (1,2) i (2,1)?
Zmienne losowe: $D_{i,j}$ (dziura na polu (i,j)) oraz $B_{i,j}$ (bryza na polu (i,j)).

Wumpus



Jakie jest prawd, że w polu (1,3) jest jama jeśli wiatr poczuliśmy w polu (1,2) i (2,1)?
Zmienne losowe: $D_{i,j}$ (dziura na polu (i,j)) oraz $B_{i,j}$ (bryza na polu (i,j)).