Graph Embedding Problem

Romain Raveaux

LI lab (EA 6300) – Université de Tours

October 18, 2013

International Master of Research in Computer Science: Computer Aided Decision Support

Laboratoire d'Informatique
EA 6300

POLYTECH

UNIVERSITÉ
FRANÇOIS - RABELAIS
TOURS

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
    - Graph-Based representation
    - Graph classification
    - Graph comparison
    - Image analysis

About Me I

- Romain Raveaux
- Teacher at the university of Tours. Polytech'Tours
- Researcher at the Computer Science Laboratory
- Researcher Activity:
  - Graph-Based representation
  - Graph classification
  - Graph comparison
  - Image analysis

About Me II

- Main publications : Referenced International Journal
  - Romain Raveaux et al. Structured representations in a content based image retrieval context. Journal of Visual Communication and Image Representation, Volume 24, Issue 8, November 2013, Pages 1252-1268.
  - Romain Raveaux et al. A local evaluation of vectorized documents by means of polygon assignments and matching. IJDAR 15(1): 21-43 (2012)
  - Romain Raveaux et al. Learning graph prototypes for shape recognition. Computer Vision and Image Understanding 115(7): 905-918 (2011)
  - Romain Raveaux et al. A graph matching method and a graph matching distance based on subgraph assignments. Pattern Recognition Letters 31(5): 394-406 (2010)

About Me II

- Main publications : Referenced International Journal
  - Romain Raveaux et al. Structured representations in a content based image retrieval context. Journal of Visual Communication and Image Representation, Volume 24, Issue 8, November 2013, Pages 1252-1268.
  - Romain Raveaux et al. A local evaluation of vectorized documents by means of polygon assignments and matching. IJDAR 15(1): 21-43 (2012)
  - Romain Raveaux et al. Learning graph prototypes for shape recognition. Computer Vision and Image Understanding 115(7): 905-918 (2011)
  - Romain Raveaux et al. A graph matching method and a graph matching distance based on subgraph assignments. Pattern Recognition Letters 31(5): 394-406 (2010)

About Me II

- Main publications : Referenced International Journal
  - Romain Raveaux et al. Structured representations in a content based image retrieval context. Journal of Visual Communication and Image Representation, Volume 24, Issue 8, November 2013, Pages 1252-1268.
  - Romain Raveaux et al. A local evaluation of vectorized documents by means of polygon assignments and matching. IJDAR 15(1): 21-43 (2012)
  - Romain Raveaux et al. Learning graph prototypes for shape recognition. Computer Vision and Image Understanding 115(7): 905-918 (2011)
  - Romain Raveaux et al. A graph matching method and a graph matching distance based on subgraph assignments. Pattern Recognition Letters 31(5): 394-406 (2010)

About Me II

- Main publications : Referenced International Journal
  - Romain Raveaux et al. Structured representations in a content based image retrieval context. Journal of Visual Communication and Image Representation, Volume 24, Issue 8, November 2013, Pages 1252-1268.
  - Romain Raveaux et al. A local evaluation of vectorized documents by means of polygon assignments and matching. IJDAR 15(1): 21-43 (2012)
  - Romain Raveaux et al. Learning graph prototypes for shape recognition. Computer Vision and Image Understanding 115(7): 905-918 (2011)
  - Romain Raveaux et al. A graph matching method and a graph matching distance based on subgraph assignments. Pattern Recognition Letters 31(5): 394-406 (2010)

About Me II

- Main publications : Referenced International Journal
  - Romain Raveaux et al. Structured representations in a content based image retrieval context. Journal of Visual Communication and Image Representation, Volume 24, Issue 8, November 2013, Pages 1252-1268.
  - Romain Raveaux et al. A local evaluation of vectorized documents by means of polygon assignments and matching. IJDAR 15(1): 21-43 (2012)
  - Romain Raveaux et al. Learning graph prototypes for shape recognition. Computer Vision and Image Understanding 115(7): 905-918 (2011)
  - Romain Raveaux et al. A graph matching method and a graph matching distance based on subgraph assignments. Pattern Recognition Letters 31(5): 394-406 (2010)

What is Graph Embedding

Transform graphs into numeric feature vectors and exploit computational strengths of state of the art statistical pattern recognition.

What is Graph Embedding

Formally, a graph embedding is a function $G \rightarrow \mathbb{R}^n$ mapping graphs from an arbitrary graph domain to a vector space.

What is Graph Embedding

There are at least 2 way for computing $G \rightarrow \mathbb{R}^n$:

- Explicit through feature extraction or dissimilarities
- Implicit through graph kernel

What is a kernel

A kernel, in this context, is a symmetric continuous function that maps
$K : [a, b] \times [a, b] \rightarrow \mathbb{R}$
where symmetric means that $K(x, s) = K(s, x)$. K is said to be non-negative definite .

What is a kernel

The inner product of two vectors a = (a1, a2, ..., an) and b = (b1, b2, ..., bn) is defined as:
$\langle \mathbf{a} \cdot \mathbf{b} \rangle = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

Implicit graph embedding

Definition

Graph Kernel Let G be a (finite or infinite) set of graphs. Function $k : G \times G \rightarrow \mathbb{R}$ is called a graph kernel if there exists a possibly infinite-dimensional Hilbert space F and a mapping $\phi : G \rightarrow f$ such that

- $k(g, g') = \langle \varphi(g), \varphi(g') \rangle$
- $\forall g, g' \in G$ where $\langle ., . \rangle$ denotes a dot product in F.

Kernel Trick

M. Aizerman, E. Braverman, and L. Rozonoer, "Theoretical foundations of the potential function method in pattern recognition learning", Automation and Remote Control, vol. 25, 1964, p. 821-837

Kernel Trick

Definition

Kernel Trick

- Let $\vec{x}$ and $\vec{y}$ be two vectors in $\mathbb{R}^n$
- Let $\varphi(\vec{x})$ and $\varphi(\vec{y})$ be two functions projecting $\vec{x}$ and $\vec{y}$ into $\mathbb{R}^m$.
- with $n \leq m$
- $k(\vec{x}, \vec{y}) = \langle \varphi(\vec{x}), \varphi(\vec{y}) \rangle$
- An explicit representation for φ is not required.
- It suffices to know that $\mathbb{R}^m$ is an inner product space

Kernel Trick

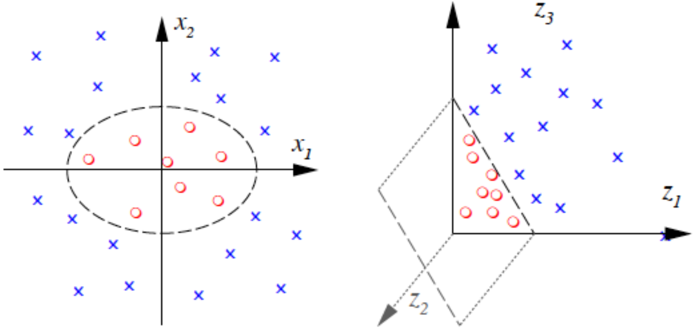
Example:

- Let $\vec{x}$ and $\vec{y}$ be two vectors in $\mathbb{R}^2$
- Let $\vec{x} = (x_1, x_2)$ and $\vec{y} = (y_1, y_2)$
- Let $\varphi(\vec{x})$ and $\varphi(\vec{y})$ be two functions projecting $\vec{x}$ and $\vec{y}$ into $\mathbb{R}^3$.
- $k(\vec{x}, \vec{y}) = \langle \vec{x}, \vec{y} \rangle^2$
- $k(\vec{x}, \vec{y}) = x_1^2 y_1^2 + 2x_1 y_1 x_2 y_2 + y_2^2 y_2^2$
- $k(\vec{x}, \vec{y}) = \langle (x_1^2, \sqrt{2}x_1 x_2, x_2^2), (y_1^2, \sqrt{2}y_1 y_2, y_2^2) \rangle$
- $k(\vec{x}, \vec{y}) = \langle \varphi(\vec{x}), \varphi(\vec{y}) \rangle$
- $\varphi(\vec{x}) = (x_1^2, \sqrt{2}x_1 x_2, x_2^2)$

Kernel Trick

Example:

$$\Phi : R^2 \rightarrow R^3$$
$$(x_1, x_2) \mapsto (z_1, z_2, z_3) := (x_1^2, \sqrt(2) x_1 x_2, x_2^2)$$



Graph Kernel Families

- Diffusion kernels (from similarity matrix)
- Convolution kernel
- Walk kernel (from adjacency matrix)

Explicit graph embedding

$\varphi : G \rightarrow \mathbb{R}^n$

- φ is explicitly computed
  - Feature extraction
  - Using dissimilarities

Explicit graph embedding using dissimilarities

Definition

Dissimilarity graph embedding

- Let us assume a graph domain G is given. If $\tau = \{g_1, \ldots, g_n\} \in G$ is a training set with n graphs and
- $P = \{P_1, \ldots, P_n\} \subseteq \tau$ is a prototype set with n graphs, the mapping
- $\varphi_n^p : G \to \mathbb{R}^n$
- is defined as the function
- $\varphi_n^p = (d(g, P_1), \ldots, d(g, P_n))$

Explicit graph embedding using features

- Let g be a graph $\in G$
- $\varphi : G \rightarrow \mathbb{R}^n$
- Let $\vec{v} = \varphi(g)$
- v_1 = Number of nodes
- v_2 = Number of edges
- v_3 = Number of paths of length 2
- v_4 = Number of path of length 3
- $v_5 = \ldots$

Conclusion

- The key idea of these new solutions is to map the structural patterns into a vector or a dot product space.
- Structural pattern recognition allows us to represent objects in terms of their parts, which is a clear advantage over statistical pattern recognition.
- But it lacks of suitable algorithmic tools for pattern classification and clustering.
- On the other hand, statistical pattern recognition offers a wealth of mathematical tools for classification, clustering, and similar tasks, but it is restricted in its representational power.