

KONTEXTOVÉ MAPY A JEJICH APLIKACE

Ing. Michal Vašinek, Vedoucí práce: Ing. Jan Platoš PhD.

michal.vasinek@vsb.cz, jan.platos@vsb.cz



MOTIVACE

V roce 1949 definoval Claude Shannon pojem entropie, neboli množství informací obsažených v datech. Tento pojem se stal základem odvětví kompresních algoritmů. Shannonem definovaná entropie popisuje data v jednom určitém stavu a představuje tak statickou veličinu. V této práci se snažím nabídnout odlišný pohled na data a jejich entropii. Data jsou zde považována za popis stavu nějakého dynamického systému, který se může vyvíjet pouze vratnými změnami. Aktuální stav systému se nachází v jednom z konečného počtu dostupných stavů ψ_n a každý stav je charakteristický svou entropií $H(\psi_n)$.

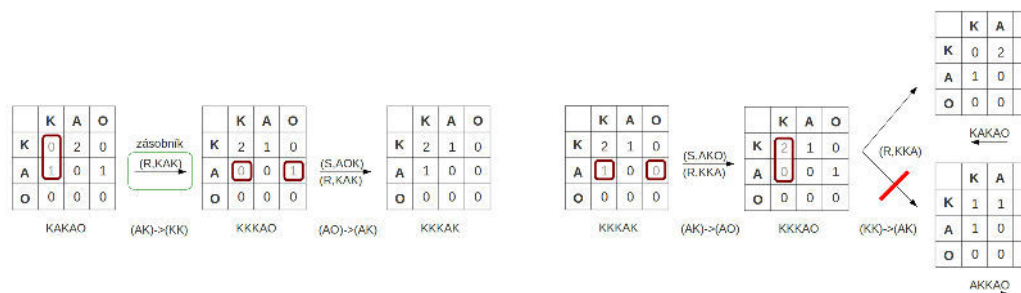
KONTEXTOVÉ MAPY A JEJICH VLASTNOSTI

Kontextová mapa je datová struktura zachycující návaznosti jednotlivých datových elementů a je vhodným nástrojem pro popis dynamických systémů. Datové elementy obsažené v kontextové mapě se vůči sobě nacházejí v několika možných relacích, tzv. relacích provázanosti, ty pak definují množinu dostupných, vratných transformací. V práci představené algoritmy se snaží převádět systém, pomocí transformací, do stavů s nižší entropií. Relací provázanosti lze však také využít pro stanovení podobnosti mezi jednotlivými elementy.

VYUŽITÍ

Vlastnosti kontextových map nacházejí uplatnění především při kompresi a analýze dat. Díky schopnosti kontextových transformací snižovat entropii ve vstupních datech, je transformační metoda vhodná pro předzpracování souborů a následně aplikaci entropických kompresních algoritmů, jako například Huffmanova nebo aritmetického kódování.

PŘÍKLAD KONTEXTOVÉ TRANSFORMACE A K NÍ INVERZNÍ TRANSFORMACE



VLIV TRANSFORMACÍ NA SNÍŽENÍ ENTROPIE A POČTU UZLŮ VE VYBRANÝCH SOUBORECH CALGARY CORPUSU

