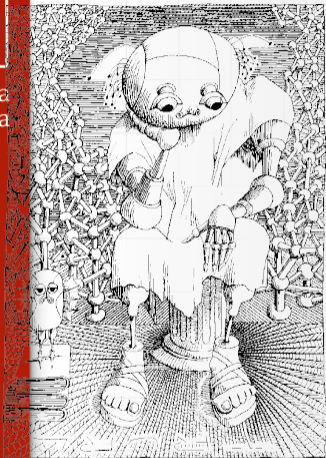




Politechnika
Wroclawska



Technologie Informacyjne – o co chodzi?

wer. 25 z drobnymi modyfikacjami!

Wojciech Myszka

2022-10-17 08:08:59 +0200



HR EXCELLENCE IN RESEARCH



Część I

Zamiast wstępu

Zamiast wstępu I

1. Podstawowa forma konsultacji: **stacjonarne** (Możliwe są konsultacje zdalne — Zoom — po uprzednim uzgodnieniu terminu).
2. Ja urzęduję w pokoju 120 budynku B1. W przypadku bardzo pilnej (to znaczy poza terminami konsultacji) konieczności kontaktu osobistego należy wcześniej umówić się za pomocą e-mail: wojciech.myszka@pwr.edu.pl
3. Do kontaktów z państwem będę używał systemu USOS i poczty elektronicznej.
4. Obecność na wykładach dla pierwszego semestru studiów są obowiązkowa, stąd na zajęciach będzie pojawiała się lista obecności...
5. Zaliczenie na podstawie quizu (ostatnie zajęcia) i innych quizów (na e-portalu)
6. Materiały pomocnicze są tu:
 - 6.1 <https://kmim.wm.pwr.edu.pl/myszka/dydaktyka/technologie-informacyjne/>

Zamiast wstępu II

6.2 Na e-portal <https://eportal.pwr.edu.pl/>

- ▶ kurs W10RAP-SI0072W # W # Technologie informacyjne - Wykład

Informacje, właściwie, dla Rodziców I

(Choć zdaję sobie sprawę, że Państwo są pełnoletni!)

1. Jest takie stare indiańskie przysłowie:
Można konia doprowadzić do wodopoju, ale nie można go zmusić do picia.
2. To na Was leży obowiązek **uczenia się**. Nauczyciele spełniają tylko rolę pomocniczą.
3. Dawajcie mi znać, jeżeli sprawy, o których mówię są dla was wszystkich **absolutnie oczywiste**. Podobnie, jeżeli coś będzie **niezrozumiałe**.
4. Pamiętajcie jednak o tym, że moim zadaniem jest w jakimś sensie **ustandaryzowanie** waszej wiedzy o komputerach i informatyce.



Część II

Informacje podstawowe

Program zajęć I

1. Program. Wymagania. Historia rozwoju systemów komputerowych.
2. Elementy systemu komputerowego. Czynniki wpływające na wydajność.
3. Instrukcje wewnętrzne, logika binarna, podstawowe operacje arytmetyczne.
4. Oprogramowanie: podstawowe, systemowe, użytkowe.
5. Systemy operacyjne: historia rozwoju, podstawowe pojęcia.
6. Algorytmy. Podstawowe konstrukcje algorytmiczne (przeгляд, podział zadania, programowanie dynamiczne, rekurencja, ...).
7. Złożoność obliczeniowa. „Trudne” zadania.
8. Języki programowania. Historia. Przykłady.
9. Parę słów o prawie i Internecie.
10. Test

1. Praktycznie wszyscy studenci mieli już kontakt z komputerem i znają podstawy obsługi podstawowych aplikacji (edytor tekstu, arkusz kalkulacyjny, przeglądarka internetowa, komunikator).
2. Celem zajęć jest usystematyzowanie i uporządkowanie ich wiedzy oraz używanej terminologii.
3. Do dyspozycji mamy **tylko** audytoryjne sale wykładowe, rzutnik...
4. Wykład przedstawi historię rozwoju sprzętu oraz oprogramowania podstawowego (System Operacyjny).
5. Nieco uwagi skupimy na zagadnieniach funkcjonowania komputera, a zwłaszcza sposobie prowadzenia obliczeń i konsekwencjach z tego płynących.



Technologie Informacyjne II

Założenia

6. Ważnymi tematami wykładu będą pojęcia algorytmu, zadanie algorytmizacji (i automatyzacji) czynności, informacje o programowaniu i językach programowania oraz proste przykłady programowania. Mam nadzieję, że informacje te będą przydatne podczas kolejnego kursu poświęconego programowaniu.
7. Ostatnim elementem wykładu jest blok tematów związanych z sieciami komputerowymi, usługami sieciowymi oraz niebezpieczeństwami (i pokusami) korzystania z komputera w sieci.
8. Wykład kończy test.

9. Nie są wykluczone testy w trakcie wykładu.



Literatura I



Biernat J., *Architektura komputerów*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2005.



Harel D., *Komputery-spółka z o.o.: czego komputery naprawdę nie umieją robić*, Ludzie, Komputery, Informacja, Wydawnictwa Naukowo-Techniczne, Warszawa 2002.



Harel D., Feldman Y., *Rzecz o istocie informatyki: algorytmika*, Klasyka informatyki, Wydawnictwa Naukowo-Techniczne, Warszawa 2001, 2002, 2008.







Knuth D.E., *Sztuka programowania*, Klasyka Informatyki, WNT, Warszawa 2001, iSBN 83-204-2539-5.



Komorowski W., *Krótki kurs architektury i organizacji komputerów*, Mikom, Warszawa 2004.

Literatura II

-  Kopertowska M., *Arkusze kalkulacyjne*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych. Poziom Zaawansowany.
-  Kopertowska M., *Bazy danych*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych moduł 5.
-  Kopertowska M., *Grafika menedżerska i prezentacyjna*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych moduł 6.
-  Kopertowska M., *Przetwarzanie tekstów*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych moduł 3.



Literatura III

-  Kurose J.F., *Sieci komputerowe: od ogółu do szczegółu z internetem w tle*, Wydawnictwo Helion, Gliwice 2006.
-  Lévénéz É., *Computer languages history* 2021, <http://www.levenez.com/lang/>.
-  Lévénéz É., *Unix history* 2021, <http://www.levenez.com/unix/>.
-  Lévénéz É., *Windows history* 2021, <http://www.levenez.com/windows/>.
-  Nowakowski Z., *Użytkowanie komputerów*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych moduł 2.



Literatura IV



Pascal B., *Rozprawy i myśli*, rozdz. Machina Arytmetyczna, Instytut Wydawniczy PAX, Warszawa 1962, 1–13, dostępne jako

<https://kmim.wm.pwr.edu.pl/myszka/wp-content/uploads/sites/2/2020/10/pascalina.pdf>.



Pigott D., *The encyclopedia of computer languages*,

<https://web.archive.org/web/20110408094022/http://hop1.murdoch.edu.au/>,

URL <http://hop1.murdoch.edu.au/> 2006.








Sikorski W., *Podstawy technik informatycznych*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych moduł 1.









Silberschatz A., *Podstawy systemów operacyjnych*, Wydawnictwa Naukowo-Techniczne, Warszawa 2006.

Literatura V





-  Stallings W., *Systemy operacyjne: struktura i zasady budowy*, Wydawnictwo Naukowe PWN, Warszawa 2006.
-  Wirth N., *Algorytmy + struktury danych = programy*, Klasyka informatyki, Wydawnictwa Naukowo-Techniczne 2004.
-  Wojciechowski A., *Usługi w sieciach informatycznych*, Wydawnictwo Naukowe PWN, Mikom, Warszawa 2006, europejski Certyfikat Umiejętności Komputerowych moduł 7.
-  Wróblewski P., *Algorytmy: struktury danych i techniki programowania*, Helion, Gliwice 2010.
-  Wróblewski P., *Algorytmy: struktury danych i techniki programowania: algorytmika nie tylko dla informatyków*, Helion, Gliwice 2003.

Literatura VI

-  Zuber R., *Metody numeryczne i programowanie*, WSziP 1975, fragmenty:
<https://kmim.wm.pwr.edu.pl/myszka/wp-content/uploads/sites/2/2020/10/zuber.pdf>.
-  Gleick J., *Informacja. bit — wszechświat — rewolucja*, Wydawnictwo Znak, Kraków 2012.
-  Patterson J.R.C., *Modern microprocessors – A 90-minute guide!*, URL
<http://www.lighterra.com/papers/modernmicroprocessors/> 2015.
-  Feynman R.P., *Feynman Lectures on Computation*, CRC Press 2018.
-  Kopaliński W., *Słownik wyrazów obcych i zwrotów obcojęzycznych*, xx wyd., Wiedza Powszechna, Warszawa 1990.
-  Lem S., *Cyberiada*, Verba, Chotomów 1991.




Literatura VII

-  Piotr Ciesielski, Jerzy Vetulani, *Błogostawiona mutacja genu FoxP2*, wyborcza.pl, 2011, http://wyborcza.pl/1,76842,9801332,Blogoslawiona_mutacja_genu_FoxP2.html.
-  Vopson M.M., *The mass–energy–information equivalence principle*, AIP Advances, 9, 9, 2019, 095206, URL <https://doi.org/10.1063%2F1.5123794>.
-  Vopson M.M., *The information catastrophe*, AIP Advances, 10, 8, 2020, 085014, URL <https://doi.org/10.1063%2F5.0019941>.
-  Landauer R., *The physical nature of information*, Physics Letters A, 217, 4-5, 1996, 188–193, URL <https://doi.org/10.1016%2F0375-9601%2896%2900453-7>.



Landauer R., *Irreversibility and heat generation in the computing process*, IBM Journal of Research and Development, 5, 3, 1961, 183–191, URL <https://doi.org/10.1147%2Frd.53.0183>.

- ▶ Obecność 
- ▶ Test(y)
- ▶ Zadania domowe...?
- ▶ ...



Materiały do zajęć

[https:](https://kmim.wm.pwr.edu.pl/myszka/dydaktyka/technologie-informacyjne/)

[//kmim.wm.pwr.edu.pl/myszka/dydaktyka/technologie-informacyjne/](https://kmim.wm.pwr.edu.pl/myszka/dydaktyka/technologie-informacyjne/)

Choć w dzisiejszych czasach zamiast zapisywać/zapamiętywać adresy trzeba nauczyć się tak formułować zapytania do wyszukiwarek internetowych, żeby trafić na potrzebną stronę...

Powinno wystarczyć „[wojciech myszka dydaktyka](https://kmim.wm.pwr.edu.pl/myszka/dydaktyka/)”. Być może trzeba będzie dodać również pwr.



Część III

Informacja



Nazwa wykładu to dwa wyrazy (definicje przytaczam, za [28]).

Technologie

technika środki, sposoby i czynności związane z wytwarzaniem dóbr materialnych. || **technikum** (l.mn. technika) średnia szkoła zawodowa, której absolwenci otrzymują tytuł technika. || **TECHNO-** w złożeniach: sztuka; rzemiosło; techniczny; technologiczny; stosowany. || **technokracja** rządy specjalistów technicznych w państwie, w społeczeństwie a. rządy na zasadach ustalonych przez techników. || **technologia** (nauka o metoda(ch) przeróbki i obróbki materiałów (t. mechaniczna, chemiczna). || **Technology Assessment** ang. [wym. teknolodzi esesment] Ocena techniki, proces pozwalający na bardziej świadomy wybór spośród wielu społ., ekon. i środowiskowych możliwości, pojawiających się dzięki postępowi nauki i techniki; przewidywanie skutków drugiego rzutu zastosowania nowych technik. || **technostruktura** pracownicy nowoczesnej korporacji kapitalistycznej, którzy — dzięki swoim kompetencjom — biorą udział w grupowym procesie podejmowania decyzji w przedsiębiorstwie.
— gr. *technikos* 'kunsztowny; artystyczny; biegły' od *techné* 'sztuka; nauka; rzemiosło; biegłość; chytryść'; por. polsi: *psycho-*; *sociotechnika*; *technet*; zob. *-kracja*; *-log-*; *struktura*.



Nazwa wykładu to dwa wyrazy (definicje przytaczam, za [28]).

Technologie

przez techników. || **technologia** (nauka o) metoda(ch) przeróbki i obróbki materiałów (t. *mechaniczna, chemiczna*). ||
Technology Assessment and

Nazwa wykładu to dwa wyrazy (definicje przytaczam, za [28]).

Technologie

przez techników. || **technologia** (nauka o) metoda(ch) przeróbki i obróbki materiałów (t. mechaniczna, chemiczna). || **Technology Assessment and**

Informacyjne

Informacja wiadomość, wieść, nowina, rzecz zakomunikowana, zawiadomienie, komunikat; pouczenie, powiadomienie, zakomunikowanie o czymś; dane; pokój, okienko, stanowisko, gdzie się udziela informacji; (ilość informacji) miara wiedzy o jakimś zdarzeniu, uzyskanej w wyniku przeprowadzenia określonego eksperymentu; mat., cyber. miara braku entropii (nieokreśloności), miara organizacji systemu; (l. genetyczna, dziedzinna) biol. zapis genetycznych właściwości organizmu zawarty w cząsteczkach DNA każdej komórki. || **informacja** teoria — dyscyplina mat. badająca, jaka ilość informacji zawiera się w pewnym zbiorze wiadomości (np. w zdaniu, książce, obrazie TV, zapisie kodowym), a także analizująca procesy przekazywania informacji. || **informatyka** techniki i metody przetwarzania informacji; dyscyplina nauki i techniki zajmująca się org. powstawania i przebiegu informacji, technologią i metodyką jej przekształcania, zwł. za pomocą techniki obliczeniowej; nauka o komputerach; połączenie informacji z automatyką. — łac. informatio "wyobrażenie; wizerunek; pomysł" od informare "formować; wymyśleć; spisać; informować"; zob. in- 3; formacja.



Technologie Informacyjne

Nazwa wykładu to dwa wyrazy (definicje przytaczam, za [28]).

Technologie

przez techników. || **technologia** (nauka o) metoda(ch) przeróbki i obróbki materiałów (t. *mechaniczna, chemiczna*). ||
Technology Assessment ang

Informacyjne

informacja - wiadomość, wieść, nowina, rzecz zakomunikowana, zawiadomienie, komunikat; pouczenie, powiadomienie, zakomunikowanie o czymś; dane; pokój, okienko, stanowisko, gdzie się udziela informacji; (*ilość informacji*) miara wiedzy o jakimś zdarzeniu, uzyskanej w wyniku przeprowadzenia określonego eksperymentu; *mat., cyber.* miara



Technika

technika środki, sposoby i czynności związane z wytwarzaniem dóbr materialnych. || **technikum** (lmn. technika) średnia szkoła zawodowa, której absolwenci otrzymują tytuł technika. || **TECHNO-** w złożeniach: sztuka; rzemiosło; techniczny; technologiczny; stosowany. || technokracja rządy specjalistów technicznych w państwie, w społeczeństwie a. rządy na zasadach ustalonych przez techników. || **technologia** (nauka o) metoda(ch) przeróbki i obróbki materiałów (*t. mechaniczna, chemiczna*). || **Technology Assessment** ang., ocena techniki, — proces pozwalający na bardziej świadomy wybór spośród wielu społ., ekon. i środowiskowych możliwości, pojawiających się dzięki postępowi nauki i techniki; przewidywanie skutków drugiego rzutu zastosowania nowych technik. || **technostruktura** pracownicy nowoczesnej korporacji kapitalistycznej, którzy — dzięki swoim kompetencjom — biorą udział w grupowym procesie podejmowania decyzji w przedsiębiorstwie. — gr. *technikós* ‘kunsztowny; artystyczny; biegły’ od *téchnê* ‘sztuka; nauka; rzemiosło; biegłość; chytryść’; por. poli-: psycho-; socjotechnika; technet; zob. -kracja; -log-; struktura.

Informacja

informacja wiadomość, wieść, nowina, rzecz zakomunikowana, zawiadomienie, komunikat; pouczenie, powiadomienie, zakomunikowanie o czymś; dane; pokój, okienko, stanowisko, gdzie się udziela informacji; (*ilość informacji*) miara wiedzy o jakimś zdarzeniu, uzyskanej w wyniku przeprowadzenia określonego eksperymentu; *mat., cyber.* miara braku entropii (nieokreśloności), miara organizacji systemu; (*i. genetyczna, dziedziczna*) biol. zapis genetycznych właściwości organizmu zawarty w cząsteczkach DNA każdej komórki. ||

informacji teoria — dyscyplina mat. badająca, jaka ilość informacji zawiera się w pewnym zbiorze wiadomości (np. w zdaniu, książce, obrazie TV, zapisie kodowym), a także analizująca procesy przekazywania informacji. || **informatyka** techniki i metody przetwarzania informacji; dyscyplina nauki i techniki zajmująca się org. powstawania i przebiegu informacji, technologią i metodyką jej przekształcania, zwł. za pomocą techniki obliczeniowej; nauka o komputerach; połączenie informacji z automatyką.

— łac. *informatio* 'wyobrażenie wizerunek; pomysł' od *informare* 'formować; wymyślać; spisać; informować'; zob. informacja.



Jak mierzyć szczęście I

Zanim zajmiemy się pomiarem informacji zastanówmy się jak mierzyć takie dziwne rzeczy, jak na przykład szczęście?

Zaczynam ponownie od cytatu.

[...] Wprowadził tedy Kontemplatorowi w brzuch znaczny zegar z połączoną strzałką, który wyskalował w jednostkach szczęśliwości i nazwał je hedonami lub hedami w skrócie. Za jeden hed przyjął tę ilość ekstazy, jakiej się doznaje, gdy przebędzie się cztery mile w bucie z gwoździem wystającym, a potem gwoździeń się usunie. Pomnożył drogę przez czas, podzielił przez zadziorność gwoźdźcia, przed nawias wyprowadził współczynnik pięty zmęczonej i tak mu się udało przelożyć szczęście na układ centymetr–gram–sekunda. Tym się trochę pocieszył. Wpatrując się w poplamiony oliwą fartuch roboczy Trurla, który się krzątał przy nim, Kontemplator, zależnie od kąta nachylenia i ogólnego oświetlenia, doznawał od 11,8 do 18,9 hedów na plamę, łatę i sekundę. Na dobre się uspokoił konstruktor. Obliczył zaraz, że jeden kilohed to tyle, ile starcy doznali, podglądając Zuzannę w kąpieli, że megahed — to radość



Jak mierzyć szczęście II

skazańca w porę odciętego od stryczka, a widząc, jak wszystko daje się doskonale wyliczyć, postać zaraz jedną z pośledniejszych machin laboratoryjnych po Klapaucjusza.[...]

Cytat zaczerpnięty z Cyberiady Stanisława Lema [29, (opowiadanie Kobyszcze)].

KoBySzczę to skrót od **K**ontemplator **B**ytu **S**zczęśny.



Jak zmierzyć ilość informacji? I

- ▶ Zaczniemy nie od informacji, ale od jej braku.
Wyobraźmy sobie taką sytuację:
 - ▶ Mamy osobę oczekującą na pewną, bardzo dla niej ważną, wiadomość.
 - ▶ Mamy do czynienia z sytuacją, że wiadomość jest bardzo prosta: możliwe są tylko dwie wartości komunikatu: A i B¹.

Kiedy niepewność co do otrzymanej odpowiedzi jest największa?

- ▶ Jeżeli jedna z wiadomości jest bardziej prawdopodobna (na przykład 99% szans, że otrzymamy wiadomość B) — nie ma wielkiej niepewności (oczywiście możemy się czasami bardzo ucieszyć gdy wypadnie A) — w zasadzie mamy pewność, że będzie to B...
- ▶ Gdy oba komunikaty są jednakowo prawdopodobne — nasza niepewność jest największa.

W teorii informacji **entropia** to właśnie miara naszej niepewności.

¹Na przykład badanie lekarskie pozwalające stwierdzić płeć podczas ciąży. Pomijam celowo sytuacje niezmiernie rzadkie. Bliźniaki — raz na 85 ciąż, trojaczki raz na 85² ciąż, itd.

Entropia

Założmy, że mamy n możliwych komunikatów. Prawdopodobieństwo otrzymania komunikatu o numerze i jest $p(i)$.

Entropia wyraża się wzorem (pochodzącym od Shannona)

$$H = - \sum_{i=1}^n p(i) \log p(i).$$

Minus jest po to, żeby zaznaczyć, że jest to nasza niewiedza (brak informacji)?

Nie. Chodzi o to, żeby wyliczona wartość była dodatnia (logarytm z wartości mniejszych od zera jest ujemny).

W zależności od tego jaka jest podstawa logarytmu użytego we wzorach, jednostką jest:

- ▶ **bit** albo **shannon** gdy podstawa to 2,
- ▶ **nat** gdy podstawa to liczba Eulera e ,
- ▶ **hartley** gdy podstawa to 10.



Entropia

Szczególne przypadki

Jeżeli wszystkie komunikaty są jednakowo prawdopodobne $\left(p(i) = \frac{1}{n}\right)$ wzór się upraszcza

$$H = - \sum_{i=1}^n \frac{1}{n} \log \frac{1}{n} = - \frac{1}{n} \sum_{i=1}^n \log \frac{1}{n} = - \frac{1}{n} \left(n \log \frac{1}{n} \right) = \quad (1)$$
$$- (\log(1) - \log(n)) = \log(n)$$

Dla skupienia uwagi

► $\log_2(10) = 3,32$

► $\log_2(26) = 4,7$



Entropia

Szczególny przypadek: $n = 2$

$$H_b(p) = -p \log_2 p - (1 - p) \log_2(1 - p).$$

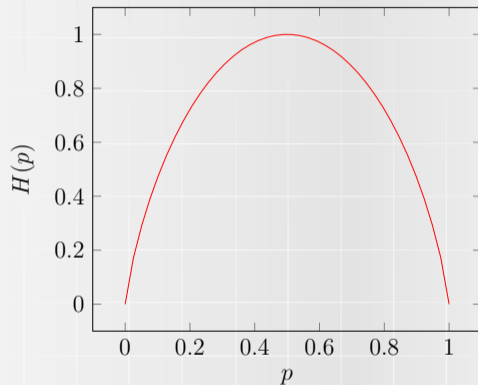
I teraz H wyrażane jest w bitach. Otrzymujemy 1 bit informacji gdy możliwe są dwa jednakowo prawdopodobne komunikaty.

$$H_b = -0.5 \log_2(0.5) - (1 - 0.5) \log_2(1 - 0.5) = 0.5 + 0.5 = 1$$



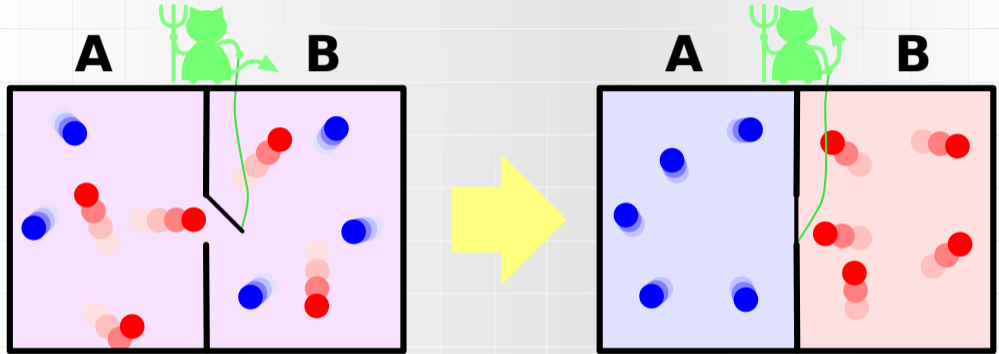
Entropia

$H(p)$



Rysunek: Wykres ilości informacji w zależności od prawdopodobieństwa w przypadku dwu możliwości wyboru

Demon Maxwella



Rysunek: Demon Maxwella w akcji

Zasada Landauera I

W latach 60. **Rolf Landauer** wykazał, że istnieje fizyczna granica minimalnego wydatku energetycznego koniecznego do wykasowania jednego bitu informacji. Wynosi ona

$$k_b \cdot T \cdot \ln 2$$

gdzie k_b jest stałą Boltzmanna ($1,38 \times 10^{-23}$ J/K), a T temperaturą otoczenia. Zatem skasowanie jednego bitu związane jest ze stratą (dyssypacją) energii (wydzieleniem ciepła).

Oznacza to, że w pewnej chwili w układach klasycznych nie będzie można zmniejszyć produkcji ciepła przez element. Przy rosnącym zagęszczeniu elementów i wzroście częstotliwości taktowania układy te będą produkować coraz więcej ciepła.

- ▶ teoria Landauera spotkała się z „różnym” przyjęciem przez naukowców;
- ▶ część z nich rozpoczęła badania eksperymentalne;
- ▶ dziś — raczej — nie podlega to dyskusji.

Od początku I

Niech X będzie zbiorem niezależnych zdarzeń

$$X = \{x_1, x_2, \dots, x_n\}$$

o rozkładzie prawdopodobieństwa P

$$P = \{p_1, p_2, \dots, p_n\}$$

Każde zdarzenie x_i ma prawdopodobieństwo $p_i \geq 0$ i takie, że $\sum p_i = 1$.

Według Shannona średnia informacja (liczba bitów) zdarzenia z obserwacji zbioru zdarzeń X wynosi:

$$H(X) = - \sum_{j=1}^n p_j \log_2 p_j$$

Od początku II

Jeżeli $p_j = 1/n$, wówczas $H(X) = \log_2 n$. Gdy obserwować będziemy N zbiorów zdarzeń (albo N -krotnie zbiór zdarzeń X) to średnia ilość informacji (bitów) wydobyta z tej obserwacji wyniesie $N \cdot H(X)$.

To już wiemy!

Liczba możliwych „stanów” (liczność zbioru zdarzeń)

$$\Omega = 2^{N \cdot H(X)}$$

pozwała to na wprowadzenia pojęcia entropii (w rozumieniu Boltzmann) tego systemu

$$S = k_b \cdot \ln(\Omega) = k_b \cdot \ln(2^{N \cdot H(X)}) = N \cdot k_b \cdot H(X) \cdot \ln(2)$$

gdzie $k_b = 1,38064 \times 10^{-23}$ J/K to stała Boltzmann.

Gdy nasz system może być jedynie w dwu możliwych, jednakowo prawdopodobnych stanach, to potrzebujemy jedynie jednego bitu do

zakodowania takiej informacji. Zatem entropia w tym przypadku może być wyliczona jako $S = k_b \cdot \ln(2)$.

Gdy używamy jakiegoś fizycznego systemu (komputera?), który podlega wszystkim ograniczeniom termodynamiki, do generowania informacji **musi** istnieć jakieś zależność pomiędzy, tworzeniem, manipulowaniem, kasowaniem informacji a termodynamiką.

I dalej I

Wyobraźmy sobie izolowane „urządzenie” pracujące jako pamięć o pojemności N bitów. Może ono być w jednym z 2^N (mikro)stanów, i jego początkowa entropia informacyjna wynosi:

$$S_{\text{info}}^i = Nk_b \ln(2)$$

Na **całkowitą** entropię systemu składa się oprócz entropii *informacyjnej* również entropia *fizyczna* (S_{phys}^i) związana z „urządzeniem” jako pewnym bytem fizycznym.

Jeżeli teraz wykonamy operację **nieodwracalnego wykasowania** informacji, to znaczy zmiany stanu każdego bitu z wartości 0 lub 1 do wartości **nieokreślonej** (jakaś wartość nie będąca ani zerem ani jedynką). Coś takiego jest łatwe do wyobrażenia sobie w przypadku, na przykład, pamięci magnetycznych.

Skoro operacja kasowania **zmniejsza** entropię informacyjną (likwidacja pojedynczego bitu zmniejsza naszą niepewność!) przeprowadzając system ze stanu S_{info}^i przez skasowanie N bitów, do stanu S_{info}^f i

I dalej II

$$\Delta S_{\text{info}} = S_{\text{info}}^f - S_{\text{info}}^i = -Nk_b \ln(2)$$

Jeżeli teraz uwzględnimy entropię całego systemu (która musi spełniać prawa fizyki i nie może w czasie maleć) to

$$\Delta S_{\text{tot}} = \Delta S_{\text{phys}} + \Delta S_{\text{info}} \geq 0$$

Zatem zmniejszenie entropii informacyjnej **musi** wiązać się ze zwiększeniem entropii części „fizycznej” poprzez zwiększenie temperatury; zatem $\Delta Q/T = \Delta S_{\text{phys}} \leq Nk_b \cdot \ln(2)$.



Rozszerzona zasada Landauera I

Skoro kasowanie informacji związane jest z wydzielaniem ciepła, to jej **tworzenie** wymaga wykonania jakiejś pracy $W \geq k_b T \ln(2)$ na każdy bit tworzonej informacji.

Jeżeli już stworzymy bit informacji, (to jeżeli nic straszego się nie dzieje) pozostanie on w tym stanie bez końca bez straty energii.

Nożna to wytłumaczyć zakładając, że po stworzeniu bitu, otrzymuje on pewną masę m_{bit} . Korzystając z zasady równoważności energii i masy, można wyliczyć masę bitu:

$$m_{\text{bit}} = \frac{k_b T \ln(2)}{c^2}$$

c , to (oczywiście) szybkość światła, a T , to temperatura, w której przechowywany jest bit informacji...

Daleko idącymi konsekwencjami tego założenia będzie zastąpienie zasady równoważności energii i masy zasadą równoważności energii, masy i informacji.



Rozszerzona zasada Landauera II

Skoro we wzorze występuje temperatura (T) to w zerze absolutnym nie może istnieć żadna informacja. Natomiast w temperaturze pokojowej ($T = 300K$) szacunkowa masa jednego bitu wyniesie $\sim 3,19 \times 10^{-38}$ kg. Powyższe na podstawie [31, 34, 33, 32].

Nadmiarowość informacji

Wróćmy do naszego przykładu. Oto dziadek czeka na informację o płci dziecka.

- ▶ W pewnym uproszczeniu możemy przyjąć, że możliwe są tylko dwie odpowiedzi.
- ▶ Ale komunikat (E-mail, SMS) będzie zawierał informację zwykle zapisaną tak: **chłopiec** albo **dziewczynka**.
- ▶ Do jej przestania zostanie zużyte znacznie więcej bitów

Jeżeli założyć, że każda litera kodowana jest na ośmiu bitach (co nie musi być prawdą) **dziewczynka** to 88 bitów, a **chłopiec** — 56.

A zwyczajowy komunikat może być jeszcze dłuższy.

Czemu tak marnotrawimy przepustowość nośnika informacji?

Nadmiarowość informacji

Wróćmy do naszego przykładu. Oto dziadek czeka na informację o płci dziecka.

- ▶ W pewnym uproszczeniu możemy przyjąć, że możliwe są tylko dwie odpowiedzi.
- ▶ Ale komunikat (E-mail, SMS) będzie zawierał informację zwykle zapisaną tak: **chłopiec** albo **dziewczynka**.
- ▶ Do jej przestania zostanie zużyte znacznie więcej bitów

Jeżeli założyć, że każda litera kodowana jest na ośmiu bitach (co nie musi być prawdą) **dziewczynka** to 88 bitów, a **chłopiec** — 56.

A zwyczajowy komunikat może być jeszcze dłuższy.

Czemu tak marnotrawimy przepustowość nośnika informacji?

To jest zadanie domowe!

Nadmiarowość wyglądu znaków

Ald IIIa K0ld

Kot ma Ale

Rysunek: Nadmiarowość kształtu liter

Nadmiarowość zawarta w tekście

Czy da się przeczytać ten tekst?

„Nie ma zncazeina kojnoleść ltier skadającytch się na dane sołwo. Wanże jset tyklo to, żbey pirewsza i osatntia lteria były na soiwm mijsecu; ptzosaote mgoą być w cakowtitym niedztaie, a mimo to nadal nie ponwinśimy mieć prombleów z pczerzytaenim tego tesktu”.

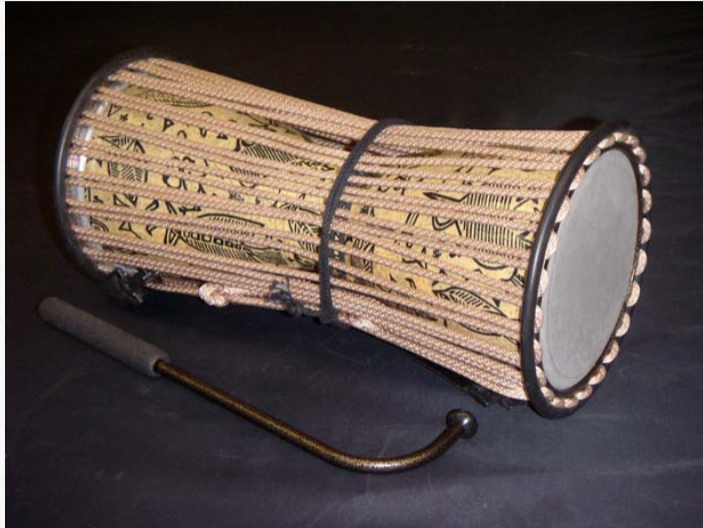
Spowodowane jest to tym, że (na ogół) czytając nie literujemy, tylko patrzymy na wyraz jak na pewną całość...

Za [30] (ale podobne teksty można również znaleźć również dla języka angielskiego).



Talking Drums

Gadające bębny



Risunek: "Gadające" bębny

Podczas eksploracji sub-sacharyjskiej Afryki, podróżnicy dowiedzieli się, że miejscowa ludność może komunikować się na odległość z użyciem bębnów. Nie bardzo rozumieli jak to się odbywa, choć wiedzieli o użyciu, na przykład, trąbki do przekazywania prostych komunikatów na polu bitwy.

Mniej-więcej w tym czasie Samuel Morse pracował nad swoim telegrafem, który podobnie jak bęben był w stanie przekazywać tylko dwa rodzaje sygnałów., Natomiast komunikacja za pomocą bębnów nie odbywała się z użyciem kodów (jak w przypadku kodu Morse'a, o czym będzie nieco później).

Bębniący starali się raczej odtwarzać za pomocą bębnów *fonemy* języka mówionego. Jest oczywiste, że można było w ten sposób odzwierciedlić jedynie bardzo ograniczony zasób słów. Co więcej zwracać uwagę trzeba było na intonację, i akcenty. Ten sam przekaz mógł znaczyć:

- ▶ alambaka boli [_ _ -- _ _ _] „obserwował brzeg rzeki”
- ▶ alambaka boli [-- -- _ _] „ugotował swoją teściową”

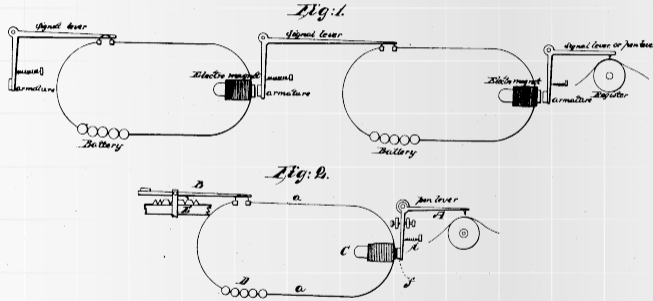
- ▶ Oprócz rytmu i akcentów, do poprawnej komunikacji potrzebne było jeszcze sporo nadmiarowych informacji. Każde słowo uzupełniane było krótką frazą precyzującą o co chodzi. Dwa uderzenia w wyżej brzmiący koniec bębna [—] mogły oznaczać słowo sango (ojciec), songe (księżyc), koko (ptactwo), fele (gatunek ryb).
 - ▶ Bębniąc o księżycu było coś takiego: *songe li tange li manga* („księżyc, który spogląda na ziemię”);
 - ▶ gdy chodziło o ptactwo to było coś takiego *koko alongo la bokiokio* („drób, małe ptaki, które gdaczą”).
- ▶ Każde niejednoznaczne słowo pojawiało się w chmurze innych, powiązanych jakoś słów, które były dosyć automatycznie — po zidentyfikowaniu znaczenia — odrzucane.

Telegraf Morse'a

- ▶ Przesyłanie informacji za pomocą dwu sygnałów (w sposób „binarny”).
- ▶ Medium: przewód elektryczny
- ▶ Informacje zapisywane na taśmie w postaci dłuższych i krótszych kresek (impulsów prądu).
- ▶ Pierwotnie miały być przesyłane wyłącznie cyfry i liczby.
- ▶ Słowa jako pozycje ze słownika — ale był to bardzo nieefektywny sposób.

Telegraf Morse'a

Example, 11.



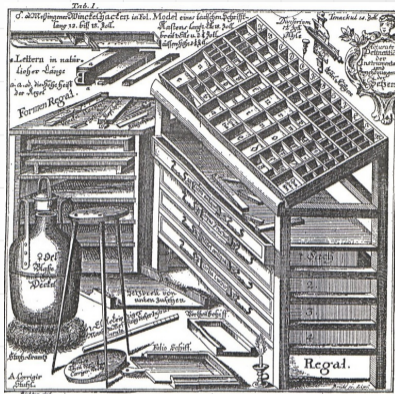
No. 117.

S. F. B. MORSE.
ELECTROMAGNETIC TELEGRAPH.

REISSUED JUNE 13, 1848.

4 SEETS-SEET 4.

Rysunek: Fragment patentu Morse'a numer 1647 z roku 1848 przedstawiający koncepcję telegrafu, a w szczególności sposób zwiększania zakresu i mechanizm drukowania komunikatu



Rysunek: Kaszta drukarska: w czasach Gutenberga i bardziej współczesna



Częstość występowania liter

W języku angielskim (pierwsza liczba to ilość czcionek w kaszcie, druga częstość)

- ▶ „e” — 12 000; 12,7%
- ▶ „t” — 9 000; 9,1%
- ▶ ...
- ▶ „z” — (tylko) 200; 0,01%

Dla porównania, w języku polskim (tylko częstość):

- ▶ „a” — 8,91%
- ▶ „i” — 8,21%
- ▶ „o” — 7,75%
- ▶ „e” — 7,66%
- ▶ ...
- ▶ „x” — 0,02%

W efekcie zdecydowano, że najczęściej występujące litery będą miały najkrótsze kody.



Kod Morse'a

A	·—	J	·— — —	S	...	1	·— — — —
B	—...	K	— · —	T	—	2	·· — — —
C	— · · ·	L	·· · ·	U	·· —	3	·· · · —
D	— · ·	M	— —	V	·· · —	4	·· · · —
E	·	N	— ·	W	· — —	5	·· · · ·
F	·· · ·	O	— — —	X	— · · —	6	— · · · ·
G	— · ·	P	·· · · ·	Y	— · — —	7	— · · · ·
H	·· · ·	Q	— · · —	Z	— · · ·	8	— · — · ·
I	··	R	·· ·	0	— — — — —	9	— · — · · ·

- ▶ W czasach gdy nie istniała jeszcze kultura pisana, przekaz tradycji (informacji) odbywał się w sposób „mówiony” (czy może raczej śpiewany albo deklamowany): Każdy przekaz uzupełniany był dodatkowymi słowami, aby komunikat miał swój rytm albo, żeby łatwiej było go deklamować lub śpiewać.
- ▶ Takie postępowanie pozwalało też utrzymać przekaz w niezmienionej formie (każda zmiana powodowała zmianę rytmu, albo psuła rymy).

Czym jest kompresja danych?

Utworzyłem plik o długości 22000 bajtów składający się z tysiąca identycznych linii o zawartości

chłopczyk dziewczynka

Następnie użyłem kilku popularnych programów kompresujących aby go skompresować

- ▶ gzip — 121 bajtów
- ▶ xz — 148 bajtów
- ▶ rar — 129 bajtów
- ▶ 7zip — 193 bajty

Czemu tak jest?



Czym jest kompresja danych (cd)?

Tym razem plik o tej samej długości (22000) ale wypełniony losowymi literami.

Tym razem

- ▶ gzip — 18264 bajtów
- ▶ xz — 18500 bajtów
- ▶ rar — 18494 bajtów
- ▶ 7zip — 18549 bajty

Czym jest kompresja danych (cd)?

Tym razem plik o tej samej długości (22000) ale wypełniony losowymi literami.

Tym razem

- ▶ gzip — 18264 bajtów
- ▶ xz — 18500 bajtów
- ▶ rar — 18494 bajtów
- ▶ 7zip — 18549 bajty

W losowym tekście nie ma (prawie żadnej) nadmiarowości. Stąd nie ma co kompresować...

Kompresja

Zaletą użytych programów jest to, że kompresja jest **odwracalna**.
Skompresowany plik można „rozpakować” odzyskując pierwotną zawartość.
Jest to kompresja bezstratna.

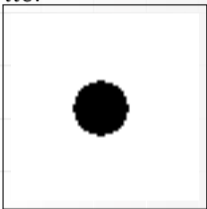
- ▶ Jej istota (w dużym uproszczeniu) polega na tym, że wyszukuje się wszystkie powtarzające się fragmenty (słowa) pliku i zastępuje je „krótszymi” słowami.
- ▶ „chłopczyk” staje się symbolem α .
- ▶ „dziewczynka” — β .
- ▶ Odstęp między wyrazami i znak nowej linii pozostają bez zmian.
- ▶ Zamiast 22 znaków mamy 4: $\alpha\beta<nl>$
- ▶ Sekwencje takich samych słów zastępujemy 1 słowem i informacją o liczbie powtórzeń

Kompresja stratna

Oto prosty obrazek: kropka na białym tle.

Kompresja stratna

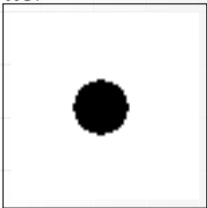
Oto prosty obrazek: kropka na białym tle.





Kompresja stratna

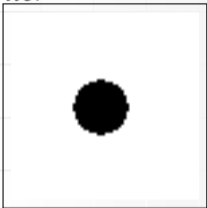
Oto prosty obrazek: kropka na białym tle.



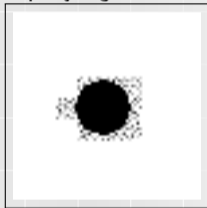
Skonwertowałem go do formatu jpg, a następnie wszystkie punkty, które nie są idealnie białe „podkręciłem” żeby lepiej było widać.

Kompresja stratna

Oto prosty obrazek: kropka na białym tle.

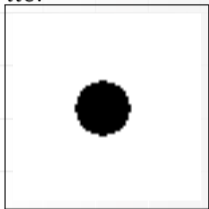


Skonwertowałem go do formatu jpg, a następnie wszystkie punkty, które nie są idealnie białe „podkręciłem” żeby lepiej było widać.

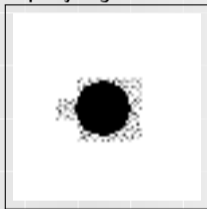


Kompresja stratna

Oto prosty obrazek: kropka na białym tle.



Skonwertowałem go do formatu jpg, a następnie wszystkie punkty, które nie są idealnie białe „podkręciłem” żeby lepiej było widać.



Ilość „błędów” zależy od stopnia kompresji. Podobnie wielkość skompresowanego pliku. Większy stopień kompresji — mniejszy plik i więcej błędów.



Przykład rzeczywisty

Na kolejnych dwu slajdach przykład zdjęcia o rozmiarach 2600×1733 piksele, zawierającego 4,5 miliona pikseli. Każdy piksel to trzy bajty.

Pierwsze zdjęcie to „oryginał” o długości 2384010 bajtów (2,3 M).

Drugie to zdjęcie przetworzone, zajmuje tylko 94980 bajtów (96 K).

Na obu zdjęciach porównać można zniekształcenia wprowadzone podczas konwersji. Hint: użyj zoomu!

Można też obejrzeć wersję interaktywną na [stronie WWW](#).



Przykład rzeźbywisty





Przykład rzeczywisty





Kompresja stratna — bezstratna

▶ grafika

- ▶ JPG, JPEG
- ▶ DjVu
- ▶ Wavelet (falki)

▶ wideo

- ▶ MPEG [124]
- ▶ H.264
- ▶ Ogg Theora

▶ muzyka

- ▶ MP3
- ▶ AAC
- ▶ WMA

▶ ogólne

- ▶ RLE (BMP)
- ▶ DEFLATE (ZIP, PNG)

▶ grafika

- ▶ PNG
- ▶ GIF

▶ wideo

- ▶ QuickTime
- ▶ H.264 loosess

▶ muzyka

- ▶ FLAC
- ▶ ALAC (Apple Loosles)

Celowa informacja nadmiarowa

- ▶ Suma kontrolna
 - ▶ bit parzystości
 - ▶ suma kontrolna
 - ▶ CRC (Cyclic Redundancy Check — cykliczny kod nadmiarowy)
- ▶ Suma MD5

Celowa informacja nadmiarowa

- ▶ Suma kontrolna
 - ▶ bit parzystości
 - ▶ suma kontrolna
 - ▶ CRC (Cyclic Redundancy Check — cykliczny kod nadmiarowy)
- ▶ Suma MD5

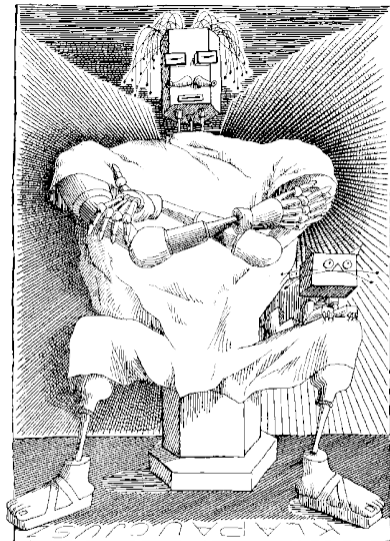
Zadanie domowe — przeczytać o różnych metodach sprawdzania integralności danych.



Rysunek: Po lewej stronie „standardowa” taśma dalekopisu, po prawej taśma ośmio-ścieżkowa, używana przez komputery



Rysunek: Tu można sprawdzić czy dziurki na C13 pochodzą od taśmy — powinna być parzysta liczba w każdym rzędku czy raczej od sera



Rysunek: Po stronie lewej Trurl, po prawej Klapaucjusz według Daniela Mroza, genialnego ilustratora, w tym i dzieł Lema

Prezentacja złożona z użyciem systemu $\text{\LaTeX} 2_{\epsilon}$ i pakietów **beamer** i **make4ht** wykorzystuje oficjalny szablon prezentacji Politechniki Wrocławskiej

<https://pwr.edu.pl/uczelnia/informacje-ogolne/materialy-promocyjne/logotyp>.

Wykorzystano krój pisma **lwona**.

W prezentacji wykorzystano podobizny Trurla i Klapaucjusza pobrane ze strony <http://pl.cyberiada.wikia.com/wiki/Cyberiada> autorstwa Daniela Mroza.

Warto również zerknąć tu: <http://www.google.com/logos/lem/>