

## Wnioskowanie w warunkach niepewności : Teoria Dempstera-Shafera

### 1. Wstęp

Wśród wielu dziedzin jakimi zajmuje się informatyka ważną pozycję zajmuje problematyka sztucznej inteligencji. Obecne systemy przypominają działaniem człowieka. Informatycy zajmujący się sztuczną inteligencją doszli do wniosku, że systemy inteligentne muszą być przygotowane do przetwarzania informacji, czyli powinny gromadzić informacje o otaczającym je świecie i wyciągać z zebranej wiedzy wnioski. Właśnie z tym związany jest największy problem sztucznej inteligencji a mianowicie uzyskanie odpowiedzi na pytanie, w jaki sposób systemy inteligentne powinny prowadzić wnioskowanie.

Możemy wyróżnić wiele wzorców wnioskowania stosowanych w systemach inteligentnych. Są to między innymi:

- wnioskowanie probabilistyczne
- wnioskowanie rozmyte
- wnioskowanie indukcyjne
- wnioskowanie dedukcyjne
- wnioskowanie niemonotoniczne

Podział ten jest zrobiony ze względu na teorie na których opierają się poszczególne wnioskowania. Ponieważ rzeczywistości nigdy nie da się dokładnie opisać to najczęściej informacje dotyczące otaczającego nas świata są nieprecyzyjne, niepełne lub niepewne. Od rodzaju informacji, jakimi dysponujemy zależy sposób wnioskowania. Dlatego też przyjęliśmy następującą klasyfikację metod wnioskowania:

- metody przetwarzania informacji niepewnej
- metody przetwarzania informacji niepełnej

Oczywiście podział ten jest umowny i roboczy a metody obu grup można łączyć.

Wśród metod z pierwszej grupy przeważa podejście numeryczne. Możemy tu wymienić m.in. schematy wielowartościowe, rozmyte, probabilistyczne, teorię Dempstera-Shafera.

W niniejszej pracy właśnie omówimy teorię **Dempstera-Shafera** nazywaną również teorią funkcji przekonania lub matematyczną teorią ewidencji. W obecnej postaci opracowali ją Dempster i Shafer. Jest ona jednym z dobrze teoretycznie poznanych modeli niepewności zajmujących się przyjmowaniem wielu wartości danego atrybutu jednocześnie. Pokazuje jedną z dróg stosowania matematycznego prawdopodobieństwa przy ocenie subiektywnej. Jest uogólnieniem bayesowskiej teorii subiektywnego prawdopodobieństwa. Gdy stosujemy bayesowską teorię do oceny ilościowej zapytania, to musimy przypisać prawdopodobieństwa możliwym odpowiedziom na to pytanie. Teoria DS jest bardziej elastyczna. Stopnie przekonania można przypisać zarówno indywidualnym odpowiedziom, jak i grupom odpowiedzi, przy czym suma stopni przekonania przypisanych odpowiedziom indywidualnym może być mniejsza, niż stopień przekonania grupy. Przypisanie stopni odpowiedziom i ich grupom nazywa się rozkładem przekonania lub „funkcja przekonania”. Teoria ta umożliwia wyprowadzenie stopni przekonania dla danego pytania z prawdopodobieństw dla pytania

pokrewnego, przy czym stopnie te nie muszą mieć matematycznych własności prawdopodobieństwa. Oferuje również szereg ważnych korzyści metodologicznych takich jak: zdolność do reprezentowania ignorancji w sposób prosty i bezpośredni, spójność z klasyczną teorią prawdopodobieństwa, zgodność z logiką boolowską oraz złożoność obliczeniową. Teorię DS stosuje się w wypadkach :

- niepełnej wiedzy
- aktualizacji przekonań
- składania ewidencji

Teoria ta może służyć jako narzędzie do reprezentacji niekompletnej wiedzy statystycznej ponieważ może objąć statystykę zbiorów losowych.

Praktyczne zastosowania teorii DS obejmują m.in.: integrację informacji pochodzących z różnych źródeł przy identyfikacji obiektu, modelowania wyszukiwania informacji w bazie danych, rozpoznawanie planu działania, zagadnienia diagnostyki technicznej w warunkach zawodności elementów pomiarowych oraz ma zastosowanie w medycynie. Poza tym wykazano również, że problem obliczania niezawodności sieci jest szczególnie przypadkiem obliczeń w teorii DS.

W systemach ekspertowych dokonał się ogromny postęp jeżeli chodzi o wykorzystanie funkcji przekonania DS do reprezentacji niepewności wiedzy. Opracowano efektywne metody wnioskowania w oparciu o reprezentację wieloatrybutowych rozkładów przekonań w postaci tzw. drzew Markowa. Przez drzewo Markowa rozumiemy drzewo nieskierowane, którego węzły reprezentują zbiory atrybutów spełniających pewne warunki sąsiedztwa. Węzłom drzewa Markowa przypisuje się tzw. *składowe funkcji przekonania*, przy czym łączny rozkład przekonań reprezentowany jako złożenie dempsterowskie (złożenie tak zwaną regułą Dempstera) składowych funkcji przekonania. W drzewie Markowa możemy wnioskować w przód, wstecz, poszukiwać uzasadnienia itp.

Niestety, dla zastosowania w systemach ekspertowych istnieje poważna luka w badaniach nad tą teorią: nie są dotychczas znane metody pozyskiwania wiedzy dla tego modelu z danych ani metody jej weryfikacji wobec danych. To znaczy, że ani struktury drzewa Markowa ani składowych funkcji przekonania nie można pozyskać z danych empirycznych. Pomimo iż często mówi się, że rozkład przekonań DS jest uogólnieniem rozkładu prawdopodobieństwa, nie udało się zaadaptować znanych w/w metod konstrukcji probabilistycznych sieci bayesowskich z danych do konstrukcji drzew Markowa czy też innych reprezentacji graficznych DS.

Jednym z najbardziej istotnych zagadnień w przetwarzaniu niepewności jest rozróżnienie pomiędzy niepewnością a niewiedzą. Właśnie rozróżnienie tych dwóch rzeczy jest celem teorii Dempstera-Shafera.

## 2. Definicje

Na początku można zauważyć iż teoria Dempstera-Shafera odrzuca jeden z aksjomatów wnioskowania probabilistycznego:

$$P(\neg A) = 1 - P(A)$$

Przykładem odrzucenia powyższego aksjomatu jest sytuacja gdy nic nie wiemy odnośnie zdarzenia  $A$  i  $\neg A$ . Poprawniej jest wtedy założyć, że oba prawdopodobieństwa są równe zero. Konsekwencją odrzucenia

powyższego aksjomatu przez teorię Dempstera-Shafera jest skojarzenie z każdą hipotezą dwóch a nie jednej wartości:  $Bel(A)$  i  $Pl(A)$ . Określają one dolną i górną granicę stopnia przekonania w prawdziwość naszej hipotezy  $A$ .

W teorii Bayesa jest wymagana całkowita specyfikacja modelu probabilistycznego. W przypadku braku pełnej specyfikacji, zwolennicy tej teorii stosują metody przybliżone, uzupełniające model probabilistyczny. Teoria Dempstera-Shafera jest metodą alternatywną, uznającą model częściowo wyspecyfikowany, w której nie uzupełnia się brakującej specyfikacji. Wyznaczane są tam prawdopodobieństwa z jakimi dane hipotezy można udowodnić na podstawie posiadanej informacji, a nie tak jak w modelu Bayesa prawdziwości tych hipotez. Obecnie znanych jest wiele metod wykraczających poza model prawdopodobieństwa. Większość z nich opiera się na tzw. miarach monotonicznych czyli takich, które przekształcają  $F$  w odcinek  $[0,1]$  spełniających następujące aksjomaty:

$$g(f) = 0, g(t) = 1, \text{ jeśli } (a \Rightarrow b) = t, \text{ to } g(a) \leq g(b)$$

Funkcją spełniającą te aksjomaty jest wprowadzona przez Shafera funkcja przekonania (belief function).

Teoria Dempstera-Shafera powstała na podstawie pracy Dempstera, a Shafer wprowadza zbiór  $A \subseteq \Theta$  elementów ogniskowych. Są to zdania o których posiadamy jakieś informacje. Elementy z  $A$  nie muszą być zdaniami elementarnymi oraz wzajemnie wyłączającymi się. Dostępne informacje o elementach  $A$  są zapisywane w postaci rozkładu bazowego prawdopodobieństwa, który reprezentuje częściowe przekonania.

### Def. 1

Niech  $\Theta$  będzie ramą rozróżniającą,  $m$  funkcją  $m: 2^\Theta \rightarrow [0,1]$  taką, że dla każdego zbioru  $A \subseteq \Theta$ .

$$m(A) \geq 0$$

$$m(\emptyset) = 0,$$

$$\sum_{X \in \Theta} m(A) = 1$$

Taka funkcja  $m$  nazywa się podstawowym przyporządkowaniem prawdopodobieństwa dla  $\Theta$ .

### Def. 2

Niech  $\Theta$  będzie ramą rozróżniającą,  $m$  podstawowym przyporządkowaniem prawdopodobieństwa dla  $\Theta$ ; Zbiór  $A$  nazywamy elementem ogniskowym w  $m$  jeżeli

$$m(A) > 0.$$

Rdzeń  $m$  (core of  $m$ ) jest zbiorem wszystkich elementów ogniskowych w  $m$ .

Definicja podstawowego przyporządkowania prawdopodobieństwa jest bardzo podobna do klasycznej definicji prawdopodobieństwa, można więc powiedzieć, że jest rozszerzeniem tej definicji dla zbiorów.

**Def. 3**

Dla zbioru hipotez możemy określić łączną wiarę w jego prawdziwość nazywaną funkcją przekonania (ang. belief function)

Niech  $\Theta$  będzie ramą rozróżniającą,  $m$  podstawowym przyporządkowaniem prawdopodobieństwa dla  $\Theta$ , wtedy funkcją przekonania związaną z tym  $m$  jest funkcja

$Bel: 2^\Theta \rightarrow [0,1]$  taką, że

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

dla każdego  $A \subseteq \Theta$ .

Funkcja przekonania mierzy wiarygodność poszlak na rzecz  $A$ .

**Def. 4**

Na ile poszlaki przemawiające przeciw  $A$  są silne mówi funkcja wiarygodności (ang. plausibility function)

Niech  $\Theta$  będzie ramą rozróżniającą,  $m$  podstawowym przyporządkowaniem prawdopodobieństwa dla  $\Theta$ , wtedy funkcją wiarygodności odpowiadającą  $m$  nazywamy funkcją  $Pl: 2^\Theta \rightarrow [0,1]$  taką, że

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

dla każdego  $A \subseteq \Theta$ .

**Zależności między funkcją przekonania a funkcją wiarygodności:**

$$Pl(A) = 1 - Bel(\neg A),$$

$$Bel(A) + Bel(\neg A) \leq 1,$$

$$Pl(A) + Pl(\neg A) \geq 1,$$

$$Bel(A) \leq Pl(A),$$

$$Bel(A \vee B) \geq Bel(A) + Bel(B) - Bel(A \wedge B),$$

$$Pl(A \wedge B) \geq Pl(A) + Pl(B) - Pl(A \vee B).$$

Pewność (niepewność) danego zadania  $A \subseteq \Theta$  może więc być reprezentowana przez odcinek

$$[Bel(A), Pl(A)]$$

Przedział ten nazywamy przedziałem przekonania dla  $A$ . Początkowo przedział przekonania jest równy  $[0, 1]$  co oznacza brak danych rozpatrywanej hipotezy, a w miarę napływu informacji przedział zawęża się.

**Def. 5**

Funkcja łączenia dwóch niepewności dla tego samego zbioru nazywana jest *regulą kombinacji Dempstera*.

Niech  $\Theta$  będzie ramą rozróżniającą,  $m_1$  i  $m_2$  podstawowe przyporządkowania prawdopodobieństwa na  $\Theta$ , wtedy  $m_1 \oplus m_2$  jest funkcją  $m_1 \oplus m_2: 2^\Theta \rightarrow [0,1]$  taką, że

$$m_1 \oplus m_2(\emptyset) = 0$$

$$m_1 \oplus m_2(C) = \frac{\sum_{A \cap B = C} m_1(A) * m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A) * m_2(B)}$$

dla wszystkich  $A \subseteq \Theta$  i  $A \neq \emptyset$ . Dla tak zdefiniowanej funkcji odpowiadającą funkcją przekonania jest  $Bel_1 \oplus Bel_2: 2^\Theta \rightarrow [0,1]$ , taka, że

$$Bel_1 \oplus Bel_2(A) = \sum_{B \subseteq A} m_1 \oplus m_2(B)$$

Należy zauważyć, że tak sformułowana reguła łącząca wiadomości zawarte w dwu podstawowych przyporządkowaniach prawdopodobieństwa nie pasuje do każdej sytuacji. Jeżeli przyjmujemy model świata „otwartego” tzn. dopuszczamy istnienie innych hipotez oprócz  $\Theta$ , normalizacja w tym wzorze jest dyskusyjna.

**Def. 6**

Niech  $B$  będzie zdarzeniem  $B \subseteq \Theta$ , a  $Bel$  funkcją przekonania nad  $\Theta$ . Wtedy przez **warunkową funkcję przekonania**  $Bel(\cdot|B)$  względem zdarzenia  $B$  rozumie się funkcję przekonania

$$Bel(\cdot|B) = Bel \oplus Bel_B$$

gdzie  $Bel_B$  jest prostą deterministyczną funkcją przekonania o słuszności  $B$ .

**3. Przykłady**

Poniżej przedstawiono kilka przykładów funkcji dotyczących teorii Dempstera-Shafera:

**1.**

Prosta funkcja przekonania wygląda następująco:

$$m(A) = s, \quad m(\Theta) = 1 - s, \quad m(B) = 0, \quad \text{gdy } B \neq A, B \neq \Theta,$$

czyli

$$Bel(B) = \begin{cases} 0, & A \notin B \\ s, & A \subseteq B \neq \Theta \\ 1, & B = \Theta \end{cases}$$

2.

Warunkowa funkcja przekonania, założmy, że istniał rozkład  $m_1$  dla  $\Theta$  określający funkcję przekonania  $Bel^1$ . Aktualnie dowiadujemy się jedynie, że jest prawdziwy zbiór  $E \in \Theta$ . Mamy więc nowy rozkład bazowy przydzielający wartość 1 zbiorowi  $E$ . Funkcja przekonania będąca kombinacją tych dwóch funkcji jest nazwaną *warunkową* dla danej funkcji  $Bel^1$ :

$$Bel(A|E) = \frac{Bel^1(A \cup E') - Bel^1(E')}{1 - Bel^1(E')}.$$

3.

Rozpatrzmy przykład:

Niech  $a, b, c$  będą trzema wzajemnie wyłączającymi się alternatywami oraz

$$m_1(a) = 0; \quad m_1(b) = 0,1; \quad m_1(c) = 0,9;$$

$$m_2(a) = 0,9; \quad m_2(b) = 0,1; \quad m_2(c) = 0;$$

Rozkłady te reprezentują dwa skrajnie niezgodne źródła informacji. Stosując regułę łączenia otrzymamy:

$$m(a) = 0; \quad m(b) = 1; \quad m(c) = 0;$$

Otrzymany wynik jest bardzo ostry, rozkład  $m$  wskazuje na całkowitą pewność alternatywy  $b$ .

Z informacji wejściowych dowiadujemy się, że  $a$  albo  $c$  są niemożliwe i dlatego też  $b$  stało się jedyną możliwą alternatywą mimo, że wszystkie źródła wykluczało tę alternatywę. Oznacza to, że ograniczono uwagę do zbyt małego zbioru. Należałoby szukać nowej możliwości i zbiór  $\Theta$  rozszerzyć. Oznaczmy tę możliwość jako  $e$  i stosując rozszerzenie  $A = \{a, b, c, e\}$  otrzymamy:

$$\begin{array}{ll} m_1(a,e) = 0, & m_2(a,e) = 0,9, \\ m_1(b,e) = 0,1, & m_2(b,e) = 0,1, \\ m_1(c,e) = 0,9, & m_2(c,e) = 0. \end{array}$$

A więc mamy:

$$\begin{array}{l} m(a) = 0,01, \\ m(b) = m(c) = 0, \\ m(e) = 0,99 \end{array}$$

Gdy mamy do czynienia z modelem „zamkniętym” (gdymy koniecznie musimy wybierać z  $\{a, b, c\}$ ) na podstawie  $m_1$  i  $m_2$  wtedy wybieramy  $b$ . Jednak w dalszym ciągu brakuje wskazania na informacje niezgodne. Shafer wprowadził wagę niezgodności informacji dla  $m_1, m_2$  i ma ona postać:

$$\text{Con}(m_1, m_2) = \log \frac{1}{\sum_{A \cap B \neq \emptyset} m_1(A) * m_2(B)}$$

Dla podanego przykładu waga niezgodności wynosi:

$$\text{Con}(m_1, m_2) = \log(100)$$

#### 4.

Rozpatrzmy teraz przykład w celu przybliżenia wzorów określających funkcję przekonania. Zakładając, że zmienna  $x$  może przyjmować wartości  $a, b, c$ . Mamy następujący zbiór  $\Theta = \{a, b, c, ab, ac, bc, abc\}$ . Na podstawie dostępnej informacji można podzielić stopnie pewności następującym podzbiорom  $\{a, b, c, ab, bc, abc\}$ . Przyjmijmy rozkład bazowego prawdopodobieństwa:

$$\begin{aligned} m(a) &= 0,2, \\ m(b) &= 0,1, \\ m(c) &= 0,1, \\ m(ab) &= 0,2, \\ m(bc) &= 0,3, \\ m(abc) &= 0,1 \end{aligned}$$

W tym przypadku funkcje Bel i Pl będą miały następujące wartości:

$$\begin{aligned} \text{Bel}(a) &= m(a) = 0,2, \\ \text{Bel}(b) &= m(b) = 0,1, \\ \text{Bel}(c) &= m(c) = 0,1, \\ \text{Bel}(ab) &= m(a) + m(b) + m(ab) = 0,5, \\ \text{Bel}(ac) &= m(a) + m(c) = 0,3, \\ \text{Bel}(bc) &= m(b) + m(c) + m(bc) = 0,5, \\ \text{Bel}(abc) &= m(a) + m(b) + m(c) + m(ab) + m(bc) + m(abc) = 1, \\ \text{Pl}(a) &= m(a) + m(ab) + m(abc) = 0,5, \\ \text{Pl}(b) &= m(b) + m(ab) + m(bc) + m(abc) = 0,7, \\ \text{Pl}(c) &= m(c) + m(bc) + m(abc) = 0,5, \\ \text{Pl}(ab) &= m(a) + m(b) + m(ab) + m(bc) + m(abc) = 0,9, \\ \text{Pl}(ac) &= m(a) + m(c) + m(bc) + m(abc) = 0,7, \\ \text{Pl}(bc) &= m(b) + m(c) + m(ab) + m(bc) + m(abc) = 0,8, \\ \text{Pl}(abc) &= m(a) + m(b) + m(c) + m(ab) + m(bc) + m(abc) = 1 \end{aligned}$$

Zdanie „ $x$  jest w  $ac$ “, mimo iż nie ma bezpośredniej wspierającej informacji ma niezerowe funkcje Bel, Pl.

5.

Rozpatrzmy zadanie diagnozowania choroby. Zbiór wszystkich możliwych diagnoz będzie w tym przypadku ramą rozróżniającą.  $\Theta = \{\text{grypa, katar, alergia, zapalenie płuc}\}$

Dla pacjenta, o którym nie ma żadnych wiadomości, wszystkie choroby są możliwe, dlatego początkowo definiujemy podstawowe przyporządkowanie prawdopodobieństwa jako:

$$m_0(X) = \begin{cases} 1 & \text{dla } X = \Theta \\ 0 & \text{w przeciwnym wypadku} \end{cases}$$

Rdzeń  $m$  jest równy  $\Theta$ . Po otrzymaniu dodatkowych informacji może się okazać, że katar lub alergia mają podstawowe przyporządkowanie prawdopodobieństwa równe 0,3 (taką informacją może być np. podwyższona temperatura). Powstaje wtedy nowa funkcja  $m_1$

$$m_1(X) = \begin{cases} 0,7 & \text{dla } X = \Theta \\ 0,3 & \text{dla } X = \{\text{katar, alergia}\} \\ 0 & \text{w przeciwnym wypadku} \end{cases}$$

Może okazać się, że pacjent nie cierpi na alergię, wtedy funkcja opisująca ten przypadek  $m_2$  ma postać:

$$m_2(X) = \begin{cases} 0,2 & \text{dla } X = \Theta \\ 0,8 & \text{dla } X = \{\text{grypa, katar, zapalenie}_{\text{płuc}}\} \\ 0 & \text{w przeciwnym wypadku} \end{cases}$$

6.

Oto inny przykład funkcji określonej na rozkładzie  $m$  w celu uwydatnienia różnych własności informacji:

$$Q(A) = \sum_{A \subseteq B, B \in \Theta} m(B)$$

Owa funkcja dla wartości z przykładu 4 będzie miała następujące wartości:

$$Q(a) = m(a) + m(ab) + m(abc) = 0,5$$

$$Q(b) = m(b) + m(ab) + m(bc) + m(abc) = 0,7$$

$$Q(c) = m(c) + m(bc) + m(abc) = 0,5$$

$$Q(ab) = m(ab) + m(abc) = 0,3$$

$$Q(ac) = m(abc) = 0,1$$

$$Q(bc) = m(bc) + m(abc) = 0,4$$



$$Q(abc) = m(abc) = 0,1$$

Wykazuje się, że rozkład bazowego prawdopodobieństwa  $m$ , z którego wyprowadza się funkcje przekonania jest jednoznaczny i może być odtworzony na podstawie funkcji przekonania. To samo można właśnie wykazać dla funkcji  $Q$ .

Funkcje  $Bel$ ,  $Pl$  oraz  $Q$  reprezentują różne makrowłasności zbioru  $\Theta$  zawarte w rozkładzie  $m$ .

W przypadku teorii Dempstera-Shafera, tak jak w metodach probabilistycznych wprowadza się miarę zawartości informacji. Powinien to być funkcjonał, przekształcający funkcje przekonania, spełniający następujące własności:

- wartość powinna być liczbą rzeczywistą nieujemną,
- w przypadku całkowitej ignorancji zawartość informacyjna powinna być równa zero,
- jeśli dwie proste funkcje przekonania  $Bel^1$ ,  $Bel^2$  mają identyczne wartości  $s$  odpowiednio dla zbiorów  $A_1$ ,  $A_2$ , gdzie  $A_1 \subset A_2$ , to zawartość informacyjna związana z  $Bel^1$  powinna być większa niż związana z  $Bel^2$ ,
- zawartość informacyjna związana z dwoma niesprzecznymi rozkładami powinna być addytywna.

Jedną z propozycji takiego funkcjonału jest

$$H(Bel) = - \sum_{A \subseteq \Theta} \log Q(A)$$

gdzie  $Q$  jest określone przez równanie wspomniane wyżej.

Prześledźmy kilka innych własności tak określonego funkcjonału. Załóżmy, że posiadamy dwa rozkłady bazowe  $m_1$ ,  $m_2$ . Zawartość informacyjna kombinacji tych rozkładów jest następująca:

$$H(Bel) = H(Bel^1) + H(Bel^2) + (2^n - 1) \text{Con}(m_1, m_2)$$

gdzie  $n$  jest liczbą elementów zbioru  $\Theta$  i  $\text{Con}(-, -)$  zostało określone w przykładzie 3.

7.

Trzy osoby Paweł, Gawł i Kasia były obecne w pokoju, w którym popełniono zbrodnię. Jest to na razie jedyna informacja, którą zna inspektor prowadzący śledztwo. Dlatego też aby śledztwo posunęło się dalej, inspektor zdecydował oprzeć się na wyniku eksperymentu losowego. Eksperymentem tym będzie rzut kostką. Jeśli rzut będzie parzysty, winowajcą będzie kobieta, jeśli nieparzysty, mężczyzna. Załóżmy, że wiemy kto był w pokoju, gdy popełniono zbrodnię oraz wiemy o eksperymencie losowym. Nie interesuje nas jednak jaki był wynik rzutu oraz jaka była metoda wyboru pomiędzy Pawłem a Gawłem, gdyby wynik był nieparzysty. Interesuje nas jedynie jaka jest szansa, że była to kobieta lub, że był to mężczyzna. Na podstawie dostępnej informacji szansę tę można ocenić jak 1 do 1.

Dowiadujemy się następnie, że Paweł w czasie zbrodni udał się na posterunek policji i dlatego ma świetne alibi. Zastanówmy się jaka jest obecnie szansa związana z płcią winowajcy.

Wiemy, że

$$\Theta = \{\text{Paweł, Gawel, Kasia}\}$$

I na podstawie pierwszego opisu sytuacji można określić jedynie rozkład  $m_0$ :

$$m_0(\Theta) = 1.$$

Eksperyment rzutu kostką pozwala na określenie następującego rozkładu:

$$\begin{aligned} m_1(\{\text{Kasia}\}) &= 0,5, \\ m_1(\{\text{Paweł, Gawel}\}) &= 0,5. \end{aligned}$$

Łącząc te dwa rozkłady uzyskamy:

$$m_{0,1}(\{\text{Kasia}\}) = \frac{m_0(\Theta)m_1(\{\text{Kasia}\})}{m_0(\Theta)m_1(\{\text{Paweł, Gawel}\}) + m_0(\Theta)m_1(\{\text{Kasia}\})} = 0,5,$$

$$m_{0,1}(\{\text{Paweł, Gawel}\}) = \frac{m_0(\Theta)m_1(\{\text{Paweł, Gawel}\})}{m_0(\Theta)m_1(\{\text{Paweł, Gawel}\}) + m_0(\Theta)m_1(\{\text{Kasia}\})} = 0,5.$$

Następnie dowiadujemy się o idealnym alibi Pawła. Możemy tę sytuację opisać rozkładem:

$$m_2(\{\text{Gawel, Kasia}\}) = 1.$$

Chcąc uzyskać ostateczny rozkład łączymy  $m_{0,1}$  z  $m_2$ :

$$m_{0,1,2}(\{\text{Kasia}\}) = \frac{m_{0,1}(\{\text{Kasia}\})m_2(\{\text{Gawel, Kasia}\})}{m_{0,1}(\{\text{Kasia}\})m_2(\{\text{Gawel, Kasia}\}) + m_{0,1}(\{\text{Paweł, Gawel}\})m_2(\{\text{Gawel, Kasia}\})}$$

$$m_{0,1,2}(\{\text{Gawel}\}) = \frac{m_{0,1}(\{\text{Paweł, Gawel}\})m_2(\{\text{Gawel, Kasia}\})}{m_{0,1}(\{\text{Kasia}\})m_2(\{\text{Gawel, Kasia}\}) + m_{0,1}(\{\text{Paweł, Gawel}\})m_2(\{\text{Gawel, Kasia}\})}$$

Widzimy więc, że dodatkowa informacja nie zmieniła oceny płci winowajcy. W dalszym ciągu możemy ocenić jak 1 do 1.