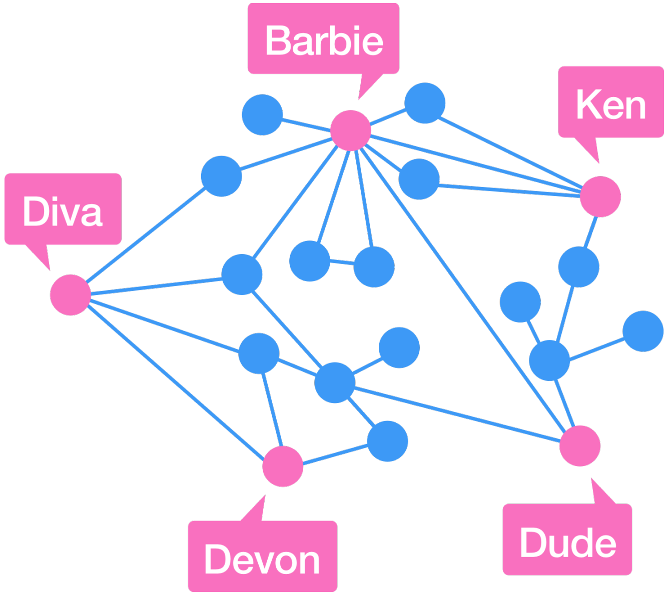
Iowa State University

Department of Computer Science

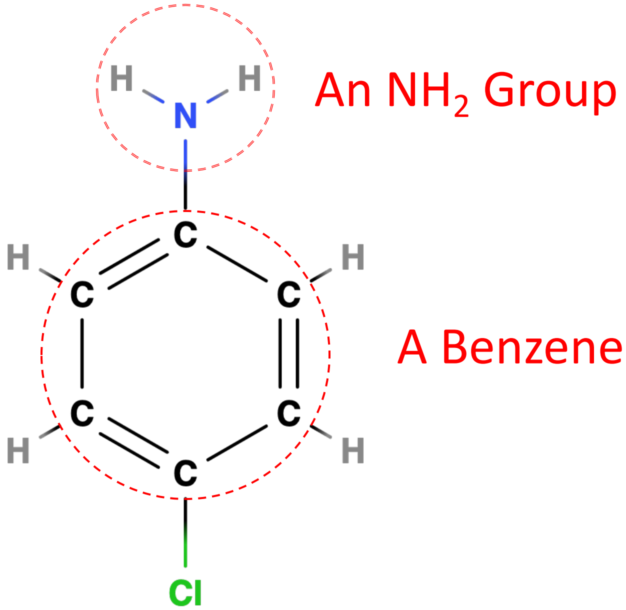
Molecular Representation Learning via Heterogeneous Motif Graph Neural Networks

Zhaoning Yu and Hongyang Gao

MOLECULAR GRAPH VS OTHER GRAPH



https://brilliant.org/wiki/social-networks/

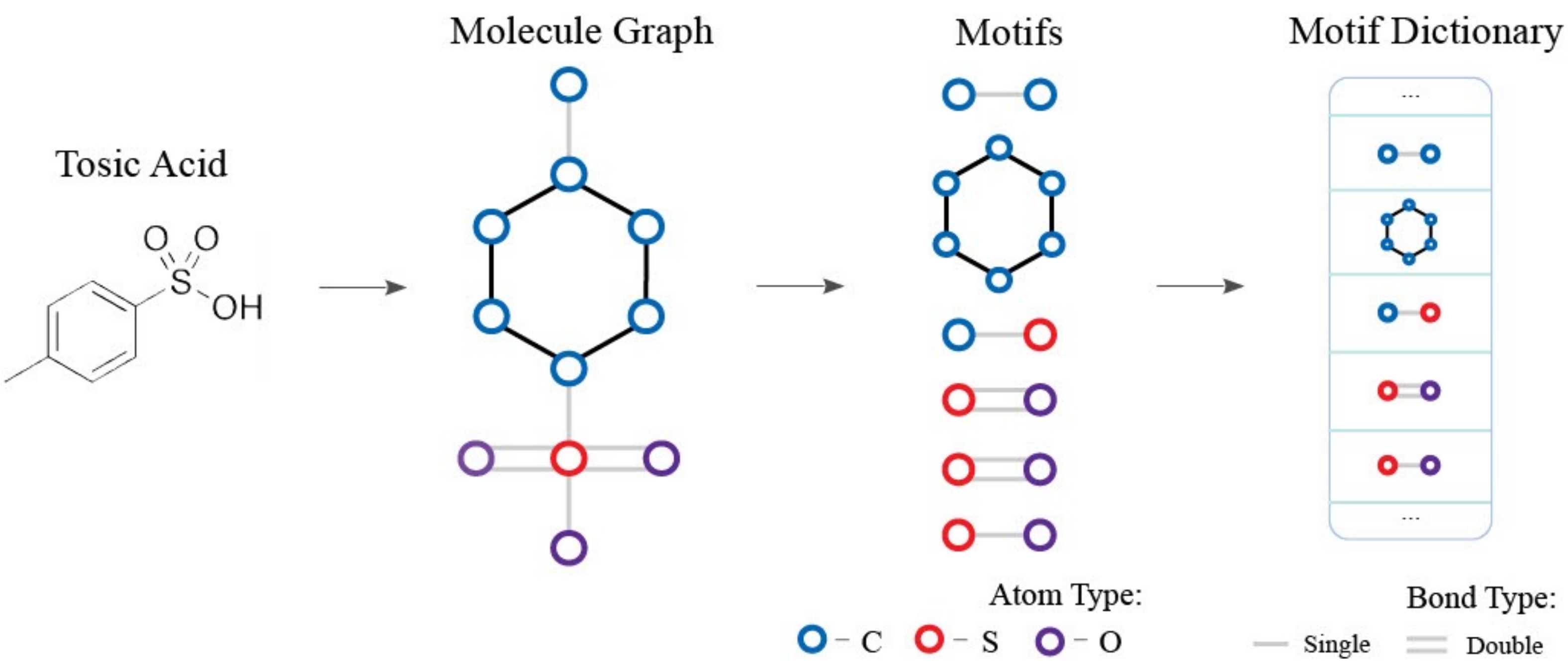


https://molview.org/

1. Common sub-graphs (motifs) in molecular graphs have special meanings
2. Most existing GNNs fail to consider motif patterns sharing among molecular graphs

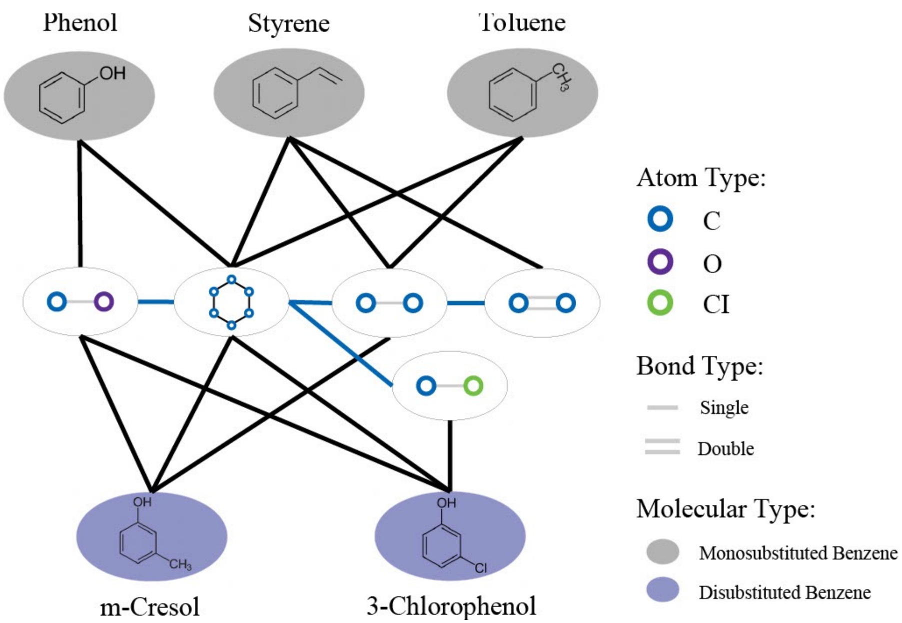
Motifs deserve more attention when designing GNNs for molecular representation learning.

MOTIF VOCABULARY



We build a motif vocabulary/dictionary by searching all molecular graphs and extract important subgraphs.

HETEROGENEOUS MOTIF GRAPH



Edge weight:

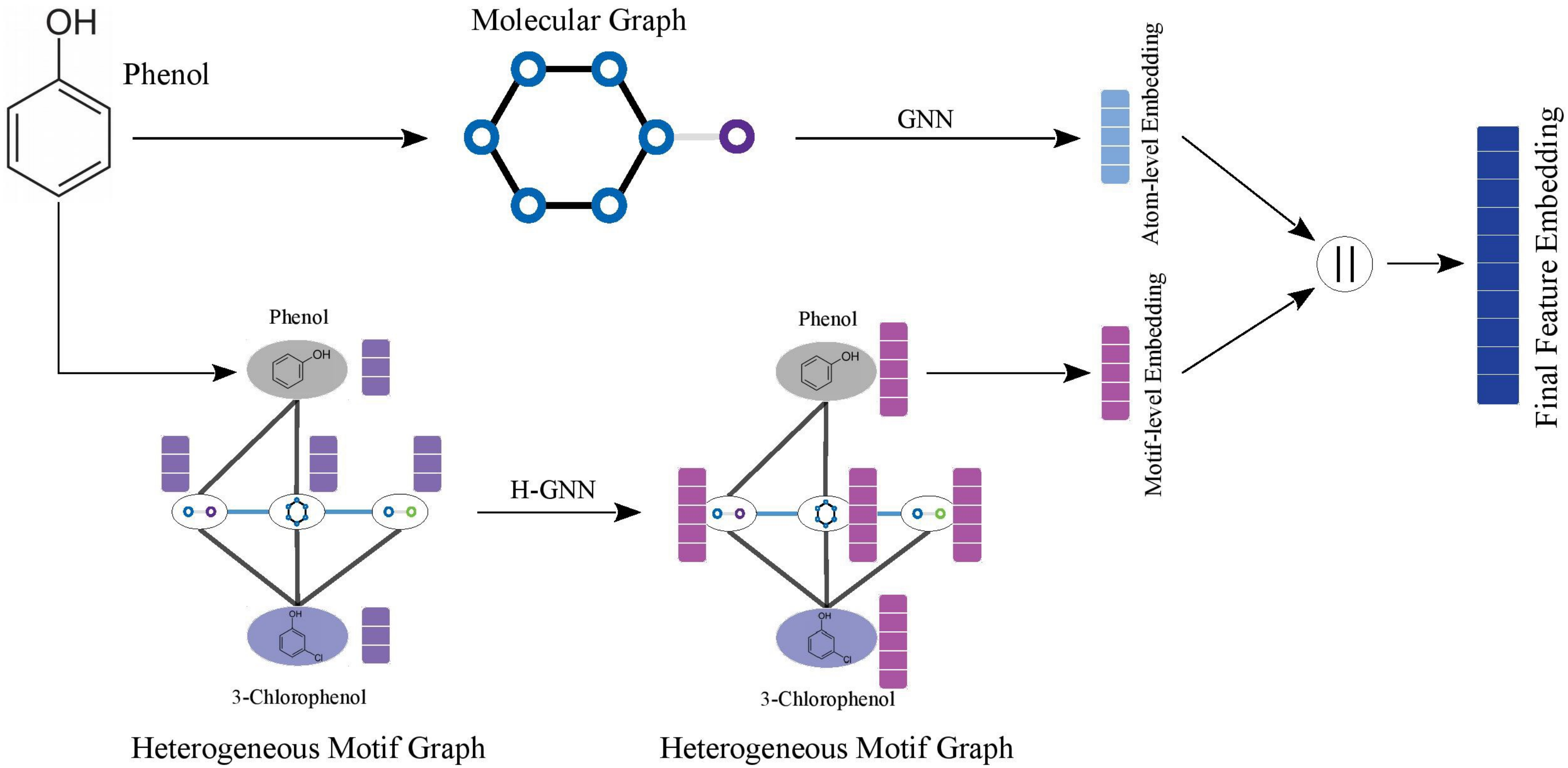
$$A_{ij} = \begin{cases} \mathrm{PMI}_{ij}, & \text{if } i, j \text{ are motifs} \\ \text{TF-IDF}_{ij}, & \text{if } i \text{ or } j \text{ is a motif} \\ 0, & \text{Otherwise} \end{cases}$$

$$\text{TF-IDF}_{ij} = C(i)_j \left(\log \frac{1+M}{1+N(i)} + 1 \right)$$

$$\mathrm{PMI}_{ij} = \log \frac{p(i,j)}{p(i)p(j)}$$

Based on the motif vocabulary, we build a heterogeneous graph that contains motif nodes and molecular nodes.

HETEROGENEOUS MOTIF GRAPH NEURAL NETWORKS



We construct a heterogeneous motif graph to learn both atom-level and motif-level representation simultaneously.

EXPERIMENTAL RESULTS

METHODS	PTC	MUTAG	NCI1	PROTEINS	MUTAGENICITY
PatchySAN	60.0 ± 4.8	92.6 ± 4.2	78.6 ± 1.9	75.9 ± 2.8	-
GCN	64.2 ± 4.3	85.6 ± 5.8	80.2 ± 2.0	76.0 ± 3.2	79.8 ± 1.6
GraphSAGE	63.9 ± 7.7	85.1 ± 7.6	77.7 ± 1.5	75.9 ± 3.2	78.8 ± 1.2
DGCNN	58.6 ± 2.5	85.8 ± 1.7	74.4 ± 0.5	75.5 ± 0.9	-
GIN	64.6 ± 7.0	89.4 ± 5.6	82.7 ± 1.7	76.2 ± 2.8	-
PPGN	66.2 ± 6.5	90.6 ± 8.7	83.2 ± 1.1	77.2 ± 4.7	-
CapsGNN	-	86.7 ± 6.9	78.4 ± 1.6	76.3 ± 3.6	-
WEGL	64.6 ± 7.4	88.3 ± 5.1	76.8 ± 1.7	76.1 ± 3.3	-
GraphNorm	64.9 ± 7.5	91.6 ± 6.5	81.4 ± 2.4	77.4 ± 4.9	-
GSN	68.2 ± 7.2	90.6 ± 7.5	83.5 ± 2.3	76.6 ± 5.0	-
OURS	$\mathbf{78.8 \pm 6.5}$	$\mathbf{96.3 \pm 2.6}$	$\mathbf{83.6 \pm 1.5}$	$\mathbf{79.9 \pm 3.1}$	$\mathbf{83.0 \pm 1.1}$

Thank you for listening!

Paper: arXiv:2202.00529