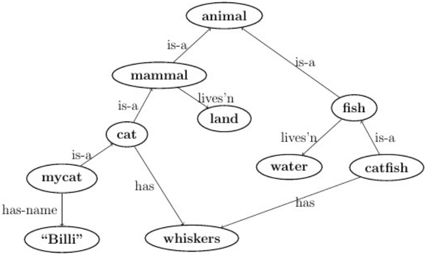
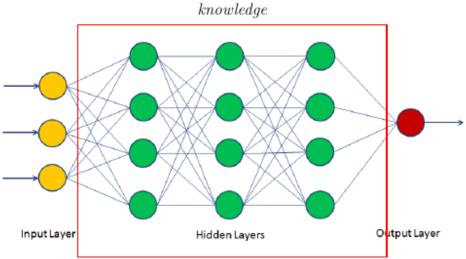
IIT Jodhpur

Biological Vision and Applications
Module 07-05: Graph Neural Networks

Hiranmay Ghosh

Explicit knowledge vs. Implicit knowledge

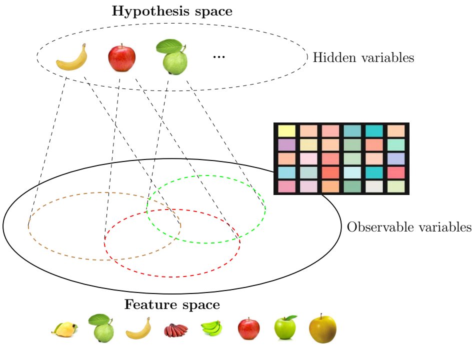




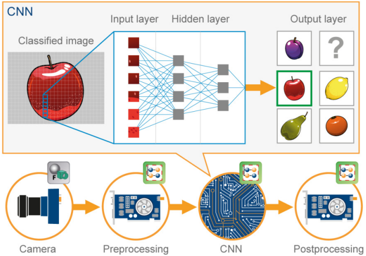
- Explicit knowledge
  - Model based
  - Inductive generalization
  - Slow
  - Good for reasoning
- Implicit knowledge
  - Emergent knowledge
  - No generalization
  - Fast
  - Good for understanding

Model-based Reasoning vs. Model-less understanding

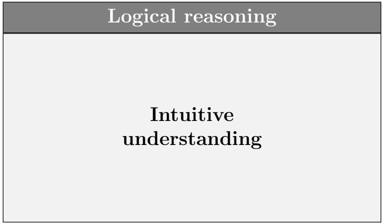
Model-based reasoning



Model-less understanding



Dual Process Theory



Human Cognitive Process

- Interaction?
- Reasoning = asking questions

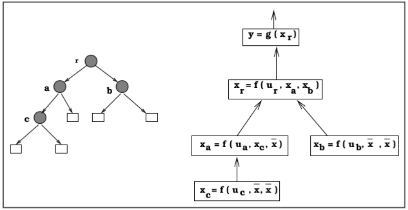
- Can we have reasoning coupled with understanding
  - "Intersection of deep learning and structured approaches"
  - Inductive generalization is the key to AI
  - Structure with emergent knowledge

Position paper from DeepMind, Google Brain, MIT, University of Edinburgh (2018)

Comments on the position paper

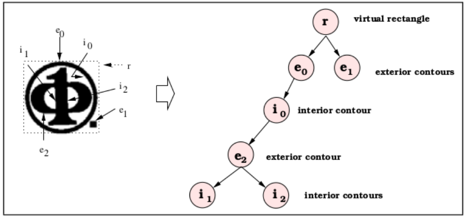
Combining structural analysis and machine learning

Part-based recognition of logos (1998)



- Functions f(), g() realized as NN
  - Trained with large data set
- Identical property descriptors u_x

IIT-J IIT-D IEEE IEEE-CompSoc ACM

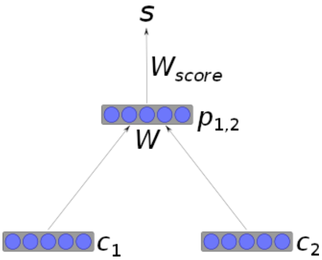


- Fairs better than a single NN

Frasconi, et al. A General Framework for Adaptive Processing of Data Structures

Frasesconi, et al. Logo Recognition by Recursive Neural Networks

Recursive Neural Network (RvNN)



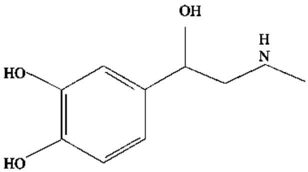
- c_1, c_2 are inputs
  - Vectors of dimension n
- p_{12} is the output
  - Also a vectors of dimension n
- $p_{12} = \mathbf{tanh}(W[c_1; c_2])$, where
  - W is a learned $n \times 2n$ weight matrix

- Limitations
  - Data must be organized as a binary tree
  - Data flow is one way – leaf to root

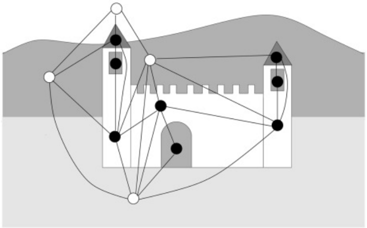
Graph Neural Network

Data is represented as a graph

- Graph focussed Application
  - What is the property of the molecule ?
  - $output = \tau(G)$

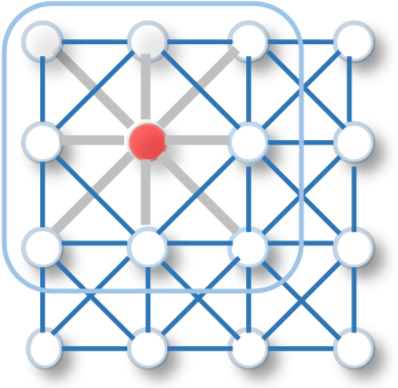


- Node focussed Application
  - What does each of the nodes represent ?
  - $output = \tau(G, n)$

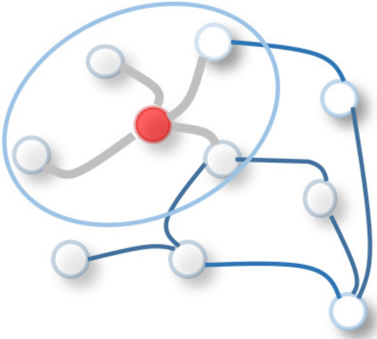


Scarselli, et al. Graph Neural Network (2009) *

2D Convolution vs. Graph Convolution



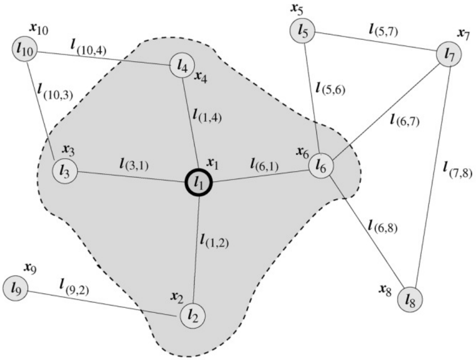
2D Convolution



Graph Convolution

Convolutional Graph Neural Network

Processing model

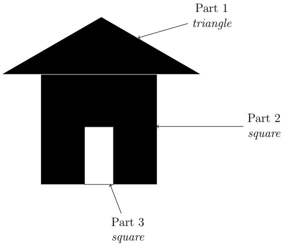


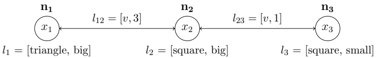
$$x_1 = f_w(l_1, \underbrace{l_{(1,2)}, l_{(3,1)}, l_{(1,4)}, l_{(6,1)}}_{l_{co[1]}}, \underbrace{x_2, x_3, x_4, x_6}_{x_{ne[1]}}, \underbrace{l_2, l_3, l_4, l_6}_{l_{ne[n]}})$$

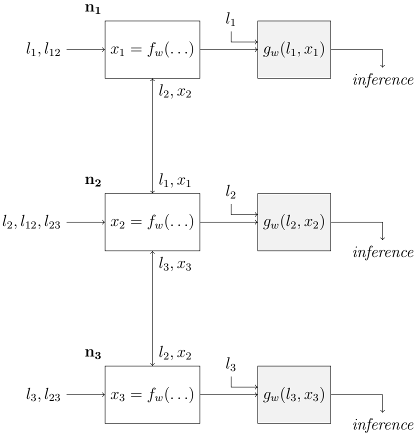
- Nodes have identical property (feature) descriptors (l_n)
  - e.g. color, texture, shape
- Edges have identical property descriptors (l_{mn})
  - e.g. distance between the center of gravities of the nodes
- State of a node: x_n

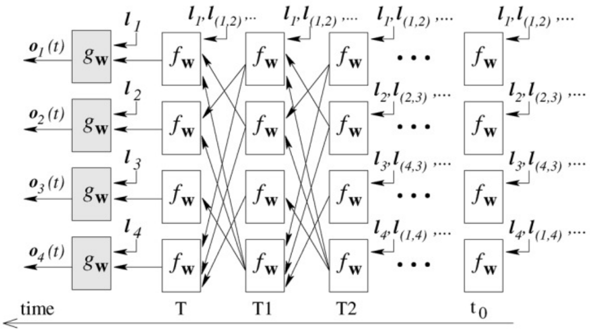
Graph Neural Network

Processing model: Example





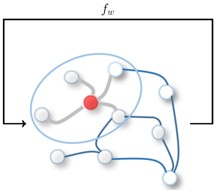


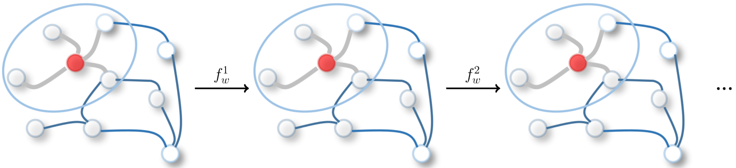


- $f_w()$ and $g_w()$ are learned with training
- Take the output after several recursions
  - Is the system guaranteed to go into a steady state after a finite number of iterations?

Recursive & Convolutional Graph Neural Network

Rec-GNN & Conv-GNN





Recursive GNN:

Same weights

at every time-stamp

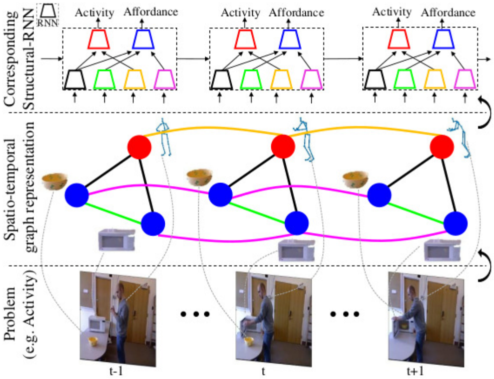
Convolutional GNN:

Different weights at different time-stamps

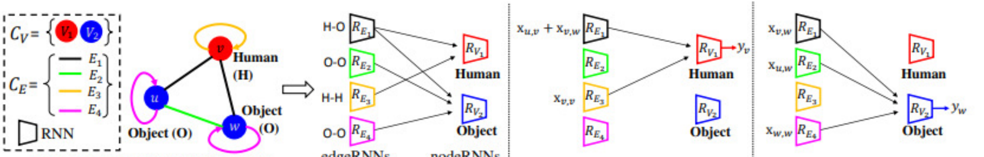
A compact review of ... GNN

Structured RNN (S-RNN)

GNN: Spatio-Temporal adaptation



Construction of S-RNN for a video shot



(a) Spatio-temporal graph with colors indicating sharing of factors
(b) Corresponding S-RNN
(c) Forward-pass for human node v
(d) Forward-pass for object node w

CVPR-16 Paper *, (Presentation Video) *

Applications

- Rec-GNN / Conv-GNN (still images - 2D/3D)
  - Fine-grained classification (e.g. bird species)
  - 3D point cloud processing (LiDAR)
- S-RNN (video / motion picture)
  - Human action recognition
  - Human/Robot - Object Interaction
  - Human Motion Modeling

Limitations and Future Research

- Scalability
  - Graph size needs to be limited
  - Number of nodes / number of edges
- Heterogeneity of graphs
  - Presently graphs are assumed to be homogeneous
- Dynamicity
  - Graph structure changing over time
- Model depth
  - Performance (accuracy) drops with depth

Quiz

Quiz 07-05

End of Module 07-05