Polynomial Time Pattern Matching Algorithm for Ordered Graph Patterns

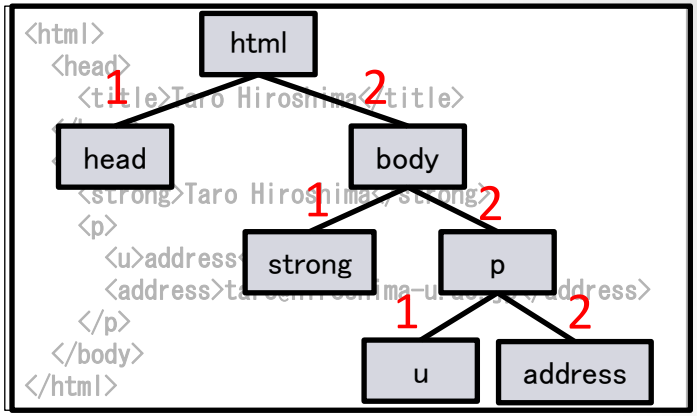
Takahiro Hino[1], Yusuke Suzuki[1],
Tomoyuki Uchida[1], Yuko Itokawa[2]

[1]Department of Intelligent Systems, Hiroshima City University, Japan
[2]Faculty of Psychological Science, Hiroshima International University, Japan

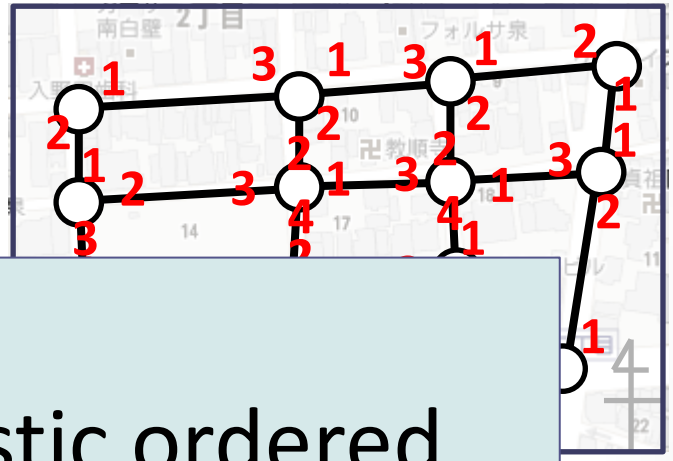
Outline

1. Motivation
2. Preliminaries
   - Ordered graph
   - Ordered graph pattern
   - Ordered graph pattern language
3. Polynomial Time Pattern Matching Algorithm for Ordered Graph Patterns
4. Conclusions & future work

Motivation

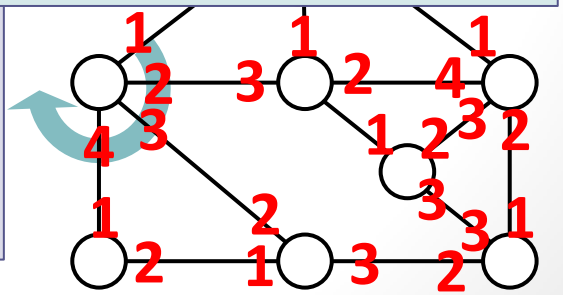


Background:
Structured data such as Web pages, TEX sources, CAD and MAP are modeled by graphs each of whose vertices has ord



Purpose:
Discovery of characteristic ordered graph patterns common to such data

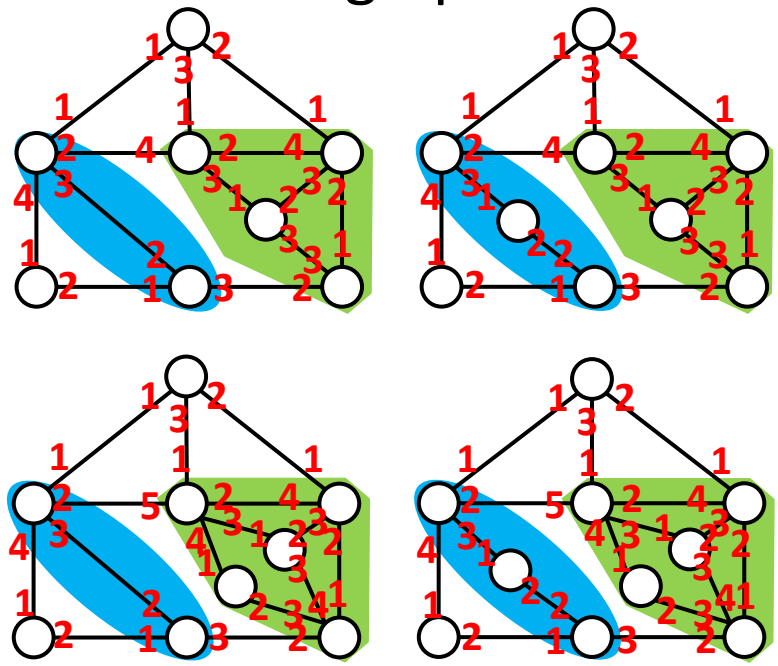
Ordered graph[Jiang and Bunke,1998]:
Each vertex has a unique order based on neighboring vertices.

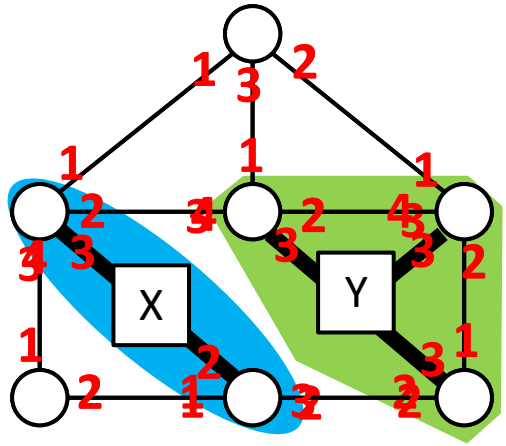


Our Approach

Ordered Graph pattern

as a knowledge representation of common structure of ordered graph structured data.





Ordered graph pattern

Other knowledge representations

1. Logic programming having ordered graph patterns as arguments
2. Decision diagram having ordered graph patterns as attributes

Our approach

The purpose of this research

is to provide a hypothesis checking process for knowledge discovery.

Efficient matching algorithm to solve the following problem:

Does an ordered graph pattern g represent structural character of an ordered graph G?

Related works:

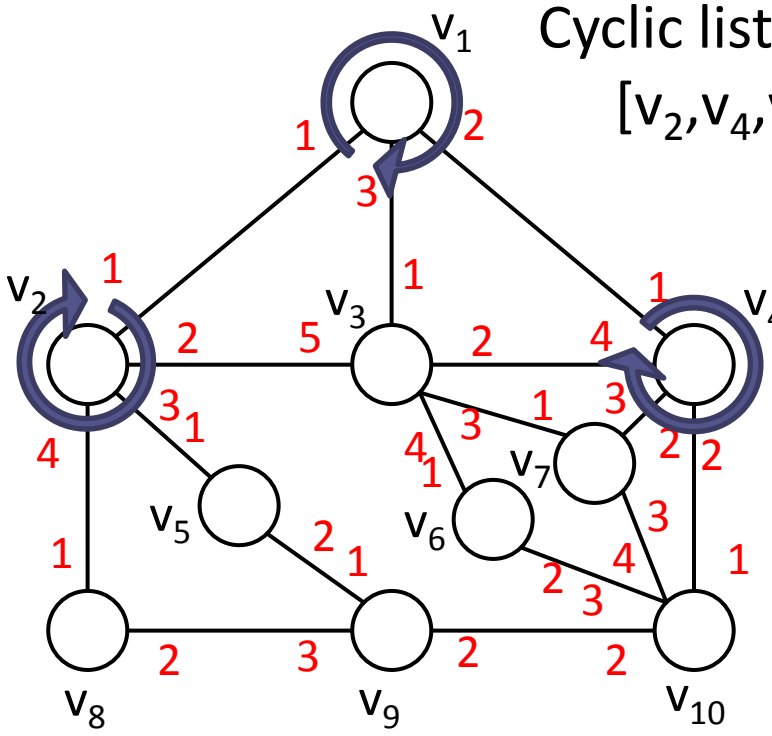
Ordered graph isomorphism[X. Jiang and H. Bunke,1998]

Matching algorithm for planar map[S. Kawamoto et al.,2010]

Ordered graph

Each vertex has a unique order based on neighboring vertices.

Cyclic list of v_2
$[v_1, v_3, v_5, v_8]$



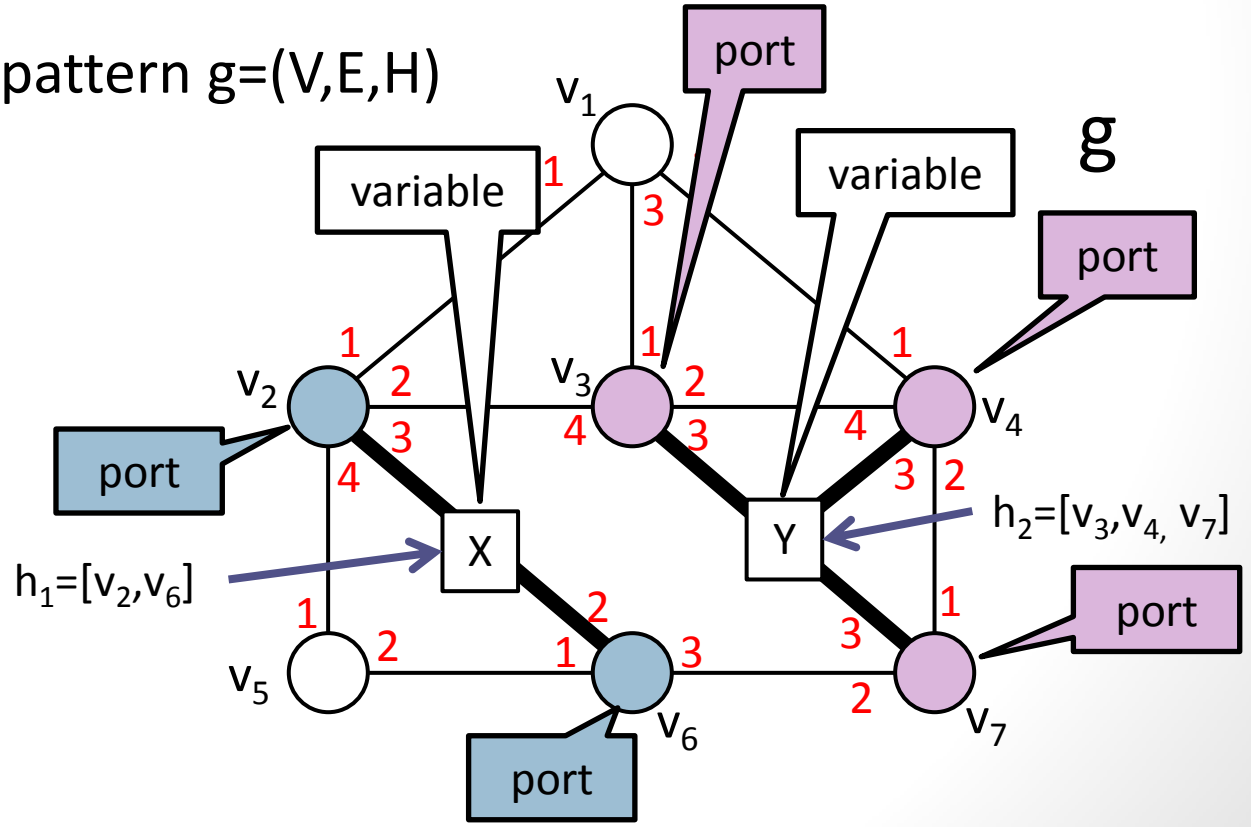
Cyclic list of v_1
$[v_2, v_4, v_3]$

Cyclic list of v_4
$[v_1, v_{10}, v_7, v_3]$

Ordered graph pattern

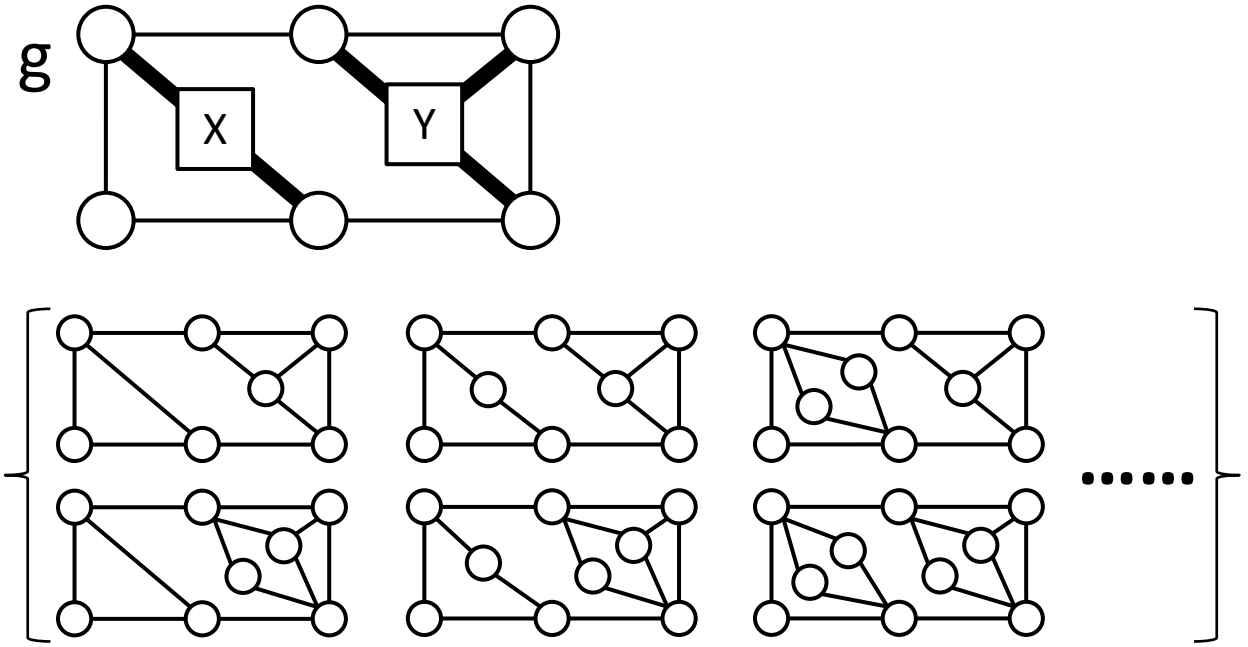
Ordered graph pattern has structural variables and each of whose vertices has an order on neighboring vertices and structural variables.

Ordered graph pattern g=(V,E,H)
V: vertex set
E: edge set
H: variable set



Ordered graph pattern language

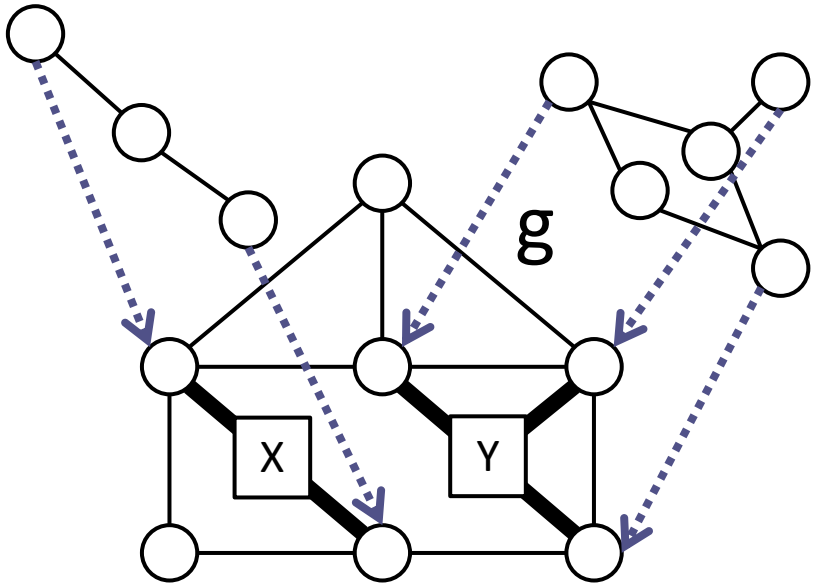
the set of all ordered graphs obtained from an ordered graph pattern by arbitrary substitutions.

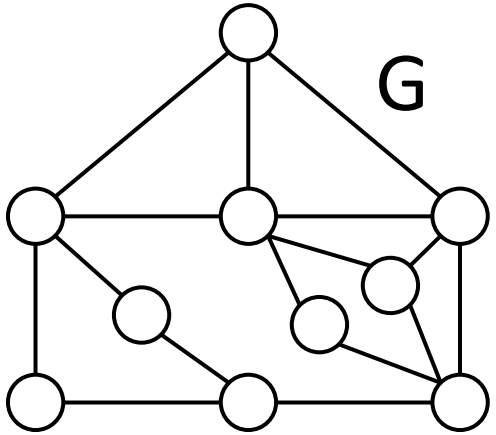


Membership problem

Instance : An ordered graph G and
an ordered graph pattern g.

Question: Does L(g) contain G?





Yes, L(g) contains G

Polynomial Time Matching Algorithm for Ordered Graph Patterns

Matching algorithm

Input: ordered graph G, ordered graph pattern g.

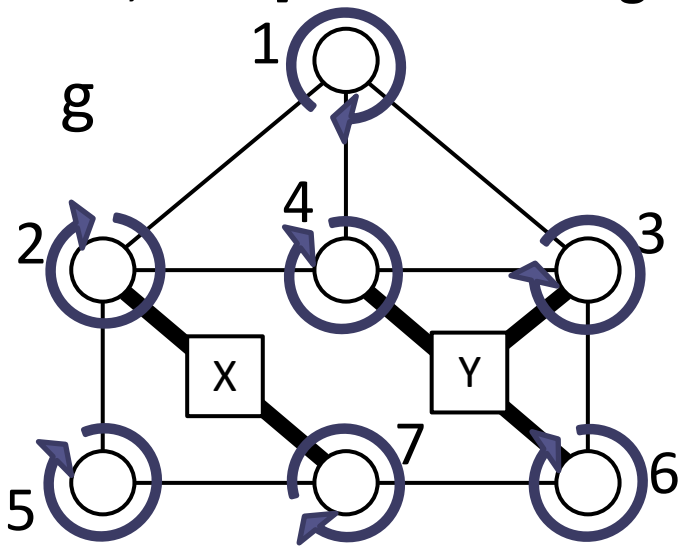
begin

1. Let u be a vertex of g and {u, v} in cyclic list of u;
2. C :=Coding({u,v} g);
3. for each w in V (G) do
4. foreach e in cyclic list of w do
5. if CodeMatch(C,e,G) and StructureMatch (C,g,G)
6. then return true;
7. return false

end.

Making the code of ordered graph pattern

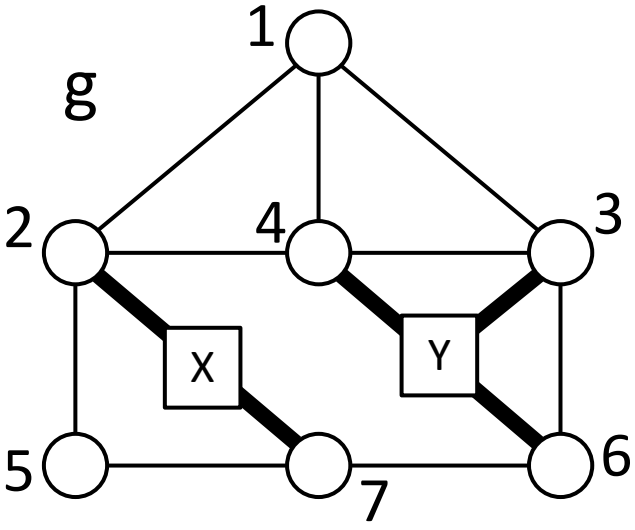
This procedure visits all vertices of the input ordered graph pattern g and attaches the ID of each vertex to the code of g by modifying polynomial time coding algorithm [X.Jiang and H.Bunke,1998] for ordered graphs.

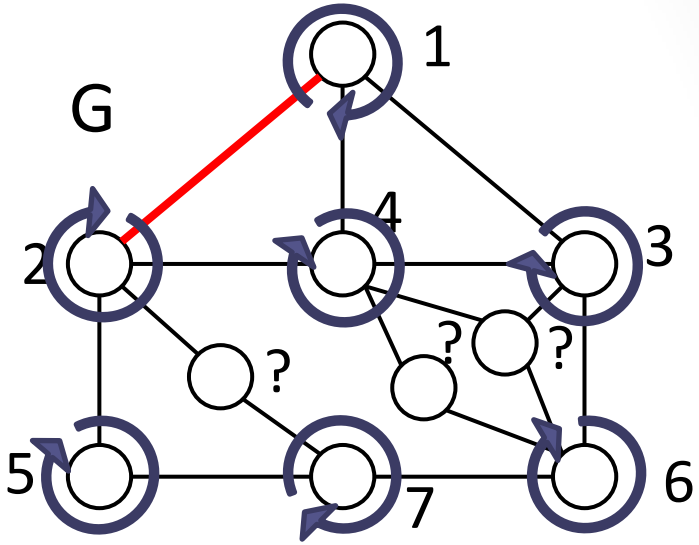


Code(g)=#234#14x5#16y4#13y2#27#37y#5x6#

CodeMatch

This procedure checks whether or not the code of g matches a code of G.





Code(g)=#234#14x5#16y4#13y2# 27#37y# 5x6#

Code(G)=#234#14?5#16?4#13??2#27#37??#5?6#

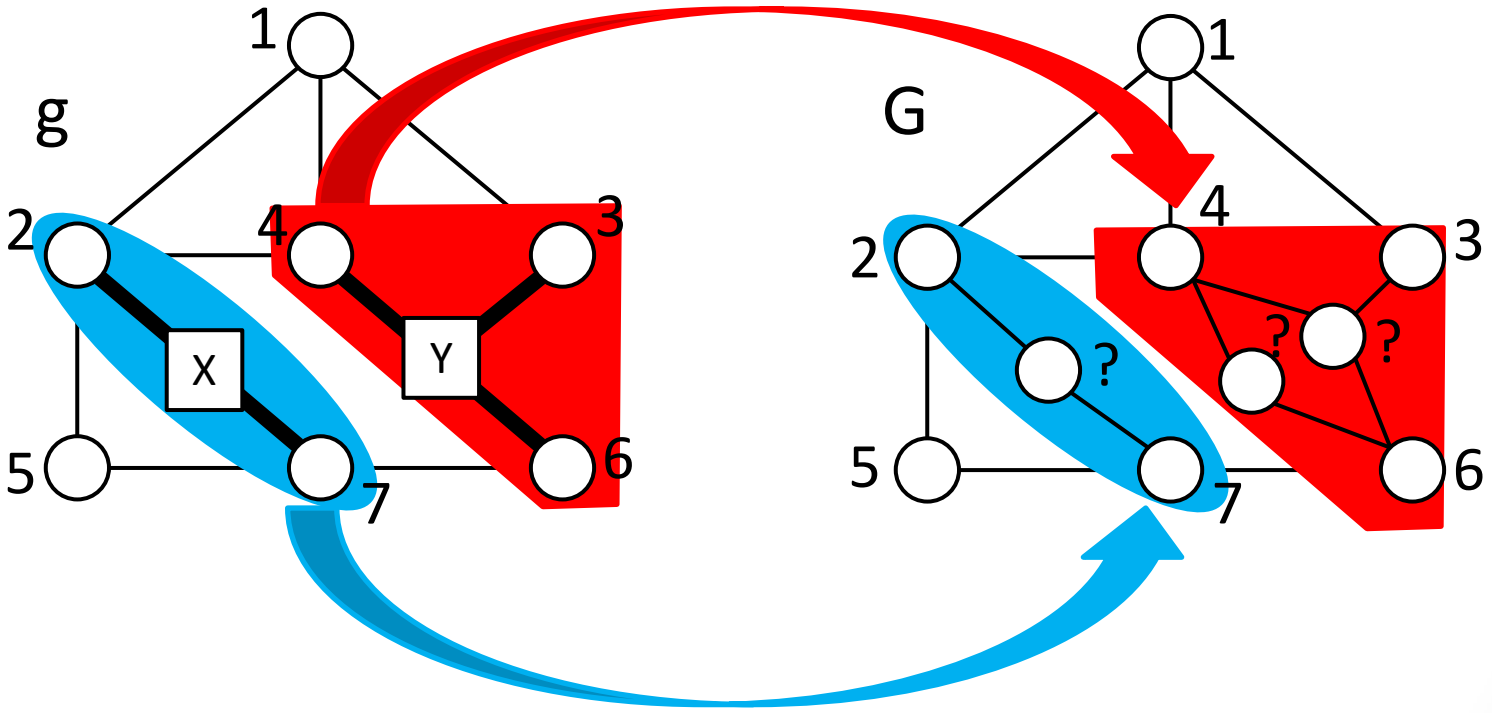
Replace ?+ with x or y

Code(G)=#234#14x5#16y4#13y2#27#37y#5x6#

Match

StructureMatch

This procedure StructureMatch determines whether or not g and G match on graph structure.



Focus on variables of g and edge induced subgraphs defined by the set of edges incident to the vertices labeled with '?' in G.

Membership problem

Theorem 1

For an ordered graph pattern g and an ordered graph G, the membership problem for g and G is solvable in polynomial time.

The total time of the algorithm is $O(|V(g)|*|E(G)|^2)$.

$|V(g)|$:The number of vertices of ordered graph pattern g.

$|E(G)|$:The number of edges of ordered graph G.

Conclusion

We have proposed an ordered graph pattern as a new graph pattern and a polynomial time matching algorithm to solve the membership problem for ordered graph patterns.

Future work

- If we can present a polynomial time algorithm for solving the MINL problem, we can prove that the class of ordered graph pattern languages is polynomial time inductively inferable from positive data.
- Applications for graph mining from graph structured data such as Web pages, TEX files, CAD data, MAP.