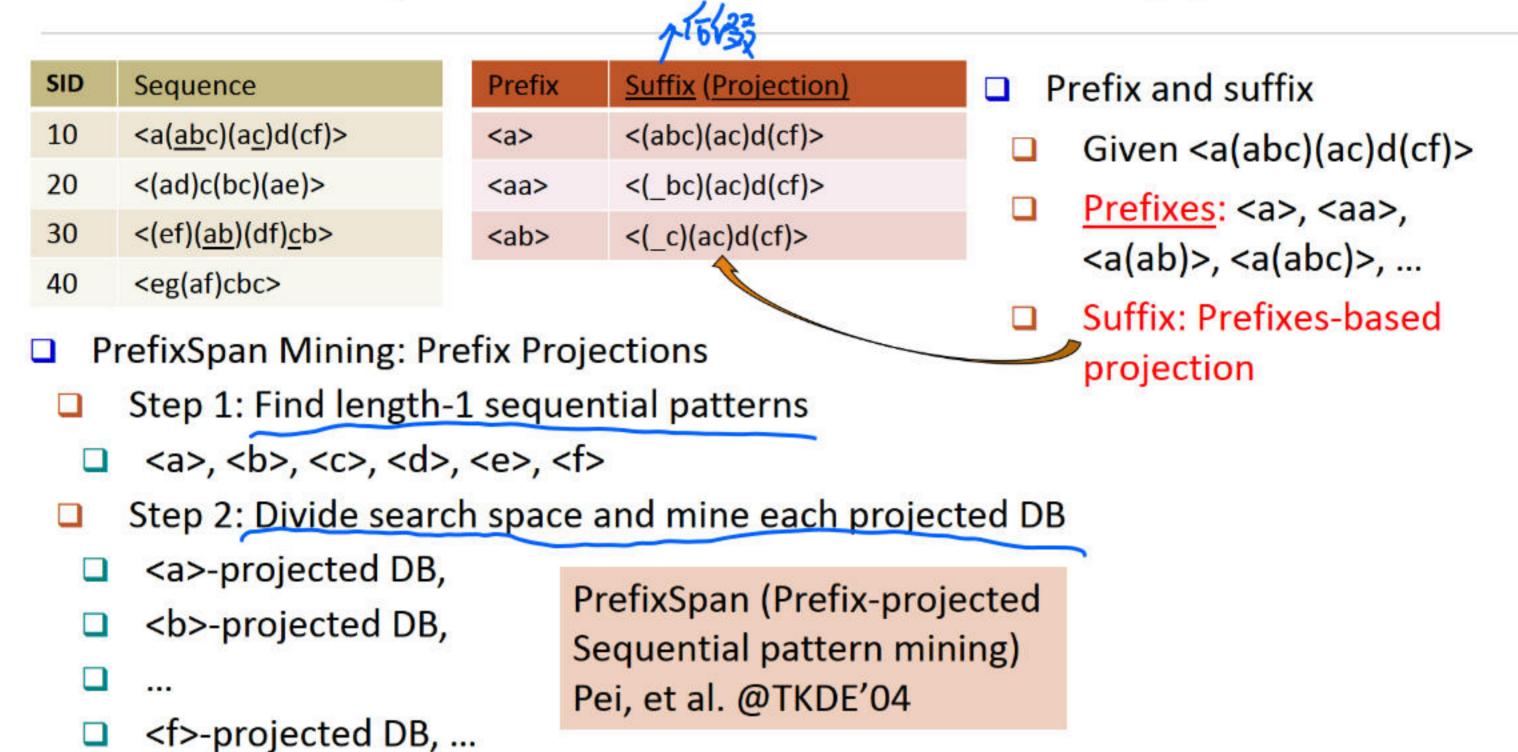
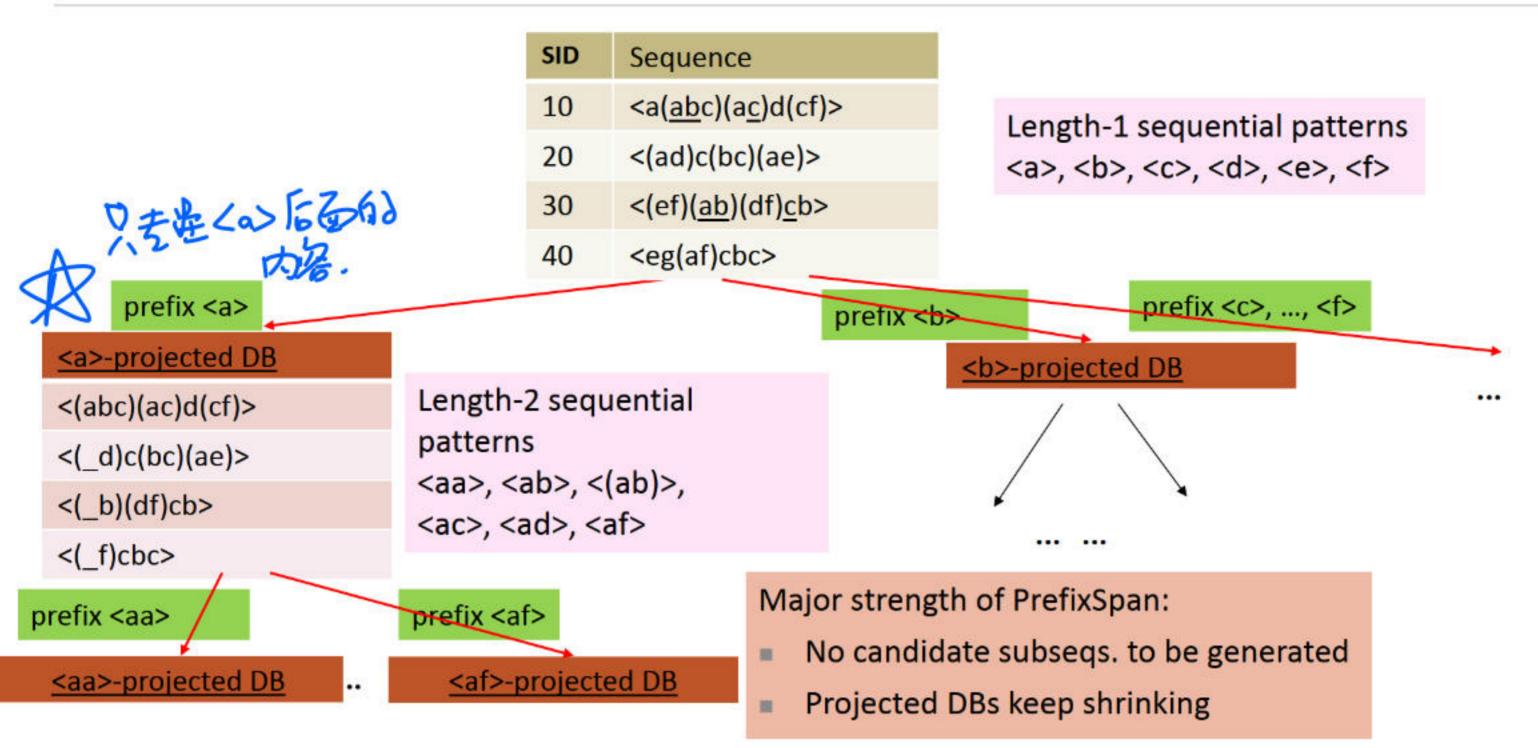


PrefixSpan: A Pattern-Growth Approach



PrefixSpan: Mining Prefix-Projected DBs



Implementation Consideration: Pseudo-Projection vs. Physical Projection

- Major cost of PrefixSpan: Constructing projected DBs
 Suffixes largely repeating in recursive projected DBs
 When DB can be held in main memory, use pseudo projection
 No physically copying suffixes
 Pointer to the sequence
 Offset of the suffix
 But if it does not fit in memory
 Physical projection
 S|<ab>: (, 4)

 s|<ab>: (_c)(ac)d(cf)>
- Suggested approach:
 - Integration of physical and pseudo-projection
 - Swapping to pseudo-projection when the data fits in memory