MODELOWANIE RZECZYWISTOŚCI

Daniel Wójcik
Instytut Biologii Doświadczalnej PAN

d.wojcik@nencki.gov.pl

tel. 5892 424

http://www.neuroinf.pl/Members/danek/swps/

Podręcznik

Iwo Białynicki-Birula
Iwona Białynicka-Birula

ISBN: 83-7255-103-0
Data wydania: 6 maja 2002

Informacja i niepewność

- Matematyczna teoria informacji zajmuje się pojemnością kanału transmisji informacji, zupełnie abstrahuje od znaczenia, wartości i sensu przekazywanej informacji.
- Informacja i niepewność to dwie strony tego samego medalu: zdobywając informację usuwamy niepewność i na odwrót, tracąc informację powiększamy niepewność.
- Im większa niepewność co do poszukiwanego wyniku, tym więcej informacji zdobywamy poznając ten wynik.

Bit

- Miarą informacji jest **bit –** skrót od binary digit. Jest to miara informacji otrzymanej w odpowiedzi na **elementarne pytanie**, to jest pytanie na które odpowiedź może brzmieć tylko „tak” lub „nie”.
- Większe jednostki to **bajt**, **kilobajt**, **megabajt**, itd.
- UWAGA:
  kilometr to 1000=10^3 metrów
  kilobajt to 1024 = 2^{10} bajtów

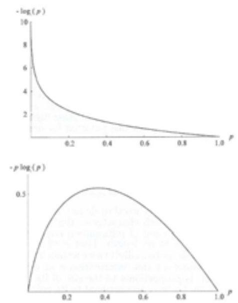
Wzór Shannona

- Zawartość informacyjna przekazu złożonego z n znaków wyrażona jest przez prawdopodobieństwa występowania tych znaków p_i

$$H = -\sum_{i=1}^{n} p_i \log_2 (p_i)$$

- Wielkość H oznacza informację mierzoną w bitach. Nazywa się ją **entropią informacyjną**.

P log(P)



Własności informacji

- Informacja, jaka może być zawarta w danym ciągu znaków jest proporcjonalna do długości tego ciągu. (Informacja jest wielkością **ekstensywną**)
- Przyjmijmy, że informacja jest zapisana w alfabecie binarnym (0,1)
- Słowem binarnym jest ciąg zer i jedynek o długości N. Liczba N mierzy objętość nośnika informacji. Informacja zawarta w słowie jest proporcjonalna do N.

Dlaczego logarytm

- Wybieramy jednostkę informacji tak, że

  Informacja H = Długość_słowa_binarnego

  Przy takim wyborze słowo złożone z jednego znaku niesie jeden bit informacji.

- Istnieje 2^N słów binarnych o długości N znaków. Zatem

  Długość_słowa = $\log_2$ (Liczba_słów)

Dlaczego logarytm cd

- Jeżeli prawdopodobieństwo wystąpienia każdego słowa jest takie samo, to wynosi ono

  p=1/Liczba_słów

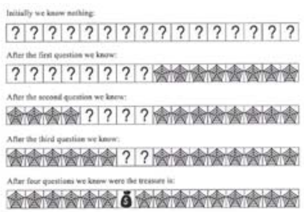
- Wobec tego informacja zawarta w pojedynczym słowie wynosi

  $H = - \log_2 (p)$

Gra w 20 pytań

- Miara informacji jest równa liczbie pytań potrzebnych do odgadnięcia słowa
- Rozważmy uproszczoną sytuację, kiedy jest 2^N równoprawdopodobnych słów o długości N. Ponumerujmy je wszystkie liczbami naturalnymi od 1 do 2^N.
- Mamy 2^N zakrytych komórek. W jednej z nich jest „skarb". Znalezienie „skarbu" jest tym samym co odgadnięcie słowa.

Najprostsza strategia



20 pytań cd

- Liczba pytań potrzebna do uzyskania pełnej informacji równa jest początkowej niepewności
- Na ogół prawdopodobieństwa wystąpienia różnych słów nie są takie same.
- Kiedy mamy odgadnąć słowo postaci KU_A nie wiemy, czy jest to KUFA, KULA, KUMA, KUNA, KUPA czy KURA.
- Kiedy mamy słowo postaci ŚW_T, to nie ma problemu.

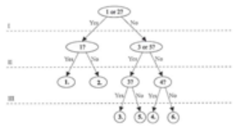
Rozkład 1

- Rozważmy skarb ukryty w jednej z 4 komórek, z prawdopodobieństwami $p_1 = 0.5$, $p_2 = 0.25$, $p_3 = 0.125$, $p_4 = 0.125$
- Pierwotna strategia daje średnio 2 pytania do osiągnięcia sukcesu
- Lepsza strategia:
  - Czy skarb jest w pierwszej komórce?
  - Czy skarb jest w drugiej komórce?
  - Czy skarb jest w trzeciej komórce?
- Średnio prowadzi ona do
  1*0.5+2*0.25+3*0.25=7/4 < 2 pytań
  zatem jest lepsza od bisekcji.

Rozkład 2

- Rozważmy rozkład 6 komórkowy:
  $p_1 = 1/3$, $p_2 = 1/5$, $p_3 = 1/5$,
  $p_4 = 2/15$, $p_5 = 1/15$, $p_6 = 1/15$

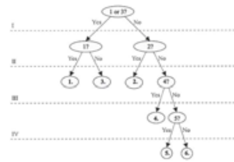
Średnia liczba pytań wynosi tu 37/15



Rozkład 2 – druga strategia

- Rozważmy rozkład 6 komórkowy:
  $p_1 = 1/3$, $p_2 = 1/5$, $p_3 = 1/5$,
  $p_4 = 2/15$, $p_5 = 1/15$, $p_6 = 1/15$

Średnia liczba pytań wynosi tu 36/15



Optymalna strategia – algorytm Huffmana

- Z początkowego rozkładu $p^0_1, p^0_2, \ldots, p^0_n$ wybieramy dwa najmniej prawdopodobne zdarzenia p^0_i oraz p^0_j.
- Łączymy je w jedno o prawdopodobieństwie p^1_k, mamy nowy rozkład $p^1_1, p^1_2, \ldots, p^1_{n-1}$
- Powtarzamy procedurę n-1 razy
- W ten sposób, od dołu, powstaje optymalne drzewo pytań.

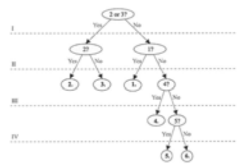
Przykład działania strategii

- Zaczynamy od rozkładu $p^0_1=1/3$, $p^0_2=1/5$, $p^0_3=1/5$, $p^0_4=2/15$, $p^0_5=1/15$, $p^0_6=1/15$
- Łączymy p^0_5 i p^0_6 w p^1_5. Dostajemy: $p^1_1=1/3$, $p^1_2=1/5$, $p^1_3=1/5$, $p^1_4=2/15$, $p^1_5=2/15$
- Łączymy p^1_4 i p^1_5 w p^2_4. Dostajemy: $p^2_1=1/3$, $p^2_2=1/5$, $p^2_3=1/5$, $p^2_4=4/15$
- Łączymy p^2_2 i p^2_3 w p^3_3. Dostajemy: $p^3_1=1/3$, $p^3_2=4/15$, $p^3_3=6/15$
- Łączymy p^3_1 i p^3_2 w p^4_2. Dostajemy: $p^4_1=9/15$, $p^4_2=6/15$

Rozkład 2 – trzecia strategia

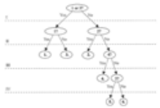
- Rozważmy rozkład 6 komórkowy:
  $p_1 = 1/3$, $p_2 = 1/5$, $p_3 = 1/5$,
  $p_4 = 2/15$, $p_5 = 1/15$, $p_6 = 1/15$

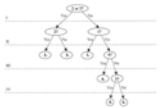
Średnia liczba pytań wynosi tu również 36/15



Porównanie dwóch ostatnich strategii

- Rozważmy rozkład 6 komórkowy:
  $p_1 = 1/3$, $p_2 = 1/5$, $p_3 = 1/5$,
  $p_4 = 2/15$, $p_5 = 1/15$, $p_6 = 1/15$





Średnia liczba pytań wynosi dla obu strategii 36/15

Średnia informacja

- Nasze rozważania pokazują, że entropia Shannona mierzy średnią informację obliczoną w przypadku, gdy znane są wszystkie prawdopodobieństwa elementarne.
- Ogólne twierdzenie o bezszumowym kodowaniu możemy sformułować tak:

  Nie istnieje strategia o średnio mniejszej liczbie pytań niż entropia Shannona

Doświadczenia Hymana

- R. Hyman pokazał, że czas reakcji na bodźce o określonej zawartości informacji jest proporcjonalny do entropii Shannona.
- Program **Hyman**

Układ dynamiczny

- *Układ*: zbiór obiektów, których zmiany w czasie chcemy badać
- *Dynamiczny*: zmieniający się w czasie zgodnie z pewną regułą

Typy układów dynamicznych

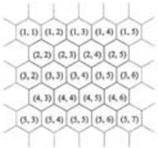
- Układy deterministyczne
  - Układy autonomiczne
  - Układy zależne od czasu
- Układy losowe (stochastyczne)

Przykład 1: Wyborcy

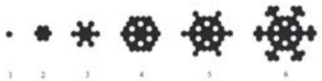
- *Stan układu*: Preferencje wyborcze grupy respondentów
- *Ewolucja*: Zmiany preferencji następują na skutek kontaktów międzyludzkich

Przykład 2: Płatek śniegu

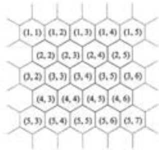
- Automat komórkowy na sieci o strukturze plastra miodu
- Ten sam stan początkowy zawsze daje ten sam stan końcowy (układ deterministyczny)



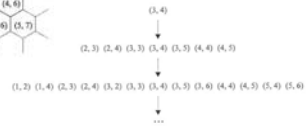
Płatek śniegu



Reguła: Nowy fragment powstaje w komórce sieci, która ma dokładnie jednego zmarzniętego sąsiada



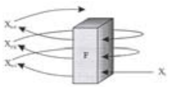
Płatek śniegu



Ewolucja układu dynamicznego

- Rozwój w czasie układu dynamicznego opisuje równanie ewolucji w postaci wzoru iteracyjnego

$$X_{n+1} = F_p(X_n)$$



- X_n i X_{n+1} opisują stan układu w chwili *n* i *n+1*
- F_p określa jakim zmianom ulega stan układu w kolejnych krokach iteracji

Przykład 3: algorytm Newtona

- Jak policzyć pierwiastek z a?
- Należy wybrać dowolną liczbę i iterować ją zgodnie z regułą

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

- Dlaczego?
  Jeżeli $x < \sqrt{a}$ to $\sqrt{a} * x < a$, a więc $\frac{a}{x} > \sqrt{a}$
  i na odwrót.

Przykład 3: pierwiastek z 2

- Weźmy $x_0=2$
- Wtedy
  - $x_1=1{,}5$
  - $x_2=1{,}416$
  - $x_3=1{,}414215$
  - $X_4=1{,}414213562374$
  - $x_5=1{,}4142135623730950488016 89$
  - ...
- I tak dalej

Kilka pojęć

- *Punkt stały*: stan układu, który się nie zmienia z czasem
- *Atraktor*: stan, do którego układ dąży z czasem
- *Dorzecze (basen) atraktora*: zbiór stanów początkowych, które dążą do danego atraktora

Przykład basenów atraktora

- Rozważmy odwzorowanie płaszczyzny

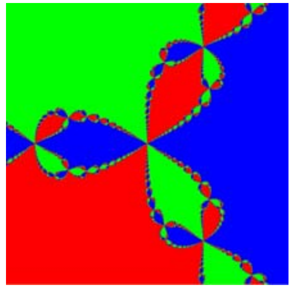
$$x_{t+1}=\frac{2}{3}x_t+\frac{1}{3}\frac{x_t^2-y_t^2}{(x_t^2+y_t^2)^2}$$

$$y_{t+1}=\frac{2}{3}y_t-\frac{1}{3}\frac{2\,x_t\,y_t}{(x_t^2+y_t^2)^2}$$

- To odwzorowanie ma trzy atraktory:

$$(1,0),\left(\frac{-1}{2},\frac{\sqrt{3}}{2}\right),\left(\frac{-1}{2},\frac{-\sqrt{3}}{2}\right)$$

Baseny pierwiastków jedynki



Model Ehrenfesta: psy i pchły

- N pcheł siedzi na dwóch psach: na Azorze A_n, a na Burku B_n
- Co chwilę jedna z pcheł podejmuje decyzję o przeskoku z jednego psa na drugiego
- Jak zmienia się rozkład pcheł?
  $A_{n+1} = A_n - 1$ z prawdopodobieństwem A_n/N
  $A_{n+1} = A_n + 1$ z prawdopodobieństwem $(N-A_n)/N$

Determinizm a losowość

- Większość procesów w przyrodzie ma składowe deterministyczną i losową
- Jeżeli „szum" jest nieduży, modelujemy układ jako ściśle deterministyczny
- Jeżeli „szum" jest dominujący, modelujemy układ jako proces stochastyczny

W ogólności

- Jesteśmy bez szans...

Chaos deterministyczny

- Proste iteracje odwzorowań:
  - Funkcja liniowa
  - Funkcja logistyczna
- Układy z ciągłym czasem
- Wpływ topologii na ewolucję

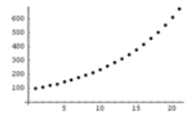
Przykład: funkcja liniowa

- Rozważmy populację much. Oznaczmy liczbę much w roku *n*-tym przez x_n. Załóżmy, że na każdą muchę w pokoleniu n średnio przypada R much w pokoleniu następnym. Wtedy ewolucja much jest dana przez:

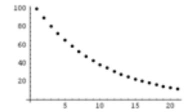
$$x_{n+1} = R\, x_n$$
$$x_n = R^n x_0$$

Przykład: funkcja liniowa

- Stały wzrost (*R*>1)



- Stały rozpad (*R*<1)



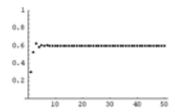
Przykład: funkcja logistyczna

Model liniowy jest zbyt dużym uproszczeniem. Wzrost populacji zwykle ograniczony jest przez czynniki zewnętrzne: drapieżców, ograniczone zasoby, itp. Biorąc pod uwagę skończoną ilość dostępnego pożywienia dostajemy model zwany logistycznym. Jeżeli wzrost R jest pomniejszony o wielkość proporcjonalną do x_n to dostajemy:

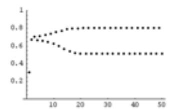
$$x_{n+1} = (R - a\, x_n)\, x_n$$
$$y_{n+1} = r(1 - y_n)\, y_n$$

Szereg czasowy

- W obszarze regularnym (r = 2.5)

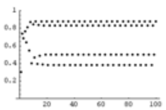


- W obszarze regularnym (r = 3.2)

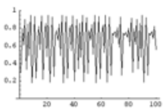


Szereg czasowy

- W obszarze regularnym (r = 3.5)

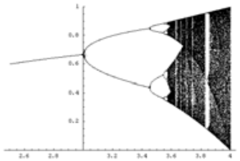


- W obszarze chaotycznym (r = 3.7)



Wykres bifurkacyjny

- Dla mapy logistycznej



Chaos deterministyczny

• Czuła zależność od warunków początkowych: Rysunek pokazuje różnicę między historią dwóch populacji różniących się początkowo o 0.0000001