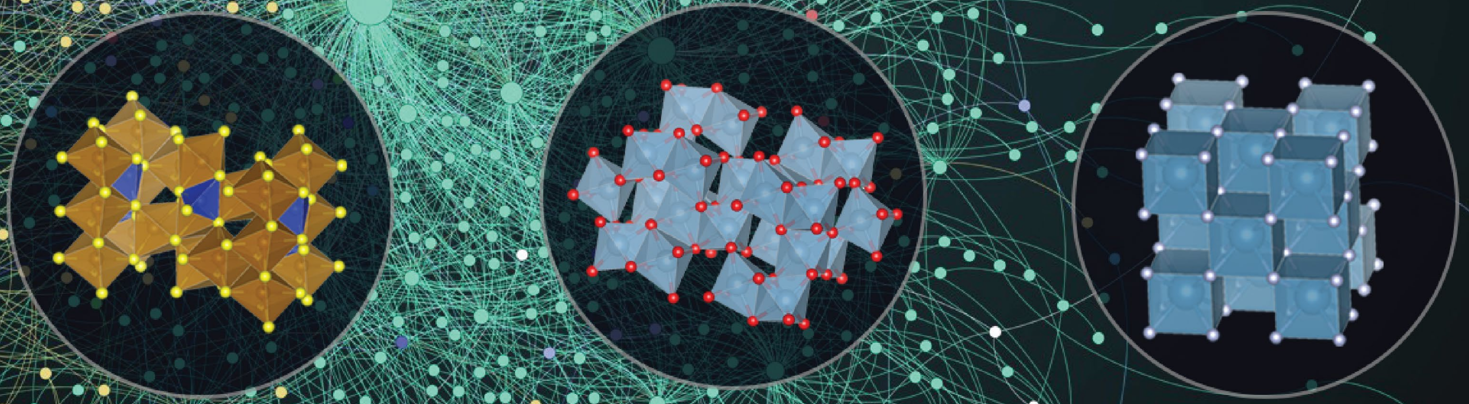
Crystal Structure graphs



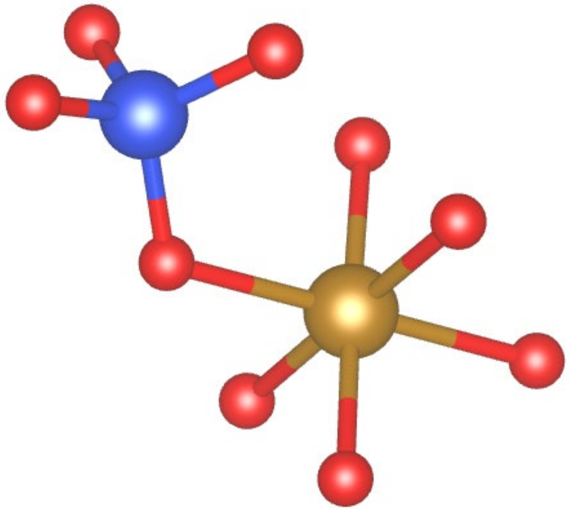
Graphs are a convenient and powerful way to represent information

Nodes: email users
Edges: email message

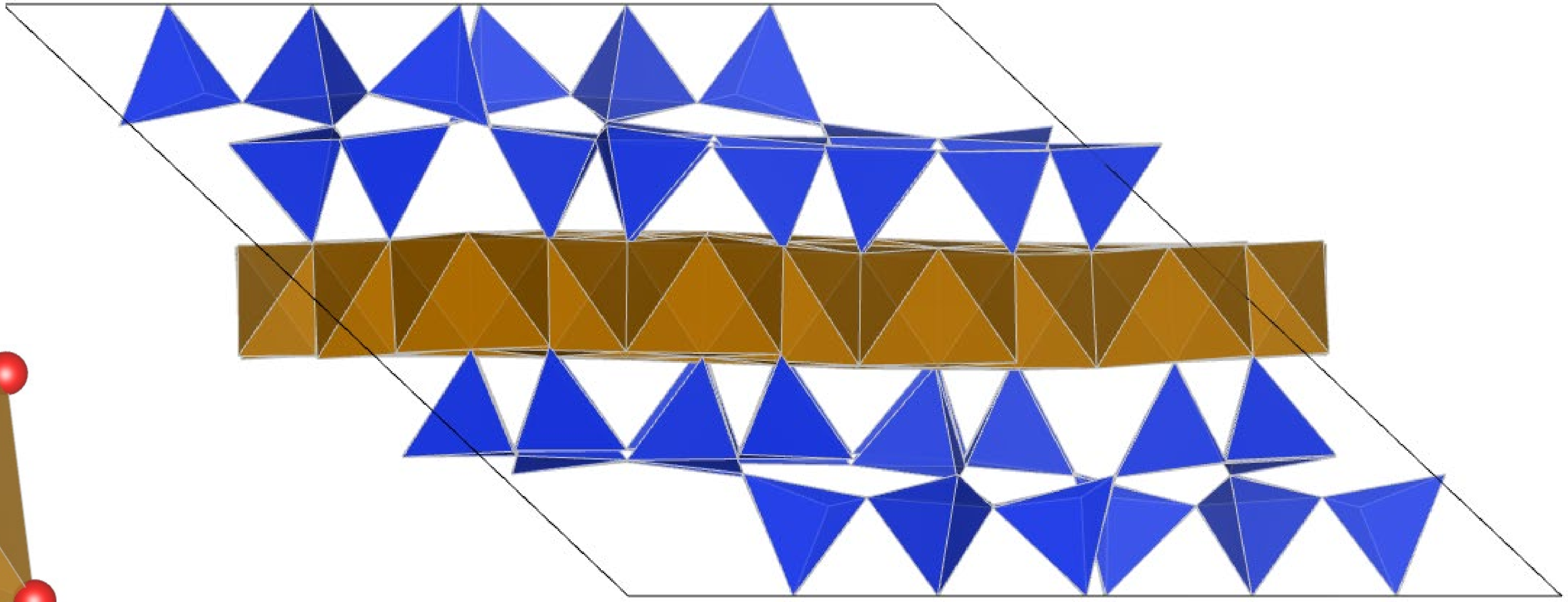
Nodes: people
Edges: relationships

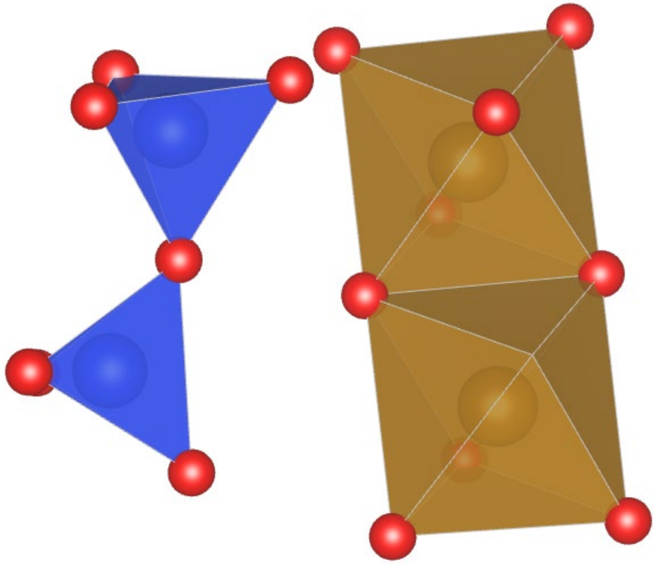
Nodes: papers
Edges: citations

Crystals and molecules are great applications of graph networks

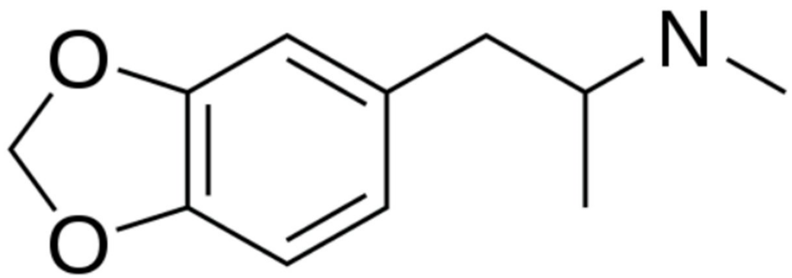


Nodes: atoms
Edges: bonds





Representing molecular Graphs



String-based Representation

SMILES: CC(CC1=CC2=C(C=C1)OCO2)NC

DeepSMILES: CCCC=CC=CC=C6))OCO5)))))))NC

SELFIES: [C][C][B1_1][P][C][C][=C][C][=C]
[B1_1][B1_1][C][=C][R1][B1_2]
[O][C][O][R1][B1_3][N][C]

In general:
Strings can encode **Turing machines** (i.e. every possible classical computation), thus are more powerful representations than simple matrices and vectors.

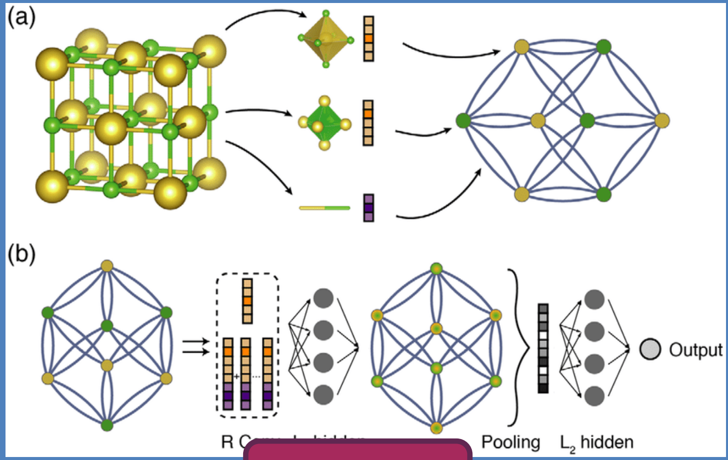
<u>Examples</u>: 3D structure, non-covalent bonds,
physical constraints, …

Adjacency-Matrix-based Representation

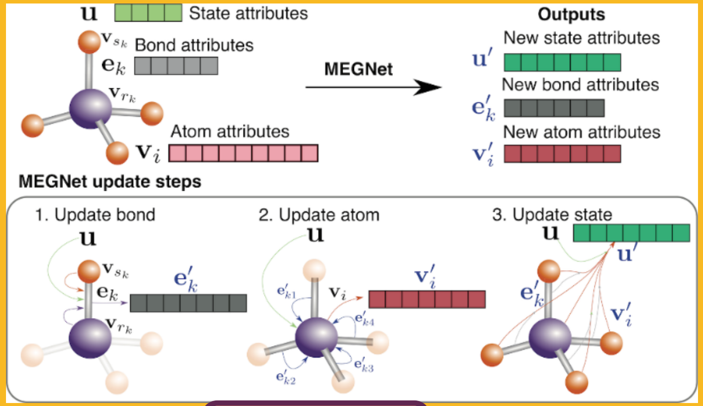
$$\begin{array}{cc}
\begin{pmatrix} C \\ O \\ C \\ O \\ C \\ C \\ C \\ C \\ C \\ C \\ C \\ N \\ C \\ C \end{pmatrix} &
\begin{pmatrix}
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{pmatrix}
\end{array}$$

https://twitter.com/MarioKrenn6240/status/1348814674415001600?s=20

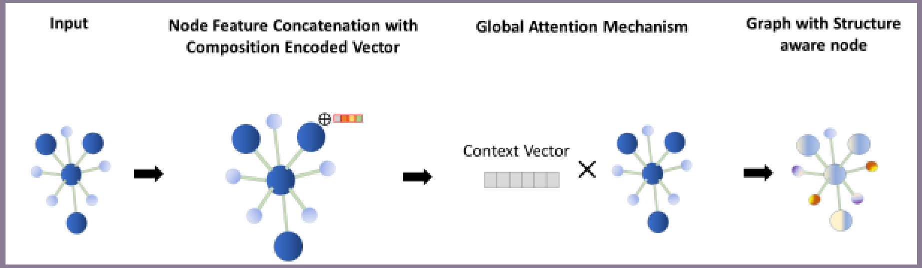
Some of the most powerful materials informatics tools are graph networks



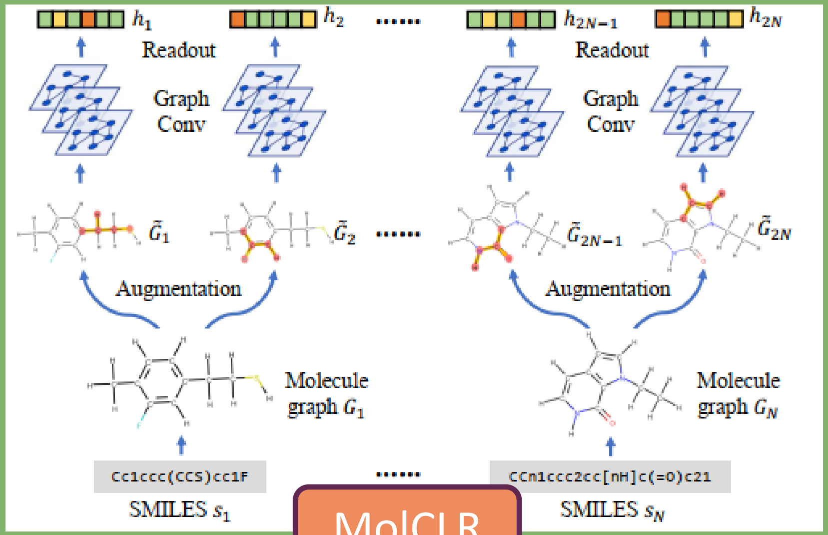
CGCNN



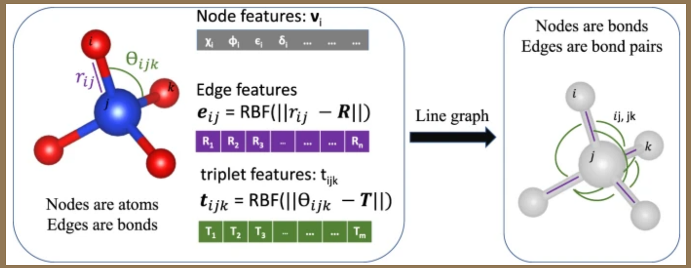
MEGNet



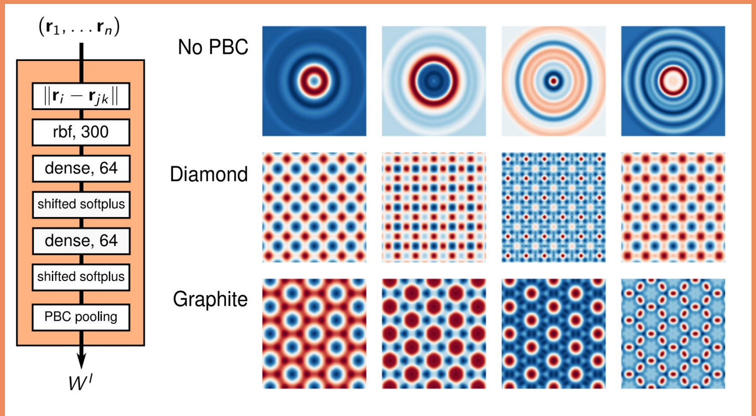
GATGNN



MolCLR



ALIGNN



SchNet

Do we really need new graph networks?

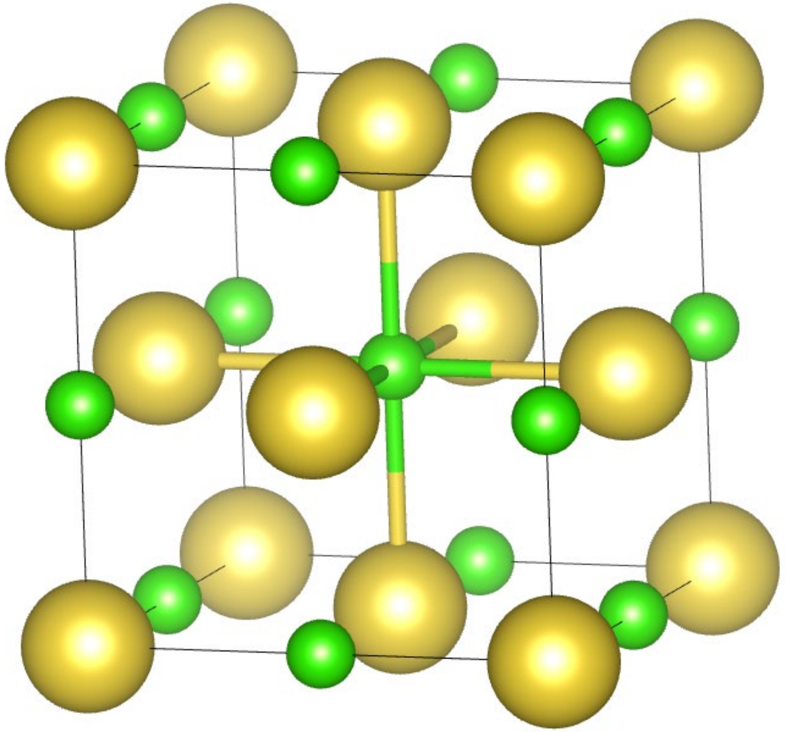
Crystal Topological Graph Neural Network

Do we really need new graph networks?

Current GNNs are…
Too connected
Too slow
Too large

Current graph networks find neighbors within some volume, and then fully connect nodes

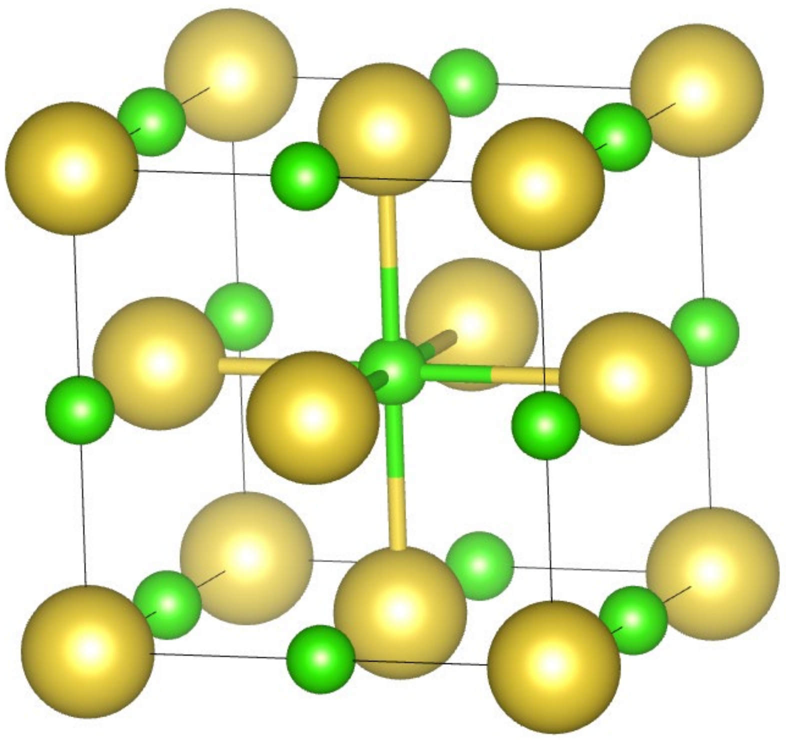
$2Å < x < 3Å$



6 neighbors

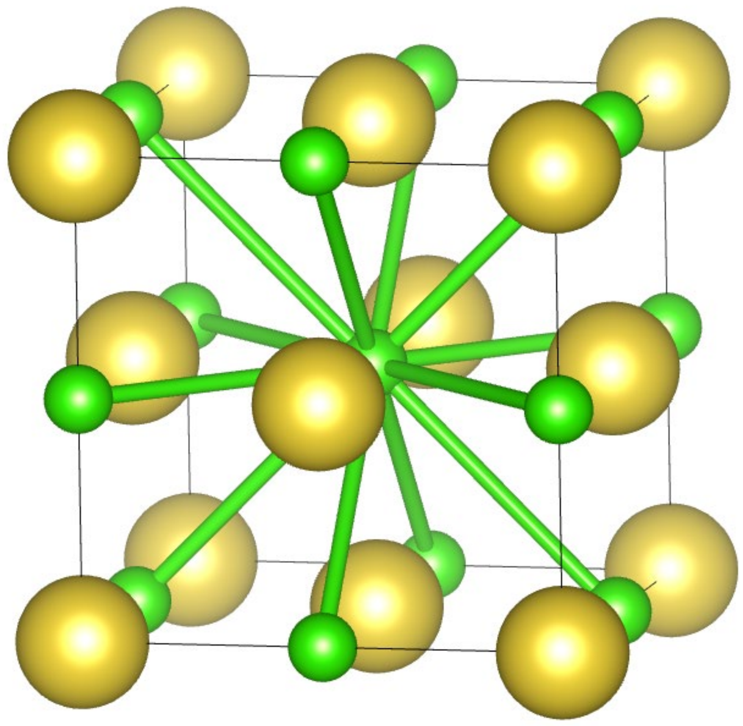
Current graph networks find neighbors within some volume, and then fully connect nodes

$2Å < x < 3Å$



6 neighbors

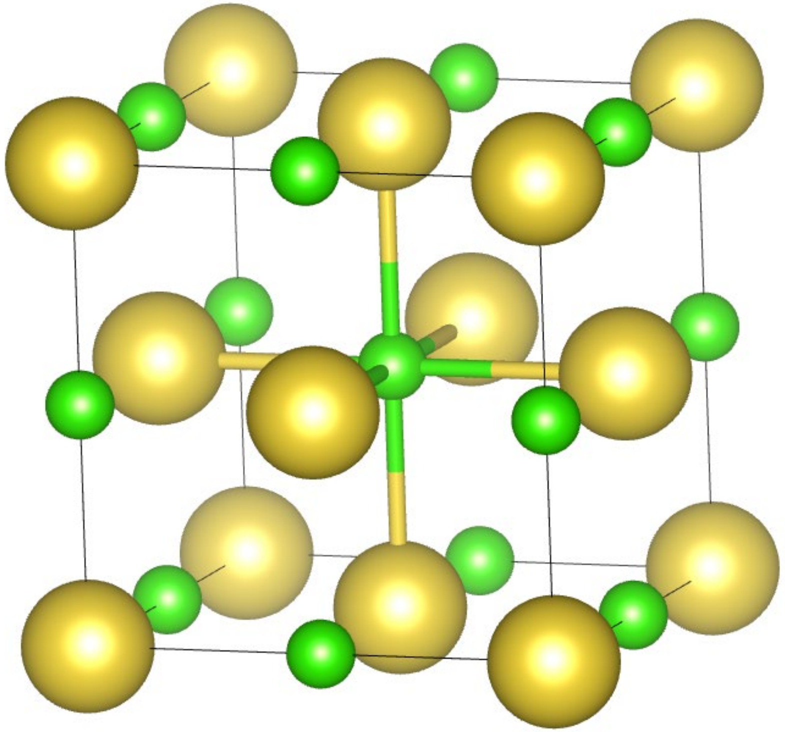
$3Å < x < 4Å$



12 neighbors

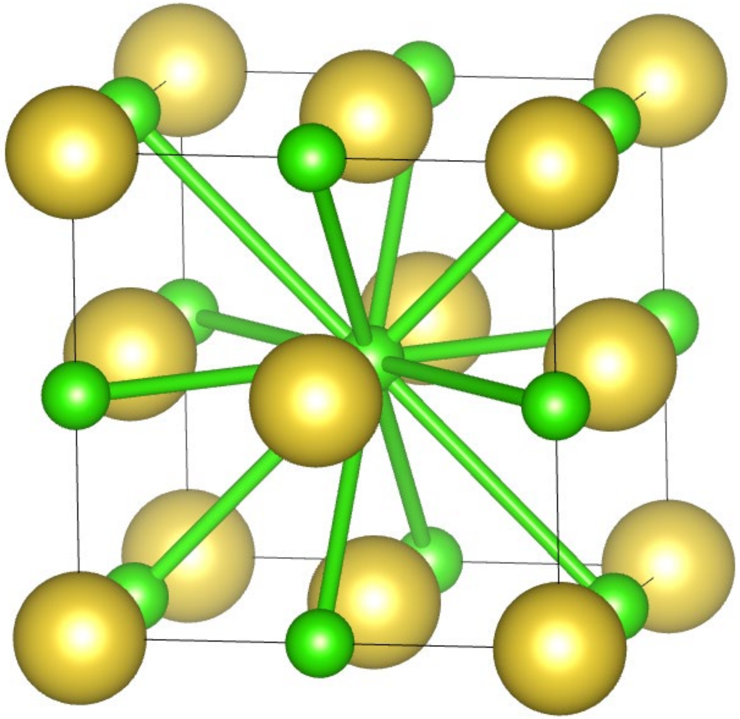
Current graph networks find neighbors within some volume, and then fully connect nodes

$2\text{Å} < x < 3\text{Å}$



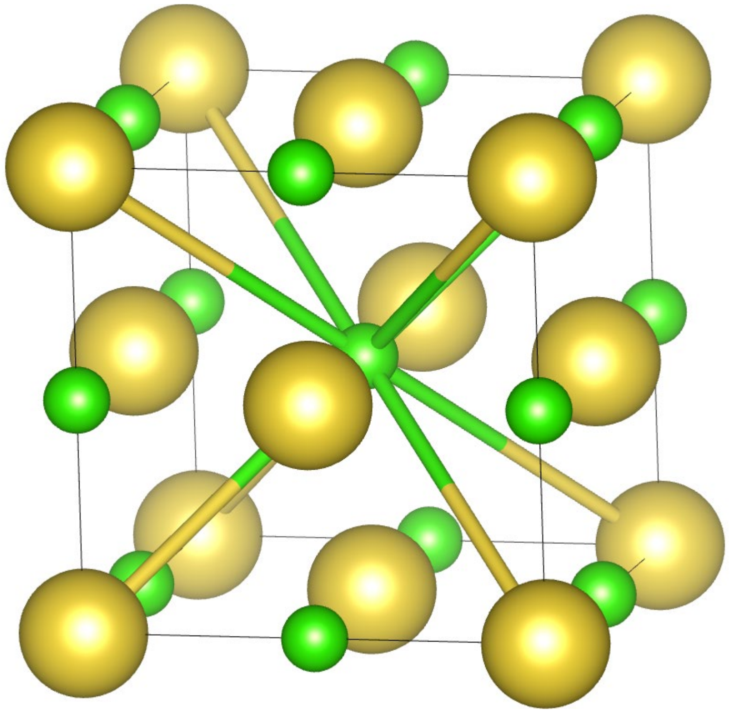
6 neighbors

$3\text{Å} < x < 4\text{Å}$



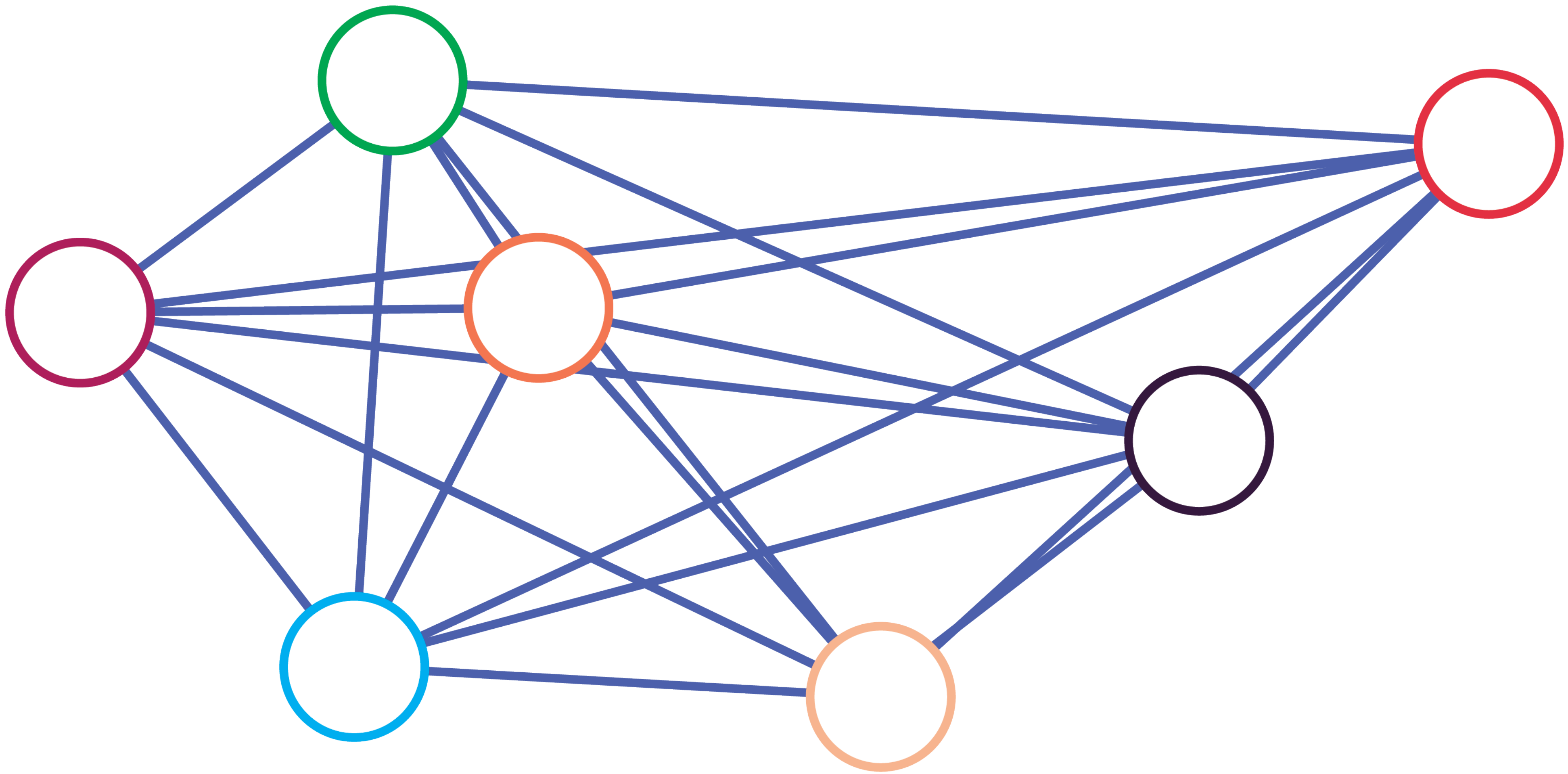
12 neighbors

$4\text{Å} < x < 5\text{Å}$

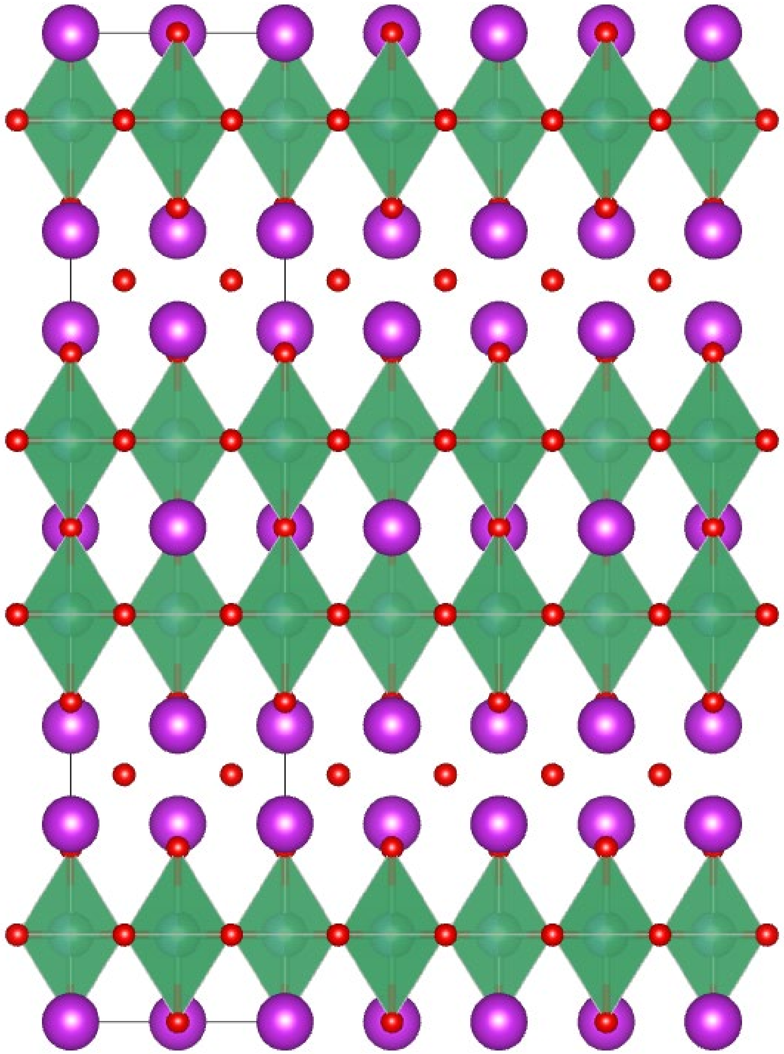


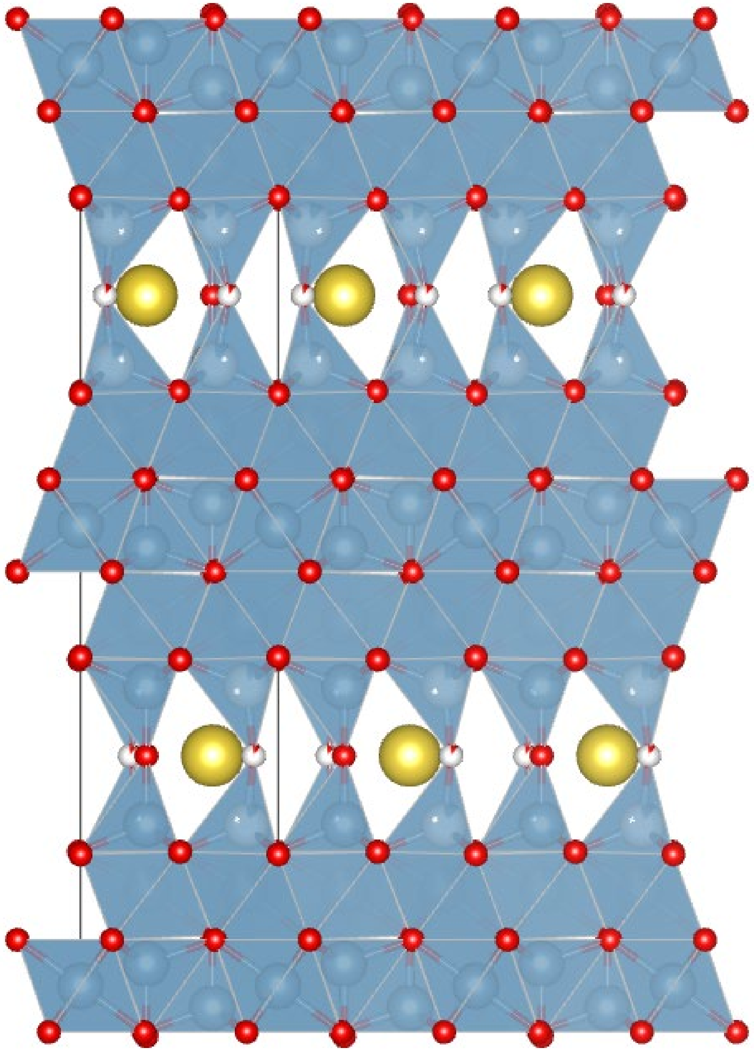
8 neighbors

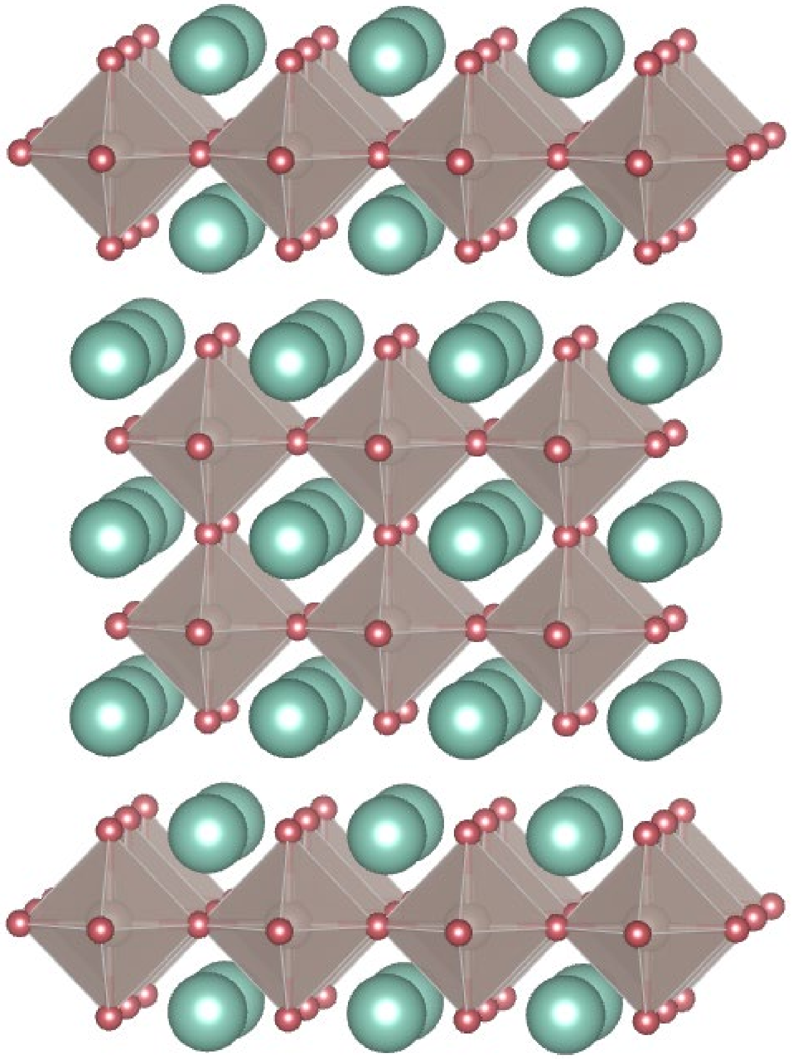
Current graph networks are fully connected



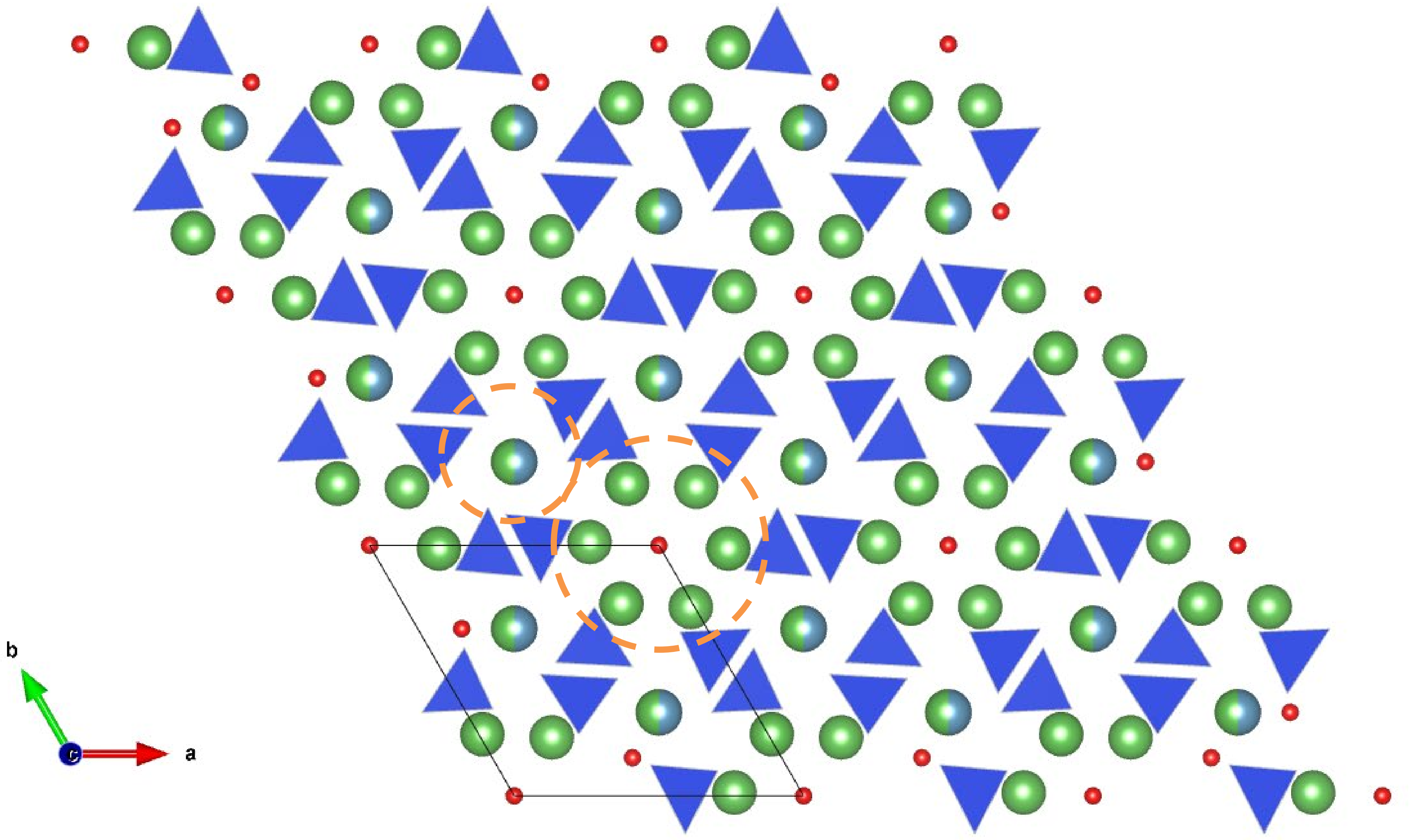
Many properties rely on specific interactions between atoms in a structure



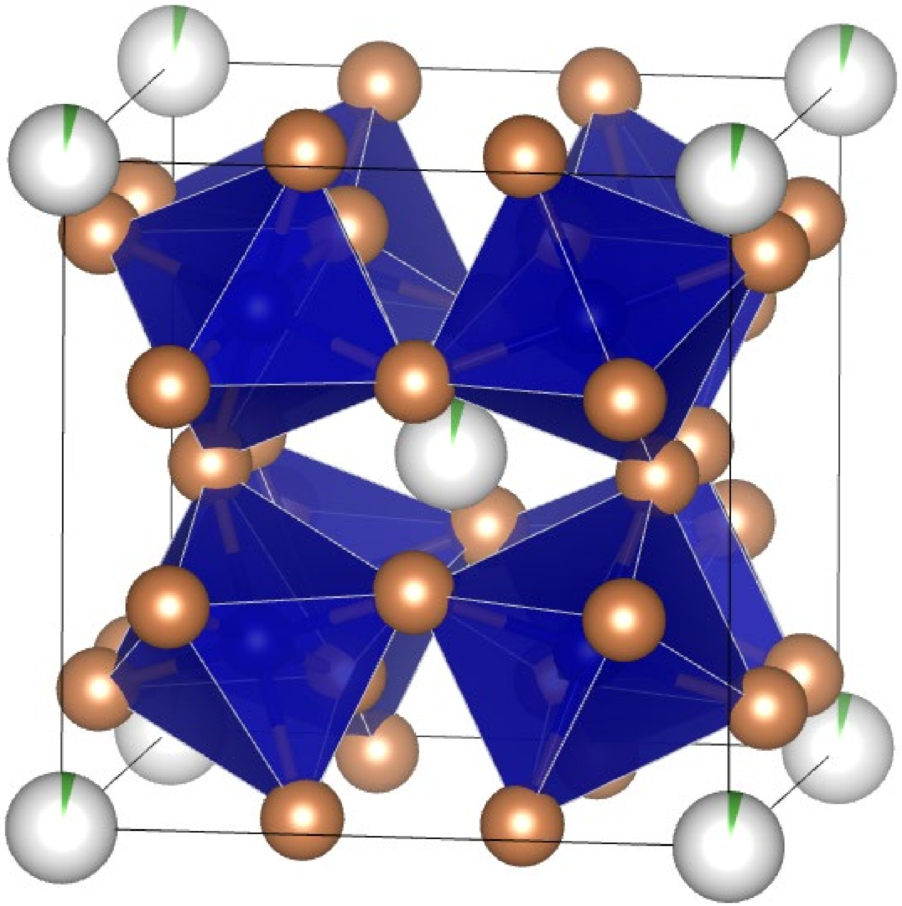


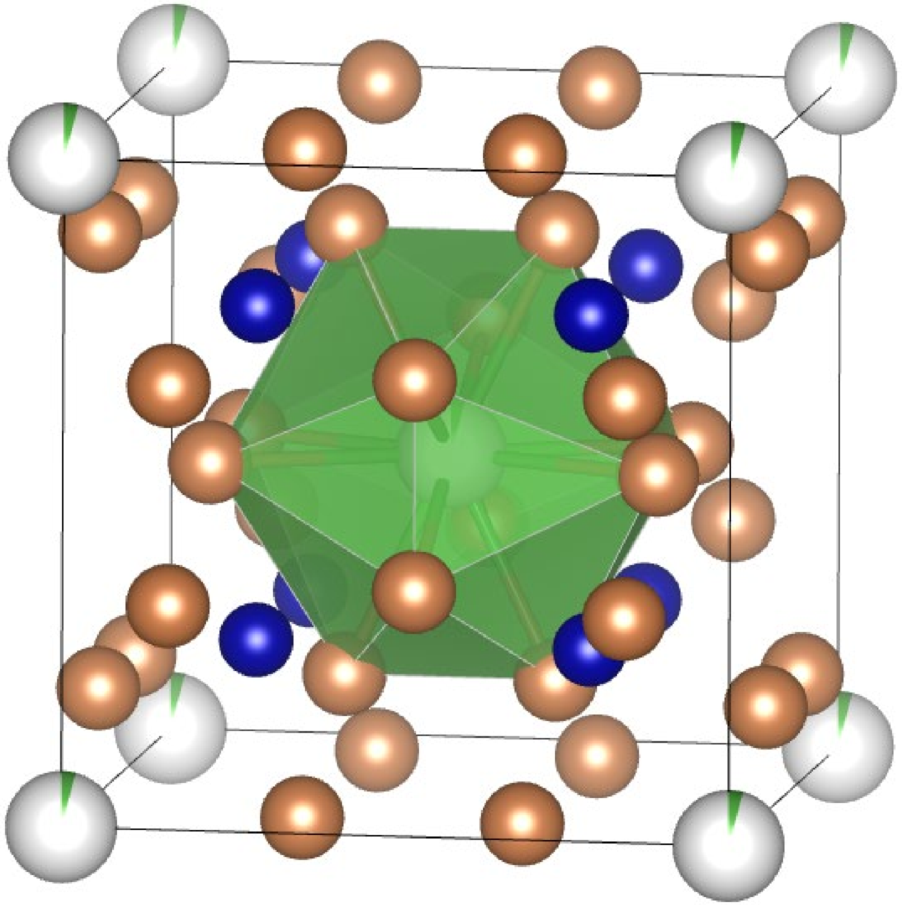


Many properties rely on specific interactions between atoms in a structure

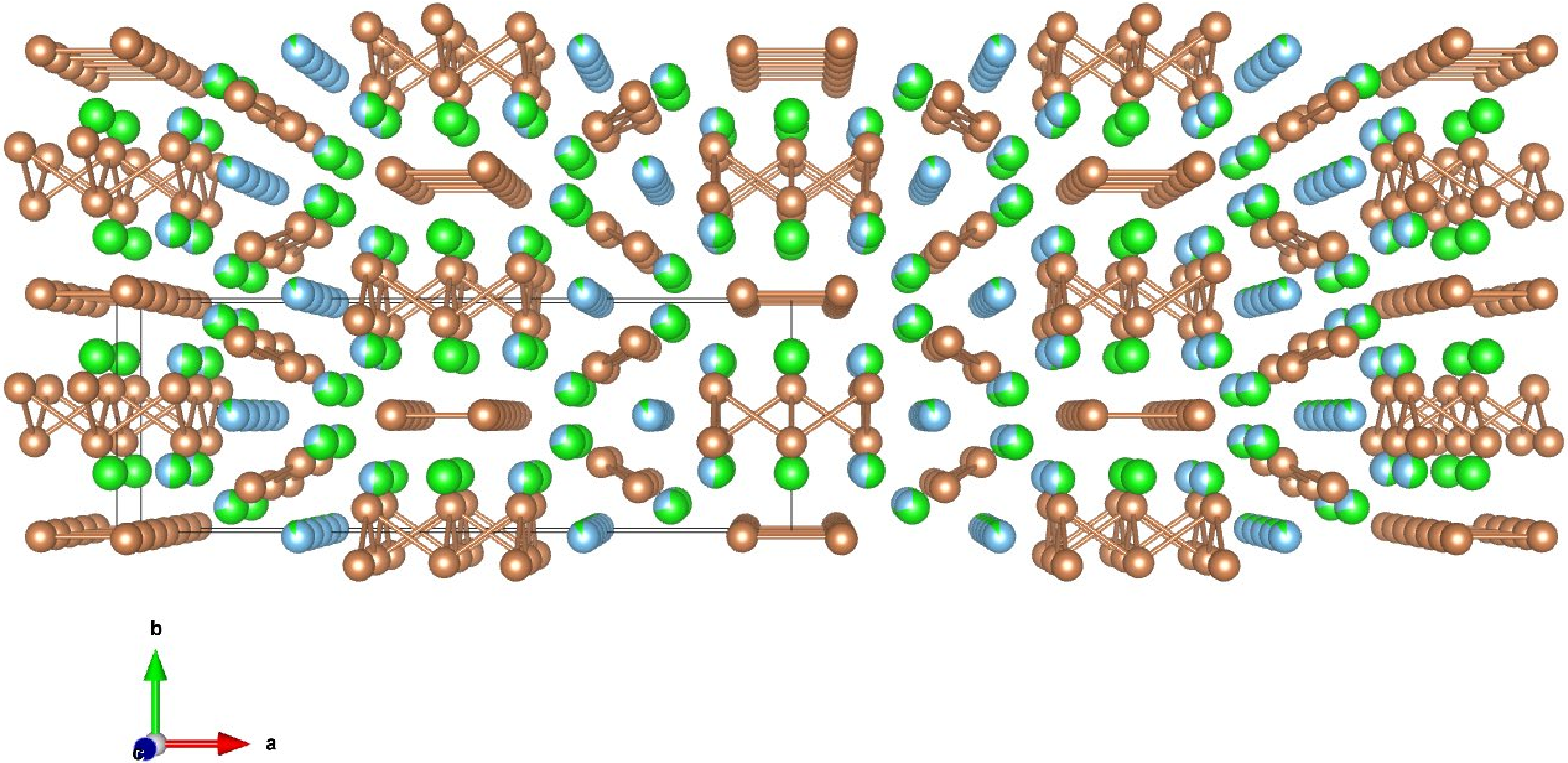


Many properties rely on specific interactions between atoms in a structure

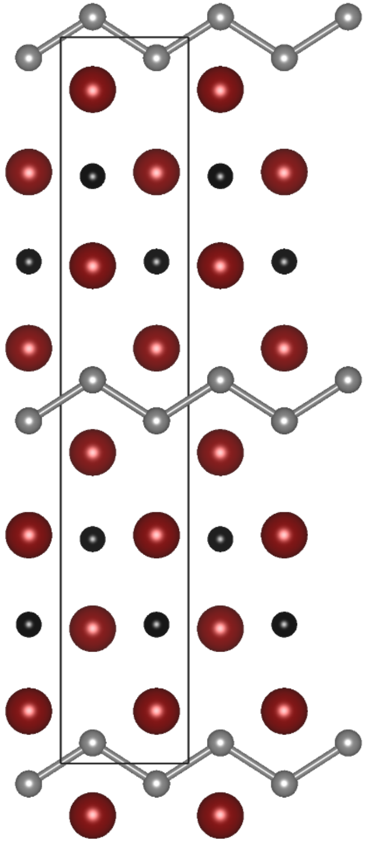


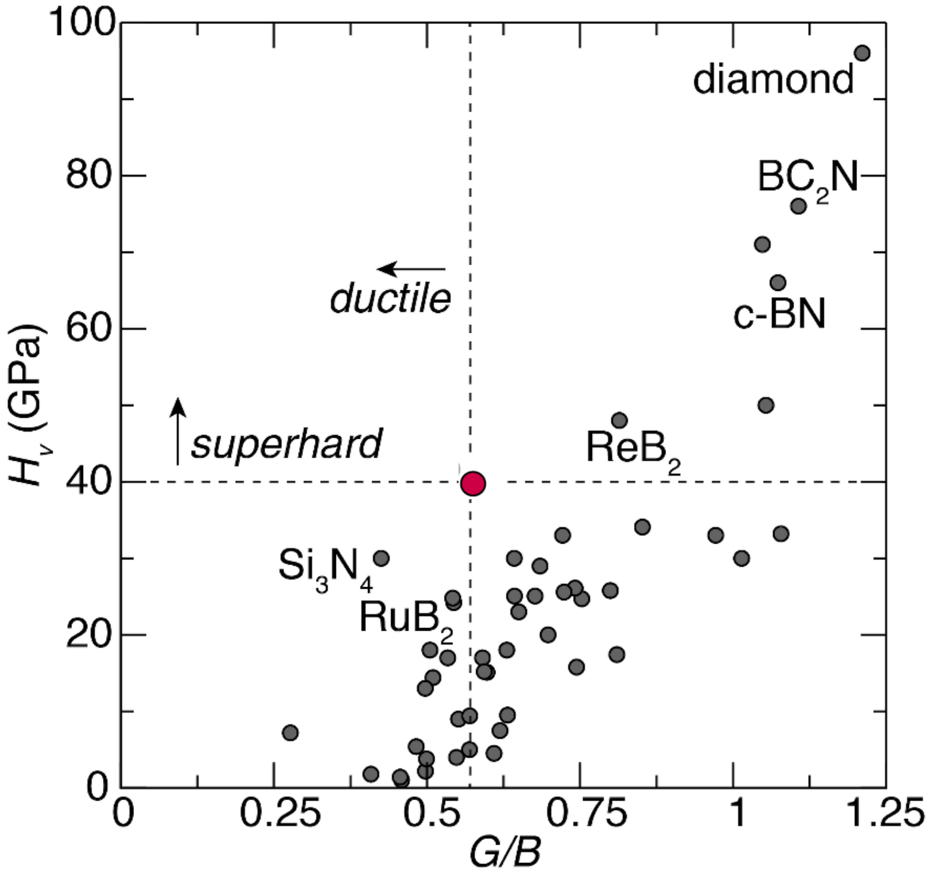


Many properties rely on specific interactions between atoms in a structure



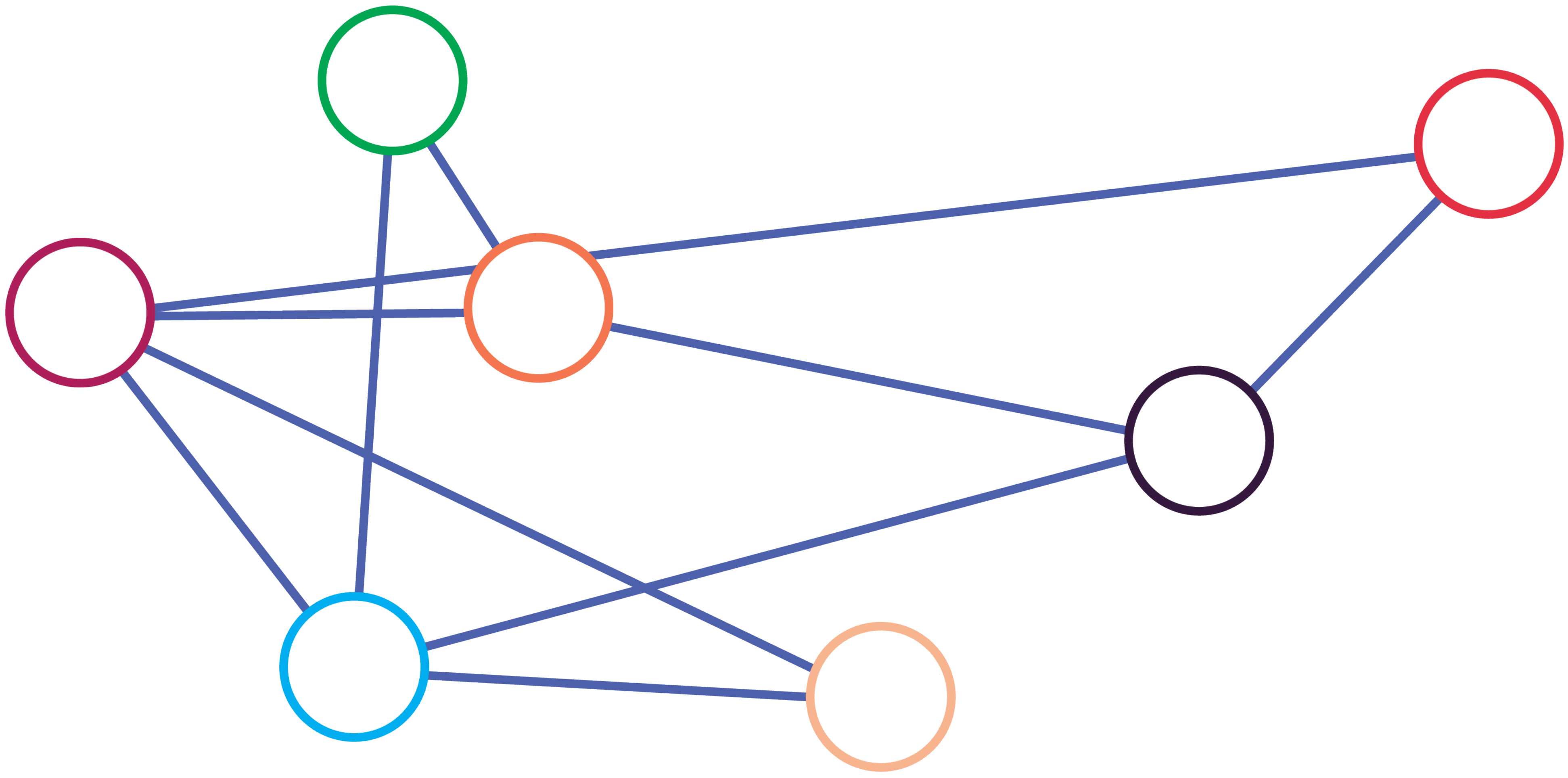
Many properties rely on specific interactions between atoms in a structure



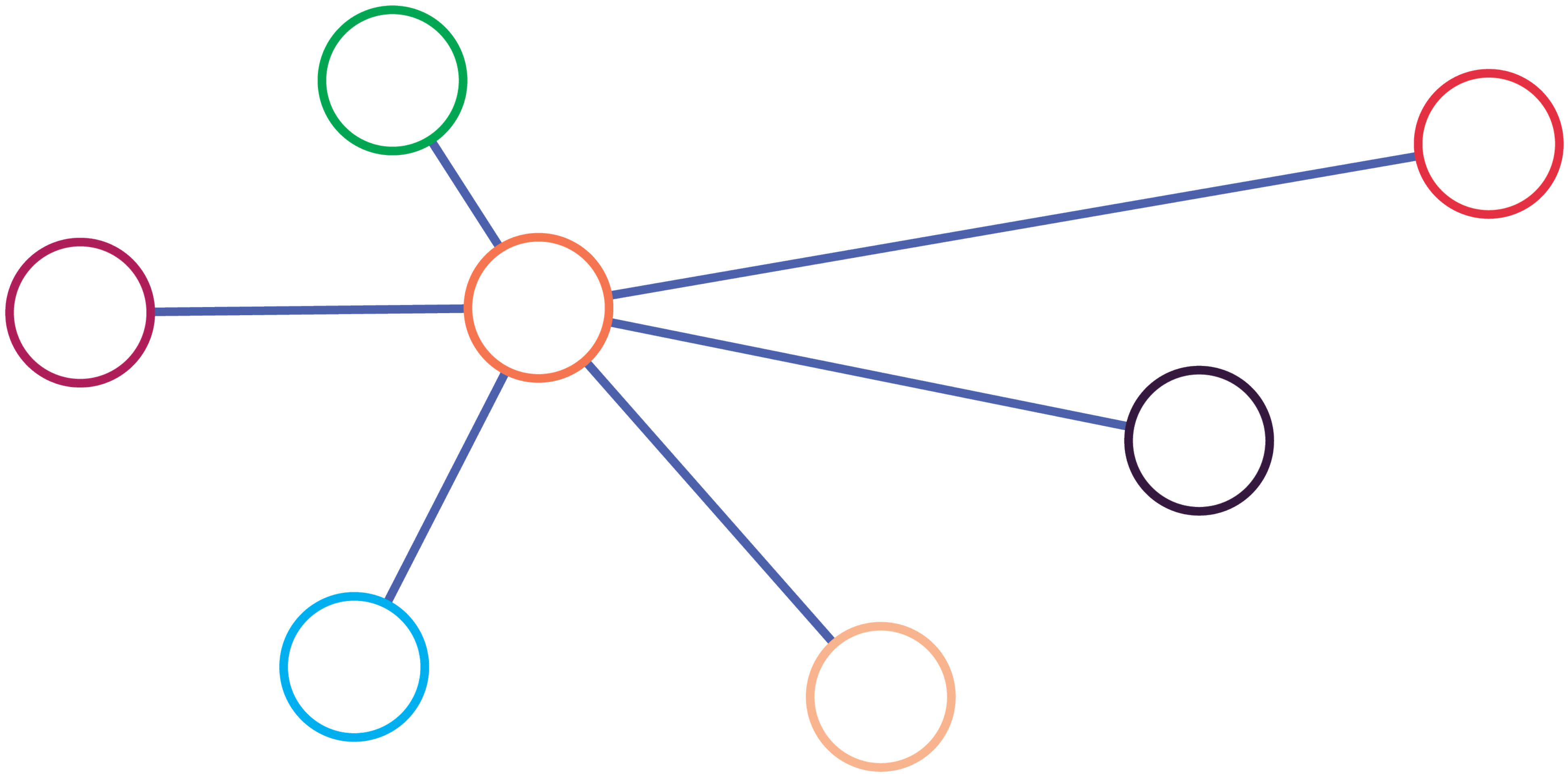


Mansouri et al. (2019) *Chem. Mater.*
Mansouri et al. (2018) *JACS*
Parry et al. (2019) *J Mat Chem A*

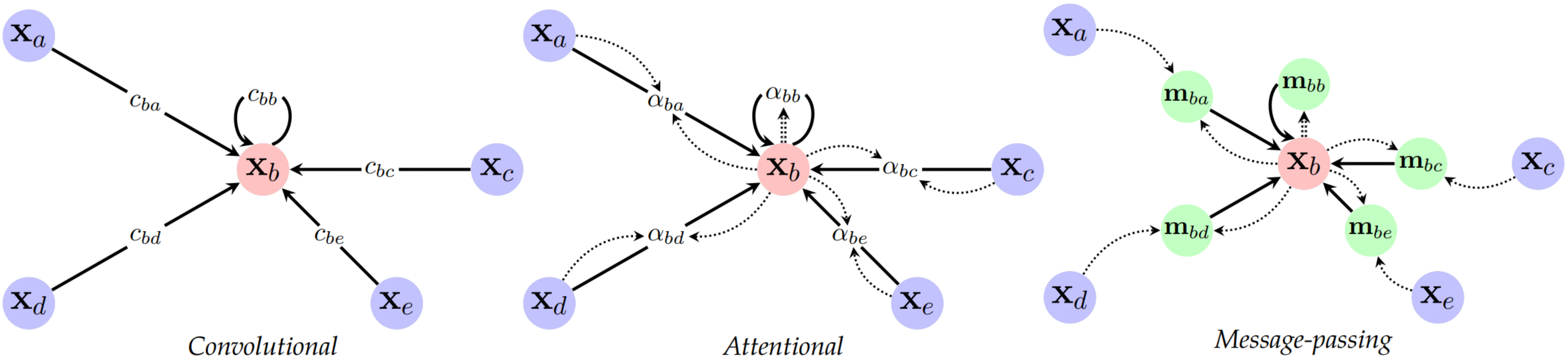
Sparsely connected graphs offer unique nodal information and faster speeds



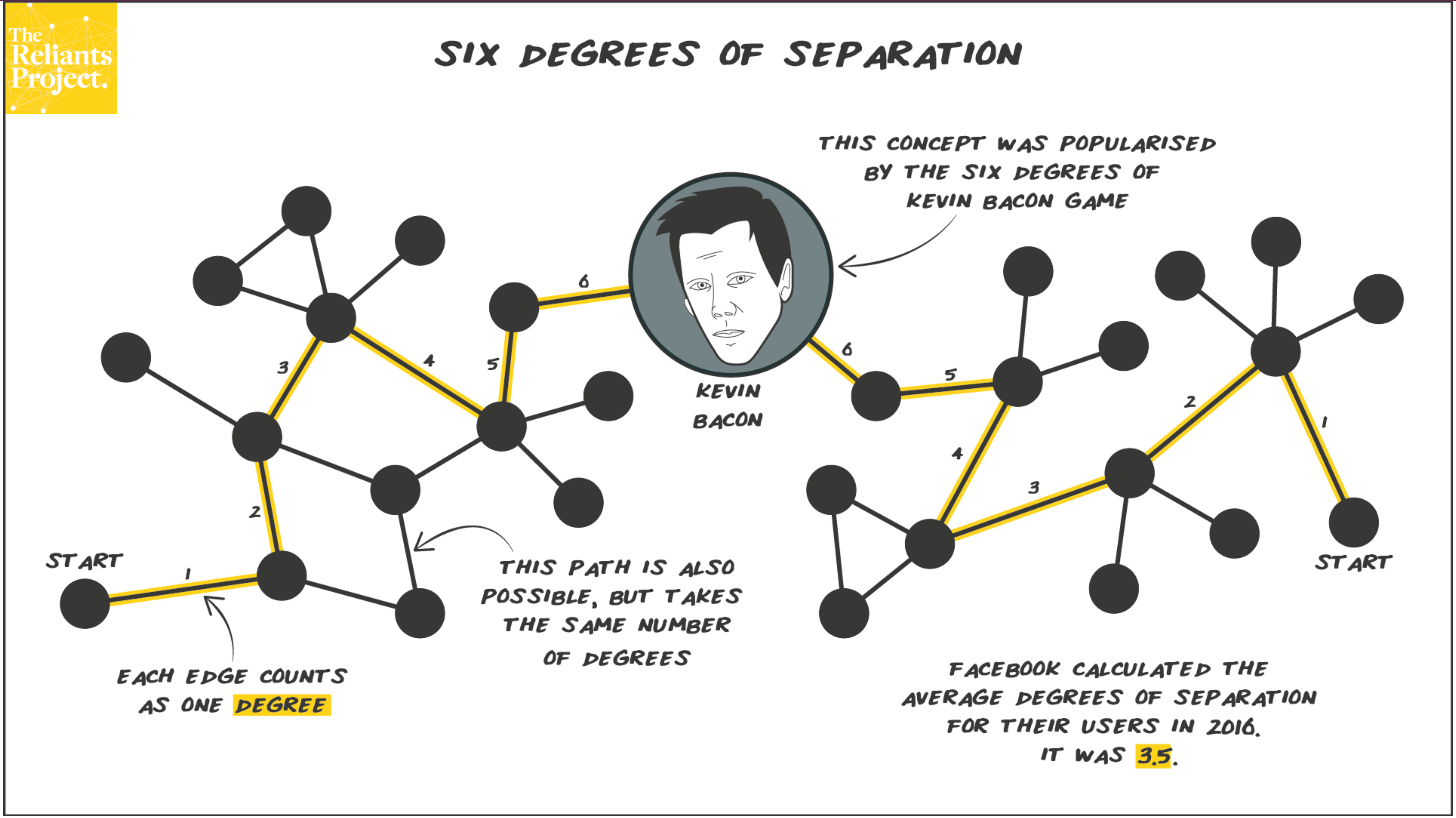
Connectivity impacts aggregation, attention, and message passing!



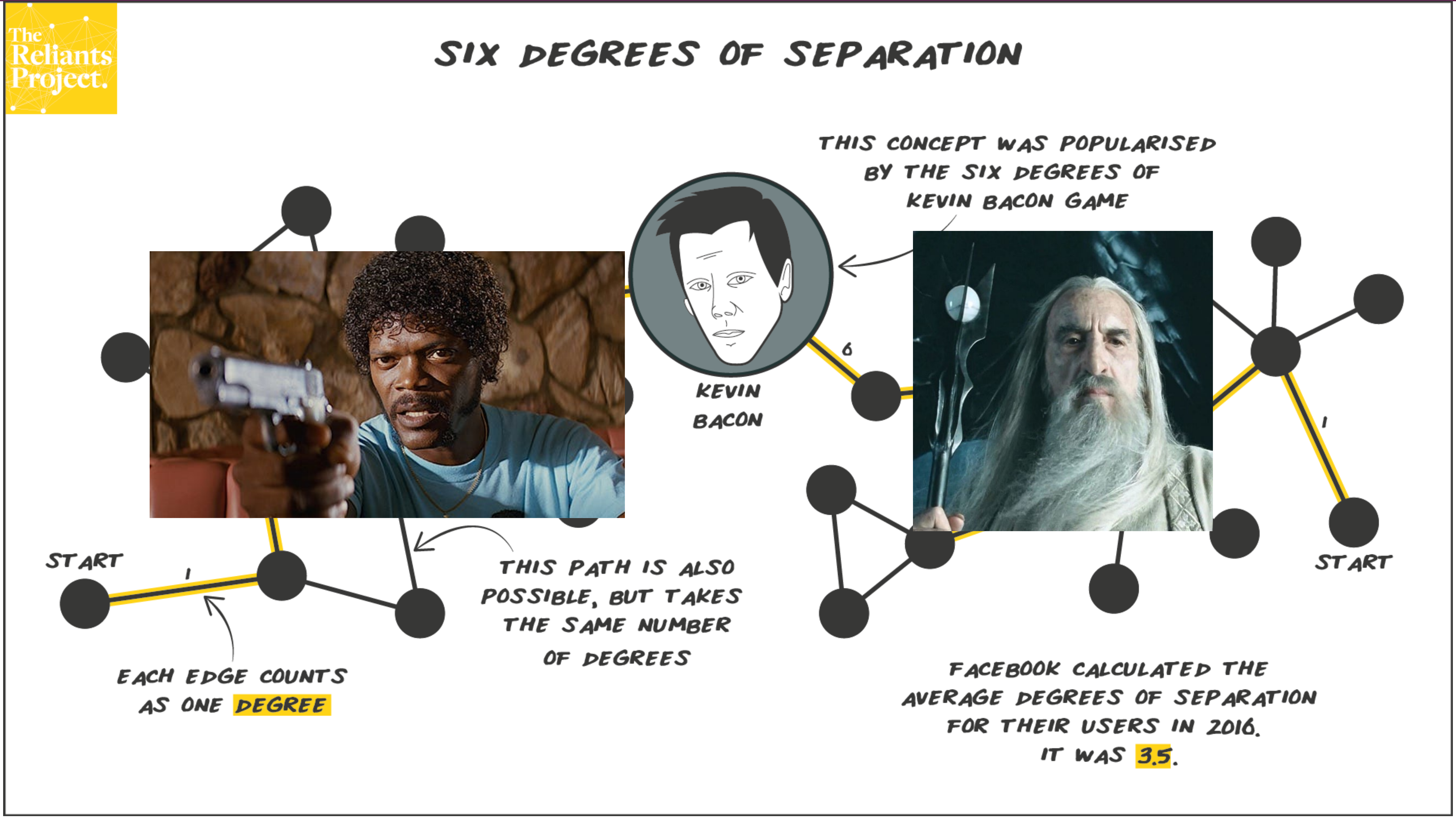
Connectivity impacts convolution, attention, and message passing!



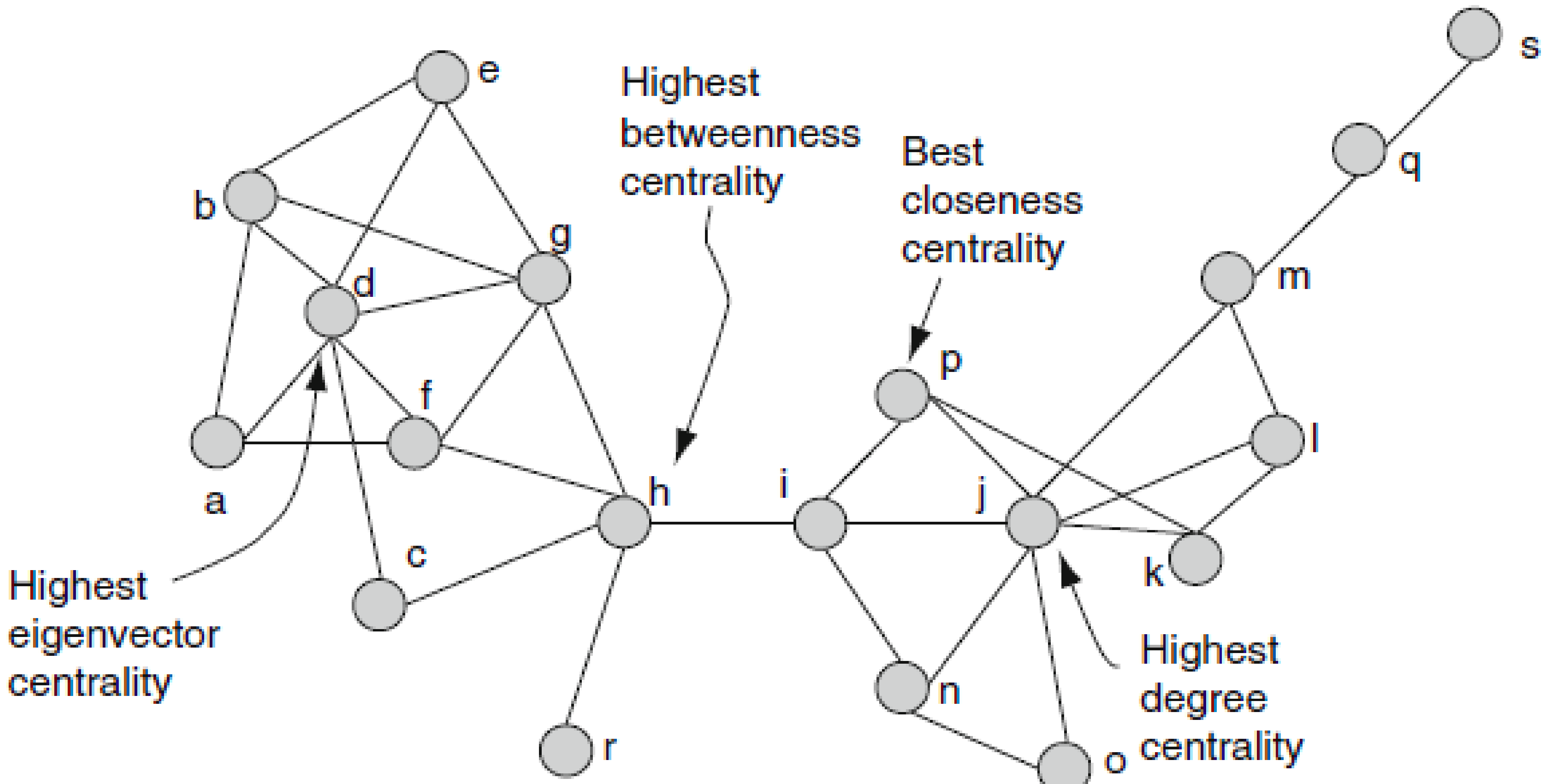
Many centralities can be considered for a given network



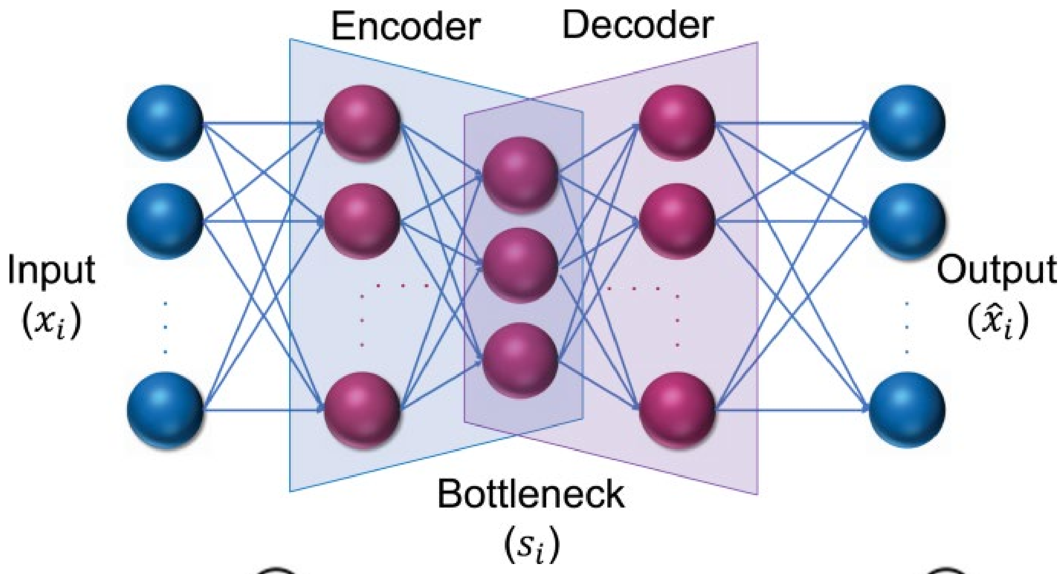
Many centralities can be considered for a given network



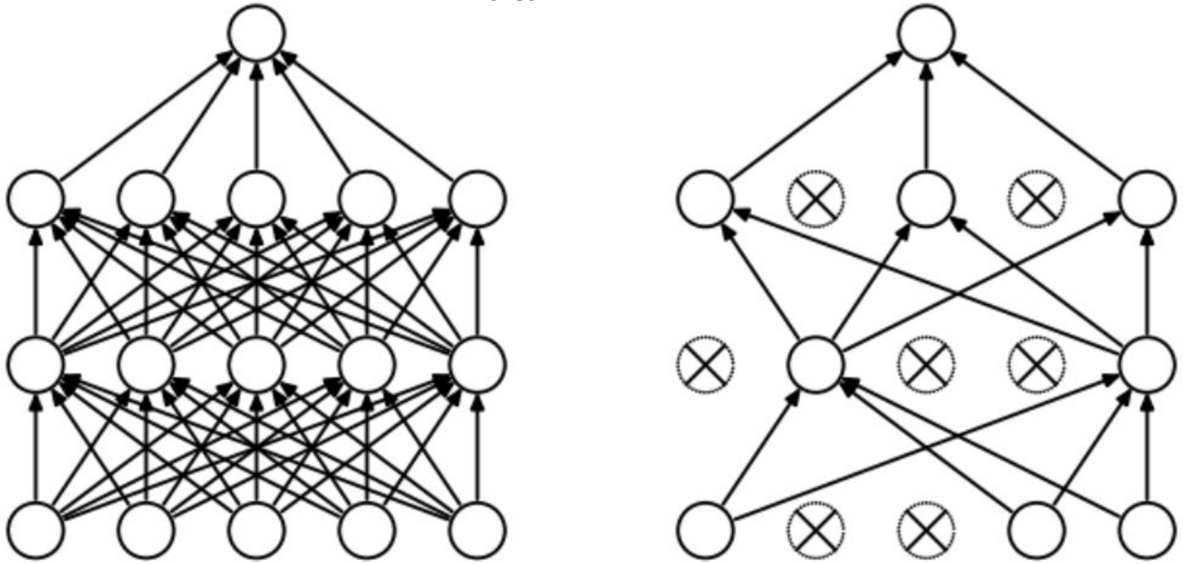
Many centralities can be considered for a given network



Is feature reduction sufficient? Or do we need something akin to dropout?



Feature reduction via auto-encoders



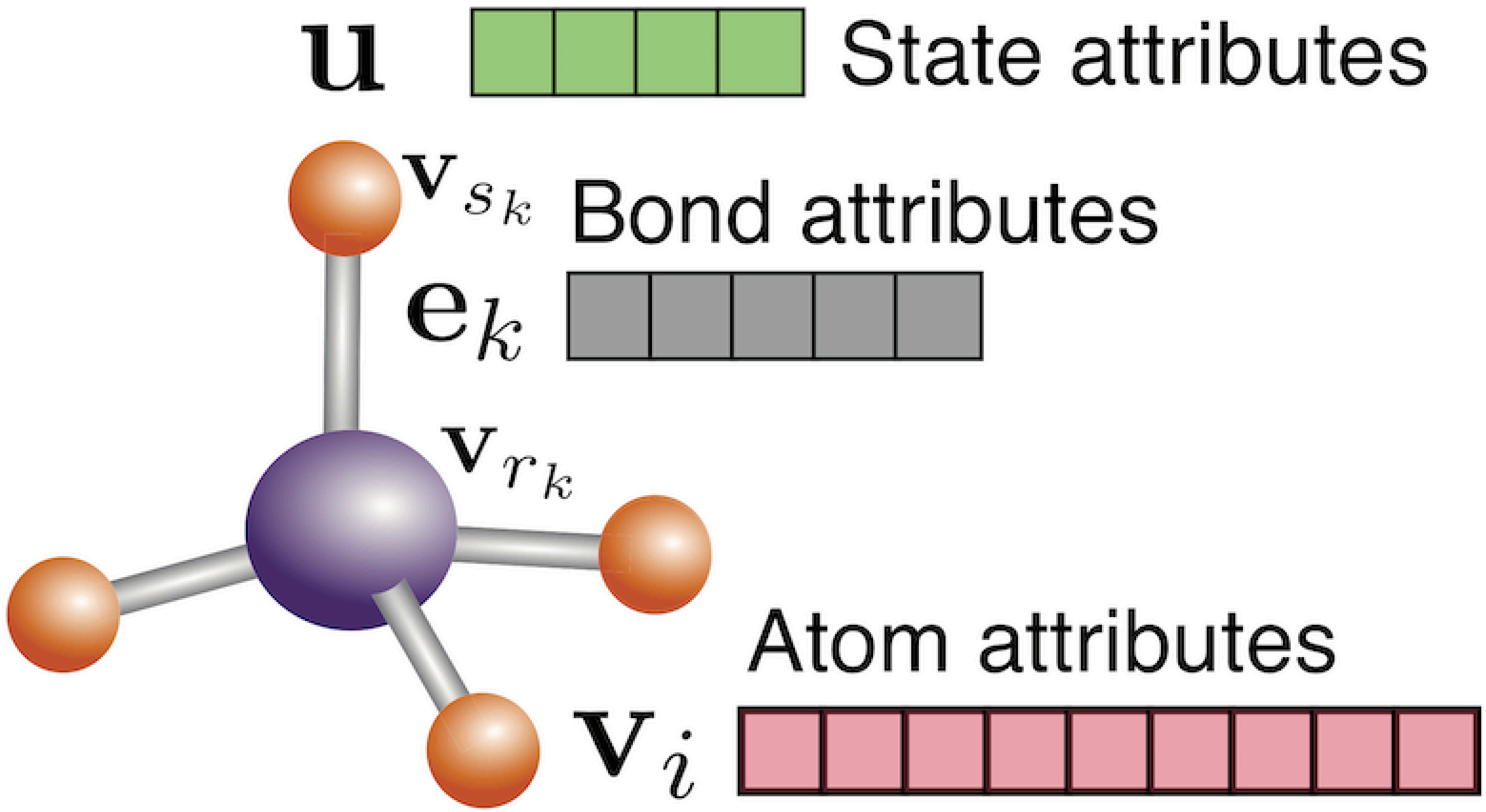
(a) Standard Neural Net

(b) After applying dropout.

Dilution (dropout) regularization

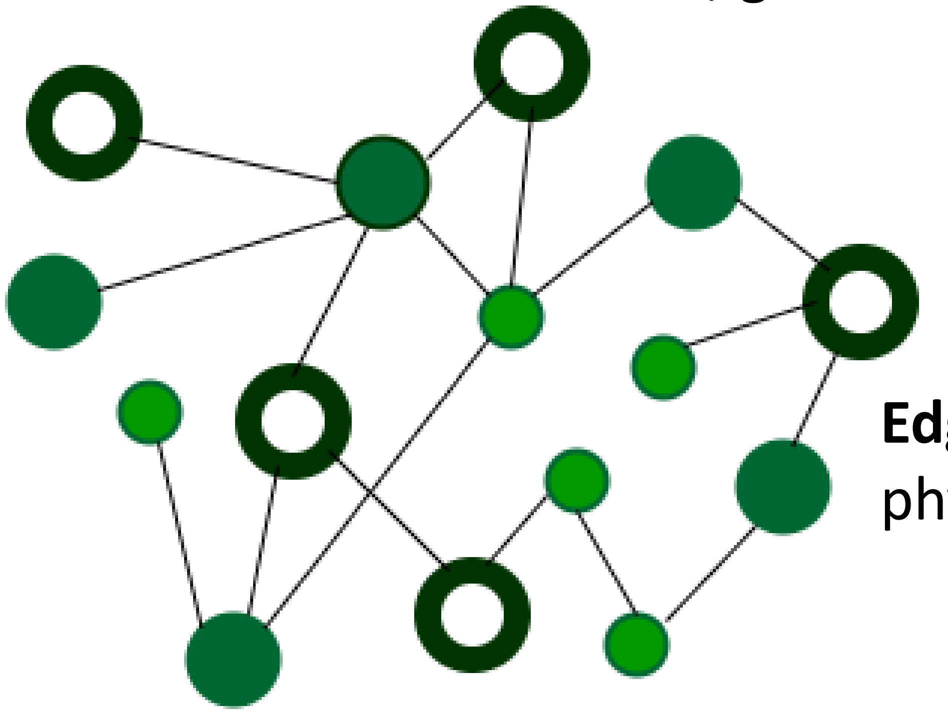
Features are becoming quite large for many GNNs

Features are becoming quite large for many GNNs



Is there a better, simpler graph representation?

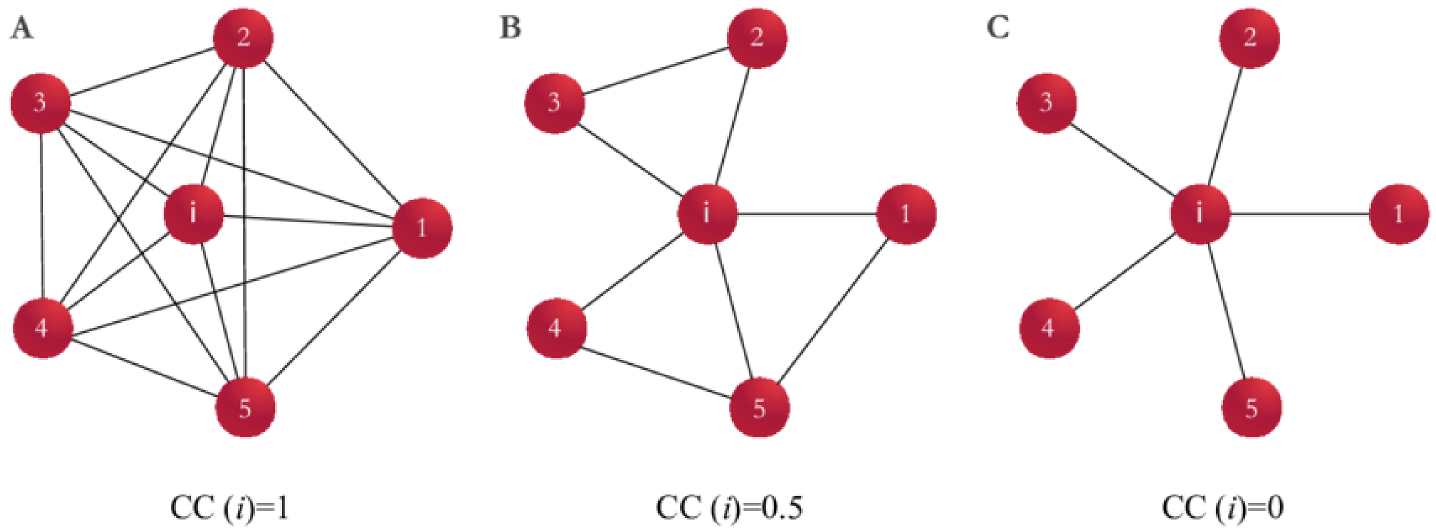
Nodes: atoms with limited features (quantum numbers, ground state energy)

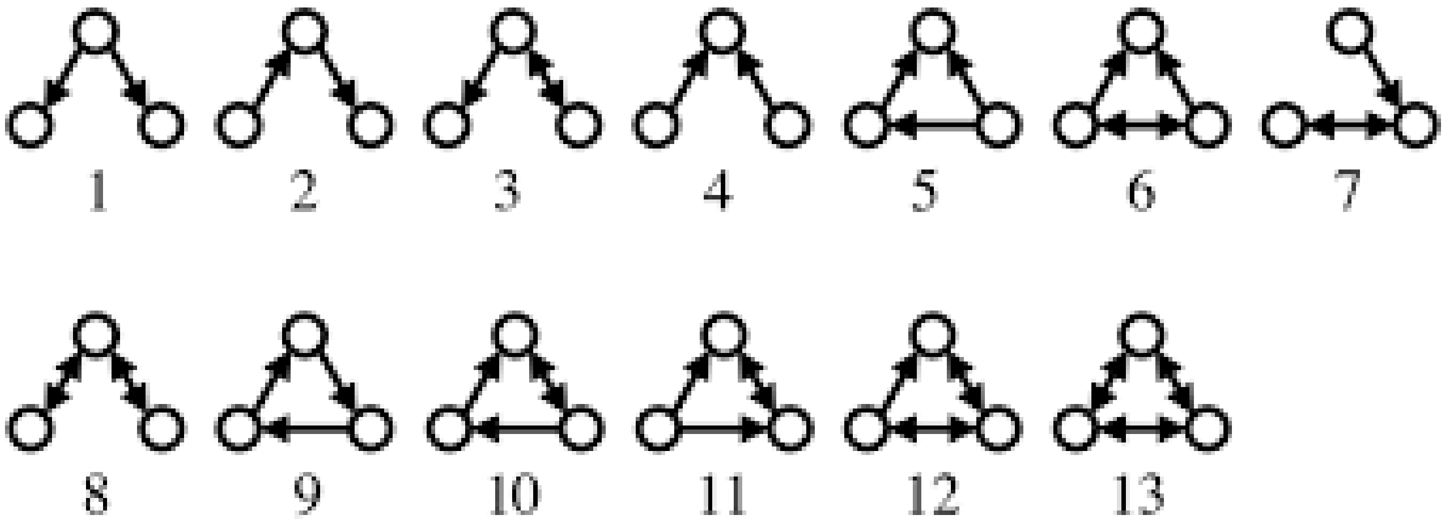


Edges: bonds, limited connectivity, based on physically meaningful parameters

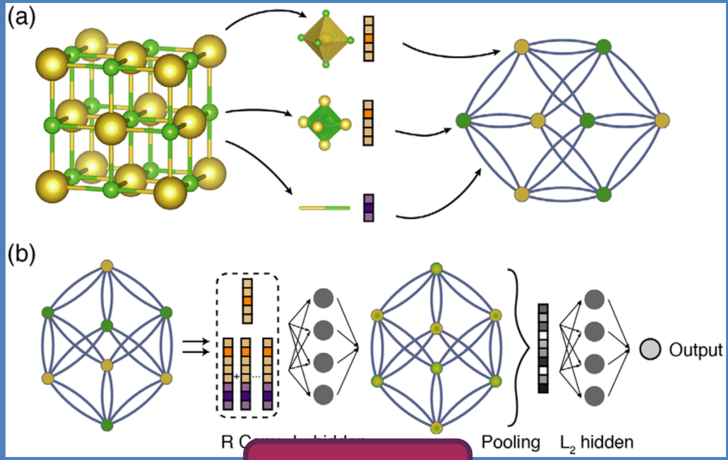
Graph: centrality, clustering coefficient, connectivity, motifs, symmetry

Graph level attributes can focus on unique aspects of sparse connections

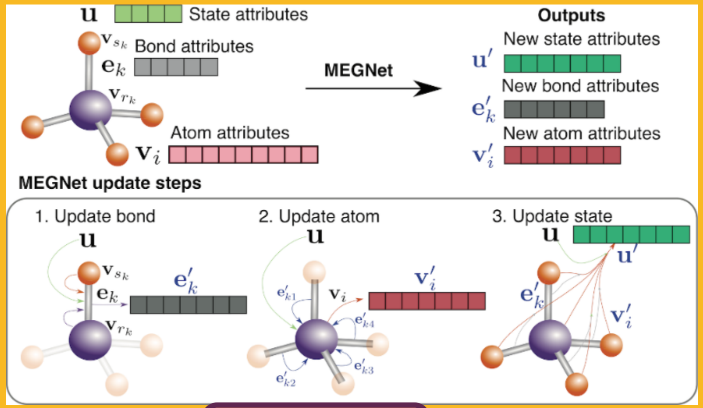




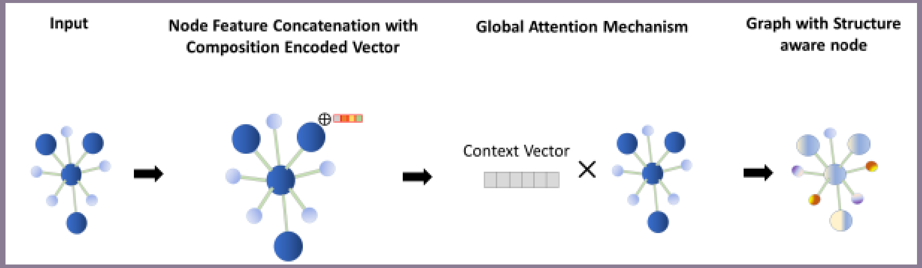
Case studies on specific tools in a future video



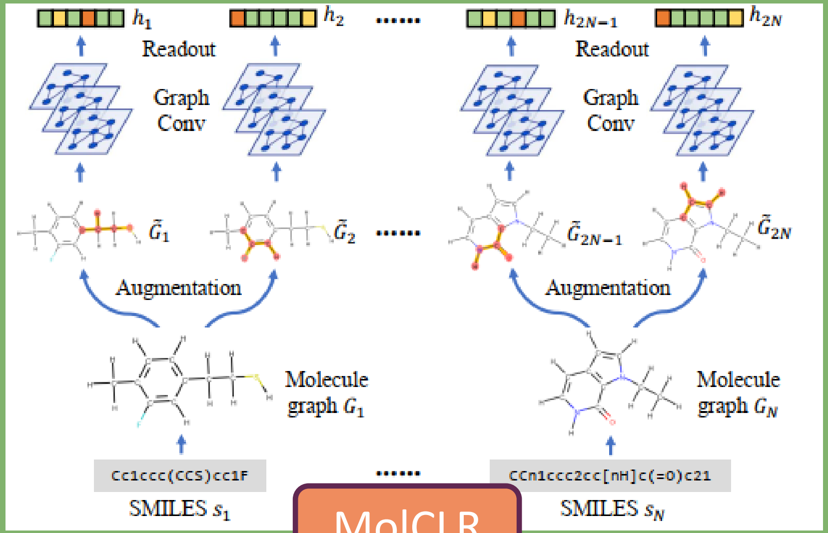
CGCNN



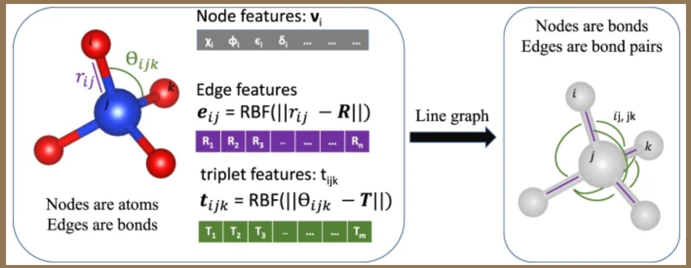
MEGNet



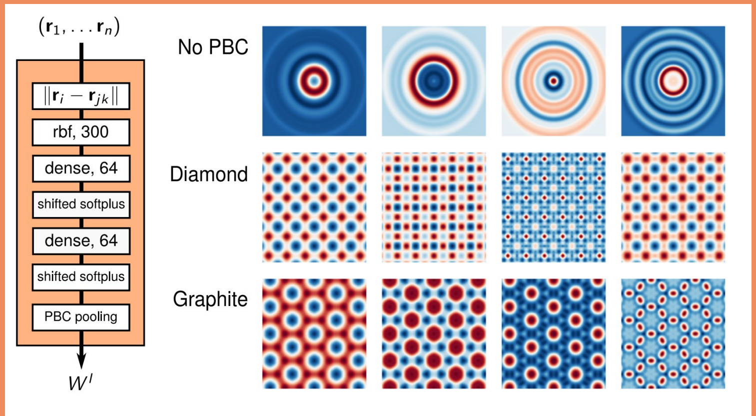
GATGNN



MolCLR



ALIGNN



SchNet


2 pt statistics