Mocnina částečně uspořádané množiny

Ing. Emilie Šeptáková

Katedra informatiky, FEI, VŠB – Technická Univerzita Ostrava, 17. listopadu 15,
708 33, Ostrava–Poruba
Emilie.Septakova @vsb.cz

Abstrakt. V příspěvku popisuji novou metodu pro vyhledávání vzoru ve tvaru stromu ve stromech na základě mocniny částečně uspořádané množiny (posetu). Toto vyhledávání by se dalo implementovat pro vyhledávání na základě podobnosti v XML dokumentech. Tato metoda umožňuje uspořádat výsledky přirozeným zpùsobem. Toto přirozené uspořádání je hlavní výhodou prezentované metody oproti jiným metodám. Další výhodou této metody je, že přirozeným zpùsobem zachycuje strukturu zpracovávaných XML dokumentů.

Klíčová slova: poset, částečně uspořádaná množina, mocnina, Hasseův diagram

1 Úvod - motivace

Jak roste počet dokumentů vytvořených v jazyce XML, roste také zájem o efektivní vyhledávání informací v nich uložených. Data, uložená v XML dokumentu, mají hierarchickou strukturu, která se dá zakreslit jako graf - strom. Je možné vyhledávat informace ve tvaru stromu v jiném stromu stejně jako např. v řetězcích? Cílem je využít mocninu posetu pro vyhledání vzoru ve tvaru stromu a přirozeném uspořádání výsledku hledání.

2 Definice částečně uspořádané množiny

Definice 1. *Podmnožina* $\boldsymbol{S}$ *kartézského součinu X3X se nazývá částečně uspořádaná na neprázdné množině X jestliže :*

1. $\forall x \in X: (x, x) \in \boldsymbol{S}$ (R)
2. $\forall x, y \in X$: *jestliže* $(x, y) \in \boldsymbol{S}$ *a* $(y, x) \in \boldsymbol{S}$, *pak* $x = y$ (AS)
3. $\forall x, y, z \in X$: *jestliže* $(x, y) \in \boldsymbol{S}$ *a* $(y, z) \in \boldsymbol{S}$, *pak* $(x, z) \in \boldsymbol{S}$ (T)

Kde (R) - reflexivnost, (AS) - antisymetrie, (T) - tranzitivnost.
Uspořádanou dvojici $\boldsymbol{S} := (X, \leq)$ nazýváme *částečně uspořádanou množinou* (posetem), jestliže $\leq$ je částečné uspořádání na X. X je nosič posetu, relace $\leq$ se nazývá uspořádání množiny X. Prvky $x, y, z \in X$ nazýváme vrcholy nebo body nebo elementy. O elementech můžeme říct, že jsou porovnatelné nebo neporovnatelné.
Příklad 1: posety:
1. X := {a, b, c, d} a

S := { (a, a), (b, b), (c, c), (d, d), (a, b), (b, c), (b, d), (a, c), (a, d)}
S je částečně uspořádaná množina na X

2. **Antiřetězec** – značíme např. **4**
   X := {a, b, c, d}   **4** := { (a, a), (b, b), (c, c), (d, d)}
3. **Řetězec** - značíme např. C_4
   X := {a, b, c, d} C_4 := { (a, a), (b, b), (c, c), (d, d), (a, b), (b, c), (c, d)}
4. **N poset**
   X := {a, b, c, d}   **N**:= { (a, a), (b, b), (c, c), (d, d), (a, c), (b, c), (b, d)}
   Kardinalita - počet elementů posetu - |N| =4
5. **S** = (X, ≤), X := {1, 2, 3, 4, 6, 12} a ≤ znamená x je dělitelem y
   S := { (1, 1), (2, 2), (3, 3), (4, 4), (6, 6), (12, 12), (1, 2), (1, 3), (2, 4), (2, 6), (4, 12), (6, 12) }, elementy 3 a 4 jsou neporovnatelné

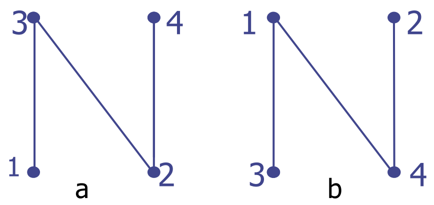
2.1 Jak reprezentovat poset

Poset můžeme reprezentovat několika způsoby. Nejčastější a nejvíc vypovídající je *Hasseův diagram* – bude vysvětleno dále, lze použít i *maticí sousednosti Hasseova diagramu, intervalové uspořádání a semiuspořádání, úhlové a kruhové uspořádání.*

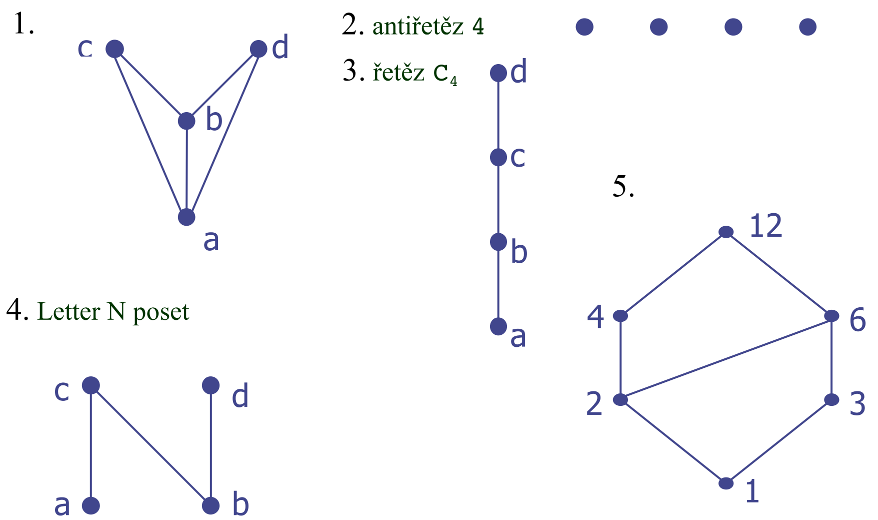
2.2 Hasseův diagram

Hasseův diagram je grafické vyjádření posetu ve tvaru grafu. Elementy posetu jsou zakresleny jako vrcholy grafu. Dva vrcholy spojené hranou jsou porovnatelné. Vrcholy jsou označené posloupností písmen nebo číslic.

Definice 2. *Označení vrcholů posetu L je bijektivní zobrazení L: {1, ..., n} →V(X), kde V(X) je množina všech vrcholů posetu (X, ⊴) o kardinalitě n.*
Definice 3. *Konečný poset (X, ≤, L) je přirozeně označený, jestliže x < y ve V(X), pak $L^{-1}(x) < L^{-1}(y)$ v {1,...,n}, doména L.*
Definice 4. *Každý poset má přirozené označení.*
Lineární rozšíření posetu (X, ≤) je poset (X, ≤*) takové, že pro x ≤ y, pak x ≤* y a poset (X, ≤*) je řetězec.



Obr. 1. Přirozeně označený poset (a), nepřirozeně označený poset (b)



Obr. 2.. Příklady posetů a jejich grafického vyjádření pomocí Hasseova diagramu (viz. příklady posetů na předchozí stránce)

3 Morfismy částečně uspořádaných množin

Definice 5. *Mějme částečně uspořádané množiny* ***S****=(X, ⊴) a* ***P*** *=(Y, ⊴). Zobrazení f z X do Y* ***zachovává uspořádání*** *jestliže ∀x, y ∈ X: x ≤ y, pak f(x) ≤ f(y).*

Výsledkem zobrazení f(X) **řetězce** X, které zachovává uspořádání, je také řetězec.

Definice 6. *Mějme částečně uspořádané množiny* ***S****=(X, ⊴) a* ***P****=(Y, ≤'). Zobrazení f: X do Y, jedna ku jedné, které zachovává uspořádání, je nazýváno* ***izomorfismem****, jestliže f^{-1} také zachovává uspořádání.*
Označujeme X ≅ Y nebo X = Y. Izomorfismus je relace ekvivalence na třídě posetů. **Automorfismus** je izomorfismus částečně uspořádané množiny na sebe. Počet všech automorfismů posetu (X, ≤) označujeme **X!**

Nechť f je zobrazení jedna ku jedné posetu S=(X, ≤) na P=(Y, ≤') zachovávající uspořádání. Pak f je isomorfizmus právě tehdy, když ∀x, y ∈ (X, ≤): x ≤ y právě tehdy, když f(x) ≤' f(y).

3.1 Mocnina posetu

Definice 7. *Mějme dvě částečně uspořádané množiny X a Y. Mějme množinu* $\boldsymbol{Y^X}$ *všech zobrazení zachovávajících uspořádání z X do Y. Definujme binární relaci na*

$\boldsymbol{Y^X}$: $f \le g \quad \forall x \in X$: $f(x) \le g(x)$, kde f, $g \in \boldsymbol{Y^X}$. Pak ($\boldsymbol{Y^X}$, $\le$) je mocnina částečně uspořádané množiny Y na X.

Příklady.

$C_3^{\underline{1}} \cong C_3,\ R^{C2}, C_3^{\ C2,}\ N^{C3}$

3.2 Kardinalita mocniny posetu

Zobrazit mocninu posetu je velmi těžké, z důvodů velkého množství vrcholů již při malém počtu vrcholů vzoru i cíle. Počet vrcholů mocniny posetu (kardinalita) lze spočítat podle následující definice a vzorce.

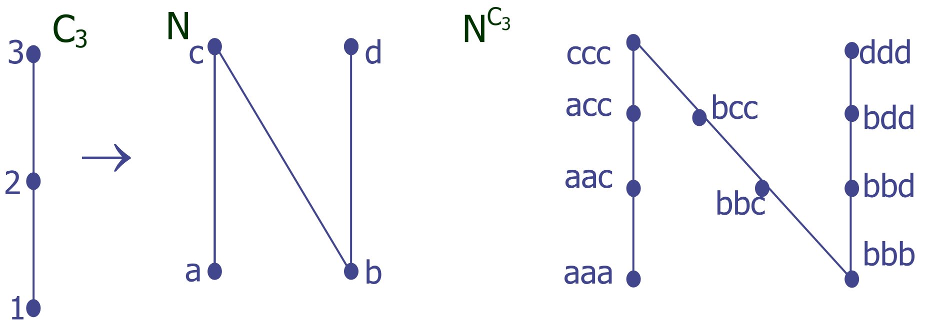
Definice 8. *Každé zobrazení f zachovávající uspořádání lze jednoznačně rozložit $f = f1 \circ f2$, kde $f2\colon X \to Imf$ je zobrazení množiny na množinu zachovávající uspořádání a $f1 : Imf \to Y$ je zobrazení jedna ku jedné zachovávající uspořádání.*

Kardinalitu mocniny posetu vypočteme podle níže uvedeného vzorce.

$$|\,Y^X| = \sum \frac{e(\mathrm{X}, \mathrm{Im}f)\ i(\mathrm{Im}f, \mathrm{Y})}{\mathrm{Im}f\,!} \qquad (1)$$

kde e(X, Imf) je počet různých uspořádání zachovávajících zobrazení na Imf, g: X $\to$ Imf a i(Imf, Y) je počet různých uspořádání zachovávajících zobrazení jedna ku jedné f: Im$f \to$ Y a X! je počet všech automorfismů posetu (X, $\le$). Imf jsou spojité, uspořádání zachovávající obrazy X - plné podposety (full subposet).

Příklad 3.

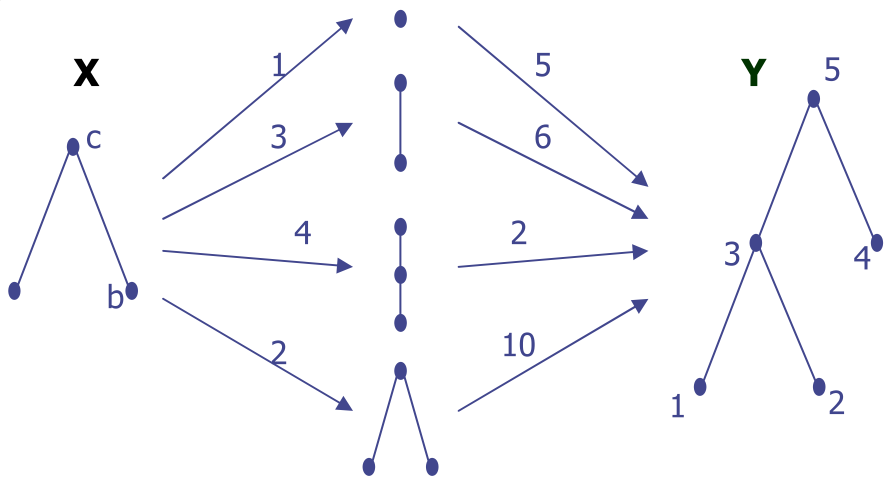


Obr.3 Hasseův diagram mocniny N^{C3}

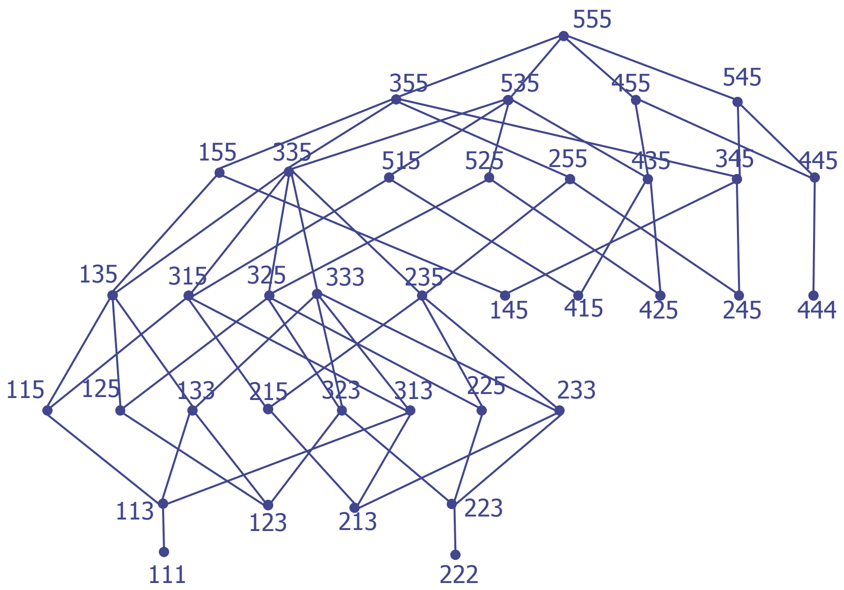
Počet všech možných 3 prvkových variací s opakováním ze 4 prvků je 64, z toho je 20 těch, pro které platí zachování uspořádání f(1) $\le$ f(2) $\le$ f(3), z toho je 10 výsledných zobrazení do N (viz Obr.3), ale žádné zobrazení není prosté, tj do třech

Příklad 4.

Y^X



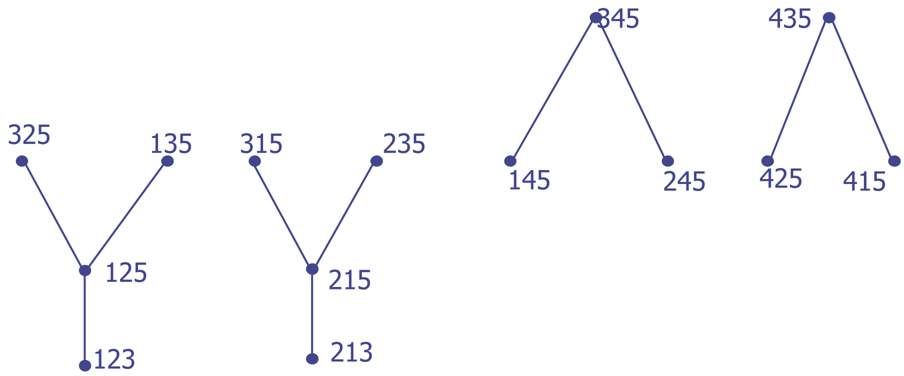
Obr. 4. Výpočet kardinality mocniny posetu Y^X pomocí rozkladu



Obr. 3. Hasseův diagram mocniny Y^X

Výpočet konkrétní kardinality posetu mocniny Y^X (viz Obr. 4)

$|Y^X|=1.5 + 3.6 + 4.2 + 2.10/2 = 37$



Obr. 4. Vrcholy Hasseova diagramu z Obr. 3, které odpovídají prostému zobrazení X do Y

Reference

1. Neggers, J., Kim, H. S. Basic. *Wordl Scientific Publishing Co. Pte. Ltd*. 1998 Singapore. ISBN 981-02-3589-5.
2. Šeptáková, E., Snášel, V., Ochodková, E. Vyhledávání na základě podobnosti v XML dokumentech. sborník konference ZNALOSTI 2003. Ostrava 2003. ISBN 80-248-0229-5.
3. Beran, L. *Uspořádané množiny*. Škola mladých matematiků, ÚV matematické olympiády nakladatelství Mladá fronta, 1978

Annotation:
In this paper I describe a new approach to searching in XML documents. This method is based on exponentiation of graphs and ordered sets. This method allows to order results in natural manner and this is the main advantage against the other methods. Another advantage of presented method is its natural concept of realizing XML documents structure.