Graph-based Methods in Pattern Recognition and Document Image Analysis (GMPRDIA)

Tutorial at the 15th IAPR International Conference on Document Analysis and Recognition (ICDAR2019)

Saturday 21st September 2019, University of Technology Sydney (UTS)

http://gmprdia.univ-lr.fr

GMPRDIA 2019
Organizing Committee and Speakers

Muhammad Muzzamil LUQMAN

L3i, La Rochelle University, France

Pau RIBA

CVC, Barcelona, Spain

Anjan DUTTA

University of Exeter, UK

http://gmprdia.univ-lr.fr

Centre de Visió per Computador

Informatique Image Interaction Univ. La Rochelle

UNIVERSITY OF EXETER

ICDAR 2019

15th International Conference on Document Analysis and Recognition

20-25 September 2019 | International Convention Centre Sydney, Australia

GMPRDIA 2019
Organizing Committee and Speakers

Muhammad Muzzamil LUQMAN

L3i, La Rochelle University, France

- Research Scientist (Permanent)
- Ph.D. in Computer Science from François Rabelais University of Tours (France) and Autonoma University of Barcelona (Spain).
- Ph.D. thesis titles “Fuzzy Multilevel Graph Embedding for Recognition, Indexing and Retrieval of Graphic Document Images”.
- Research interests
  - Structural Pattern Recognition
  - Document Image Analysis
  - Camera-Based Document Analysis and Recognition
  - Graphics Recognition
  - Artificial Intelligence / Machine Learning
- http://pageperso.univ-lr.fr/muhammad_muzzamil.luqman

Informatique Image Interaction
Univ. La Rochelle

GMPRDIA 2019
Organizing Committee and Speakers

Pau RIBA
CVC, Barcelona, Spain

- Ph.D. student in Computer Science from the Computer Vision Center (CVC, Barcelona) under supervision of Josep Llados (since October 2016)
- Research interests:
    - Graph-based representation for visual objects
    - Graph-based algorithms for solving various tasks in Computer Vision Pattern Recognition and Machine Learning
    - Machine Learning
- http://www.cvc.uab.es/people/priba/
    - https://github.com/priba

Centre de Visió per Computador

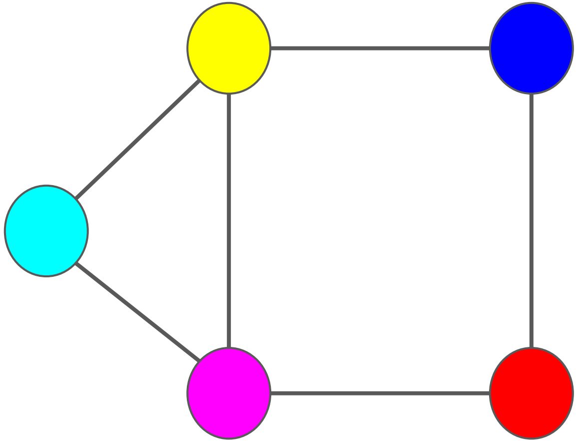
GMPRDIA 2019
Organizing Committee and Speakers

Anjan DUTTA
University of Exeter, UK

- Lecturer (Assistant Professor) in Computer Vision & Machine Learning
- Until July 2019, he was a Marie-Curie postdoctoral fellow under the P-SPHERE project at the Computer Vision Centre, Barcelona, Spain.
- Ph.D. in Computer Science from the Universitat Autònoma de Barcelona (UAB) in the year of 2014.
- Ph.D. thesis titled “Inexact Subgraph Matching Applied to Symbol Spotting in Graphical Documents”
- Research interests
  - graph-based representation for visual objects
  - graph-based algorithms for solving various tasks in Computer Vision, Pattern Recognition and Machine Learning
- https://sites.google.com/site/2adutta/home

Question 1/3
How a graph is represented in computer memory?

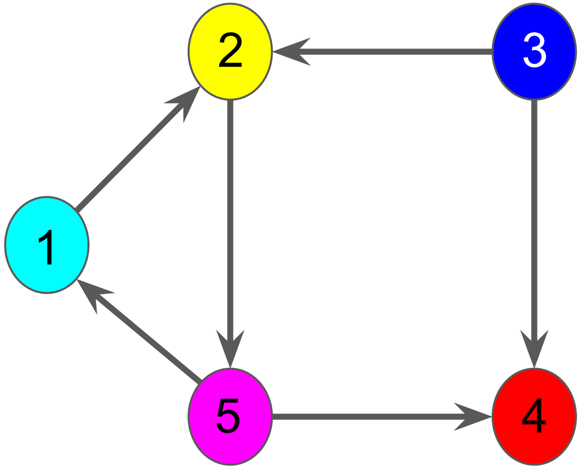
- Type of graph representation in computer memory
- There are two ways:
  - Sequential representation
  - Linked representation



Question 1/3
How a graph is represented in computer memory?

Sequential representation

- Adjacency matrix

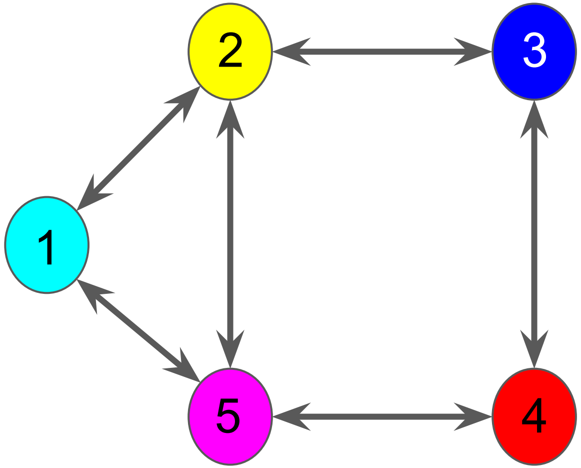


$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Question 1/3
How a graph is represented in computer memory?

Sequential representation

- Adjacency matrix

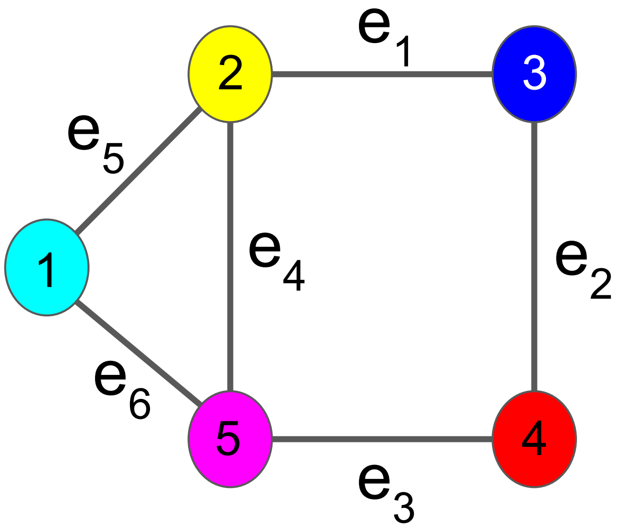


$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Question 1/3
How a graph is represented in computer memory?

Sequential representation

- Incidence matrix

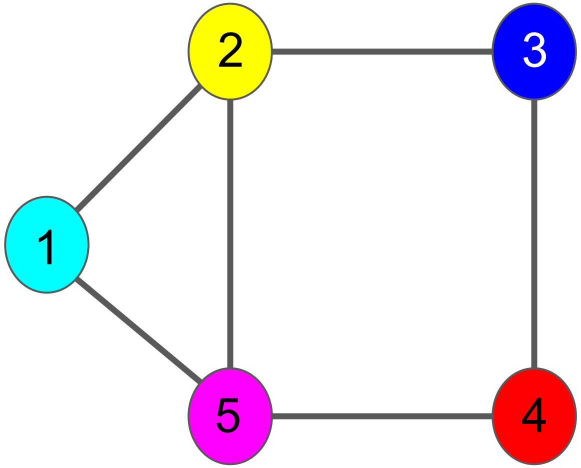


$$I = \begin{array}{c} \begin{array}{cccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Question 1/3
How a graph is represented in computer memory?

Linked representation

- Adjacency list



$$A = \begin{array}{llll} \mathbf{1} & 2 & 5 & \\ \mathbf{2} & 1 & 3 & 5 \\ \mathbf{3} & 2 & 4 & \\ \mathbf{4} & 3 & 5 & \\ \mathbf{5} & 1 & 2 & 4 \end{array}$$

Question 2/3
How graphs are stored on disk?


```xml
<?xml version="1.0" encoding="UTF-8" ?>
<descriptors generationTime="1421079803.32" nbRegion="15">
 <region compactness="0.292119565217" idRegion="0" nbPixel="539">
  <colorRgb b="198" g="198" r="206"/>
  <colorLab L="204" a="129" b="124"/>
  <boundingBox height="23" width="32" x="109" y="165"/>
  <moments>0.182159670549,0.00546902073257,5.53115452371e-05,3.23028031097e-06,2.35327007343e-11,-1.04512269946e-07,3.62022171482e-11</moments>
  <adjacentRegion>
   <id id="3">
    <Percentage val="0.547266107331"/>
   </id>
   <id id="8">
    <Percentage val="0.61328705673"/>
   </id>
   <id id="9">
    <Percentage val="0.538285151261"/>
   </id>
   <id id="10">
    <Percentage val="0.324987337731"/>
   </id>
   <id id="11">
    <Percentage val="0.231954249729"/>
   </id>
   <id id="13">
    <Percentage val="0.371791495215"/>
   </id>
   <id id="14">
    <Percentage val="0.244609544714"/>
   </id>
  </adjacentRegion>
 </region>
...
```


SSGCI competition (http://ssgci.univ-lr.fr)

Le, T.N., Luqman, M.M., Dutta, A., Héroux, P., Rigaud, C., Guérin, C., Foggia, P., Burie, J.C., Ogier, J.M., Lladós, J. and Adam, S., 2018. Subgraph spotting in graph representations of comic book images. Pattern Recognition Letters, 112, pp.118-124.

Question 3/3
In which languages can i program/code a graph-based method?

Matlab

MatlabBGL, Graph and Network algorithms, GAIMC, ...

Python

Networkx, igraph, ...

C/C++

Boost Graph Library, ...

and many others ...

Saturday 21st september 2019
09h00 – 12h30

Part-1

- **A historic perspective of graph-based methods in PR & DIA**
- **Neural Networks on graphs and modern trends in graph-based PR & DIA**

Coffee break (10h30 - 11h00)

Part-2

- Applications of Graph Neural Networks
    - Learning Graph Distances
    - Table Detection
- Hands-on
    - Deep Graph Library

http://gmprdia.univ-lr.fr

Centre de Visió per Computador

Informatique Image Interaction
Univ. La Rochelle

UNIVERSITY OF EXETER

ICDAR 2019
15th International Conference on Document Analysis and Recognition
20-25 September 2019 | International Convention Centre Sydney, Australia

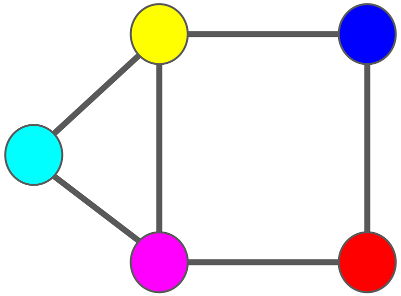
Structural and Statistical Pattern Recognition

	Pattern Recognition	
	Structural	Statistical
Data structure	symbolic data structure	numeric feature vector
Representational strength	Yes	No
Fixed dimensionality	No	Yes
Sensitivity to noise	Yes	No
Efficient computational tools	No	Yes

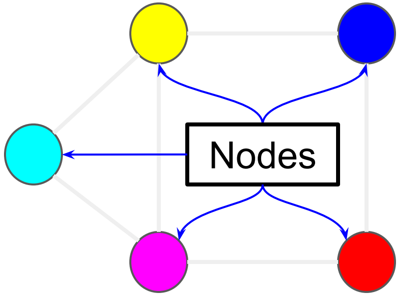
How images (and/or other types of content) are represented by graphs?

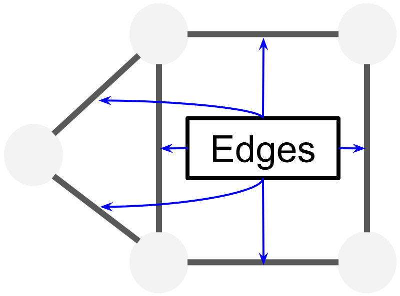
Graph

- A *graph* $G = (V, E)$ is a mathematical structure for representing relationships.

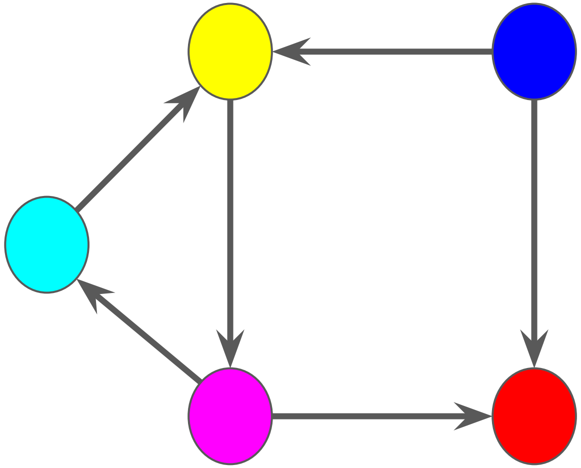


- A graph $G = (V, E)$ consists of a set of *nodes* V connected by *edges* E.

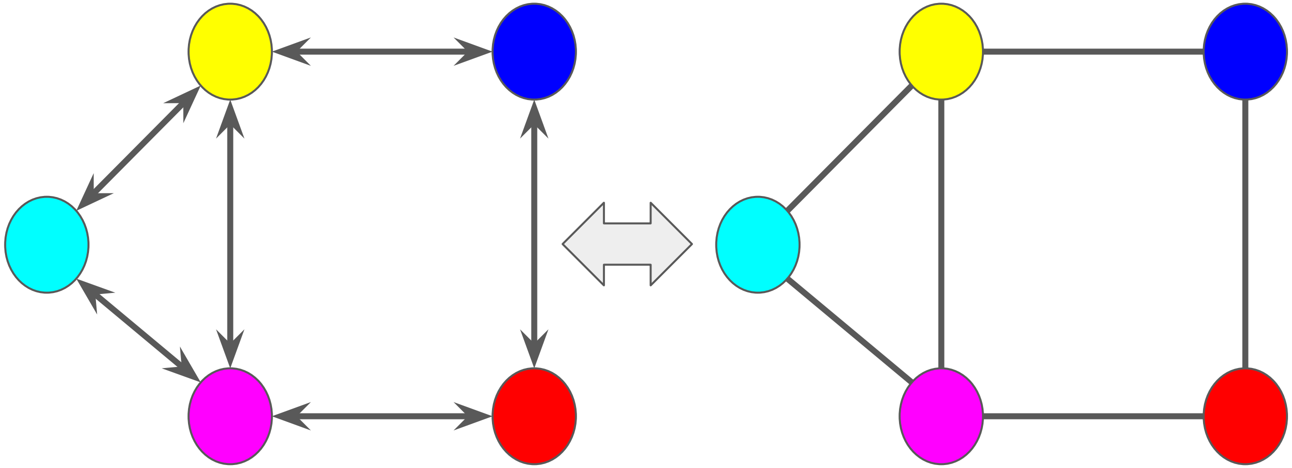




Directed and Undirected Graph



Directed Graph

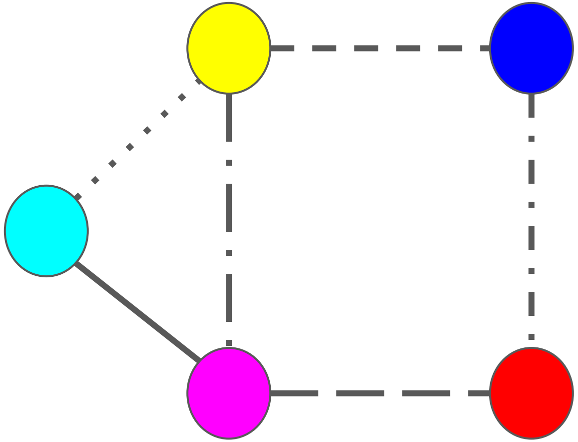


Undirected Graph

Attributed Graph

An attributed Graph is a 4-tuple $G = (V, E, \alpha, \beta)$

- Set of nodes V
- Set of edges $E \subseteq V \times V$
- Node attribute function $\alpha : V \to L_V$
- Edge attribute function $\beta : E \to L_E$



Graph Representation: Issues to Consider

Graph representation of objects depends on:

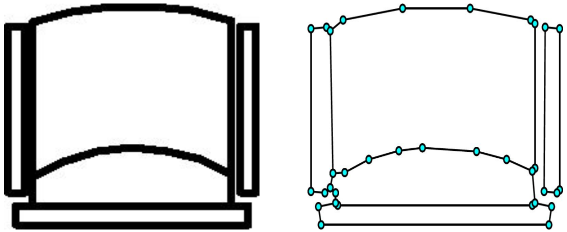
1. Problem definition
2. Type of solution / methodology
3. Stability and noise tolerance

Discriminant units of information in an underlying image for representing it by a graph

- Critical Points
- Line Segments
- Homogeneous Regions
- Keypoints
- Convex Regions
- etc.

Critical Points

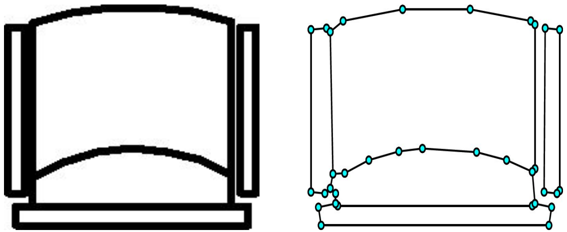
- Critical points from skeleton or edge analysis as nodes.
- Type of edges:
  - Adjacency
  - Proximity
  - k-NN
  - Delaunay triangulation
- Example
  - Symbol spotting by hashing serialized subgraphs.
  - Critical points as nodes and their connections as edges.



A. Dutta, J. Lladós, and U. Pal. A symbol spotting approach in graphical documents by hashing serialized graphs. In PR, vol. 46, no. 3, pp. 752-768, 2013.

Line Segments

- Line segments from skeleton or edge analysis as nodes.
- Type of edges:
  - Adjacency
  - Proximity
  - k-NN
  - Delaunay triangulation
- Example
  - Subgraph matching applied to symbol spotting.
  - Each line segment as a node and upto 3 nearest neighbors are joined to form edges.



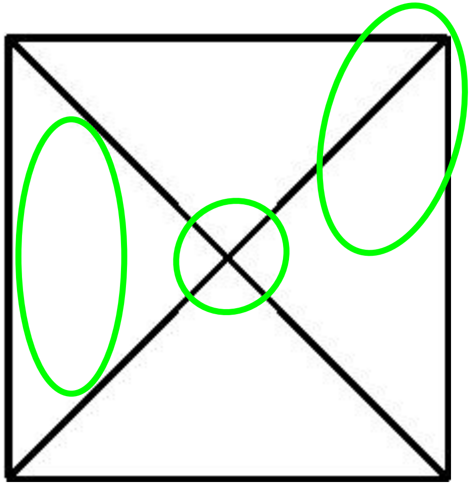
A. Dutta, J. Lladós, H. Bunke and U. Pal. “A Product graph based method for dual subgraph matching applied to symbol spotting". GREC, 2014.

Homogeneous Regions

- Regions either existing or generated by a preprocessing stage as nodes.
- Type of edges:
  - Adjacency
  - Proximity
  - Delaunay triangulation
- Example
  - SSGCI competition, ICPR 2016.
  - RAG of cartoon characters
  - Subgraph spotting

Keypoints

- Detected keypoints using some off-the-shelf algorithm as nodes.
- Type of edges:
  - Proximity
  - k-NN
  - Delaunay triangulation
- Example
  - Symbol recognition.
  - Shape context of detected SIFT interest points.



T. H. Do, S. Tabbone, O. R. Terrades. “Sparse representation over learned dictionary for symbol recognition". SP, pp. 36-47, 2016.

Example: Skeleton Graph

- Skeleton graph
- Each junction or end point as a node of the graph
- Edges are created following the skeleton

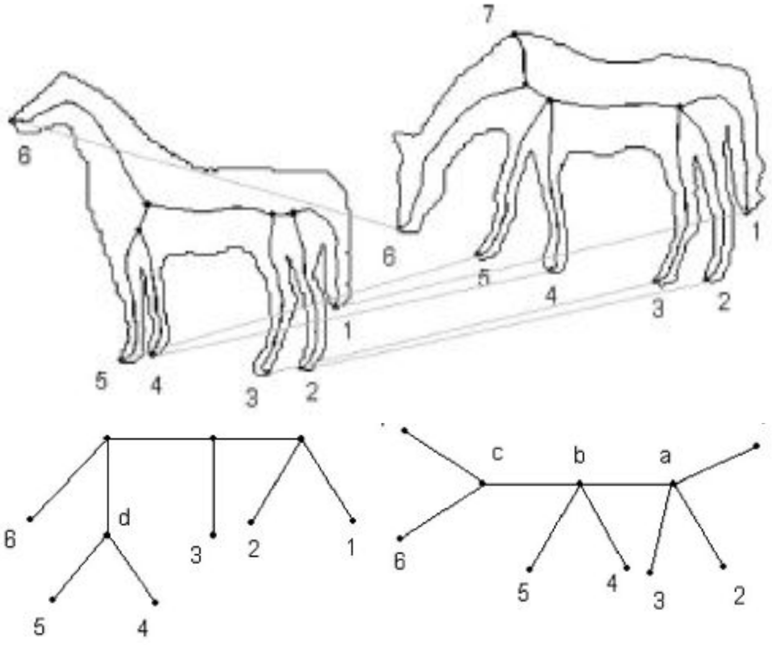


Figure credit: Bai and Latecki PAMI 2008

X. Bai and L. J. Latecki. Path Similarity Skeleton Graph Matching. IEEE TPAMI, vol. 30, no. 7, 1282-1292, 2008.

Example: Region Adjacency Graph

- Region adjacency graph
- Each white region as a node in the graph
- Each pair of adjacent nodes is connected by an edge

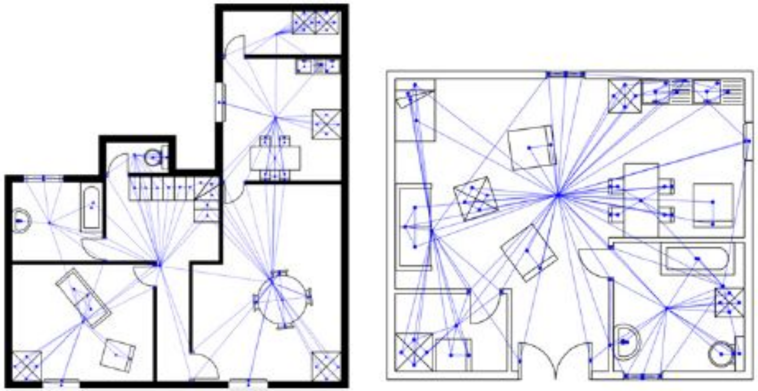


Figure credit: Le Bodic et al 2012

P. L. Bodic, P. Héroux, S. Adam and Y. Lecourtier. An integer linear program for substitution-tolerant subgraph isomorphism and its use for symbol spotting in technical drawings. PR, vol. 45, no. 12, pp. 4214-4224, 2012.

Example: Graph of convexities

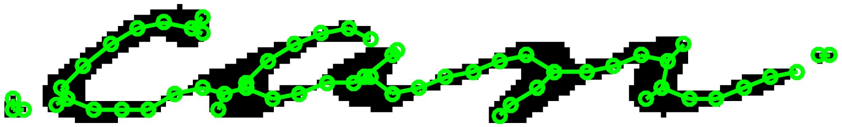
- Convex part segmentation
- Each convex part as node
- Nearest nodes are joined as edges

Figure credit: Riba et al, PRL 2017

P. Riba, J. Lladós, A. Fornés, A. Dutta. Large-scale graph indexing using binary embeddings of node contexts for information spotting in document image databases. PRL, vol. 87, pp. 203 - 211, 2017.

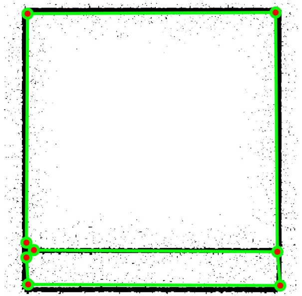
Example: Graph of critical points

- Critical points, grid etc as nodes.
- Adjacent nodes on the writing are joined.
- Normalized coordinates as node attributes



Histograph dataset
(http://www.histograph.ch/)

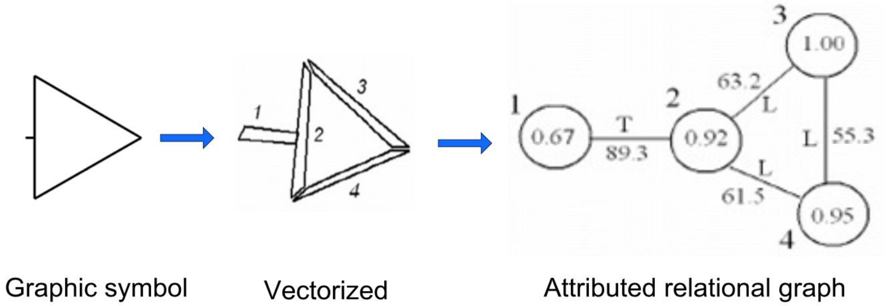
- Critical points as nodes.
- Adjacent nodes on the symbol are joined.
- Coordinate as node attributes.
- Line type as edge attributes.



GREC dataset
(http://www.fki.inf.unibe.ch/databases)

Example: Vecto-Quad graph representation

- Graph representation developed for line drawings
- Each node in the graph represents a line in underlying image
- Thin lines are termed as vectors
- Thick lines or filled shapes are termed as quadrilaterals
- Connections between the vectors/quadrilaterals are represented by edges
- Attributes on nodes as well as edges

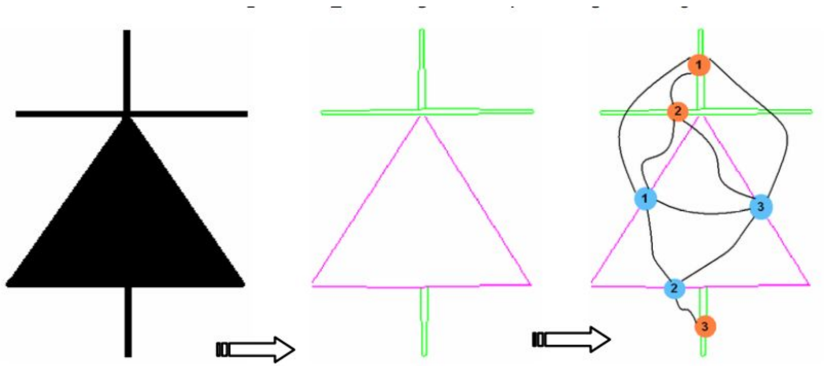


R. Qureshi, J. Ramel, H. Cardot, and P. Mukherji, "Combination of symbolic and statistical features for symbols recognition," in IEEE ICSCN, 2007, pp. 477–482.

J.Y. Ramel, N. Vincent, H. Emptoz, "A structural Representation for understanding line-drawing images", InternationalJournalonDocumentAnalysisandRecognition, vol.3(2),2000,pp.58- 66.

Example: Vecto-Quad graph representation

- Vectors and Quadrilaterals representation well adapted to the underlying line-drawing images

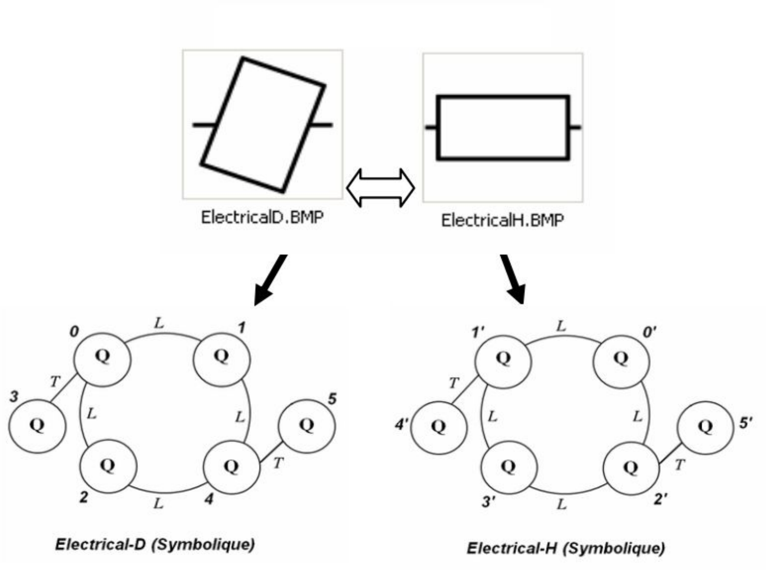


R. Qureshi, J. Ramel, H. Cardot, and P. Mukherji, *“Combination of symbolic and statistical features for symbols recognition,” in IEEE ICSCN, 2007, pp. 477–482.*

J.Y. Ramel, N. Vincent, H. Emptoz, "A structural Representation for understanding line-drawing images", InternationalJournalonDocumentAnalysisandRecognition, vol.3(2),2000,pp.58- 66.

Example: Vecto-Quad graph representation

- Graph-based representations have built-in rotation invariance

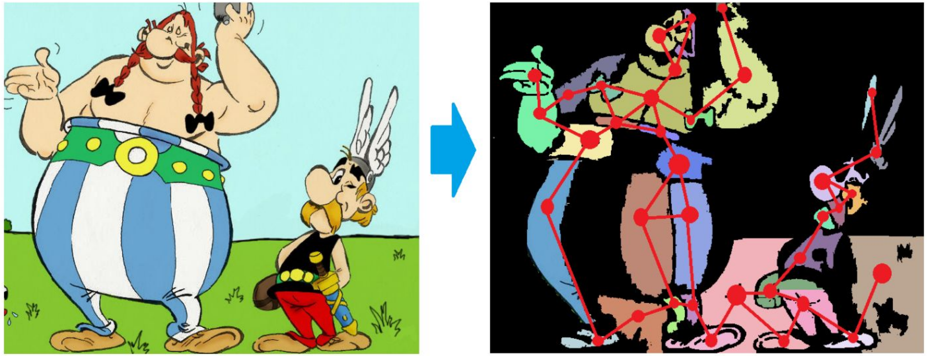


R. Qureshi, J. Ramel, H. Cardot, and P. Mukherji, "Combination of symbolic and statistical features for symbols recognition," in IEEE ICSCN, 2007, pp. 477–482.

J.Y. Ramel, N. Vincent, H. Emptoz, "A structural Representation for understanding line-drawing images", InternationalJournalonDocumentAnalysisandRecognition, vol.3(2),2000,pp.58- 66.

Example: MSER-regions based graph representation

- Graph representation developed for colored comic images
- Each node in graph represents an MSER region in underlying image
- Spatial relations between MSER regions are represented by edges in graph
- Attributes on nodes as well as edges



Thanh-Nam Le, Muhammad Muzzamil Luqman, Jean-Christophe Burie, Jean-Marc Ogier: Content-based comic retrieval using multilayer graph representation and frequent graph mining. ICDAR 2015: 761-765

M. M. Luqman, H. N. Ho, J.-c. Burie, and J.-M. Ogier, "Automatic indexing of comic page images for query by example based focused content retrieval," in 10th 1APR International Workshop on Graphics Recognition, United States, Aug. 2013.

Example: MSER-regions based graph representation

- Multilayer graph representation
  - Color layer
  - Hu-moments layer
  - Compactness layer

Thanh-Nam Le, Muhammad Muzzamil Luqman, Jean-Christophe Burie, Jean-Marc Ogier: Content-based comic retrieval using multilayer graph representation and frequent graph mining. ICDAR 2015: 761-765

M. M. Luqman, H. N. Ho, J.-c. Burie, and J.-M. Ogier, "Automatic indexing of comic page images for query by example based focused content retrieval," in 10th 1APR International Workshop on Graphics Recognition, United States, Aug. 2013.

Learning Graph Representation

- Learning graph that best represent an image for matching to another relevant image
- Fully connected graph of detected key points
- Learning node and edge parameters that prioritize a set of nodes for a particular structure

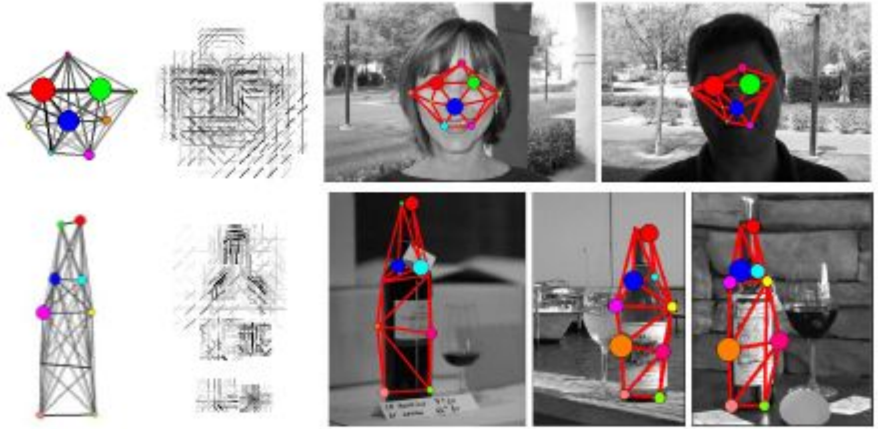


Figure credit: Cho et al 2013

M. Cho, K. Alahari and J. Ponce. Learning Graphs to Match. ICCV, 2013.

How we can/used to solve Pattern Recognition problems using graphs?

A very general overview of historical evolution of graph-based solutions to Pattern Recognition

- Graph matching (isomorphism)

  [Messmer, 1995] [Sonbaty and Ismail, 1998]

- Graph Edit Distance (GED)

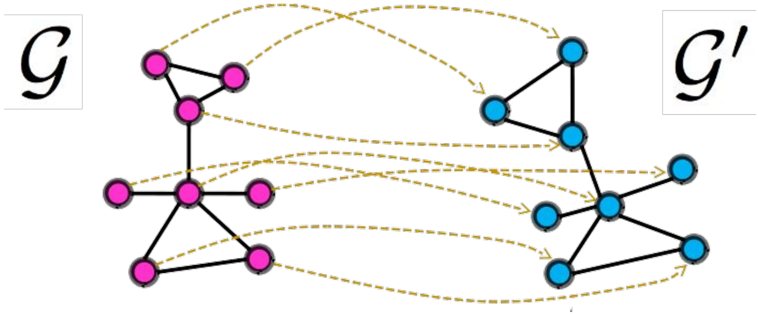
  [Bunke and Shearer, 1998] [Neuhaus and Bunke, 2006]

- Graph EMbedding (GEM)

  [Luqman et al., 2009] [Sidere et al., 2009] [Gibert et al., 2011]

Graph matching

Finding matches (isomorphism) between two graphs.

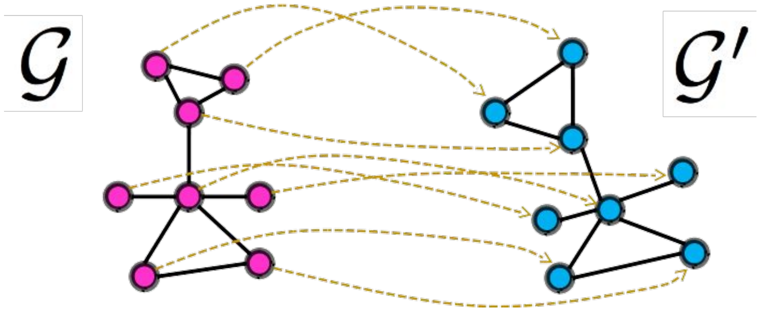


$$\mathbf{X} \in \{0, 1\}^{nn'}$$

- $\mathbf{X}_{ia} = 1$ if node i in G corresponds to node a in G'
- $\mathbf{X}_{ia} = 0$ otherwise

Graph matching

Maximizing the matching score ***S***

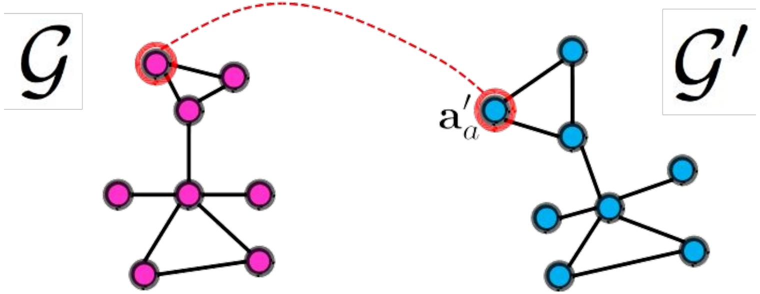


$$\mathbf{X}^* = \arg\max_{\mathbf{X}} S(\mathcal{G}, \mathcal{G}', \mathbf{X})$$

$$\text{s.t.} \begin{cases} \mathbf{X} \in \{0,1\}^{nn'} \\ \sum_{i=1}^{n} \mathbf{X}_{ia} \leq 1, \sum_{a=1}^{n'} \mathbf{X}_{ia} \leq 1 \end{cases}$$

Graph matching

How to measure the matching score S ?

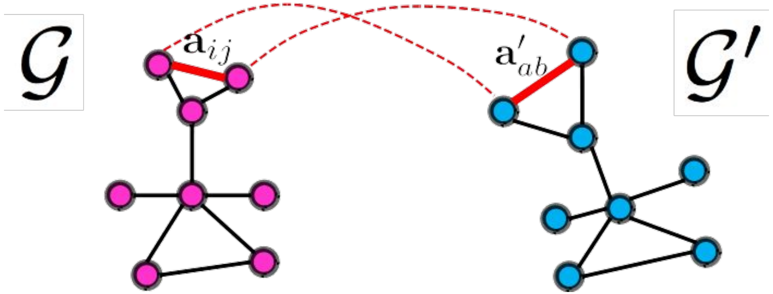


$$S_V(\mathbf{a}_i, \mathbf{a}'_a)$$

- Each node and each edge has its own attribute
- Node similarity function

Graph matching

How to measure the matching score S ?

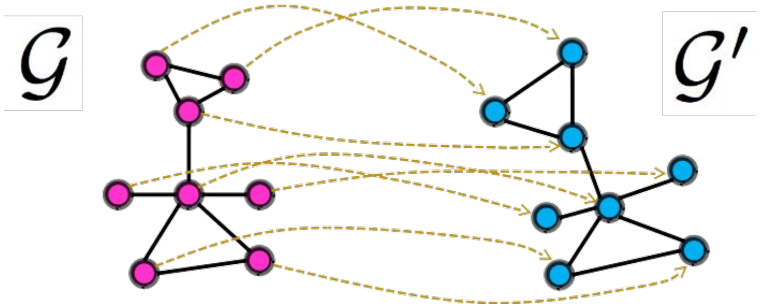


$$S(\mathcal{G}, \mathcal{G}', \mathbf{X}) = \sum_{\mathbf{X}_{ia}=1} S_V(\mathbf{a}_i, \mathbf{a'}_a) + \sum_{\mathbf{X}_{ia}=1, \mathbf{X}_{jb}=1} S_E(\mathbf{a}_{ij}, \mathbf{a'}_{ab})$$

- Sum of S_V and S_E values for the assignment $\mathbf{X}$.

Graph matching

How to measure the matching score S?



- $\mathbf{X}_{ia} = 1$ if node i in $\mathcal{G}$ corresponds to node a in $\mathcal{G}'$
- $\mathbf{X}_{ia} = 0$ otherwise

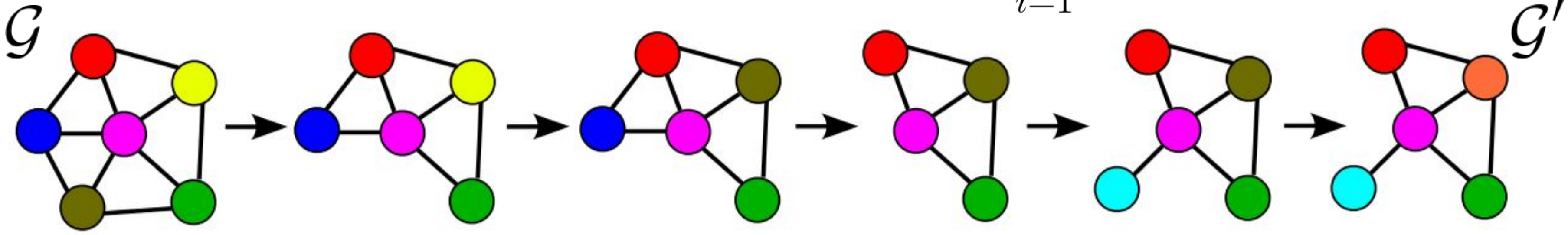
Advances in graph matching

- Quadratic assignment problem
  - NP-hard, thus exact solution is infeasible
- Advances in approximate (inexact) algorithms
  - Error-tolerant (inexact) graph matching
  - Relaxation and Projection

Graph edit distance

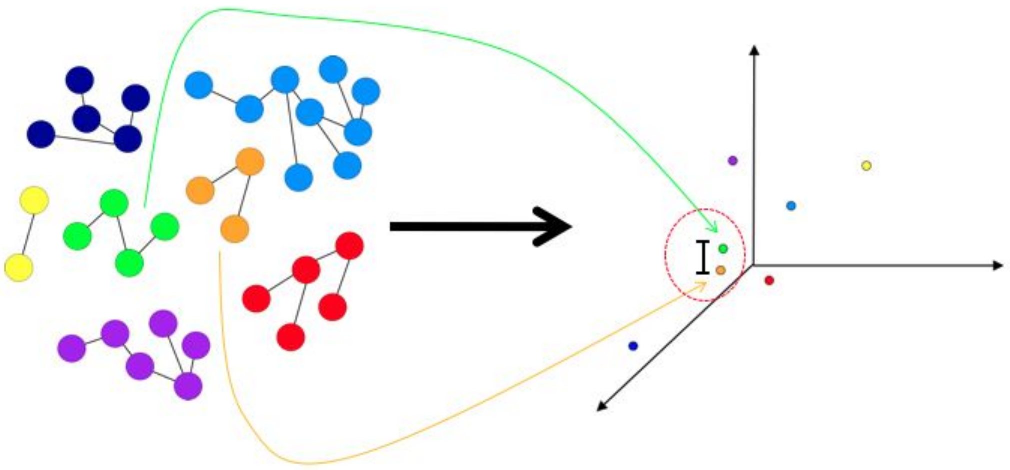
- A measure of similarity between two graphs.
- Node and edge insertion, deletion, substitution.
- Summation of the edit costs

$$\mathrm{GED}(\mathcal{G}, \mathcal{G}') = \min_{(e_1,\ldots,e_k)\in\mathcal{P}(\mathcal{G},\mathcal{G}')} \sum_{i=1}^{k} c(e_i)$$



A. Sanfeliu, K. S. Fu. A distance measure between attributed relational graphs for pattern recognition. IEEE TSMC, vol. 13, no. 3, 1983.

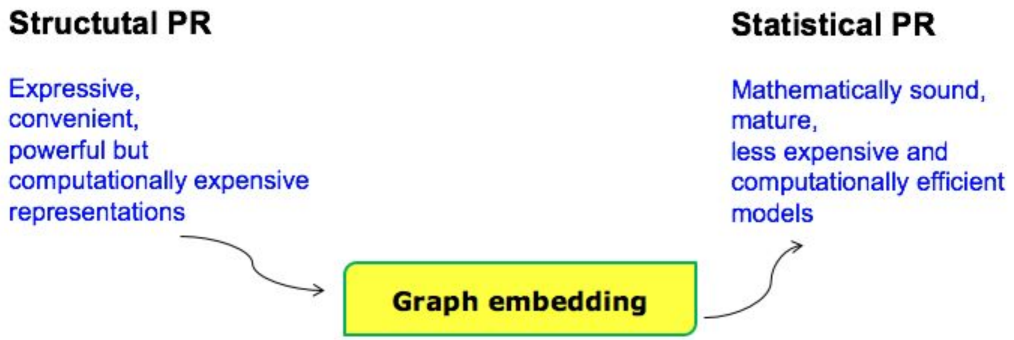
Graph embedding



By mapping a high dimensional graph into a point in suitable vector space, graph embedding permits to perform the basic mathematical computations which are required by various statistical pattern recognition techniques, and offers interesting solutions to the problems of graph clustering and classification.

Luqman, M. M. (2012). Fuzzy Multilevel Graph Embedding for Recognition, Indexing and Retrieval of Graphic Document Images. Ph.D. thesis. University of Tours, France and Autonoma University of Barcelona, Spain.

Graph embedding



Luqman, M. M. (2012). Fuzzy Multilevel Graph Embedding for Recognition, Indexing and Retrieval of Graphic Document Images. Ph.D. thesis. University of Tours, France and Autonoma University of Barcelona, Spain.

Graph embedding

Graph probing based methods

[Wiener, 1947] [Papadopoulos et al., 1999] [Gibert et al., 2011] [Sidere et al., 2012]

Graph embedding

Dissimilarity based methods

[Pekalska et al., 2005] [Ferrer et al., 2008] [Riesen, 2010] [Bunke et al., 2011]

Graph embedding

Graph feature extraction based methods

- Node information
- Edge information
- Structure
- Topology
- Geometry
- Node/Edge neighborhood information

Muhammad Muzzamil Luqman, Jean-Yves Ramel, Josep Lladós, Thierry Brouard: Fuzzy multilevel graph embedding. Pattern Recognition 46(2): 551-565 (2013)

Nicholas Dahma, Horst Bunke, Terry Caelli, Yongsheng Gao. Efficient subgraph matching using topological node feature constraints, Pattern Recognition 48 (2015) 317330.

Hana Jarraya, Muhammad Muzzamil Luqman, Jean-Yves Ramel: Improving Fuzzy Multilevel Graph Embedding Technique by Employing Topological Node Features: An Application to Graphics Recognition. GREC 2015: 117-132

Graph embedding

Graph feature extraction based methods - FMGE

Multilevel analysis of graph

Graph Level Information **[macro details]**	Structural Level Information **[intermediate details]**	Elementary Level Information **[micro details]**
✓ Graph order ✓ Graph size	✓ Node degree ✓ Homogeneity of subgraphs in graph	✓ Node attributes ✓ Edge attributes

Graph order	Graph size	Embedding of node degree	Embedding(s) of subgraph(s) homogenity	Embedding(s) of node attribute(s)	Embedding(s) of edge attribute(s)

Muhammad Muzzamil Luqman, Jean-Yves Ramel, Josep Lladós, Thierry Brouard: Fuzzy multilevel graph embedding. Pattern Recognition 46(2): 551-565 (2013)

Graph embedding

- Graph kernels can be intuitively understood as functions measuring the similarity of pairs of graphs
- Graph kernels allow kernelized learning algorithms such as support vector machines to work directly on graphs, without having to do feature extraction to transform them to fixed-length, real-valued feature vectors
- Laplacian Graph Kernel, Treelet Kernel, Random Walk Kernel, Graphlet Kernel, etc.

Donatello Conte, Jean-Yves Ramel, Nicolas Sidere, Muhammad Muzzamil Luqman, Benoit Gaüzère, Jaume Gibert, Luc Brun, Mario Vento: A Comparison of Explicit and Implicit Graph Embedding Methods for Pattern Recognition. GbRPR 2013: 81-90

A very general overview of historical evolution of graph-based solutions to Pattern Recognition

- Graph matching (isomorphism)

  [Messmer, 1995] [Sonbaty and Ismail, 1998]

- Graph Edit Distance (GED)

  [Bunke and Shearer, 1998] [Neuhaus and Bunke, 2006]

- Graph EMbedding (GEM)

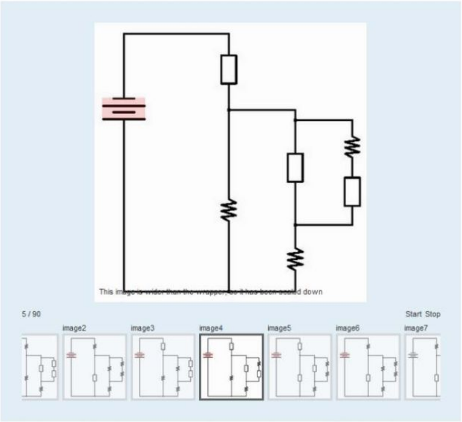
  [Luqman et al., 2009] [Sidere et al., 2009] [Gibert et al., 2011]

What kind of Pattern Recognition problems have been solved by using graphs?

- Graph similarity
- Graph classification
- Graph clustering
- Graphics detection / localization / recognition / classification / clustering / spotting
- Chemical molecules recognition / classification / clustering
- Fingerprint recognition
- Handwriting recognition
- Signature recognition / verification
- Document image segmentation / classification / clustering / indexing
- QBE and CBIR in document images
- Focused retrieval in document images
- etc.

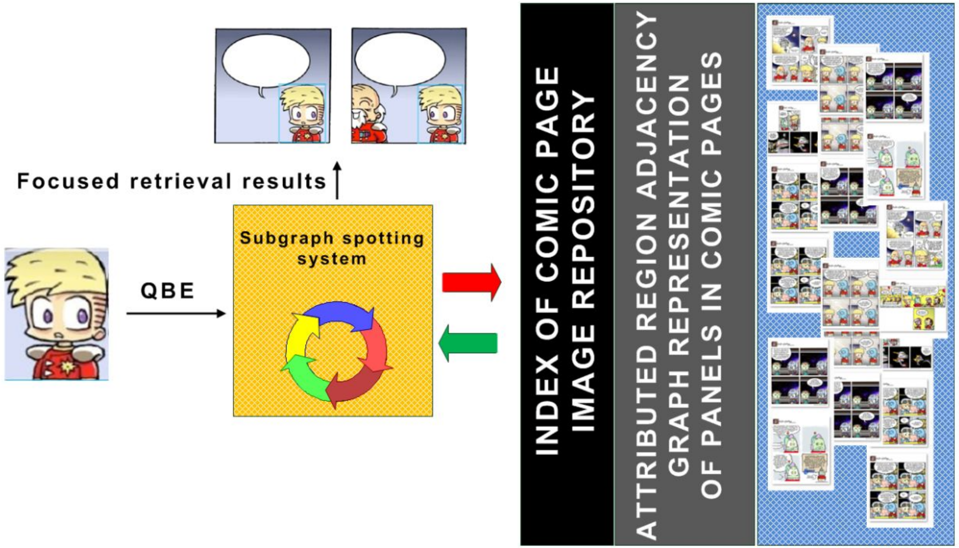
Subgraph Spotting through Explicit Graph Embedding: An Application to Content Spotting in Graphic Document Images

(a) Query image.



Luqman, M. M., Ramel, J. Y., Lladós, J., & Brouard, T. (2011). Subgraph spotting through explicit graph embedding: An application to content spotting in graphic document images. International Conference on Document Analysis and Recognition, ICDAR, 870–874.

Automatic indexing of comic page images for query by example based focused content retrieval



Luqman, M. M., Ho, H. N., Burie, J., & Ogier, J. (2013). Automatic indexing of comic page images for query by example based focused content retrieval. In Tenth IAPR International Workshop on Graphics RECognition (GREC) (pp. 153–157).

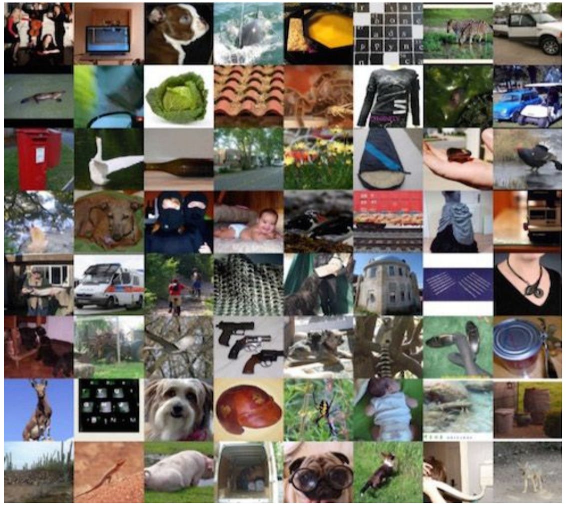
Content-based Comic Retrieval Using Multilayer Graph Representation and Frequent Graph Mining

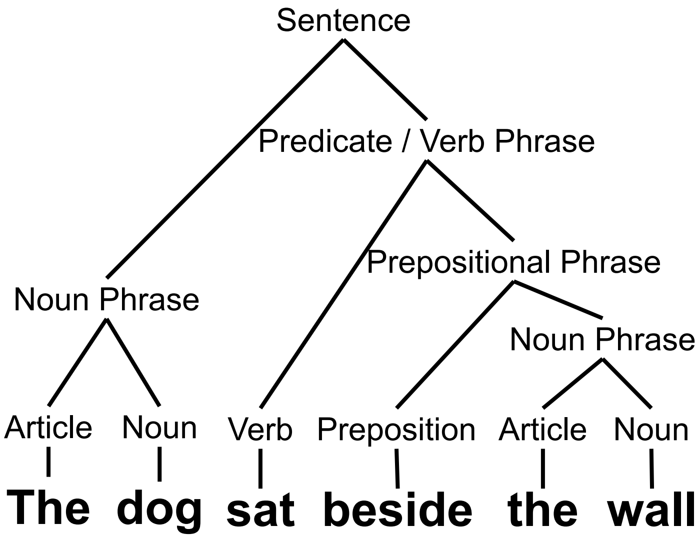
Le, T., Luqman, M. M., Burie, J., & Ogier, J. (2015). Content-based Comic Retrieval Using Multilayer Graph Representation and Frequent Graph Mining. 13th International Confrence on Document Analysis and Recognition - ICDAR’15, 15–19.

How has the success story of deep learning influenced the graph-based methods of Pattern Recognition?

Success story of deep learning

IMAGENET





Natural Language Processing (NLP)

Speech Data

Evolution of deep learning

DeepMind
AlphaGo
IMAGENET
Microsoft

First NIPS

Perceptron
Rosenblatt

Backprop
Werbos

SVM
Vapnik

CNN
LeCun

Autoencoder
LeCun, Hinton

facebook
AI Research

Speech
Recognition

1958 1959 1982 1987 1995 1997 1998 1999 2006 2010 2012 2014 2015 2016

Visual cortex
Hubel & Wiesel

Neurocognitron
Fukushima

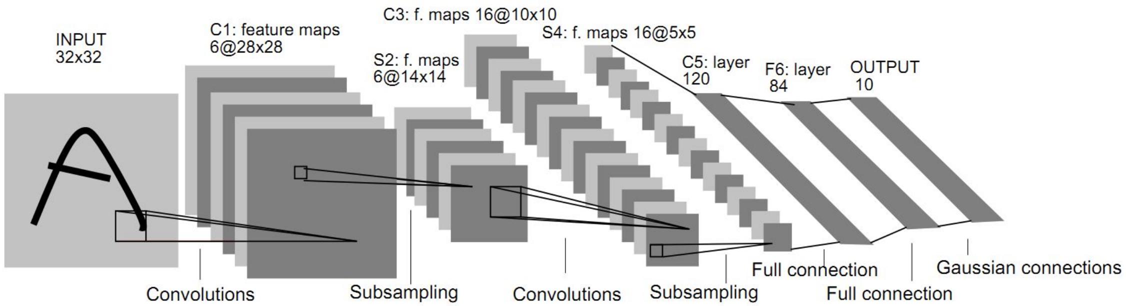
RNN / LSTM
Schmidhuber

NVIDIA
GEFORCE
First GPU

ImageNet
breakthrough
Krizhevsky

TensorFlow
OpenAI
TESLA
Autonomous
cars

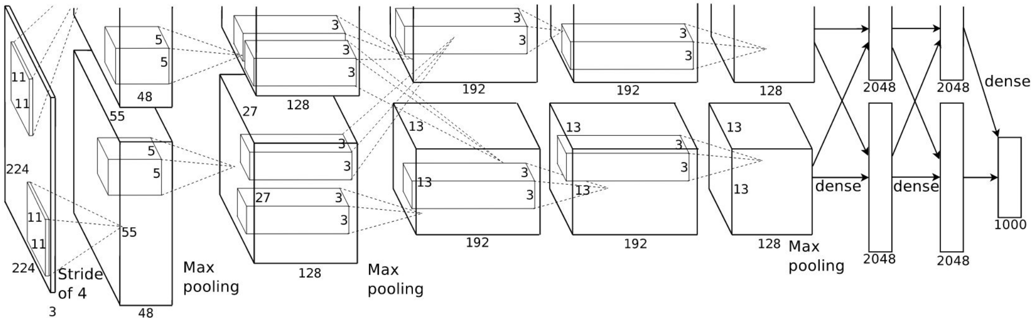
CNN: LeNet 5



- 3 convolutional + 1 fully connected layer
- 1M parameters
- Training set: MNIST 70K images
- Trained on CPU
- `tanh` non-linearity

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. IEEE, 1998.

CNN: AlexNet

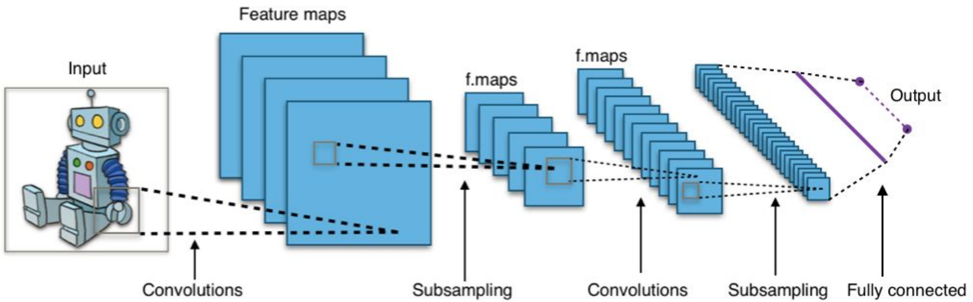


- 5 convolutional + 3 fully connected layer
- 60M parameters
- Trained on ImageNet 1.5M images
- Trained on GPU
- ReLU non-linearity
- Dropout regularization

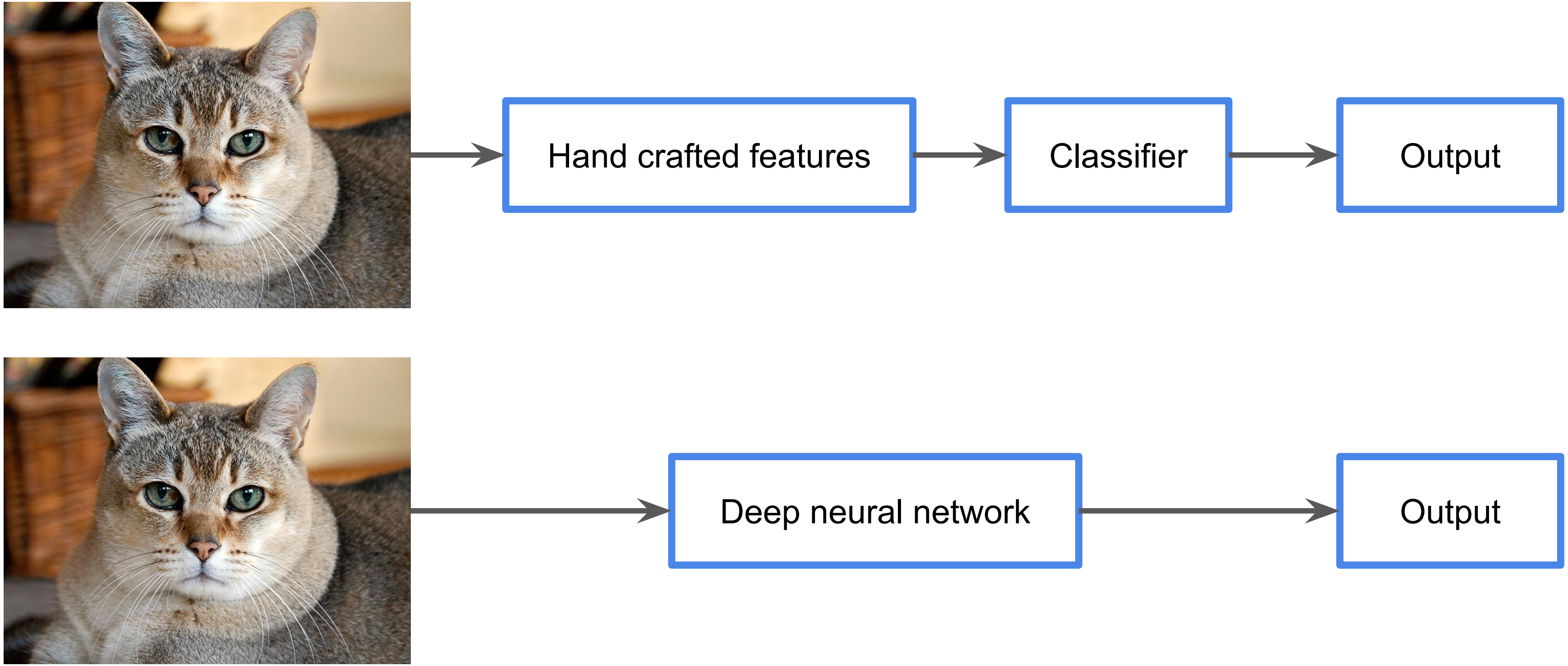
A. Krizhevsky, I. Sutskever and G. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. NIPS, 2012.

Convolutional neural network

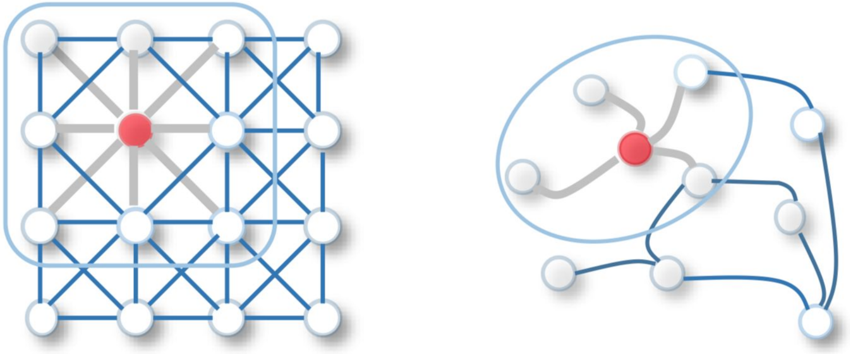
- Hierarchical compositionality
- Weight sharing
- Big data
- Computational power



Traditional vs “deep” learning



a graph convolution can be generalized from a standard 2D convolution



(a) 2D Convolution. Analogous to a graph, each pixel in an image is taken as a node where neighbors are determined by the filter size. The 2D convolution takes the weighted average of pixel values of the red node along with its neighbors. The neighbors of a node are ordered and have a fixed size.

(b) Graph Convolution. To get a hidden representation of the red node, one simple solution of the graph convolutional operation is to take the average value of the node features of the red node along with its neighbors. Different from image data, the neighbors of a node are unordered and variable in size.

Fig. 1: 2D Convolution vs. Graph Convolution.

Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu, "A Comprehensive Survey on Graph Neural Networks.", arXiv 2019.

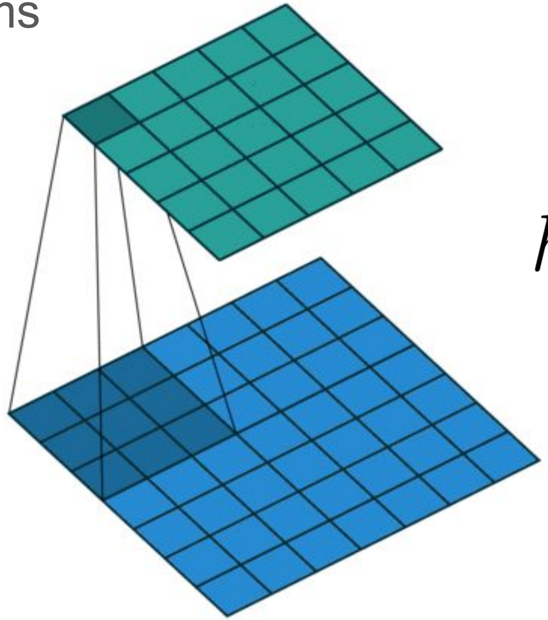
CNN: Message passing in a grid graph

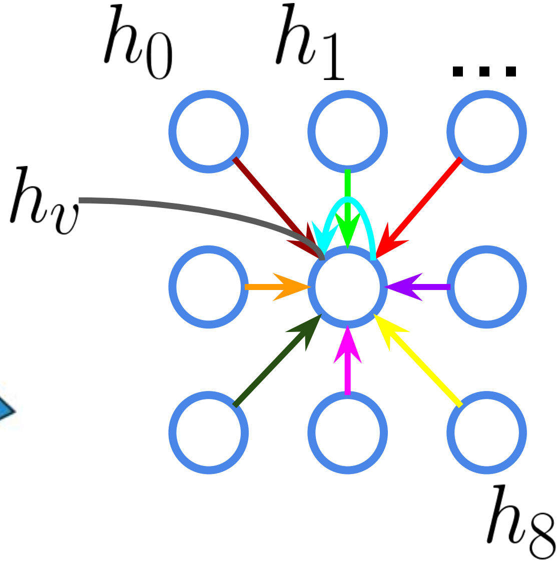
- Individual message transforms

$$W_v h_v$$

- Sum everything up

$$\sum_v W_v h_v$$





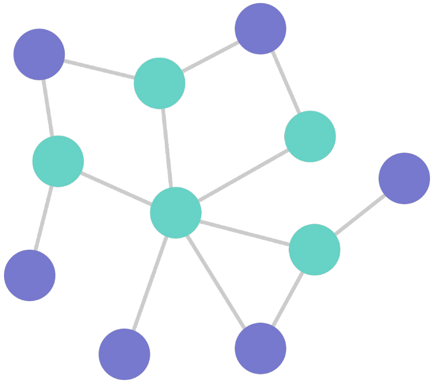
- Full update

$$h_v^{t+1} = \sigma(W_0^t h_0^t + W_1^t h_1^t + \ldots + W_v^t h_v^t + \ldots + W_8^t h_8^t)$$

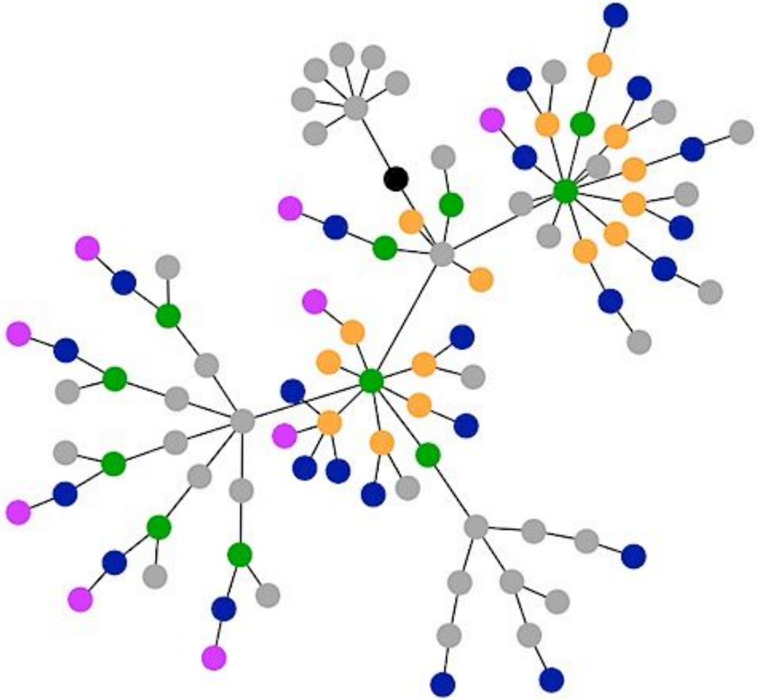
Animation by V. Dumoulin

Graph structured data

What if the data look like this?



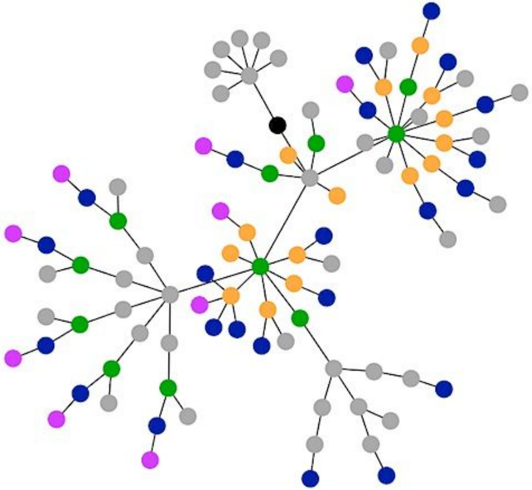
or this:

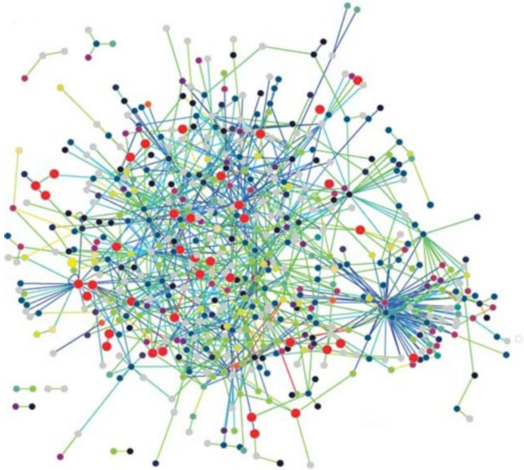


Graph structured data

Real world examples:

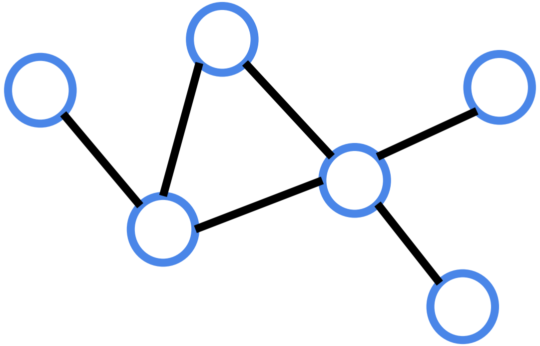
- Social networks
- World wide web
- Protein interaction networks
- Telecommunication networks
- Knowledge graphs
- ...



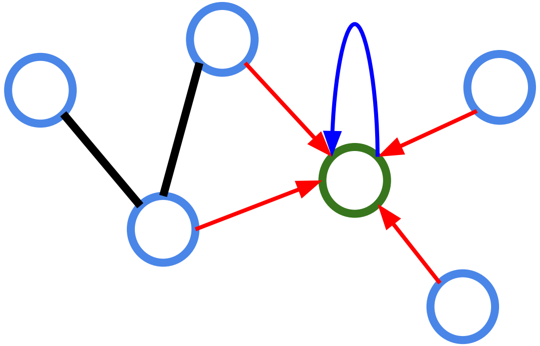


Message passing on graphs

Consider this undirected graph:



Calculate update for node in green:



Update rule:

$$h_v^{t+1} = \sigma\left(h_v^t W_0^t + \frac{1}{c_{vw}} \sum_{w \in \mathcal{N}_v} h_w^t W_1^t\right)$$

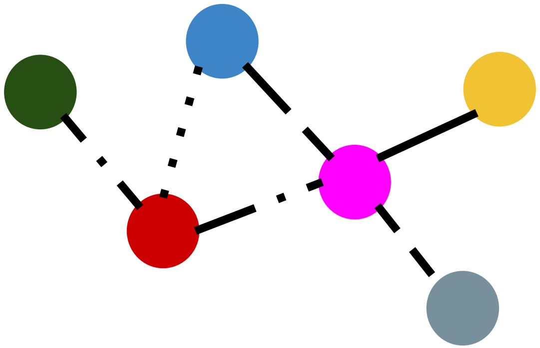
More general or simpler function also can be chosen

1. J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, G. E. Dahl. Neural Message Passing for Quantum Chemistry. ICML, 2017.
2. T. Kipf, M. Welling, Semi-Supervised Classification with Graph Convolutional Networks. ICLR, 2017

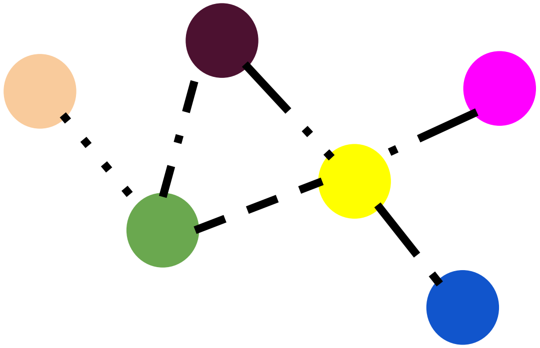
Several iteration of message passing

Node and edge updation:

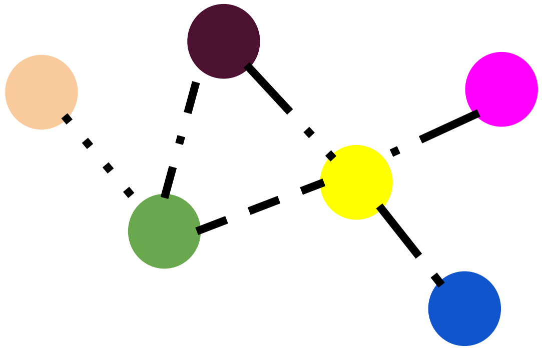
Initial stage:



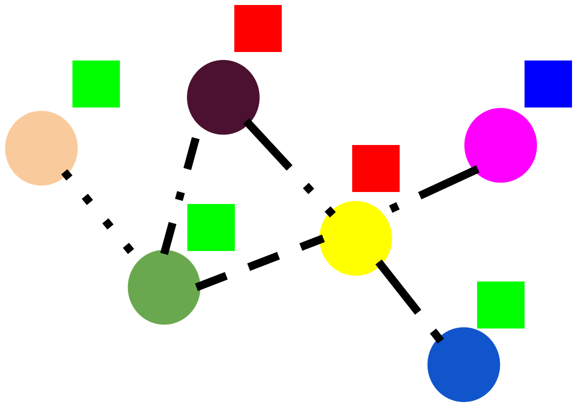
Final stage:



Graph wise classification



Node wise classification



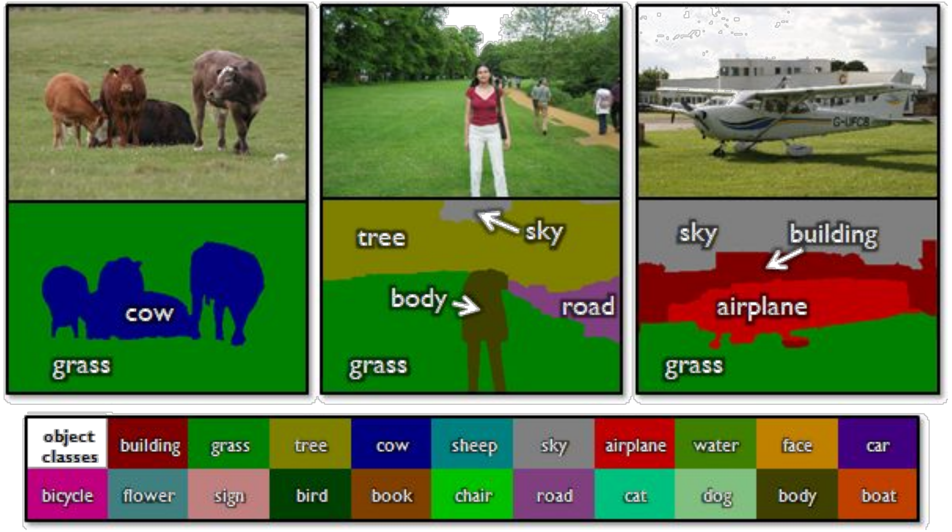


Figure credit: Shotton et al IJCV 2007

Neural Message Passing

Message function:

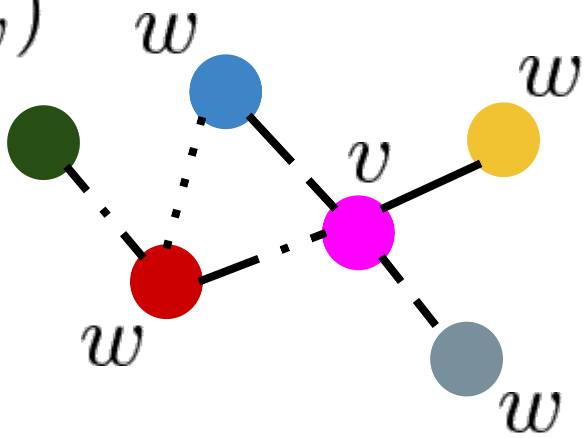
$$m_v^{t+1} = \sum_{w \in \mathcal{N}_v} M(h_v^t, h_w^t, e_{vw})$$

Update function:

$$h_v^{t+1} = U(h_v^t, m_v^{t+1})$$

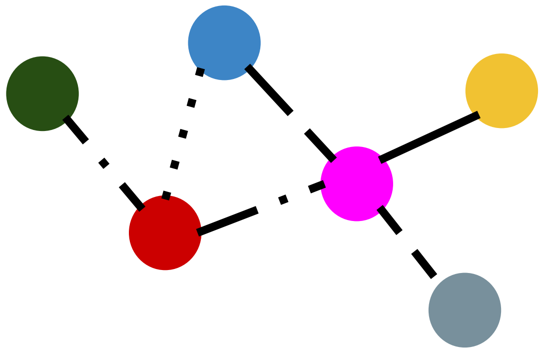
Readout function:

$$\hat{y} = R(\{h_v^T \mid v \in G\})$$



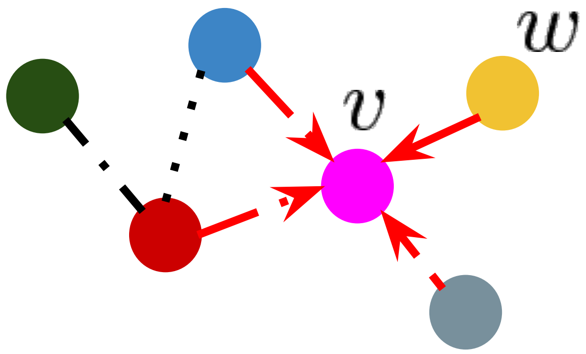
J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, G. E. Dahl. Neural Message Passing for Quantum Chemistry. ICML, 2017.

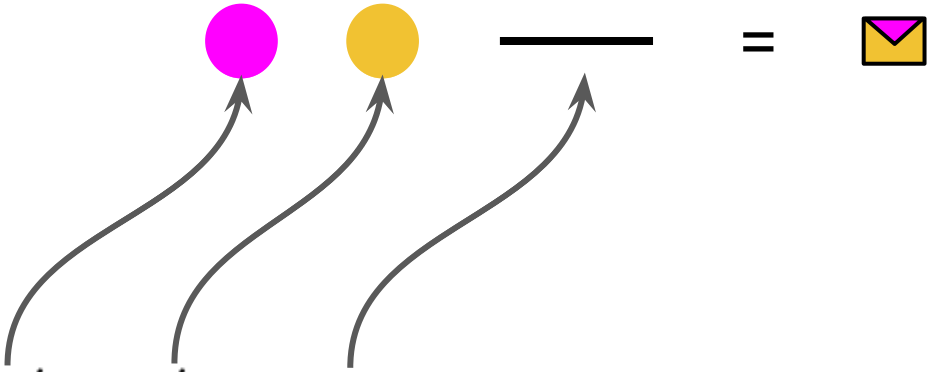
Running Example



Message passing

Message function:

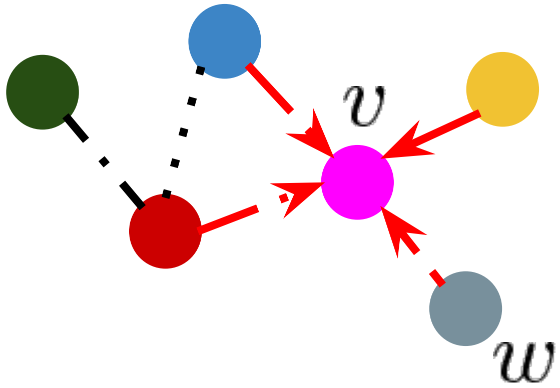


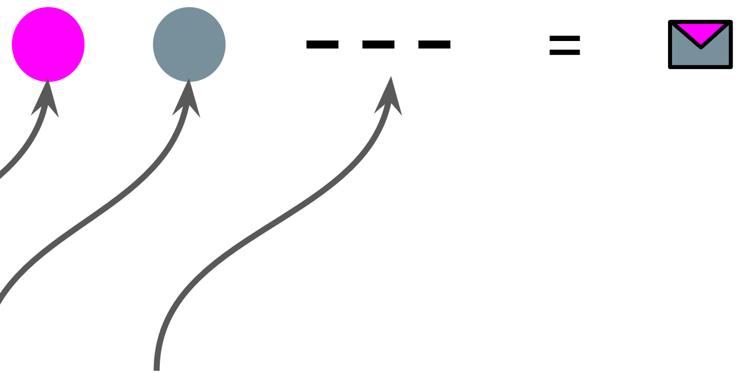


$$M(h_v^t, h_w^t, e_{vw})$$

Message passing

Message function:

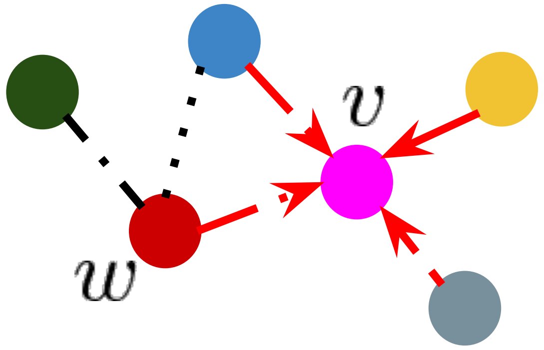


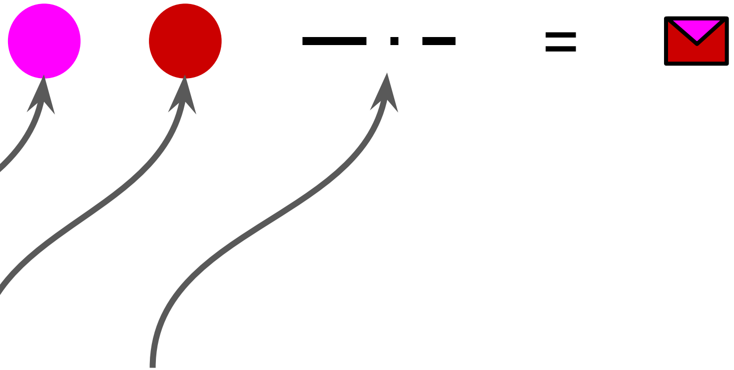


$$M(h_v^t, h_w^t, e_{vw})$$

Message passing

Message function:

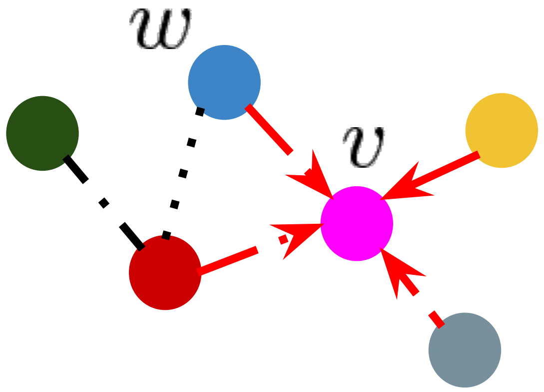


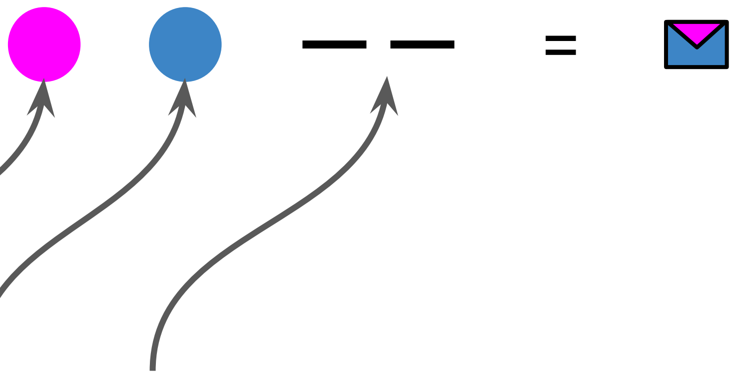


$$M(h_v^t, h_w^t, e_{vw})$$

Message passing

Message function:

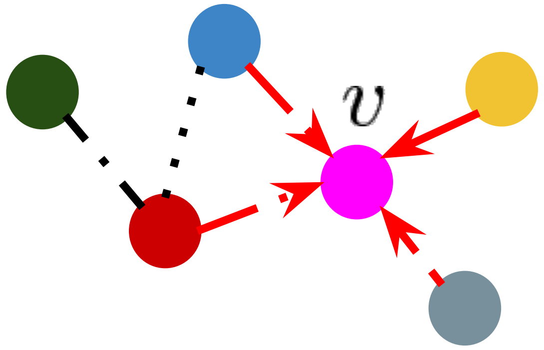


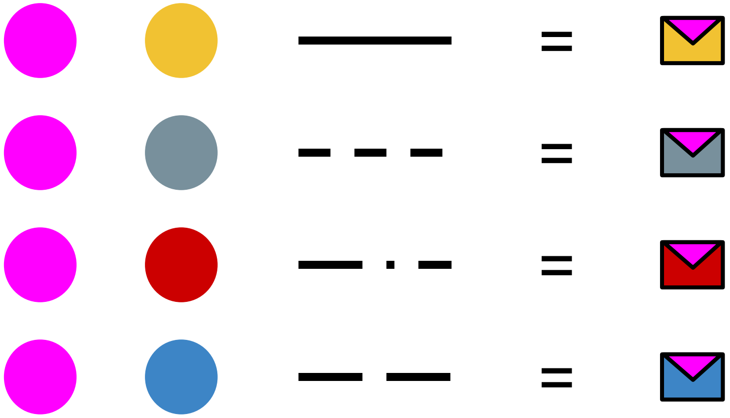


$$M(h_v^t, h_w^t, e_{vw})$$

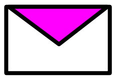
Message passing

Message function:



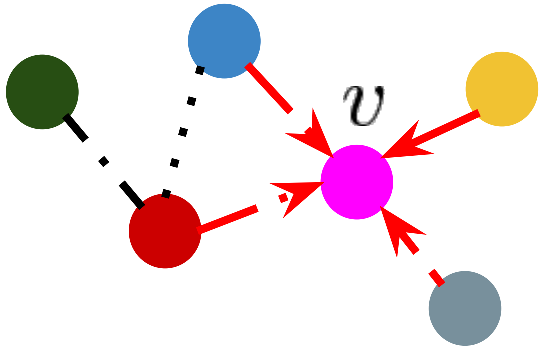


$$m_v^{t+1} = \sum_{w \in \mathcal{N}_v} M(h_v^t, h_w^t, e_{vw})$$ =



Message passing

Example message function:



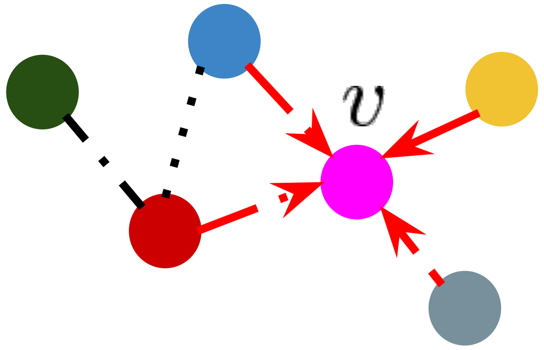
$$M(h_v^t, h_w^t, e_{vw}) = \frac{A_{vw}}{\sqrt{\deg(v)\deg(w)}} h_w^t$$

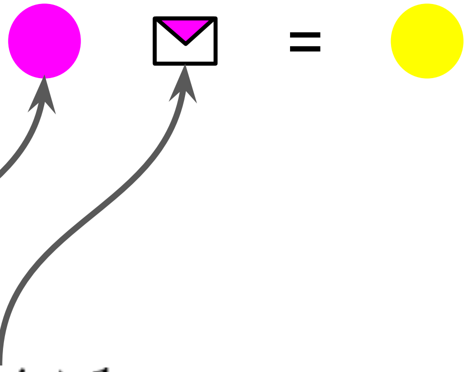
where h_v is the hidden state of the node v, e_{vw} is edge feature of vw, and A_{vw} is a learned matrix.

T. Kipf and M. Welling. Semi-Supervised Classification with Graph Convolutional Networks, ICLR 2017.

Message passing

Update function:

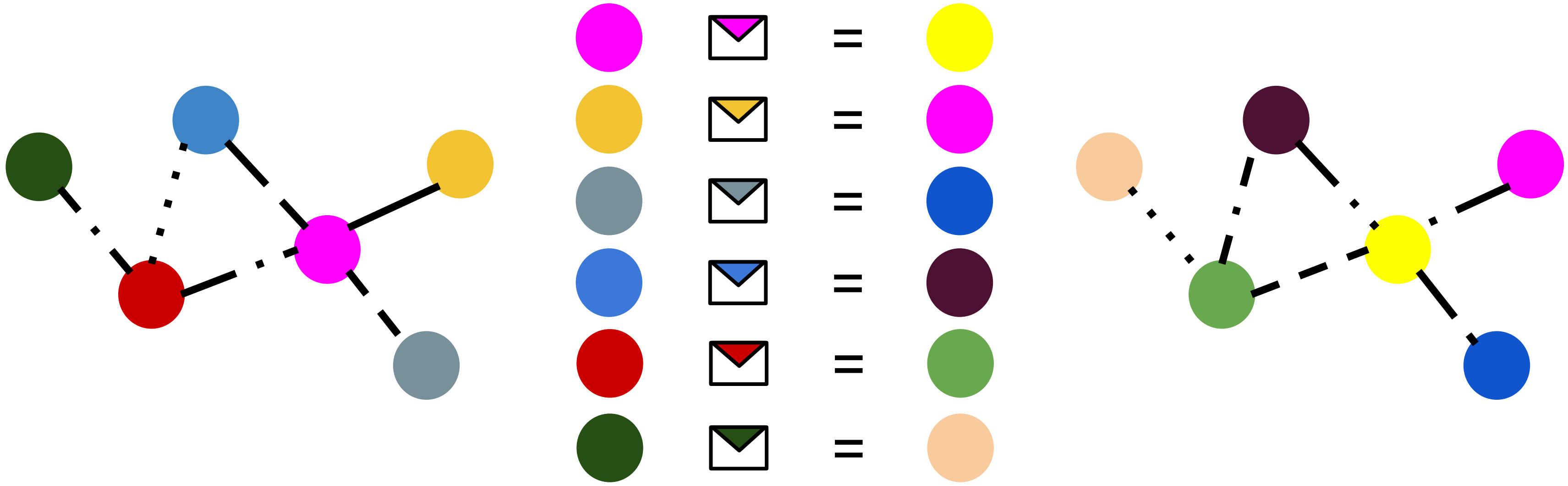




$$U(h_v^t, m_v^{t+1})$$

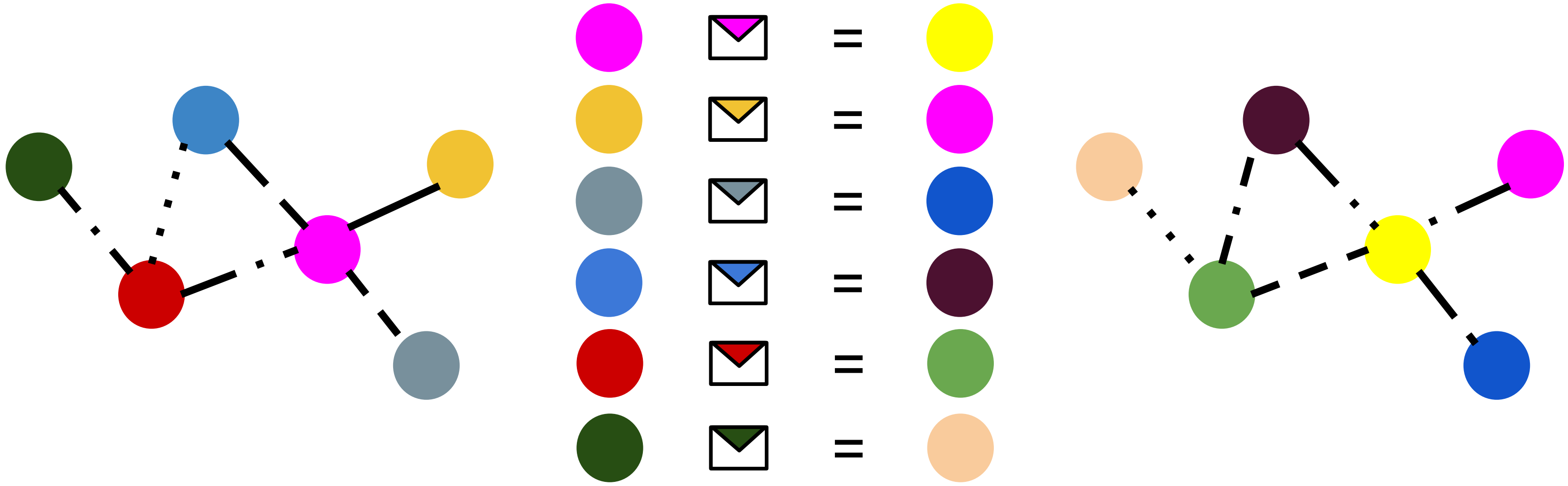
Message passing

Update function:



Message passing

Example update function:



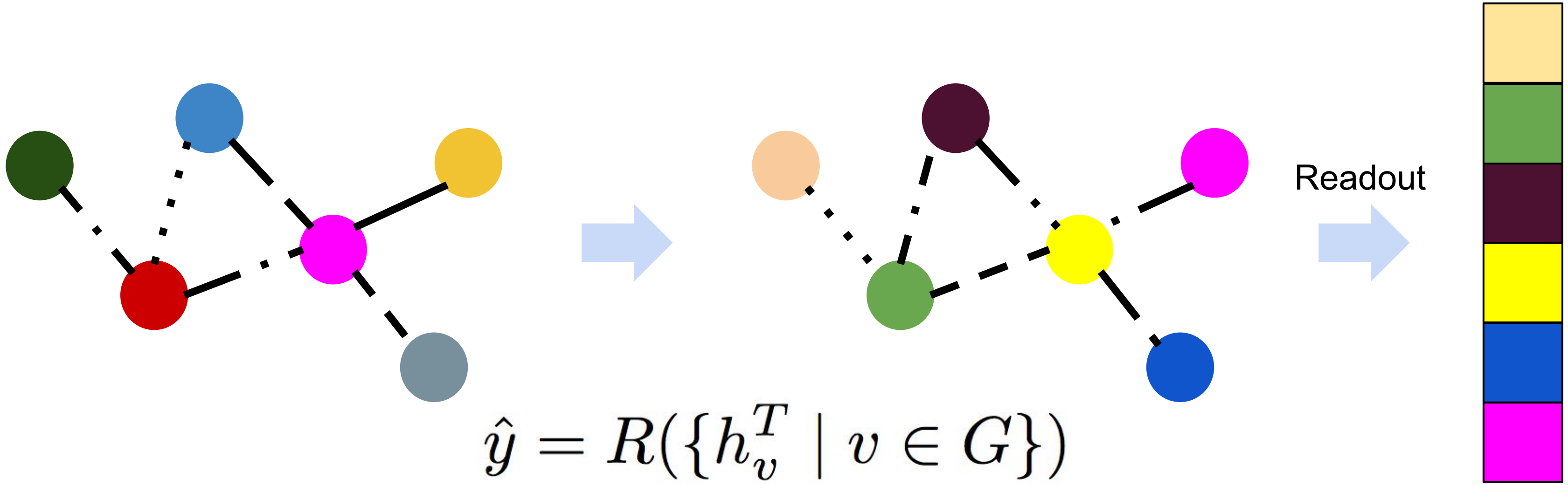
$$U_t(h_v^t, m_v^{t+1}) = \sigma(W^t m_v^{t+1})$$

where W^t are learned matrices one for each time step, σ is a non-linearity function such as ReLU (Rectified Linear Unit)

T. Kipf and M. Welling. Semi-Supervised Classification with Graph Convolutional Networks, ICLR 2017.

Readout

Readout function:



$$\hat{y} = R(\{h_v^T \mid v \in G\})$$

Example:

$$R = f(\sum_v h_v^t)$$

This readout function sums the current hidden states of all the nodes and computes an output through a learnable neural network f.

Convolutional Networks on Graphs

- Message Function

$$m_v^{t+1} = M(h_v^t, h_w^t, e_{vw}) = (h_w^t, e_{vw})$$

- Update Function

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) = \sigma(H_t^{deg(v)} m_v^{t+1})$$

- Readout Function

$$\hat{y} = R = f(\sum_{v,t} \mathrm{softmax}(W_t h_v^t))$$

where $(.,.)$ denotes concatenation, H_t^N are learned matrices one for each time step t and degree edge label, f is a neural network and σ is a non-linearity function such as ReLU

D. Duvenaud, D. Maclaurin, J. A. Iparraguirre, R. G. Bombarelli, T. Hirzel, A. Aspuru-Guzik, R. P. Adams. Convolutional Networks on Graphs for Learning Molecular Fingerprints, NIPS 2015.

Gated Graph Sequence Neural Networks

- Message Function

$$m_v^{t+1} = M(h_v^t, h_w^t, e_{vw}) = A_{e_{vw}} h_w^t$$

- Update Function

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) = \mathrm{GRU}(h_v^t, m_v^{t+1})$$

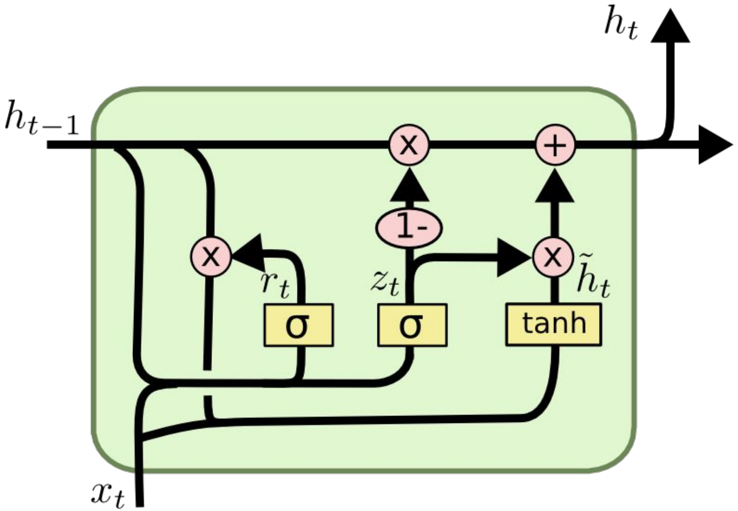
- Readout Function

$$\hat{y} = R = \sum_v \sigma(i(h_v^t, h_v^0) \odot (j(h_v^t))$$

where $A_{e_{vw}}$ is a learned matrix one for each discrete edge label, GRU is Gated Recurrent Unit, i, j are neural networks and $\odot$ is element wise multiplication, σ is a non-linearity function such as ReLU

Y. Li, D. Tarlow, M. Brockschmidt and R. Zemel. Gated graph sequence neural networks. ICLR, 2016.

GRU



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Interaction Networks

- Message Function

$$m_v^{t+1} = M(h_v^t, h_w^t, e_{vw}) = f(h_v^t, h_w^t, e_{vw})$$

- Update Function

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) = g(h_v^t, x_v, m_v^{t+1})$$

- Readout Function

$$\hat{y} = R = f(\sum_v h_v^t)$$

where f, g represent neural networks, $(.,.)$ denotes concatenation, x_v is an external vector representing some outside influence to the node v.

P. W. Battaglia, R. Pascanu, M. Lai, D. Rezende, K. Kavukcuoglu. Interaction Networks for Learning about Objects, Relations and Physics, NIPS, 2016.

Molecular Graph Convolutions

- Message Function

$$m_v^{t+1} = M(h_v^t, h_w^t, e_{vw}) = e_{vw}^t$$

- Update Function

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) = \alpha(W_1(\alpha(W_0 h_v^t), m_v^{t+1}))$$

- Readout Function

$$\hat{y} = R = \alpha(W_4(\alpha(W_2, e_{vw}^t), \alpha(W_3(h_v^t, h_w^t))))$$

where $(.,.)$ denotes concatenation, W_i are learned weight matrices, α is the ReLU activation.

S. Kearnes, K. McCloskey, M. Berndl, V. Pande, P. Riley, Molecular Graph Convolutions: Moving Beyond Fingerprints, JCAMD, vol. 30, no. 8, 2016.

Convolutional and Locally Connected Neural Networks

- Message Function

$$m_v^{t+1} = M(h_v^t, h_w^t, e_{vw}) = C_{vw}^t h_w^t$$

- Update Function

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) = \sigma(m_v^{t+1})$$

where C_{vw} are parameterized by the eigenvectors of the graph Laplacian L and the other parameters of the model, σ is a non-linearity function such as ReLU

1. Defferrard et al., Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering, NIPS 2016.
2. Bruna et al., Spectral Networks and Locally Connected Networks on Graphs, ICLR 2014.

Graph Convolutional Networks

- Message Function

$$m_v^{t+1} = M(h_v^t, h_w^t, e_{vw}) = \frac{A_{vw}}{\sqrt{\deg(v)\deg(w)}} h_w^t$$

- Update Function

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) = \sigma(W^t m_v^{t+1})$$

where A_{vw} is a learnable parameter, W^t are learned matrices one for each time step, σ is a non-linearity function such as ReLU

T. Kipf and M. Welling. Semi-Supervised Classification with Graph Convolutional Networks, ICLR 2017.

Recommended Reading

Tutorials:

- Geometric Deep Learning, Tutorial, CVPR, 2017. http://geometricdeeplearning.com/
- Deep Learning on Graphs with Graph Convolutional Networks. http://deeploria.gforge.inria.fr/thomasTalk.pdf

List of papers:

- Gilmer et al., Neural Message Passing for Quantum Chemistry, 2017. https://arxiv.org/abs/1704.01212
- Kipf et al., Semi-Supervised Classification with Graph Convolutional Networks, ICLR 2017. https://arxiv.org/abs/1609.02907
- Defferrard et al., Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering, NIPS 2016. https://arxiv.org/abs/1606.09375
- Bruna et al., Spectral Networks and Locally Connected Networks on Graphs, ICLR 2014. https://arxiv.org/abs/1312.6203
- Duvenaud et al., Convolutional Networks on Graphs for Learning Molecular Fingerprints, NIPS 2015. https://arxiv.org/abs/1509.09292
- Li et al., Gated Graph Sequence Neural Networks, ICLR 2016. https://arxiv.org/abs/1511.05493
- Battaglia et al., Interaction Networks for Learning about Objects, Relations and Physics, NIPS 2016. https://arxiv.org/abs/1612.00222
- Kearnes et al., Molecular Graph Convolutions: Moving Beyond Fingerprints, 2016. https://arxiv.org/abs/1603.00856

Recommended Reading

Source Code / Repositories:

- Neural Message Passing for Computer Vision: https://github.com/priba/nmp_qc
- Graph Convolutional Networks in TensorFlow: https://github.com/tkipf/gcn
- Graph Convolutional Networks in PyTorch: https://github.com/tkipf/pygcn
- PyTorch implementation of graph ConvNets: https://github.com/xbresson/graph_convnets_pytorch
- Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering: https://github.com/mdeff/cnn_graph

Other material:

- Blog post on Graph Convolutional Networks: http://tkipf.github.io/graph-convolutional-networks

Saturday 21st september 2019
09h00 – 12h30

Part-1

- A historic perspective of graph-based methods in PR & DIA
- Neural Networks on graphs and modern trends in graph-based PR & DIA

Coffee break (10h30 - 11h00)

Part-2

- Applications of Graph Neural Networks
  - Learning Graph Distances
  - Table Detection
- Hands-on
  - Deep Graph Library

http://gmprdia.univ-lr.fr

Centre de Visió per Computador

Informatique Image Interaction
Univ. La Rochelle

UNIVERSITY OF EXETER

ICDAR 2019
15th International Conference on Document Analysis and Recognition
20-25 September 2019 | International Convention Centre Sydney, Australia

Saturday 21st september 2019
09h00 – 12h30

Part-1

- A historic perspective of graph-based methods in PR & DIA
- Neural Networks on graphs and modern trends in graph-based PR & DIA

Coffee break (10h30 - 11h00)

Part-2

- **Applications of Graph Neural Networks**
  - **Learning Graph Distances**
  - **Table Detection**
- **Hands-on**
  - **Deep Graph Library**

http://gmprdia.univ-lr.fr

Centre de Visió per Computador

Informatique Image Interaction
Univ. La Rochelle

UNIVERSITY OF EXETER

ICDAR 2019
15th International Conference on Document Analysis and Recognition
20-25 September 2019 | International Convention Centre Sydney, Australia

Application:

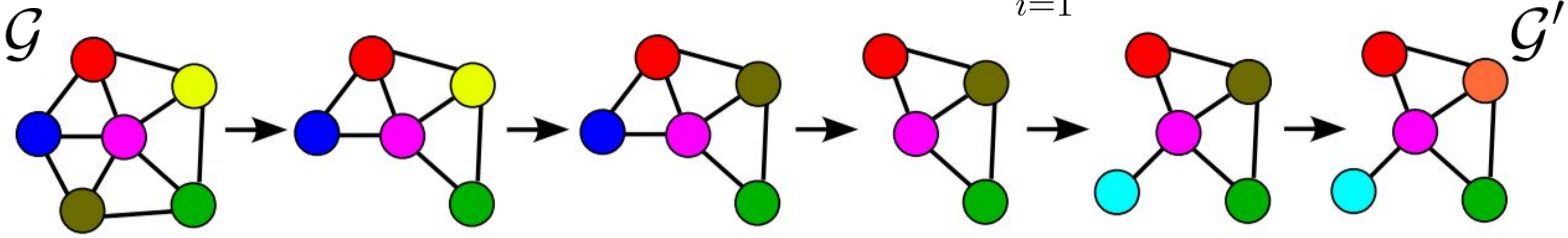
Learning Graph Distances

Riba *et al.* **Learning Graph Distances with Message Passing Neural Networks**. In ICPR, 2018

Graph edit distance (Reminder)

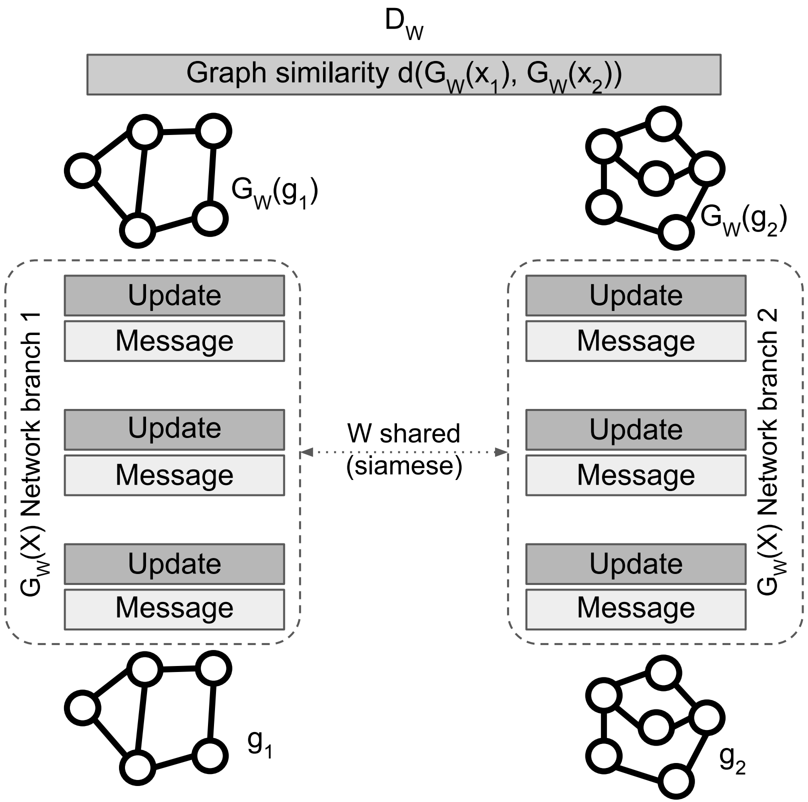
- A measure of similarity between two graphs.
- Node and edge insertion, deletion, substitution.
- Summation of the edit costs

$$\mathrm{GED}(\mathcal{G}, \mathcal{G}') = \min_{(e_1,\ldots,e_k) \in \mathcal{P}(\mathcal{G},\mathcal{G}')} \sum_{i=1}^{k} c(e_i)$$



A. Sanfeliu, K. S. Fu. A distance measure between attributed relational graphs for pattern recognition. IEEE TSMC, vol. 13, no. 3, 1983.

Architecture



Graph similarity

- Hausdorff Distance

$$\mathrm{H}(A, B) = \max\left(\max_{a \in A} \inf_{b \in B} d(a, b), \max_{b \in B} \inf_{a \in A} d(a, b) \right)$$

- Chamfer Distance

$$\hat{\mathrm{H}}(A, B) = \sum_{a \in A} \inf_{b \in B} d(a, b) + \sum_{b \in B} \inf_{a \in A} d(a, b)$$

- Proposed distance.

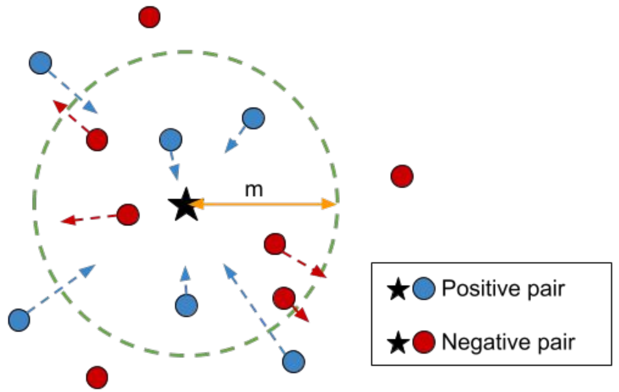
$$d(g_1, g_2) = \frac{\hat{\mathrm{H}}(V_1, V_2)}{|V_1| + |V_2|}$$

Contrastive loss

Given $D_W = d(G_W(g_1), G_W(g_2))$ where g_1 and g_2 are graphs and W a set of specific weights, the **Loss Function** is

$$l(D_W) = \frac{1}{2}\begin{cases} D_W^2, & \text{if } Y = 1 \text{ (positive pair)} \\ \{\max(0, m - D_W)\}^2, & \text{if } Y = 0 \text{ (negative pair)} \end{cases}$$

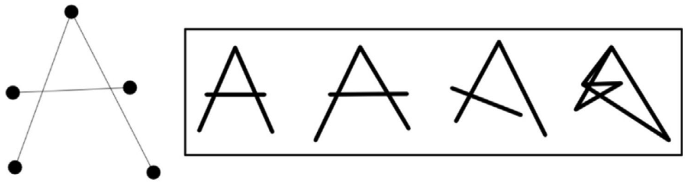
where $m=1$ is the adaptive margin.



Datasets

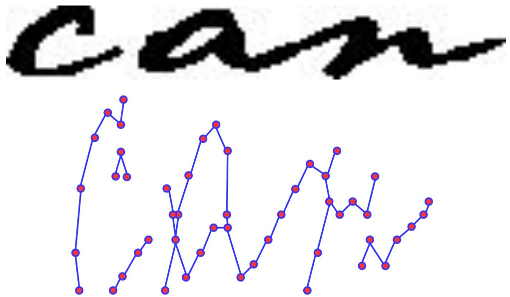
Letter

- Synthetic Graphs
- 15 classes
- 750 Graphs per class
- 3 different distortion levels



George Washington

- Handwritten words
- Several graph constructions
- 105 keywords
- 4894 instances
- **HistoGraph** (subset for classification)



Classification Letters

	LOW	MED	HIGH
BP*	99.73	94.27	89.87
HED†	**97.87**	**86.93**	**79.2**
Embedding‡	99.80	94.90	92.90
MPNN	95.04 $\pm$0.7224	83.20 $\pm$1.2189	72.27 $\pm$2.0060
Siamese MPNN	**98.08 $\pm$ 0.1068**	**89.0136 $\pm$ 0.1808**	**74.77 $\pm$ 6.4505**
Test BP	98.19 $\pm$0.1361	88.37 $\pm$0.41	79.65 $\pm$6.4345
Test HED	98.00 $\pm$0.1461	89.79 $\pm$0.3110	77.07 $\pm$5.6106

Classification Histograph

			Siamese MPNN	
Subset	**BP***	**PSGE**†	3-NN	5-NN
Keypoint	77.62	80.42	**85.31** $\pm$ **1.6552**	82.80 $\pm$ 0.5600
Projection	81.82	80.42	**73.15** $\pm$ **2.6014**	69.65 $\pm$ 1.5064

Retrieval GeorgeWashington

Method		mAP
PHOC*		64.00
BOF HMM†		80.00
DTW	DTW'01	42.26
	DTW'08	63.39
	DTW'09	64.80
	DTW'16	68.64
Mean Ensemble BP‡		69.16
Siamese MPNN		75.85 ± 3.64

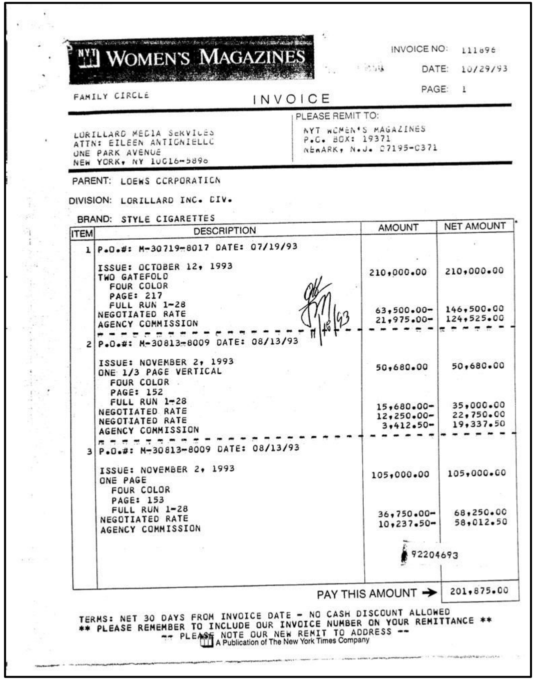
Application:

Table Detection by GNN

Riba *et al.* Table Detection in Invoice Documents by Graph Neural Networks. In ICDAR, 2019

Motivation

- Invoice Documents
- Semi-structured Documents
- Tables share **Repetitive Patterns**



WOMEN'S MAGAZINES

INVOICE NO: 111896
DATE: 10/29/93
PAGE: 1

FAMILY CIRCLE

INVOICE

LORILLARD MEDIA SERVICES
ATTN: EILEEN ANTIGNIELLO
ONE PARK AVENUE
NEW YORK, NY 10016-5896

PLEASE REMIT TO:
NYT WOMEN'S MAGAZINES
P.O. BOX: 19371
NEWARK, N.J. 07195-0371

PARENT: LOEWS CORPORATION

DIVISION: LORILLARD INC. DIV.

BRAND: STYLE CIGARETTES

ITEM	DESCRIPTION	AMOUNT	NET AMOUNT
1	P.O.#: M-30719-8017 DATE: 07/19/93		
	ISSUE: OCTOBER 12, 1993		
	TWO GATEFOLD	210,000.00	210,000.00
	FOUR COLOR		
	PAGE: 217		
	FULL RUN 1-28		
	NEGOTIATED RATE	63,500.00-	146,500.00
	AGENCY COMMISSION	21,975.00-	124,525.00
2	P.O.#: M-30813-8009 DATE: 08/13/93		
	ISSUE: NOVEMBER 2, 1993		
	ONE 1/3 PAGE VERTICAL	50,680.00	50,680.00
	FOUR COLOR		
	PAGE: 152		
	FULL RUN 1-28		
	NEGOTIATED RATE	15,680.00-	35,000.00
	NEGOTIATED RATE	12,250.00-	22,750.00
	AGENCY COMMISSION	3,412.50-	19,337.50
3	P.O.#: M-30813-8009 DATE: 08/13/93		
	ISSUE: NOVEMBER 2, 1993		
	ONE PAGE	105,000.00	105,000.00
	FOUR COLOR		
	PAGE: 153		
	FULL RUN 1-28		
	NEGOTIATED RATE	36,750.00-	68,250.00
	AGENCY COMMISSION	10,237.50-	58,012.50

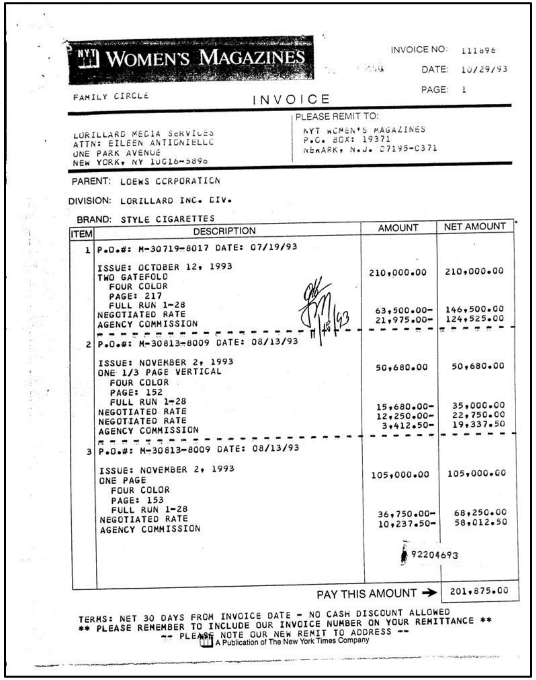
11/18/93

92204693

PAY THIS AMOUNT → 201,875.00

TERMS: NET 30 DAYS FROM INVOICE DATE - NO CASH DISCOUNT ALLOWED
** PLEASE REMEMBER TO INCLUDE OUR INVOICE NUMBER ON YOUR REMITTANCE **
-- PLEASE NOTE OUR NEW REMIT TO ADDRESS --
A Publication of The New York Times Company

Motivation



NYT WOMEN'S MAGAZINES

INVOICE NO: 111696
DATE: 10/29/93
PAGE: 1

FAMILY CIRCLE

INVOICE

LORILLARD MEDIA SERVICES
ATTN: EILEEN ANTIGNIELLO
ONE PARK AVENUE
NEW YORK, NY 10016-5896

PLEASE REMIT TO:
NYT WOMEN'S MAGAZINES
P.O. BOX: 19371
NEWARK, N.J. 07195-0371

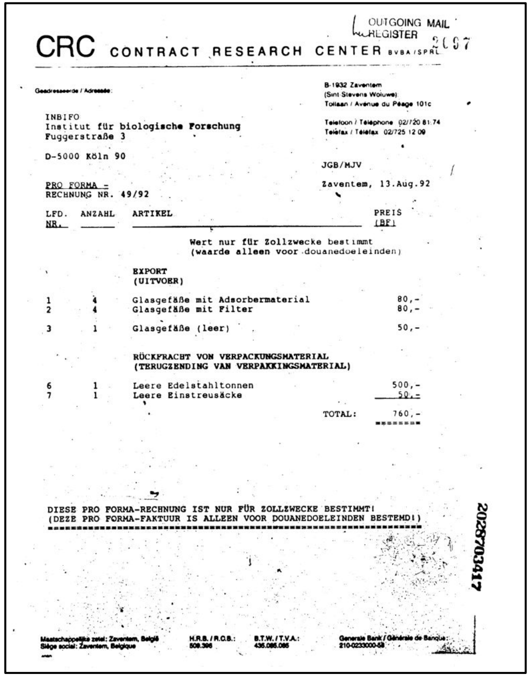
PARENT: LOEWS CORPORATION
DIVISION: LORILLARD INC. DIV.
BRAND: STYLE CIGARETTES

ITEM	DESCRIPTION	AMOUNT	NET AMOUNT
1	P.O.#: M-30719-8017 DATE: 07/19/93		
	ISSUE: OCTOBER 12, 1993 TWO GATEFOLD FOUR COLOR PAGE: 217 FULL RUN 1-28	210,000.00	210,000.00
	NEGOTIATED RATE AGENCY COMMISSION	63,500.00- 21,975.00-	146,500.00 124,525.00
2	P.O.#: M-30813-8009 DATE: 08/13/93		
	ISSUE: NOVEMBER 2, 1993 ONE 1/3 PAGE VERTICAL FOUR COLOR PAGE: 152 FULL RUN 1-28	50,680.00	50,680.00
	NEGOTIATED RATE NEGOTIATED RATE AGENCY COMMISSION	15,680.00- 12,250.00- 3,412.50-	35,000.00 22,750.00 19,337.50
3	P.O.#: M-30813-8009 DATE: 08/13/93		
	ISSUE: NOVEMBER 2, 1993 ONE PAGE FOUR COLOR PAGE: 153 FULL RUN 1-28	105,000.00	105,000.00
	NEGOTIATED RATE AGENCY COMMISSION	36,750.00- 10,237.50-	68,250.00 58,012.50

92204693

PAY THIS AMOUNT → 201,875.00

TERMS: NET 30 DAYS FROM INVOICE DATE - NO CASH DISCOUNT ALLOWED
** PLEASE REMEMBER TO INCLUDE OUR INVOICE NUMBER ON YOUR REMITTANCE **
-- PLEASE NOTE OUR NEW REMIT TO ADDRESS --
A Publication of The New York Times Company



OUTGOING MAIL REGISTER 2097

CRC CONTRACT RESEARCH CENTER BVBA/SPRL

Geadresseerde / Adressée:

B-1932 Zaventem
(Sint Stevens Woluwe)
Tollaan / Avenue du Péage 101c

INBIFO
Institut für biologische Forschung
Fuggerstraße 3

D-5000 Köln 90

Telefoon / Téléphone 02/720 81 74
Telefax / Téléfax 02/725 12 09

JGB/MJV

PRO FORMA -
RECHNUNG NR. 49/92

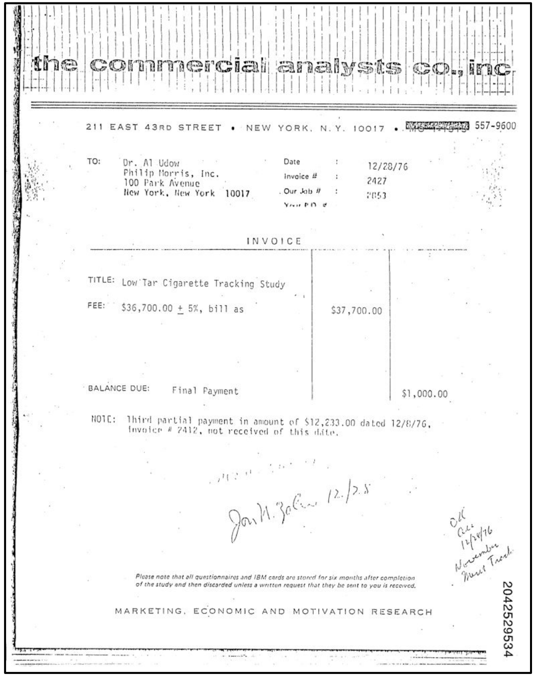
Zaventem, 13.Aug.92

LFD. NR.	ANZAHL	ARTIKEL	PREIS (BF)
		Wert nur für Zollzwecke bestimmt (waarde alleen voor douanedoeleinden)	
		EXPORT (UITVOER)	
1	4	Glasgefäße mit Adsorbermaterial	80,-
2	4	Glasgefäße mit Filter	80,-
3	1	Glasgefäße (leer)	50,-
		RÜCKFRACHT VON VERPACKUNGSMATERIAL (TERUGZENDING VAN VERPAKKINGSMATERIAL)	
6	1	Leere Edelstahltonnen	500,-
7	1	Leere Einstreusäcke	50,-
		TOTAL:	760,-

DIESE PRO FORMA-RECHNUNG IST NUR FÜR ZOLLZWECKE BESTIMMT!
(DEZE PRO FORMA-FAKTUUR IS ALLEEN VOOR DOUANEDOELEINDEN BESTEMD!)

2028703417

Maatschappelijke zetel: Zaventem, België
Siège social: Zaventem, Belgique
H.R.B. / R.C.B.: 509.396
B.T.W. / T.V.A.: 435.095.095
Generale Bank / Générale de Banque: 210-0233000-58



the commercial analysts co., inc.

211 EAST 43RD STREET • NEW YORK, N.Y. 10017 • [redacted] 557-9600

TO: Dr. Al Udow
Philip Morris, Inc.
100 Park Avenue
New York, New York 10017

Date : 12/28/76
Invoice # : 2427
Our Job # : 2053
Your P.O. # :

INVOICE

TITLE:	Low Tar Cigarette Tracking Study	
FEE:	$36,700.00 ± 5%, bill as	$37,700.00
BALANCE DUE:	Final Payment	$1,000.00

NOTE: Third partial payment in amount of $12,233.00 dated 12/8/76, Invoice # 2412, not received of this date.

Please note that all questionnaires and IBM cards are stored for six months after completion of the study and then discarded unless a written request that they be sent to you is received.

MARKETING, ECONOMIC AND MOTIVATION RESEARCH

2042529534

Graph Construction

NYT WOMEN'S MAGAZINES

INVOICE NO: 111896
DATE: 10/29/93
PAGE: 1

FAMILY CIRCLE

INVOICE

LORILLARD MEDIA SERVICES
ATTN: EILEEN ANTIGNIELLO
ONE PARK AVENUE
NEW YORK, NY 10016-5896

PLEASE REMIT TO:
NYT WOMEN'S MAGAZINES
P.O. BOX: 19371
NEWARK, N.J. 07195-0371

PARENT: LOEWS CORPORATION

DIVISION: LORILLARD INC. DIV.

BRAND: STYLE CIGARETTES

ITEM	DESCRIPTION	AMOUNT	NET AMOUNT
1	P.O.#: M-30719-8017 DATE: 07/19/93		
	ISSUE: OCTOBER 12, 1993 TWO GATEFOLD FOUR COLOR PAGE: 217 FULL RUN 1-28	210,000.00	210,000.00
	NEGOTIATED RATE	63,500.00-	146,500.00
	AGENCY COMMISSION	21,975.00-	124,525.00
2	P.O.#: M-30813-8009 DATE: 08/13/93		
	ISSUE: NOVEMBER 2, 1993 ONE 1/3 PAGE VERTICAL FOUR COLOR PAGE: 152 FULL RUN 1-28	50,680.00	50,680.00
	NEGOTIATED RATE	15,680.00-	35,000.00
	NEGOTIATED RATE	12,250.00-	22,750.00
	AGENCY COMMISSION	3,412.50-	19,337.50
3	P.O.#: M-30813-8009 DATE: 08/13/93		
	ISSUE: NOVEMBER 2, 1993 ONE PAGE FOUR COLOR PAGE: 153 FULL RUN 1-28	105,000.00	105,000.00
	NEGOTIATED RATE	36,750.00-	68,250.00
	AGENCY COMMISSION	10,237.50-	58,012.50
		92204693	

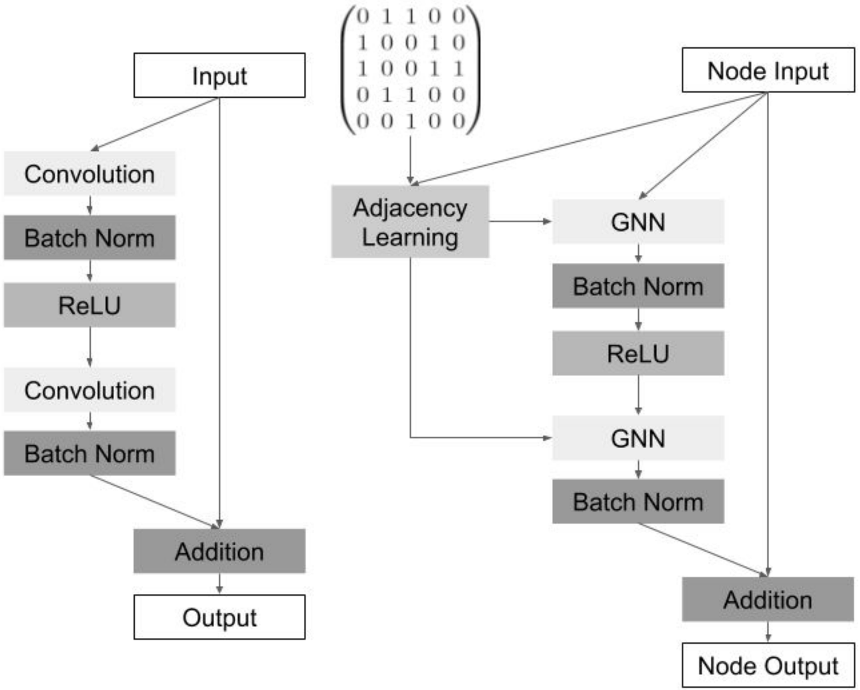
PAY THIS AMOUNT → 201,875.00

TERMS: NET 30 DAYS FROM INVOICE DATE - NO CASH DISCOUNT ALLOWED
** PLEASE REMEMBER TO INCLUDE OUR INVOICE NUMBER ON YOUR REMITTANCE **
-- PLEASE NOTE OUR NEW REMIT TO ADDRESS --

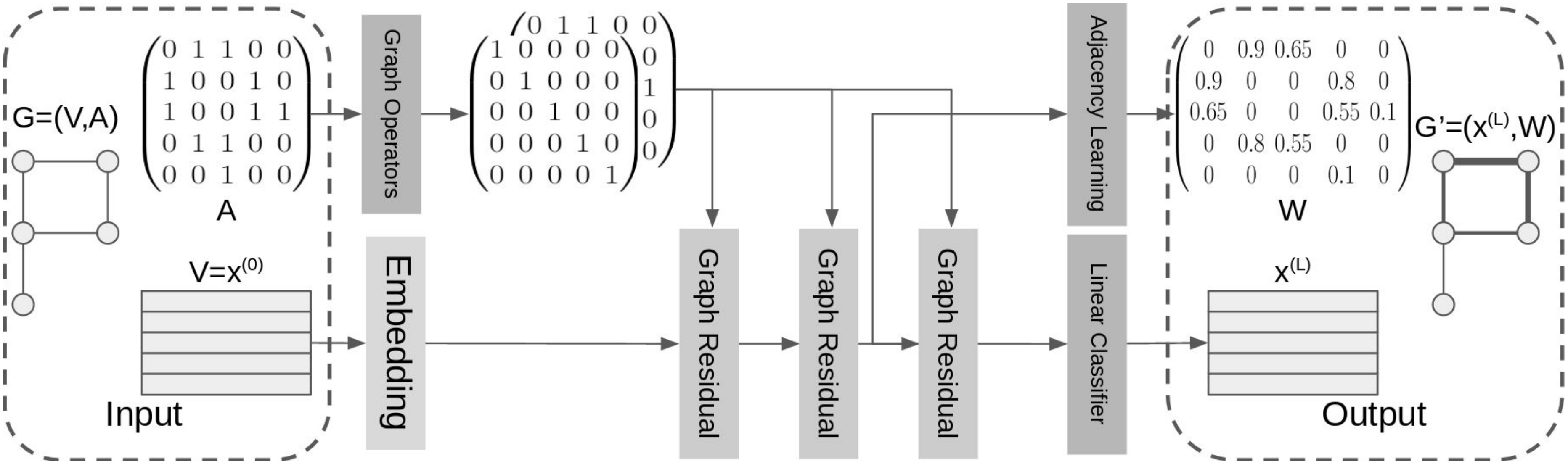
A Publication of The New York Times Company

Graph Residual Block

- Follows the idea of ResNet
- GNN layer with skip connection
- Edge weights are learned at the beginning of the block



Architecture



Objective functions

- **Node classifier**: Linear classifier with Softmax operation

- **Edge classifier**: Binary Cross entropy

- Edge weights are learned at the beginning of the block

Table detection

- Discard 0’ed edges
- Subgraphs with nodes classified as Table are considered
- Confidence score of these subgraphs are thresholded for the final decision

Datasets

CON-ANONYM

- 960 documents
- 8 region annotation
- Common car invoices
- Not publicly available

RVL-CDIP

- Overall 25,000 images
- 5 region annotation
- Selected 518 invoice class
- Publicly available

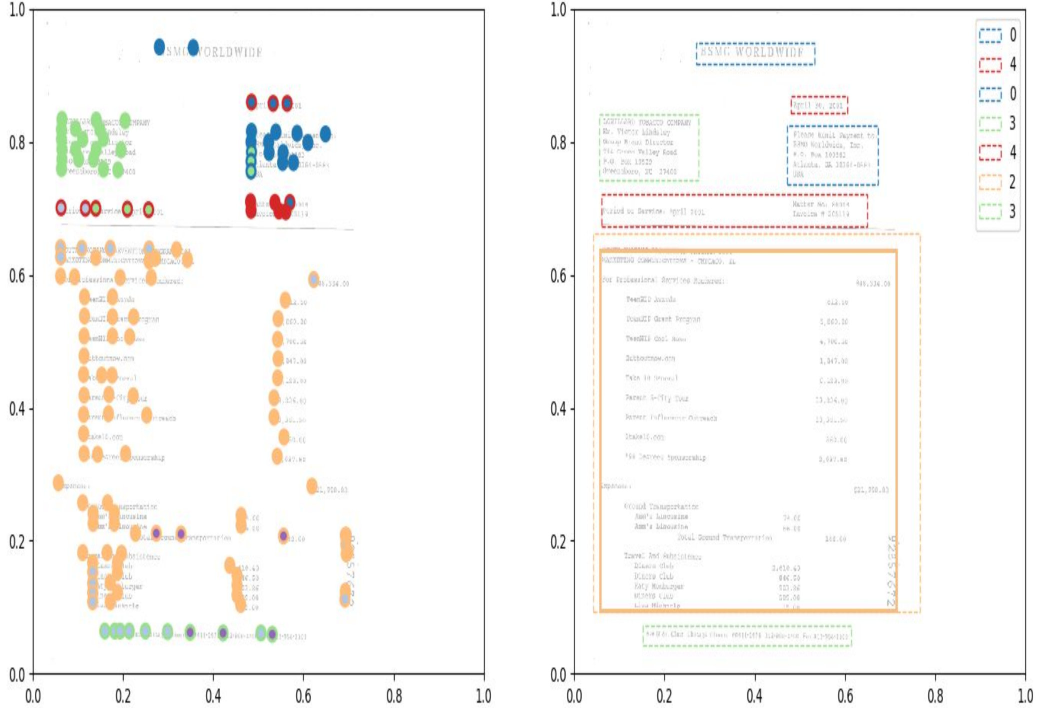
Node classification

Task	CON-ANONYM			RVL-CDIP		
	All	Table	Edge	All	Table	Edge
Pow 2	82.8	96.4	–	57.8	80.9	–
+ Edge	84.2	97.0	93.4	58.2	79.1	84.1
Pow 5	82.7	96.2	–	56.5	82.3	–
+ Edge	84.5	97.2	93.4	62.3	83.9	84.0

Table Detection

Task	CON-ANONYM			RVL-CDIP		
	F1-Score	Precision	Recall	F1-Score	Precision	Recall
Pow 2	69.4	65.8	73.4	28.6	23.9	35.4
+ Edge	70.8	65.2	77.6	30.8	26.7	36.5
Pow 5	68.4	65.3	71.8	22.6	20.0	26.0
+ Edge	73.7	78.4	69.5	30.8	25.2	39.6

Qualitative



Graph Neural Networks

Neural Message Passing (Reminder)

Message function:

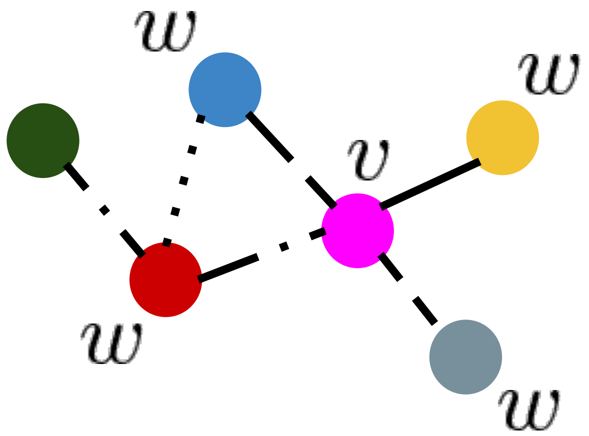
$$m_v^{t+1} = \sum_{w \in \mathcal{N}_v} M(h_v^t, h_w^t, e_{vw})$$

Update function:

$$h_v^{t+1} = U(h_v^t, m_v^{t+1})$$

Readout function:

$$\hat{y} = R(\{h_v^T \mid v \in G\})$$



J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, G. E. Dahl. Neural Message Passing for Quantum Chemistry. ICML, 2017.

Simple Message Passing Layer

Let us consider a graph $G = (V, A)$ where V is the set of nodes and A the adjacency matrix

DON'T PANIC!

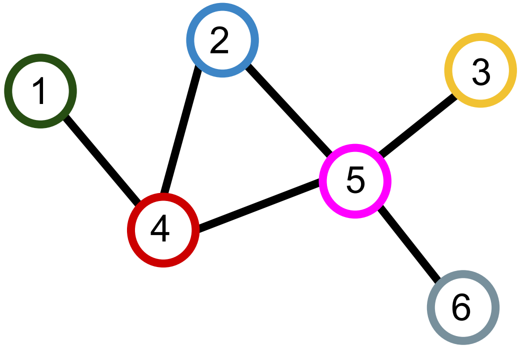
$$x^{(k+1)} = G_c(x^{(k)}) = \rho\left(\sum_{B \in A^{(k)}} B x^{(k)} \theta_B^{(k)}\right)$$

V. Garcia and J. Bruna. *Few-Shot Learning with Graph Neural Networks*. ICLR, 2018

Formalization

$$\boxed{x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} B x^{(k)} \theta_B^{(k)}\right)}$$

$$G = (V, A)$$

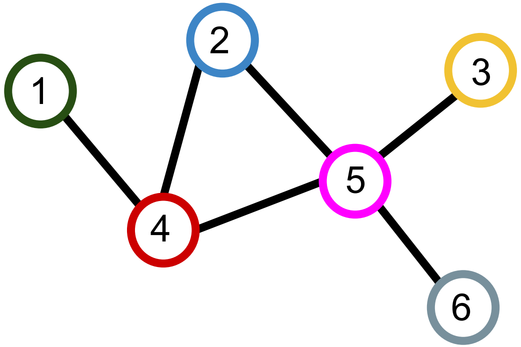


$$V = \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 3 & 2 \\ 1 & 1 \\ 0 & 1 \\ 3 & 1 \end{pmatrix} \quad A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} B x^{(k)} \theta_B^{(k)}\right)$$

Formalization

$$G = (V, A)$$



$$\mathcal{A} = \{I_{|V|}, A\}$$

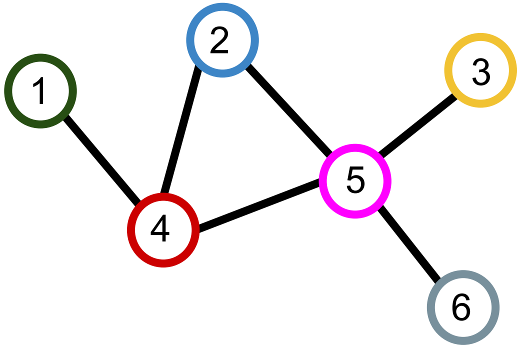
$$x^{(0)} = V \in \mathbb{R}^{|V| \times d_0}$$

$$x^{(1)} = \rho\left(I_{|V|} x^{(0)} \theta_{I_{|V|}}^{(0)} + A x^{(0)} \theta_A^{(0)}\right)$$

Formalization

$$x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} Bx^{(k)}\theta_B^{(k)}\right)$$

$$G = (V, A)$$



$$\mathcal{A} = \{I_{|V|}, A\}$$

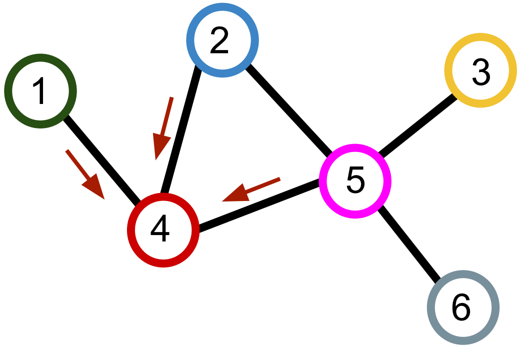
$$x^{(0)} = V \in \mathbb{R}^{|V| \times d_0}$$

$$x^{(1)} = \rho\left(x^{(0)}\theta_{I_{|V|}}^{(0)} + Ax^{(0)}\theta_A^{(0)}\right)$$

$$x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} B x^{(k)} \theta_B^{(k)}\right)$$

Formalization

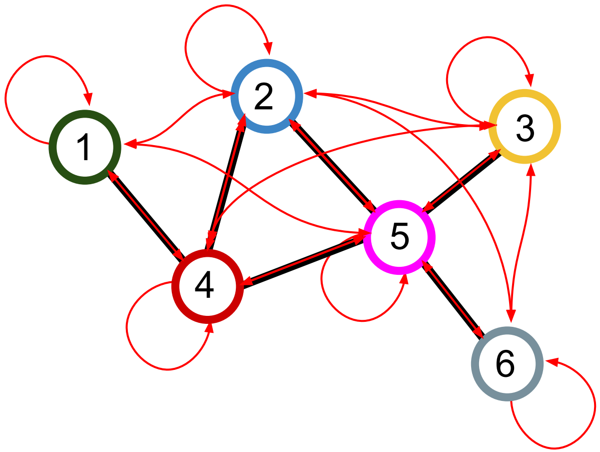
$$G = (V, A)$$



$$Ax^{(0)} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 3 & 2 \\ 1 & 1 \\ 0 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \\ \boxed{3} & 3 \\ 9 & 4 \\ 0 & 1 \end{pmatrix}$$

Formalization

$$x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} B x^{(k)} \theta_B^{(k)}\right)$$

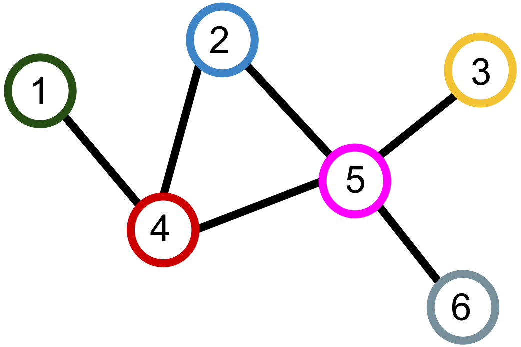


But what other operators can we use in $\mathcal{A}$?

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 3 & 1 & 0 \\ 1 & 1 & 0 & 1 & 4 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \quad \ldots$$

Formalization

$$\boxed{x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B\in\mathcal{A}^{(k)}} Bx^{(k)}\theta_B^{(k)}\right)}$$



But can we learn the operator we use in $\mathcal{A}$?

$$\phi_k(B) = \begin{pmatrix} \phi_k(B)_{0,0} & \cdots & \phi_k(B)_{0,5} \\ \vdots & \ddots & \vdots \\ \phi_k(B)_{5,0} & \cdots & \phi_k(B)_{5,5} \end{pmatrix}$$

$$\phi_k(B)_{i,j} = \begin{cases} 0 & \text{if } B_{i,j} = 0 \\ \sigma\left(\mathrm{MLP}_{\tilde{\theta}}\left(\left|x_i^{(k)} - x_j^{(k)}\right|\right)\right) & \text{otherwise} \end{cases}$$

Simple Message Passing Layer

Let us consider a graph $G = (V, A)$ where V is the set of nodes and A the adjacency matrix

$$x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} B x^{(k)} \theta_B^{(k)}\right)$$

V. Garcia and J. Bruna. *Few-Shot Learning with Graph Neural Networks*. ICLR, 2018

Frameworks

Deep Learning Frameworks

PyTorch

- Rapid prototyping in Research
- Dynamic computational graphs
- Debugging

TensorFlow

- Large-scale deployments
- Cross-platform and embedded deployment
- Static computational graphs

Graph Neural Networks Libraries

PyTorch

PyTorch geometric

- Pytorch-geometric
- Deep Graph Library

DeepGraphLibrary

TensorFlow

- Graph Nets (DeepMind)

Graph Neural Networks Libraries

PyTorch geometric

- Fast re-implementation of existing models
- Faster

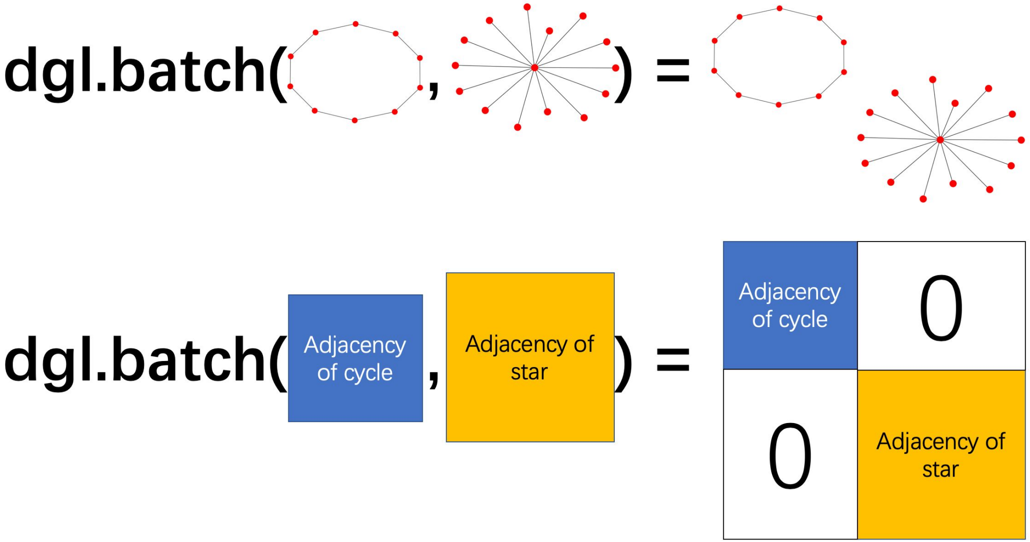
M. Fey and J.E. Lenssen. [Fast graph representation learning with PyTorch Geometric](). ICLR Workshop on Representation Learning on Graphs and Manifolds, 2019

DeepGraphLibrary

- Higher-level abstraction (auto-batching)
- No need to worry with sparse matrix multiplication

M. Wang *et al*. [Deep Graph Library: Towards Efficient And Scalable Deep Learning on Graphs](). ICLR Workshop on Representation Learning on Graphs and Manifolds, 2019

DeepGraphLibrary



M. Wang *et al*. [Deep Graph Library: Towards Efficient And Scalable Deep Learning on Graphs](). ICLR Workshop on Representation Learning on Graphs and Manifolds, 2019

Scratch Implementation

colab

Implementation

colab

Recommended Reading

- Jie Zhou, Ganqu Cui, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, Maosong Sun, "Graph Neural Networks: A Review of Methods and Applications.", arXiv 2018.
- Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu, "A Comprehensive Survey on Graph Neural Networks.", arXiv 2019.
- Ziwei Zhang, Peng Cui, Wenwu Zhu, "Deep Learning on Graphs: A Survey.", arXiv 2018.
- Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, Pierre Vandergheynst, "Geometric Deep Learning: Going beyond Euclidean data.", IEEE SPM 2017.
- Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, George E. Dahl, "Neural Message Passing for Quantum Chemistry.", ICML 2017

Recommended Reading

- “Geometric Deep Learning” http://geometricdeeplearning.com/
  - Workshops: ICCV, ECCV, BMVC, …
  - Tutorials: CVPR, NIPS, ECCV, SIGGRAPH, ...
- Steeve Huang, “Hands-on Graph Neural Networks with PyTorch & PyTorch Geometric”
- “DeepGraphLibrary Tutorial”

Discussion and Closing

- Are graphs still relevant?
- Are graph-based methods still useful for Pattern Recognition and Document Image Analysis?
- What are the current trends and next steps?

Centre de Visió per Computador

Informatique Image Interaction
Univ. La Rochelle

UNIVERSITY OF EXETER

ICDAR 2019
15th International Conference on Document Analysis and Recognition
20-25 September 2019 | International Convention Centre Sydney, Australia