Where we are

Background (15 min)

Graph models, subgraph isomorphism, subgraph mining, graph clustering

Exploratory Graph Analysis (40 min)

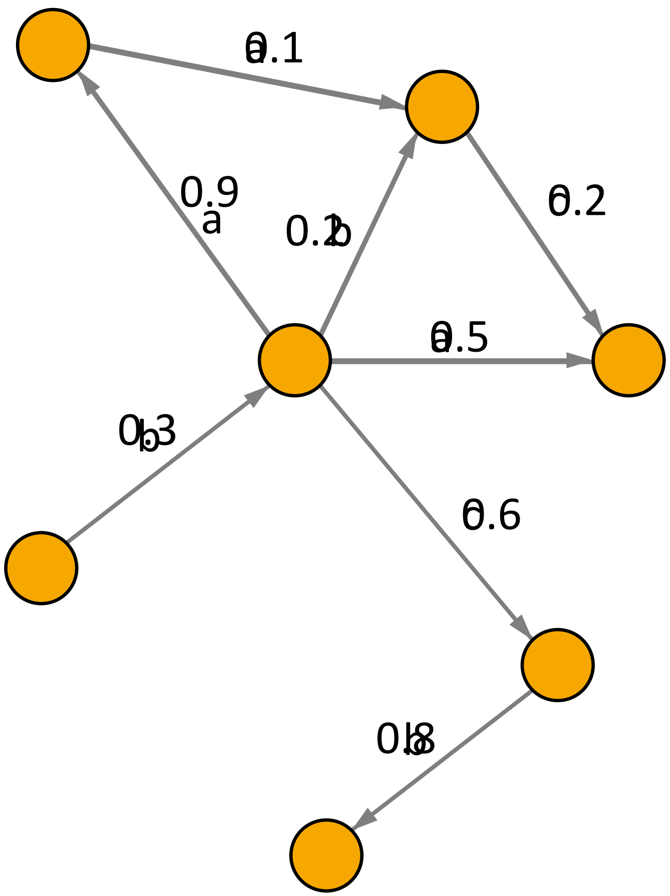
Focused Graph Mining (40 min)

Refinement of Query Results (40 min)

Machine Learning and Visualization (40 min)

Challenges and discussion

Graphs



$G = (V, E, l)$ / $G = (V, E, p)$

Vertices

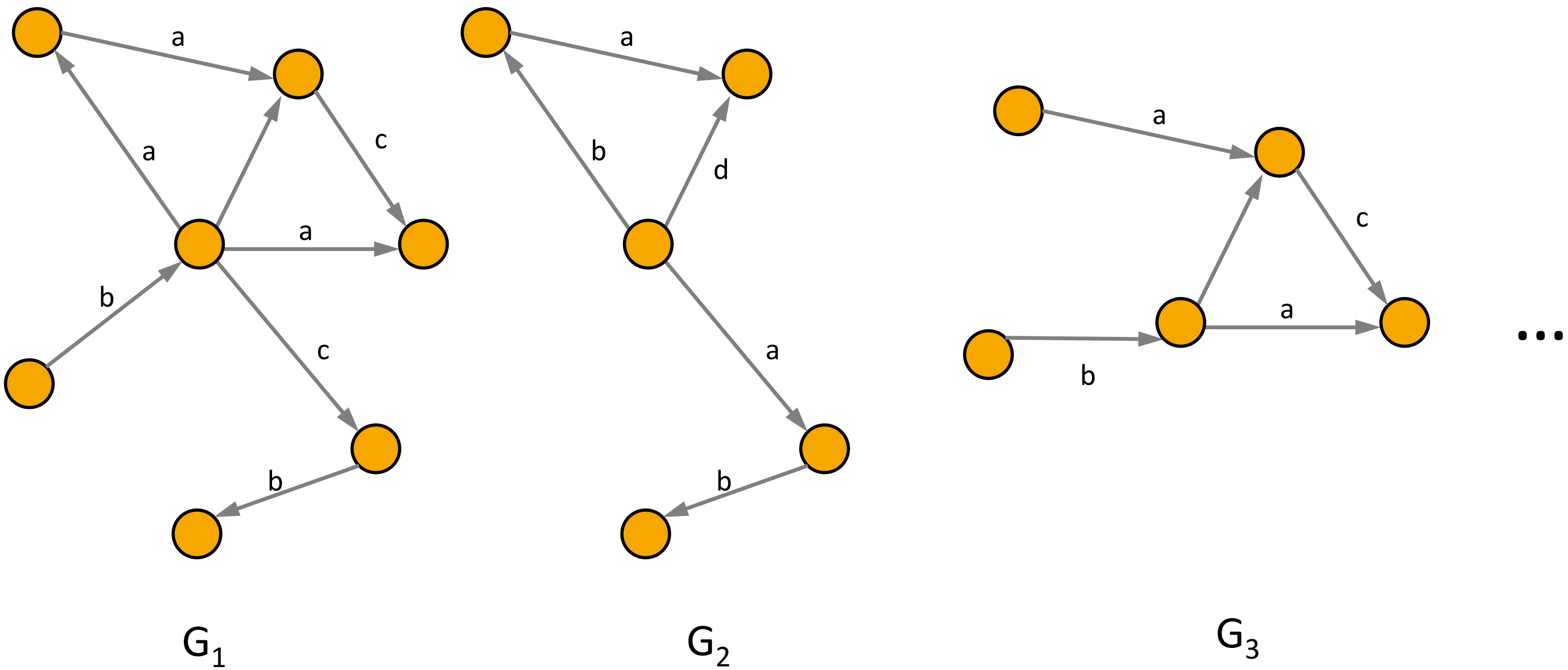
Edges

Label / Probability function

$l: V \cup E \to \Sigma$

- Undirected Graphs
  - Co-authorship, Pipes, Biological
- Directed graphs
  - Follows, Roads, co-citations …
- Labeled Graphs
  - Knowledge graphs, social, …
- Probabilistic graphs
  - Causal graphs

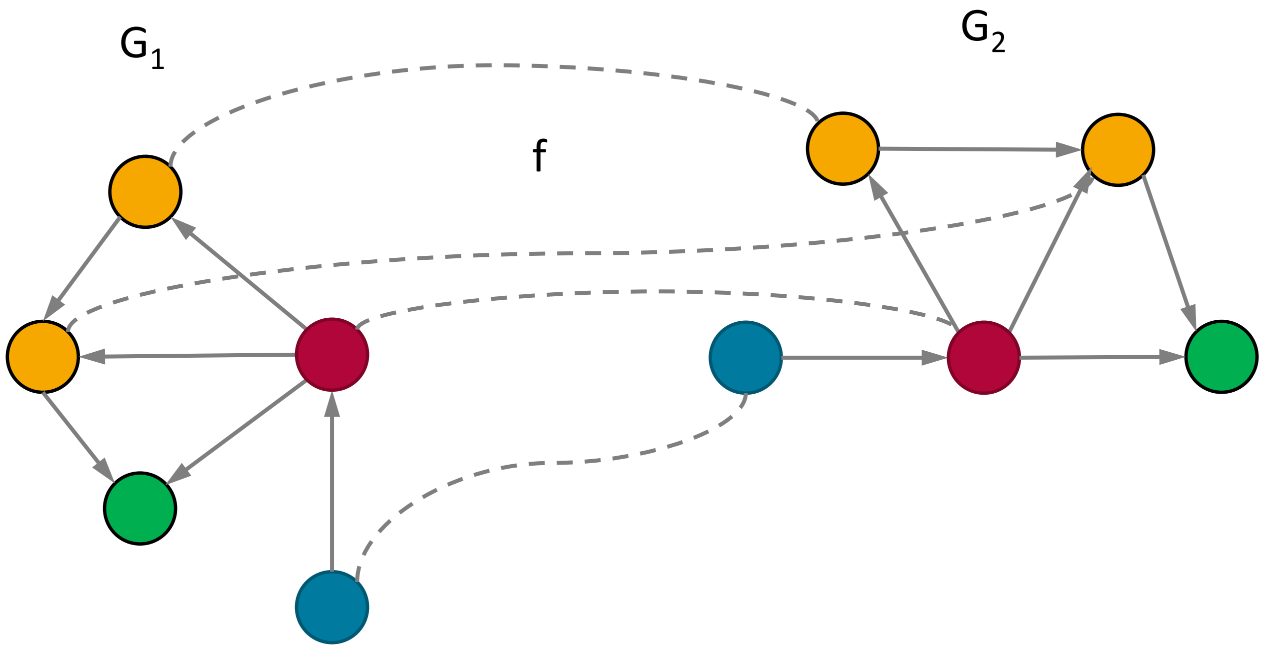
Graph databases (set of graphs)



$$D = \{G_1, G_2, \ldots, G_n\}, G_i = \langle V_i, E_i, l_i \rangle, l_i: E_i \cup V_i \rightarrow \Sigma$$

Set of small labeled graphs
Chemical compounds, Business models, 3D objects

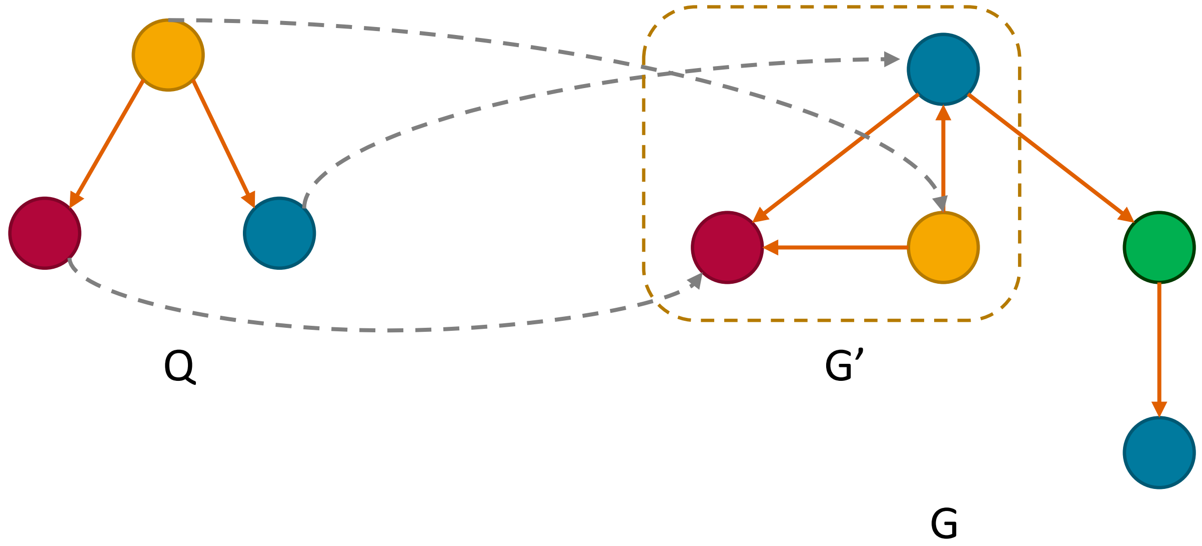
Graph Isomorphism



Given two graphs, $G_1 : \langle V_1, E_1, l_1 \rangle$, $G_2 : \langle V_2, E_2, l_2 \rangle$ G_1 is isomorphic G_2 iff exists a **bijective** function $f : V_1 \rightarrow V_2$ s.t.:

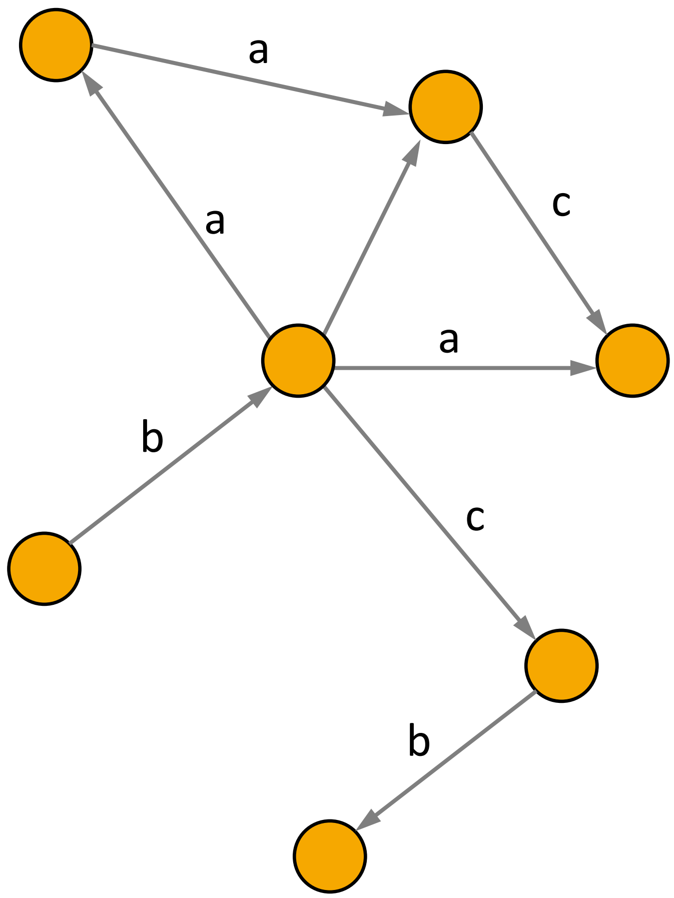
1. For each $v_1 \in V_1, l(v_1) = l(f(v_1))$
2. $(v_1, u_1) \in E_1$ iff $\left(f(v_1), f(u_1)\right) \in E_2$

Subgraph Isomorphism



A graph ,$Q\colon\langle V_Q, E_Q, l_Q\rangle$ is subgraph isomorphic to a graph $G\colon\langle V, E, l\rangle$ if exists a subgraph $G' \sqsubseteq G$, isomorphic to Q

Graph Clustering and Community Detection



Given: graph with nodes, edges, labels

$$G = (V, E, l)$$

Vertices   Edges   Labeling function $l: V \cup E \rightarrow \Sigma$

Discover: a partitioning of communities

$$C = \{C_1, C_2, C_3, \ldots, C_k\}$$

- **Optimize a given quality criterion** Q(C), e.g. ***Modularity*** or other measures
- Is an **NP-hard problem** to find the optimal partitioning