

Graph Convolutional Networks

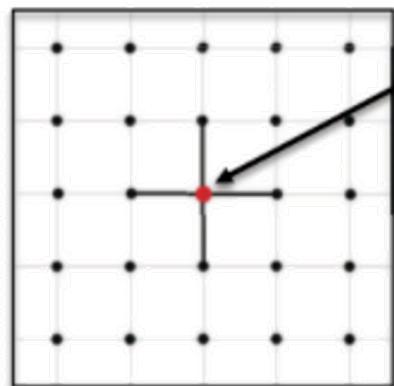
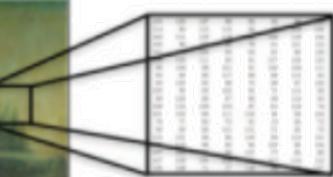
Krishnendu Mukherjee

Advisor: Dr. Yamil J. Colón
ML Reading Group
August 2nd, 2021

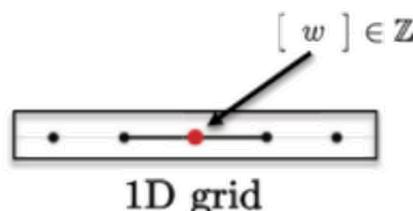


Graph Domains

Data Domain

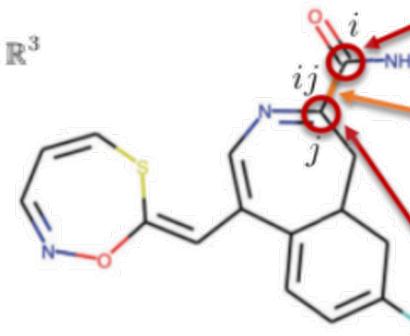


2D grid



1D grid

Graph Domain



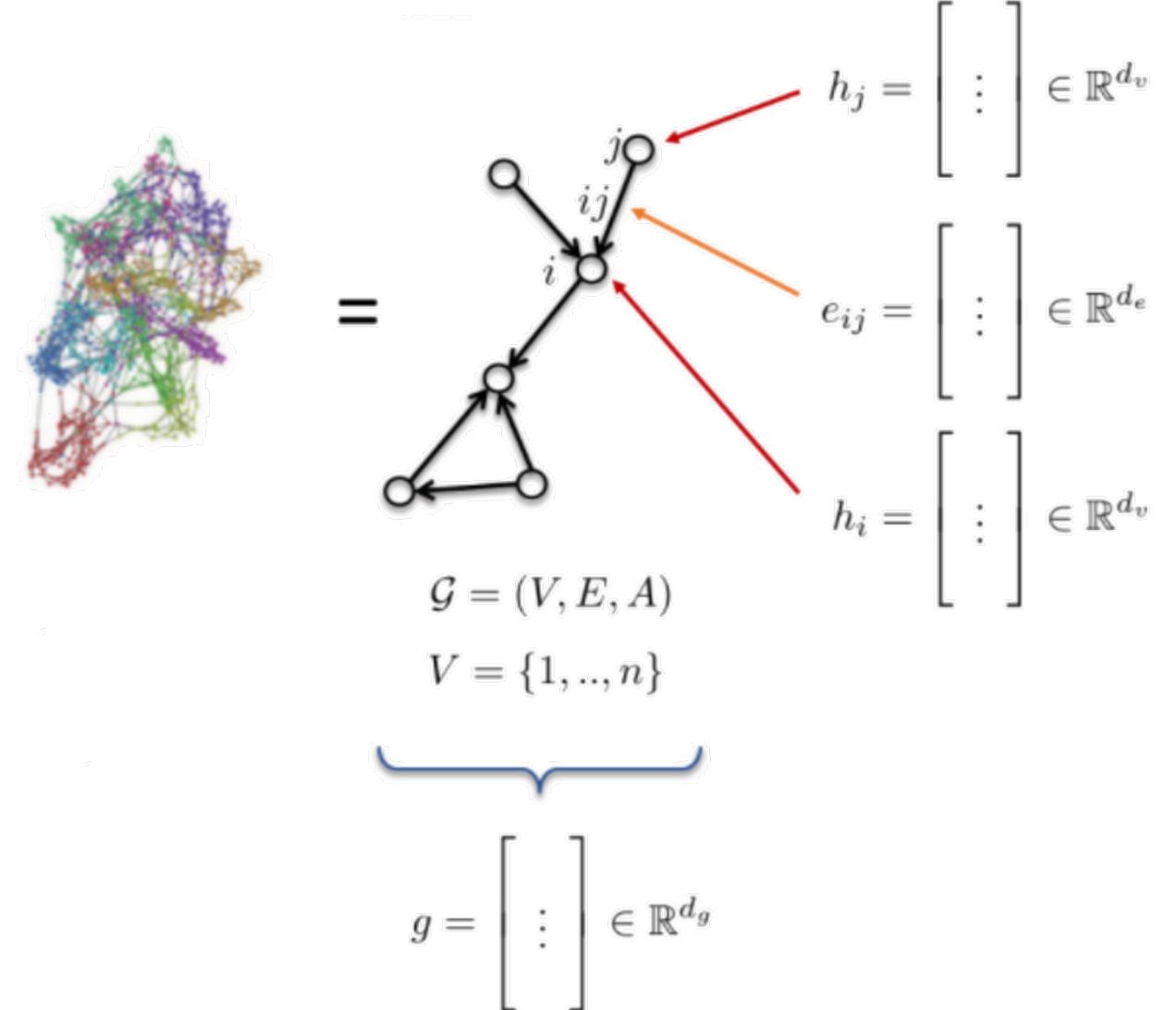
Quantum Chemistry
(novel molecules for drugs
and materials)

$$\text{Atom}_i = \begin{bmatrix} \text{type} \\ \text{coordinates} \\ \text{charge} \end{bmatrix}_i \in \mathbb{R}^{d_v}$$
$$\text{Bond}_{ij} = \begin{cases} 1 & \text{if } ij \text{ bond} \\ 0 & \text{otherwise} \end{cases}$$
$$\text{Atom}_j = \begin{bmatrix} \text{type} \\ \text{coordinates} \\ \text{charge} \end{bmatrix}_j \in \mathbb{R}^{d_v}$$
$$\text{Bond}_{ij} = \begin{bmatrix} \text{type} \\ \text{energy} \end{bmatrix}_{ij} \in \mathbb{R}^{d_e}$$

D RICCARDO ROMEO, ANDREA D'ONOFRIO, AND TIZIANO ROMEO
Although we use “interventions” in a general way for “why,” “inter” uses the word in a more limited sense. By “intervention” Judea Pearl means “for what purposes?” Judea Pearl has many ways to “Why did this happen?”
“why” — “From what causes?” In the past and “For what purposes?”
“why” — “From what causes?” In the past and “For what purposes?”
“Intervention” clearly emphasizes the latter sense, which is why “
Intervention” and its partner “Intervention” reflect the basic tendency of
English speakers to use “intervention” for “intervention.” However, no representative
meaning, such as cause and effect.
“INTERVENTION” IS NOT NEEDED TO USE IT. A COUPLE OF OTHER WORDS
WOULD DO JUST AS WELL.
“Intervention” is indeed hard, although the latter is, as many things, merely a
matter of definition. In fact, the term “intervention” has been used in many different
environments, those phrases have served generations with complete
fidelity.
“What is a number?” is the less specific of the two phrases, and also the
less common. Just like many other words, it is different from the point of view of a linguist.
“What is a theorem?” is a good example. From a mathematician’s point of view, “theorem”

Graph Definitions

- Graph G is defined as:
 - Vertices V
 - Edges E
 - Adjacency Matrix A
- Graph features are:
 - Node features: h_i, h_j (atom type)
 - Edge features: e_{ij}
 - Graph features: g (molecular energy)



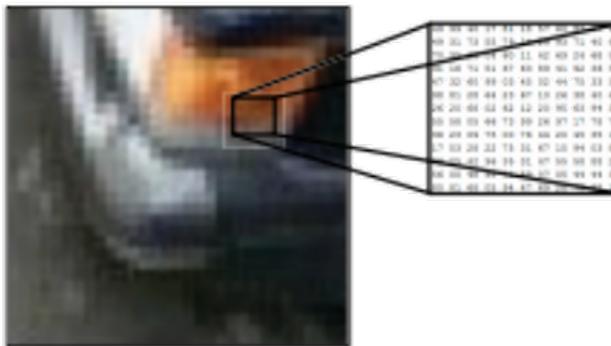
What is a Convolution

$$h^l \times w^l = h^{l+1}$$

Features at l^{th}
layer

Pattern/Kernel
(learned by
backpropagation)

Convolved
feature at $(l+1)^{\text{th}}$
layer



h^ℓ

Image/Hidden
features

*



=

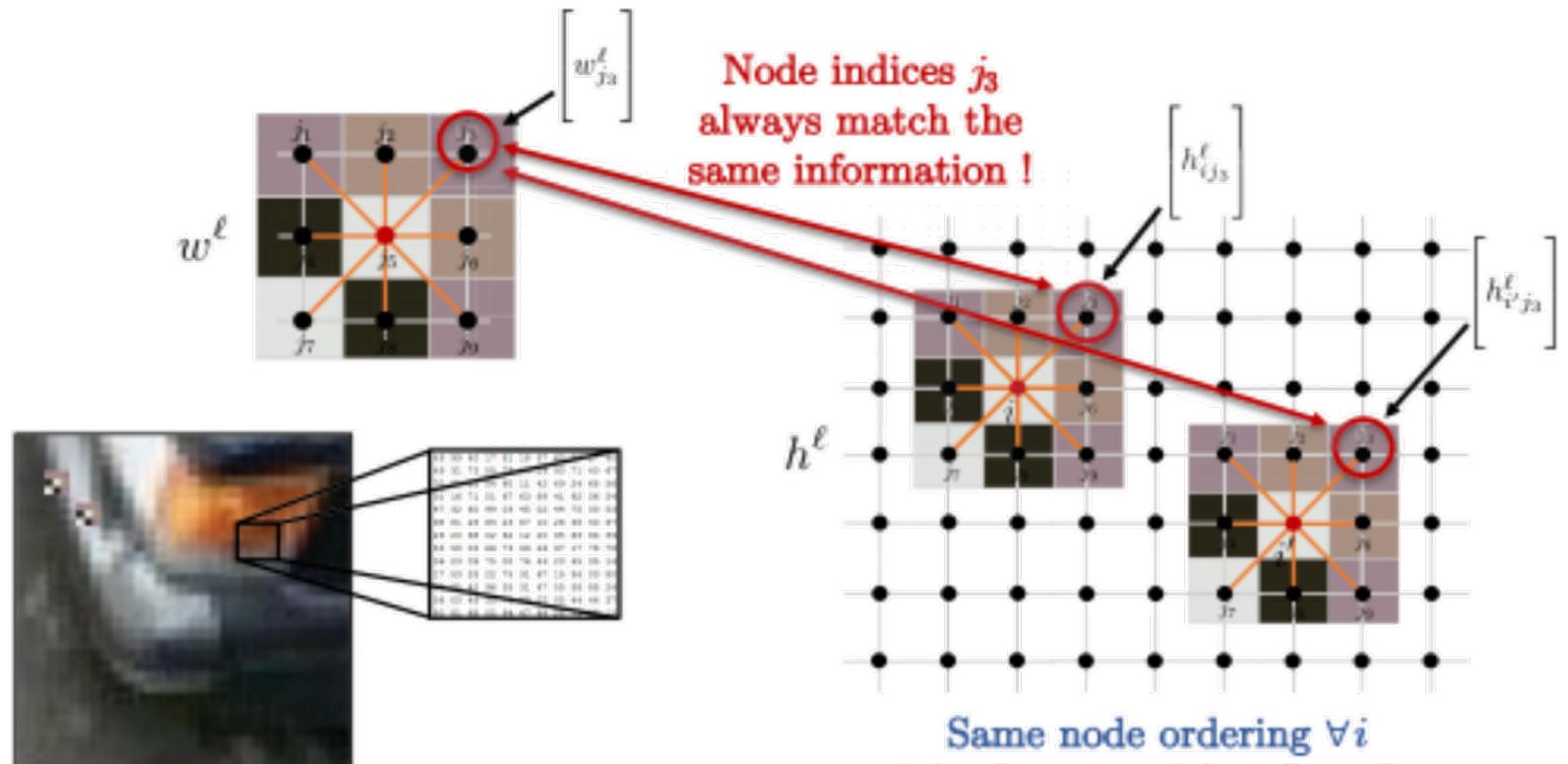


$h^{\ell+1}$

Pattern/kernel
(learned by
backpropagation)

Convolution as template matching

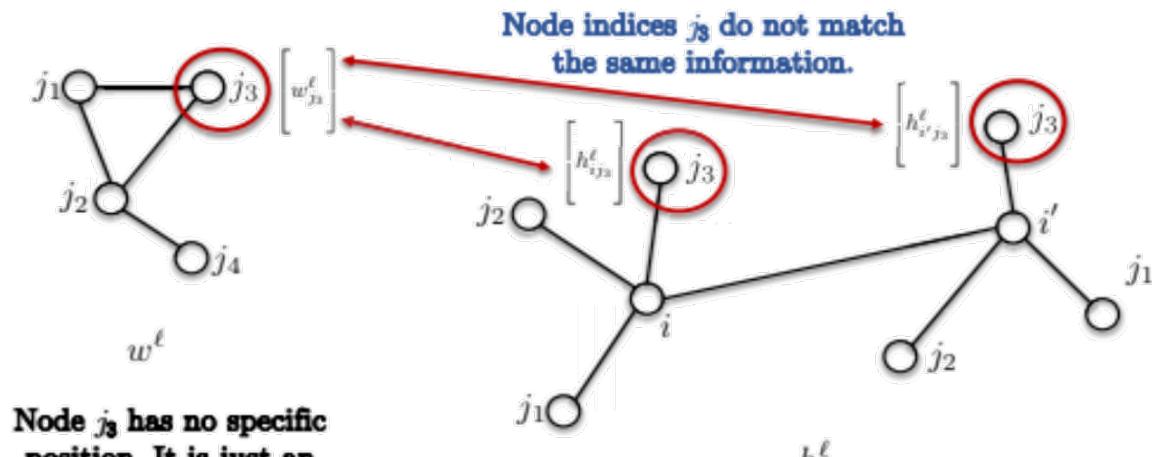
- The template/kernel w_i is overlaid (as defined by the strides) across the image from top-left to bottom right
- Different kernel convolutions are done to represent an image in terms of kernels/templates
- Kernels might become combination of kernels as we proceed just like ANNs



Processors

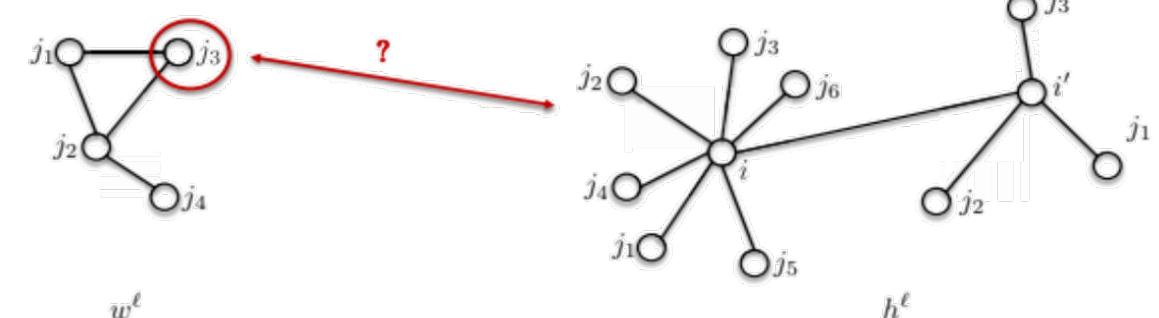
Problems with extending template matching to Graphs

- No ordering of nodes
- Neighborhood sizes are different



Node j_3 has no specific position. It is just an (arbitrary) index.

No node ordering on graphs :
The correspondence of nodes is lost on graphs.
There is no up, down, right and left on graphs.



Xavier Bresson

Convolution theorem to graphs

$$\mathcal{F}(w * h) = \mathcal{F}(w) \odot \mathcal{F}(h)$$

- Use convolution theorem for spectral ConvNets

Convolution
Theorem

$$\begin{aligned}
 w * h &= \underbrace{\mathcal{F}^{-1}}_{\Phi} (\underbrace{\mathcal{F}(w)}_{\Phi^T w = \hat{w}} \odot \underbrace{\mathcal{F}(h)}_{\Phi^T h}) \\
 &= \underbrace{\Phi}_{n \times n} \left(\underbrace{\hat{w}}_{n \times 1} \odot \underbrace{\Phi^T h}_{n \times 1} \right) \\
 &= \Phi \left(\underbrace{\hat{w}(\Lambda)}_{n \times n} \underbrace{\Phi^T h}_{n \times 1} \right) \\
 &= \Phi \hat{w}(\Lambda) \Phi^T h \\
 &= \hat{w}(\underbrace{\Phi \Lambda \Phi^T}_{\Delta}) h \\
 &= \hat{w}(\Delta) \underbrace{h}_{n \times 1}
 \end{aligned}$$

$$w * h = \mathcal{F}^{-1}(\mathcal{F}(w) \odot \mathcal{F}(h))$$

$$\Delta (n \times n) = \Phi^T \Lambda \Phi$$

- Δ is the graph Laplacian operator
- Φ contains column vectors and those are also called Fourier functions
- Λ is a diagonal matrix with Laplacian Eigenvalues

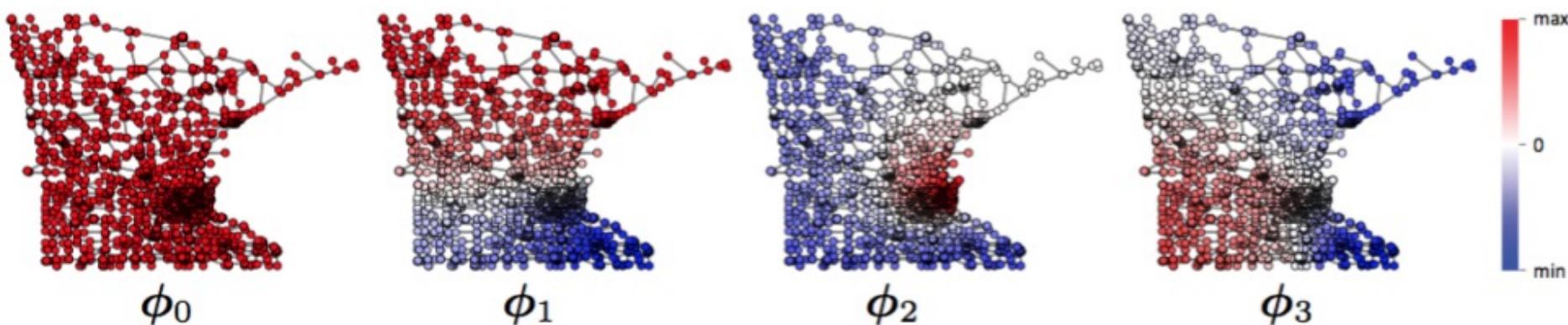
Fourier functions in Graph domain

- Use convolution theorem for spectral ConvNets

Convolution
Theorem

$$w * h = \mathcal{F}^{-1}(\mathcal{F}(w) \odot \mathcal{F}(h))$$

$$\Delta (n \times n) = \Phi^T \Lambda \Phi$$



First Laplacian eigen vectors of a graph

Physical interpretation of terms in the Convolution Theorem

- Use convolution theorem for spectral ConvNets

$$\mathcal{G} = (V, E, A) \rightarrow \Delta_{n \times n} = I - D^{-1/2} A D^{-1/2}$$

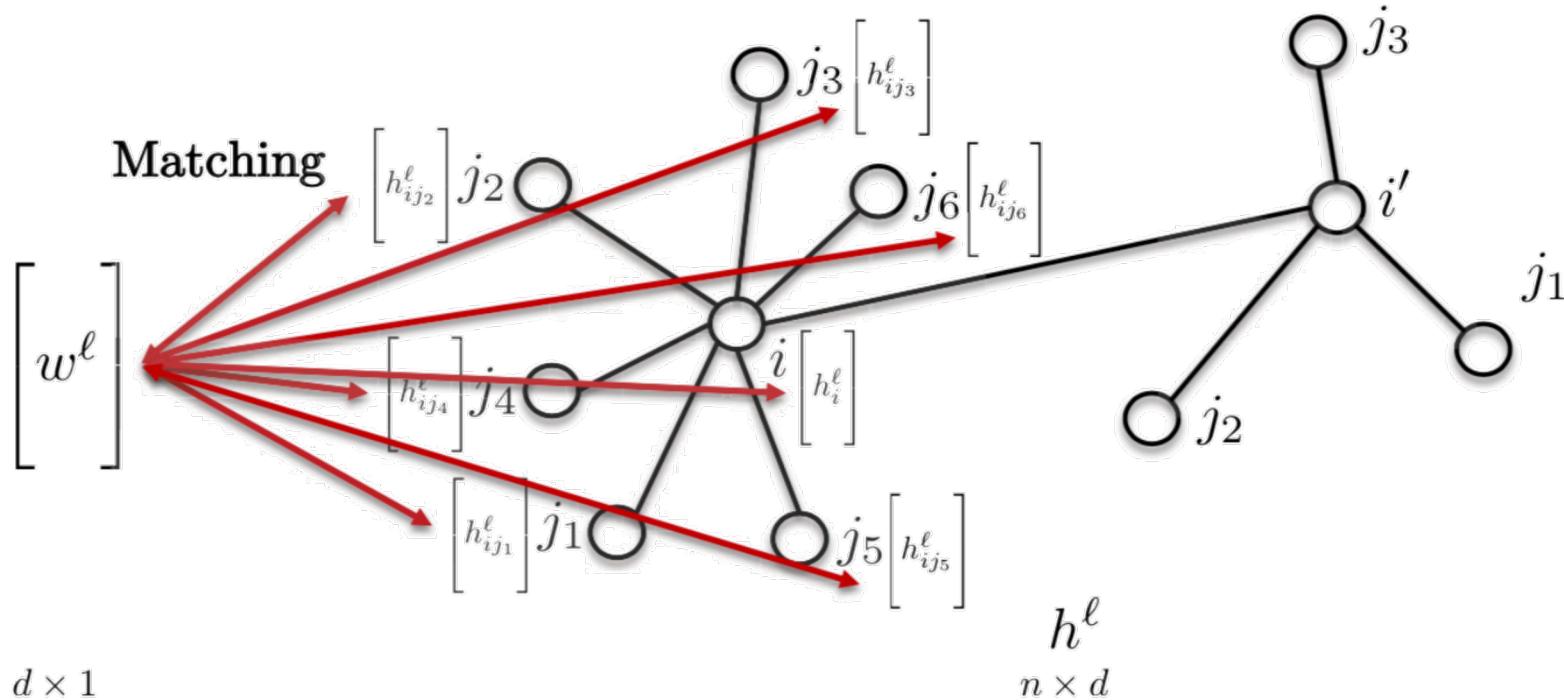
where $D_{n \times n} = \text{diag} \left(\sum_{j \neq i} A_{ij} \right)$

- Matrix A is the adjacency matrix.
- I is the identity matrix
- D is the diagonal matrix, and each element on the diagonal is the degree of the node

The Laplacian shows → the measurement of smoothness of graph or the difference between local values h_i and its neighborhood values of h_j 's.

Template matching for Spatial Graph ConvNets

- Using one template vector w , instead of matrix (like 3×3)
- For each feature (in d) we would have a vector/template w



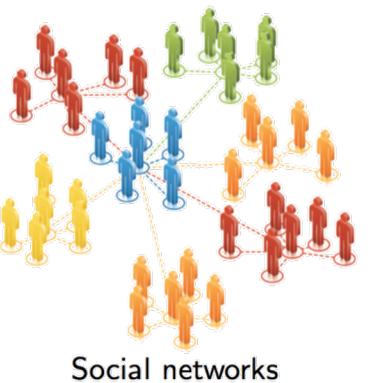
$$h_i^{\ell+1} = \eta \left(\sum_{j \in \mathcal{N}_i} \underbrace{\langle w^\ell, h_{ij}^\ell \rangle}_{\text{scalar}} \underbrace{(h_{ij}^\ell)^T w^\ell}_{\text{One feature}} \right)$$

$$h_i^{\ell+1} = \eta \left(\sum_{j \in \mathcal{N}_i} \underbrace{W^\ell h_{ij}^\ell}_{\substack{d \times 1 \\ d \text{ features}}} \right)$$

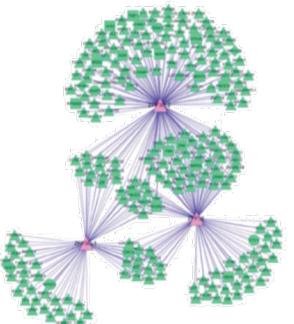
$$h^{\ell+1} = \eta \left(\underbrace{A h^\ell W^\ell}_{\substack{n \times d \\ d \times d \\ n \times d \\ \text{Vectorial representation}}} \right)$$

Conclusion

- GCNs generalize CNNs to data on graphs
- The convolution operator needs to be redesigned for graphs → 2 ways, spectral GCNs (Convolution theorem) (cheaper) and Spatial GCNs (expensive)

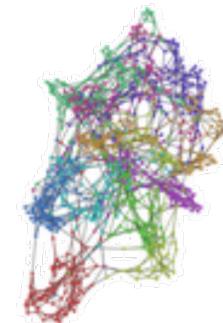


Social networks

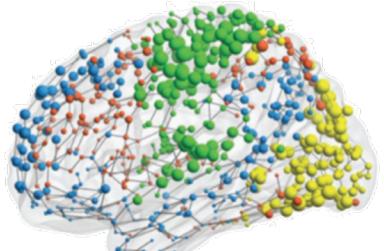


Regulatory networks

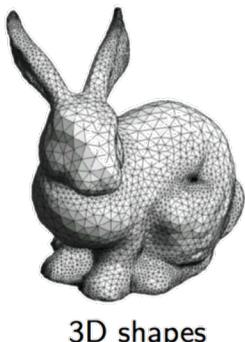
=



Graphs/
Networks



Functional networks



3D shapes



THANKS FOR YOUR
ATTENTION
&
ANY
QUESTIONS?