Teoretyczne podstawy informatyki

Wykład 3b:
Kombinatoryka

http://hibiscus.if.uj.edu.pl/~erichter/Dydaktyka2010/TPI-2010

Kombinatoryka i prawdopodobieństwo

- Często spotykamy się z problemem obliczenia wartości wyrażającej prawdopodobieństwo zajścia określonych zdarzeń.
- Dziedzina matematyki zajmująca się tą tematyką to **kombinatoryka**.
- Pojęcia związane z próbami szacowania prawdopodobieństwa występowania zdarzeń definiuje **teoria prawdopodobieństwa**.

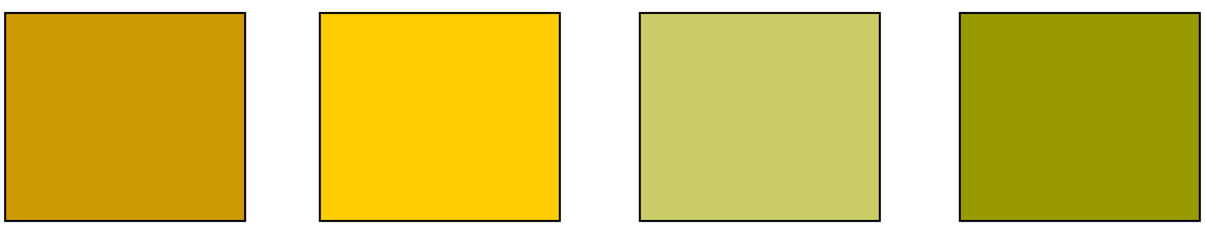
- Zacznijmy od kombinatoryki...

Wariacje z powtórzeniami

- Jednym z najprostszych, ale też najważniejszych problemów jest analiza listy elementów, z których każdemu należy przypisać jedną z wartości należących do stałego zbioru.
- Należy określić możliwą liczbę różnych przyporządkowań (*wariacji z powtórzeniami*) wartości do elementów.
- Przykład:

  4 kwadraty, każdy można pokolorować jednym z 3 kolorów.

  Ile możliwych pokolorowań? 3 • 3 • 3 • 3 = 34 = 81



Wariacje z powtórzeniami

- Mamy listę **n**-elementów. Istnieje zbiór **k**-wartości z których każda może być przyporządkowana do jakiegoś elementu. Przyporządkowanie jest listą **n** wartości (n_1, n_2,...n_n). Gdzie każda z n_1, n_2,...n_n jest jedną z wartości **k**.
- Istnieje k^n różnych przyporządkowań.
- **Twierdzenie:**
  S(n): liczba możliwych sposobów przyporządkowania dowolnej z k wartości do każdego z n elementów wynosi k^n.

Wariacje z powtórzeniami

- **Podstawa:**

  Przypadek podstawowy to **$n=1$**. Jeżeli mamy **1** element możemy wybrać dla niego dowolną spośród **k** wartości. Istnieje wiec **k** różnych przyporządkowań. Ponieważ **$k^1=k$**, podstawa indukcji jest prawdziwa.

- **Indukcja:**

  Załóżmy że **$S(n)$** jest prawdziwe i rozważmy **$S(n+1)$**,określające że istnieje **k^{n+1}** możliwych przyporządkowań jednej z **k** wartości do każdego z **$n+1$** elementów.

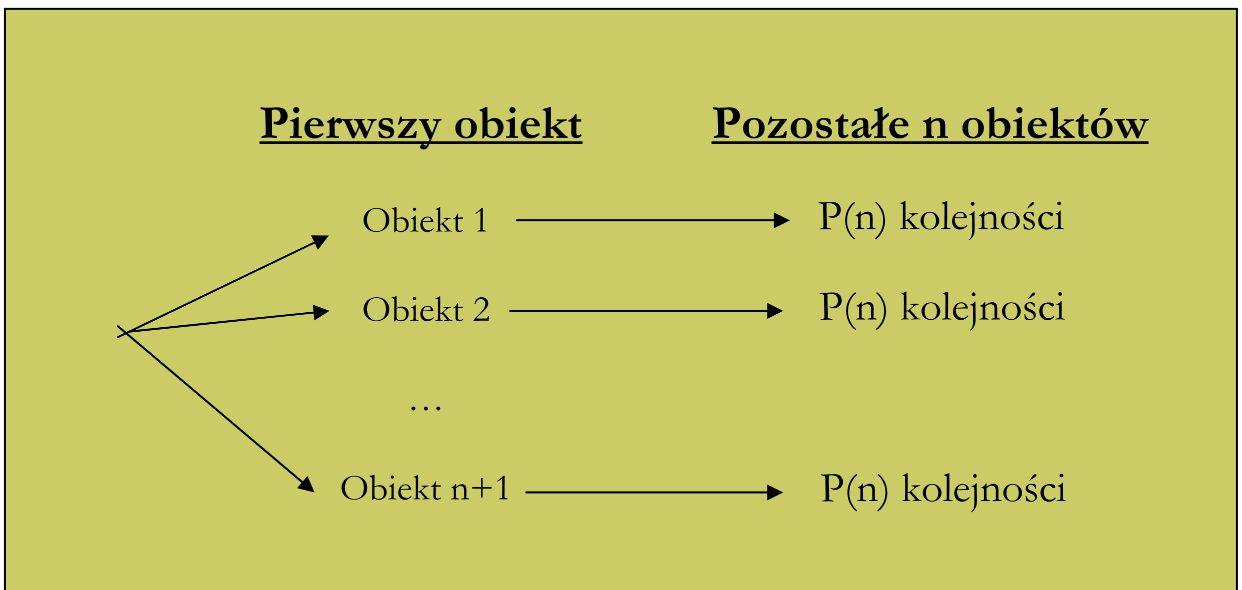
  Wiemy, że istnieje **k** możliwości doboru wartości dla pierwszego elementu. Zgodnie z hipoteza indukcyjną, istnieje **k^n** przyporządkowań wartości do pozostałych **n** elementów. Łączna liczba przyporządkowań wynosi
  $k \bullet k^n = k^{n+1}$. Cnd.

Permutacje

- Mając **n** różnych obiektów, na ile różnych sposobów można je uporządkować w jednej lini?
- Takie uporządkowanie nazywamy permutacją.
- Liczbę permutacji **n** obiektów zapisujemy jako **P(n)**.

Jak obliczyć P(n+1)?

- Problem: mamy **n+1** obiektów (**a_1, a_2,…,a_n, a_{n+1}**) które mają zostać posortowane.
- Ilość możliwych wyników sortowania jest **P(n+1)**



Permutacje n+1 obiektów

Jak obliczyć P(n+1)?

- **Twierdzenie**:
  P(n) = n! dla wszystkich **n ≥ 1**

- **Podstawa**:
  Dla **n=1**, **P(1)=1** określa że istnieje jedna permutacja dla jednego obiektu.

- **Indukcja**:
  Załóżmy ze **P(n) = n!**
  Wówczas wg. naszego twierdzenia: **P(n+1)=(n+1)!**
  Rozpoczynamy od stwierdzenia że **P(n+1)=(n+1) • P(n)**
  Zgodnie z hipotezą indukcyjną **P(n)=n!**, zatem **P(n+1)=(n+1) • n!**
  Zatem **P(n+1)=(n+1) • n! =(n+1) • n • (n-1) • (n-2) • ... • 1 = (n+1)!**, czyli nasze twierdzenie jest poprawne. Cnd.

- Jednym z interesujących zastosowań wzoru na liczbę permutacji jest dowód na to że **algorytmy sortujące musza działać w czasie co najmniej proporcjonalnym do (n log n), dla n elementów do posortowania**, chyba że wykorzystują jakieś specjalne własności sortowanych elementów.

Wariacje bez powtórzeń

- Niekiedy chcemy wybrać tylko niektóre spośród elementów zbioru i nadać im określony porządek.
- Uogólniamy opisaną poprzednio funkcję P(n) reprezentującą liczbę permutacji, aby otrzymać **dwuargumentową funkcje P(n,m),** którą definiujemy jako ilość możliwych sposobów wybrania **m elementów z n-elementowego zbioru**, przy czym **istotną** rolę odgrywa **kolejność wybierania elementów**, natomiast nieważne jest uporządkowanie elementów nie wybranych.
- Zatem **P(n) = P(n,n).**

Wariacje bez powtórzeń

Przykład:

- Ile istnieje sposobów utworzenia sekwencji **m** liter ze zbioru **n** liter, jeżeli żadna litera nie może występować więcej niż raz?
- Na sam początek możemy zauważyć warunek, by zadanie miało sens: $n \geq m$
- Pierwszą literę możemy wybrać na n sposobów (wybieramy ze zbioru **n**-elementowego), drugą na **n-1** sposobów (gdyż nie możemy wybrać tej samej litery co poprzednio), trzecią na **n-2** sposoby…
- Ostatnią na **n-(m-1)** sposobów.

Twierdzenie: $P(n,m) = n \cdot (n-1) \cdot \ldots \cdot n-(m-1)$ **dla wszystkich** $m \leq n$

Twierdzenie: $P(n,m) = \dfrac{n!}{(n-m)!}$ **dla wszystkich** $m \leq n$

Kombinacje

- Kombinacja to każdy podzbiór zbioru skończonego. Kombinacją ***m***-elementową zbioru ***n***-elementowego **A** nazywa się każdy ***m***-elementowy podzbiór zbioru **A** (**0≤*m*≤*n***). Używa się też terminu "kombinacja z ***n*** elementów po ***m*** elementów" lub wręcz "kombinacja z ***n* po *m***".
- Taką funkcję zapisujemy jako: $\binom{n}{m}$

$$\binom{n}{m} = \frac{P(n,m)}{P(m)} = \frac{n!}{(n-m)! \bullet m!}$$

Wyznaczanie liczby kombinacji

- **Rekurencyjny algorytm: (ilustruje tzw. trójkąt Pascala)**

- **Podstawa**: $\binom{n}{0}$ =1 dla dowolnego n≥1.
  Oznacza to że istnieje tylko jeden sposób wybrania **zero** elementów ze zbioru **n**-elementowego – wybranie niczego.
  Także $\binom{n}{n}$ =1, ponieważ jedynym sposobem wybrania **n**-elementów ze zbioru **n**-elementowego jest wybranie ich wszystkich.

- **Indukcja:** Jeśli O< m< n, to $\binom{n}{m}=\binom{n-1}{m}+\binom{n-1}{m-1}$.
  Oznacza to, że jeżeli chcemy wybrać **m** elementów ze zbioru **n**-elementowego, możemy albo:
  - nie wybrać pierwszego elementu, po czym wybrać m elementów z pozostałych n-1 elementów. Taką liczbę możliwości wyraża $\binom{n-1}{m}$.
  - wybrać pierwszy element, po czym wybrać m-1 elementów z pozostałych n-1 elementów. Taka liczbę możliwości wyraża $\binom{n-1}{m-1}$.

Trójkąt Pascala

- Rekurencję przy obliczaniu liczby kombinacji często ilustruje się przy pomocy **trójkąta Pascala**.
- $\binom{n}{m}$ = **(m+1)** liczba w **(n+1)** wierszu

podstawa

indukcja

1

1 1

1 2 1

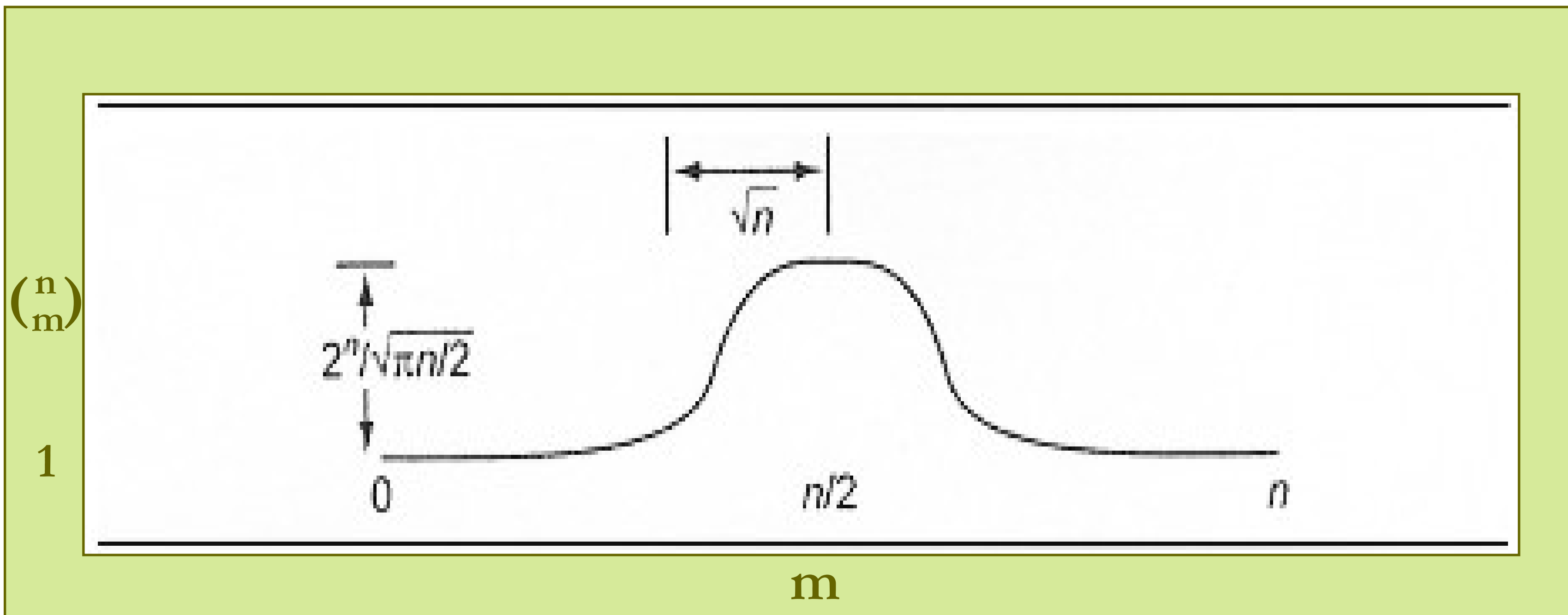
1 3 3 1

1 4 6 4 1

$\binom{4}{2}$ = 4!/(2! x 2!) = 6

Interesujące własności funkcji $\binom{n}{m}$

- To również są współczynniki rozkładu dwuwyrazowego wielomianu (dwumianu) $(x+y)^n$
- $\sum_{m=0}^{n} \binom{n}{m} = 2^n$
- Wykres funkcji $\binom{n}{m}$ dla stałej dużej wartości **n**:



Permutacje z powtórzeniami

- **Twierdzenie**:
  Jeżeli istnieje **n** elementów podzielonych na **k** grup o rozmiarach równych odpowiednio $i_1, i_2, i_3, \ldots, i_k$ gdzie elementy jednej grupy są identyczne, ale elementy różnych grup różnią się od siebie, liczba uporządkowań tych elementów wynosi
  $S(k) = n! / \prod_{j=1}^{k} i_j!$

- **Podstawa**:
  Dla **k=1**, istnieje tylko jedna grupa zawierająca identyczne elementy, które możemy uporządkować tylko w jeden sposób, niezależnie od liczności tego zbioru. Jeśli **k=1**, to $i_1=n$, zatem $S(1)=n!/n!=1$ jest prawdziwe.

- **Indukcja**:
  Załóżmy że **S(k)** jest prawdziwe i rozważmy sytuacje, w której mamy **k+1** grup. Niech ostatnia grupa składa się z $m=i_{k+1}$ elementów, występujących na **m** pozycjach,
  z których możemy je wybierać na $\binom{n}{m}$ sposobów.
  Stosując hipotezę indukcyjną otrzymujemy że $S(k+1)= \binom{n}{m} \cdot (n-m)! / \prod_{j=1}^{k} i_j!$ co łatwo można przekształcić (pamiętając że $m=i_{k+1}$) do postaci:
  $S(k+1)= n! / \prod_{j=1}^{k+1} i_j!$ a więc cnd.

- Jest to typowy problem przy układania anagramów

Łączenie reguł kombinatorycznych

- Typowy problem kombinatoryczny wymaga łączenia przedstawionych reguł (cegiełek) w bardziej skomplikowane struktury.
- Techniki których używamy to:
  - prowadzenie obliczeń jako sekwencji wyborów;
  - prowadzenie obliczeń jako różnicy innych obliczeń (np. wszystkich wyborów – nieprawidłowych wyborów);
  - prowadzenie obliczeń jako sumy rozwiązań dla podprzypadków które są wzajemnie rozłączne.