Lecture 9: Applications of Graph Neural Networks

Relational Learning

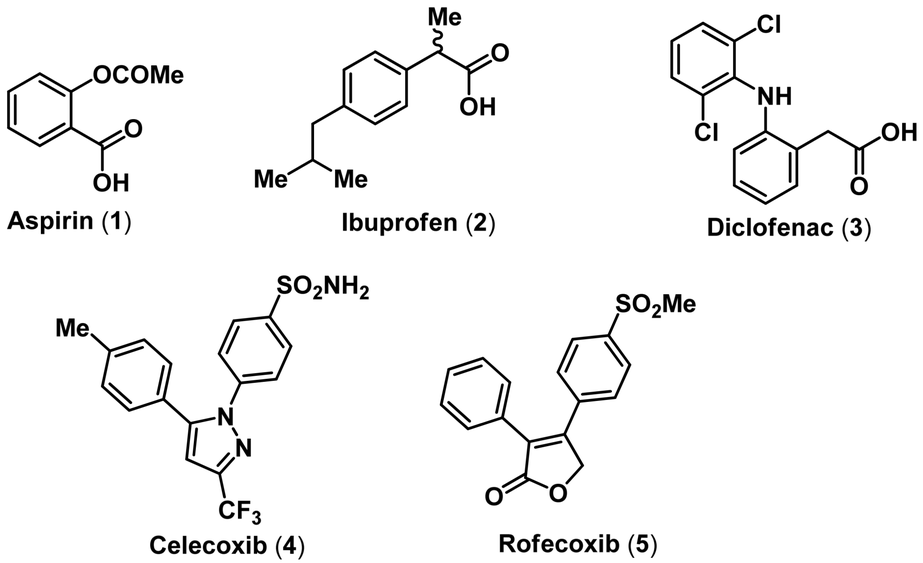
İsmail İlkan Ceylan Advanced Topics in Machine Learning, University of Oxford 07.02.2021

Overview

- Biomedical data: Molecule, interactome, complex relationships
- Drug discovery
- Protein folding
- Particle physics
- Combinatorial optimization and reasoning
- Computer vision: Scene graphs and question answering
- Recommender systems
- Traffic forecasting
- Summary of the relational learning theme

Biomedical Data

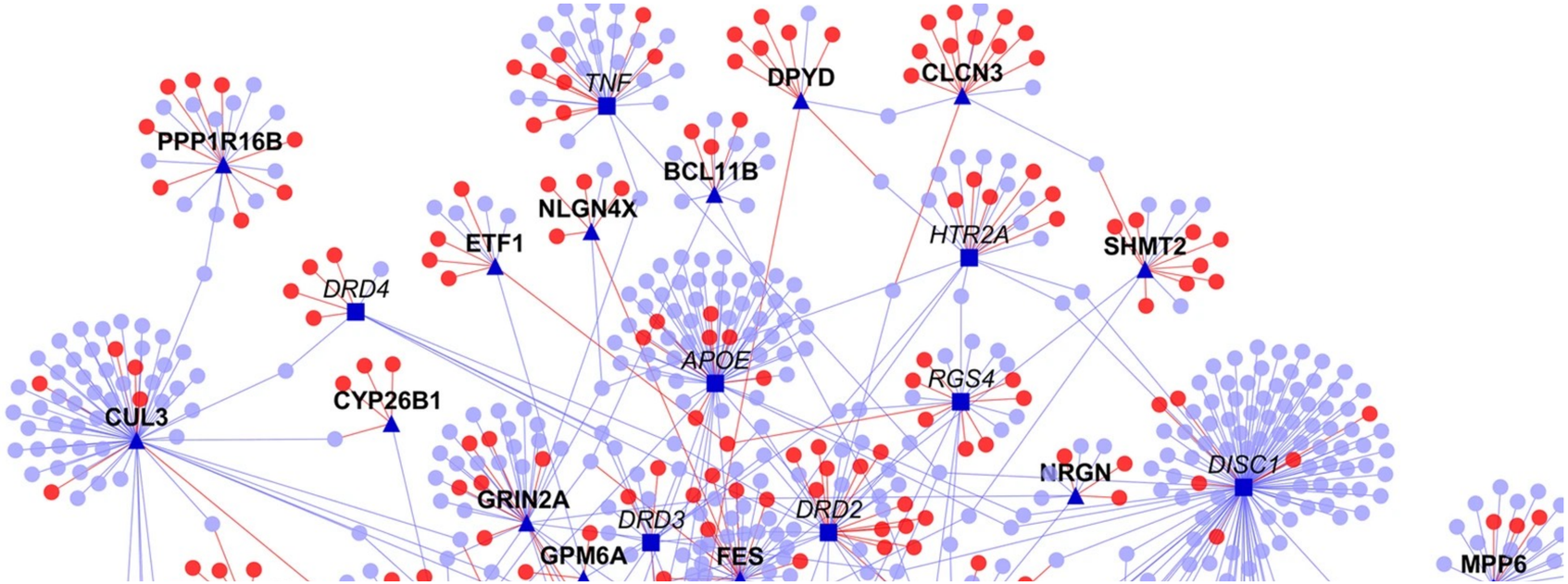
Biomedical Data: Molecular Scale



Molecules (Rao et al, 2013): Figure shows the molecule structure of NSAID drugs. “Me" is an abbreviation for "methyl" (CH3).

Molecular scale: Small molecule drugs can be represented as graphs relating their constituent atoms and chemical bonding structure. Complex molecules, such as proteins can be represented as graphs capturing spatial and structural relationships between their amino acid residues.

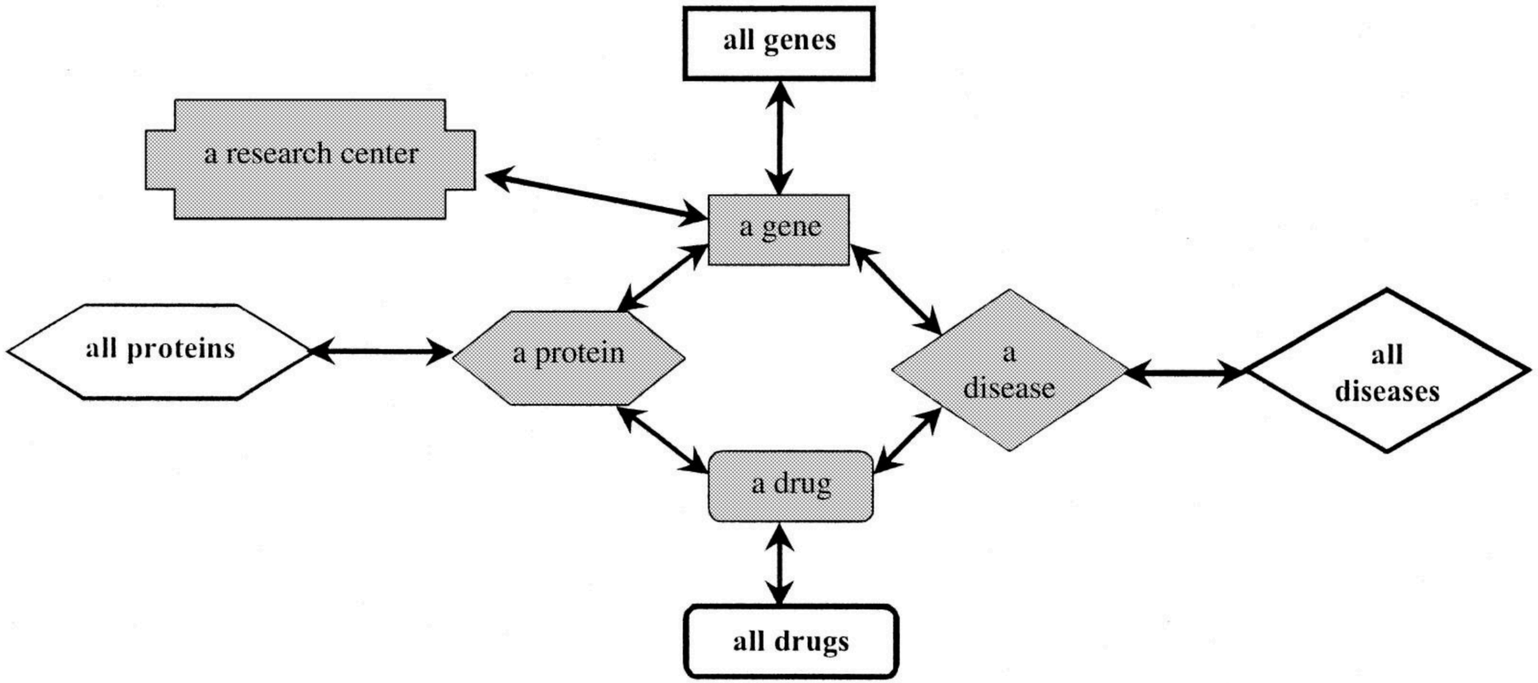
Biomedical Data: Intermediary Scale



Excerpt from Schizophrenia interactome (Ganapathiraju et al, 2016): Genes are shown as nodes and PPIs as edges connecting the nodes. Schizophrenia-associated genes are shown as dark blue nodes, novel interactors as red color nodes and known interactors as blue color nodes. Red edges are the novel interactions, whereas blue edges are known interactions.

Intermediary scale: An interactome defines a set of molecular interactions in a particular cell — They can be represented as graphs, e.g., protein–protein interaction graphs.

Biomedical Data: Abstract Scale



PharmGKB (Hewett et al., 2002): Abstract, complex relationships among the objects, including 'expresses', as in 'a gene expresses a protein': 600+ different relationships.

Abstract scale: KGs can represent the complex relationships between drugs, side effects, diagnosis, associated treatments, and test results etc.

Drug Discovery

Timeline of Drug Development

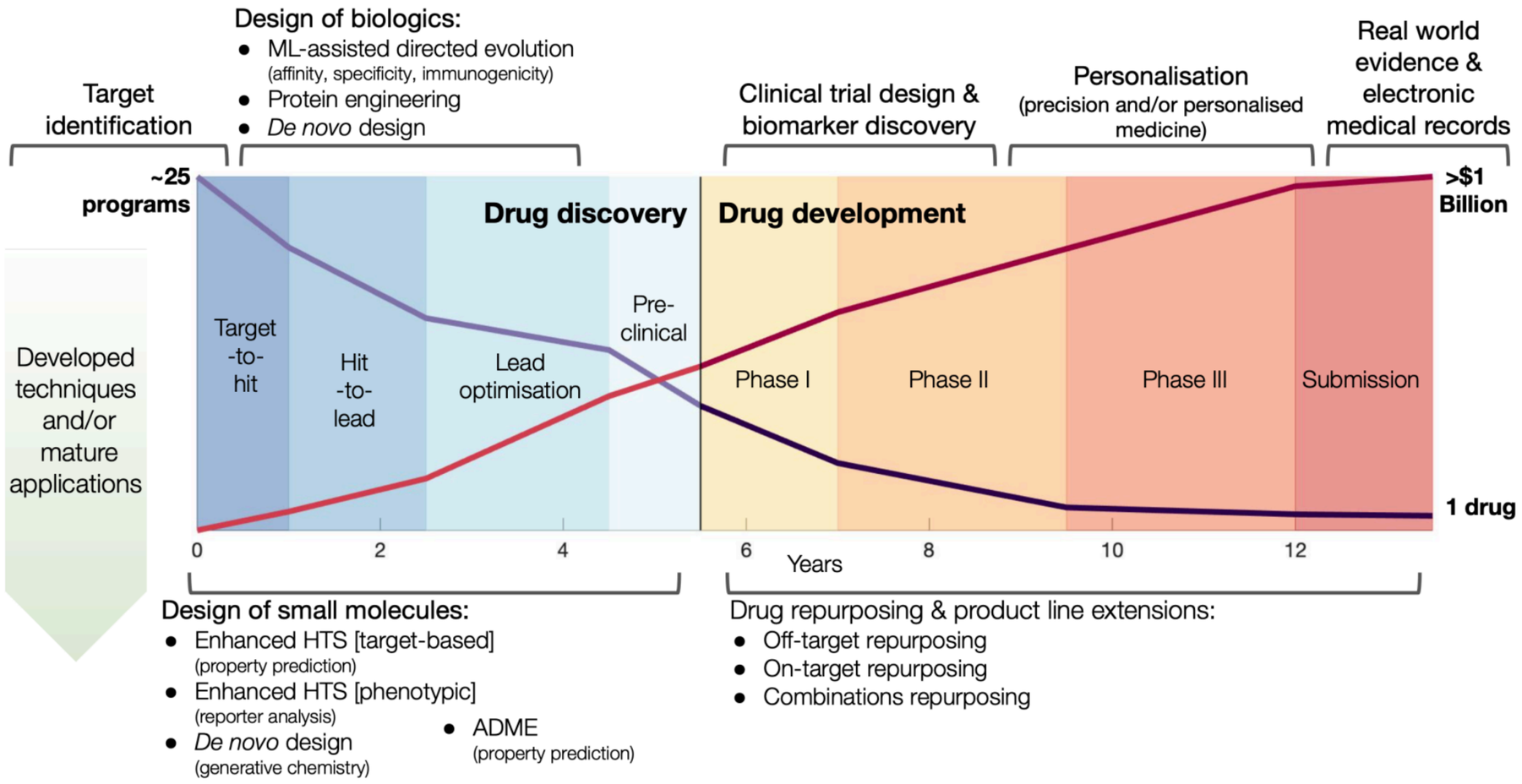


Figure (Gadoulet et al, 2021): Timeline of drug development. Drug discovery is a long and expensive process - Great interest in applying computational methods to enhance drug discovery.

Drug Development Applications

Relevant application	Reference	Method type	Task level	ML approach	Data types	Exp. val?
4.1 Target identification						
—	[47]	Geometric (§3.2)	Node-level	Unsupervised	Di, Dr, GA	
4.2 Design of small molecules therapies						
Molecular property prediction	[21]	GNN (§3.4)	Graph-level	Supervised	Dr	
	[101]	GNN (§3.4)	Graph-level	Supervised	Dr	
	[22]	GNN (§3.4)	Graph-level	Supervised	Dr	
Enhanced high throughput screens	[50]	GNN (§3.4)	Graph-level	Supervised	Dr	✓
De novo design	[102]	GNN (§3.4)	Graph-level	Unsupervised	Dr	
	[48]	Factorisation (§3.3)	Graph-level	Semi-supervised	Dr	✓
4.3 Design of new biological entities						
ML-assisted directed evolution	—	—	—	—	—	
Protein engineering	[49]	GNN (§3.4)	Subgraph-level*	Supervised	PS	
De novo design	[103]	GNN (§3.4)	Graph-level	Supervised	PS	✓
4.4 Drug repurposing						
Off-target repurposing	[104]	Factorisation (§3.3)	Node-level	Unsupervised	Dr, PI	
	[105]	GNN (§3.4)	Graph-level	Supervised	Dr, PS	
On-target repurposing	[106]	Factorisation (§3.3)	Node-level	Unsupervised	Dr, Di	
	[107]	GNN (§3.4)	Node-level	Supervised	Dr, Di	
	[108]	Geometric (§3.2)	Node-level	Unsupervised	Dr, Di, PI, GA	
Combination repurposing	[109]	GNN (§3.4)	Node-level	Supervised	Dr, PI, DC	
	[110]	GNN (§3.4)	Graph-level	Supervised	Dr, DC	✓

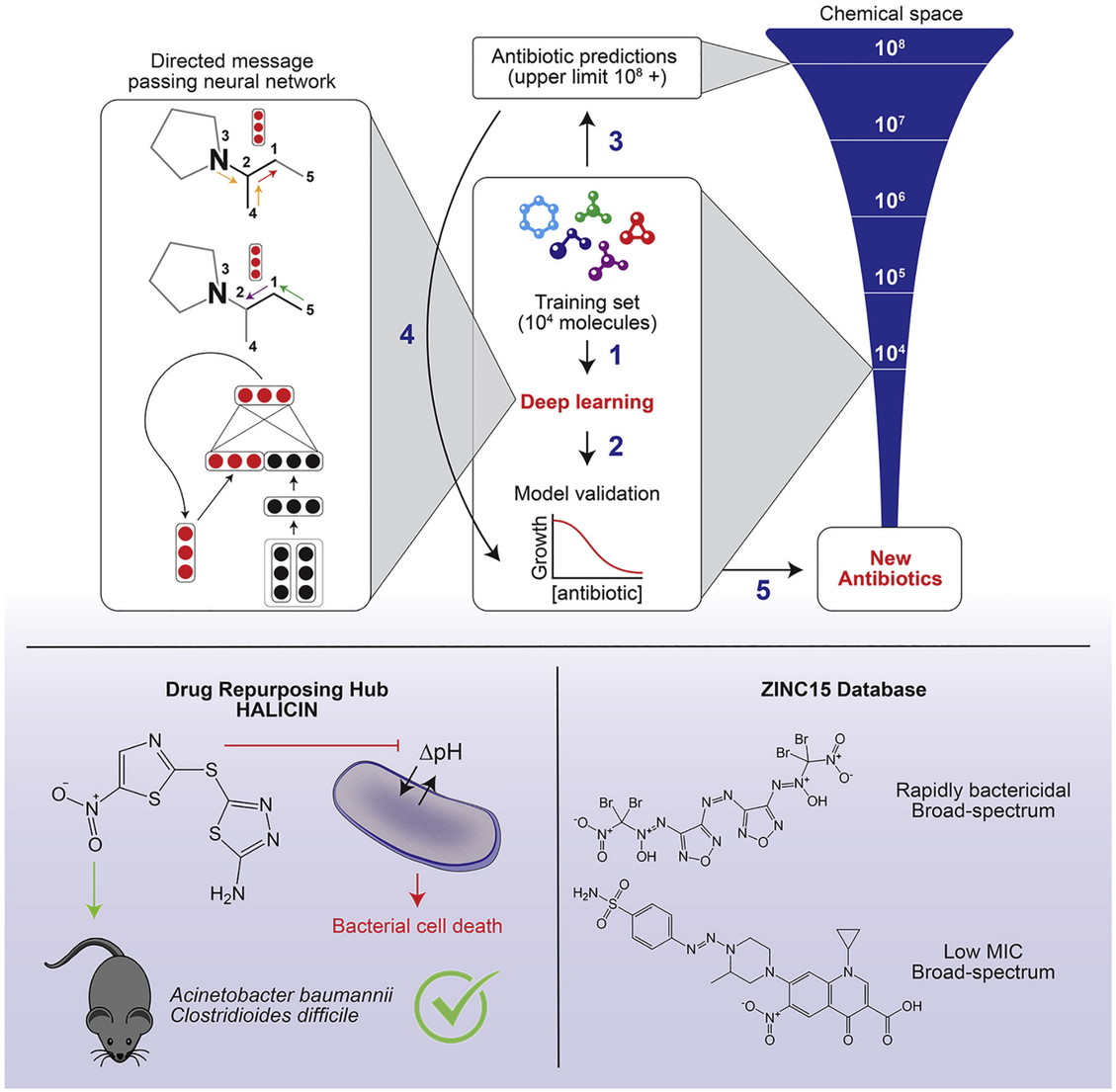
Table (Gadoulet et al, 2021): Applications of GNNs in drug discovery. The acronyms stand for Dr: Drugs, DC: Drug combinations, PS: Protein, PI: Protein interactions, GA: Gene annotations, Di:Diseases, respectively.

A Message Passing Approach to Antibiotic Discovery

Context: Very few antibiotics developed recently, and most of those newly approved antibiotics are slightly different variants of existing drugs.

A message passing approach: Stokes et al., (2020) use MPNNs for antibiotic discovery:

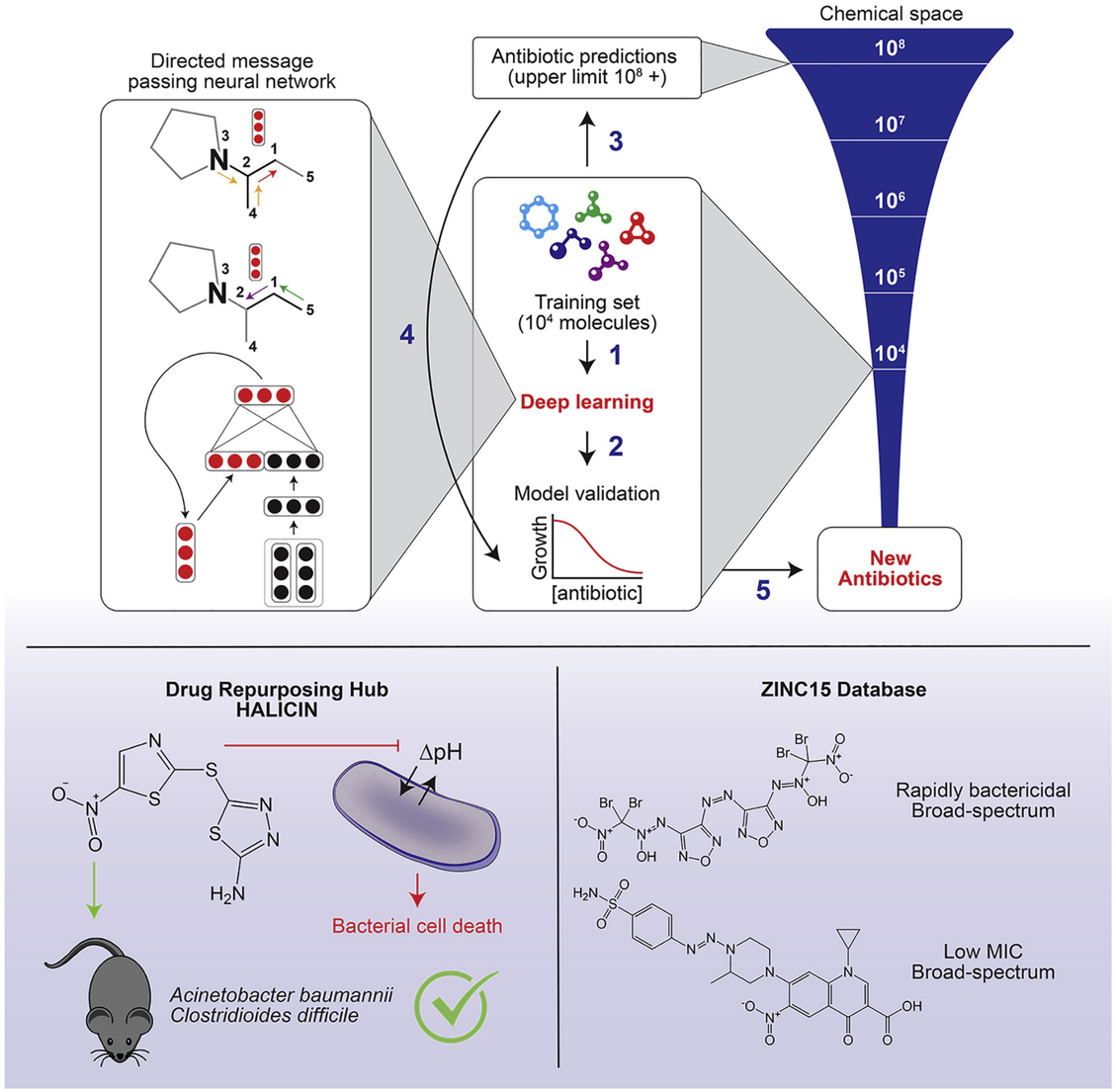
- MPNN model trained to predict a molecular property: antibacterial activity.
- Model applied to various chemical databases and candidate compounds are selected.
- Novel molecules are identified with antibacterial activity against pathogens.



A Message Passing Approach to Antibiotic Discovery

A message passing approach: Stokes et al., (2020) use MPNNs for antibiotic discovery:

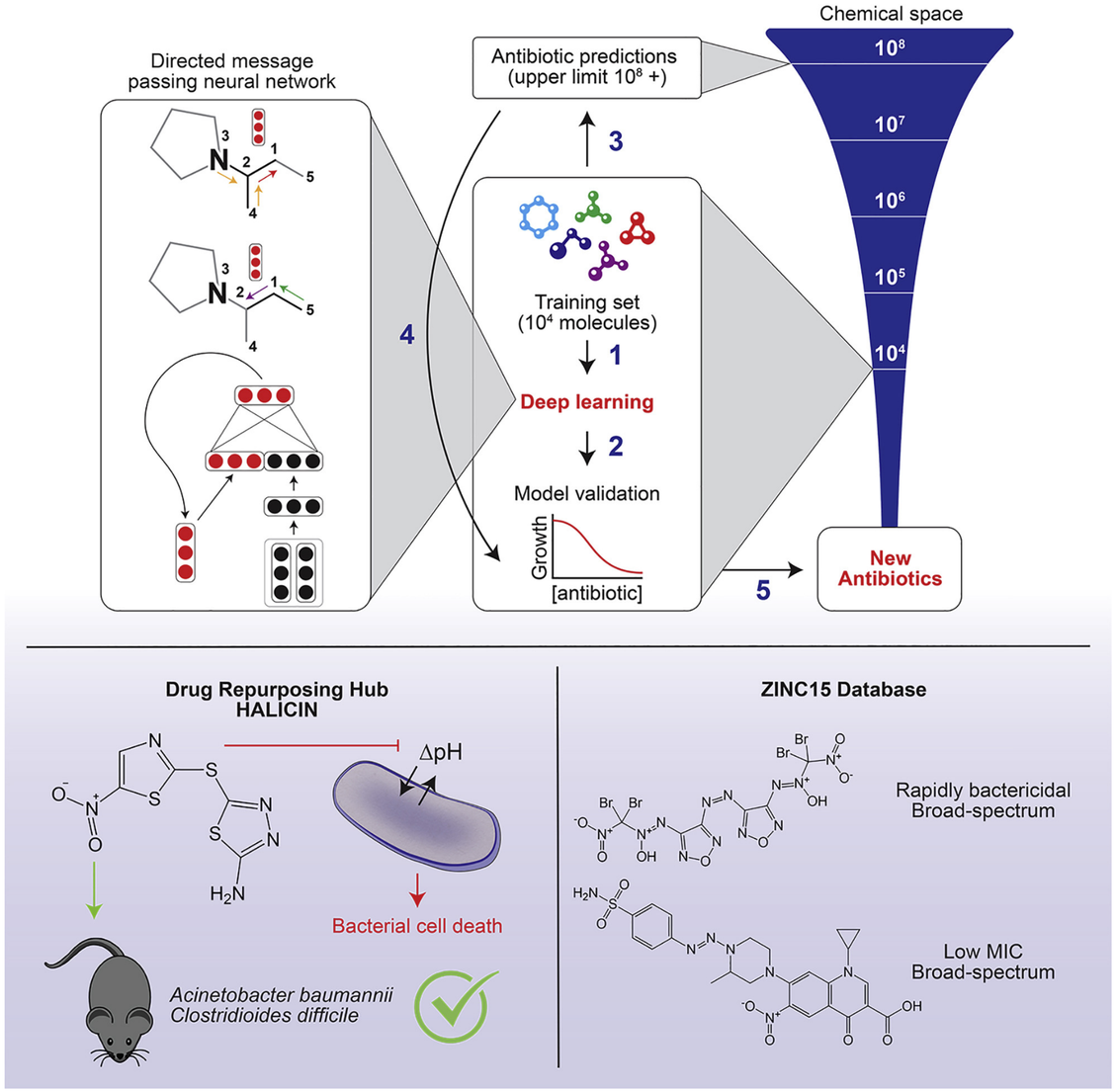
- Train MPNN model to predict the inhibition of the growth of E. coli using a collection of 2,335 diverse molecules.
- Apply the model to multiple chemical libraries and rank the candidate compounds according to the model's predicted score.
- Select a list of promising candidates that can potentially inhibit the growth of E. coli.



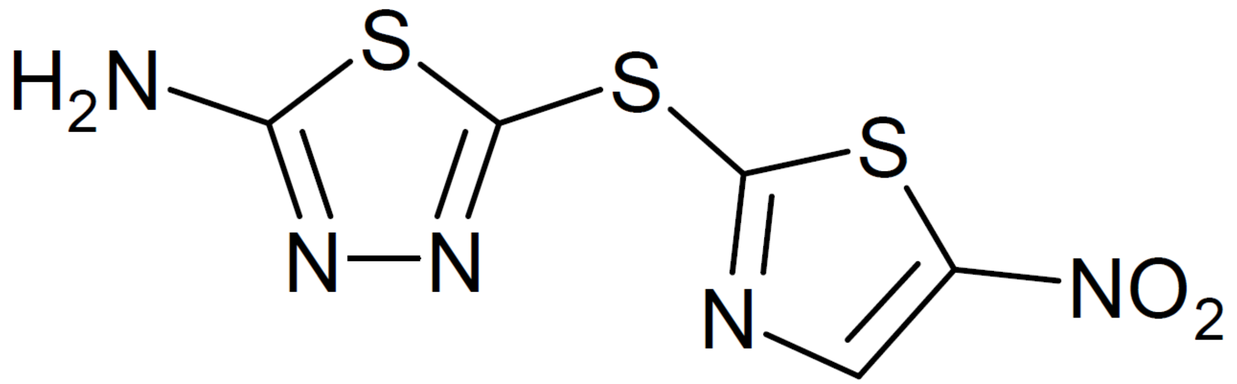
A Message Passing Approach to Antibiotic Discovery

Overview of the results:

- Halicin is identified from Drug Repurposing Hub as a potent inhibitor of E. coli growth.
- Experimental investigations revealed that halicin displays growth inhibitory properties against a wide spectrum of pathogens.
- 8 additional antibacterial compounds discovered based on 23 compounds identified from the ZINC15 database (> 107 million molecules).
- Two of these molecules displayed potent broad-spectrum activity and could overcome an array of antibiotic-resistance determinants in E. coli.



A Message Passing Approach to Antibiotic Discovery



"Indeed, modern neural molecular representations have the potential to: (1) decrease the cost of lead molecule identification because screening is limited to gathering appropriate training data, (2) increase the true positive rate of identifying structurally novel compounds with the desired bioactivity, and (3) decrease the time and labor required to find these ideal compounds from months or years to weeks."

(Stokes et al., 2020)

Protein Folding

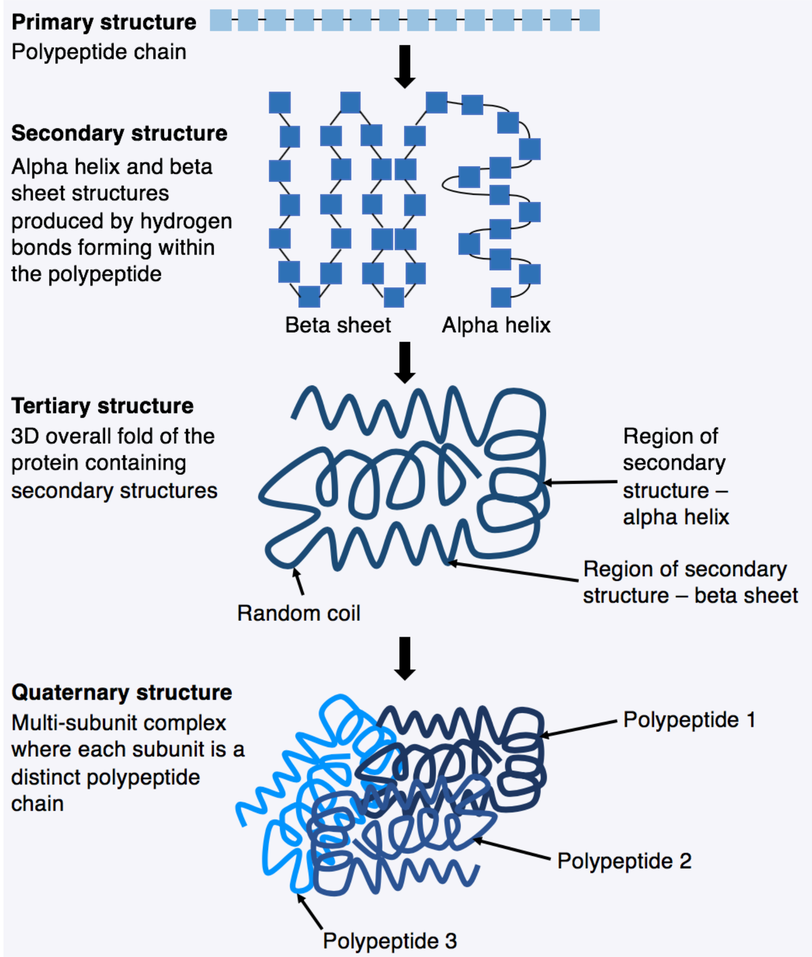
Protein Folding

Protein folding: Chains of amino acids spontaneously fold to form the 3D structures of the proteins.

Computational task: Determine the 3D structure of a protein from a sequence of amino acids.

Highly challenging: Depends on the thermodynamics of the interatomic forces, etc.

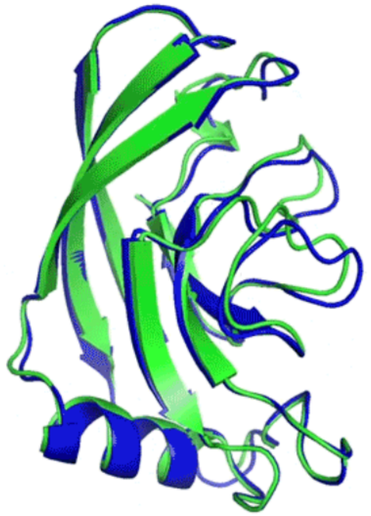
Competitions: To predict the native structure of a protein from its amino acid sequence is a problem of great interest.



By Kep17 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=87932134

Protein Folding: AlphaFold

T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)

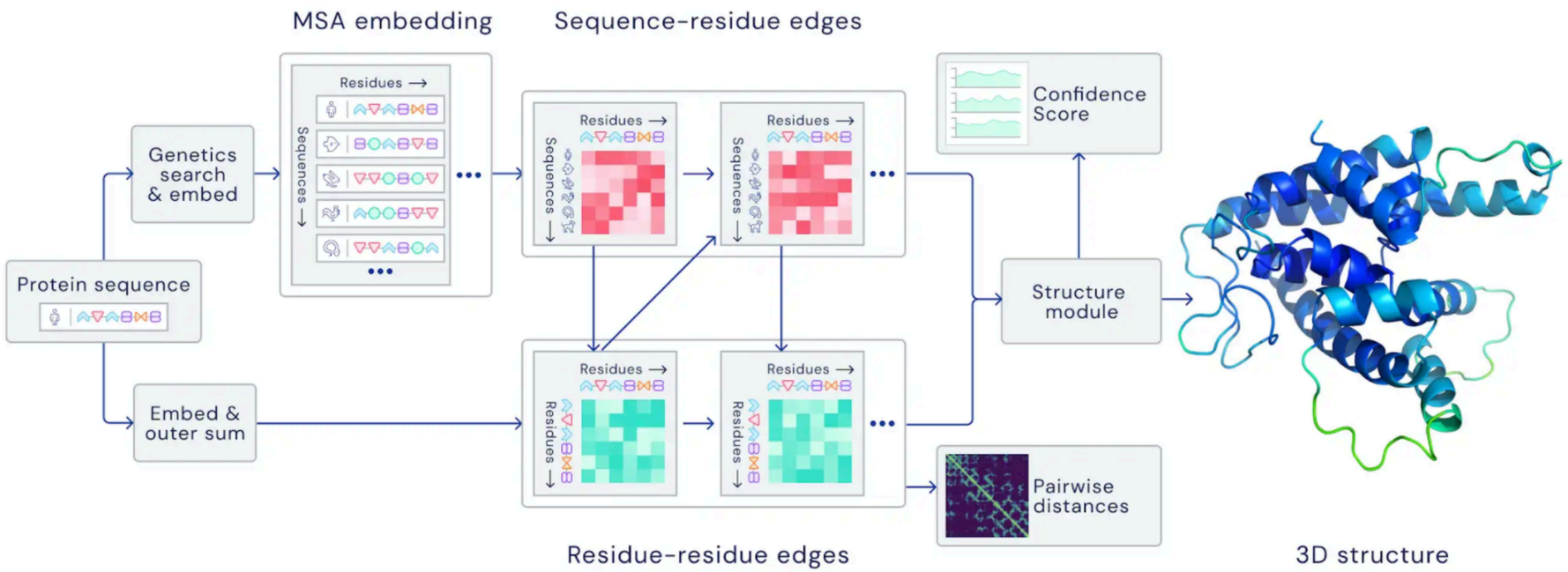


T1049 / 6y4f
93.3 GDT
(adhesin tip)

Experimental result
Computational prediction

A breakthrough by a Deepmind team, read more at:
https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

AlphaFold



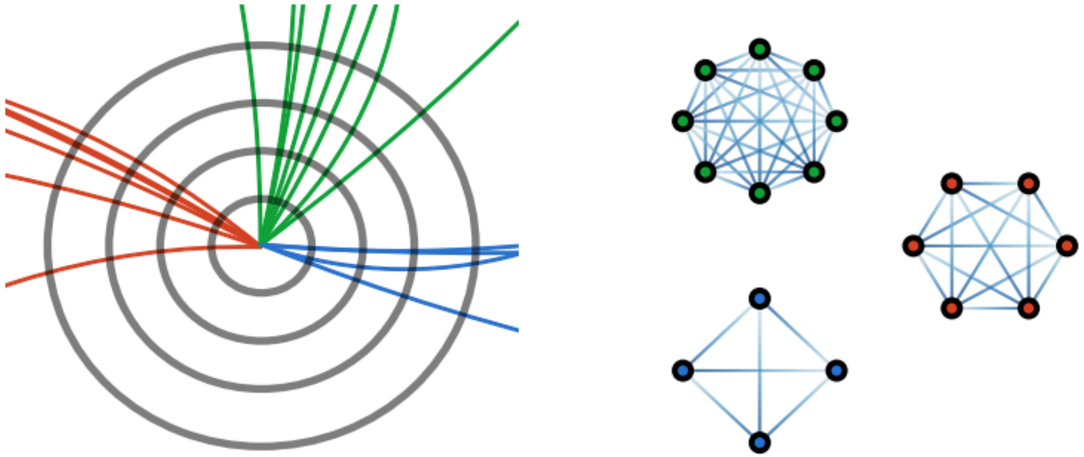
Key ideas behind AlphaFold (Jumper et al., 2021):

Encoding as a spatial graph: Amino acids as nodes, and proximity between amino acids as edges.

GNN approach: Train to predict the new positions of amino acids - allowing to predict the 3D structure.

Particle Physics

Particle Physics



Data in particle physics: Often represented by sets and graphs - GNNs offer key advantages.

We follow the survey by Shlomi et al., (2021), and highlight some applications of GNNs in particle physics.

Particle Physics

Jets: Sprays of stable particles that stem from multiple successive interaction and decays of particles, originating from a single initial object, i.e., quark, gluon, W-boson, top-quark, or Higgs boson.

Jet classification: Identify the original object that gave rise to the jet — a very important task in particle physics.

Jets as graphs: View a jet as a graph, where nodes are particles (with features) and edges represent interactions, and apply graph classification.

Event classification: Predicting the physics process at the origin of the recorded data, e.g., to classify signals in the IceCube neutrino observatory (Choma et al., 2018).

Many other problems are of interest, e.g., jet clustering.

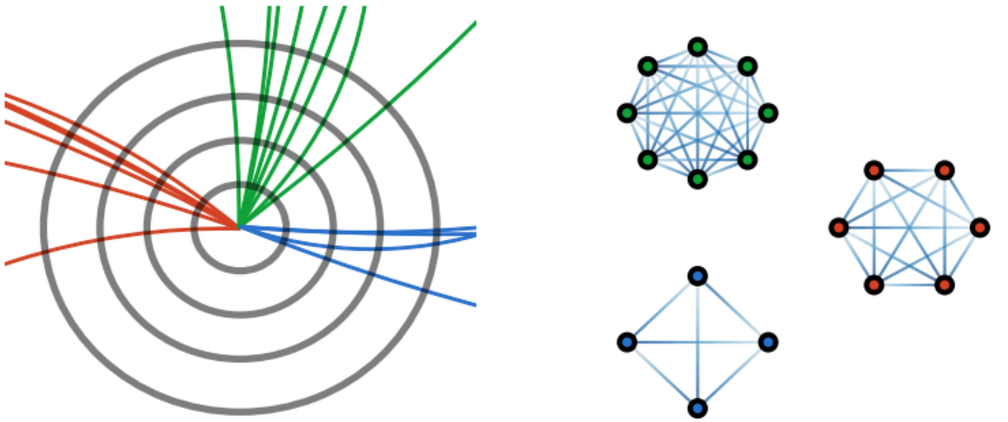


Figure (Shlomi et al., 2021), depicting jet classification based on the particles associated to the jet.

Combinatorial Optimization and Reasoning

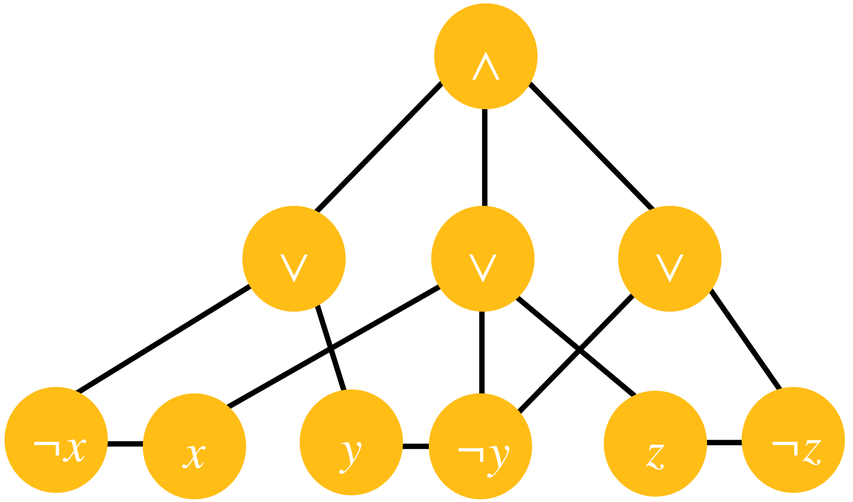
Reasoning Capacity of Graph Neural Networks

What is the reasoning capacity of GNNs?

Not plausible for GNNs to solve NP-hard problems.

Example: Can GNNs learn to solve (small) SAT instances with single-bit supervision (Selsam et al., 2018)?

- Represent each propositional formula as a graph.
- Produce training data, based on existing SAT solvers.
- Train the GNN to predict satisfiability of novel formulas.



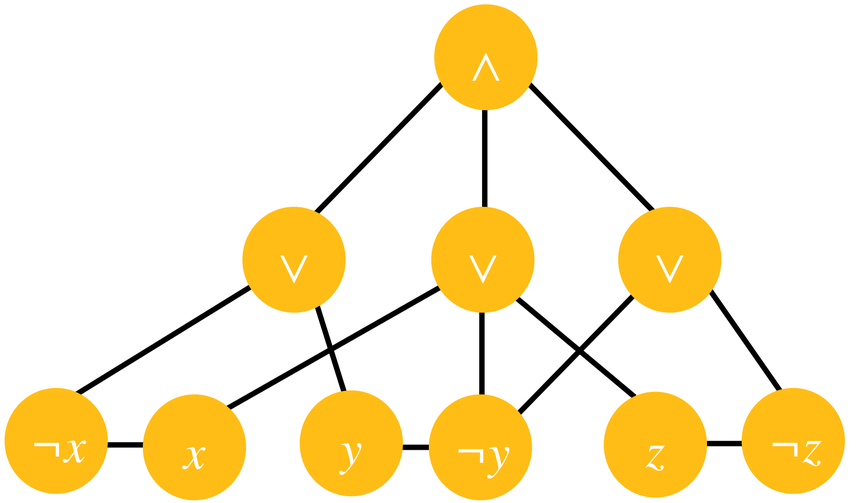
$$\phi = (\neg x \vee y) \wedge (x \vee \neg y \vee z) \wedge (\neg y \vee \neg z)$$

Reasoning Capacity of Graph Neural Networks

GNNs are a good choice:

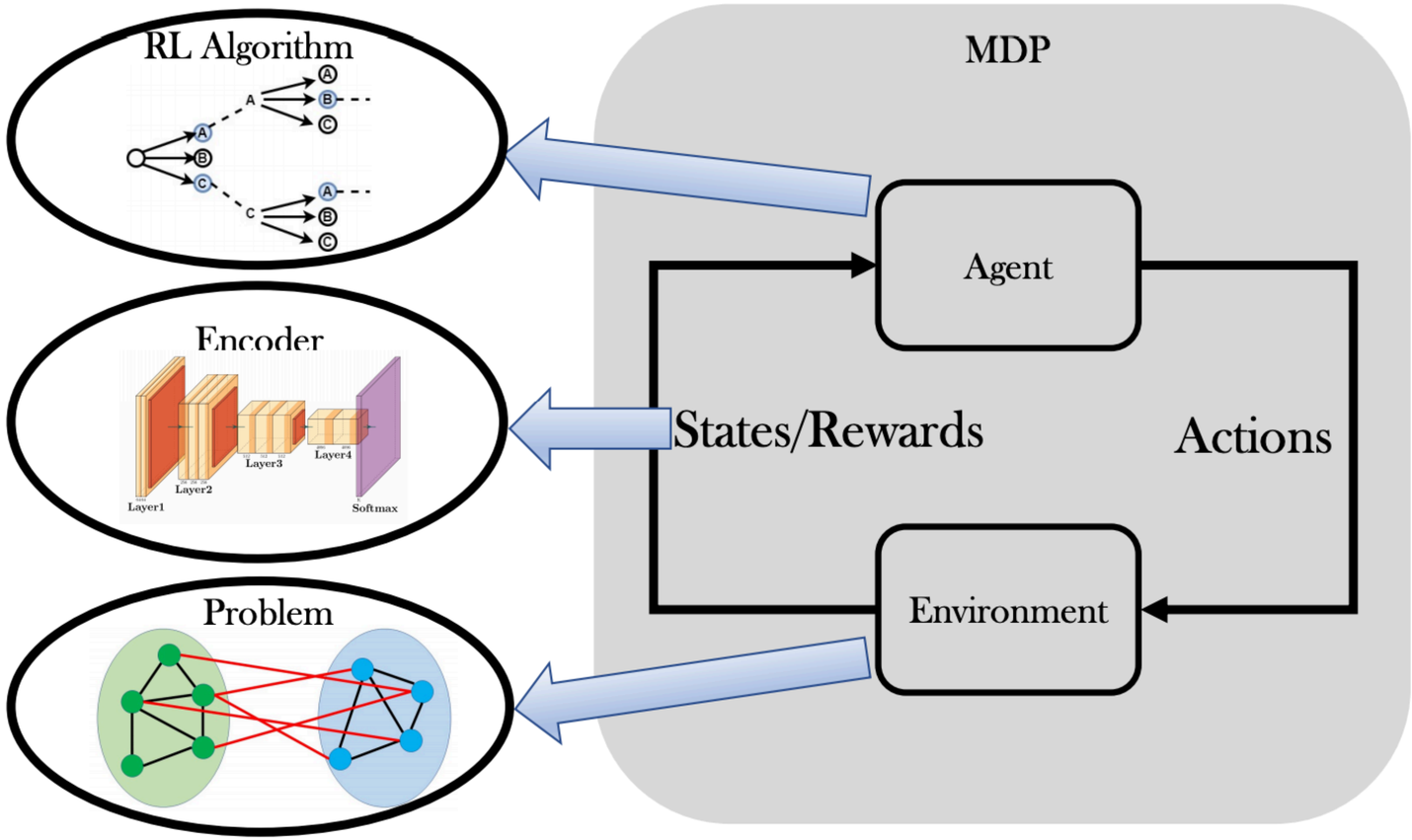
- Explicit structural encoding of an input formula.
- Permutation-invariance, e.g.,
  $(x \wedge y) \vee (\neg x \wedge \neg y) \equiv (\neg x \wedge \neg y) \vee (x \wedge y)$.
- Naming-invariance, e.g.,
  $(x \wedge y) \vee (\neg x \wedge \neg y) \equiv (\neg z \wedge \neg u) \vee (z \wedge u)$,
- Strong inductive bias, given by formula distinguishability.
- Separate representations for logical operators $\wedge$, $\vee$.

Many other problems, such as TSP, #SAT, etc. are investigated.



$$\phi = (\neg x \vee y) \wedge (x \vee \neg y \vee z) \wedge (\neg y \vee \neg z)$$

Graph Neural Networks and Combinatorial Problems



GNNs alone are limited for such problems, and a line of work combines the power of GNNs with reinforcement learning for solving combinatorial problems. Figure (Mazyavkina et al., 2020) shows the pipeline.

Graph Neural Networks and Combinatorial Problems

Idea: Model the problem as a sequential decision-making process, e.g., MDP, where the agent interacts with the environment by performing a sequence of actions in order to find a solution.

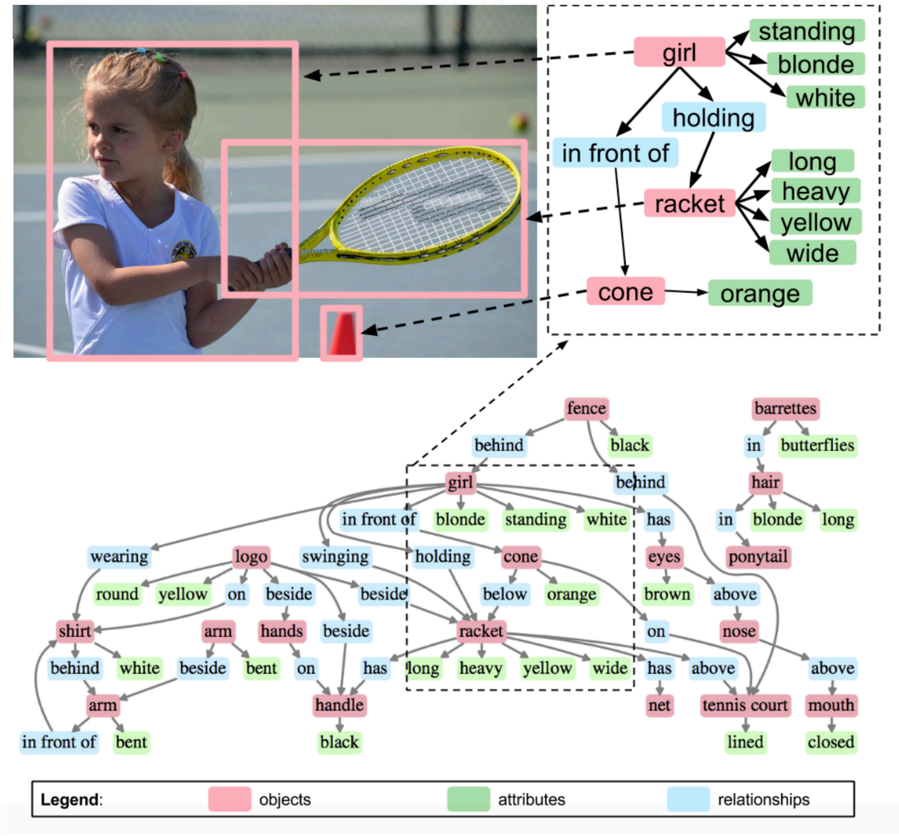
Goal: An agent acting in MDP tries to find a policy function that maps states into actions, while maximizing the expected cumulative discounted sum of rewards, i.e., finding an optimal policy.

A typical run is as follows:

- Formulate the combinatorial problem (e.g., Max-Cut), as an MDP.
- Use a reinforcement learning algorithm (e.g., Monte-Carlo Tree Search) to move the environment to the next state (e.g., removing a vertex from a solution set).
- Encode states with a GNN: Map states to the actions' values (e.g., replacing simulation step in MCTS).
- Once the model is trained, the agent can search the solutions for unseen instances of the problem.

Computer Vision: Scene Graphs and Question Answering

Scene Graphs



Scene graph (Johnson et al., 2015): A data structure that describes the contents of a scene, encoding object instances, attributes of objects, and relationships between objects.

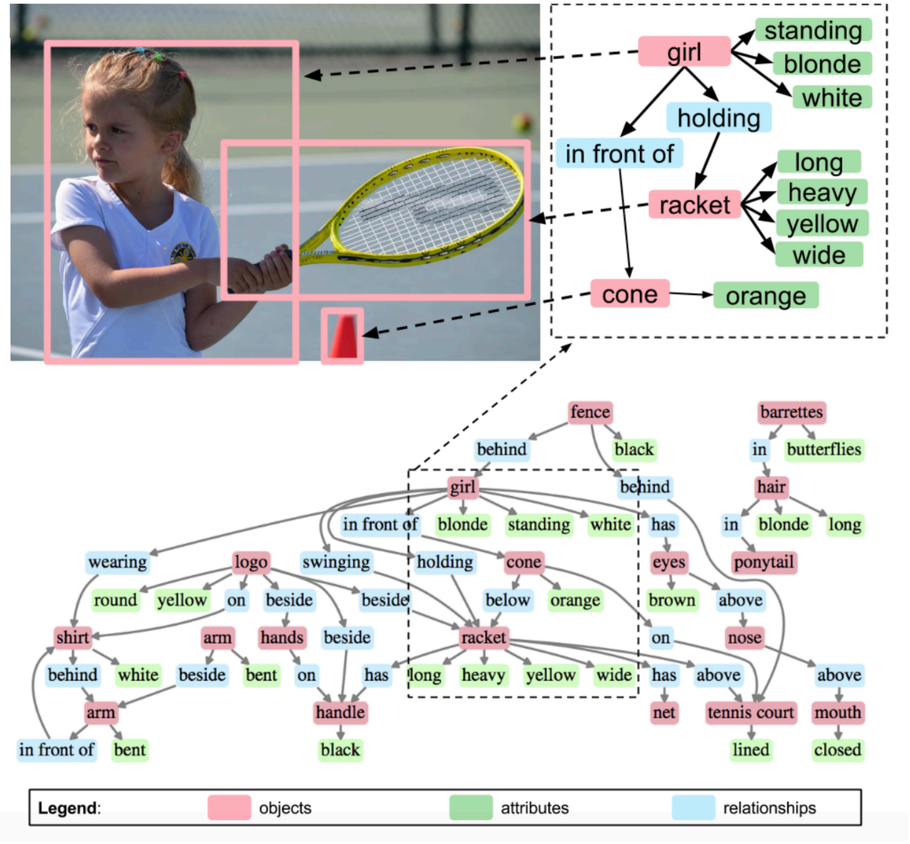
Scene Graphs

Retrieving images/videos by describing their contents is an exciting application of computer vision.

Example: A system may allow people to search for images by specifying not only objects ("man", "boat") but also structured relationships ("man on boat") and attributes ("boat is white") involving these objects.

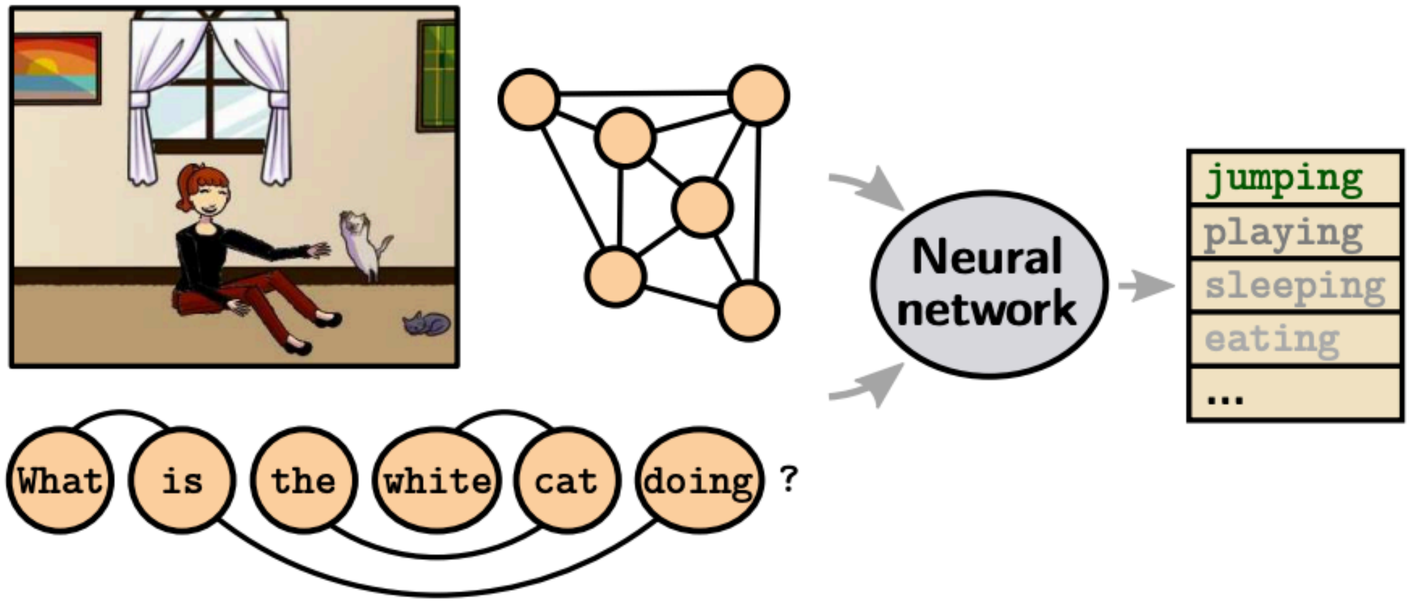
Structured data: Explicitly represent and reason about the objects, attributes, and relationships in images.

Tasks: GNNs are used in both generating scene graphs and for high-level tasks that one would be interested in performing on them, e.g., visual question answering.



Scene graph (Johnson et al., 2015).

Visual Question Answering



Encode the input scene as a graph representing the objects and their spatial arrangement. Encode the input question as a graph representing words and their syntactic dependencies. Train a neural network to reason over these representations, and to produce a suitable answer as a prediction (Tenet et al., 2016).

Visual Question Answering

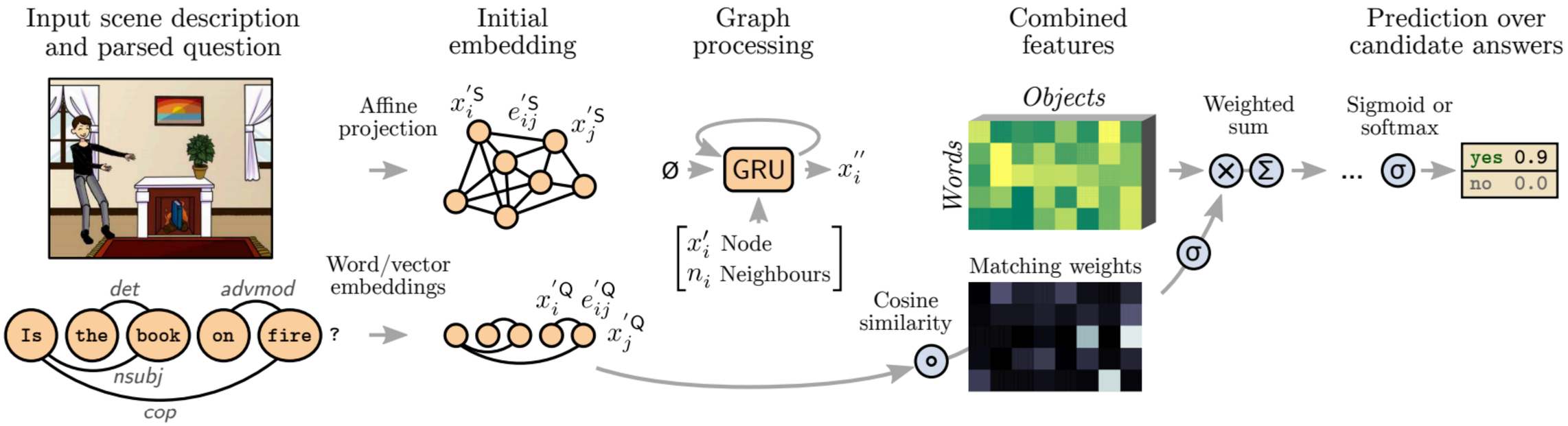


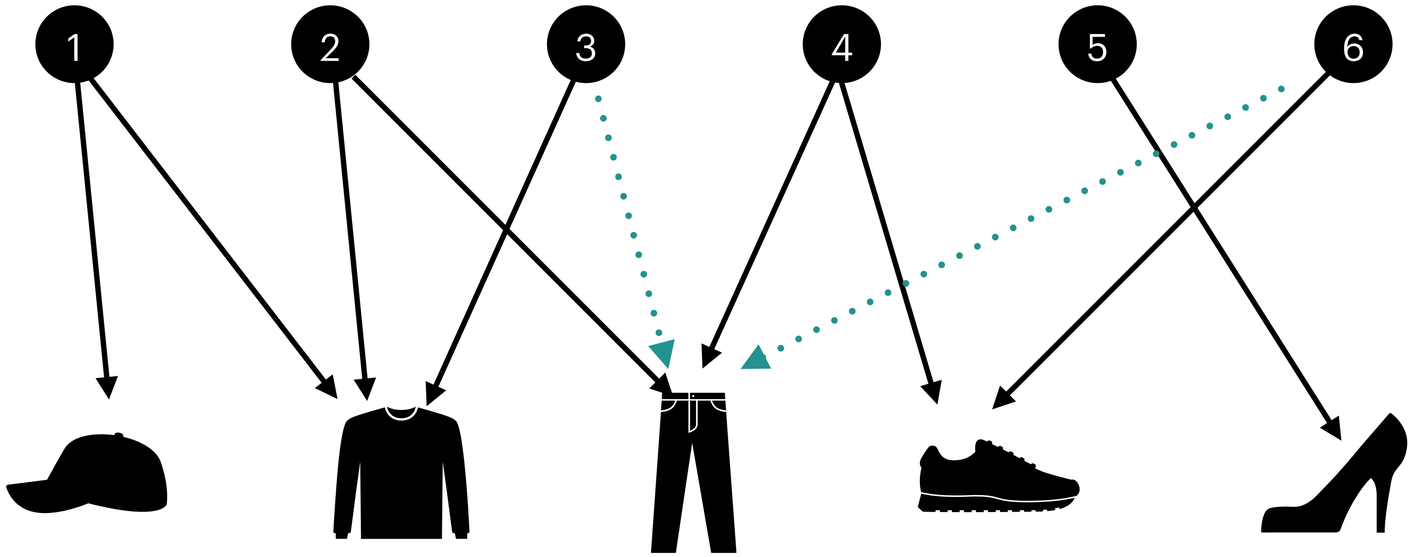
Figure (Tenet et al., 2016) illustrating a pipeline for visual question answering using gated GNNs.

Recommender Systems

Recommender Systems

Users

Items



Interactions

Traffic Forecasting

Road Networks

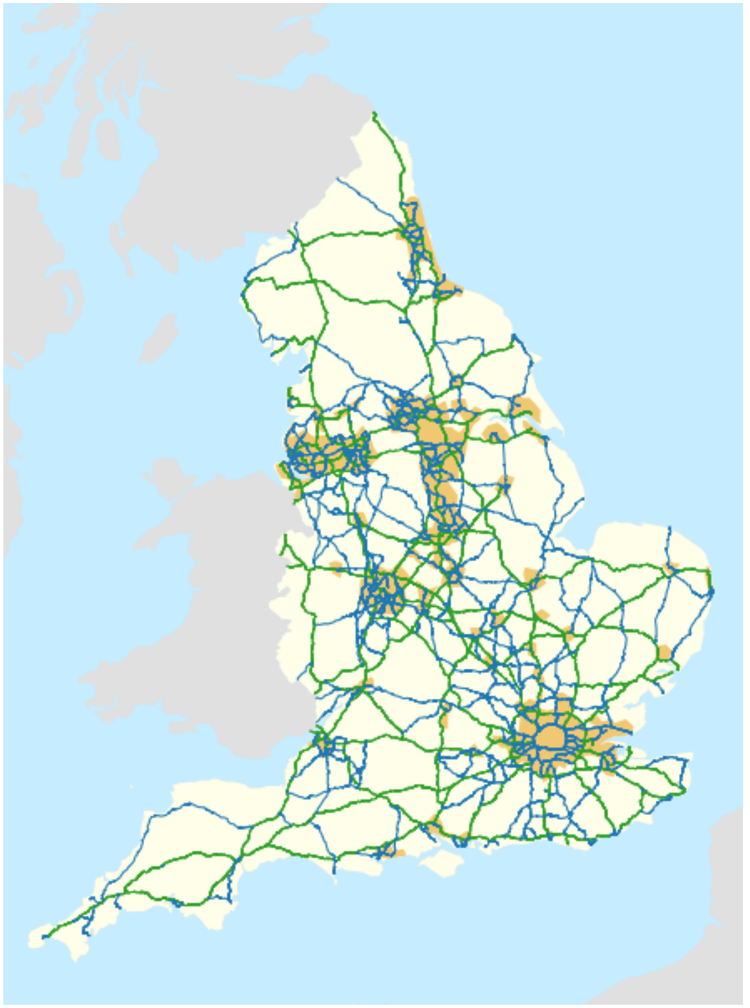
Road networks: Real-world road networks can be at many scales and are of relational nature.

Complex system: Real-world traffic is complex as it is affected by live traffic conditions, historical traffic patterns, rush hours, road quality, speed limits, accidents, and closures.

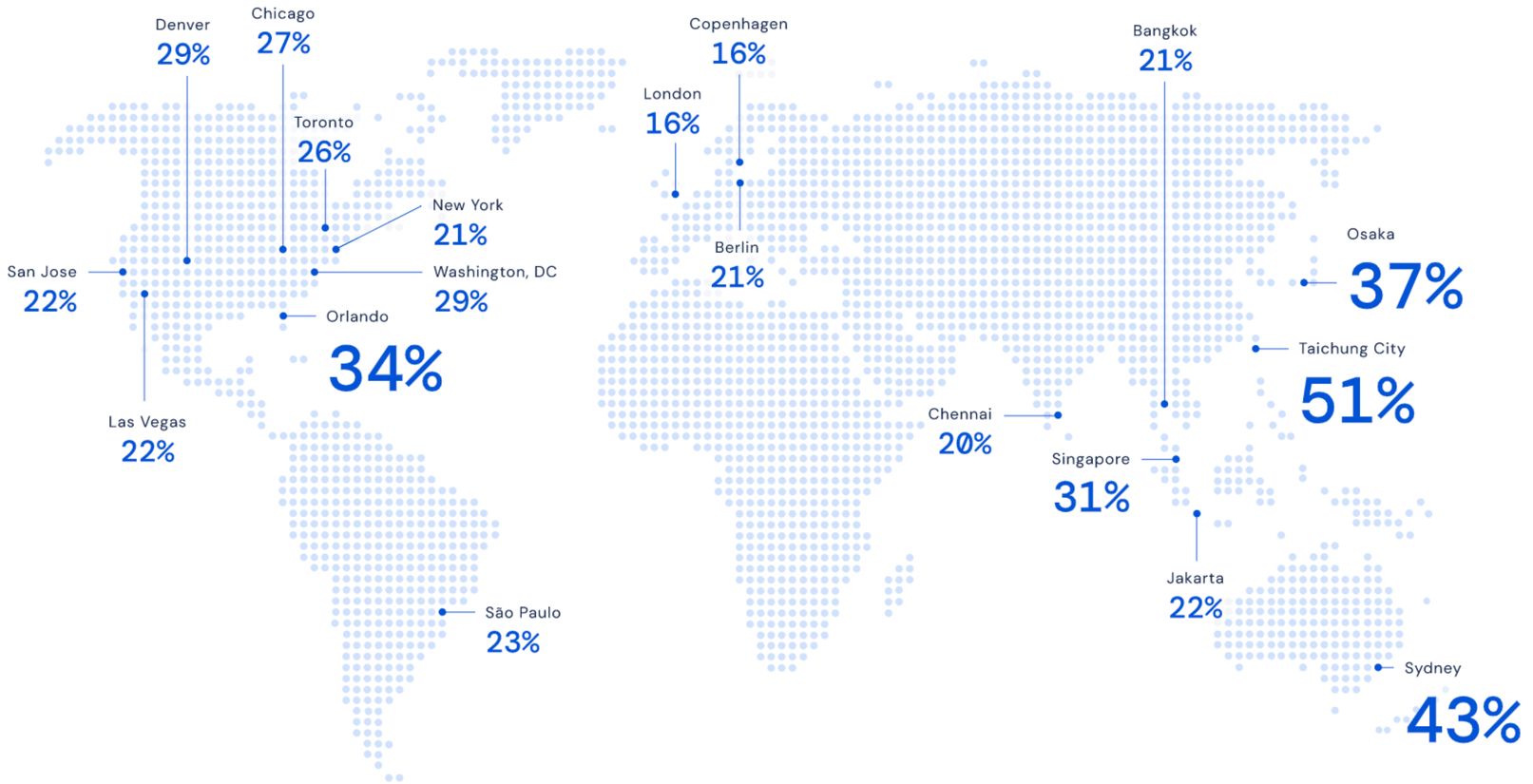
Structural constraints: Some routes may be invalid, and traffic regulations need to be taken into account.

Tasks: Calculating routes, estimating arrival times...

Spatiotemporal domain: Account for location and time, while incorporating relational learning biases.

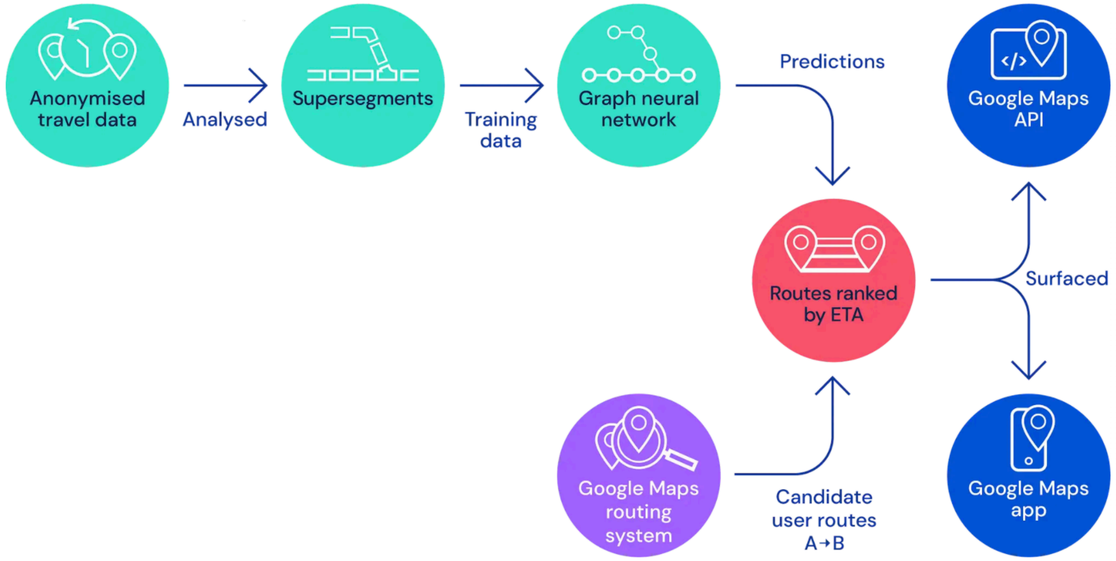


Improving Predictions for Estimated Time of Arrival



DeepMind partnered with the Google Maps to improve the accuracy of real time ETAs by up to 50 %. Read more at: https://deepmind.com/blog/article/traffic-prediction-with-advanced-graph-neural-networks

Improving Predictions for Estimated Time of Arrival



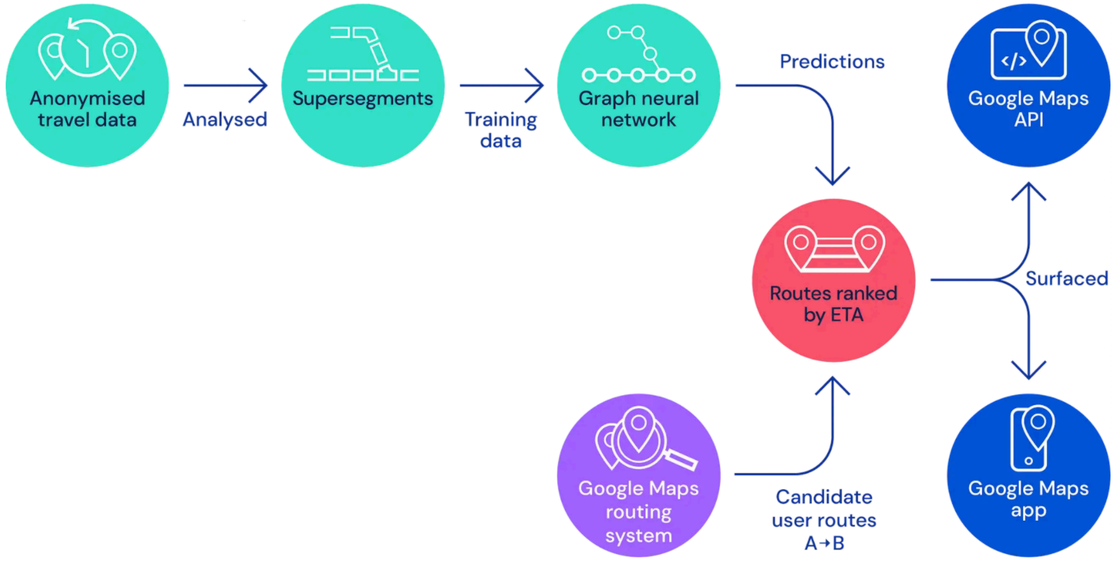
A GNN based approach to improve ETA.

Approach based on organizing the road network into super-segments: Multiple adjacent segments of road that share significant traffic volume:

(1) **Route analyzer**: Processes terabytes of traffic information to construct super-segments.

(2) **A GNN model**: Optimized with multiple objectives to predict the travel time for each super-segment.

Improving Predictions for Estimated Time of Arrival



A GNN based approach to improve ETA.

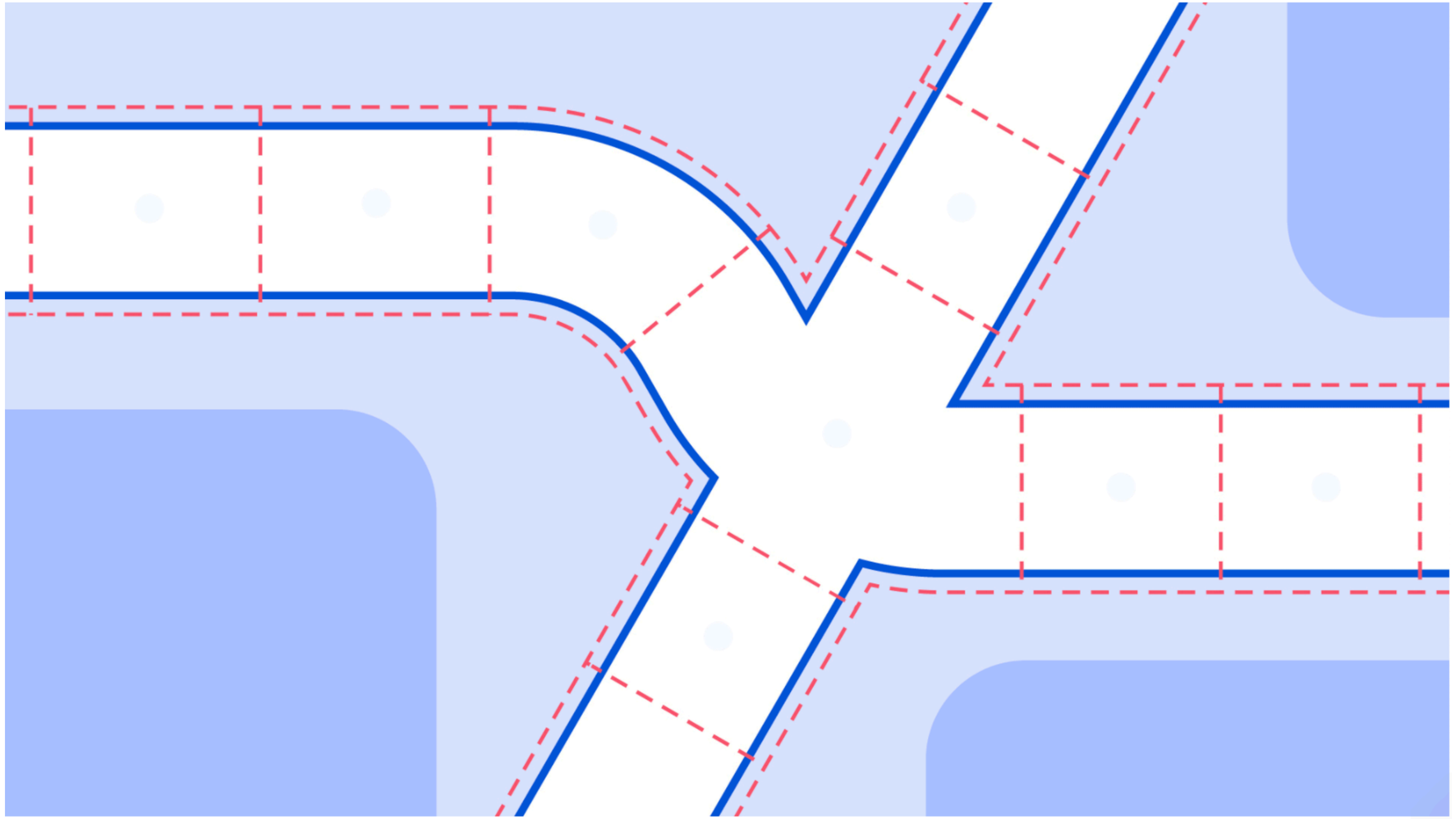
Why is a GNN good choice?

(1) **Super-segments**: Each super-segment is composed of dynamically sized number of segments

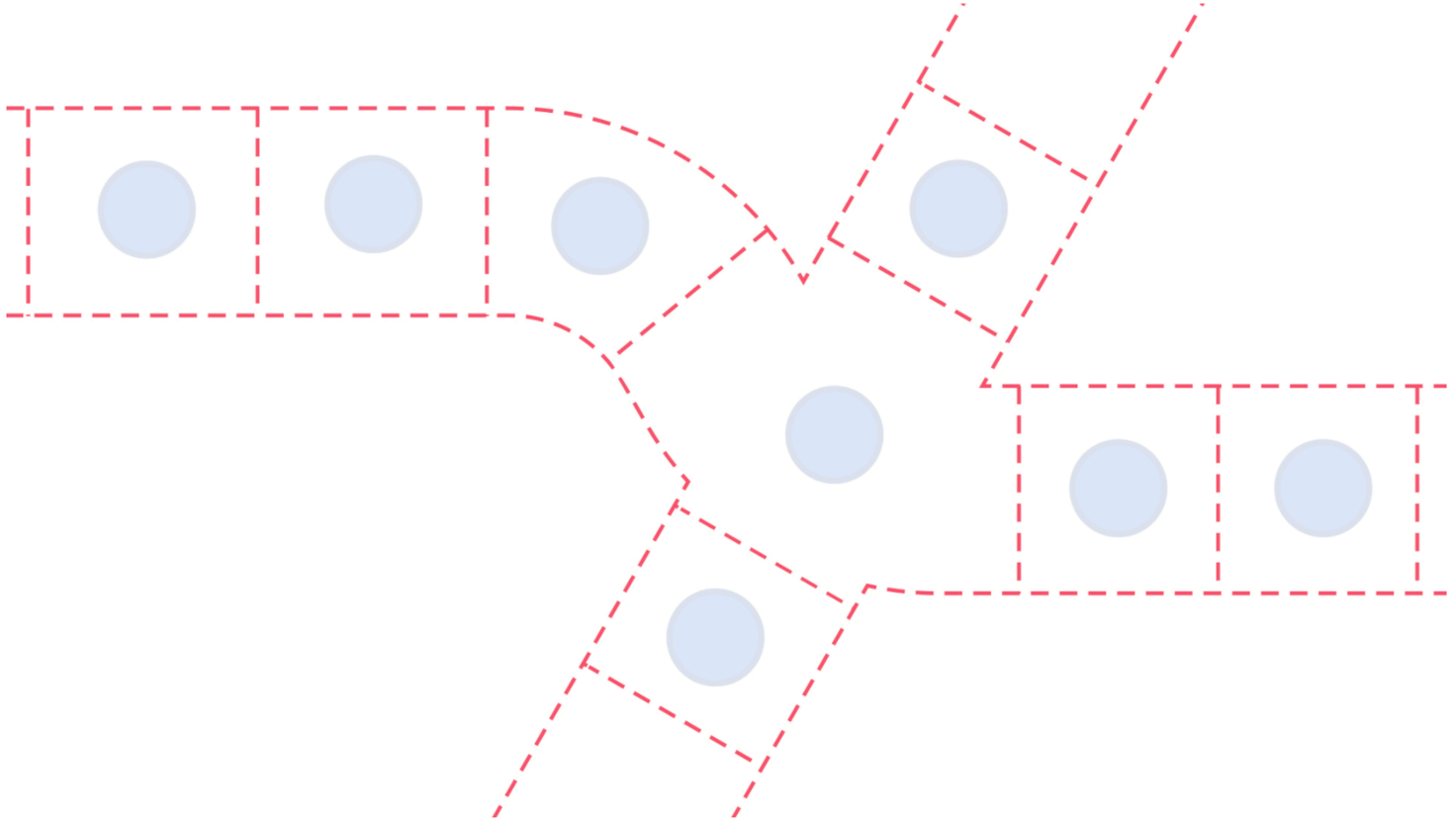
(2) **Naive idea**: Use a model for each super-segment - At scale, millions of these models are needed!

(3) **Quest for a single model**: One model that can handle each super-segment.

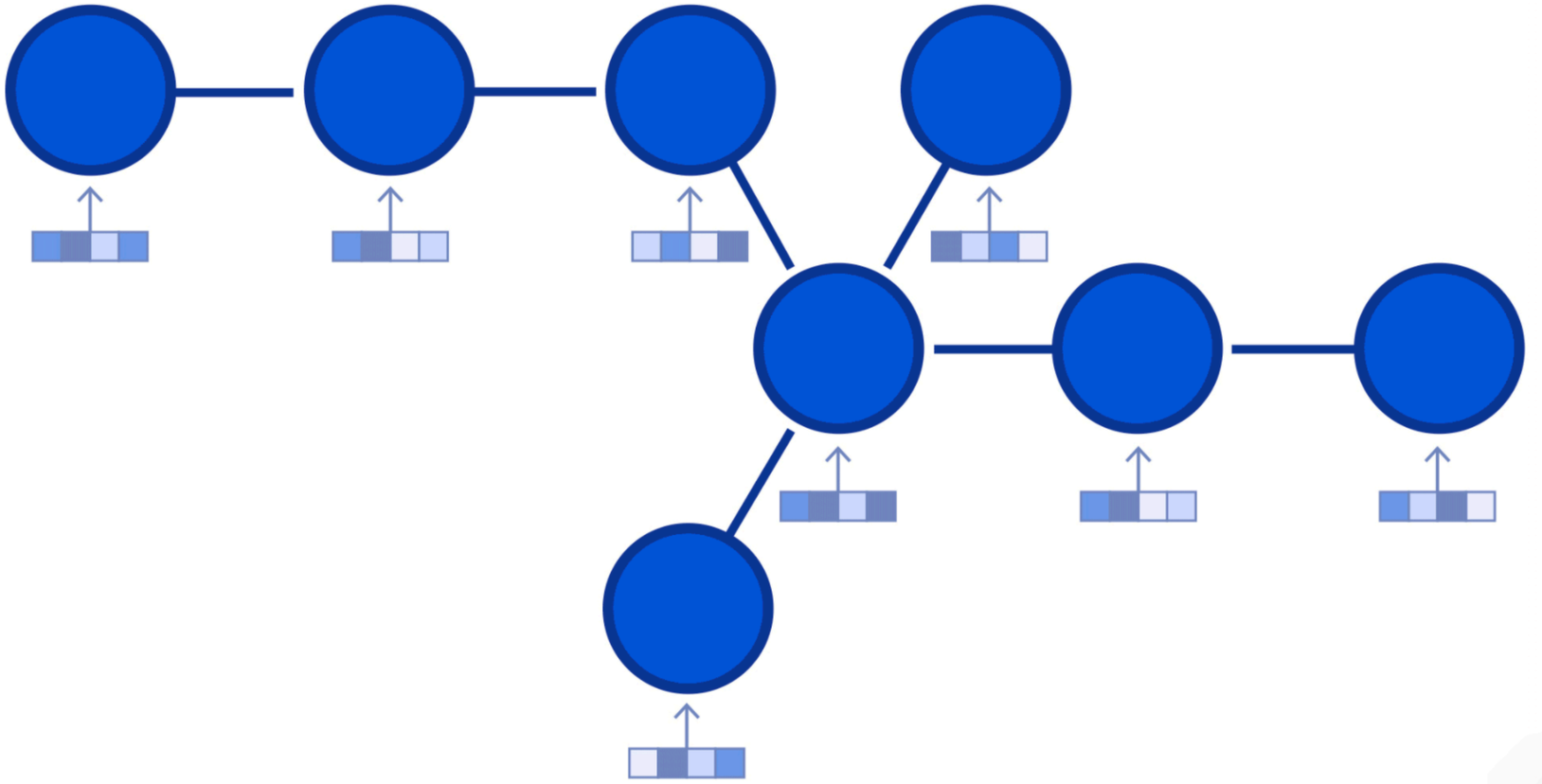
Improving Predictions for Estimated Time of Arrival



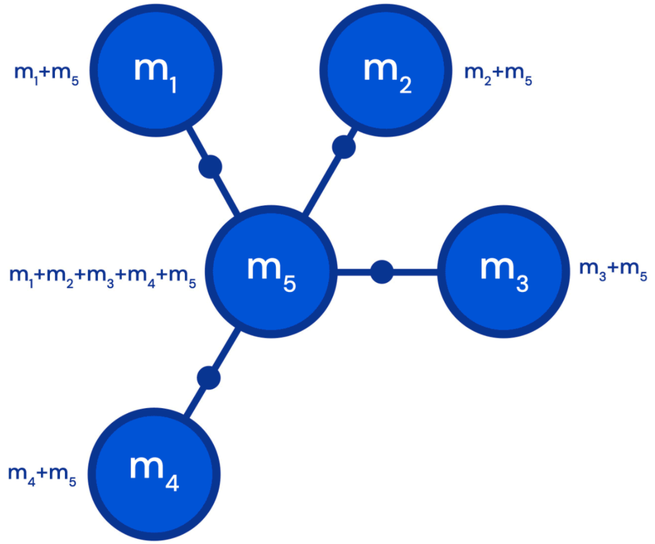
Improving Predictions for Estimated Time of Arrival



Improving Predictions for Estimated Time of Arrival



Improving Predictions for Estimated Time of Arrival



Powerful encoding: Not only traffic ahead or behind, but also along adjacent and intersecting roads.

Model capacity: Each super-segment, which can be of varying length and of varying complexity - from simple two-segment routes to routes containing hundreds of nodes - shall be processed by the same model.

Empirical gains: Expanding to include adjacent roads that are not part of the main road gives improvements, e.g., a jam on a side street can affect traffic on a larger road.

Summary

- Biomedical data: inherently relational
- Drug discovery: target identification, small molecule therapies, drug repurposing
- AI-assisted antibiotic discovery: Halicin
- Protein folding: From sequences to 3D structures
- Particle physics: jet classification, event classification
- Combinatorial optimization: reinforcement learning and GNNs
- Computer vision: Scene graphs and visual question answering
- Recommender systems
- Traffic forecasting: improving estimated arrival time with graph encodings
- Plethora of recent applications, e.g., assisting mathematicians to conjecture and prove theorems in knot theory.

Summary of the Relational Learning Theme

- Lecture 1. **Relational data & node embeddings**: Model properties, inductive capacity, expressiveness, evaluation.
- Lecture 2. **Knowledge graph embedding models**: translational, bilinear, box embedding models, and beyond.
- Lecture 3. **Graph neural networks**: motivation, permutation-invariance, permutation-equivariance, message passing neural networks, generalizations, graph representation learning tasks.
- Lecture 4. **Message passing neural network architectures**: GGNN, GCN, GAT, GIN, rGCN.
- Lecture 5. **Expressive power of message passing neural networks**: graph isomorphism, 1-WL equivalence, logical characterization, limitations.
- Lecture 6. **Higher-order graph neural networks**: k-GNN, invariant/equivariant graph networks, PPGNs, homophily, heterophily.
- Lecture 7. **Message passing neural networks and randomization**: universality, permutation-invariance, evaluation.
- Lecture 8. **Generative graph neural networks**: variational, adversarial, autoregressive.
- Lecture 9. **Applications of graph neural networks**: life sciences, combinatorial optimization, traffic forecasting, computer vision, etc.

Thanks!

Good luck with your projects...

References

- T. Gaudelet , B. Day , A. Jamasb, J. Soman , C. Regep , G. Liu1 , J. B. R. Hayter , R. Vickers , C. Roberts, J. Tang, D. Roblin, T. L. Blundell, M. M. Bronstein, and J. P. Taylor-King. Utilising Graph Machine Learning within Drug Discovery and Development, *arXiv*, 2021.

- Ganapathiraju, M. K., Thahir, M., Handen, A., Sarkar, S. N., Sweet, R. A., Nimgaonkar, V. L., Loscher, C. E., Bauer, E. M., & Chaparala, S. (2016). Schizophrenia interactome with 504 novel protein-protein interactions. *NPJ schizophrenia*, *2*, 16012.

- Rao, P.P., Kabir, S.N., & Mohamed, T.S. (2010). Nonsteroidal Anti-Inflammatory Drugs (NSAIDs): Progress in Small Molecule Drug Development. *Pharmaceuticals, 3*, 1530 - 1549.

- Hewett M, Oliver DE, Rubin DL, et al. Pharmgkb: the pharmacogenetics knowledge base. *Nucleic Acids Res* 2002;30(1):163–5.

- Jonathan M. Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M. Donghia, Craig R. MacNair, Shawn French, Lindsey A. Carfrae, Zohar Bloom-Ackerman, Victoria M. Tran, Anush Chiappino-Pepe, Ahmed H. Badran, Ian W. Andrews, Emma J. Chory, George M. Church, Eric D. Brown, Tommi S. Jaakkola, Regina Barzilay, James J. Collins. A Deep Learning Approach to Antibiotic Discovery. *Cell*, 2020; 180 (4): 688

- J. Shlomi, P. Battaglia, J. Vlimant. Graph Neural Networks in Particle Physics. *Machine Learning: Science and Technology*, 2021.

References

- N. Choma, F. Monti, L. Gerhardt, T. Palczewski, Z. Ronaghi, M. Prabhat, W. Bhimji, M. Bronstein, S. Klein, J. Bruna. Graph Neural Networks for IceCube Signal Classification. *ICMLA*, 2018.
- D. Selsam, M. Lamm, B. Bunz, P. Liang, D.L. Dill, L. de Moura. Learning a SAT solver from single-bit supervision, *ICLR*, 2019.
- J. Johnson, R. Krishna, M. Stark, L. Li, D. A. Shamma, M. S. Bernstein, L. Fei-Fei. Image Retrieval using Scene Graphs, *CVPR*, 2015.
- D. Teney, L. Liu, A. Hengel. Graph-Structured Representations for Visual Question Answering, *CVPR,* 2016.
- A. Agarwal, A. Mangal,Vipul. Visual Relationship Detection using Scene Graphs: A Survey, *arXiv*, 2020.
- N. Mazyavkina , S. Sviridov, S. Ivanov and E. Burnaev. Reinforcement Learning for Combinatorial Optimization: A Survey, *arXiv*, 2020.
- Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596,** 583–589 (2021).