Graph kernels

Luc Brun

GREYC – CNRS UMR 6072, University of Caen, ENSICAEN
Luc.Brun@ensicaen.fr

In close collaboration with:
Francois Xavier Dupé (LIF)
Benoit Gauzere (GREYC)
Didier Villemin (LCMT)
Donatello Conte (MIVIA)
Mario Vento (MIVIA)

Séminaire LORIA 2012

Motivation

Structural Pattern Recognition

- 🙂 Rich description of objects
- 🙁 Poor properties of graph's space does not allow to readily generalize/combine sets of graphs

Statistical Pattern Recognition

- 🙁 Global description of objects
- 🙂 Numerical spaces with many mathematical properties (metric, vector space, ...).

Motivation

Analyse large famillies of structural and numerical objects using a **unified** framework based on pairwise similarity.

Kernels : Definition

- A kernel k is a **symmetric** similarity measure on a set χ

$$\forall (x,y) \in \chi^2,\ k(x,y) = k(y,x)$$

- k is said to be **definite positive** (d.p.) iff k is symmetric and iff:

$$\left.\begin{array}{lcl} \forall (x_1,\dots,x_n) & \in & \chi^n \\ \forall (c_1,\dots,c_n) & \in & \mathbb{R}^n \end{array}\right\} \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j k(x_i,x_j) \geq 0$$

- $K = (k(x_i,x_j))_{(i,j)\in\{1,\dots,n\}}$ is the Gramm matrix of k. k is d.p. iff:

$$\forall c \in \mathbb{R}^n - \{0\},\ c^t K c \geq 0$$

Kernels and scalar products

Aronszajn 1950 :

A kernel k is d.p. on a space χ
if and only if
it exists

- one Hilbert space $\mathcal{H}$ and
- a function $\varphi : \chi \rightarrow \mathcal{H}$

such that:

$$k(x, y) = < \varphi(x), \varphi(y) >$$

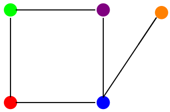
Outline

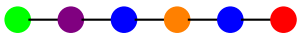
Kernel and strutured data

The kernel trick provides an implicit embedding whose metric is defined from our similarity criterion (the kernel).

Walk kernels

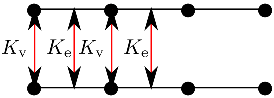
- Walks: Let $G = (V, E)$. $W = (v_1, \dots, v_n)$ is a walk iff $(v_i, v_{i+1}) \in E, \forall i \in \{1, \dots, n-1\}$.



 is a walk

- Kernel between walks

$$K(h,h') = \begin{cases} 0 & \text{if } |h| \neq [h'| \text{ and} \\ K_v(v_1, v'_1).\Pi_{i=1}^{|h|} K_e(e_i, e'_i) K_v(v_{i+1}, v'_{i+1}) & \text{otherwize} \end{cases}$$



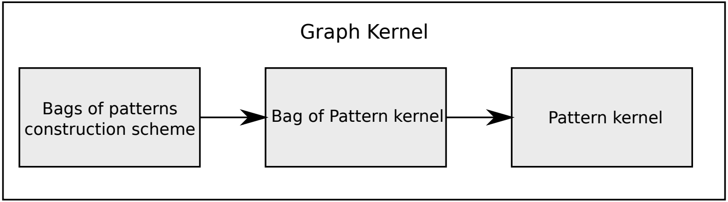
- Walk kernels :

$$K(G_1, G_2) = \sum_{h \in \mathcal{W}(G_1)} \sum_{h' \in \mathcal{W}(G_2)} K(h,h') \lambda_{G_1}(h) \lambda_{G_2}(h')$$

Finite Bag kernels

$$\left.\begin{array}{ccc} G & \rightarrow & B(G) \\ G' & \rightarrow & B(G') \end{array}\right\} K(G,G') = K(B(G),B(G'))$$

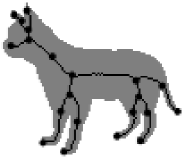
Three independent step to design a graph kernel.



Questions :

Which type of graph best capture object's properties ?
Shape recognition:

- Shape described by skeletons encoded by graphs $G = (V, E, \omega)$.
  - V:
    - extremities/intersection of branches or
    - abrupt changes of the radius along a branch.
  - E:
    - branches of the skeleton.
    - Each edge is weighted by a function ω encoding the length of the boundary which contributed to the branch [Torsello04].



Chemoinformatic: $G = (V, E, \mu, \nu)$

- V set of atoms, μ atom's type,
- E set of bounds, ν bound's type.

Questions :

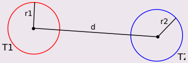
- Which type of graph best capture object's properties ?
- Which pattern should we choose ?
  - Paths,
  - Sub trees,
  - Sub graphs.
  - Sub combinatorial maps,...
- How to build a bag which minimizes the loss of information ?
  - All patterns up to a given size ?
  - More relevant patterns ?
    - Shape recognition: A priori information $\rightarrow$ covering problem
    - Chemoinformatics: Multiple kernel learning with one kernel per pattern.
- How to define a kernel between bags

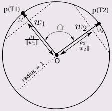
Bag of paths kernels
shape recognition

Haussler99

$$K_{\text{mean}}(T_1, T_2) = \frac{1}{|T_1|}\frac{1}{|T_2|}\sum_{t\in T_1}\sum_{t\in T_2} K_{\text{pattern}}(t, t'),$$

More complex kernels





(Desobry 2005)

$$K'(x, y) = \frac{K(x,y)}{\sqrt{K(x,x)K(y,y)}}$$
$$K'(x, x) = \|\varphi'(x)\|^2 = 1$$

Normalized kernel

Weighted mean kernel

$$\begin{array}{lcl} K_{\text{weighted}}(T_1, T_2) & = & \frac{1}{|T_1|}\frac{1}{|T_2|}\sum_{t\in T_1}\sum_{t'\in T_2}\lambda_{T_1}(t)\lambda_{T_2}(t')K_{\text{pattern}}(t, t'), \\ \lambda_{T_i}(t) & = & <\varphi(t), \mu_{T_i}>^d \end{array}$$

Bag of paths kernels
chemoinformatic

Basic

$$K_{\text{mean}}(G_1, G_2) = \sum_{t \in T(G_1) \cap T(G_2)} K(f_t(G_1), f_t(G_2)),$$

Basic MKL

$$K_{\text{BMKL}}(G_1, G_2) = \sum_{t \in T(G_1) \cap T(G_2)} d_t K(f_t(G_1), f_t(G_2)),$$

Infinite MKL

$$K_{\text{IMKL}}(G_1, G_2) = \sum_{t \in \mathcal{T}} \sum_{t' \in \mathcal{T}, t' \leq t} d_{t,t'} K(f_t(G_1), f_{t'}(G_2)),$$

where $\mathcal{T} = T(G_1) \cup T(G_2)$

Pattern kernels
How to compare patterns

- Classical approach:

$$K_{\text{classic}}(t,t') = \begin{cases} 0 \quad \text{if } t \not\sim t', \\ \sum_{\varphi \in Isom(t,t')} \left(\prod_{i=1}^{|V|} K_v(\mu(v_i), \varphi(\mu(v_i' i))) \right) \\ \left(\prod_{i=1}^{|E|} K_e(\nu(e_i), \nu(e_i')) \right) \end{cases}$$

- Pb: Similar but not identical patterns are not comparable.
- Solutions :
  - Shape recognition : Use rewrittings
  - Chemoinformatics : Use edit distance.

Conclusion

- Our aim : Studdy properties of structure space and combine structural pattern recognition and machine learning/optimisation methods.

Thank you for your attention !

Questions ?

References

Indefinite Kernels Feature Space Interpretation of SVM with Indefinite Kernels, Bernard Haasdonk, PAMI 27(4) 2005

Kernels and Graph edit distance Bridgining the Gap Between Graph Edit Distance and kernel machines, Michel Neuhauss and Horst Bunke (Machine perception, AI, vol. 68).

Convolution Kernel Convolution Kernels on Discrete Structures TR: UCSC-CRL-99-10, David Haussler

Diffusion kernels Diffusion Kernels on graphs and Other Discrete Structures, Risi Imre Kondor, John Lafferty. ICML 2002

Kernels and infinite Bags

- Graph kernels, Journal of Machine Learning Research 11(2010) 1201-1242, S.V.N. Vishwanathan, Nicol N. Schraudolph, Risi Kondor, Karsten M. Borgwadt
- Marginalized Kernels Between Labeled Graphs, H. Kashima, K. Tsuda and A. Inokuchi, ICML-2003
- Graph kernels based on tree patterns for molecules, Pierre Mahé and Jean-Philippe Vert, Machine Learning 75(1) 3-35

Kernels and finite Bags

- Dupe, F. -X. & Brun, L.. Edition within a graph kernel framework for shape recognition. In Graph Based Representation in Pattern Recognition 2009 , pages 11-21 2009 .
- Next talk.