Bit-vector encoding and matrix decomposition

Lukas Havrlant

DEPARTMENT OF COMPUTER SCIENCE
PALACKÝ UNIVERSITY, OLOMOUC

Contents

Binary matrix

- A binary matrix is a matrix A such that $A_{i,j} \in \{0, 1\}$
- What about a product of two binary matrices?

Example

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 1 + 1 \cdot 1 & 0 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

It does not seem right...

Product of two binary matrices

We need another definition for a product of two binary matrices.

Definition (Product of two binary matrices)

Given an $n \times k$ matrix A and $k \times m$ matrix B, the product $A \circ B$ is defined as

$$(A \circ B)_{i,j} = \bigvee_{l=1}^{k} A_{il} \cdot B_{lj}$$

Example

$$\begin{aligned}\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix} \cdot \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix} &= \begin{pmatrix}0\cdot 1 \vee 1 \cdot 1 & 0 \cdot 1 \vee 1 \cdot 0\\ 1 \cdot 1 \vee 1 \cdot 1 & 1 \cdot 1 \vee 1 \cdot 0\end{pmatrix} = \\ &= \begin{pmatrix}\max(0\cdot 1, 1 \cdot 1) & \max(0 \cdot 1, 1 \cdot 0)\\ \max(1 \cdot 1, 1 \cdot 1) & \max(1 \cdot 1, 1 \cdot 0)\end{pmatrix} = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}\end{aligned}$$

Binary matrix decomposition

- In general: a matrix decomposition is a factorization of a matrix into a product of matrices.
- Binary matrix decomposition is a factorization of a matrix into a product of (two) binary matrices.
- Problem description: The input is a $n \times m$ binary matrix I and the output are $n \times k$ binary matrix A and $k \times m$ binary matrix B such that $I = A \circ B$ and k is as small as possible.
- k is number of factors and it is well-known Schein rank in boolean matrix theory.

Interpretation of a binary matrix decomposition

What is the interpretation of decomposition of a binary matrix? (rows = patients, columns = symptoms, 1 = patient has symptom)

$$I = \begin{pmatrix} 1&1&1&0&1&0&0&0\\ 1&1&0&0&0&1&0&1\\ 0&1&0&0&1&0&1&0\\ 1&1&0&0&0&1&0&1\\ 1&1&1&0&1&0&0&0\\ 0&1&0&0&1&0&1&0\\ 0&1&0&0&1&0&1&0\\ 0&0&0&0&0&0&1&0\\ 1&1&1&0&1&0&0&0\\ 0&0&0&0&0&0&1&0\\ 1&1&1&0&1&0&0&0\\ 1&1&0&0&0&1&0&1 \end{pmatrix}. \quad A_{\mathcal{F}} = \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&1\\ 0&1&0&0\\ 1&0&0&0\\ 0&0&1&1\\ 0&0&1&1\\ 0&0&0&1\\ 1&0&0&0\\ 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0 \end{pmatrix}, \quad B_{\mathcal{F}} = \begin{pmatrix} 1&1&1&0&1&0&0&0\\ 1&1&0&0&0&1&0&1\\ 0&1&0&0&1&0&1&0\\ 0&0&0&0&0&0&1&0 \end{pmatrix}$$

Symptoms (columns):

1: headache	5: cold
2: fever	6: stiff neck
3: painful limbs	7: rash
4: swollen glands in neck	8: vomiting

Maximal rectangle in a binary matrix

The question is: how to efficiently compute a decomposition of a binary matrix? We can use approach using formal concept analysis.

Definition (Formal context)

The formal context is a tripple $\langle X, Y, I \rangle$, where X is a set of "objects", Y is a set of "attributes" and $I \subseteq X \times Y$ is realation "object x has attribute y".

For example: X is a set of patients, Y is a set of symptoms.

Definition (Rectangle)

A rectangle $R = \langle A, B \rangle$ is a subset of I and there exists $A \subseteq X, B \subseteq Y$ such that $R = A \times B$.

$$\begin{pmatrix} \mathbf{1} & \mathbf{1} & 0 & 1 \\ \mathbf{1} & \mathbf{1} & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} \mathbf{1} & \mathbf{1} & 0 & 1 \\ \mathbf{1} & \mathbf{1} & 1 & 0 \\ \mathbf{1} & \mathbf{1} & 0 & 0 \end{pmatrix}$$

Rectangle   Maximal rectangle

Covering of a relation I

Let $\mathcal{R}$ be a set of all maximal rectangles in a context $\langle X, Y, I\rangle$. Clearly:

$$\bigcup \mathcal{R} = I.$$

Problem: finding a minimal set $\mathcal{F} \subseteq \mathcal{R}$ such that $\bigcup \mathcal{F} = I$.

$$I = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad I_1 = \begin{pmatrix} \mathbf{1} & \mathbf{1} & 0 \\ \mathbf{1} & \mathbf{1} & 0 \\ 0 & 1 & 1 \end{pmatrix}, I_2 = \begin{pmatrix} 1 & \mathbf{1} & 0 \\ 1 & \mathbf{1} & 0 \\ 0 & \mathbf{1} & 1 \end{pmatrix}, I_3 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & \mathbf{1} & \mathbf{1} \end{pmatrix}.$$

Now: $I_1 \cup I_2 \cup I_3 = I$, but also $I_1 \cup I_3 = I$. The set $\{I_1, I_3\}$ is the minimal set which covers a relation I.

Algorithm for computing a decomposition of binary matrix

Let $\mathcal{F} = \{\langle A_1, B_1\rangle, \langle A_2, B_2\rangle, \dots, \langle A_k, B_k\rangle\}$ is a minimal set of rectangles which covers relation I of a context $\langle X, Y, I\rangle$. Then we can compute $n \times k$ matrix $A_{\mathcal{F}}$ and $k \times m$ matrix $B_{\mathcal{F}}$:

$$(A_{\mathcal{F}})_{il} = \begin{cases} 1 & \text{if } i \in A_l \\ 0 & \text{if } i \in A_l \end{cases}, \qquad (B_{\mathcal{F}})_{lj} = \begin{cases} 1 & \text{if } j \in B_l \\ 0 & \text{if } j \in B_l \end{cases}$$

The lth column of $A_{\mathcal{F}}$ consists of the characteristic vector of A_l and the lth row of $B_{\mathcal{F}}$ consists of the characteristic vector of B_l.

From the previous slide: we have

$$\mathcal{F} = \{I_1, I_3\} = \{\langle\{1,2\},\{1,2\}\rangle, \langle\{3\},\{2,3\}\rangle\}$$

We will get:

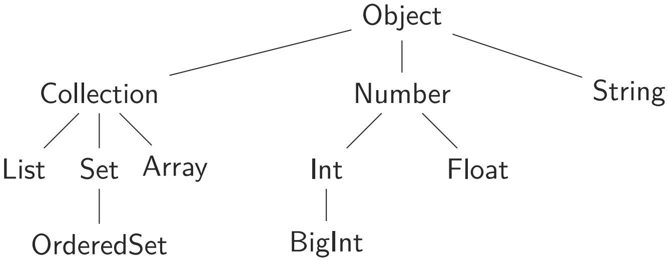
$$A_{\mathcal{F}} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, B_{\mathcal{F}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad A_{\mathcal{F}} \circ B_{\mathcal{F}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Conclusion 1

The main point: if we have a minimal set $\mathcal{F}$ of rectangles which covers the relation I, we can easily compute a decomposition of binary matrix I.

Bit-vector encoding of partially ordered set

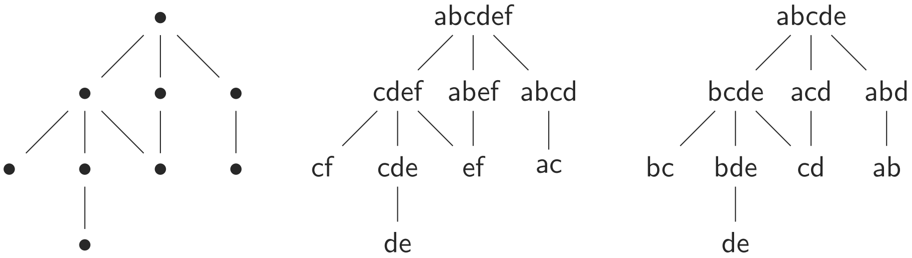
- Partially ordered set (poset) is a reflexive, antisymmetric, and transitive binary relation.
- How to represent a poset in a computer? How to efficiently compute if $x \leq_P y$ for a poset P? Or $x \wedge_P y$? What is it good for?



Suppose we have a function with one parameter of type `Collection`. Can we call this function with an argument of type `OrderedSet`? Is `OrderedSet` subtype of `Collection`? In other words – is `OrderedSet` $\leq$ `Collection`?

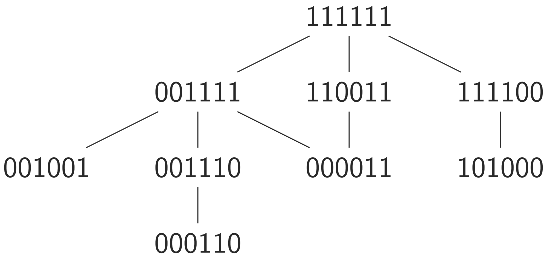
Bit-vector encoding

- There are different approaches of how to encode a poset P. Some of them:
  1. Graph representation: every element knows his upper/lower neighbors.
  2. A binary matrix representation: $A_{i,j} = 1$ iff $x_i \leq x_j$.
  3. Bit-vector encoding.
  4. ...
- Bit-vector encoding: we add a label (a set) to each element such that $x \leq_P y$ iff label$(x) \subseteq$ label(y).



Implementation of a bit-vector encoding

- Bit-vector encoding is very easy to implement.
- We can represent a (sub)set as a characteristic vector.
- The query $x \leq_P y$ is then equal to a bit-operation $x \& y = x$ (logical and).



Question: What is the size of a minimal set which can be used for encoding a poset P?

Order embedding

Definition (Order embedding)

Given two posets P and S, an order embedding f is a map $f : P \to S$ such that

$$\forall x, y \in P : \quad x \leq_P y \text{ iff } f(x) \leq_S f(y).$$

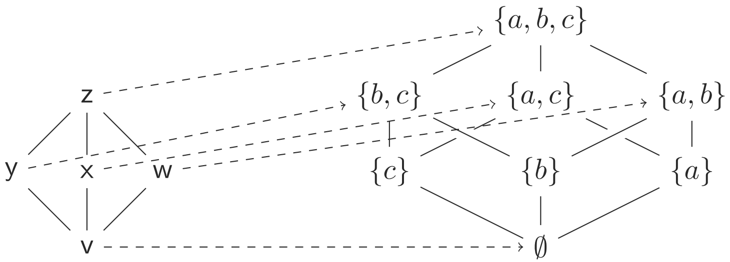


Figure : Example of an order embedding f from a poset P to a poset S

2-dimension

For every set S the power set $\mathcal{P}(S)$ always forms a poset $\langle \mathcal{P}(S), \subseteq \rangle$.

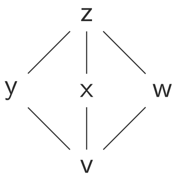
Definition (2-dimension)

The 2-dimension of a poset P is the size of a minimal set S such that P can be embedded into a poset $\langle \mathcal{P}(S), \subseteq \rangle$. We will denote it $\dim_2 P$.

The $\dim_2 P$ is also equal to the size of a minimal set which can be used for encoding a poset P. We can see the embedding map as a label function.

A poset as a context

We can convert a poset $\mathbf{P} = \langle P, \leq \rangle$ into a formal context $C = \langle P, P, \leq \rangle$.



$\longrightarrow$

	z	y	x	w	v
z	1	0	0	0	0
y	1	1	0	0	0
x	1	0	1	0	0
w	1	0	0	1	0
v	1	1	1	1	1

A wild idea: Is it possible to compute a 2-dimension of P using a set of rectangles covering the relation $\leq$? (Spoiler: no!)

Concept lattice

Let $C = \langle X, Y, I \rangle$ be a formal context and $\mathcal{B}(X, Y, I)$ the set of all maximal rectangles in C. $\mathcal{B}(X, Y, I)$ with operation $\leq_C$ defined as

$$\langle A, B \rangle, \langle C, D \rangle \in \mathcal{B}(X, Y, I) : \quad \langle A, B \rangle \leq_C \langle C, D \rangle \text{ iff } A \subseteq C \quad (\text{or } B \supseteq D)$$

forms a (concept) lattice $\langle \mathcal{B}(X, Y, I), \leq_C \rangle$.

Theorem

For a poset $\mathbf{P} = \langle P, \leq \rangle$*:*

$$\dim_2 \mathbf{P} = \dim_2 \langle \mathcal{B}(P, P, \leq), \leq_C \rangle$$

Corollary: We can compute $\dim_2 \langle \mathcal{B}(P, P, \leq), \leq_C \rangle$ (or $\mathcal{B}(P, P, \leq)$, in short) instead of $\dim_2 \mathbf{P}$.

Ferrers 2-dimension

Definition (Ferrers 2-dimension)

Let $L = \langle \mathcal{B}(X, Y, I), \leq_C \rangle$ be a concept lattice. Let $\mathcal{R}$ be a minimal set of maximal rectangles which covers a relation $X \times Y \setminus I$. Then Ferrers 2-dimension of a lattice L is $|\mathcal{R}|$. We will denote it $\mathrm{fdim}_2\, C$.

Ferrers 2-dimension $=$ number of maximal rectangles which cover "a complementary matrix" ($=$ zeros, not ones). Example:

$$\begin{array}{cccccc}
 & z & y & x & w & v \\
z & 1 & \boxed{0} & 0 & 0 & \boxed{0} \\
y & 1 & 1 & 0 & 0 & 0 \\
x & 1 & \boxed{0} & 1 & 0 & \boxed{0} \\
w & 1 & \boxed{0} & 0 & 1 & \boxed{0} \\
v & 1 & 1 & 1 & 1 & 1
\end{array}$$

This matrix has Ferrers 2-dimension 3, because we need 3 rectangles to cover "zeros".

2-dimension and Ferrers 2-dimension

(It's not a surprise that. . .)

Theorem

For a poset $\mathbf{P} = \langle P, \leq \rangle$*:*

$$\dim_2 \mathbf{P} = \operatorname{fdim}_2 \mathcal{B}(P, P, \leq).$$

Proof.

Ten pages in FCA by Ganter&Wille. . . $\square$

Conclusion 2

1. The size of a minimal set for encoding a poset P is equal to $\dim_2 P$.
2. $\dim_2 P = \mathrm{fdim}_2\, \mathcal{B}(P, P, \leq)$.
3. $\mathrm{fdim}\, \mathcal{B}(P, P, \leq)$ is equal to a size of a minimal set of rectangles $\mathcal{R}$ which cover complementary relation $>$.
4. From conclusion 1: the set $\mathcal{R}$ can be used to decompose a binary matrix $>$

Final conclusion: The size of a minimal set for encoding a poset P is equal to Schein rank of a matrix $>$. (Schein rank = number of factors in the decomposition)

In progress: It should be possible to compute an encoding of a poset P using algorithms for decomposition matrix.

Thank you for your attention!