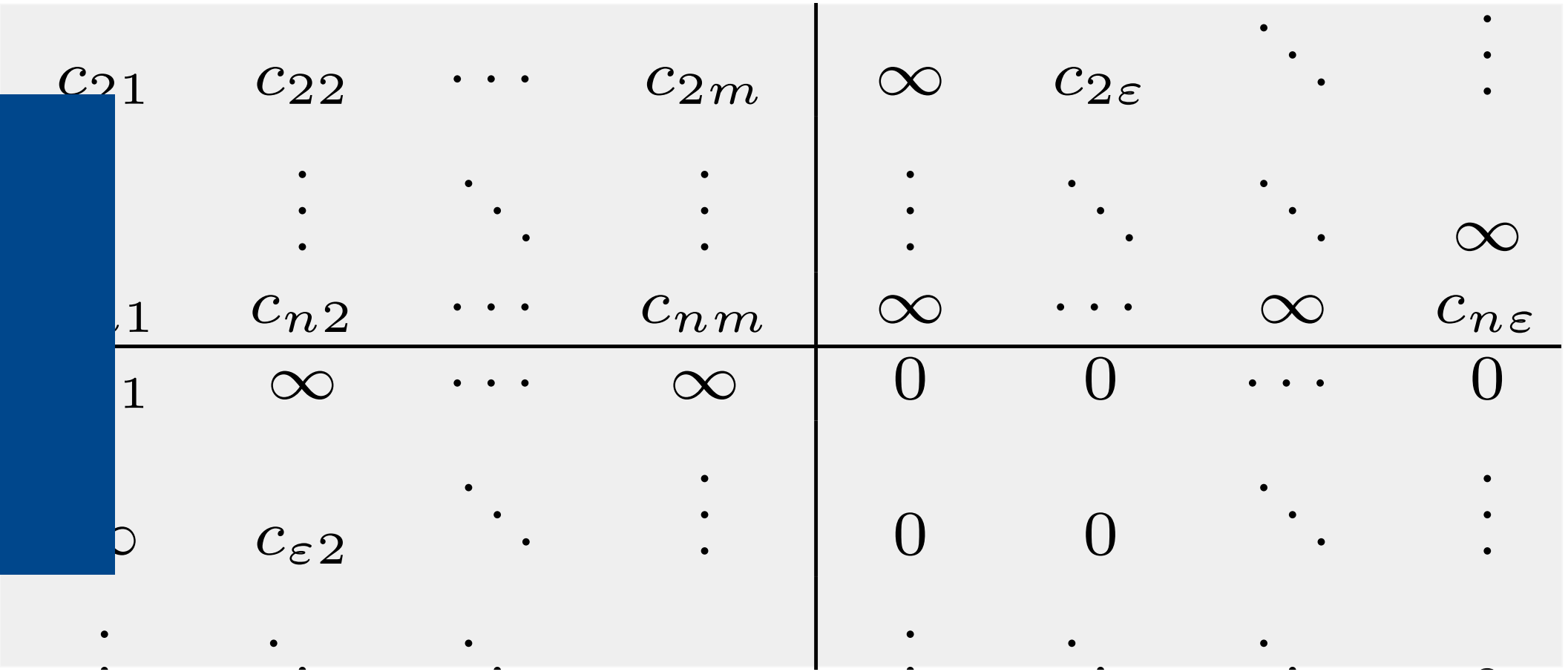
# GBR 2017

**A Survey on Applications of Bipartite Graph Edit Distance**

Michael Stauffer, Thomas Tschachtli, Andreas Fischer, and Kaspar Riesen

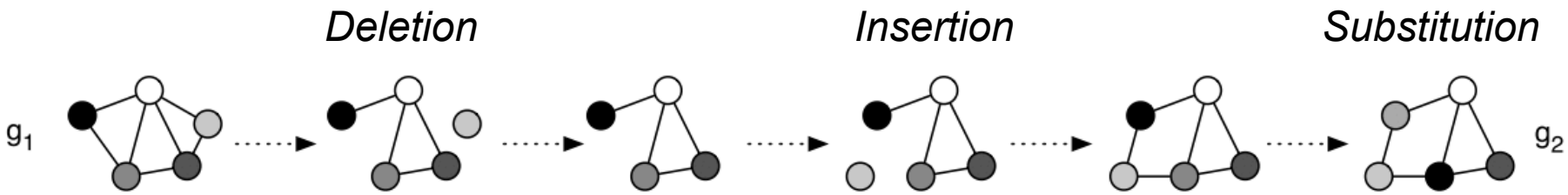


## Content

- Bipartite Graph Edit Distance (BP)
- Applications of BP
- Extensions of BP
- Conclusion + Future Work
- Q+A

**Bipartite Graph Edit Distance** – Graph Edit Distance

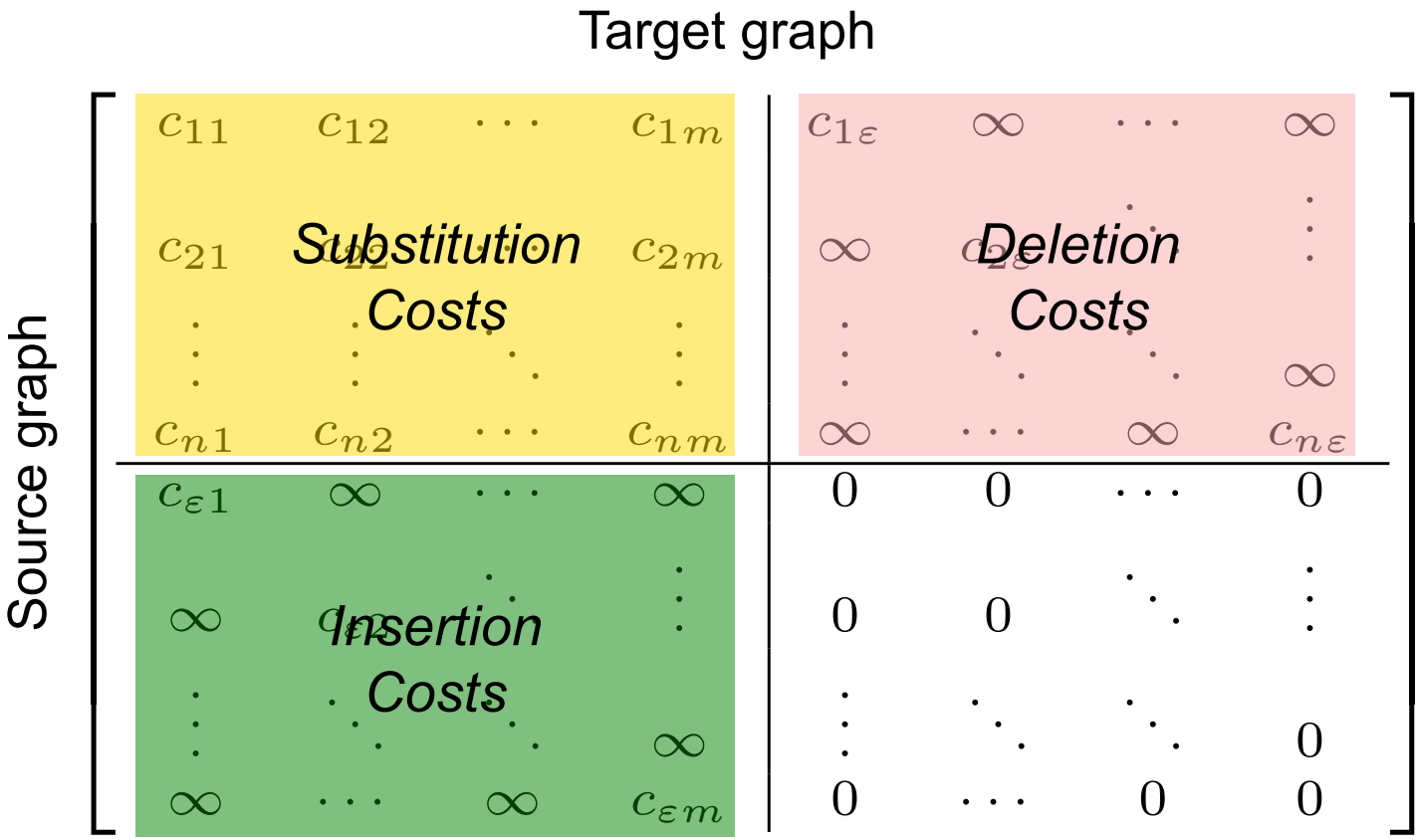
Graph Edit Distance (GED) is a popular measure for inexact graph matching, it measures the **minimum amount of distortion** that is needed to **transform one graph into another**.




Exact GED is exponential in the number of nodes of the involved graphs and is therefore applicable on rather small graphs only.

➡ Bipartite Graph Edit Distance (BP) approximates GED in polynomial time

**Bipartite Graph Edit Distance** – Approximation for GED




Approximative GED via **L**inear **S**um **A**ssignment **P**roblem (i.e. Hungarian method)

Cost matrix with node (and local edge) edit costs

## **Applications of BP** – Taxonomy

In the last ten years, the BP framework has been citied more than 360 times.
In the following we present a selection of BP applications.

- Image Analysis
- Handwritten Document Analysis
- Biometrics
- Bio- and Chemoinformatics
- Knowledge and Process Management
- Malware Detection
- Other Applications

## **Applications of BP** – Image Analysis

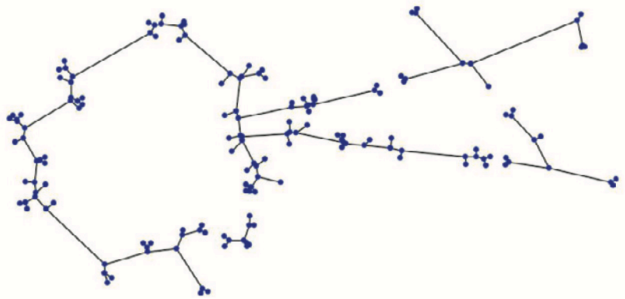
**A Graph-based approach for Kite recognition**

K. Madi, H. Seba, H. Kheddouci, O. Barge, Pattern Recognition Letters (2017)

Automatic recognition of historical stone structures (so called Kites) in large satellite image databases using graph-based representations and matchings.

Kite Image

Skeletonised Kite Image



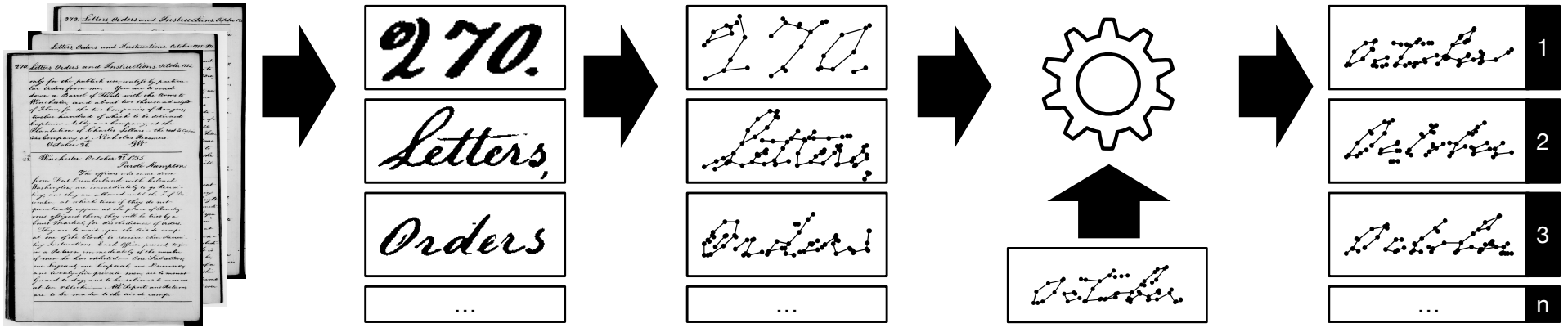
Graph

## **Applications of BP** – Handwritten Document Analysis

**Graph-based Keyword Spotting in Historical Handwritten Documents**

M.Stauffer, A. Fischer, K. Riesen, S+SSPR (2016)

Arbitrary retrieval of keywords (represent by a query graph) in historical handwritten documents (represented by a set of document graphs).



## **Applications of BP** – Biometrics

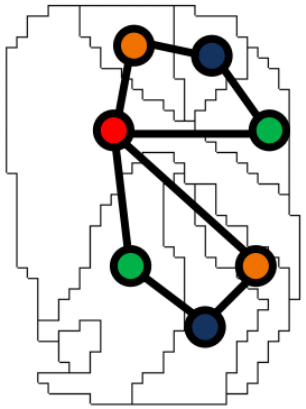
**Graph-based fingerprint classification using orientation field in core area**

Y. Choi, G. Kim, IEICE Electronics Express (2010)

Automatic identification of fingerprint images by means of graph-based representations and matching procedures.

Fingerprint image

Core region and segmented image



Graph

**Applications of BP** – Bio- and Chemoinformatics

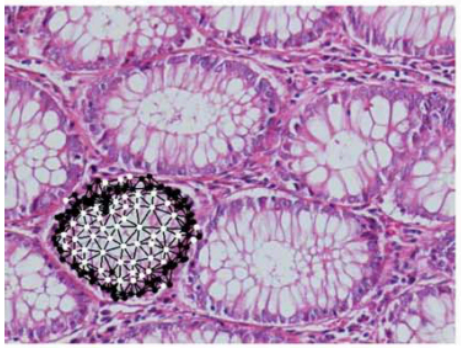
**A Hybrid Classification Model for Digital Pathology Using Structural and Statistical Pattern Recognition**

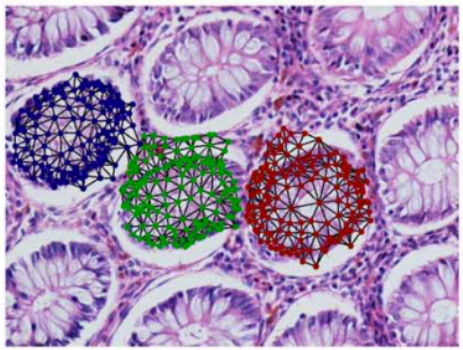
E. Ozdemir, C. Gunduz-Demir, IEEE Transactions on Medical Imaging (2010)

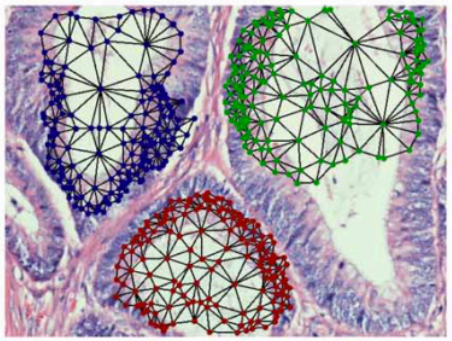
Identification of cancerous tissues by means of graph-based representation of query and tissue images.

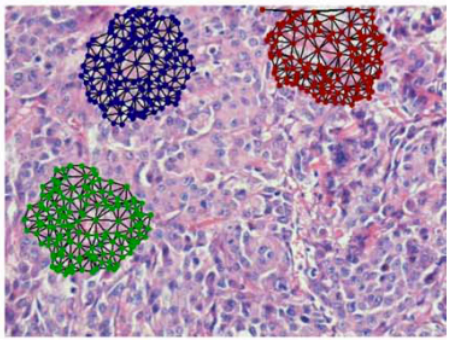
Query Graph

Tissue Graphs









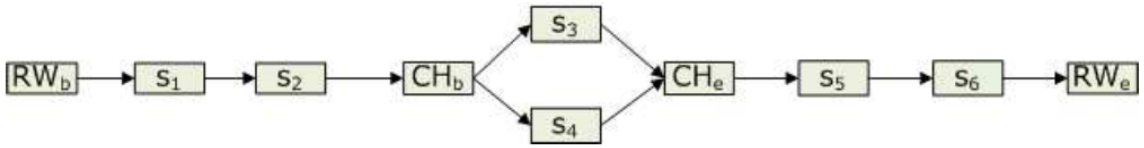
*Normal* *Low-grade* *High-grade*

**Applications of BP** – Knowledge and Process Management

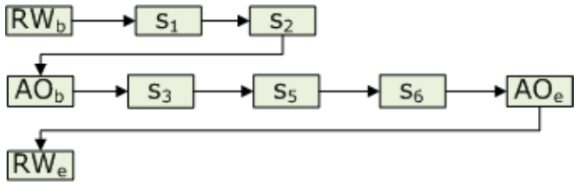
## Graph-based matching of composite OWL-S services

A. Cuzzocrea, J.L. De Coi, M. Fisichella, D. Skoutas, International Conference on Database Systems for Advanced Applications (2011)

Retrieval (discover) of similar composite web services (OWL-S) by means of graph matching. Nodes represent process states and atomic services, while edges are used to represent the control flow.




Query Graph



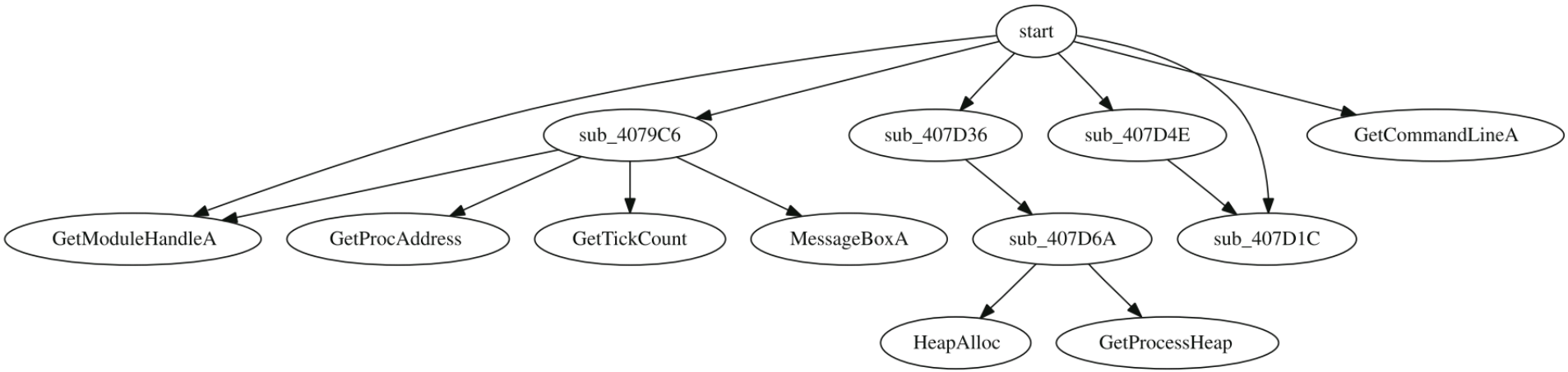

Next Process Graph

## **Applications of BP** – Malware Detection

**Malware Classification based on Call Graph Clustering**

J. Kinable, O. Kostakis, Computer Virology (2011)

Function calls of malicious software is represented by means of call graphs. Thus, similar malware can be found by means of graph matching and clustering.

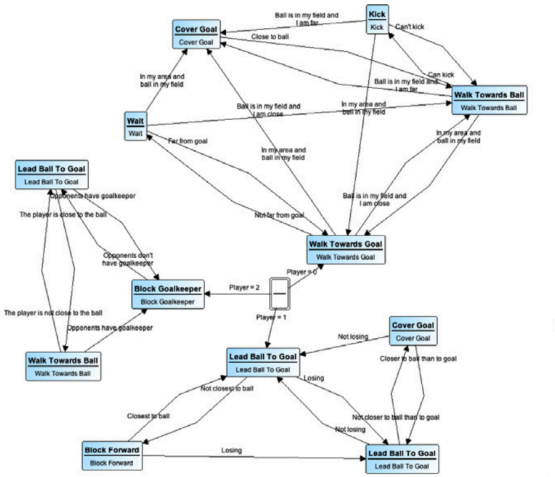



Malware Call Graph

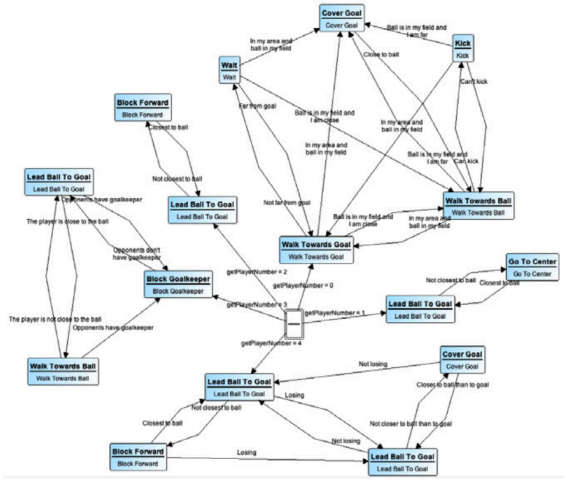
## **Applications of BP** – Other Applications

**Supporting sketch-based retrieval from a library of reusable behaviours**
G. Flórez-Puga, P.A. González-Calero, G. Jiménez-Díaz, B. Díaz-Agudo, Expert Systems with Applications (2013)

Retrieval of behaviour sketches for non-player characters in computer games.




Query Graph




Resulting Graph

## Extensions of BP

The BP framework has not only been applied, but also extended with respect to speed, accuracy, approximation, and problem domain.

Selected papers of each extension domain are given below (this list makes no claim to be complete)

**Speed**

[1] W. Jones, A. Chawdhary, A. King, Revisiting Volgenant-Jonker for Approximating Graph Edit Distance, in: GbR, 2015: pp. 98–107.

[2] F. Serratosa, Fast computation of Bipartite graph matching, Pattern Recognit. Lett. 45 (2014) 244–250.

[3] A. Fischer, C.Y. Suen, V. Frinken, K. Riesen, H. Bunke, Approximation of graph edit distance based on Hausdorff matching, Pattern Recognit. 48 (2015) 331–343.

## Extensions of BP

### Speed

[4] K. Riesen, M. Ferrer, A. Fischer, H. Bunke, Approximation of Graph Edit Distance in Quadratic Time, in: GbR, 2015: pp. 3–12.

[5] F. Serratosa, Speeding up Fast Bipartite Graph Matching Through a New Cost Matrix, Int. J. Pattern Recognit. Artif. Intell. 29 (2015) 1550010.

### Accuracy

[6] V. Carletti, B. Gaüzère, L. Brun, M. Vento, Approximate Graph Edit Distance Computation Combining Bipartite Matching and Exact Neighborhood Substructure Distance, in: GbR, 2015: pp. 188–197.

[7] B. Gaüzère, S. Bougleux, K. Riesen, L. Brun, Approximate Graph Edit Distance Guided by Bipartite Matching of Bags of Walks, in: S+SSPR, 2014: pp. 73–82.

[8] F. Serratosa, X. Cortés, Graph Edit Distance: Moving from global to local structure to solve the graph-matching problem, Pattern Recognit. Lett. 65 (2015) 204–210.

## Extensions of BP

### Accuracy

[9] X. Cortés, F. Serratosa, C.F. Moreno-García, On the Influence of Node Centralities on Graph Edit Distance for Graph Classification, in: GbR, 2015: pp. 231–241.

### Approximation

[10] Z. Abu-Aisheh, R. Raveaux, J.-Y. Ramel, Anytime graph matching, Pattern Recognit. Lett. 84 (2016) 215–224.

### Problem Domain

[11] S. Bougleux, L. Brun, V. Carletti, P. Foggia, B. Gaüzère, M. Vento, Graph edit distance as a quadratic assignment problem, Pattern Recognit. Lett. 87 (2017) 38–46.

### Others

[12] C.F. Moreno-García, F. Serratosa, X. Cortés, Consensus of Two Graph Correspondences Through a Generalisation of the Bipartite Graph Matching, in: GbR, 2015: pp. 87–97.

## Conclusion + Future Work

- Graph Edit Distance (GED) is a flexible paradigm for inexact graph matching that suffers from exponential time complexity.
- BP approximates GED in cubic time. This makes GED applicable in diverse complex domains.
- The present paper reviews applications of BP stemming from seven domains (image analysis, biometrics, chemoinformatics, …)
- In future work we plan to give a survey on extensions of BP with respect to the basic method.

## Q+A

?!