Sequential Patterns

- Find the occurrences of the patterns

  - <{F}{D}>

  - <{A} {C} {D} >

  - <{B}{A}>

  - <{A}{F} {D}>

  - <{A}{E}>

  in the following input sequence

$$< \quad \{A,B\} \quad \{A,C,D\} \quad \{C,F\} \quad \{A,D\} \quad \{A,B,D\} \quad \{E\} \quad \{A,B,F\} \quad \{D\} >$$

| t=0 | t=1 | t=2 | t=3 | t=4 | t=5 | t=6 | t=7 |

with (i) no temporal constraints and (ii) max-gap = 3

Given the following dataset of sequences

ID	Sequence
1	a b → a → b
2	b → a → c d
3	a → b
4	a → a → b d

 Generate sequential patterns if min_sup = 35%

Sequential patterns with GSP

GSP Exercise - solution

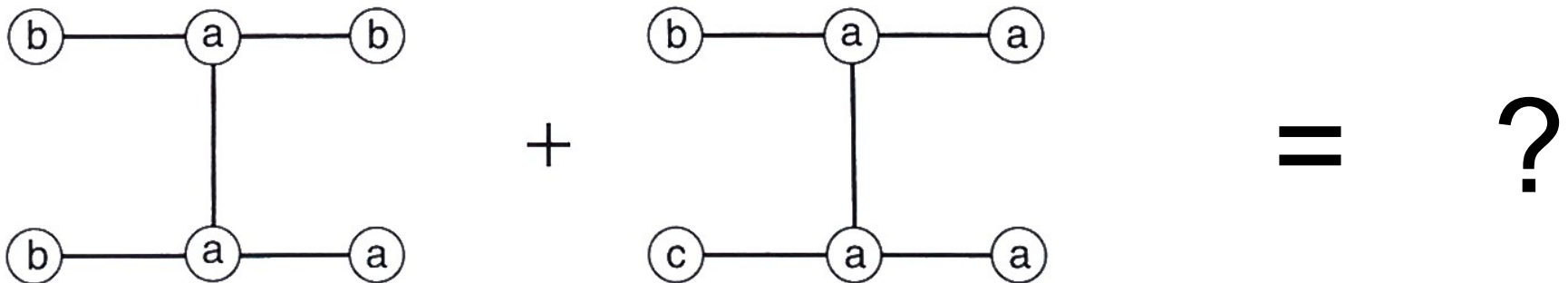| Sequential pattern	Support
a	100 %
b	100 %
d	50 %
a → a	50 %
a → b	75 %
a → d	50 %
b → a	50 %
a → a → b	50 %

Graph mining
Candidates generation

- Draw all candidate subgraphs obtained by joining the following pairs of graphs through edge-growing candidate generation
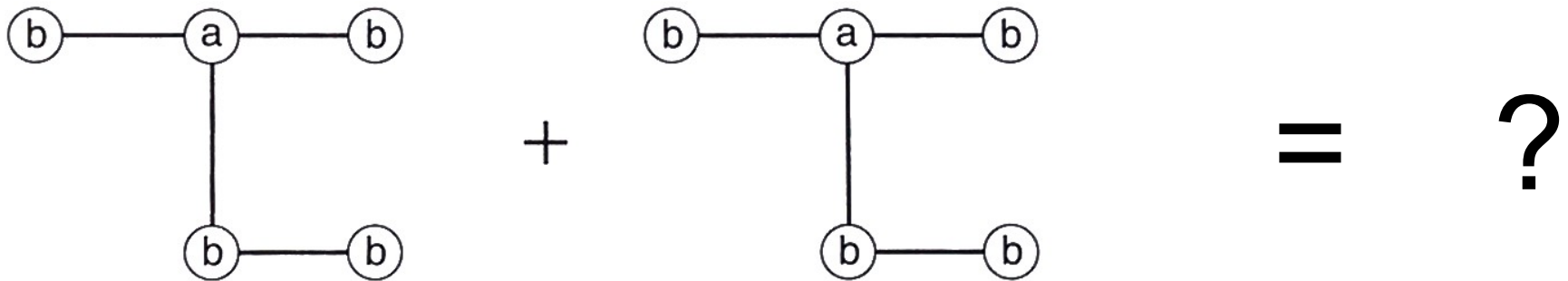
(a)

(b)

Given the following dataset of time series:

ID	Time series
A	< 19, 14, 19, 19, 26, 29, 38, 30 >
B	< 15, 10, 0, 2, 4, 7, 1, 9 >
C	< 19, 11, 20, 27, 18, 25, 15, 19 >
D	< 12, 3, 12, 19, 18, 24, 27, 31 >

compute the matrix of distances among all pairs of time series adopting a Dynamic Time Warping distance, constrained with a "Sakoe-Chiba Band " of size r=2, i.e. the maximum misalignment allowed in the matching is of 2 positions.

Times series: DTW