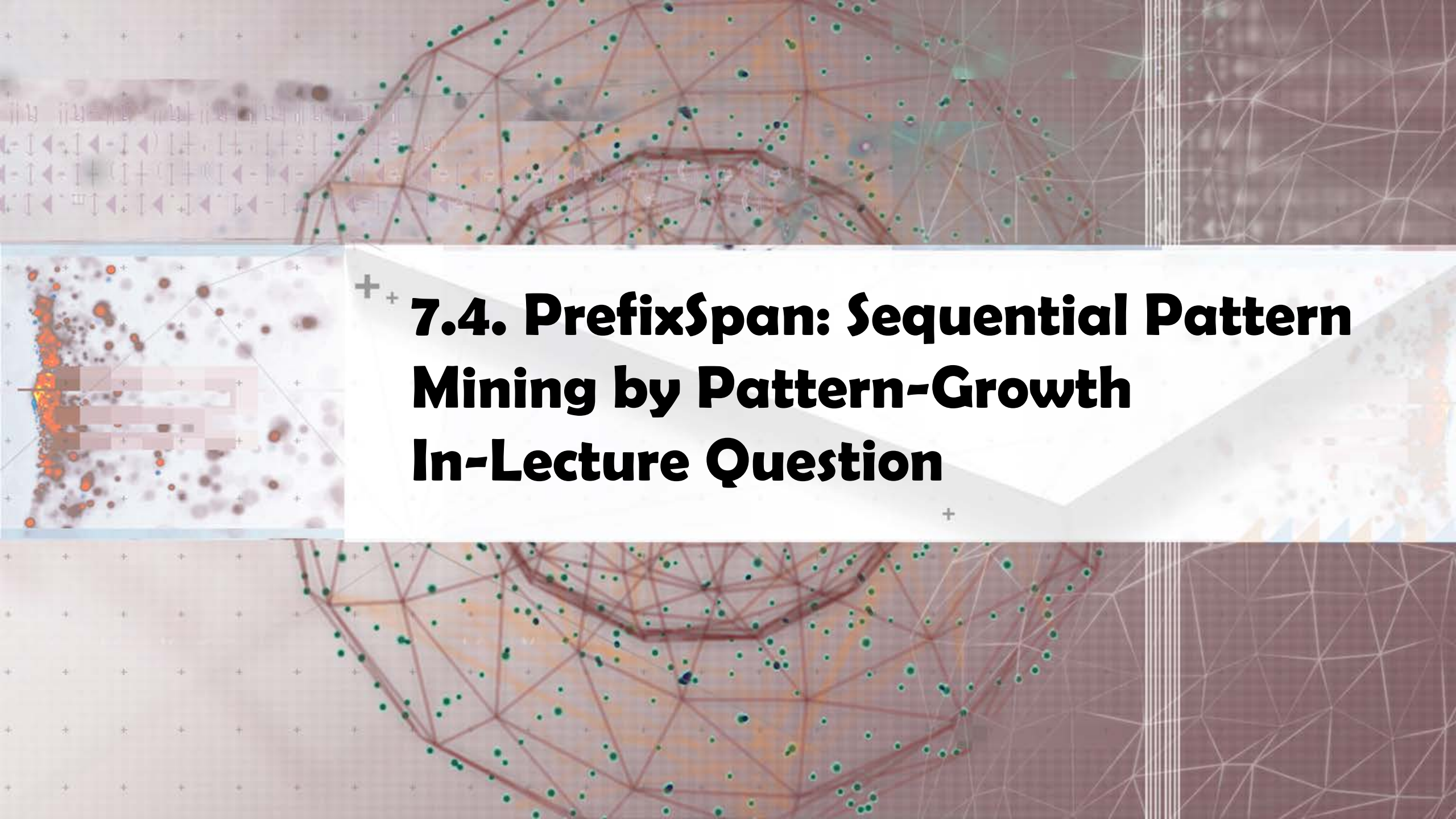

7.4. PrefixSpan: Sequential Pattern Mining by Pattern-Growth In-Lecture Question

Question

For the given sequence database, which of the following belongs to the <f>-projected database?

- <(c_)>
- <(ab)(df)cb>
- <cb>
- <(_a)cbc>

SID	Sequence
10	<a(abc)(ac)d(cf)>
20	<(ad)c(bc)(ae)>
30	<(ef)(ab)(df)cb>
40	<eg(af)cbc>

2

Question

For the given sequence database, which of the following belongs to the <f>-projected database?

- <(c_)>
- <(ab)(df)cb>
- <cb>
- <(_a)cbc>

SID	Sequence
10	<a(abc)(ac)d(cf)>
20	<(ad)c(bc)(ae)>
30	<(ef)(ab)(df)cb>
40	<eg(af)cbc>

- **Answer:** <(ab)(df)cb>
- **Explanation:** All the sequences in the database will be represented as follows in the <f>-projected database.
  - SID-10: <>
  - SID-20: <>
  - SID-30: <(ab)(df)cb>
  - SID-40: <cbc>
- By simple comparison, we can see that B belongs to the <f>-projected database.

3