Expressivity and Generalization: Fragment-Biases for Molecular GNNs



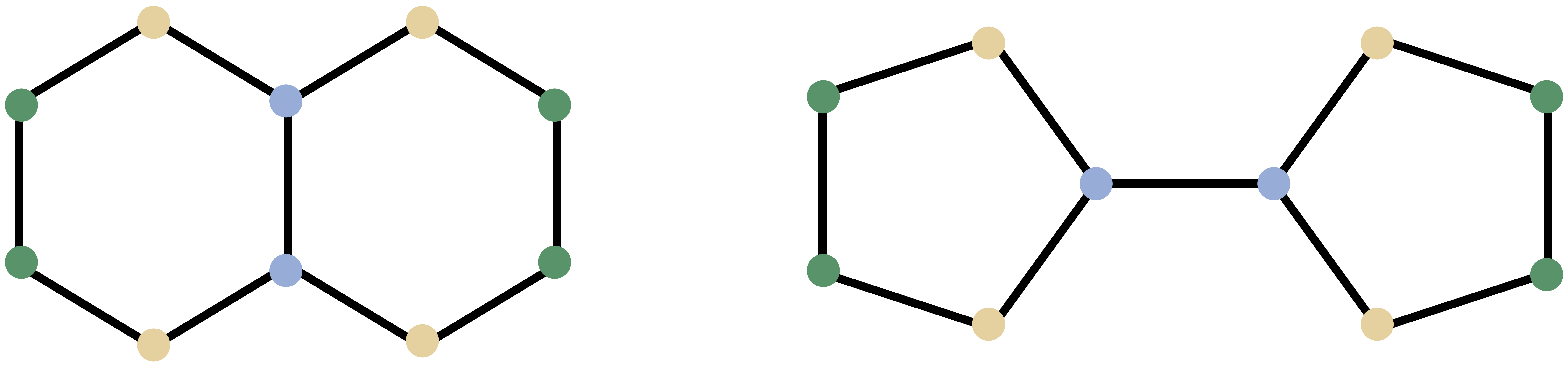


Tom Wollschläger *, **Niklas Kemper** *, Leon Hetzel, Johanna Sommer, Stephan Günnemann

Data Analytics and Machine Learning Group
Technical University of Munich

*equal contribution

Traditional GNNs limited expressivity



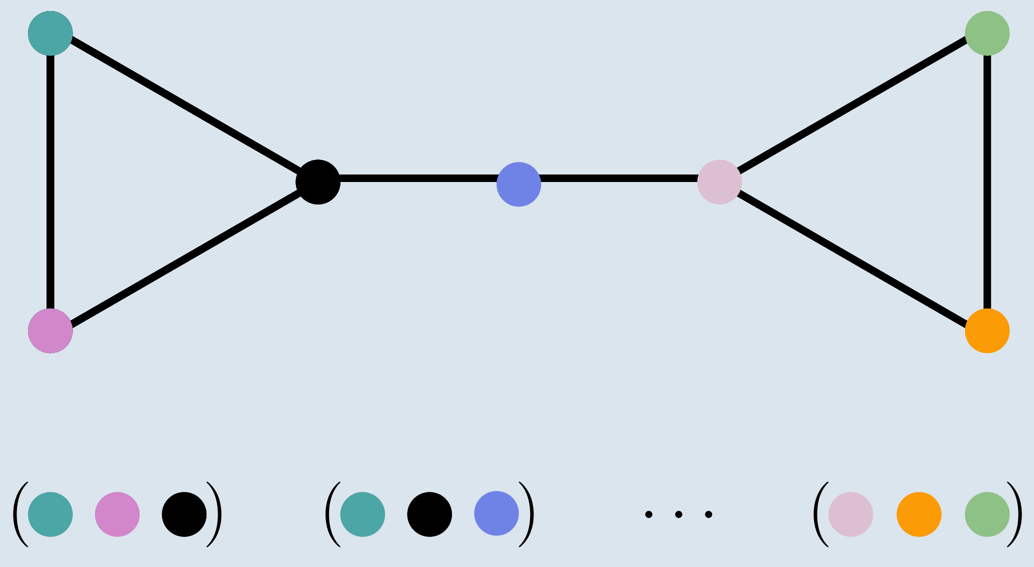
Limited expressivity:

- Distinguishing non-isomorphic graphs bounded by the Weisfeiler & Lehman test
- Blind to substructures [1]

[1] Chen et al. Can Graph Neural Networks Count Substructures?, *NeurIPS 2020*

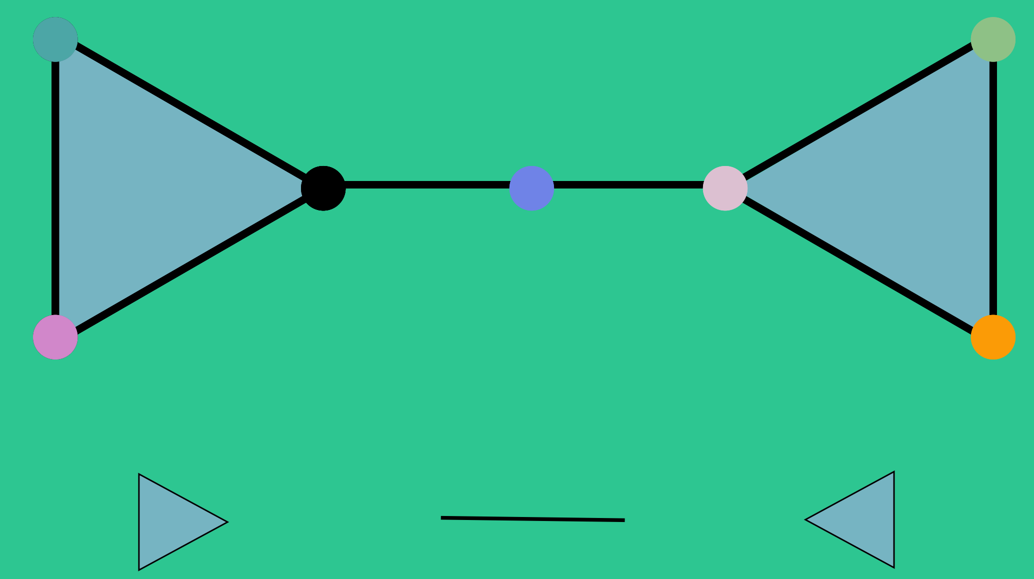
Approaches to increase expressivity

Higher-order GNNs



- ☹ Complexity
- ☺ Theory [2]
- 😐 Expressivity [2]
- ☹ Generalization [3]

Fragment-biased GNNs



- ☺ Complexity
- ☹ Theory
- ? Expressivity
- ? Generalization

[2] Zhang et. al. A Quantitative Framework for GNN Expressiveness. *ICLR 2024*

[3] Campi et. al. Expressivity of Graph Neural Networks Through the Lens of Adversarial Robustness. *CoRR 2023*

Design choices for fragment-biased GNNs

1. How should a graph be fragmented?
2. How to use fragment information in a model?

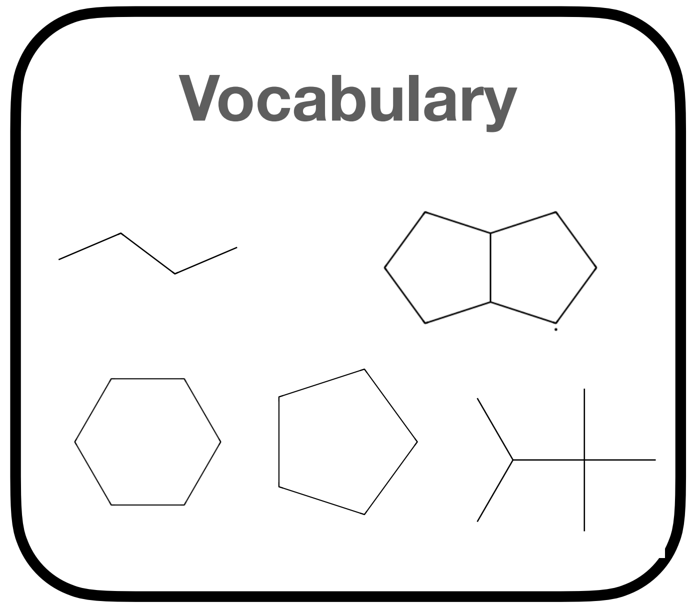
Design choices for fragment-biased GNNs

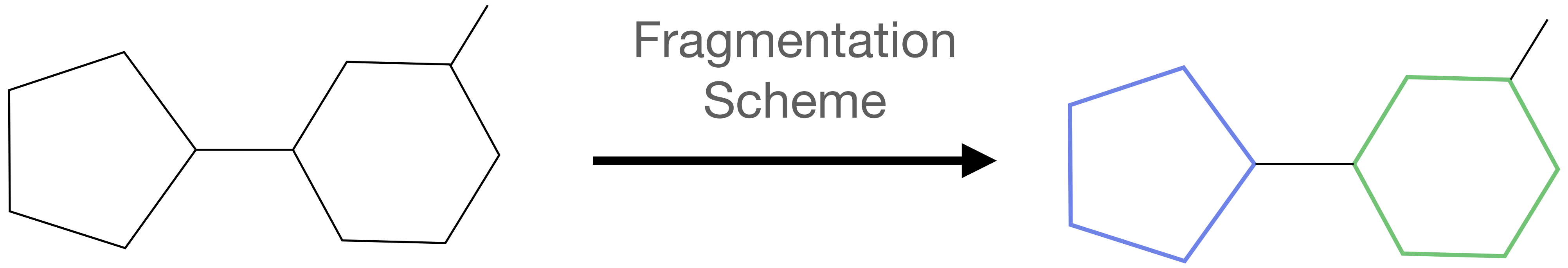
1. How should a graph be fragmented?

2. How to use fragment information in a model?

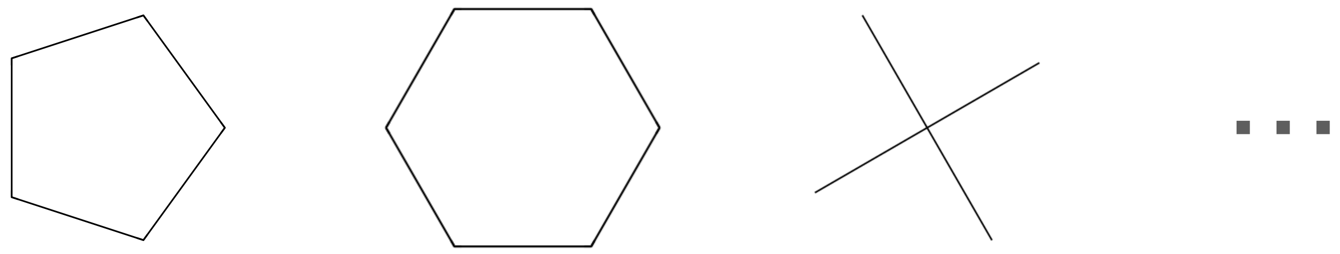
How should a graph be fragmented?

Two conflicting goals





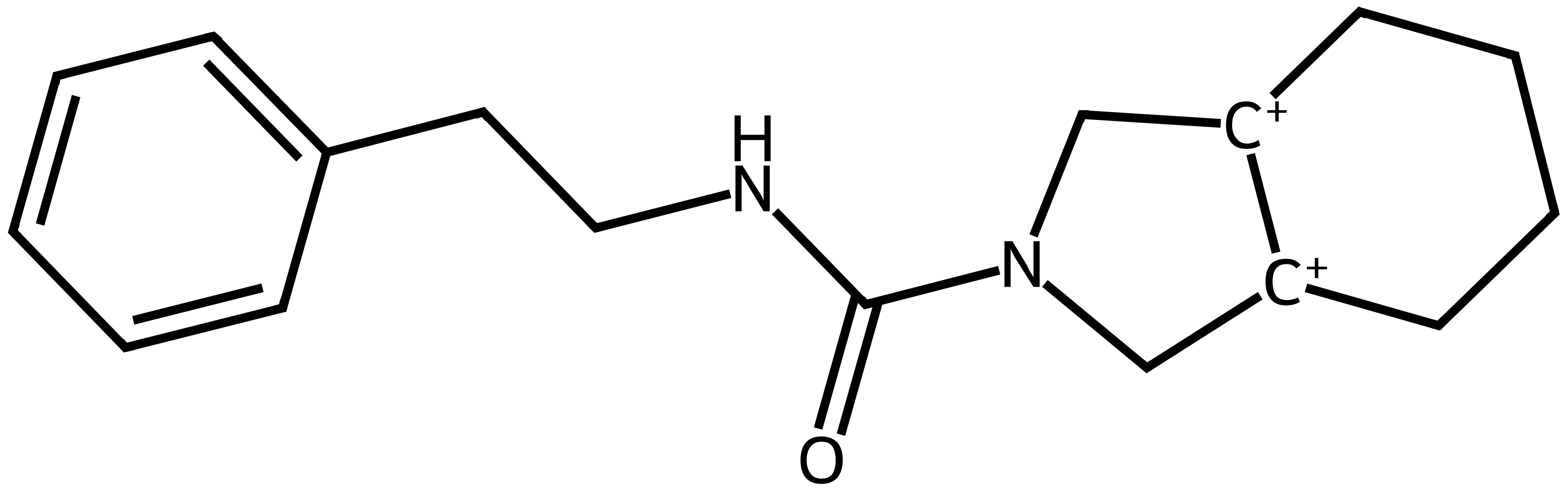
1. Fragmentation should include all *important* substructures.



2. Fragmentation should facilitate generalization across diverse graphs.

Our RingsPaths Fragmentation

Fragment the complete molecule using only small building blocks



1. Minimal Cycle Basis
2. Maximally long uninterrupted paths

➡ Fragment complete graph using only two types of substructures

Design choices for fragment-biased GNNs

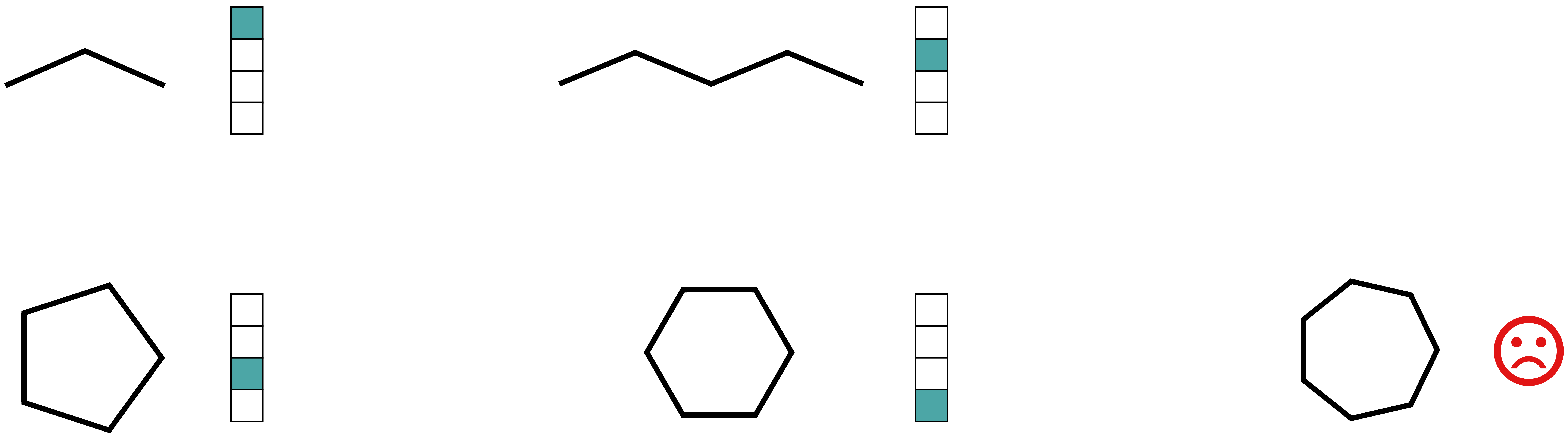
1. How should a graph be fragmented?

2. How to use fragment information in a model?

    A. How to encode fragment information?
    B. How to incorporate fragment information into the model?

One Hot Encoding

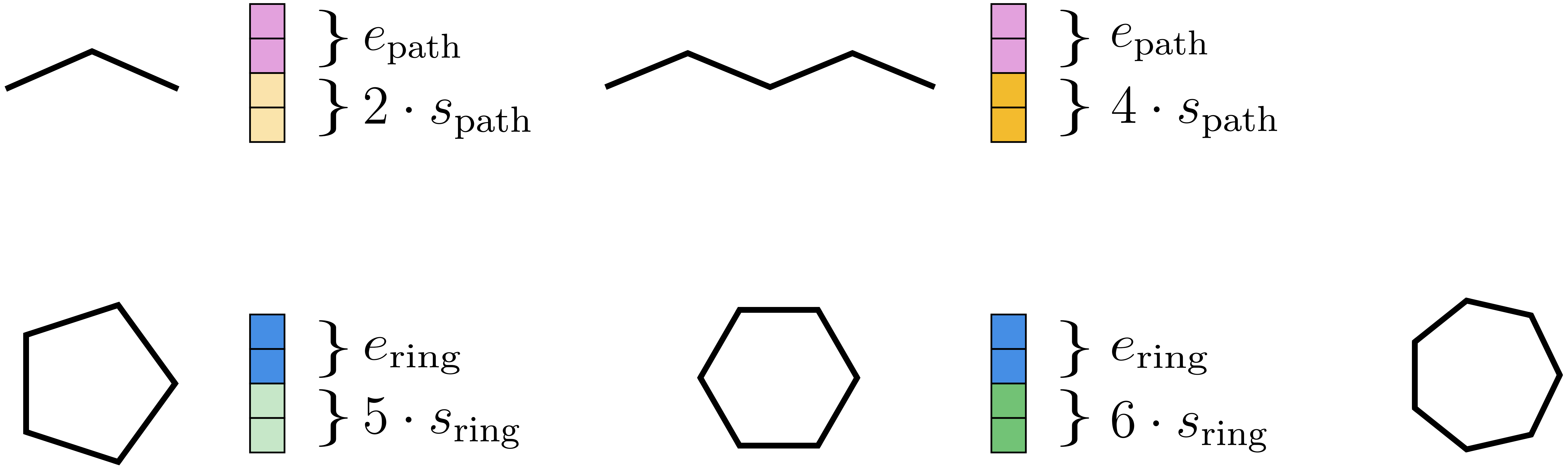
Similar Fragments — Different Encodings



- Supports only a fixed number of fragments

Ordinal Encoding

Similar Fragments — Similar Encodings



- Supports infinitely many fragment types
- Transfer knowledge between similar fragments

Design choices for fragment-biased GNNs

1. How should a graph be fragmented?

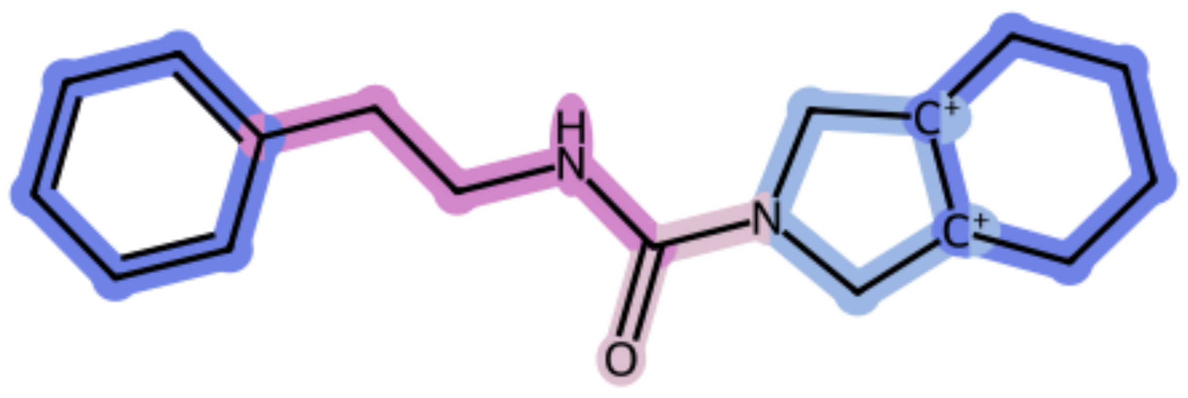
2. How to use fragment information in a model?

    A. How to encode fragment information?
    B. How to incorporate fragment information into the model?

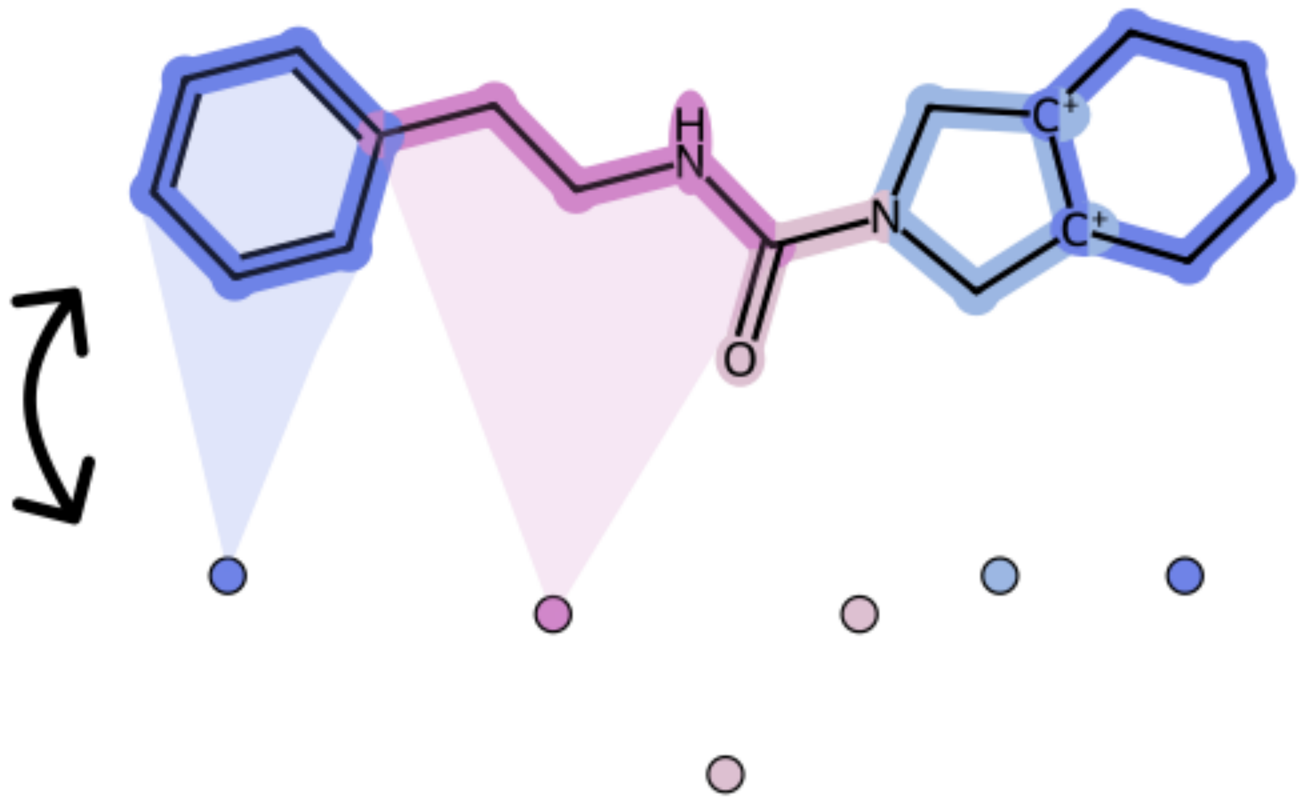
Approaches to incorporate fragment information

Difficult to compare expressivity directly

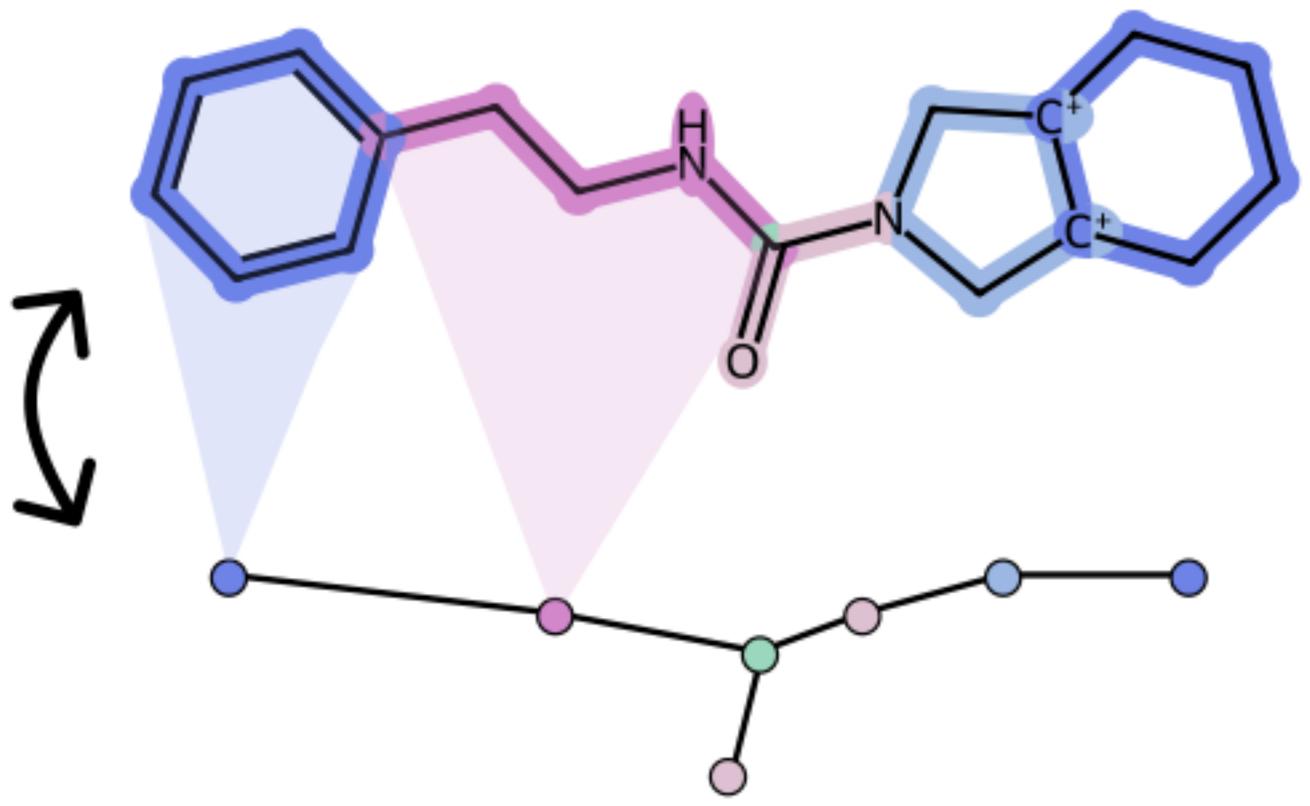
Node Features



Fragment Representation



Higher-level Graph

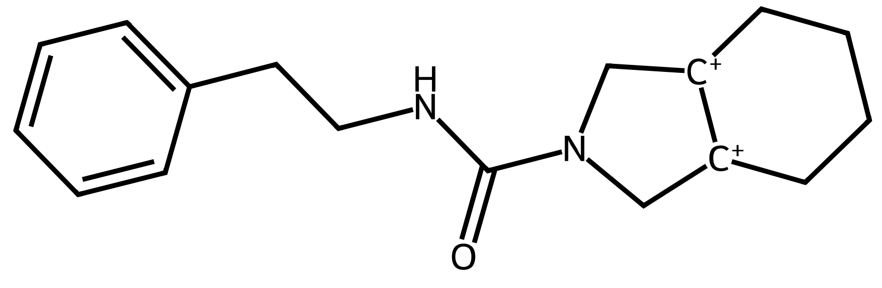


- Does higher-level abstraction come with an increase in expressivity?

New Measures of expressivity

Expressivity increases with higher-level abstraction

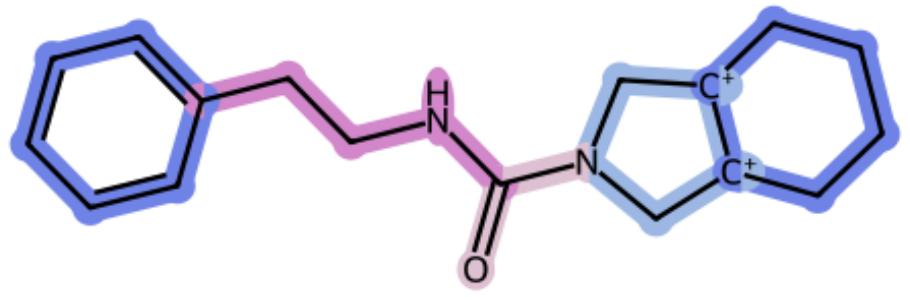
No Fragmentation



WL test

< Theorem 4.6

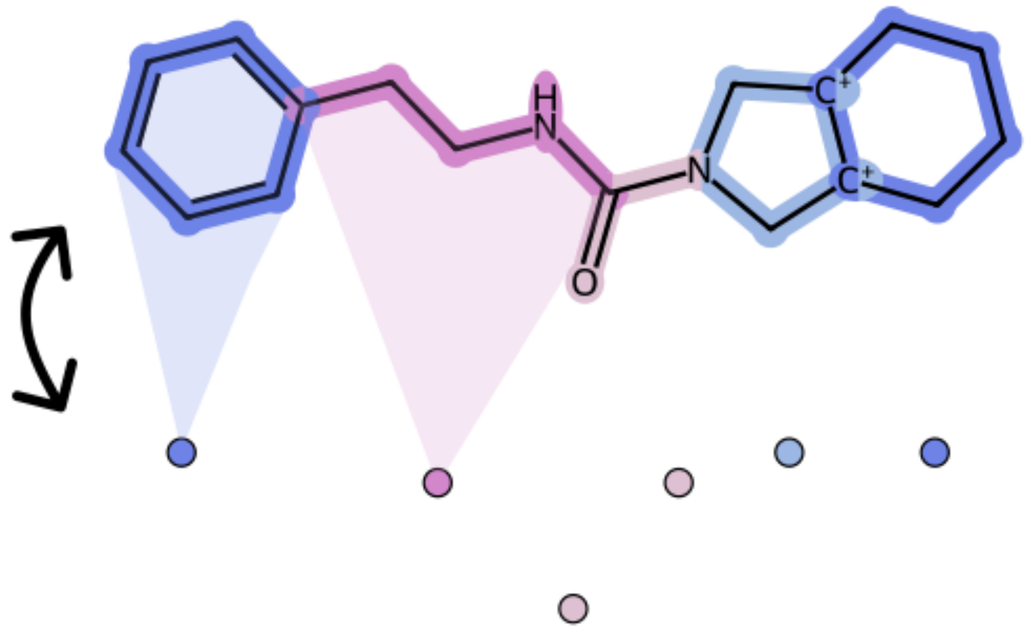
Node Features



NF-WL test

< Theorem 4.7

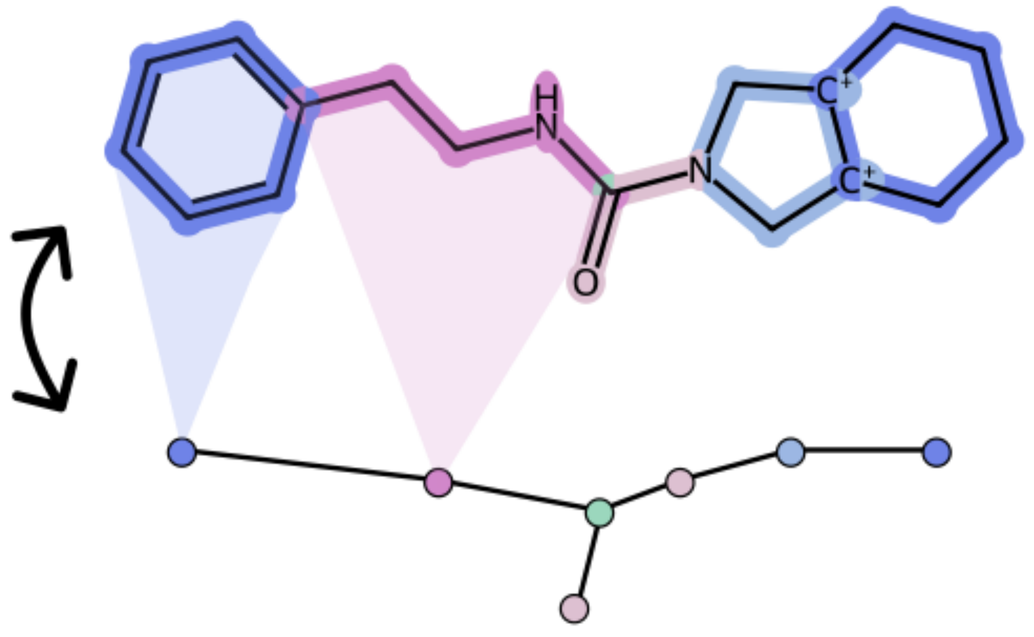
Fragment Representation



FR-WL test

< Theorem 4.8

Higher-level Graph



HLG-WL test

Expressivity strictly increases

➡ Enables comparison of existing fragment-biased GNNs

Design choices for fragment-biased GNNs

1. How should a graph be fragmented?

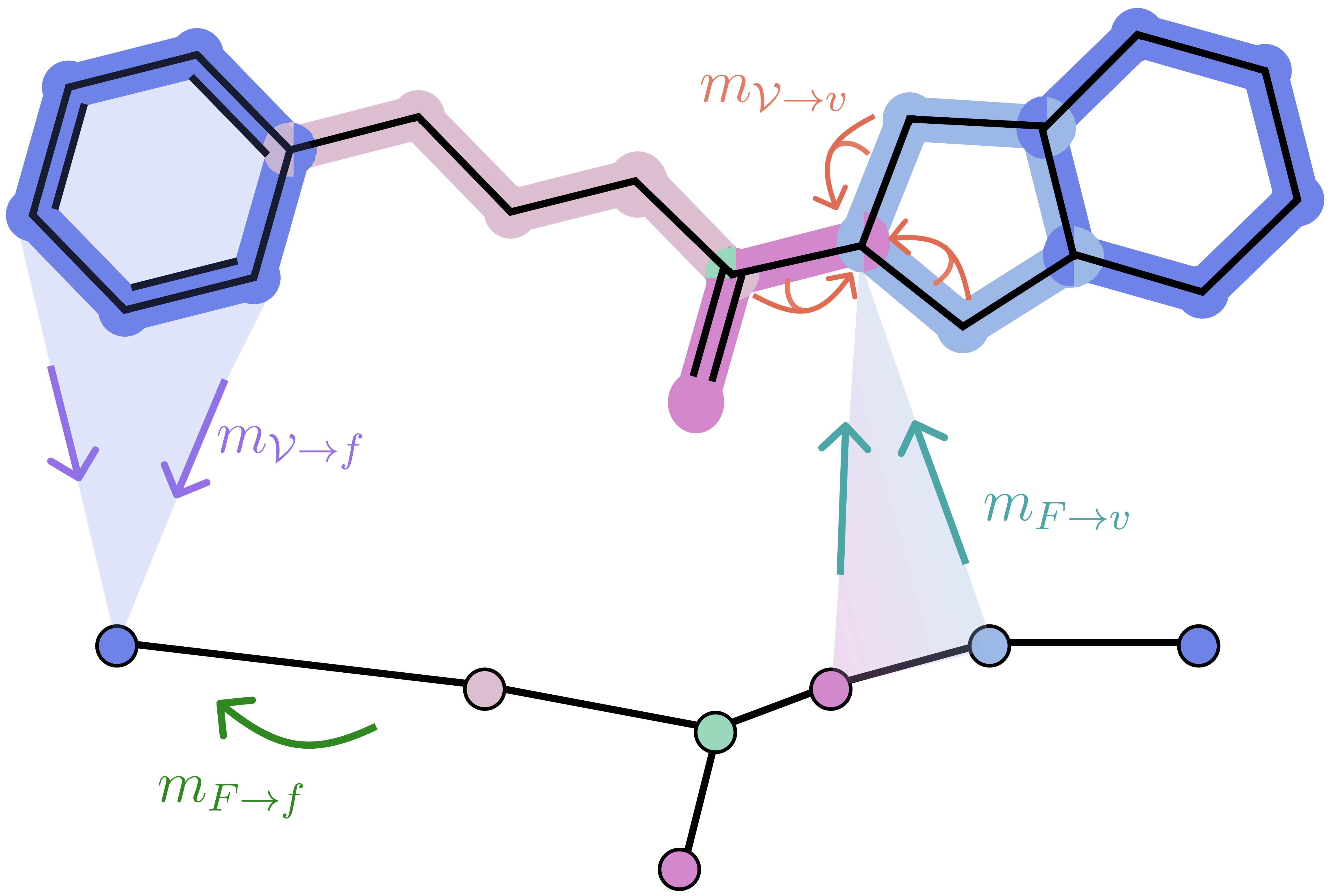
2. How to use fragment information in a model?

    A. How to encode fragment information?
    B. How to incorporate fragment information into the model?

FragNet

Overview of our model

- RingsPaths Fragmentation
- Ordinal Encoding
- Higher-level Graph

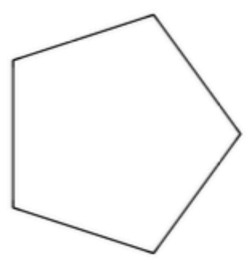
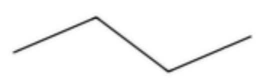


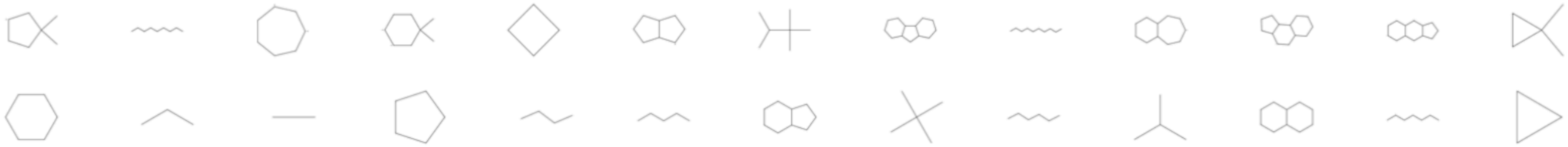
Empirical Evaluation

1. Expressivity
2. Benchmarks
3. Generalization

Expressivity

FragNet can count chemically important substructures

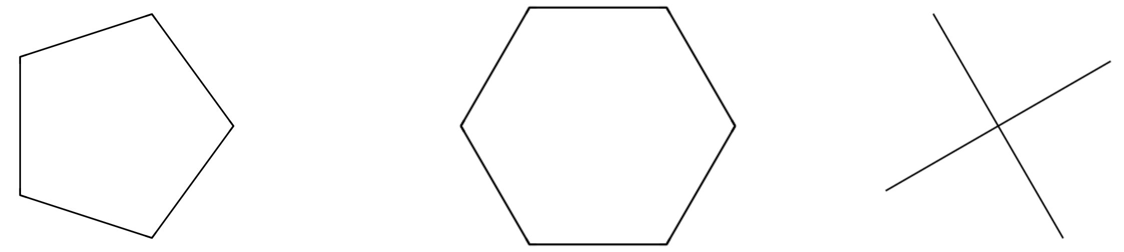
Fragment		
Counts	5629	3904
Accuracy	0.986	0.99



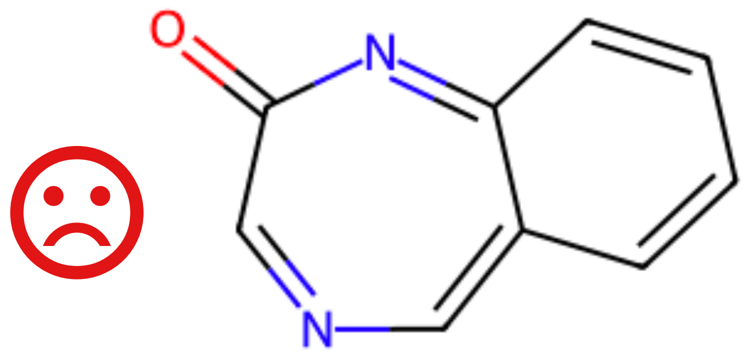
How should a graph be fragmented?

Two ~~conflicting~~ goals

1. Fragmentation should ~~include~~ **enable to learn** all *important* substructures.

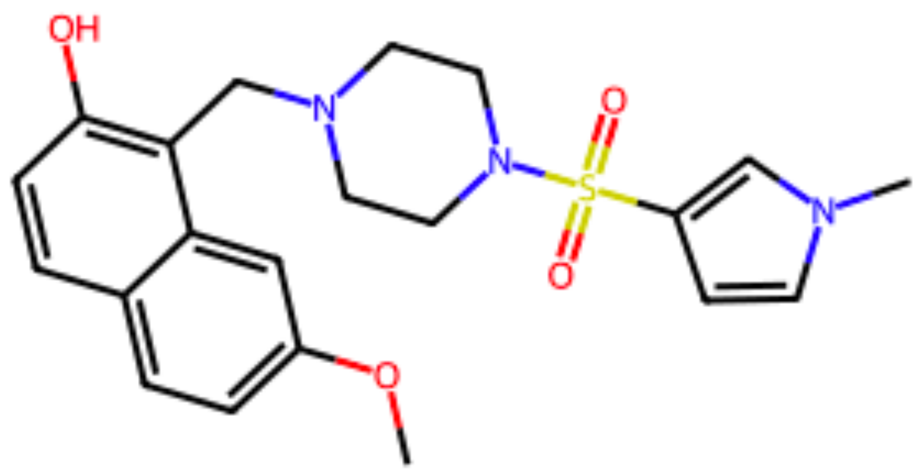


2. Fragmentation should facilitate generalization across diverse graph structures.

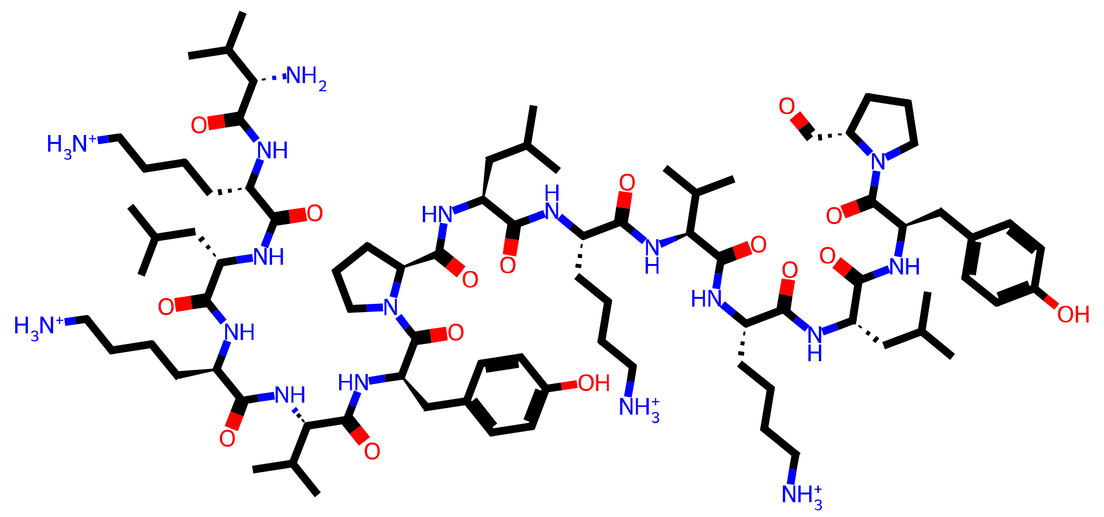


Empirical Performance

FragNet is SOTA among (fragment-biased) GNNs



ZINC



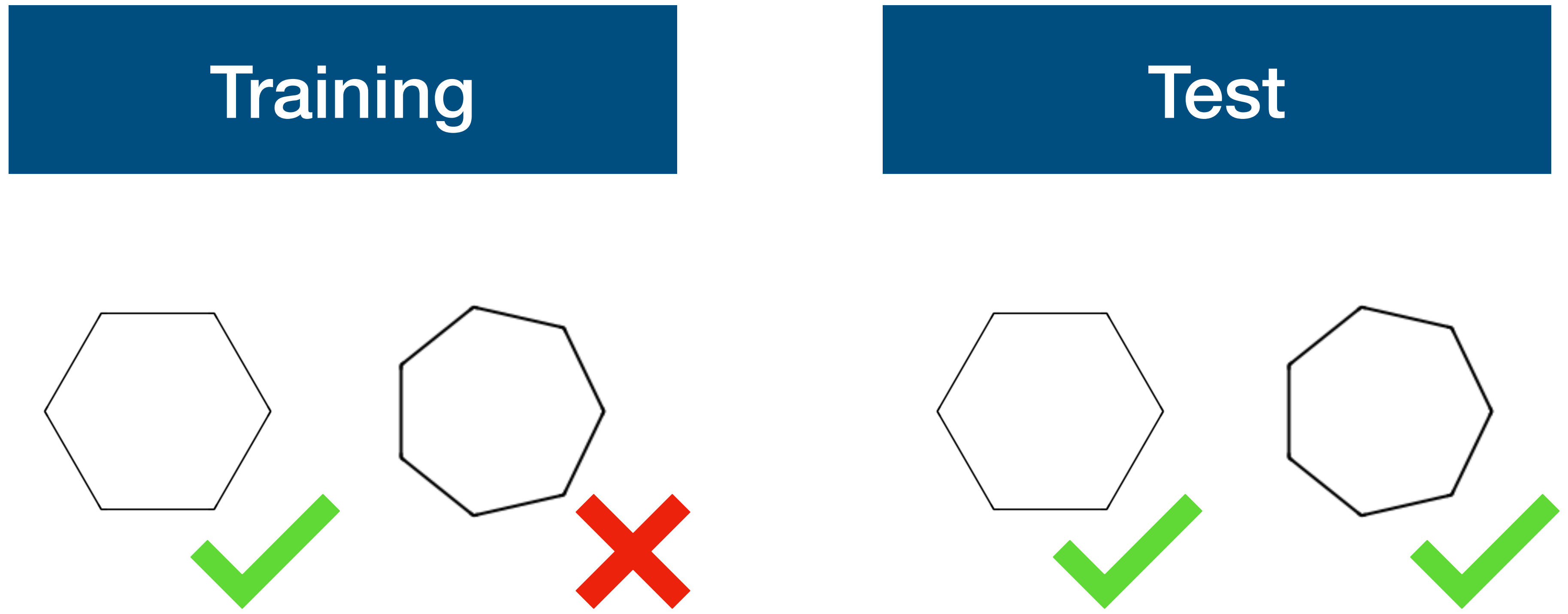
Peptides

- FragNet best (fragment-biased) GNN
- Comparable performance to state-of-the-art transformer GRIT [4] on Peptides-struct & ZINC-full

[4] Ma et al. Graph Inductive Biases in Transformers without Message Passing *ICML 2023*

Generalization: Unseen Fragments

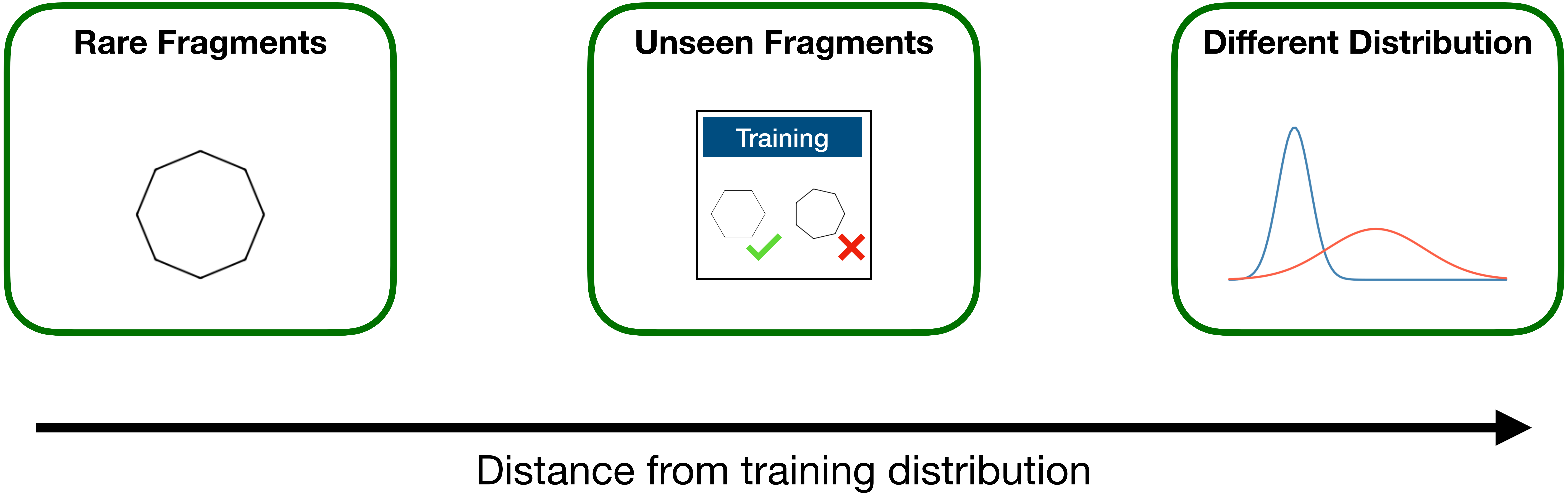
Ordinal Encoding helps to generalize to fragments not in the training set



Model	**ZINC 10k**	
	training (MAE ↓)	**test** (MAE ↓)
GRIT	0.02	0.61
Ours	0.08	**0.34**

Generalization

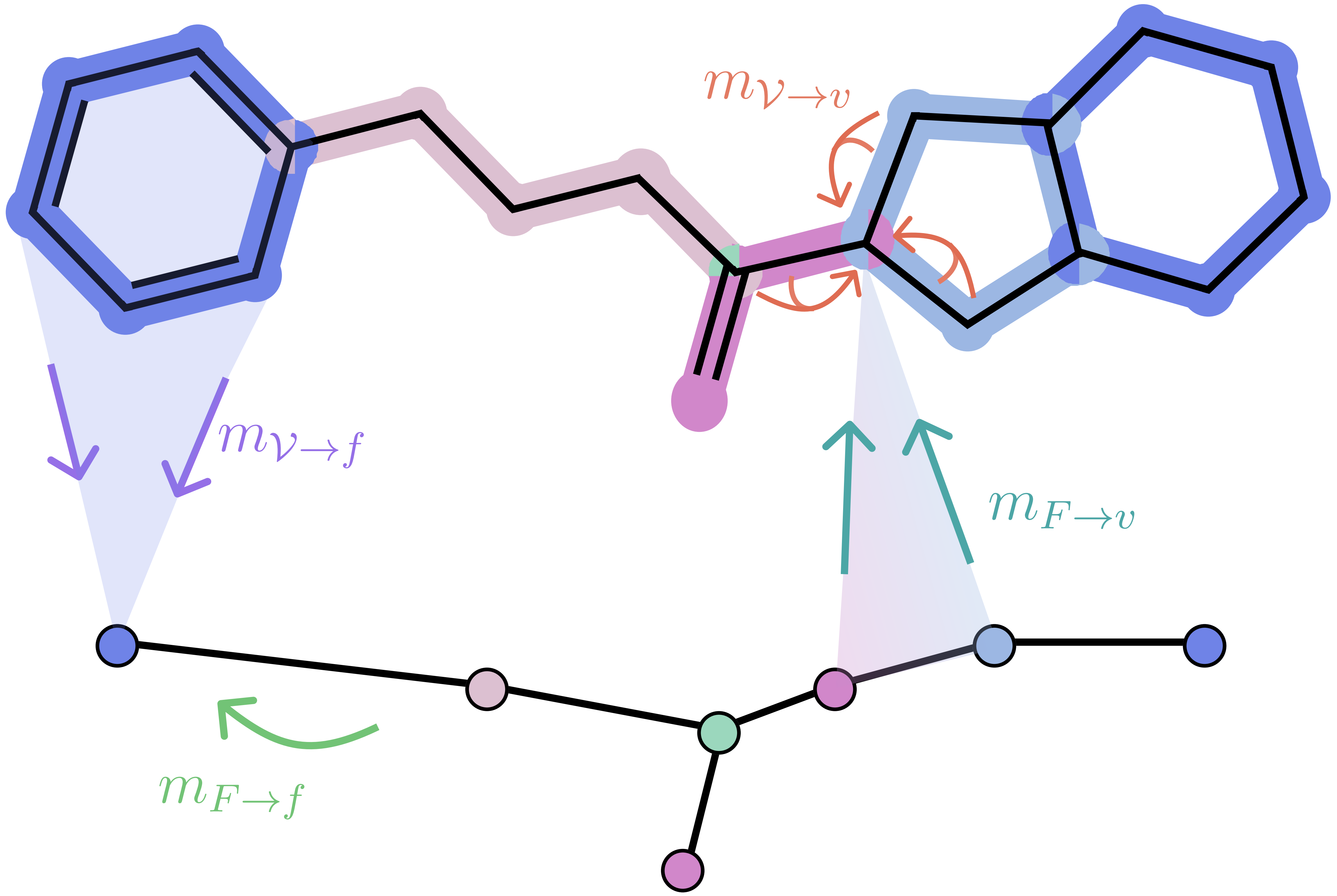
RingsPaths + Ordinal Encoding + Higher-level graph = improved generalization



FragNet

A robust and highly expressive fragment-biased GNN

Paper & Code



- Complexity
- Theory
- Expressivity
- Generalization
- SOTA performance for GNNs

References

[1] Chen et al. Can Graph Neural Networks Count Substructures?, *NeurIPS 2020*

[2] Zhang, B., Gai, J., Du, Y., Ye, Q., He, D., and Wang, L. Beyond Weisfeiler-Lehman: A Quantitative Framework for GNN Expressiveness. *ICLR 2024*

[3] Campi, F., Gosch, L., Wollschläger, T., Scholten, Y., and Günnemann, S. Expressivity of Graph Neural Networks Through the Lens of Adversarial Robustness. *CoRR 2023*

[4] Ma et al. Graph Inductive Biases in Transformers without Message Passing *ICML 2023*

[5] Degen et al. On the art of compiling and using 'drug-like' chemical fragment spaces, *ChemMedChem* 2008

[6] Bodnar et al. Weisfeiler and Lehman Go Cellular: CW Networks, *NeurIPS* 2021

[7] Bouritsas et al. Improving Graph Neural Network Expressivity via Subgraph Isomorphism Counting, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2021

[8] Zhu, J., Wu, K., Wang, B., Xia, Y., Xie, S., Meng, Q., Wu, L., Qin, T., Zhou, W., Li, H., and Liu, T.-Y. $\mathcal{O}$-GNN: incorporating ring priors into molecular modeling. September 2022.

[9] Fey et al., Hierarchical Inter-Message Passing for Learning on Molecular Graphs, *GRL+ Workhop at ICML* 2020