Preferencje i funkcja użyteczności

Wartości prawdopodobieństw obliczane przy pomocy sieci przekonań dostarczają informacji o względnych prawdopodobieństwach różnych zdarzeń, wariantów, itp. Jednak nie mówią jaką należy podjąć decyzję w oparciu o te wartości. Czy plan działania gwarantujący uzyskanie 90% celów z prawdopodobieństwem 0.95 jest lepszy niż plan gwarantujący uzyskanie 95% celów z prawdopodobieństwem 0.90 ?

Inteligentny agent musi mieć reprezentację swoich **preferencji** dla podejmowania decyzji. W tym celu będziemy stosowali pojęcie **funkcji użyteczności** (*utility*) $U(S)$ określającej które stany są korzystniejsze dla agenta. W oczywisty sposób, użyteczność jest pojęciem względnym, które można określić dla konkretnego agenta.

Zasada MEU

Agent, posiadający preferencje wyrażone wartościami użyteczności i posługujący się prawdopodobieństwem dla określania faktów o świecie i możliwych konsekwencji, działa racjonalnie jeśli wybiera akcję dające najwyższą oczekiwaną użyteczność (MEU = *Maximum Expected Utility*) uśrednioną po wszystkich możliwych wynikach tych akcji.

Oczekiwana użyteczność $EU(A)$ niedeterministycznej akcji A ze zbiorem możliwych wyników $\{\mathit{Result}_i(A)\}$ z prawdopodobieństwami $P(\mathit{Result}_i(A)|Do(A), E)$, gdzie E zawiera całą dostępną wiedzę agenta o świecie, a $Do(A)$ jest stwierdzeniem wykonania akcji A, jest dana jako:

$$EU(A|E) = \sum_i P(\mathit{Result}_i(A)|Do(A), E) \times U(\mathit{Result}_i(A))$$

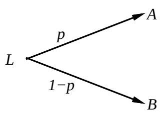
Jednak wybieranie akcji według zasady MEU nie jest łatwe. Aby tego dokonać należałoby określić wartości $P(\mathit{Result}_i(A)|Do(A), E)$, co wymaga NP-zupełnych obliczeń na sieci przekonań. Wyznaczenie wszystkich wartości $U(\mathit{Result}_i(A))$ często może wymagać nakładu pracy ze strony agenta, ponieważ może on nie znać z góry własności danego stanu, a więc i jego przydatności dla swoich celów.

Loterie i preferencje

Sytuację agenta podejmującego decyzje w warunkach niepewności będziemy nazywali **loterią** dla podkreślenia różnych możliwych wyników zdarzających się z różnymi prawdopodobieństwami. Wynik loterii może być konkretnym stanem lub kolejną loterią.

Na przykład, loterię L z dwoma możliwymi wynikami: A z prawdopodobieństwem p i B z prawdopodobieństwem $1-p$ możemy zapisać:

$$L = [p, A; 1-p, B]$$



Jako podstawę wyboru między loteriami lub stanami agent stosuje preferencje:

$A \succ B$ – A jest preferowane nad B

$A \sim B$ – nie ma wyraźnej preferencji między A i B

$A \succsim B$ – A jest preferowane nad B lub nie ma preferencji

Aksjomaty teorii użyteczności

Przyjmujemy, że preferencje agenta muszą spełniać następujące własności, zwane aksjomatami teorii użyteczności:

porządek

$$(A \succ B) \vee (B \succ A) \vee (A \sim B)$$

przechodniość

$$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$$

ciągłość

$$A \succ B \succ C \Rightarrow \exists p\ [p, A; 1-p, C] \sim B$$

(Ciągłość oznacza, że jeśli B ma preferencję pomiędzy A i C to istnieje pewne prawdopodobieństwo p, dla którego agent nie ma preferencji pomiędzy wyborem B (na pewno), a loterią pomiędzy A i C.)

podstawianie

$$A \sim B \Rightarrow [p, A; 1-p, C] \sim [p, B; 1-p, C]$$

(Jeśli nie ma preferencji pomiędzy dwoma loteriami A i B, to nie ma ich również pomiędzy dwoma innymi, bardziej złożonymi loteriami, które różnią się między sobą tylko wystąpieniem A i B. Własność ta zachodzi niezależnie od treści i prawdopodobieństw w tych innych loteriach.)

monotoniczność

$$A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1-p, B] \succsim [q, A; 1-q, B])$$

(Jeśli agent preferuje A względem B, to dla dwóch loterii z wynikami A i B preferuje również loterię, która z większym prawdopodobieństwem daje wynik A niż B.)

dekompozycja

$$[p, A;\ 1-p, [q, B;\ 1-q, C]] \sim [p, A;\ (1-p)q, B;\ (1-p)(1-q), C]$$

(Złożone loterie można zredukować do prostszych, stosując prawa prawdopodobieństwa.)

Znaczenie aksjomatów

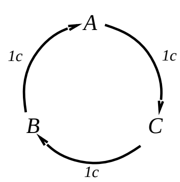
Aksjomaty są tak skonstruowane, że ich naruszenie musi być związane z zachowaniem jawnie irracjonalnym.

Wyobraźmy sobie agenta z systemem preferencji naruszającym aksjomat przechodniości: $A \succ B \succ C$ i jednocześnie $C \succ A$:

Jeśli $B \succ C$, to agent posiadający C powinien być skłonny zapłacić 1 (euro)centa aby osiągnąć B.

Jeśli $A \succ B$, to agent posiadający B powinien być skłonny zapłacić 1 (euro)centa aby osiągnąć A.

Jeśli $C \succ A$, to agent posiadający A powinien być skłonny zapłacić 1 (euro)centa aby osiągnąć C.



Takiego agenta możnaby zatem skłonić do oddania wszystkich pieniędzy jako konsekwencji posiadanych przez niego preferencji.

Funkcje użyteczności

Przedstawione aksjomaty teorii użyteczności w rzeczywistości nic nie mówią o użytecznościach, a jedynie o preferencjach agenta. Pojęciem wyjściowym do rozważań na temat podejmowania decyzji przez agentów będą ich preferencje spełniające powyższe aksjomaty.

Wiadomo że, jeśli zestaw preferencji agenta spełnia aksjomaty teorii użyteczności, to istnieje funkcja rzeczywista określona na zbiorze stanów $U : S \rightarrow \Re$, taka, że:

$$\begin{aligned} U(A) > U(B) \quad &\Leftrightarrow \quad A \succ B \\ U(A) = U(B) \quad &\Leftrightarrow \quad A \sim B \end{aligned}$$

Ta funkcja użyteczności przyjmuje dla loterii o określonych wynikach $S_1, ..., S_n$ i ich prawdopodobieństwach $p_1, ...p_n$ wartość:

$$U([p_1, S_1; ...; p_n, S_n]) = \sum_i p_i U(S_i)$$

Model racjonalnego agenta

Ponieważ wynikiem niedeterministycznych akcji agenta są loterie, zatem powyższy wzór może być użyty do obliczenia oczekiwanej użyteczności tych akcji, i jest realizacją zasady MEU wprowadzonej wcześniej wzorem:

$$EU(A|E) = \sum_i P(\mathit{Result}_i(A)|Do(A), E) \times U(\mathit{Result}_i(A))$$

Istnienie tej zasady i teorii użyteczności nie oznacza, że racjonalnie zachowujący się (czyli: maksymalizujący oczekiwaną wartość użyteczności) inteligentni agenci **jawnie** obliczają tę funkcję i jej maksima. Agenci mogą posługiwać się różnymi reprezentacjami wiedzy i sposobami obliczania swoich strategii. Jednak obserwując poczynania racjonalnie zachowującego się agenta, można zamodelować jego preferencje i funkcję użyteczności.

Teoria użyteczności w odniesieniu do pieniędzy

Rozważmy różne możliwe funkcje użyteczności w odniesieniu do stanów opisanych przez ilość posiadanych pieniędzy. Jest sensowne przyjąć, że funkcje użyteczności będą w tym przypadku monotoniczne dla konkretnych (pewnych) sum pieniędzy.

Rozważmy przykład: wygraliśmy teleturniej i mamy do wyboru nagrodę jednego miliona złotych, lub rzut monetą, i wtedy w przypadku orła otrzymamy trzy miliony, ale w przypadku reszki nic. Większość ludzi wybrałaby pewny milion na rękę (tylko co z podatkiem?), ale jeśli policzyć oczekiwaną wartość pieniężną (EMV) wariantu z rzutem monetą otrzymamy:

$$\frac{1}{2}(0\text{zł}) + \frac{1}{2}(3,000,000\text{zł}) = 1,500,000\text{zł}$$

podczas gdy EMV dla wariantu pewnego wynosi $1,000,000$zł.

Co wynika z tego przykładu? Spróbujmy policzyć użyteczności możliwych stanów wynikowych. Oznaczając przez S_k stan, w którym posiadamy początkowo k złotych mamy:

$$\begin{aligned} EU(\text{rzut\_monetą}) &= \frac{1}{2}U(S_k) + \frac{1}{2}U(S_{k+3,000,000}) \\ EU(\text{milion\_na\_rękę}) &= U(S_{k+1,000,000}) \end{aligned}$$

By określić użyteczność posiadania różnych sum pieniędzy możemy przyjąć, że początkowo większa suma gotówki jest dla nas bardziej użyteczna, np.: $U(S_k) = 5$, $U(S_{k+1,000,000}) = 8$, $U(S_{k+3,000,000}) = 10$. Wtedy otrzymamy $EU(\text{rzut\_monetą}) = 7.5$ i mamy podstawę by przyjąć oferowany nam pewny milion.

Z drugiej strony, gdybyśmy posiadali już wiele milionów, to może te wartości wyszłyby inne, i wtedy opłacałoby się może przyjąć rzut monetą w powyższej loterii.

Paradoks St.Petersburga

Otrzymujemy propozycję udziału w grze (Bernoulli, 1738), w której rzucamy monetą tak długo aż wyrzucimy reszkę, i gdy nastąpi to w n-tym rzucie to wygrywamy 2^n złotych. Na pewno opłaca się grać w tę grę. Pytanie jednak brzmi: ile bylibyśmy gotowi zapłacić za możliwość wzięcia w niej udziału? Prawdopodobieństwo wyrzucenia reszki w n-tym rzucie wynosi $1/2^n$, więc:

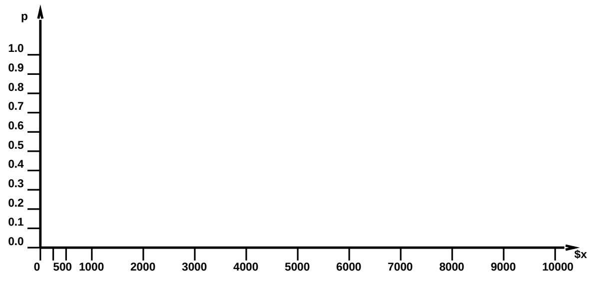
$$\text{EMV}(St.P.) = \sum_i P(\text{Reszka}_i)\text{MV}(\text{Reszka}_i) = \sum_i \frac{1}{2^i}2^i = \frac{2}{2}+\frac{4}{4}+\frac{8}{8}+\cdots = \infty$$

Czy to oznacza, że powinniśmy zapłacić dowolną (skończoną) sumę pieniędzy za możliwość udziału w takiej grze? Nie brzmi to poprawnie, i nie brzmiało również dla Bernoulliego, który zaproponował zastosowanie logarytmicznej funkcji użyteczności dla pieniędzy, $U(S_k) = \log_2 k$, dzięki czemu otrzymujemy:

$$EU(St.P.) = \sum_i P(\text{Reszka}_i)U(\text{Reszka}_i) = \sum_i \frac{1}{2^i}\log_2 2^i = \frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots = 1$$

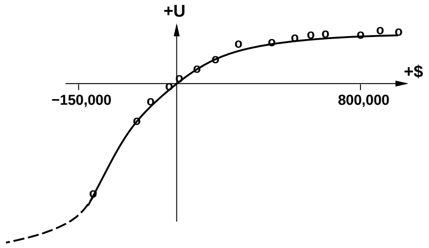
Użyteczność pieniędzy dla studenta

Dla każdej wartości x, określmy przez głosowanie, prawdopodobieństwo p, przy którym połowa grupy studenckiej preferuje loterię $[p, \$10,000; (1-p), \$0]$ ponad pewną wypłatę kwoty \$x:



Krzywe użyteczności pieniędzy

Przyjmuje się, że użyteczność pieniędzy jest logarytmo-podobną funkcją, która jest wklęsła dla dodatnich wartości monetarnych. W badaniach z lat 1960-tych wyznaczono doświadczalnie funkcję użyteczności pieniędzy dla pewnej konkretnej osoby $U(S_k) = -263.31 + 22.09 \log(k + 150,000)$:



Dla małych wartości ujemnych funkcja jest nadal wklęsła, ponieważ posiadanie długu powoduje zwykle u ludzie stan paniki. Jednak dla bardzo dużych wartości długu funkcja ta zdaje się mieć charakter wypukły, ponieważ „przykrość” z posiadania ogromnego zadłużenia nie wzrasta liniowo z poziomem tego zadłużenia.

Wracając do dodatnich wartości monetarnych, można stwierdzić, że agenci z wklęsłą funkcji użyteczności ogólnie będą preferować otrzymanie sumy oczekiwanej wygranej z loterii (na pewno), niż brać udział w tej loterii:

$$U(S_L) < U(S_{\mathsf{EMV}(L)})$$

Takie zachowanie można nazwać **ryzyko-fobią**. Obszar wypukłości funkcji użyteczności dla wielkich wartości długu można nazwać zachowaniem **ryzyko-lubnym**. W małych przedziałach funkcja użyteczności jest na ogół liniowa, i odpowiednie zachowanie jest nazywane **ryzyko-obojętnym**.

Nieracjonalność

Przyjęcie wklęsłej, logarytmo-podobnej funkcji użyteczności pieniędzy nie wyjaśnia całej psychologii podejmowania decyzji finansowych przez ludzi. Okazuje się, że systematycznie naruszają oni aksjomaty użyteczności. Na przykład, mając wybór między loteriami A i B, oraz C i D:

A:	80% wygrania \$4000	C:	20% wygrania \$4000
B:	100% wygrania \$3000	D:	25% wygrania \$3000

większość ludzi wybiera B ponad A, ale C ponad D. Jednak jeśli przyjmiemy $U(\$0) = 0$, to pierwszy wybór oznacza, że $0.8 \times U(\$4000) < U(\$3000)$, natomiast drugi wybór oznacza coś dokładnie przeciwnego.

Jednym możliwym wyjaśnieniem tego wyniku jest zwykła niezgodność zachowania się ludzi z aksjomatami użyteczności. Jednak inne wyjaśnienie jest oparte na uwzględnieniu poczucia **żalu**. Ludzie wiedzą, że w przypadku loterii A będą się czuli idiotycznie jeśli wybiorą, i następnie przegrają tę loterię, wiedząc, że mogli dokonać bezpiecznego i opłacalnego (choćby mniej) wyboru loterii B. W drugim przypadku to poczucie nie wystąpi, więc zachowanie jest racjonalne.

Znormalizowane funkcje użyteczności

Zauważmy, że aksjomaty użyteczności nie określają funkcji użyteczności jednoznacznie na podstawie samych preferencji. Na przykład, agent posługujący się funkcją użyteczności: $U'(S) = k_1 + k_2 U(S)$ gdzie k_1 i k_2 sa stałymi ($k_2 > 0$), będzie zachowywał się identycznie do agenta z funkcją $U(S)$, jeśli obaj agenci posiadają te same przekonania.

Funkcję użyteczności można zatem przeskalować liniowo i przesunąć o dowolną wartość w górę lub w dół, i zachowanie agenta nie zmieni się. Dlatego można posługiwać się **znormalizowaną funkcją użyteczności**.

Oznaczając przez $u_\perp$ użyteczność stanu „najgorszej katastrofy” $u_\perp = U(S_\perp)$ dla pierwotnej funkcji użyteczności $U(S)$, a przez $u_\top$ użyteczność stanu „największej nagrody” $u_\top = U(S_\top)$, dla znormalizowanej funkcji użyteczności U' przyjmiemy $U'(S_\perp) = 0$ i $U'(S_\top) = 1$, natomiast użyteczności stanów pośrednich $U'(S)$ określamy prosząc agenta o podanie prawdopodobieństwa p, dla którego agent nie ma preferencji pomiędzy stanem S a **loterią standardową** $[p, S_\top; (1-p), S_\perp]$

$$U'(S) \quad = \quad p, \text{ gdy } S \sim [p, S_\top; (1-p), S_\perp]$$

Podejmowanie decyzji

Bayesowskie sieci przekonań pozwalają na uzyskiwanie rozkładów prawdopodobieństw dowolnych zmiennych, przy posiadaniu informacji o dowolnej kombinacji innych zmiennych. Znając dodatkowo rozkład użyteczności, możemy tę wiedzę zastosować, z wykorzystaniem zasady MEU.

Rozważmy przykład: czy zabrać ze sobą parasol gdy wychodzimy z domu? Parasol jest użyteczny tylko wtedy, gdy pada deszcz, poza tym jest uciążliwy, i można go zgubić. Skąd jednak możemy wiedzieć, czy będzie padać? Pewną wskazówką jest prognoza pogody.

Weather	P(W)
sunny	0.7
rainy	0.3

Forecast	P(F|sunny)	P(F|rainy)
good	0.8	0.1
bad	0.2	0.9

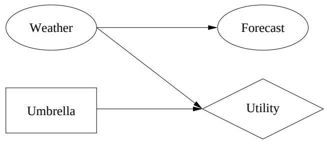
Przy okazji zauważmy, że powyższa sieć jest ciekawym przypadkiem zależności probabilistycznej, która zachodzi w kierunku odwrotnym do chronologii. Pogoda wpływa na prognozę, pomimo iż prognoza określana jest wcześniej.
Jak to możliwe? O to trzebaby zapytać meteorologów.

Diagramy wpływów

Zarówno rozważana akcja jak i użyteczności sytuacji mogą być wyrażone na grafie sieci przekonań jako specjalne węzły akcji i użyteczności. Sieć powinna zawierać połączenia z węzłów losowych i węzłów akcji do węzłów użyteczności.

Weather	P(W)
sunny	0.7
rainy	0.3

Forecast	P(F|sunny)	P(F|rainy)
good	0.8	0.1
bad	0.2	0.9



Umbrella
take
leave

Weather	sunny	sunny	rainy	rainy
Umbrella	take	leave	take	leave
Utility	20	100	70	0

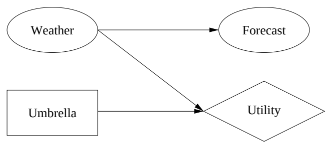
Takie rozszerzone sieci przekonań nazywane są **diagramami wpływów** (*influence diagrams*) albo też **sieciami decyzyjnymi** (*decision networks*). Niektóre narzędzia do budowy i przetwarzania sieci przekonań obsługują również takie diagramy wpływów.

Obliczanie decyzji

Najpierw rozważmy przypadek braku dodatkowych informacji o pogodzie.

Weather	P(W)
sunny	0.7
rainy	0.3

Forecast	P(F|sunny)	P(F|rainy)
good	0.8	0.1
bad	0.2	0.9



Umbrella
take
leave

Weather	sunny	sunny	rainy	rainy
Umbrella	take	leave	take	leave
Utility	20	100	70	0

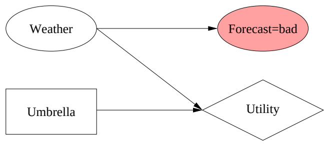
$$\begin{aligned} EU(\text{leave}) &= P(\text{sunny}) * U(\text{leave, sunny}) + P(\text{rainy}) * U(\text{leave, rainy}) \\ &= 0.7 * 100 + 0.3 * 0 \\ &= 70 \\ EU(\text{take}) &= P(\text{sunny}) * U(\text{take, sunny}) + P(\text{rainy}) * U(\text{take, rainy}) \\ &= 0.7 * 20 + 0.3 * 70 \\ &= 35 \end{aligned}$$

W tym przypadku wyższą użyteczność ma pozostawienie parasola w domu.

Załóżmy, że znana jest prognoza złej pogody. Prawdopodobieństwo pogody obliczone przez sieć wynosi teraz: $P(\text{sunny, rainy}|\text{bad}) = (0.341, 0.659)$.

Weather	P(W)
sunny	0.7
rainy	0.3

Forecast	P(F|sunny)	P(F|rainy)
good	0.8	0.1
bad	0.2	0.9



Umbrella
take
leave

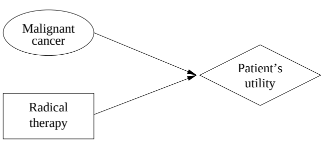
Weather	sunny	sunny	rainy	rainy
Umbrella	take	leave	take	leave
Utility	20	100	70	0

$$\begin{aligned} EU(\text{leave}|\text{bad}) &= P(\text{sunny}|\text{bad}) * U(\text{leave, sunny}) + P(\text{rainy}|\text{bad}) * U(\text{leave, rainy}) \\ &= 0.34 * 100 + 0.66 * 0 \\ &= 34 \\ EU(\text{take}|\text{bad}) &= P(\text{sunny}|\text{bad}) * U(\text{take, sunny}) + P(\text{rainy}|\text{bad}) * U(\text{take, rainy}) \\ &= 0.34 * 20 + 0.66 * 70 \\ &= 53 \end{aligned}$$

W tym przypadku właściwą decyzją jest zabranie parasola ze sobą.

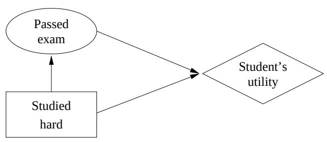
Krótkie podsumowanie — pytania sprawdzające

1. Rozważ problem pacjenta, który z pewnym prawdopodobieństwem ma nowotwór złośliwy i rozważane jest podjęcie terapii intensywnej. Sytuację opisuje pokazany poniżej diagram wpływów z częściowo wypełnioną tabelką użyteczności pacjenta.
   (a) Wyznacz i uzasadnij w racjonalny sposób brakujące użyteczności.
   (b) Oblicz wartość prawdopodobieństwa nowotworu, przy którym decyzja o terapii ulega zmianie.



Nowotwór złośliwy	Leczenie intensywne	Użyteczność pacjenta
N	N	100
N	T	
T	N	0
T	T	

2. Rozważ problem studenta, który przystępuje do sesji egzaminacyjnej i może uczyć się do egzaminu, bądź podejść bezstresowo, próbując zaliczyć „z biegu", i nie tracić czasu na uczenie się.
   (a) Rozważ dwie różne alternatywne filozofie życiowe studenta, i opisz je wartościami użyteczności w skali od 0 do 100.
   (b) Dla przykładowo wybranego przedmiotu, określ prawdopodobieństwo pierwotne zdania egzaminu, i dla wybranych filozofii życiowych studenta oblicz jego decyzje zgodne z zasadą MEU. (Traktujemy tu węzeł decyzyjny uczenia się jako reprezentujący niezależną decyzję, a zatem zdanie egzaminu jest w tym wypadku zmienną losową niezależną od uczenia się.)
   (c) Następnie, traktując uczenie się jako zmienną losową niezależną, określ warunkowy rozkład prawdopodobieństwa zdania egzaminu z wybranego poprzednio przedmiotu. Wybierz prawdopodobieństwo bezwarunkowe uczenia się i oblicz wartości oczekiwane użyteczności dla dwóch filozofii.



Uczył się	Zdał egzamin	Użyteczność studenta
N	N	
N	T	
T	N	
T	T	

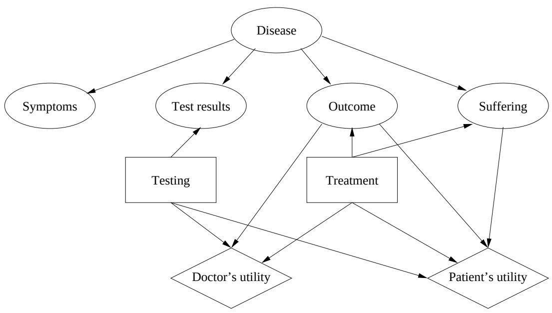
Wielokrotne akcje i wielokrotne użyteczności

W sieci przekonań jest zwykle wiele węzłów losowych, ponieważ głównym przeznaczeniem takich sieci jest uproszczenie obliczeń prawdopodobieństw warunkowych dla złożonych przypadków.

Jest za to zupełnie normalnym przypadkiem gdy w diagramie wpływów istnieje dokładnie jeden węzeł decyzyjny i jeden węzeł użyteczności. Wynika to z faktu, że diagramy wpływów przedstawiają wiedzę niezbędną do podjęcia pojedynczej decyzji. Aby określić tę decyzję jednoznacznie, najlepiej mieć jedno kryterium.

Jednak wiele węzłów akcji może znaleźć się w diagramie wpływów. Wyrażają one wtedy sytuację, gdy agent musi podjąć tylko jedną z tych decyzji, lub w jednym kroku decyzję łączną, na podstawie informacji z sieci. Diagramy wpływów nie pozwalają na poprawny wybór zestawu decyzji, gdy konsekwencje jednej z nich wpływają na kolejne.

Z drugiej strony, wielokrotne użyteczności, gdy są takie, muszą być zagregowane za pomocą jednego z modeli obliczania użyteczności wieloatrybutowej, opisanych poniżej.



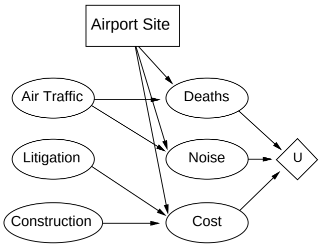
Na podstawie powyższej sieci można zapisać pacjentowi badania, lub podjąć decyzję o leczeniu, niezależnie od tego, czy wyniki badań są dostępne czy nie. Nie ma natomiast możliwości podjęcia decyzji o badaniach, i uwzględniających skutki tych badań decyzji o leczeniu, w jednym kroku.

Zwróćmy również uwagę na odrębne obliczanie użyteczności z punktu widzenia pacjenta i lekarza. Pomimo iż dla obu stron ważny jest końcowy wynik, to biorą one pod uwagę jeszcze inne czynniki, w odmienny sposób (np. prestiż lekarza i cierpienie pacjenta). Podejmowane na ich podstawie decyzje mogę być inne.

Użyteczności wieloatrybutowe

W wielu problemach praktycznych trzeba uwzględniać różne kryteria poprawności podejmowanych decyzji, albo, inaczej mówiąc, różne funkcje użyteczności, biorące pod uwagę różne atrybuty rozważanych stanów. Preferencje wynikające z uwzględnienia różnych atrybutów mogą być ze sobą sprzeczne.

Rozważmy przykład z lokalizacją lotniska. Czynnikami do uwzględnienia są: koszty gruntu, komplikacje przy jego zakupie, odległość od centrum miasta, zwiększony ruch na drogach, lokalne warunki pogodowe, i inne zagrożenia. Dla każdej możliwej lokalizacji można określić wartość krytycznych atrybutów takich jak: koszt całkowity, zagrożenie wypadkami (śmiertelnymi), uciążliwość ruchu lotniczego (hałas), itp.



Wieloatrybutowe funkcje użyteczności

Chcemy zbudować model podejmowania decyzji dla przypadków wieloatrybutowe. Będziemy oznaczać podlegające ocenie atrybuty stanu przez $X_1, X_2, ...$ a ich wartości $x_1, x_2, ...$. Przyjmiemy dla uproszczenia, że większe wartości atrybutów oznaczają wyższą jakość rozwiązania, z punktu widzenia danego atrybutu.

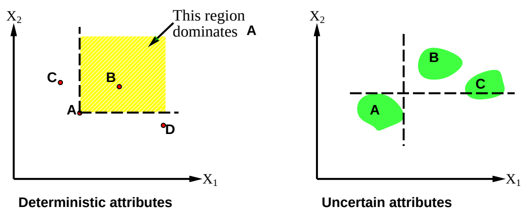
Zakładając, że poszczególne atrybuty mają funkcje użyteczności oznaczone jako: $f_1(x_1), f_2(x_2), ...$ możemy próbować wyrazić globalną funkcję użyteczności agenta wzorem:

$$U(x_1, x_2, ...x_n) = f(f_1(x_1), f_2(x_2), ..., f_n(x_n))$$

Niestety, na ogół funkcja $f()$ jest trudna do opisania w jawny sposób. Najpierw rozważymy więc szereg łatwiejszych przypadków szczególnych.

Ścisła dominacja

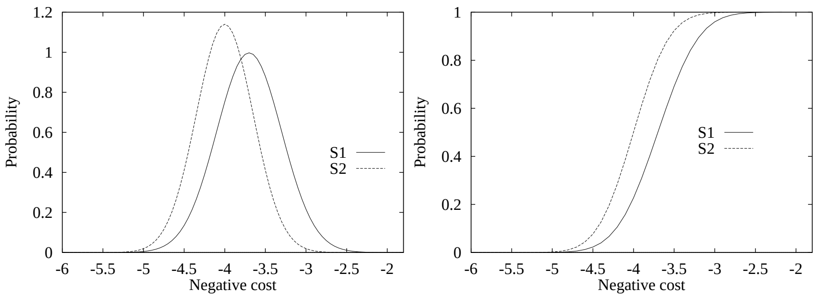
Zjawiskiem **ścisłej dominacji** nazwiemy sytuację, kiedy jeden wybór B posiada każdy z atrybutów lepszy od innego wyboru A. Na przykład, jedna lokalizacja lotniska jest tańsza, powoduje mniejsze obciążenie dla środowiska (naturalnego i ludzkiego), i jest bezpieczniejsza dla lotów. Wtedy możemy bez wahania odrzucić możliwość A z dalszych rozważań. Jednak nie można łatwo dokonać wyboru między możliwościami A a C lub A a D (diagram po lewej).



Ścisła dominacja lub jej brak może wystąpić również w warunkach niepewności, gdy dana akcja da wynik w postaci pewnej dystrybucji prawdopodobieństwa dla wartości atrybutów (diagram po prawej).

Dominacja stochastyczna

Ścisła dominacja nie zdarza się często w praktyce, i w ogólnym przypadku wartości atrybutów mogą nie być znane na pewno. Wtedy możemy posłużyć się rozkładami prawdopodobieństwa, i **dominacją stochastyczną**.



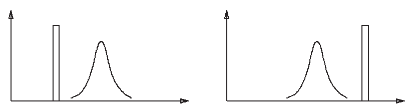
Na przykład, gdyby koszt lokalizacji lotniska w miejscu S_1 był przewidziany z dystrybucją normalną z wartością oczekiwaną \$3,700M i odchyleniem standardowym \$400M, a w miejscu S_2 z dystrybucją normalną z wartością oczekiwaną \$4,000M i odchyleniem standardowym \$350M, to S_1 dominuje stochastycznie S_2, co widać z przebiegu dystrybucji kumulacyjnych.

Jeśli agent rozpatruje dwie możliwe akcje A_1 i A_2, które prowadzą do dystrybucji prawdopodobieństw $p_1(x)$ i $p_2(x)$ na atrybucie X, to możemy powiedzieć, że A_1 dominuje stochastycznie A_2 jeśli:

$$\forall x \int_{-\infty}^{x} p_1(x')dx' \leq \int_{-\infty}^{x} p_2(x')dx'$$

W skrajnym przypadku, możemy nawet mieć dla jednego z przypadków (wyniku akcji A_2) wartość x_2 atrybutu X pewną (tzn. ze 100% prawdopodobieństwem), co wcale nie oznacza, że akcja A_2 powinna być preferowana (ani, że preferowana powinna być akcja A_1 z wynikiem w postaci rozkładu prawdopodobieństwa dla atrybutu X).

W zależności od konkretnego rozkładu akcja z niepewnym wynikiem może dominować stochastycznie akcję z wynikiem pewnym, albo na odwrót.



Użyteczności wieloatrybutowe — przypadek deterministyczny

W ogólnym przypadku żadna dominacja może nie zachodzić. Jednak atrybuty stanu X_1 i X_2 mogą być **niezależne preferencyjnie** od X_3 jeśli preferencje między $\langle x_1, x_2, x_3\rangle$ a $\langle x_1', x_2', x_3\rangle$ nie zależą od wartości x_3.

Jeśli każda para atrybutów X_i i X_j jest preferencyjnie niezależna od każdego z pozostałych atrybutów X_k, to ten zbiór atrybutów posiada własność **wzajemnej niezależności preferencyjnej** (MPI — *mutual preferential independence*).

Okazuje się, że w takim przypadku istnieje **addytywna funkcja wartości** opisująca preferencje agenta:

$$V(S) = \sum_i V_i(X_i(S))$$

W wielu przypadkach ta formuła pozwala prawidłowo modelować rzeczywiste preferencje i podejmowane decyzje.

Użyteczności wieloatrybutowe — przypadek probabilistyczny

Można uogólnić własność MPI na loterie: zbiór atrybutów **X** jest **niezależny użytecznościowo** od zbioru atrybutów **Y**, jeśli preferencje między loteriami na atrybutachz **X** są niezależne od konkretnych wartości atrybutów z **Y**. Zbiór atrybutów jest **wzajemnie niezależny użytecznościowo** (MUI — *mutually utility-independent*), jeśli każdy podzbiór jego atrybutów jest niezależny użytecznościowo od pozostałych atrybutów.

Dla atrybutów MUI zachowanie agenta może być opisane **multiplikatywną funkcją użyteczności**, którą dla przypadku trzyatrybutowego można zapisać:

$$\begin{aligned} U(S) &= k_1U_1(X_1(S)) + k_2U_2(X_2(S)) + k_3U_3(X_3(S)) \\ &+ k_1k_2U_1(X_1(S))U_2(X_2(S)) + k_2k_3U_2(X_2(S))U_3(X_3(S)) + k_3k_1U_3(X_3(S))U_1(\\ &+ k_1k_2k_3U_1(X_1(S))U_2(X_2(S))U_3(X_3(S)) \end{aligned}$$

W pewnych szczególnych przypadkach istnieje również całkowicie addytywna funkcja użyteczności.

Wartość informacji

W dotychczasowych rozważaniach przyjmowaliśmy założenie, że cała dostępna informacja jest znana agentowi podejmującemu decyzje. To założenie jest mało realistyczne. W praktycznych przypadkach podejmowania decyzji jedną z najważniejszych i jednocześnie najtrudniejszych kwestii jest na jakie pytania dotyczące problemu należy zebrać odpowiedzi.

> Np. lekarz nie ma wszystkich kluczowych informacji o pacjencie w chwili gdy zapoznaje sie z jego przypadkiem. Może zatem zlecić wykonanie pewnych badań, jednak takie zbieranie informacji jest z jednej strony kosztowne, a z drugiej wpływa na opóźnienie leczenia.

Ważność informacji zależy od dwóch czynników: (1) czy różne możliwe wyniki badań mocno wpłyną na podejmowaną decyzję, i (2) prawdopodobieństwa różnych wyników.

Teoria wartości informacji pozwala podejmować decyzje, które informacje należy zebrać.

Wartość informacji — przykład

Załóżmy, że firma wiertnicza rozważa możliwość wykupu praw wiercenia w jednym z n obszarów oceanu. Załóżmy dalej, że jest pewne, że dokładnie w jednym z obszarów położone jest złoże, którego eksploatacja przyniesie zysk C, natomiast koszt każdego obszaru wynosi C/n. Zauważmy, że wartość oczekiwana zysku EP z tej transakcji wynosi 0:

$$EP = \frac{1}{n}\left(C - \frac{C}{n}\right) + \frac{n-1}{n}\left(-\frac{C}{n}\right) = 0$$

Załóżmy dalej, że pewien geolog wie na pewno, czy w jednym konkretnym obszarze jest ropa. Jaka może być wartość tej informacji dla firmy wiertniczej?

Rozważmy przypadki. Z prawdopodobieństwem $\frac{1}{n}$ wybrany obszar zawiera ropę, i w takim przypadku firma kupi ten obszar i zarobi C minus C/n koszt praw. W przeciwnym wypadku, z prawdopodobieństwem $\frac{n-1}{n}$ obszar nie zawiera ropy, i firma wiedząc to, kupi inny obszar, zawierający ropę z prawdopodobieństwem $\frac{1}{n-1}$, i zarobi, być może, C, z oczekiwaną wartością zysku $C/(n-1)$, ponownie minus C/n:

$$EP' = \frac{1}{n}\left(C - \frac{C}{n}\right) + \frac{n-1}{n}\left(\frac{C}{n-1} - \frac{C}{n}\right) = \frac{C}{n}$$

Wniosek: informacja ma wartość, w tym przypadku równą cenie prawa do eksploatacji złoża.

Wartość dokładnej informacji

Załóżmy, że aktualna wiedza agenta to E i jego celem jest określenie najlepszej akcji α spośród wszystkich możliwych akcji A. Określamy wartość oczekiwaną (uśrednianą po różnych wynikach $\mathit{Result}_i(A)$ tej akcji) użyteczności tej akcji:

$$EU(\alpha|E) = \max_A \sum_i U(\mathit{Result}_i(A))P(\mathit{Result}_i(A)|\mathit{Do}(A), E)$$

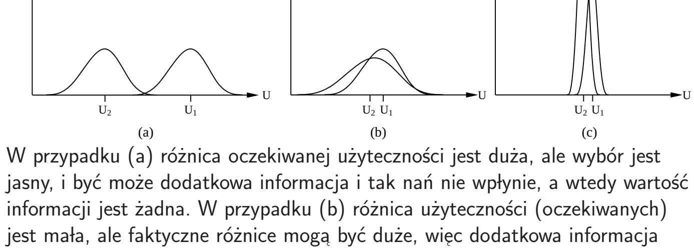
Gdyby jednak agent zdobył wiedzę o wartości pewnej zmiennej losowej E_j to wartość oczekiwana użyteczności wtedy wybranej akcji α_{E_j} byłaby:

$$EU(\alpha_{E_j}|E, E_j) = \max_A \sum_i U(\mathit{Result}_i(A))P(\mathit{Result}_i(A)|\mathit{Do}(A), E, E_j)$$

Ponieważ E_j jest zmienną losową o nieznanej wartości, to decyzję o tym, czy chcemy poznać jej wartość musimy podjąć biorąc pod uwage wszystkie możliwe jej wartości, i to co obecnie o nich wiemy. **Wartość dokładnej informacji** (VPI, *value of perfect information*) o zmiennej losowej E_j obliczamy jako:

$$\mathit{VPI}_E(E_j) = \left(\sum_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$$

Wyobraźmy sobie, że mamy do wyboru tylko dwie akcje A_1 i A_2, oraz ich użyteczności z wartościami oczekiwanymi U_1 i U_2. Zdobycie informacji E_j spowoduje, że oczekiwane użyteczności tych akcji zmienią się na U_1' i U_2'. Znając wartości U_1, U_2, oraz U_1', U_2', możemy podjąć decyzję, czy opłaca się zdobywać informację.



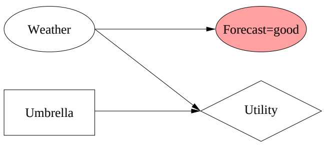
W przypadku (a) różnica oczekiwanej użyteczności jest duża, ale wybór jest jasny, i być może dodatkowa informacja i tak nań nie wpłynie, a wtedy wartość informacji jest żadna. W przypadku (b) różnica użyteczności (oczekiwanych) jest mała, ale faktyczne różnice mogą być duże, więc dodatkowa informacja może mieć istotną wartość. W przypadku (c) różnica wartości oczekiwanych jest mała, jak również wariancja obu zmiennych, więc ostateczna różnica prawdopodobnie będzie niewielka, i informacja również ma wartość znikomą.

Przykład: wartość informacji o pogodzie

Wróćmy do przykładu z deszczem i noszeniem parasola. Obliczyliśmy wcześniej:

$$\begin{aligned} MEU(\text{Umbrella}) &= \max_a EU(a) = 70 \\ MEU(\text{Umbrella}|\text{bad}) &= \max_a EU(a|\text{bad}) = 53 \end{aligned}$$

Możemy jeszcze dodatkowo obliczyć użyteczność najlepszej decyzji w przypadku dobrej pogody (w oczywisty sposób optymalną decyzją będzie wtedy „leave" ponieważ ta decyzja zwyciężyła już nawet przy braku wiedzy o pogodzie):



$$MEU(\text{Umbrella}|\text{good}) = \max_a EU(a|\text{good}) = 95$$

Aby obliczyć wartość prognozy pogody musimy znać rozkład prawdopodobieństwa dla zmiennej Forecast. Można go uzyskać odpytując sieć przekonań: $P(\text{good}, \text{bad}) = (0.59, 0.41)$. Dalej:

$$\mathit{VPI}_E(E_j) = \left(\sum_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) \right) - MEU(\alpha|E)$$

$$\begin{aligned} \mathit{VPI}(\text{Forecast}) &= P(\text{good}) EU(\alpha_{\text{good}}|\text{good}) + P(\text{bad}) EU(\alpha_{\text{bad}}|\text{bad}) - MEU(\alpha) \\ &= P(\text{good}) MEU(\text{good}) + P(\text{bad}) MEU(\text{bad}) - MEU(\alpha) \\ &= 0.59 * 95 + 0.41 * 53 - 70 \\ &= 7.78 \end{aligned}$$

Dla rozkładu użyteczności zdefiniowanego w tym problemie, wartość prognozy pogody wynosi 7.78, wyrażona w jednostkach użyteczności. Gdybyśmy mogli kupić prognozę pogody o wiarygodności przyjętej w powyższym obliczeniu, i poniżej tej ceny, to opłaca się to zrobić aby podjąć lepszą decyzję w kwestii parasola.

Własności wartości informacji

Twierdzenie: wartość informacji jest nieujemna.

Może jednak osiągnąć wartość zero, gdy np. znajomość innych faktów czyni daną informację bezużyteczną. Wartość informacji nie jest więc addytywna.

Z kolei wartość informacji o dwóch zmiennych losowych nie zależy od kolejności gromadzenia tych informacji. Jeśli znamy wartości zmiennych E_i i E_j to wnioski, które płyną z tej łącznej wiedzy nie zależą od tego kiedy i w jakiej kolejności agent te informacje pozyskał.

$$VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j)$$

Jednak wartość dwóch różnych informacji może być różna, i agent mógłby próbować obliczyć, które informacje przyniosłyby mu większy zysk (z uwzględnieniem ich kosztów).

Krótkowzroczny agent gromadzący informacje

Inteligentny agent powinien zadawać użytkownikowi pytania w sensownej kolejności, unikać zadawania pytań nieistotnych, brać pod uwagę ważność informacji względem ich kosztu, i przestać zadawać pytania kiedy to nie ma już sensu. Takie działanie można osiągnąć posługując się wartością informacji.

Możliwy algorytm agenta: wybierz informację, której zdobycie przyniesie największy oczekiwany zysk netto (zmniejszony o koszt zdobycia informacji), i jeśli ten zysk netto jest dodatni to zdecyduj się zdobywać tę informację. W przeciwnym wypadku przejdź do rzeczywistego działania.

Ten algorytm agenta jest **krótkowzroczny**, ponieważ kieruje się zyskiem z jednej tylko informacji, podczas gdy pozyskanie więcej niż jednej informacji mogłoby się okazać korzystne. Jest to w pewnym sensie analogiczne do zachłannej strategii przeszukiwania, i podobnie jak strategie zachłanne może przynieść sukces w niektórych przypadkach.

W ogólnym przypadku agent inteligentny mógłby rozważyć różne podzbiory zmiennych losowych i sekwencje pytań o nie.

Krótkie podsumowanie — pytania sprawdzające

1. Dla problemu pacjenta z chorobą nowotworową przedstawionego w pytaniu na stronie 21, i wartości użyteczności tam wybranych, oblicz wartość dokładnej informacji o nowotworze.