

Author, MSc: Adrian Szklarski: 10.2023

Space, rocket and aerospace technologies in science and programming.

Contact me:

LinkedIn: <https://www.linkedin.com/in/szklarskiadrian/>

GitHub: <https://github.com/AdrianSzklarski>

Subject: Probabilistic models/ Model probabilistyczny

<p>Algorithm analysis: is the study of the cost of an algorithm treated as a random variable. When studying an algorithm, one always analyzes the number of steps performed by the algorithm, in other words, the time cost.</p> <p>The first step of the analysis: the probabilistic space is the set of data accepted by the algorithm, so it is necessary to define these data (precise determination of the probabilistic model) and probability distribution</p> <p>Universality: the analysis carried out is based on the data of an algorithm of any size and various data so the cost of operation will depend on this data size and the specific layout of this data.</p>	<p>Analiza algorytmów: to badanie kosztu algorytmu traktowanego jako zmienna losowa. Podczas badań algorytmu zawsze analizuje się liczbę kroków wykonanych przez algorytm czyli inaczej mówiąc koszt czasowy.</p> <p>Pierwszy krok analizy: przestrzenią probabilistyczną jest zbiór danych przyjmowanych przez algorytm, zatem koniecznym jest określenie tych danych (precyzyjne ustalenie modelu probabilistycznego) oraz rozkładu prawdopodobieństwa.</p> <p>Uniwersalizm: Przeprowadzona analiza oparta jest o dane algorytmu dowolnego rozmiaru i różnych danych zatem koszt działania będzie zależał od tego rozmiaru danych i konkretnego układu tych danych.</p>
<p>n - natural number, data size d - data layout, for a given algorithm there may be many d - data layout size The data space for a fixed n is a set :</p>	<p>n - liczba naturalna, rozmiar danych d - układ danych, dla danego algorytmu może istnieć wiele d - rozmiar układu danych Przestrzeń danych dla ustalonego n to zbiór:</p>

$$D_n = \{d : |d| = n\}$$

Probability of a data system d of size n in D_n :	Prawdopodobieństwo wystąpienia układu danych d rozmiaru n w :
---	---

$$\sum_{D_n} P_{D_n} = 1$$

<p>Cost of an algorithm: if given a one-element set then the cost will be a function of one variable n, since it depends only on n. However, if this set is multi-element and given a finite probability P_{D_n} distribution then the cost for a given n is a random variable defined for D_n.</p>	<p>Koszt algorytmu: jeżeli dany jest zbiór jednoelementowy to koszt będzie funkcją jednej zmiennej n, ponieważ zależy on tylko do n. Jeżeli jednak ten zbiór jest wieloelementowy i ma dany określony rozkład prawdopodobieństwa P_{D_n} wtedy koszt dla danego n jest zmienną losową określoną dla D_n.</p>
---	--

<p>The cost is determined by the number of steps performed by the algorithm for a given d, so for a fixed n we work with a random variable T_n^{-1} that takes integer values greater than zero (positive). Thus, we have $T_n \geq 0$ for $d =n$.</p>	<p>Koszt jest określony przez liczbę kroków wykonanych przez algorytm dla danego d, zatem dla ustalonego n pracujemy ze zmienną losową T_n która przyjmuje wartości całkowite większe od zera (dodatnie). Zatem mamy $T_n \geq 0$ dla $d =n$.</p>
<p>The study of the probability of an algorithm boils down to the study of the characteristic quantities of a variable T_n, namely:</p> <ul style="list-style-type: none"> - expected value - variance - standard deviation <p>When analyzing an algorithm, one usually works with a large number of random variables indexed by size n.</p> <p>At a fixed value n characteristic values are numbers T_n from which it is difficult to extract information about the analyzed algorithm and its characteristics. In contrast, when working with algorithms, one looks for "how the algorithm behaves as the size of the data increases n, that is, its asymptotic behavior.</p>	<p>Badanie prawdopodobieństwa algorytmu sprowadza się do badania charakterystycznych wielkości zmiennej T_n czyli:</p> <ul style="list-style-type: none"> • wartości oczekiwanej • wariancji • odchylenia standardowego <p>Analizując algorytm pracuje się najczęściej z dużą ilością zmiennych losowych indeksowanych rozmiarem n.</p> <p>Przy ustalonej wartości n charakterystyczne wartości T_n to są liczby, z których trudno jest pozyskać informacje na temat analizowanego algorytmu i jego cech. Natomiast pracując z algorytmami szuka się „jak algorytm zachowuje się wraz ze wzrostem rozmiaru danych n czyli jego asymptotyczne zachowanie.</p>
<p>$(T_n)_\theta$ - expected value² $(T_n)_V$ - variance $(T_n)_\Gamma$ - standard deviation</p>	<p>$(T_n)_\theta$ - wartość oczekiwana³ $(T_n)_V$ - wariancja $(T_n)_\Gamma$ - odchylenie standardowe</p>
<p>P_{nj} - the probability T_n of it taking the value of $j \geq 0$</p> <p>$f P_n(p)$ unction that builds for T_n a series of the form:</p>	<p>P_{nj} - prawdopodobieństwo tego, że T_n przyjmie wartość $j \geq 0$</p> <p>$P_n(p)$ funkcja która buduje dla T_n szereg postaci:</p>

$$P_n(p) = \sum_{j \geq 0} P_{nj} p^j$$

MATHEMATICAL THEOREM 1:	TWIERDZENIE MATEMATYCZNE 1:
--------------------------------	------------------------------------

$$\sum_{j \geq 0} P_{nj} = 1 \quad \rightarrow \quad P_n(1) = 1$$

$$P_n(p) / \frac{d}{dt} \quad \rightarrow \quad P'_n(p) = \sum_{j \geq 0} j P_{nj} p^{j-1}$$

$$(T_n)_\theta = \sum_{j \geq 0} j P_{nj} \quad \rightarrow \quad (T_n)_\theta = P'_n(1)$$

$$P_n(p) / \frac{d^2}{d^2 t} \quad \rightarrow \quad P''_n(p) = \sum_{j \geq 0} j(j-1) P_{nj} p^{j-2}$$

- 1 Random discrete variable / Losowa zmienna dyskretna.
- 2 Random variable T_n and are functions of the natural variable n , the asymptotic properties of these functions are studied.
- 3 Zmiennej losowej T_n i są one funkcjami zmiennej naturalnej n , badane są własności asymptotyczne tych funkcji.

Thus, from the variation we get:

Z wariacji otrzymujemy zatem:

$$\begin{aligned}
 (T_n)_V &= \sum_{j \geq 0} (j - (T_n)_\Theta)^2 P_{nj} = \sum_{j \geq 0} (j^2 - 2j(T_n)_\Theta + (T_n)_\Theta^2) P_{nj} = \sum_{j \geq 0} (j^2 - 2jP'_n(1) + P'_n(1)^2) P_{nj} \\
 (T_n)_V &= \sum_{j \geq 0} (j^2 P_{nj} - 2jP'_n(1)P_{nj} + P'_n(1)^2 P_{nj}) = \sum_{j \geq 0} j^2 P_{nj} - 2P'_n(1) \sum_{j \geq 0} jP_{nj} + P'_n(1)^2 \sum_{j \geq 0} P_{nj} \\
 (T_n)_V &= \sum_{j \geq 0} j^2 P_{nj} - 2P'_n(1) \sum_{j \geq 0} jP_{nj} + P'_n(1)^2 \sum_{j \geq 0} P_{nj} \quad , \text{ by further transforming, przekształcając dalej:} \\
 (T_n)_V &= \sum_{j \geq 0} j(j-1)P_{nj} + \sum_{j \geq 0} jP_{nj} - P'_n(1)^2 \rightarrow \text{we get , otrzymujemy:}
 \end{aligned}$$

$$(T_n)_V = P''_n(1) + P'_n(1) - (P'_n(1))^2$$