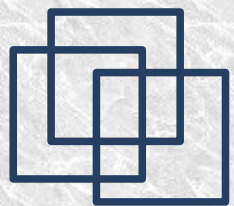


Inżynieria Wiedzy i Systemy Ekspertowe

Niepewność wiedzy

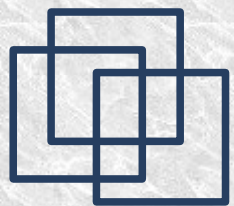
dr inż. Michał Bereta
Politechnika Krakowska

<http://torus.uck.pk.edu.pl/~beretam/>
beretam@torus.uck.pk.edu.pl



Niepewność wiedzy

- 1) Źródła niepewności wiedzy
- 2) Twierdzenie Bayesa
- 3) Współczynniki pewności (Certainty Factors)
- 4) Logika rozmyta (Fuzzy Logic)

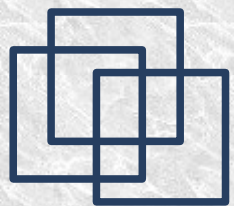


Niepewność wiedzy

Wiedza uzyskana od ekspertów jest często

- niepewna
- niekompletna
- niespójna
- **niedoskonała**

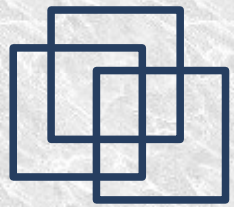
Systemy ekspertowe muszą być w stanie radzić sobie z wiedzą tego typu.



Niepewność wiedzy

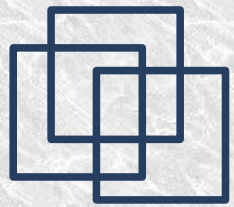
Eksperci są w stanie podejmować decyzje w oparciu o swoją często niedoskonałą wiedzę.

Systemy ekspertowe również muszą być w stanie radzić sobie z wiedzą tego typu.



Czym jest niepewność?

Niepewność można zdefiniować jako brak dokładnej wiedzy, tzn. takiej, która umożliwia podejmowanie w pełni uzasadnionych decyzji i w pełni wiarygodnych wniosków.

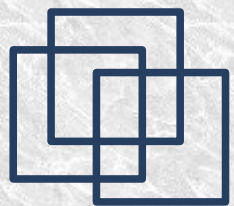


Czym jest niepewność?

Klasyczna logika zakłada, że dostępna wiedza jest doskonała oraz że zawsze jest spełnione **prawo wyłączonego środka**:

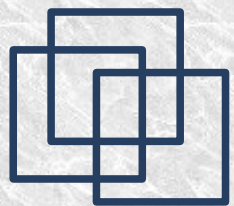
IF A jest TRUE THEN A nie jest FALSE

IF A jest FALSE THEN A nie jest TRUE



Źródła niepewności

- Nieprecyzyjność języka naturalnego
- Trudności ze sformułowaniem zależności w postaci reguł
- Łączenie wiedzy kilku ekspertów
- Niekompletność danych, na których system ma operować



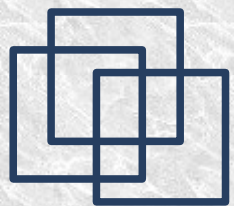
• Nieprecyzyjność języka naturalnego

- Always 99
- Very often 88
- Usually 85
- Sometimes 20
- Usualuy not 10
- Never 0

(Ray Simpson 1944)

- Always 100
- Very often 87
- Usually 79
- Sometimes 29
- Usualuy not 16
- Never 0

(Milton Hakel 1968)

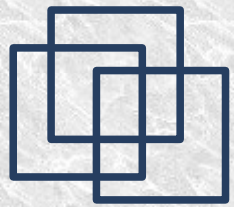


- Łączenie wiedzy kilku ekspertów

W rozwoju systemu ekspertowego PROSPECTOR brało udział **dziewięcioro** ekspertów.

Eksperci często mają odmienne zdanie na dany temat.

Ich wiedza często jest na różnym poziomie w zależności od zagadnienia w danej dziedzinie.



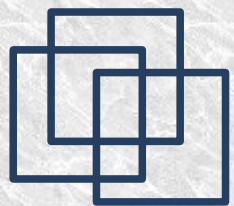
Teoria prawdopodobieństwa

- Prawdopodobieństwo zdarzenia A to proporcja przypadków, kiedy dane zdarzenie zachodzi.
- $p(A) = 0$ – całkowita niemożliwość
- $p(A) = 1$ – całkowita pewność

Jeśli wynikiem może być sukces lub porażka:

$$p(\text{sukces}) = \text{liczba_sukcesów} / \text{liczba_prób}$$

$$p(\text{porażka}) = \text{liczba_porażek} / \text{liczba_prób}$$



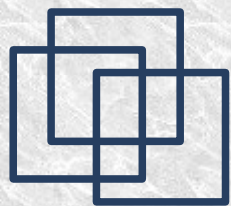
Prawdopodobieństwo warunkowe

Założmy, że zdarzenia A i B nie są wzajemnie wykluczające się.

Prawdopodobieństwo zdarzenia A może zależeć od tego, czy miało miejsce zdarzenie B.

$p(A|B)$ = liczba_wystąpień_A_oraz_B / liczba_wystąpień_B

$$p(A|B) = \frac{|A \text{ and } B|}{|B|}$$



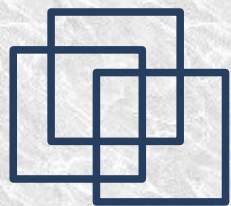
Prawdopodobieństwo warunkowe

Prawdopodobieństwo tego, że wystąpi A i B nazywamy prawdopodobieństwem łącznym tych zdarzeń.

$$p(A \cap B)$$

Zatem prawdopodobieństwo warunkowe definiujemy jako:

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$



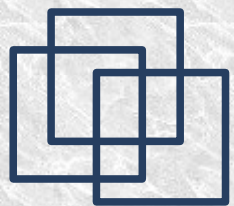
Prawdopodobieństwo warunkowe

Prawdopodobieństwo tego, że wystąpi A i B nazywamy prawdopodobieństwem łącznym tych zdarzeń.

$$p(A \cap B)$$

Zatem prawdopodobieństwo warunkowe definiujemy jako:

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$



Prawdopodobieństwo warunkowe

Analogicznie prawdopodobieństwo B pod warunkiem A:

$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$



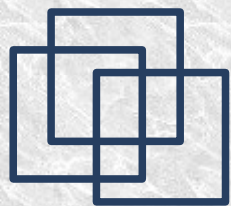
$$p(B \cap A) = p(B|A) \cdot p(A)$$

Prawdopodobieństwo łączne jest przemienne:

$$p(A \cap B) = p(B \cap A)$$



$$p(A \cap B) = p(B|A) \cdot p(A)$$



Twierdzenie Bayesa

Analogicznie prawdopodobieństwo B pod warunkiem A:

$$p(A \cap B) = p(B|A) \cdot p(A) \longrightarrow p(A|B) = \frac{p(A \cap B)}{p(B)}$$

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{p(B)}$$

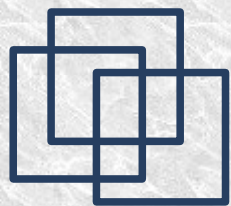
Wzór Bayesa

$p(A)$ – prawdopodobieństwo zdarzenia A

$p(B)$ – prawdopodobieństwo zdarzenia B

$p(A|B)$ – prawdopodobieństwo zdarzenia A pod warunkiem, że zaszło zdarzenie B

$p(B|A)$ – prawdopodobieństwo zdarzenia B pod warunkiem, że zaszło zdarzenie A



Prawdopodobieństwo warunkowe

W przypadku, gdy A zależy od zdarzeń $B_1, B_2, \dots, B_n$ wzajemnie się wykluczających:

$$B_i \cap B_j = \emptyset \text{ dla } i \neq j$$

wtedy

$$p(A \cap B_1) = p(A|B_1) \cdot p(B_1)$$

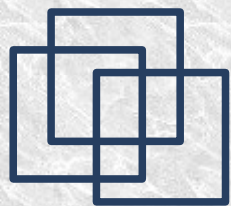
$$p(A \cap B_2) = p(A|B_2) \cdot p(B_2)$$

...

$$p(A \cap B_n) = p(A|B_n) \cdot p(B_n)$$



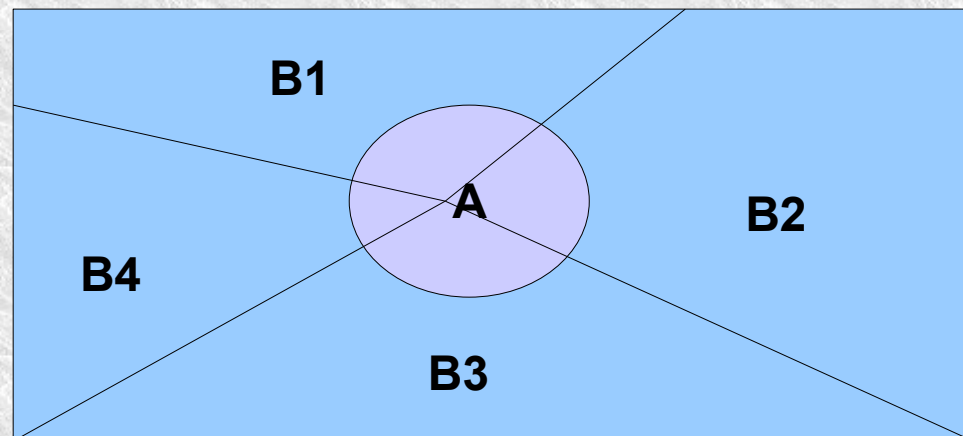
$$\sum_{i=1}^n p(A \cap B_i) = \sum_{i=1}^n p(A|B_i) \cdot p(B_i)$$



Prawdopodobieństwo całkowite

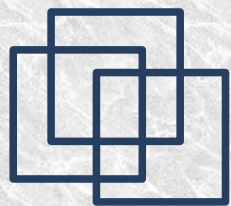
Jeśli B_i $i=1\dots n$ wyczerpują wszystkie możliwe zdarzenia:

$$B_1 \cup B_2 \cup \dots \cup B_n = S$$



S – zbiór wszystkich możliwych zdarzeń;
wtedy prawdopodobieństwo całkowite
 $p(A)$:

$$\sum_{i=1}^n p(A \cap B_i) = p(A)$$



Prawdopodobieństwo całkowite

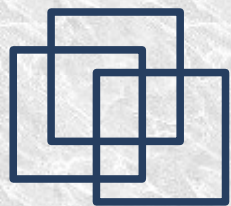
$$\sum_{i=1}^n p(A \cap B_i) = p(A)$$



$$p(A) = \sum_{i=1}^n p(A|B_i) \cdot p(B_i)$$

Jeśli A zależy jedynie od dwóch
wykluczających się zdarzeń, np. B oraz
NOT B:

$$p(A) = p(A|B) \cdot p(B) + p(A|\neg B) \cdot p(\neg B)$$



Reguła Bayesa

$$p(A) = p(A|B) \cdot p(B) + p(A|\neg B) \cdot p(\neg B)$$

Analogicznie dla B:

$$p(B) = p(B|A) \cdot p(A) + p(B|\neg A) \cdot p(\neg A)$$

Korzystając z reguły Bayesa otrzymamy:

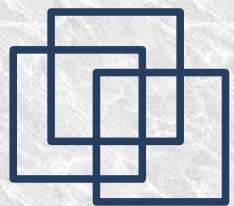
$$p(A|B) = \frac{p(B|A) \cdot p(A)}{p(B|A) \cdot p(A) + p(B|\neg A) \cdot p(\neg A)}$$



Reguła Bayesa

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{p(B|A) \cdot p(A) + p(B|\neg A) \cdot p(\neg A)}$$

Wzór ten jest podstawą do obsługi niepewności w systemach ekspertowych z wykorzystaniem metod probabilistycznych.



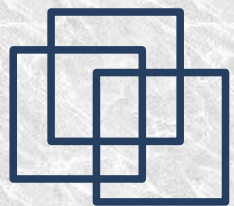
Systemy ekspertowe i wiedza niepewna

Założmy, że mamy regułę w postaci:

IF E is TRUE

THEN H is TRUE {z prawd. p }

Reguła ta mówi, że jeśli wystąpi **E** to **H** wystąpi z prawdopodobieństwem **p** .



Systemy ekspertowe i wiedza niepewna

IF E is TRUE

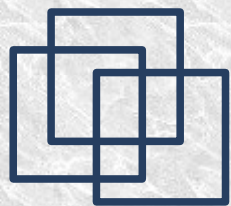
THEN H is TRUE {p}

Wiemy, że **E** wystąpiło.

Jak policzyć prawdopodobieństwo, że
wystąpiło **H** ?

Odpowiedź:

$$p(H|E) = \frac{p(E|H) \cdot p(H)}{p(E|H) \cdot p(H) + p(E|\neg H) \cdot p(\neg H)}$$



Systemy ekspertowe i wiedza niepewna

$$p(H|E) = \frac{p(E|H) \cdot p(H)}{p(E|H) \cdot p(H) + p(E|\neg H) \cdot p(\neg H)}$$

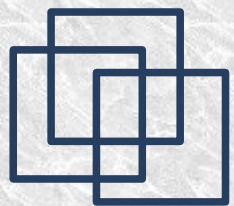
$p(H)$ – prawdopodobieństwo a priori wystąpienia H

$p(E|H)$ – prawdopodobieństwo, że jeśli wystąpi H,
to wystąpiło też E

$p(\sim H)$ – prawdopodobieństwo a priori nie
wystąpienia H

$p(E|\sim H)$ – prawdopodobieństwo, że jeśli nie
wystąpi H, to mimo wszystko wystąpiło E

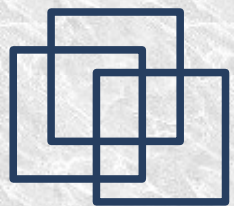
$p(H|E)$ – **prawdopodobieństwo a posteriori**
wystąpienia H, jeśli wystąpiło E



Systemy ekspertowe i wiedza niepewna

$$p(H|E) = \frac{p(E|H) \cdot p(H)}{p(E|H) \cdot p(H) + p(E|\neg H) \cdot p(\neg H)}$$

Prawdopodobieństwa **p(H)**, **p(E|H)**, **p(~H)** oraz **p(E|~H)** są określane przez eksperta bądź też szacowane statystycznie.

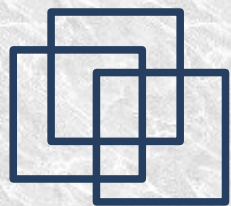


Systemy ekspertowe i wiedza niepewna

Co w przypadku wystąpienia wielu hipotez?

(Hipotezy oraz przesłanki (dowody, ang. Evidences) muszą być wzajemnie wykluczające się oraz wyczerpujące).

$$p(H_i|E) = \frac{p(E|H_i) \cdot p(H_i)}{\sum_{k=1}^m p(E|H_k) \cdot p(H_k)}$$

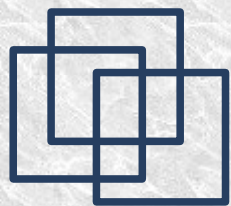


Systemy ekspertowe i wiedza niepewna

Co w przypadku wystąpienia wielu przesłanek E_i ?

(Hipotezy oraz przesłanki (dowody, ang. Evidences) muszą być wzajemnie wykluczające się oraz wyczerpujące).

$$p(H_i | E_1, E_2 \dots E_n) = \frac{p(E_1, E_2 \dots E_n | H_i) \cdot p(H_i)}{\sum_{k=1}^m p(E_1, E_2 \dots E_n | H_k) \cdot p(H_k)}$$



Systemy ekspertowe i wiedza niepewna

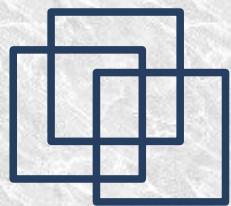
Co w przypadku wystąpienia wielu przesłanek E_i ?

(Hipotezy oraz przesłanki (dowody, ang. Evidences) muszą być wzajemnie wykluczające się oraz wyczerpujące).

$$p(H_i | E_1, E_2 \dots E_n) = \frac{p(E_1, E_2 \dots E_n | H_i) \cdot p(H_i)}{\sum_{k=1}^m p(E_1, E_2 \dots E_n | H_k) \cdot p(H_k)}$$

Powyższy wzór wymaga od nas pliczenia wszystkich prawdopodobieństw warunkowych dla wszystkich możliwych kombinacji przesłanek E_i oraz hipotez H_i .

$$p(E_1, E_2 \dots E_n | H_i)$$



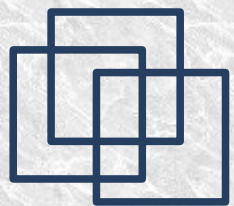
Systemy ekspertowe i wiedza niepewna

Powyższy wzór wymaga od nas pliczenia wszystkich prawdopodobieństw warunkowych dla wszystkich możliwych kombinacji przesłanek E_i oraz hipotez H_i .

$$p(E_1, E_2 \dots E_n | H_i)$$

W praktyce jest to często niemożliwe.

W przypadku szacowania na podstawie danych pojawiają się wątpliwości o jakość uzyskanych szacowań, jeśli np. przykładowych danych pokrywających dany przypadek jest mało.

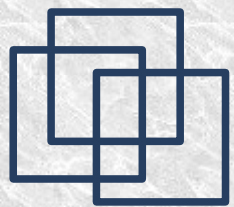


Systemy ekspertowe i wiedza niepewna

Często zakłada się zatem, że **E1**, **E2**, ..., **En** są wzajemnie niezależne.

Dwa zdarzenia **A** oraz **B** są określane jako niezależne wtedy i tylko wtedy, gdy:

$$p(A \cap B) = p(A) \cdot p(B)$$



Systemy ekspertowe i wiedza niepewna

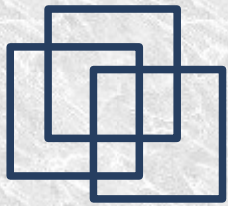
Zamiast:

$$p(H_i | E_1, E_2 \dots E_n) = \frac{p(E_1, E_2 \dots E_n | H_i) \cdot p(H_i)}{\sum_{k=1}^m p(E_1, E_2 \dots E_n | H_k) \cdot p(H_k)}$$

otrzymujemy zatem:

$$p(H_i | E_1, E_2 \dots E_n) = \frac{p(E_1 | H_i) \cdot p(E_2 | H_i) \cdot \dots \cdot p(E_n | H_i) \cdot p(H_i)}{\sum_{k=1}^m p(E_1, E_2 \dots E_n | H_k) \cdot p(H_k)}$$

Założenie takie nie jest zawsze uzasadnione i może przyczynić się do błędnego działania systemu.



Przykład

Warunki **E1**, **E2** oraz **E3** są wzajemnie warunkowo niezależne. Trzy hipotezy **H1**, **H2** oraz **H3** są wzajemnie wykluczające się oraz wyczerpujące.

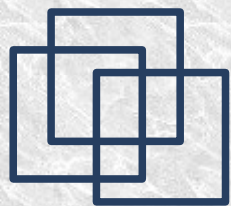
Ekspert dostarcza prawdopodobieństw a priori

$p(H1)$, $p(H2)$, $p(H3)$

oraz prawdopodobieństwa warunkowe

$p(E_i|H_j)$

Prawdopodobieństwo	Hipoteza		
	i=1	i=2	i=3
$p(H_i)$	0.4	0.35	0.25
$p(E1 H_i)$	0.3	0.8	0.5
$p(E2 H_i)$	0.9	0.0	0.7
$p(E3 H_i)$	0.6	0.7	0.9



Przykład

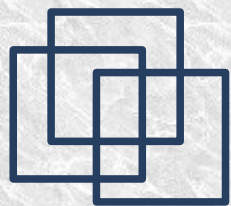
Zakładamy, że jako pierwszy zaobserwowane jest **E3** :

$$p(H_i|E_3) = \frac{p(E_3|H_i) \cdot p(H_i)}{\sum_{k=1}^3 p(E_3|H_k) \cdot p(H_k)}$$

$$p(H_1|E_3) = \frac{0.6 \cdot 0.4}{0.6 \cdot 0.4 + 0.7 \cdot 0.35 + 0.9 \cdot 0.25} = 0.34$$

$$p(H_2|E_3) = \frac{0.7 \cdot 0.35}{0.6 \cdot 0.4 + 0.7 \cdot 0.35 + 0.9 \cdot 0.25} = 0.34$$

$$p(H_3|E_3) = \frac{0.9 \cdot 0.25}{0.6 \cdot 0.4 + 0.7 \cdot 0.35 + 0.9 \cdot 0.25} = 0.32$$



Przykład

Następnie zaobserwowano **E1** :

Powinno być:

$$p(H_i|E_1, E_3) = \frac{p(E_1, E_3|H_i) \cdot p(H_i)}{\sum_{k=1}^3 p(E_1, E_3|H_k) \cdot p(H_k)}$$

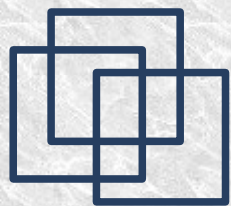
Ale E1, E2, E3 są warunkowo niezależne, zatem:

$$p(H_i|E_1, E_3) = \frac{p(E_1|H_i) \cdot p(E_3|H_i) \cdot p(H_i)}{\sum_{k=1}^3 p(E_1|H_k) \cdot p(E_3|H_k) \cdot p(H_k)}$$

$$p(H_1|E_1, E_3) = \frac{0.3 \cdot 0.6 \cdot 0.4}{0.3 \cdot 0.6 \cdot 0.4 + 0.8 \cdot 0.7 \cdot 0.35 + 0.5 \cdot 0.9 \cdot 0.25} = 0.19$$

$$p(H_2|E_1, E_3) = \frac{0.8 \cdot 0.7 \cdot 0.35}{0.3 \cdot 0.6 \cdot 0.4 + 0.8 \cdot 0.7 \cdot 0.35 + 0.5 \cdot 0.9 \cdot 0.25} = 0.52$$

$$p(H_3|E_1, E_3) = \frac{0.5 \cdot 0.9 \cdot 0.25}{0.3 \cdot 0.6 \cdot 0.4 + 0.8 \cdot 0.7 \cdot 0.35 + 0.5 \cdot 0.9 \cdot 0.25} = 0.29$$



Przykład

Następnie zaobserwowano **E2** :

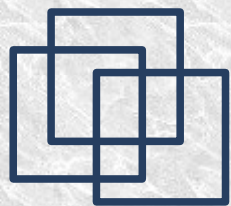
$$p(H_i | E_1, E_2, E_3) = \frac{p(E_1 | H_i) \cdot p(E_2 | H_i) \cdot p(E_3 | H_i) \cdot p(H_i)}{\sum_{k=1}^3 p(E_1 | H_k) \cdot p(E_2 | H_k) \cdot p(E_3 | H_k) \cdot p(H_k)}$$

$$p(H_1 | E_1, E_2, E_3) = \frac{0.3 \cdot 0.9 \cdot 0.6 \cdot 0.4}{0.3 \cdot 0.9 \cdot 0.6 \cdot 0.4 + 0.8 \cdot 0.0 \cdot 0.7 \cdot 0.35 + 0.5 \cdot 0.7 \cdot 0.9 \cdot 0.25} = 0.45$$

$$p(H_2 | E_1, E_2, E_3) = \frac{0.8 \cdot 0.0 \cdot 0.7 \cdot 0.35}{0.3 \cdot 0.9 \cdot 0.6 \cdot 0.4 + 0.8 \cdot 0.0 \cdot 0.7 \cdot 0.35 + 0.5 \cdot 0.7 \cdot 0.9 \cdot 0.25} = 0.0$$

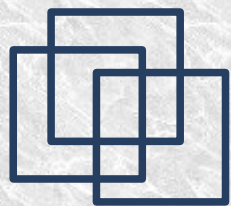
$$p(H_3 | E_1, E_2, E_3) = \frac{0.5 \cdot 0.7 \cdot 0.9 \cdot 0.25}{0.3 \cdot 0.9 \cdot 0.6 \cdot 0.4 + 0.8 \cdot 0.0 \cdot 0.7 \cdot 0.35 + 0.5 \cdot 0.7 \cdot 0.9 \cdot 0.25} = 0.55$$

Ostatecznie zatem najbardziej prawdopodobna jest hipoteza **H3**.



Reguła Bayesa

System ekspertowy **PROSPECTOR** był pierwszym systemem ekspertowym, który używał wnioskowania bayesowskiego.



Przykład 2

Prognoza pogody

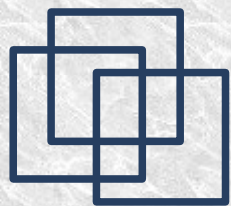
IF dziś jest deszcz

THEN jutro jest deszcz

IF dziś jest słońce

THEN jutro jest słońce

Zakładamy, że nie mając żadnych innych przesłanek, deszcz jest równie prawdopodobny jak słońce – zatem prawdopodobieństwa a priori wynoszą 0.5.



Przykład 2

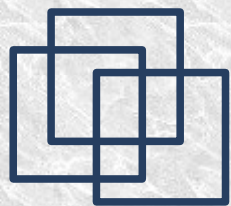
Prognoza pogody

IF dzis jest deszcz {LS=2.5, LN=0.6}

THEN jutro jest deszcz {a priori 0.5}

IF dzis jest slonce {LS=1.6, LN=0.4}

THEN jutro jest slonce {a priori 0.5}



Przykład 2

Prognoza pogody

IF dzis jest deszcz {LS=2.5, LN=0.6}

THEN jutro jest deszcz {a priori 0.5}

IF dzis jest slonce {LS=1.6, LN=0.4}

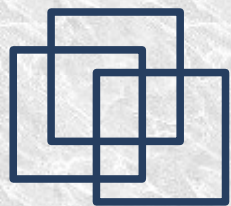
THEN jutro jest slonce {a priori 0.5}

LS (*likelihood of sufficiency*) – współczynnik wystarczalności

Miara przekonania eksperta, że **H** wystąpi, jeśli wystąpiło **E**.

$$LS = \frac{p(E|H)}{p(E|\neg H)}$$

np. dla reguły 1:
$$LS = \frac{p(\text{dzis jest deszcz} | \text{jutro jest deszcz})}{p(\text{dzis jest deszcz} | \text{jutro jest slonce})}$$



Przykład 2

Prognoza pogody

IF dzis jest deszcz {LS=2.5, LN=0.6}

THEN jutro jest deszcz {a priori 0.5}

IF dzis jest slonce {LS=1.6, LN=0.4}

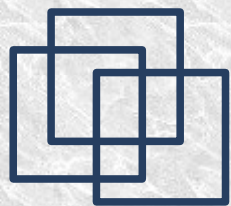
THEN jutro jest slonce {a priori 0.5}

LN (*likelihood of necessity*) – współczynnik konieczności

Miara przekonania eksperta, na ile **E** jest istotny dla wystąpienia **H**.

$$LN = \frac{p(\neg E|H)}{p(\neg E|\neg H)}$$

np. dla reguły 1: $LS = \frac{p(\text{dzis jest slonce}|\text{jutro jest deszcz})}{p(\text{dzis jest slonce}|\text{jutro jest slonce})}$



Przykład 2

Prognoza pogody

LS nie może być otrzymany z **LN** ani **LN** z **LS**.

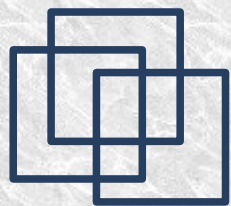
Ekspert musi dostarczyć obie wartości.

Nie jest konieczne szacowanie prawdopodobieństw warunkowych.

LN oraz **LS** mogą być szacowane bezpośrednio.

Wysokie wartości **LS** ($\mathbf{LS} \gg 1$) wskazuje na to, że reguła mocno wspiera hipotezę **H** jeśli zaobserwowano **E**.

Małe wartości **LN** ($0 < \mathbf{LN} < 1$) wskazują na to, że reguła mocno zaprzecza hipotezie **H** w przypadku braku **E**.



Przykład 2

Prognoza pogody

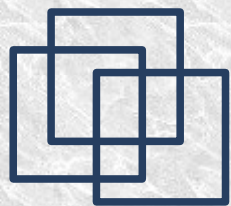
Reguła 1:

IF dziś jest deszcz {**LS=2.5**, **LN=0.6**}

THEN jutro jest deszcz {a priori **0.5**}

Reguła ta mówi, że jeśli dziś jest deszcz, to jest duże prawdopodobieństwo, że jutro również będzie deszcz (**LS=2.5**).

Ale nawet jeśli dziś nie ma deszczu, istnieje pewna niewielka szansa na to, by jutro był deszcz.
(**LN=0.6**)



Przykład 2

Prognoza pogody

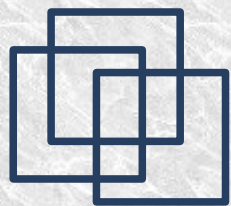
Reguła 2:

IF dzis jest slonce {**LS=1.6**, **LN=0.4**}

THEN jutro jest slonce {a priori **0.5**}

Jeśli dziś jest słońce, jest duże
prawdopodobieństwo, że jutro również będzie
słońce (**LS=1.6**).

LN=0.4 określa jakie są szanse na to, że jutro
będzie słońce jeśli dziś mamy deszcz (**E** nie jest
spełnione)



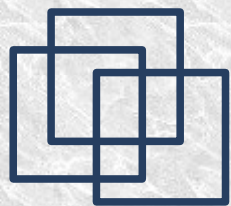
Przykład 2

Prognoza pogody

Jak przebiega rozumowanie?

W pierwszym kroku prawdopodobieństwa a priori $p(H)$ są przekształcane w ***szanse a priori*** (ang. ***prior odds***).

$$O(H) = \frac{p(H)}{1-p(H)}$$



Przykład 2

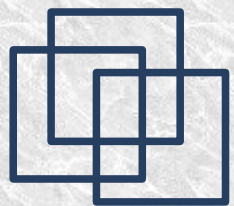
Prognoza pogody

Prawdopodobieństwa a priori są używane jedynie za pierwszym razem.

W celu uzyskania ***szans a posteriori*** (ang. ***posterior odds***) szanse a priori są uaktualniane za pomocą **LS** jeśli przesłanka reguły jest spełniona, lub za pomocą **LN** jeśli przesłanka nie jest spełniona.

$$O(H|E) = LS \cdot O(H)$$

$$O(H|\neg E) = LN \cdot O(H)$$



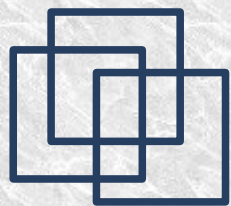
Przykład 2

Prognoza pogody

Prawdopodobieństwa a posteriori są uzyskiwane z szans a posteriori.

$$p(H|E) = \frac{O(H|E)}{1+O(H|E)}$$

$$p(H|\neg E) = \frac{O(H|\neg E)}{1+O(H|\neg E)}$$



Przykład 2

Prognoza pogody

IF **dzis jest deszcz** {LS=2.5, LN=0.6}

THEN **jutro jest deszcz** {a priori 0.5}

IF **dzis jest slonce** {LS=1.6, LN=0.4}

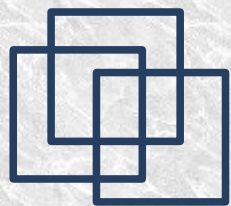
THEN **jutro jest slonce** {a priori 0.5}

Zakładamy, że dziś jest deszcz. Jaka pogoda będzie jutro?

$$O(\text{jutro jest deszcz}) = \frac{0.5}{1-0.5} = 1.0 \quad \text{Reguła 1}$$

$$O(\text{jutro jest deszcz} | \text{dzis jest deszcz}) = 2.5 \cdot 1.0 = 2.5$$

$$p(\text{jutro jest deszcz} | \text{dzis jest deszcz}) = \frac{2.5}{1+2.5} = 0.71$$



Przykład 2

Prognoza pogody

IF **dzis jest deszcz** {LS=2.5, LN=0.6}

THEN **jutro jest deszcz** {a priori 0.5}

IF **dzis jest slonce** {LS=1.6, LN=0.4}

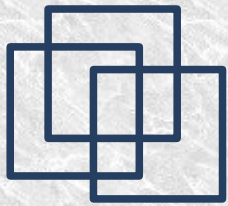
THEN **jutro jest slonce** {a priori 0.5}

Zakładamy, że dziś jest deszcz. Jaka pogoda będzie jutro?

$$O(\text{jutro jest slonce}) = \frac{0.5}{1-0.5} = 1.0 \quad \text{Reguła 2}$$

$$O(\text{jutro jest slonce} | \text{dzis jest deszcz}) = 0.4 \cdot 1.0 = 0.4$$

$$p(\text{jutro jest slonce} | \text{dzis jest deszcz}) = \frac{0.4}{1+0.4} = 0.29$$



Przykład 3

Prognoza pogody

Reguła 1

IF dziś deszcz {LS=2.5 LN=0.6}
THEN jutro deszcz {a priori 0.5}

Reguła 2

IF dziś słońce {LS=1.6 LN=0.4}
THEN jutro słońce {a priori 0.5}

Reguła 3

IF dziś deszcz
AND opady niskie {LS=10 LN=1}
THEN jutro słońce {a priori 0.5}

Reguła 4

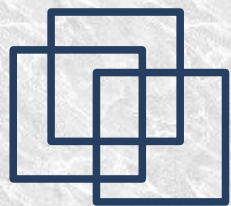
IF dziś deszcz
AND opady niskie
AND temp. niska {LS=1.5 LN=1}
THEN jutro słońce {a priori 0.5}

Reguła 5

IF dziś słońce
AND temp. wysoka {LS=2 LN=0.9}
THEN jutro deszcz {a priori 0.5}

Reguła 6

IF dziś słońce
AND temp. wysoka
AND niebo zachmurzone {LS=5 LN=1}
THEN jutro deszcz {a priori 0.5}



Przykład 3

Prognoza pogody

Jaka dziś pogoda?
Deszcz.

Reguła 1

IF dziś deszcz {LS=2.5 LN=0.6}
THEN jutro deszcz {a priori 0.5}

Reguła 1:

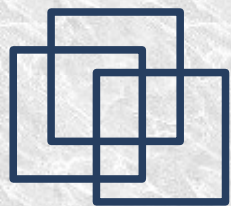
$$O(\text{jutro jest deszcz}) = \frac{0.5}{1-0.5} = 1.0$$

$$O(\text{jutro jest deszcz} | \text{dzis jest deszcz}) = 2.5 \cdot 1.0 = 2.5$$

$$p(\text{jutro jest deszcz} | \text{dzis jest deszcz}) = \frac{2.5}{1+2.5} = 0.71$$

Prognoza:

jutro deszcz {0.71}



Przykład 3

Prognoza pogody

Jaka dziś pogoda?

Deszcz.

Reguła 2:

Reguła 2

IF dziś słońce {LS=1.6 LN=0.4}

THEN jutro deszcz {a priori 0.5}

$$O(\text{jutro jest słońce}) = \frac{0.5}{1-0.5} = 1.0$$

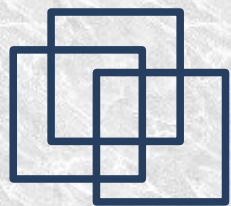
$$O(\text{jutro jest słońce} | \text{dzis jest deszcz}) = 0.4 \cdot 1.0 = 0.4$$

$$p(\text{jutro jest słońce} | \text{dzis jest deszcz}) = \frac{0.4}{1+0.4} = 0.29$$

Prognoza:

jutro deszcz {0.71}

jutro słońce {0.29}



Przykład 3

Prognoza pogody

Jakie opady?
Niskie.

Reguła 3:

Reguła 3

IF dziś deszcz
AND opady niskie {LS=10 LN=1}
THEN jutro słońce {a priori 0.5}

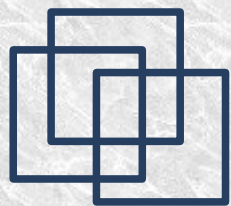
$$O(jutro\ slonce) = \frac{0.29}{1-0.29} = 0.41$$

$$O(jutro\ slonce | dzis\ deszcz \cap opady\ niskie) = 10 \cdot 0.41 = 4.1$$

$$p(jutro\ slonce | dzis\ deszcz \cap opady\ niskie) = \frac{4.1}{1+4.1} = 0.8$$

Prognoza:

jutro deszcz {0.71}
jutro słońce {0.8}



Przykład 3

Prognoza pogody

**Jaka dziś temperatura?
Niska.**

Reguła 4:

Reguła 4

**IF dziś deszcz
AND opady niskie
AND temp. niska {LS=1.5 LN=1}
THEN jutro słońce {a priori 0.5}**

$$O(jutro\ slonce) = \frac{0.8}{1-0.8} = 4$$

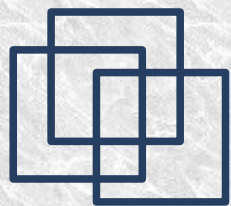
$$O(jutro\ slonce | dzis\ deszcz \cap opady\ niskie \cap temp.\ niska) = 1.5 \cdot 4 = 6$$

$$p(jutro\ slonce | dzis\ deszcz \cap opady\ niskie \cap temp.\ niska) = \frac{6}{1+6} = 0.86$$

Prognoza:

jutro deszcz {0.71}

jutro słońce {0.86}



Przykład 3

Prognoza pogody

Jaka dziś temperatura?
Niska.

Reguła 5:

Reguła 5

IF dziś słońce
AND temp. wysoka {LS=2 LN=0.9}
THEN jutro deszcz {a priori 0.5}

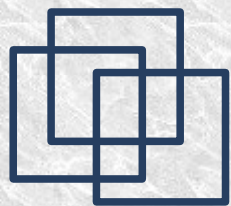
$$O(jutro\ deszcz) = \frac{0.71}{1-0.71} = 2.45$$

$$O(jutro\ deszcz | dzis\ deszcz \cap temp.\ niska) = 0.9 \cdot 2.45 = 2.21$$

$$p(jutro\ deszcz | dzis\ deszcz \cap temp.\ niska) = \frac{2.21}{1+2.21} = 0.69$$

Prognoza:

jutro deszcz {0.69}
jutro słońce {0.86}



Przykład 3

Prognoza pogody

Jaka dziś temperatura?
Niska.

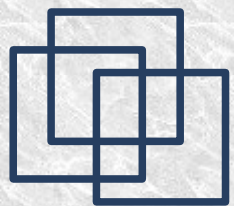
IF dziś słońce
AND temp. wysoka
AND niebo zachmurzone {LS=5 LN=1}
THEN jutro deszcz {a priori 0.5}

Reguła 6:

Reguła 6 nie zmieni naszej prognozy - dlaczego?

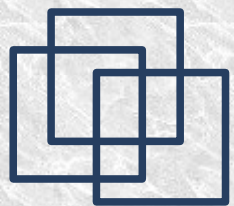
Ostatecznie prognoza:

jutro deszcz {0.69}
jutro słońce {0.86}



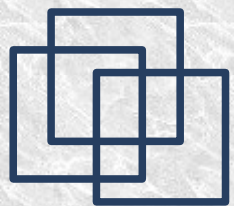
Zalety reguły Bayesa

- Dobrze rozwinięta teoria matematyczna.
- Decyzje oparte na metodzie Bayesa są optymalne (pod warunkiem, że znamy odpowiednie rozkłady prawdopodobieństw).
- Wszystkie inne metody podejmowania decyzji mogą być co najwyżej zbieżne do metody Bayesa.



Wady reguły Bayesa

- Czasami trzeba czynić nieuprawnione założenia upraszczające problem (np. założenie o niezależności bądź warunkowej niezależności między zmiennymi).
- Badania psychologiczne dowodzą, iż ludziom ciężko jest szacować prawdopodobieństwa – często prawd. a priori są niespójne z prawd. warunkowymi.



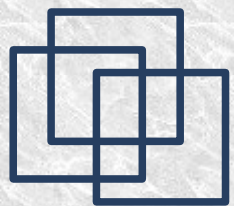
Wady reguły Bayesa

Samochód nie chce zapalić.

Jaka jest przyczyna?

IF oznaką są dziwne odgłosy

THEN rozrusznik jest zepsuty { $p=0.7$ }



Wady reguły Bayesa

$p(\text{rozrusznik nie jest zły} \mid \text{są dziwne odgłosy})$

=

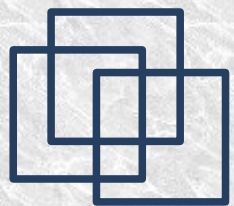
$p(\text{rozrusznik jest dobry} \mid \text{są dziwne odgłosy})$

= $1 - 0.7 = 0.3$

IF oznaką są dziwne odgłosy

THEN rozrusznik jest dobry {p=0.3}

Jest to reguła w pewnym sensie ukryta, której istnienie może wprowadzać niejasności.



Wady reguły Bayesa

Spójrzmy od innej strony na sytuację:

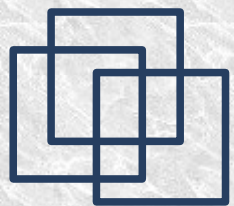
IF rozrusznik jest zepsuty

THEN są dziwne odgłosy { $p=0.85$ }

IF rozrusznik jest zepsuty

THEN nie ma dziwnych odgłosów { $p=0.15$ }

Aby użyć reguły Bayesa potrzebujemy jeszcze prawd. a priori, że rozrusznik jest zły jeśli samochód nie chce zapalić. Załóżmy, że ekspert szacuje je na 5%.

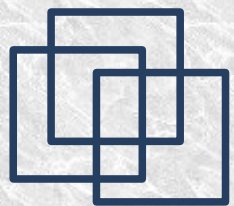


Wady reguły Bayesa

Z reguły Bayesa otrzymamy:

$$p(\text{rozrusznik jest zły} | \text{śa dziwne odgłosy}) = \frac{0.85 \cdot 0.05}{0.85 \cdot 0.05 + 0.15 \cdot 0.95} = 0.23$$

Jest to wartość dużo mniejsza niż wcześniej
oszacowane przez eksperta prawd. = 0.7.



Wady reguły Bayesa

Skąd ta rozbieżność? Różne założenia eksperta podczas szacowania prawd. a priori oraz prawd. warunkowych.

$$p(H|E) = \frac{p(E|H) \cdot p(H)}{p(E|H) \cdot p(H) + p(E|\neg H) \cdot p(\neg H)}$$

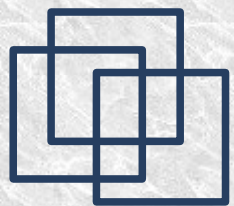
$$p(H) = \frac{p(H|E) \cdot p(E|\neg H)}{p(H|E) \cdot p(E|\neg H) + p(E|H) [1 - p(H|E)]}$$

$p(H)$ = p(rozrusznik jest zepsuty)

$p(H|E)$ = p(rozrusznik jest zepsuty | są dziwne odgłosy)

$p(E|H)$ = p(są dziwne odgłosy | rozrusznik jest zepsuty)

$p(E|\sim H)$ = p(są dziwne odgłosy | rozrusznik jest dobry)



Wady reguły Bayesa

Skąd ta rozbieżność? Różne założenia eksperta podczas szacowania prawd. a priori oraz prawd. warunkowych.

$$p(H) = \frac{0.7 \cdot 0.15}{0.7 \cdot 0.15 + 0.85 \cdot (1 - 0.7)} = 0.29$$

Jest to dużo większa wartość niż 0.05 podana wcześniej przez eksperta.