Graph Matching

Input: 2 digraphs G1 = (V1,E1), G2 = (V2,E2)

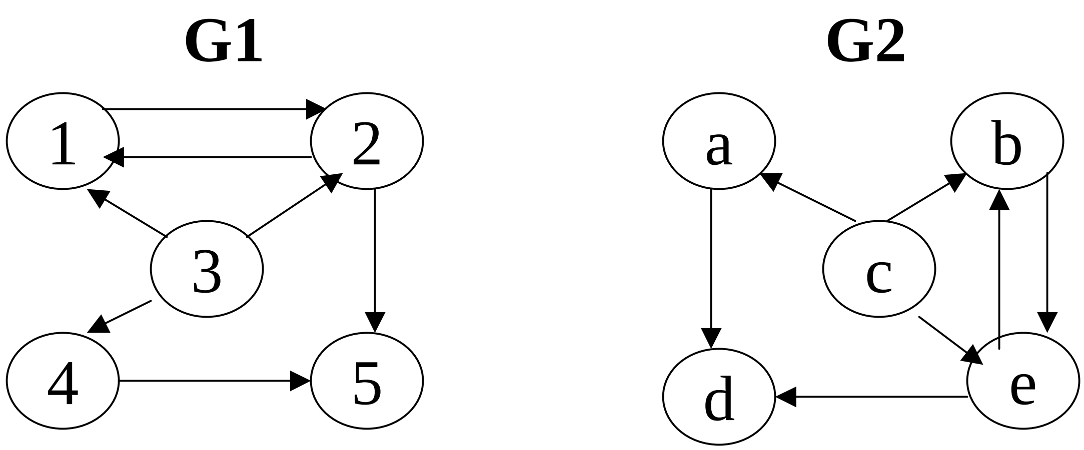
Questions to ask:

1. Are G1 and G2 isomorphic?
2. Is G1 isomorphic to a subgraph of G2?
3. How similar is G1 to G2?
4. How similar is G1 to the most similar subgraph of G2?

Isomorphism for Digraphs

G1 is isomorphic to G2 if there is a 1-1, onto mapping h: V1 $\rightarrow$ V2 such that

$$(v_i, v_j) \in E1 \text{ iff } (h(v_i), h(v_j)) \in E2$$

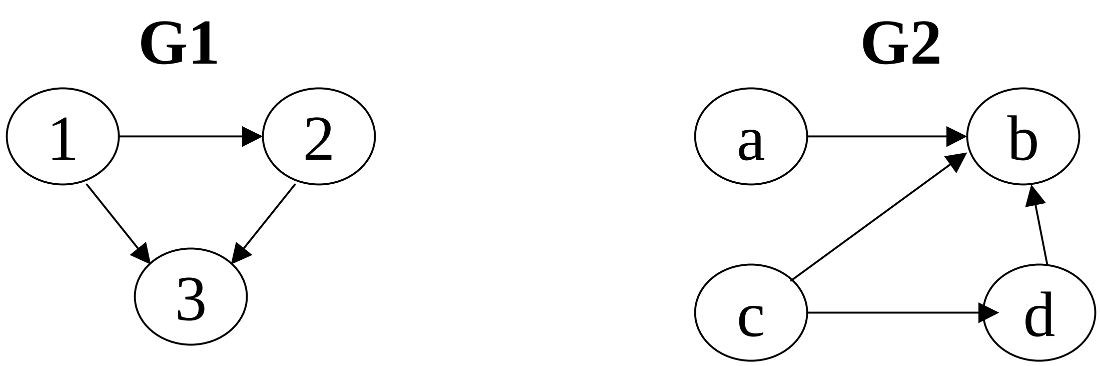


Find an isomorphism h: {1,2,3,4,5} $\rightarrow$ {a,b,c,d,e}.
Check that the condition holds for every edge.

Subgraph Isomorphism for Digraphs

G1 is isomorphic to a subgraph of G2 if there is a 1-1 mapping h: $V1 \rightarrow V2$ such that

$$(vi,vj) \in E1 \Rightarrow (h(vi), h(vj)) \in E2$$



Isomorphism and subgraph isomorphism are defined similarly for undirected graphs.

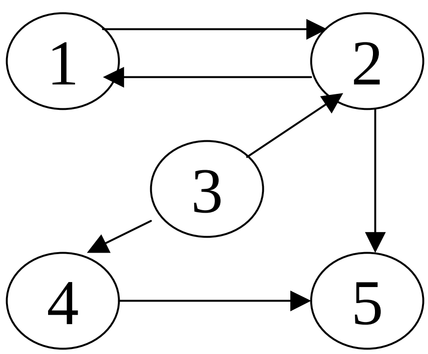
In this case, when $(vi,vj) \in E1$, either (vi,vj) or (vj,vi) can be listed in E2, since they are equivalent and both mean {vi,vj}.

Similar Digraphs

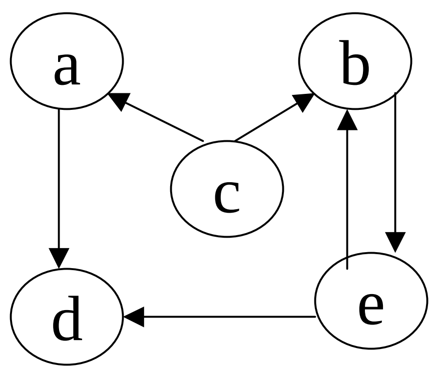
Sometimes two graphs are close to isomorphic, but have a few “errors."

Let h(1)=b, h(2)=e, h(3)=c, h(4)=a, h(5) = d.

G1



G2



(1,2)	(b,e)
(2,1)	(e,b)
X	(c,b)
(4,5)	(a,d)
(2,5)	(e,d)
(3,2)	X
(3,4)	(c,a)

The mapping **h** has **2 errors**.

(c,b) $\in$ G2, but (3,1) $\notin$ G1

(3,2) $\in$ G1, but (c,e) $\notin$ G2

Error of a Mapping

Intuitively, the error of mapping h tells us
- how many edges of G1 have no corresponding edge in G2 and
- how many edges of G2 have no corresponding edge in G1.

Let G1=(V1,E1) and G2=(V2,E2), and let $h: V1 \rightarrow V2$ be a 1-1, onto mapping.

forward error	$EF(h) = \|\{(vi,vj) \in E1 \mid (h(vi),h(vj)) \notin E2\}\|$ edge in E1 corresponding edge not in E2
backward error	$EB(h) = \|\{(vi,vj) \in E2 \mid (h^{-1}(vi),h^{-1}(vj)) \notin E1\}\|$ edge in E2 corresponding edge not in E1
total error	$Error(h) = EF(h) + EB(h)$
relational distance	$GD(G1,G2) = \min Error(h)$ for all 1-1, onto $h: V1 \rightarrow V2$

Variations of Relational Distance

1. normalized relational distance:
   Divide by the sum of the number of edges in E1 and those in E2.

2. undirected graphs:
   Just modify the definitions of EF and EB to accommodate.

3. one way mappings:
   h is 1-1, but need not be onto
   Only the forward error EF is used.

4. labeled graphs:
   When nodes and edges can have labels, each node should be mapped to a node with the same label, and each edge should be mapped to an edge with the same label.

Graph Matching Algorithms

1. graph isomorphism
* 2. subgraph isomorphism
* 3. relational distance
4. attributed relational distance (uses labels)

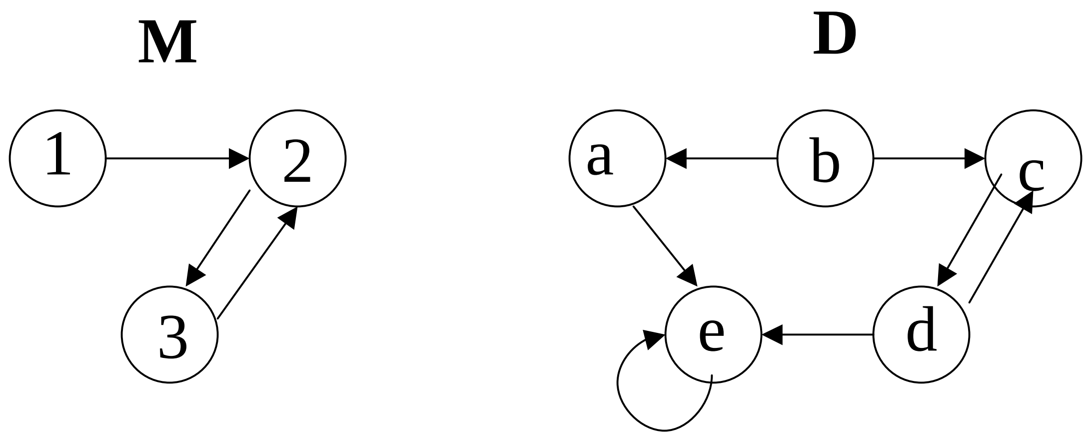
Subgraph Isomorphism

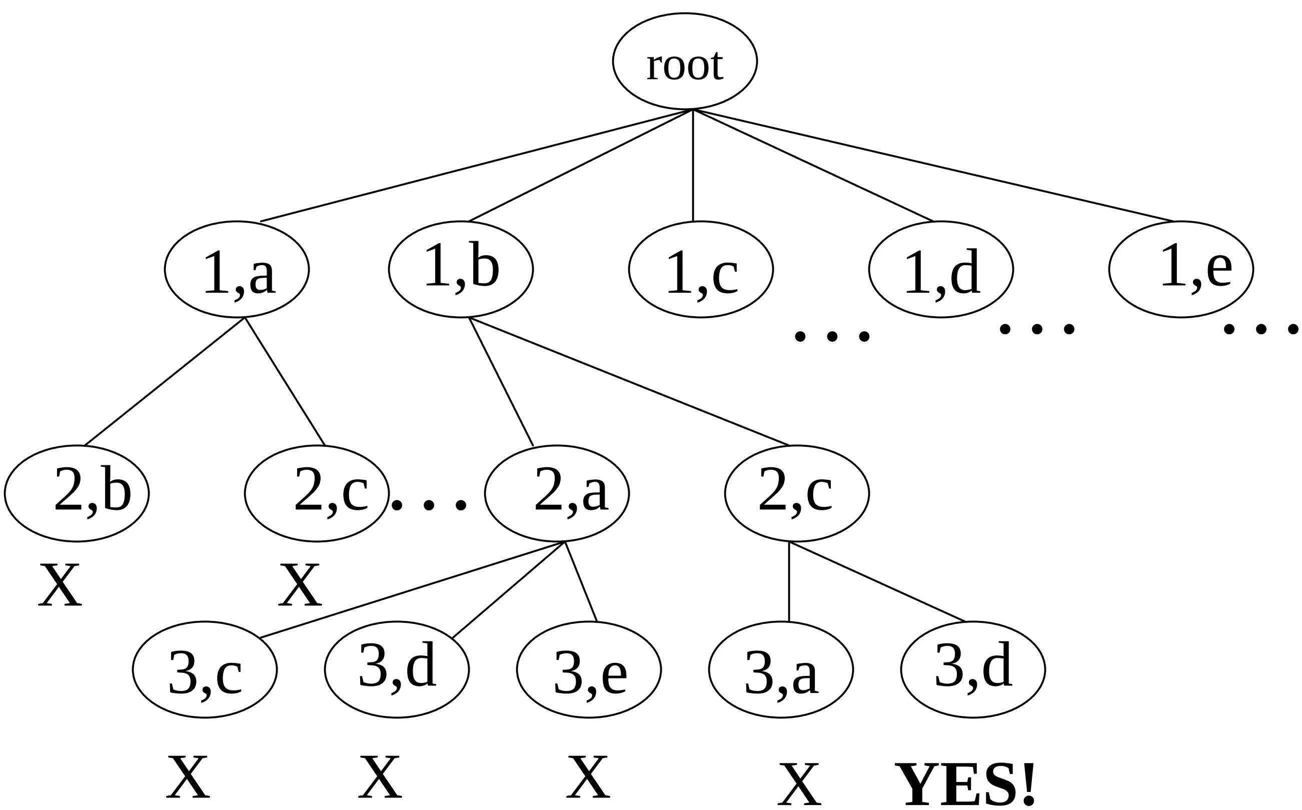
Given model graph M = (VM,EM)
data graph D = (VD,ED)

Find 1-1 mapping $h:VM \rightarrow VD$

satisfying $(vi,vj) \in EM \Rightarrow ((h(vi),h(vj)) \in ED$.

Method: Backtracking Tree Search





Treesearch for Subgraph Isomorphism in Digraphs

```
procedure Treesearch(VM, VD, EM, ED, h)
  {
  v = first(VM);
  for each w ∈ VD
    {
    h’ = h ∪ {(v,w)};
    OK = true;
    for each edge (vi,vj) in EM satisfying that
            either 1. vi = v and vj ∈ domain(h’)
                or 2. vj = v and vi ∈ domain(h’)
          if  ( (h’(vi),h’(vj)) ∉ ED )
              {OK = false; break;};

    if OK   {
                VM’ = VM – v;
                VD’  = VD – w’
                 if isempty(VM’) output(h’);
                 else Treesearch(VM’,VD’,EM,ED,h’)
                 }  }  }
```

Branch-and-Bound Tree Search

Keep track of the least-error mapping.