1, 2, 3, кодирај! – Научни осврт – Организација и трансформација информација

Организација информација

Иформације представљамо сликом, текстом, тонским записом, видеом...На диску компјутера можемо стокирати један једини текст, једну слику, један тонски запис или један видео. Међутим, ми врло често желимо да на диску стокирамо више оваквих објеката. У том случају морамо да применимо други принцип који ће нам омогућити не само да представимо информације него и да их организујемо.

На диску се налази простор у ком стокирамо велики број « 0 или 1 », обично је то осам хиљада милијарди за диск од тераоктета. За стокирање више текстова, слика,... почињемо да делимо простор диска на више зона које називмао „фолдери“ а који нам омогућују да у сваки од њих стокирамао текст, слику, звучни запис, видео,...Када они постану бројни онда и њих групишемо у облику досијеа... тј., правим структуру дрвета.
Организација информација на великој скали се може урадити и коришћењем **веза хипертекста**. Таква веза омогућује да се у првом фајлу назначе имена позиције на једном или другом диску, а организација инфорамција вазама хипертекста примењена на мрежу дискова и компјутера је идеја Weba.

Организација информација помоћу фајлова, фолдера уз коришћење хипертекста није довољна када се стокира велики број информација па се примењују други типови органзације. Прва је за структуиране информације у оквиру неке **базе података**. На пример, адресар је мала база податак која се састоји од 5 података а састоји се од имена, презимена, адресе, телефонског броја и електронске адресе. Податек у овој бази можемо користити тражењем неког од наведених података.

Мање структуиране податке можемо оставити на гомили на једном или више дискова и претраживачем налазити оно што нас интересује. Овај начин организације захтева мање рада јер се не формира одговарајућа структура али је и мањег квалитета јер захтевана информација може да се појави у више текстова.

Трансформација и манипулација информација

Информације је, захваљујући репрезентацији и трансформацији, могуће пребацивати из места у место (тј. могуће их је усмеравати кроз простор), из архива до оног ко их тражи касније (тј. усмеравати их у времену), а поготову их је могуће трансформисати. Бројни алгоритми омогућују трансформацију специфичних података. На пример, алгоритам множења се примењује само на бројеве. Међутим, три типа алгоритама се примењују на било које податке: алгоритми компресије, корекције и шифрирања.

Компресија

Видели смо да се пиксел розе бомбон представља са ***11111001100001010011110*** као и да се монохроматична слика од милион пиксела овог типа боје добија постављањем 1 милион пута блокова од 24 « 0 или 1 ». Уместо да се неки текст, као у овом случају, представи понављањем двадесет четири милиона знакова можемо користити концизнију форму којом се наглашава да ће блок ***11111001100001010011110*** бити поновљен милион пута. Овај процес је познат као компресија!

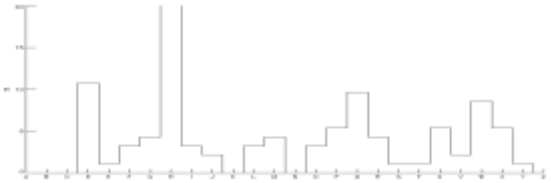
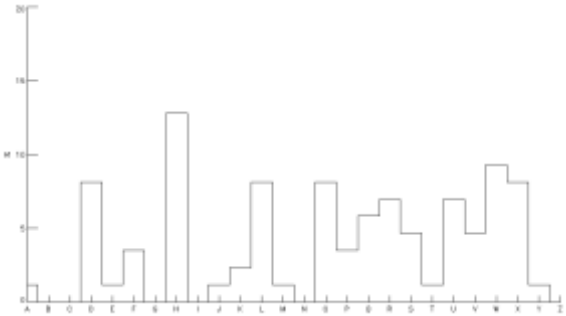
Крекција

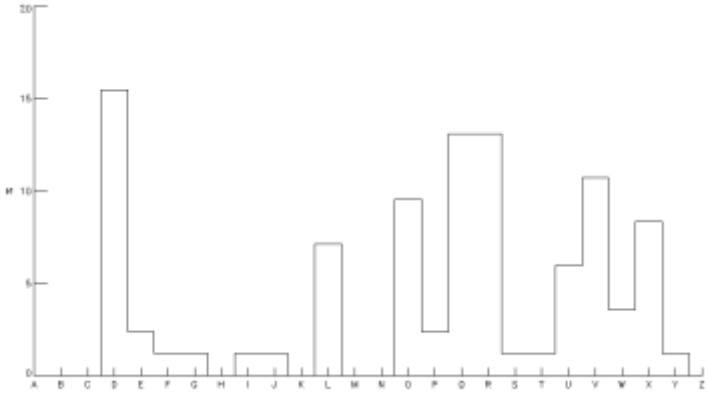
Алгоритми корекције грешака омогућују да се означи, понекад и коригује, грешка у информацији. На пример, грешка која је настала при трансмисији те информација. Једноставан метод за налажење грешке је да се понови три пута свака унитарна информација. На пример, ***101*** је трансформисана у ***111000111***. Па ако је направљена грешка, на пример у трећем триплету: ***111000110***, врло лако се налази али и коригује та грешка, јер пошто је 1 била највише пута поновљена у том триплету онда је сигурно да баш она недостаје. Овај алгоритам је неуспешан када се у истом триплету направе две грешке. Иначе овај начин тражења грешака захтева доста меморије.

Шифрирање

Коначно, алгоритмом шифрирања неки текст постаје читљив смао оном ко поседује његову шифру. Једноставан, али и не толико ефикасан, метод шифрирања је користио ЈУлије Цезар у комуникацији са својим армијама. Састојао се у померању сваког слова азбуке за одређен број места. Тако се изрека "VENI VIDI VICI" , коришћењем његове шифре, која користи померања слова за три места унапред, "YGPL YLPL YLEL". Да би се ова порука дешифрирала потребно је имати кључ те шифре. Међутим, пошто у латиници постоје само 23 слова, онда постоје само двадесет три могућности па је њено дешифровање врло лако и за оног ко нема њен кључ.

Коришћење неких уобичајених анализа може омогућити брзо налажење шифре, Наиме, у неком језику свако слово има познату фреквенцу појављивања (на пример, у француском језику се E – најчешће појављује; затом следе слова A, I, N, O, R, S, T, L , док су на другом крају, слова која се најређе користе, Z, W и K). Тако у неким случајевима појава најчешће коришћеног слова може да допринесе одгонетању шифре, па је потребно користити хистограм за одговарајући језик. У доња 3 примера (приказани су француски текстови које нећемо преводити,п.п.)кључ за шифру је оувек исти (+3, као што је користио Јулије Цезар). Очигледно је, да чак и када се уоче нека мала одступања у хистограмима, ми лако уочавамо исти мотив (исти као у тексту романа « La disparition » de G. Perec који је био написан без E).

Шифрирана порука (кључ+3)	**Аанализа фреквнеце**	**Дешифрирана порука**
MH MHWWH DYHF JUDFH PRQ IHXWUH MH IDLV OHQWHPHQW O DEDQGRQ GX JUDQG PDQWHDX TXL PH FDOIHXWUH HW MH WLUH PRQ HVSDGRQ		Je jette avec grâce mon feutre, Je fais lentement l’abandon Du grand manteau qui me calfeutre, Et je tire mon espadon (Cyrano de Bergerac, E. Rostand)
O HWUH KXPDLQ FURLUD WRXMRXUV TXH SOXV OH URERW SDUDLW KXPDLQ SOXV LO HVW DYDQFH FRPSOHAH HW LQWHOOLJHQW		L'être humain croira toujours que plus le robot paraît humain, plus il est avancé, complexe et intelligent. (Les robots de l’aube, I. Asimov)

Шифрирана порука (кључ+3)	Аанализа фреквнеце	Дешифрирана порука
LO DEDQGRQQD VRQ URPDQ VXU VRQ OLW LO DOOD D VRQ ODYDER LO PRXLOOD XQ JDQW TX LO SDVVD VXU VRQ IURQW VXU VRQ FRX		Il abandonna son roman sur son lit. Il alla à son lavabo; il mouilla un gant qu'il passa sur son front, sur son cou. (La Disparition, G. Perec)

Постоје и други знанто комплекснији методи шифрирања, тако је током Другог светског рата немачко шифрирање Енигма (одгонетнуо A. Turing са својом екипом) која се базирала такође на заменама слова, али се кључ шифрирања правилно мењао у тексту! У историји је било и других начина шифрирања, нрочито у време ратова (на пример, шифрирање ADFGVX тококом Другог светског рата).

Знатно бољи методи шифрирања постоје и користе се свакодневно за заштиту наших кредитних картица и других тајни које желимо да делимо само с ограниченим бројем људи …

<< Le codage binaire | Eclairages | Signal vs information >>