

KONG: Kernels for ordered-neighborhood graphs

NeurIPS conference (poster #122)

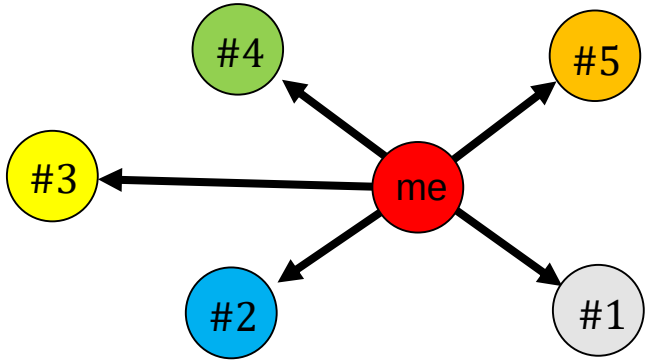
Authors: Moez Draief, Konstantin Kutzkov, **Kevin Scaman**, Milan Vojnovic
Date: November 30, 2018

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE | HUAWEI

Background

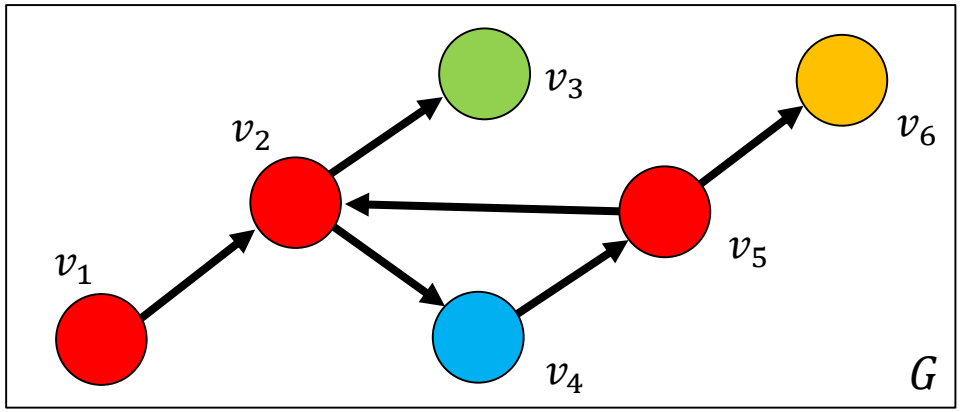
- Graphs are **highly complex objects**
  - Combinatorial nature of the object
  - Many relevant features
    - size, connectivity, density, hubs, periphery, short range patterns, large-scale structure, cliques, connected components, spectral characteristics…
  - **How to make it usable for ML problems?**

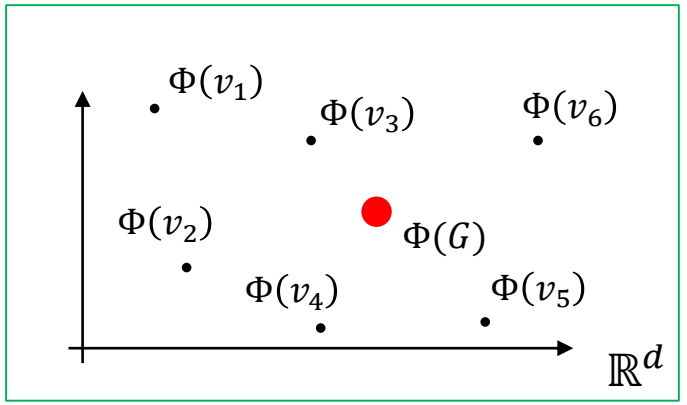
- Additional information: **ordered neighborhoods**
  - All edges may not be as important (e.g. friends on a social network)
  - Networks are often dynamic objects, changing through time
  - We may have a **ranking among neighbors**
    - Time of creation, importance, objective value, distance,…
  - **How to account for this information?**



The KONG algorithmic framework

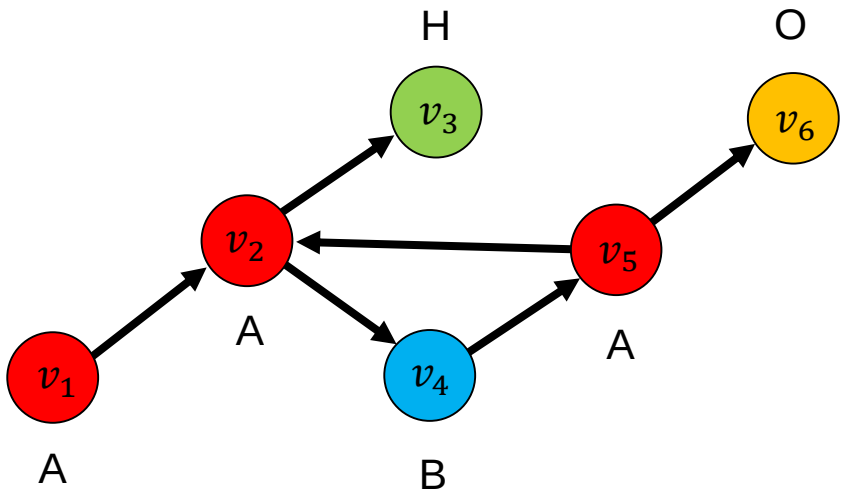
- **A scalable kernel representation for graphs**
  1) **Iterative** algorithm for node representation
     - Weisfeiler-Lehman, breadth-first search…
  2) Ordered neighborhood representation using **string kernels**
     - K-gram counting approach, order captured by selection process
  3) Refined k-gram counting using **polynomial** or **cosine** kernels
     - More powerful representation
  4) **Sketching** method for kernel approximation
     - Approximate embedding of counting vectors preserving scalar products



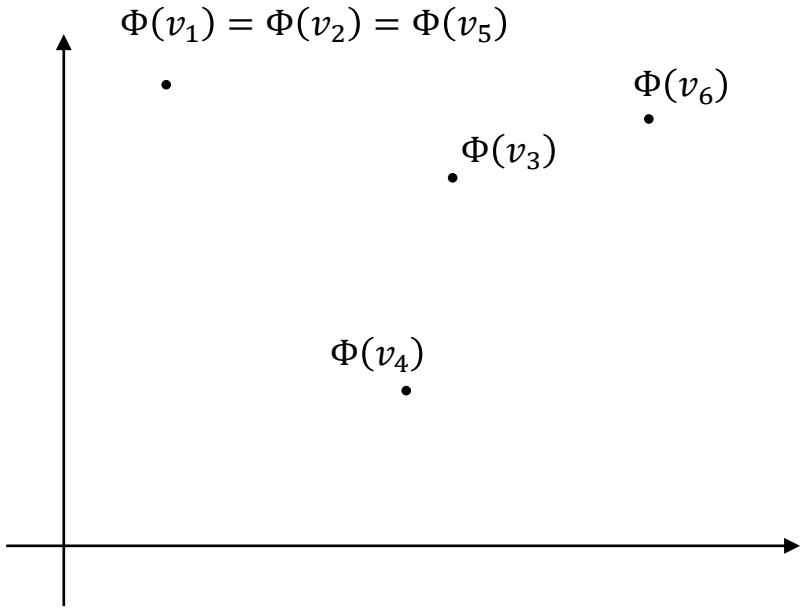


Simple example

Graph with string representations

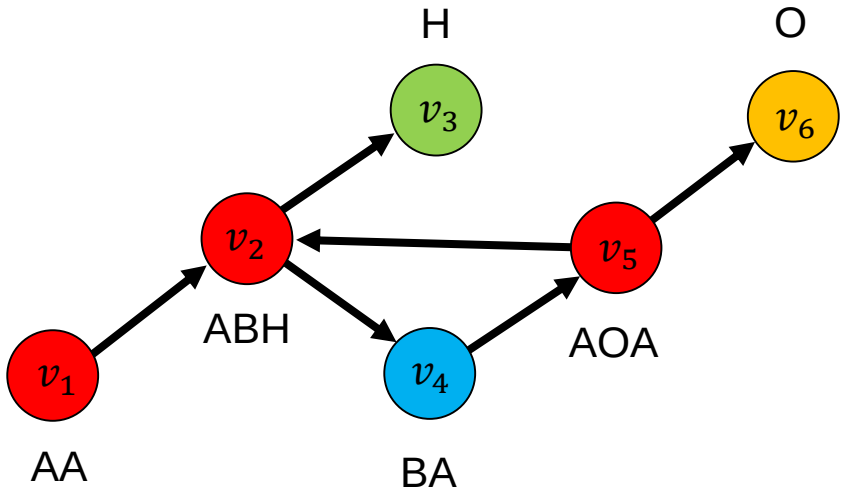


Sketching representation in $\mathbb{R}^d$

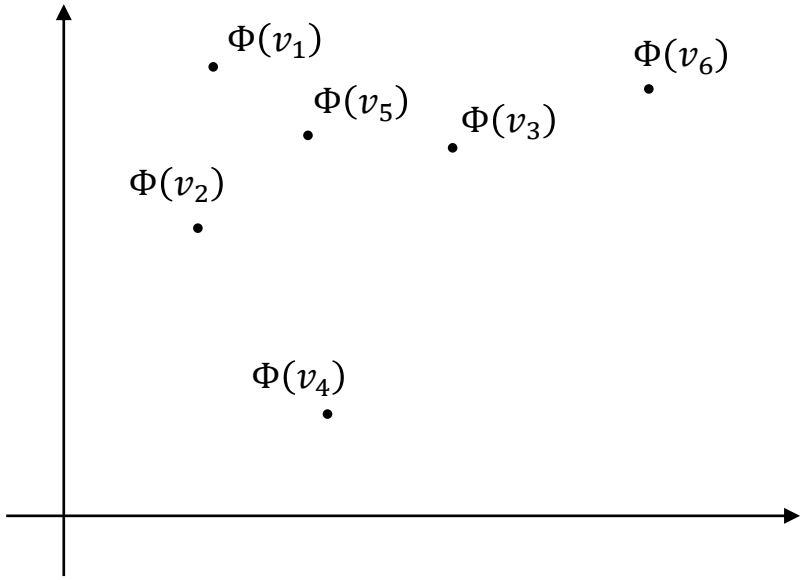


Simple example

Graph with string representations

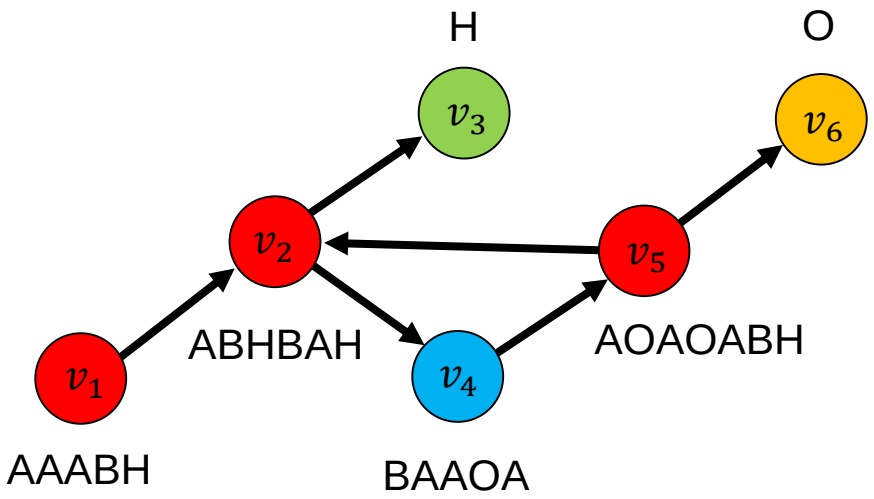


Sketching representation in $\mathbb{R}^d$

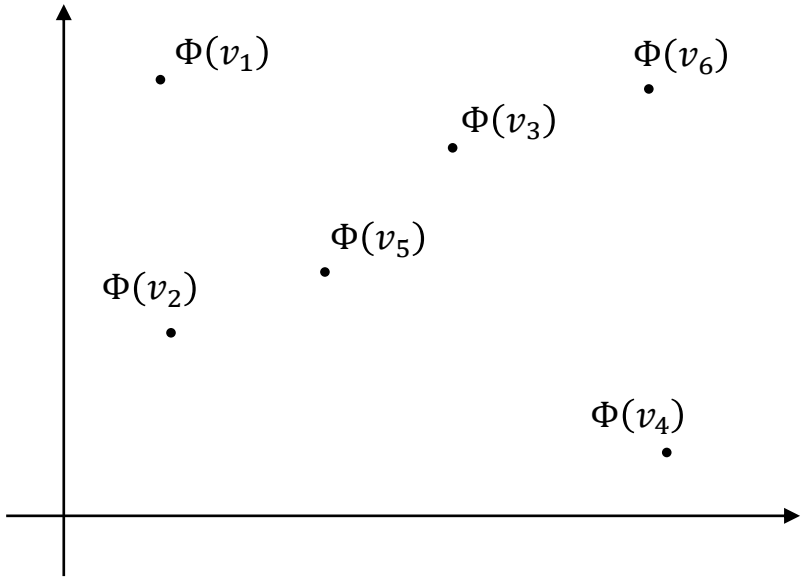


Simple example

Graph with string representations

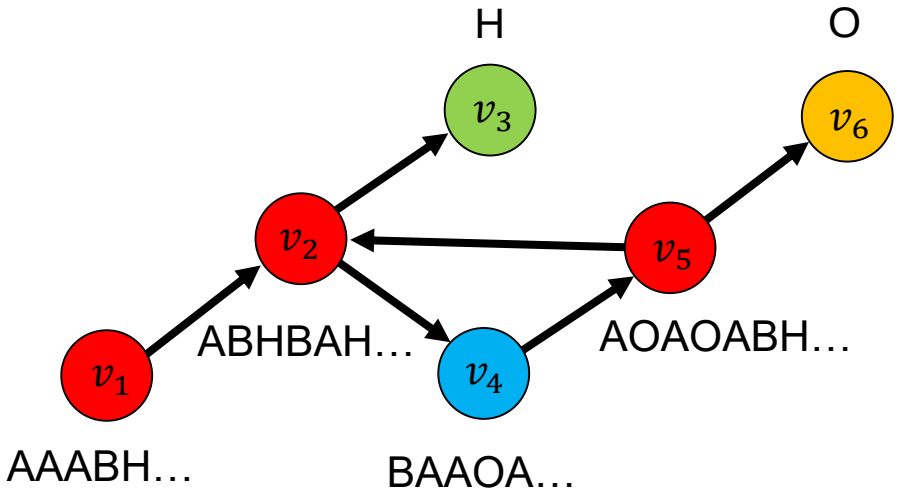


Sketching representation in $\mathbb{R}^d$

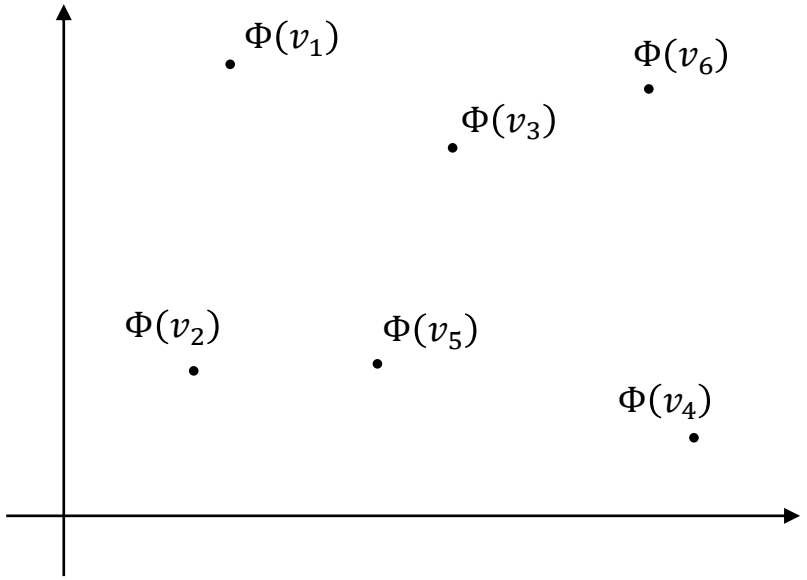


Simple example

Graph with string representations

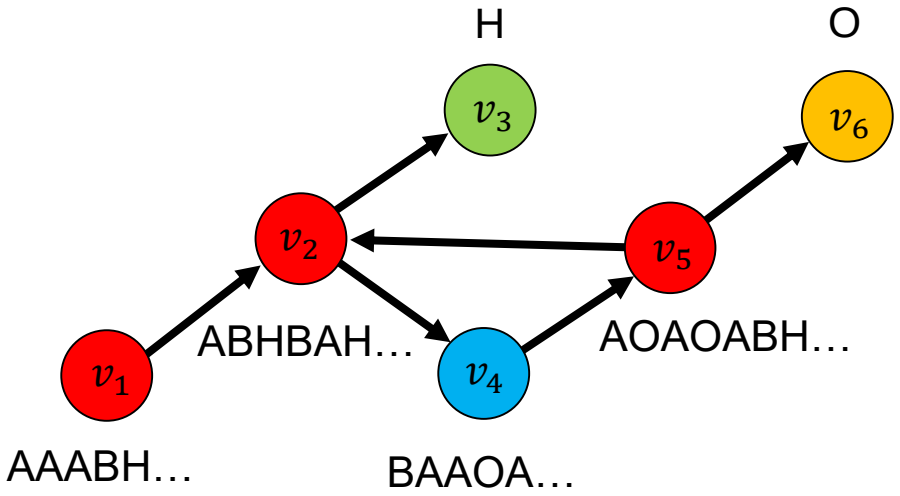


Sketching representation in $\mathbb{R}^d$

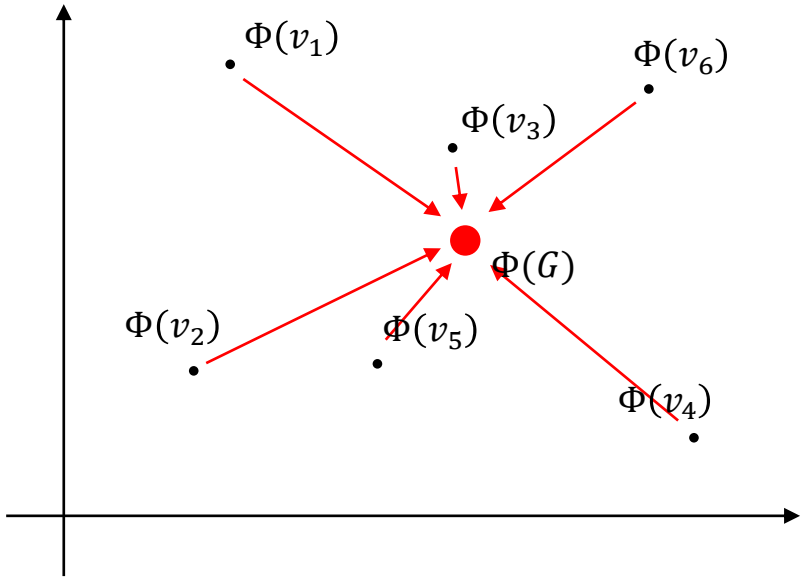


Simple example

Graph with string representations



Sketching representation in $\mathbb{R}^d$



Conclusion

- **The KONG framework: a new scalable algorithm for graphs kernels**
  - First method using **ordered neighborhoods**,
  - **Highly scalable** approach that can handle graphs with millions of nodes in seconds on a laptop in a single-threaded implementation,
  - **Flexibility** in the choice of the kernel function,
  - Outputs **vector representations**
    - Can be used by any ML algorithm for regression, classification, clustering, etc…
  - Excellent results on datasets from various domains, including
    - Anomaly detection in network flow graphs,
    - Gender prediction in recommender systems,
    - Affluence prediction in customer purchase graphs.

Poster #122