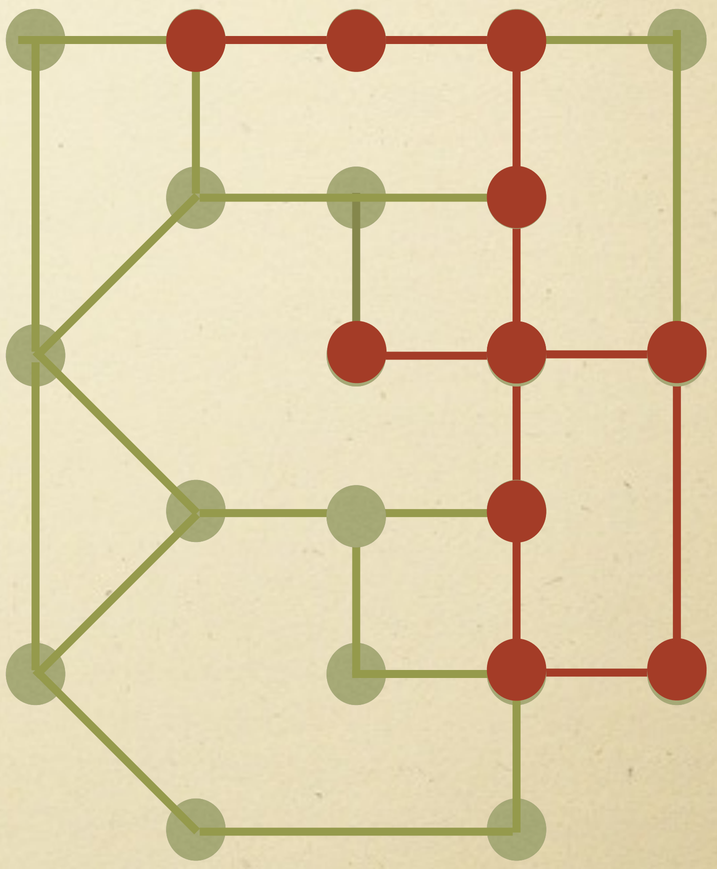
Graph homomorphism and similarity measures

Eunjung Kim, LAMSADE/CNRS
ESIGMA Kick-off meeting, 31 May 2018

Walk



Graph kernel method

graphs

(feature) vectors

mapping ϕ →

$[2,3,5,5,6,3,2,0,4]$

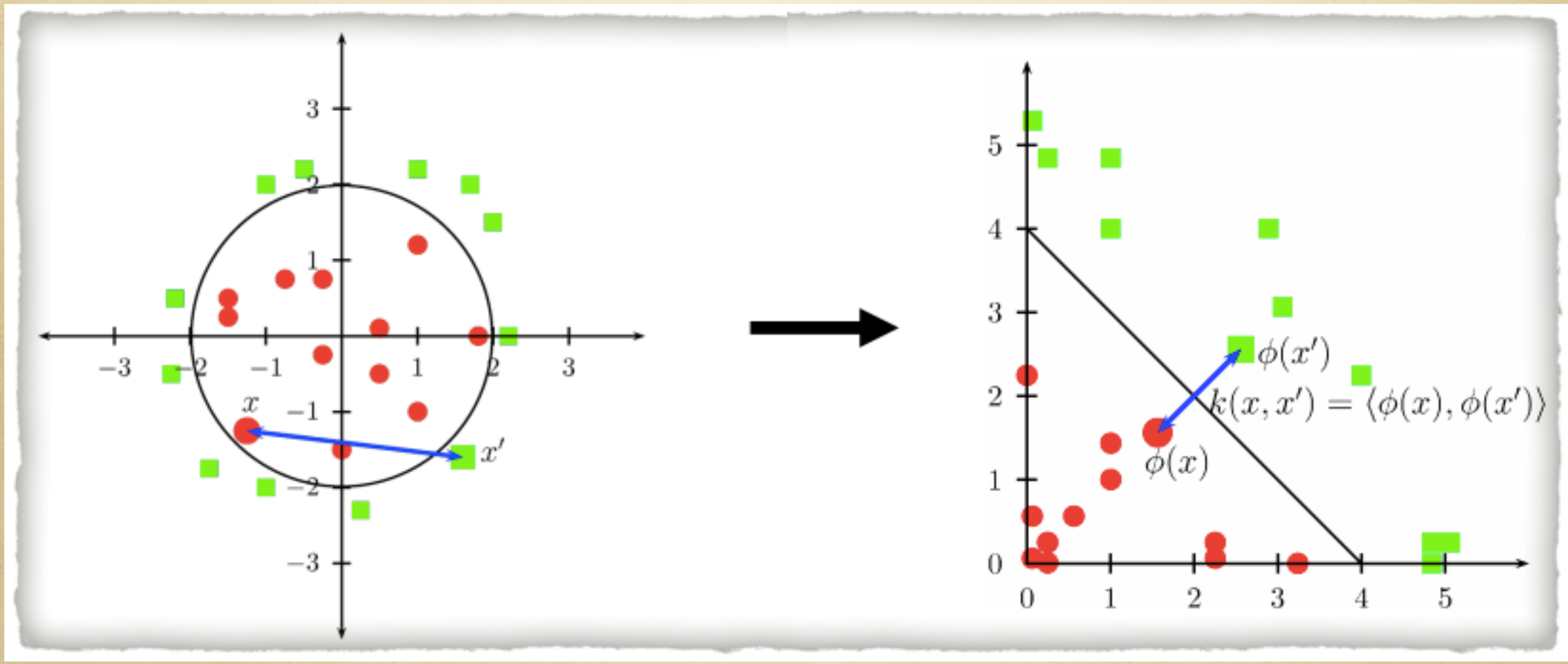
$[1,3,4,2,5,3,4,0,0]$

dist???()

$k(G,H)=\langle\phi(G),\phi(H)\rangle$

* kernel = inner product on some space,
which in turn defines a metric via $d(x,y):=\sqrt{\langle x-y,x-y\rangle}$.

- Easy computation: we can avoid actually defining an explicit mapping, computing the mapping, or doing coordinate-wise computation in the feature space.
- Good math make it work: if a function k meets some criteria (p.d.kernel), then it's an inner product of some space and a distance is well-defined. ()

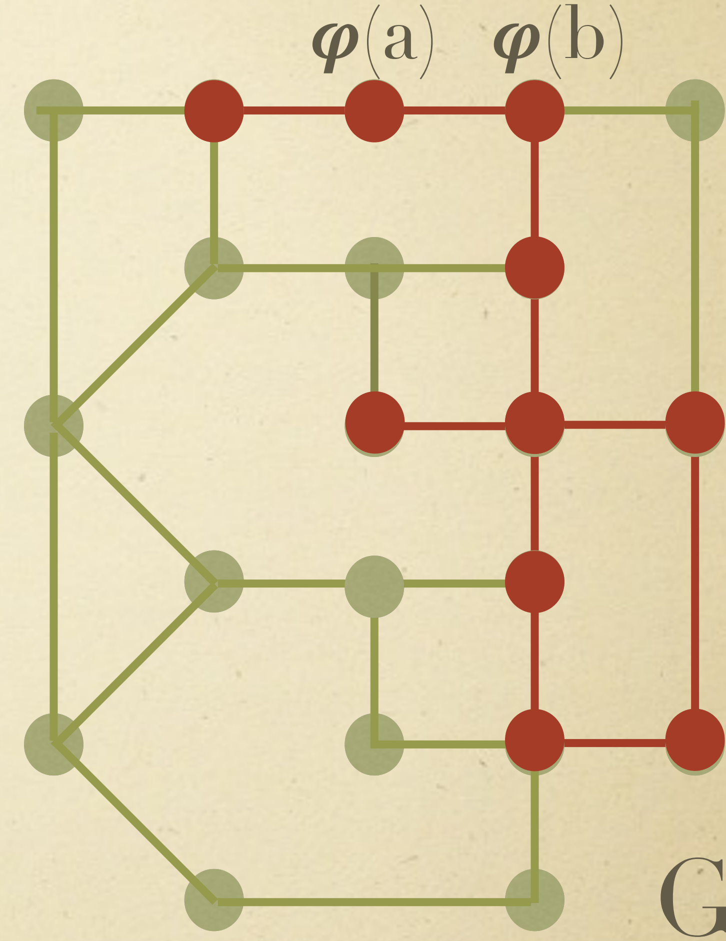


[Borgwardt and Stegle's slide "An introduction to Graph Kernes" 2010]

Graph kernel (via random walk)

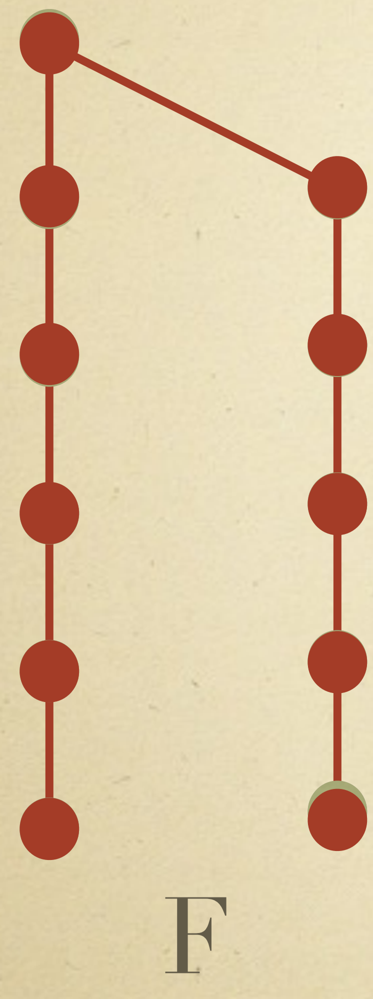
Idea, as much as I get from Ioannis' talk:

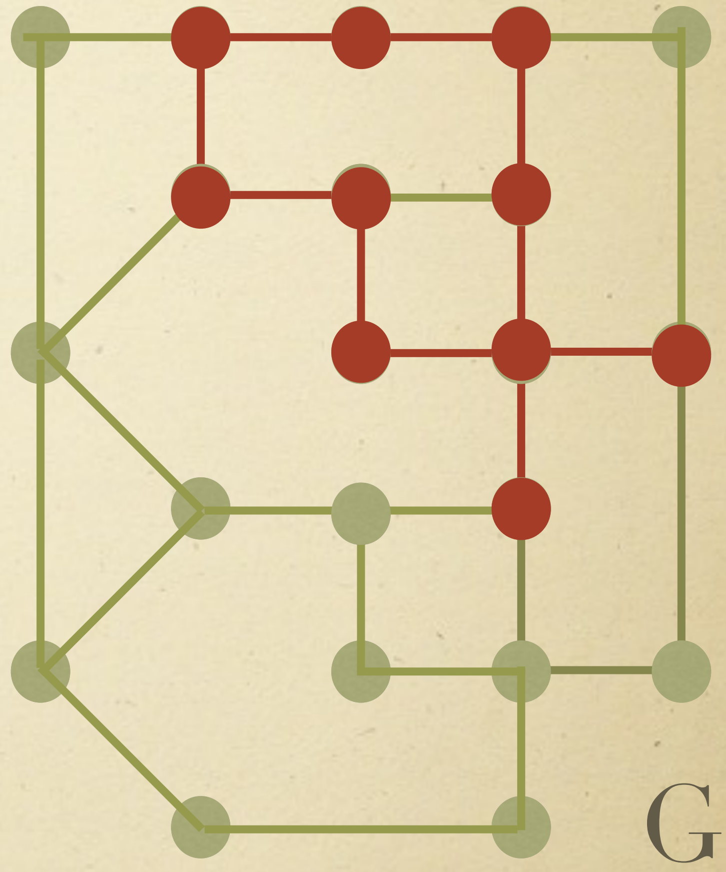
Homomorphism



mapping $\boldsymbol{\varphi}$: V(F)→V(G)
s.t. (a,b) implies ($\boldsymbol{\varphi}$(a),$\boldsymbol{\varphi}$(b))

Homomorphism





mapping φ: V(F)→V(G) s.t. (a,b) implies (φ(a),φ(b))

Isomorphism via homomorphism

G and H are isomorphic
if and only if
#HOM(F,G) = #HOM(F,H) for every graph F

László Lavász, "Operations with structures"
. Acta Mathematica Hungarica 1967

Homomorphism vector

G

HOM(G)=[#HOM(F1,G), #HOM(F2,G), #HOM(F3,G), ...]

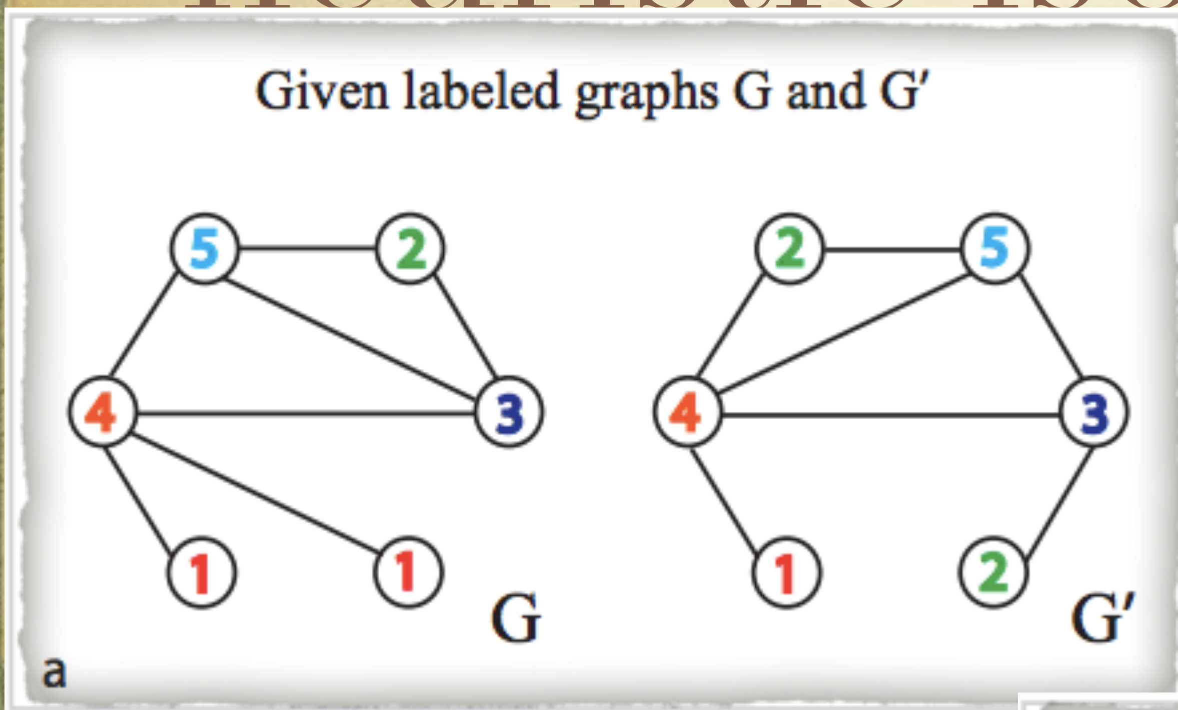
≠

H

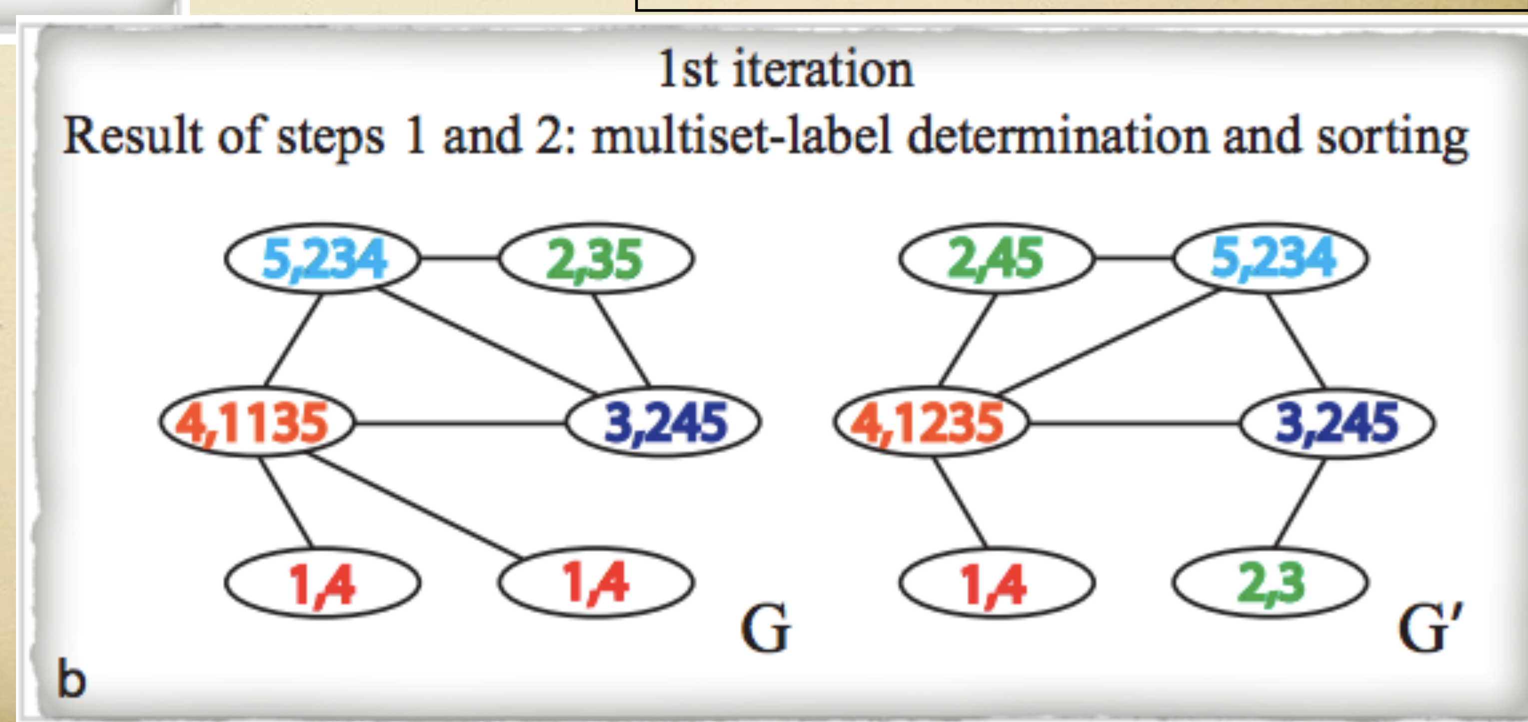
HOM(H)=[#HOM(F1,H), #HOM(F2,H), #HOM(F3,H), ...]

G and H are isomorphic
if and only if
HOM(F,G)=HOM(F,H) for every graph F

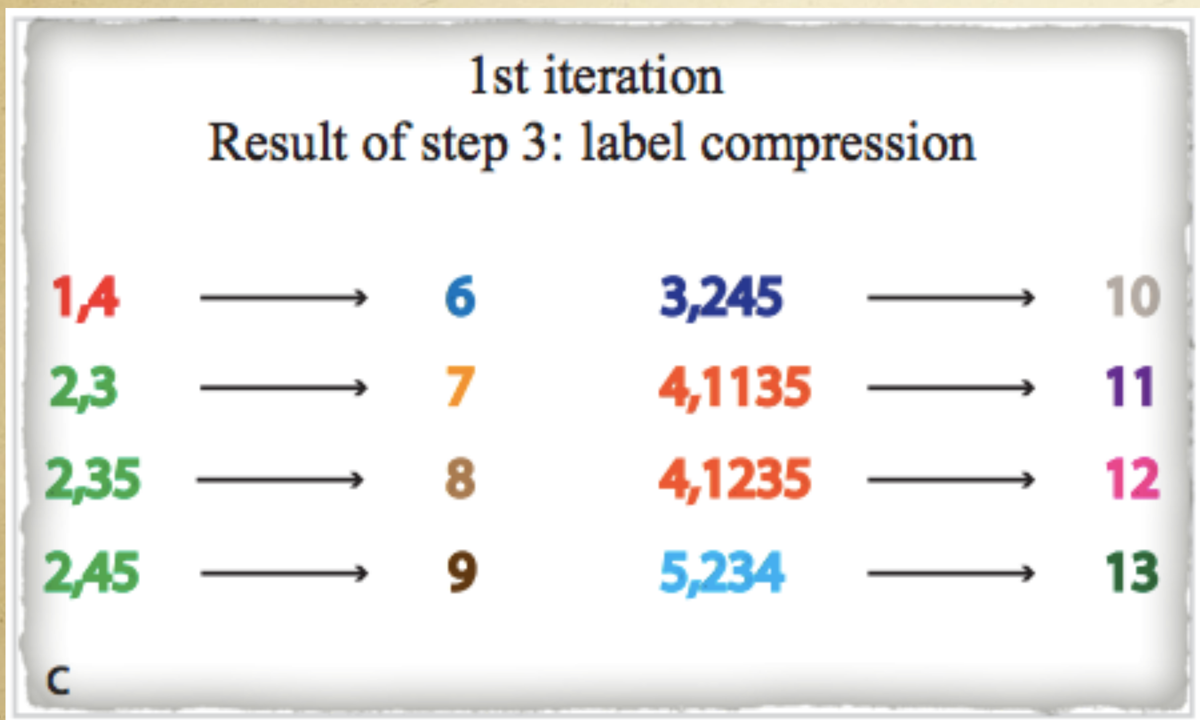
Color-refinement: heuristic isomorphism test



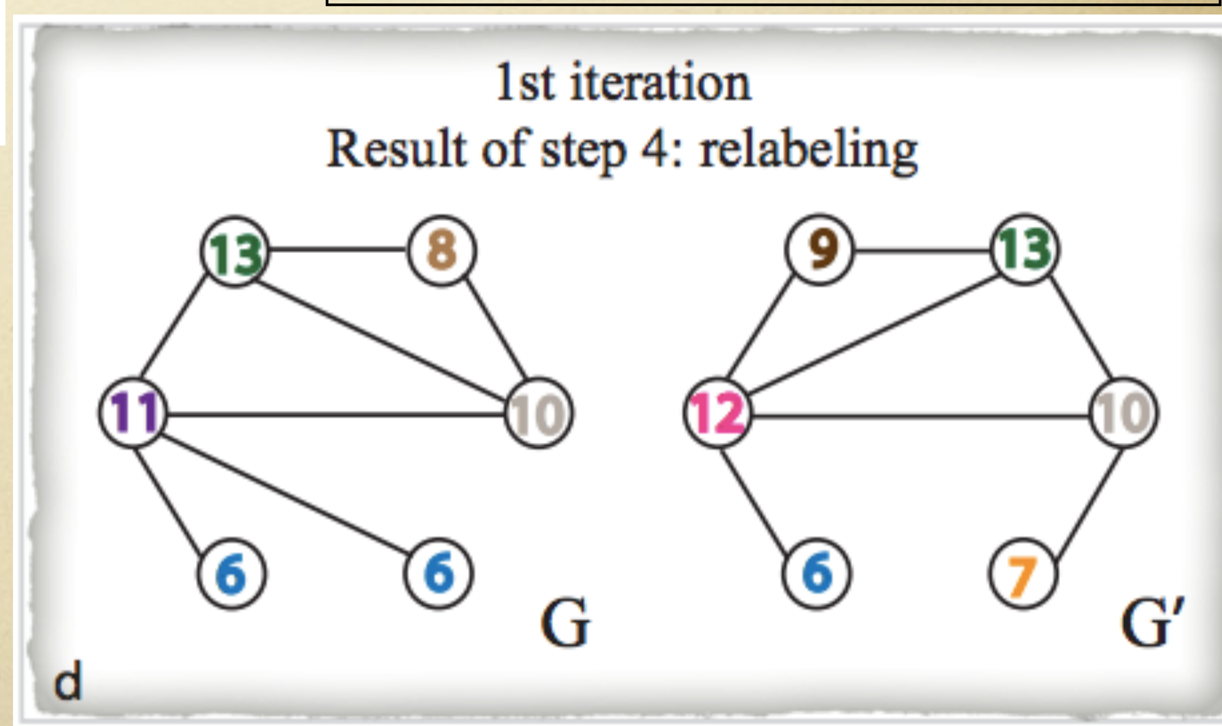
Figures excerpted from Shervashidze et. al. 2011



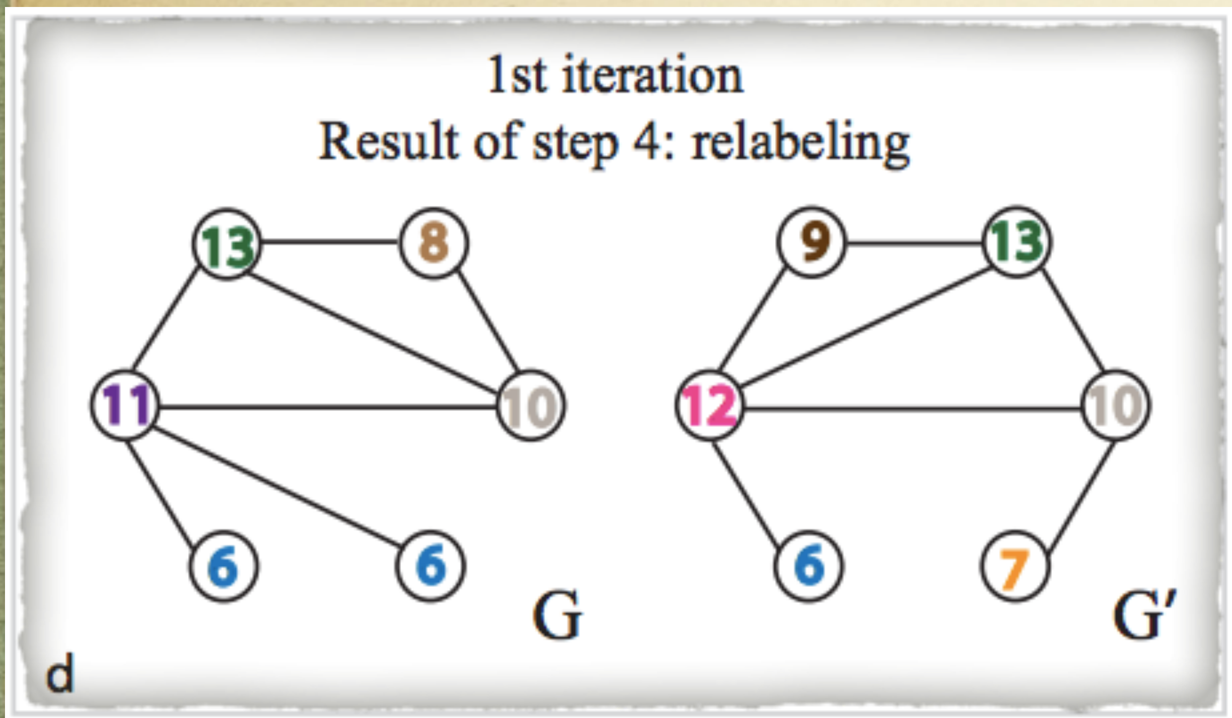
Color-refinement



Figures excerpted from Shervashidze et. al. 2011



Color-refinement



Figures excerpted from Shervashidze et. al. 2011

End of the 1st iteration

Feature vector representations of G and G′

$$\varphi^{(1)}_{WLsubtree}(G) = (2, 1, 1, 1, 1, 2, 0, 1, 0, 1, 1, 0, 1)$$

$$\varphi^{(1)}_{WLsubtree}(G') = (1, 2, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1)$$

Counts of original node labels

Counts of compressed node labels

$$k^{(1)}_{WLsubtree}(G,G') = \langle \varphi^{(1)}_{WLsubtree}(G), \varphi^{(1)}_{WLsubtree}(G') \rangle = 11.$$

e

Color-refinement

color-refinement algorithm

ultimate isomorphism test

G

H

What is this vector?

HOM(G)|all graphs

≠

HOM(H)|all graphs

Indistinguishable under color-refinement

G and H are indistinguishable under color-refinement
if and only if
HOM(F,G)=HOM(F,H) for every tree F

Dell, Grohe and Rattan, "Lovász Meets Weisfeiler and Leman"
ICALP 2018

Color-refinement from Lovász' perspective

	color-refinement algorithm	ultimate isomorphism test
G	$\mathrm{HOM}(G)\mid_{\text{all trees}}$	$\mathrm{HOM}(G)\mid_{\text{all graphs}}$
	$=$	$\neq$
H	$\mathrm{HOM}(H)\mid_{\text{all trees}}$	$\mathrm{HOM}(H)\mid_{\text{all graphs}}$

Color-refinement from Lovász' perspective

	1-Weisfeiler-Leman algorithm (color-refinement)	k-Weisfeiler-Leman algorithm	n-Weisfeiler-Leman algorithm
G	HOM(G)|all trees	HOM(G)| tw ≤ k	HOM(G)|all graphs
			$\neq$
H	HOM(H)|all trees	HOM(H)| tw ≤ k	HOM(H)|all graphs

DGR 2018

Take-home message

"State-of-the-art graph kernel methods work for a good reason, now with theoretical support."

"How can we use it?"

Future Work?