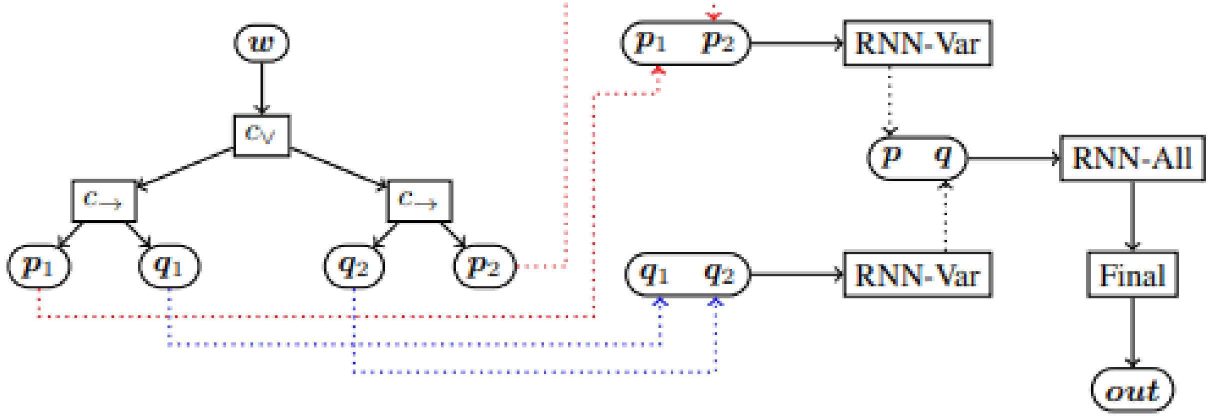
Directed Graph Networks

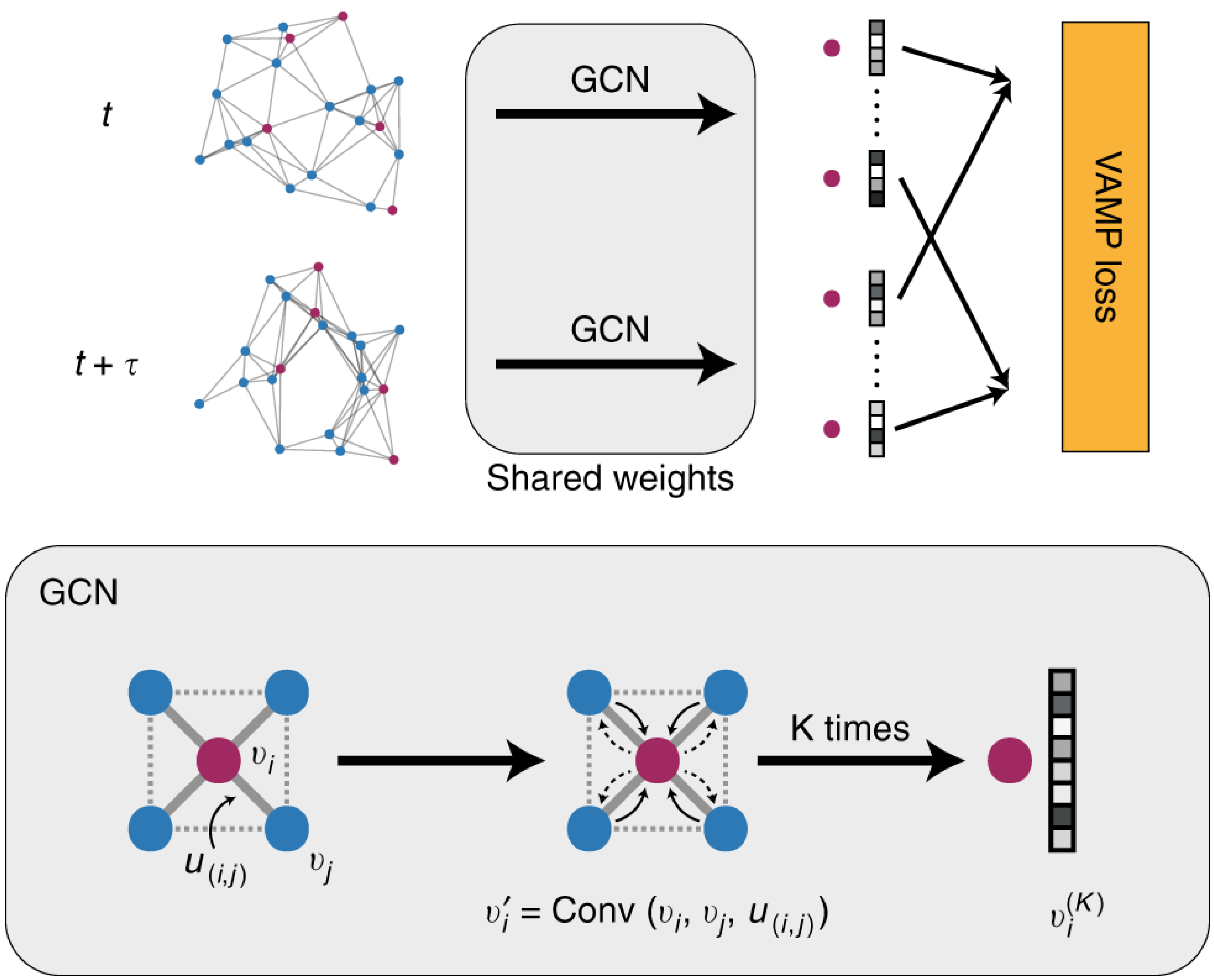
Neural Networks for Logic

- Sequence NNs: LSTM, 1-D convolutions
- Tree-like. TreeRNNs, TopDownNet
- Algorithmic assistance (cheating?): PossibleWorldNet
- (hyper-) graphs…



Graph Neural Networks

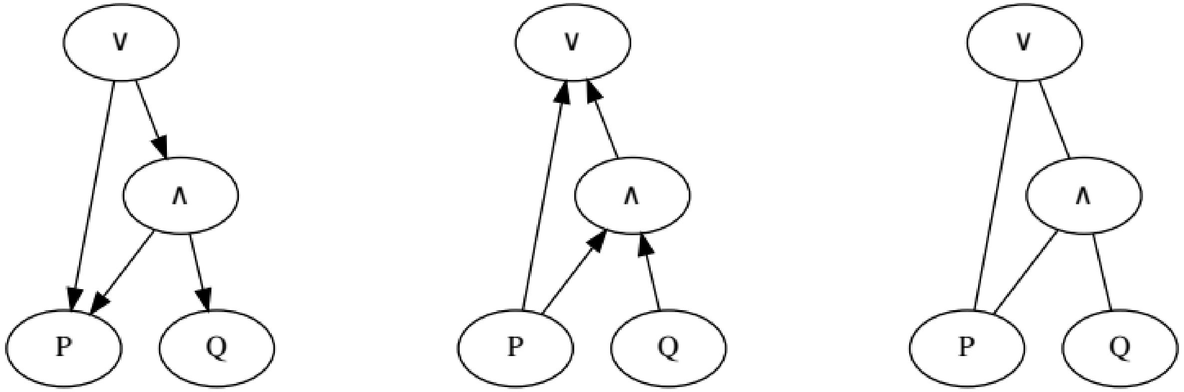
- Treat inputs as a graph (typically undirected)
- Each node has some feature vector attached: maybe edges as well
- Pass messages to neighbours to produce a new vector: “convolution”



t
t + τ
GCN
GCN
Shared weights
VAMP loss
GCN
v_i
$u_{(i,j)}$
v_j
K times
$v_i' = \mathrm{Conv}\,(v_i, v_j, u_{(i,j)})$
$v_i^{(K)}$

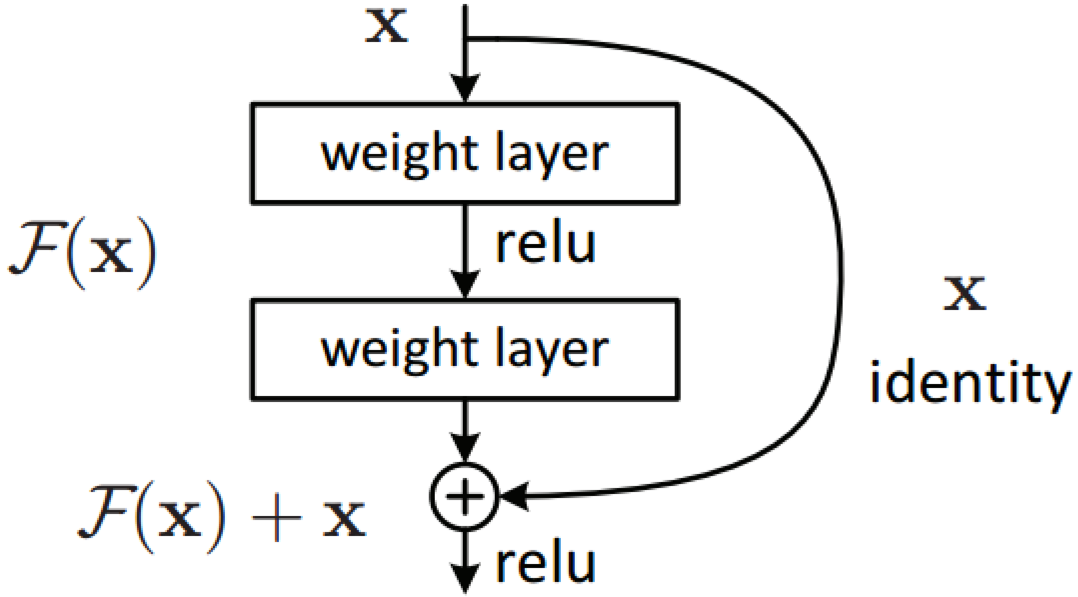
GNNs for logical formulae

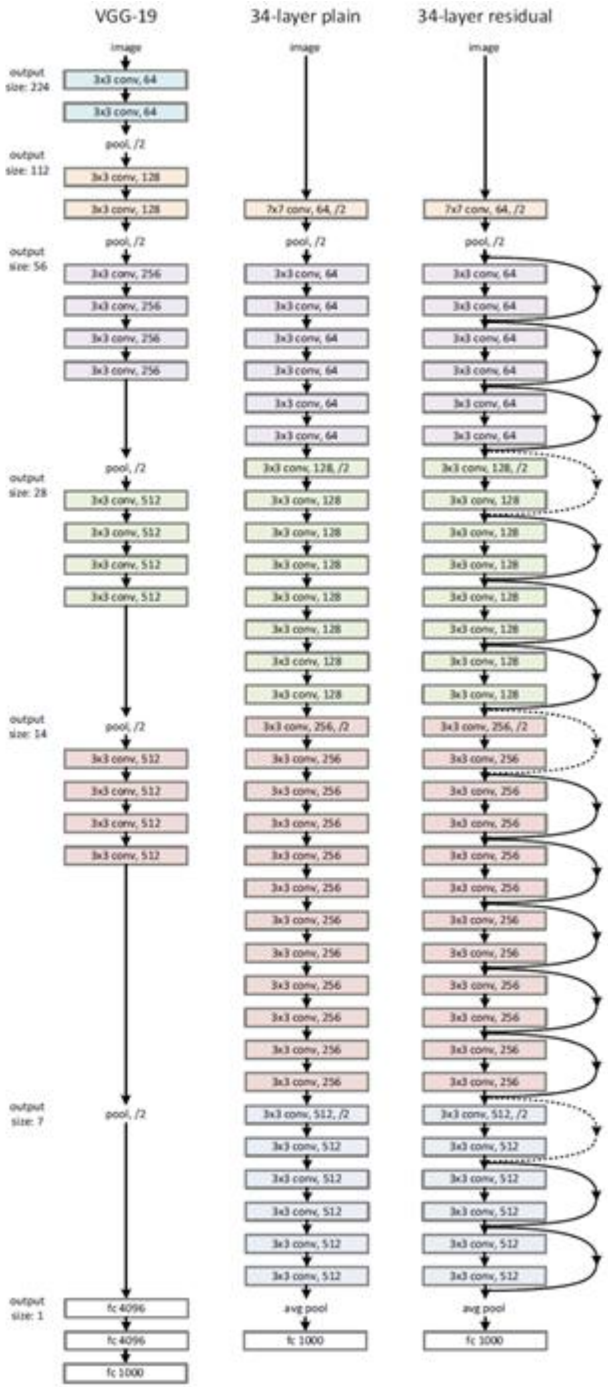
- Natural, avoid learning to parse structure
- Captures e.g. variable binding pleasantly, invariant (!) to e.g. associativity, alpha-equivalence
- Undirected graphs lose structure, directed graphs only pass messages one way?



(a) top-down only (b) bottom-up only (c) undirected

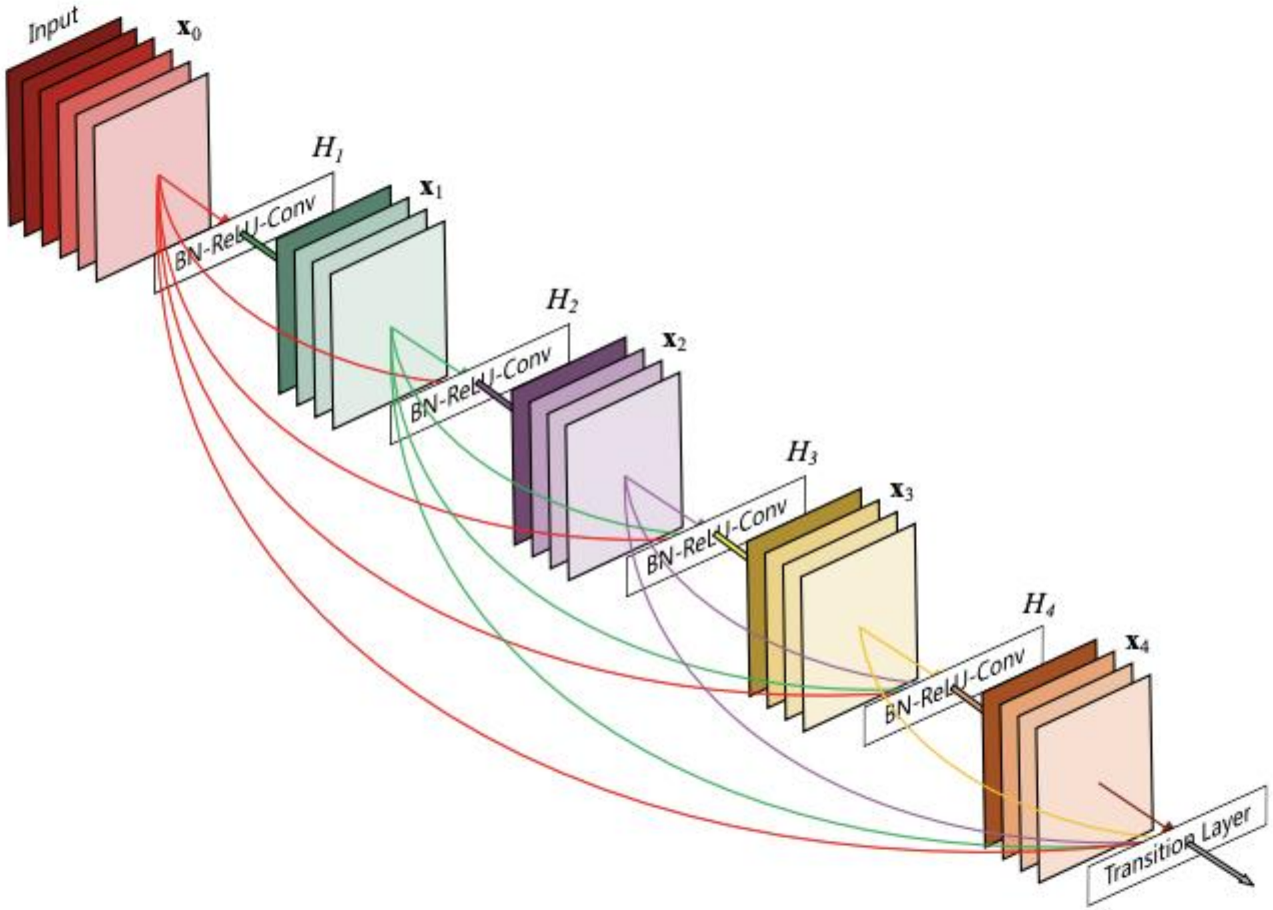
Aside: residual networks





VGG-19
34-layer plain
34-layer residual
image
output size: 224
output size: 112
output size: 56
output size: 28
output size: 14
output size: 7
output size: 1
3x3 conv, 64
pool, /2
3x3 conv, 128
3x3 conv, 256
3x3 conv, 512
fc 4096
fc 1000
7x7 conv, 64, /2
3x3 conv, 128, /2
3x3 conv, 256, /2
3x3 conv, 512, /2
avg pool

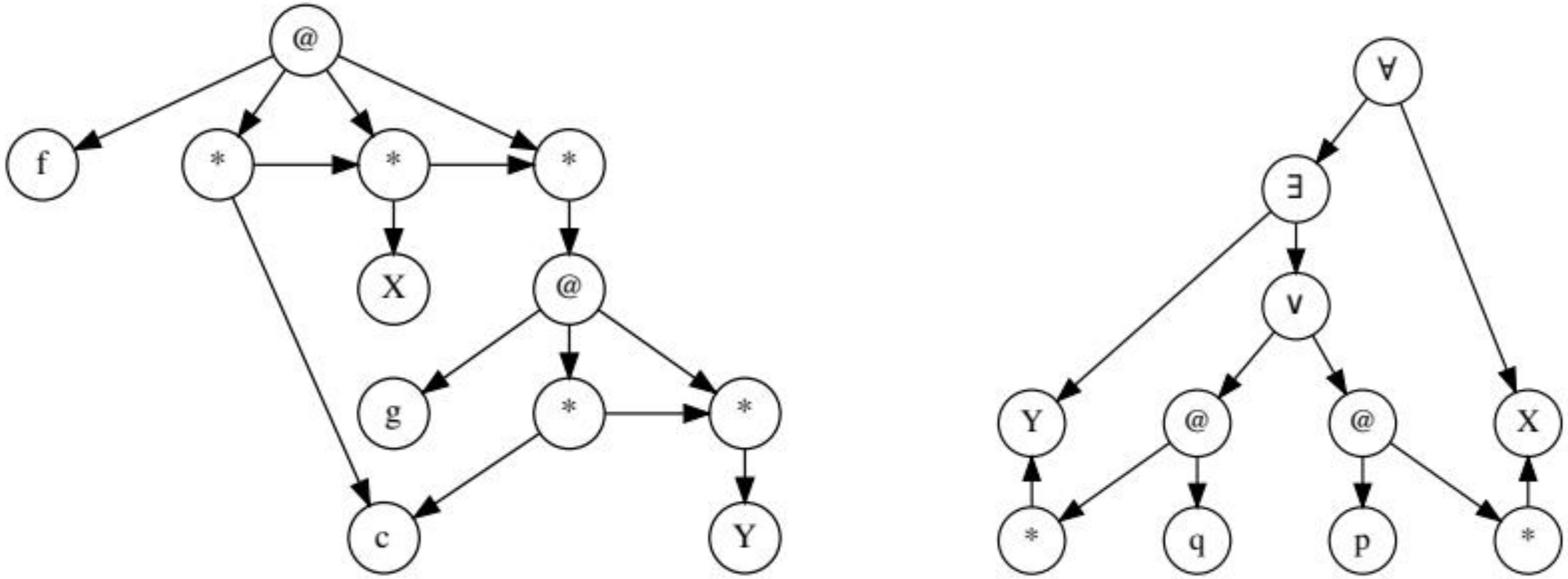
Dense Networks



Bidirectional Graph Networks

- Idea: “convolve” in *both* edge directions
- Factor-2 blowup in feature size for each layer
- Use feature-reduction layers of e.g. *DenseNet* to control blowup

Directed Graphs for FOL



(a) $f\,(c, X, g\,(c, Y))$ (b) $\forall x.\exists y.P(x) \vee Q(y)$

Figure 3: First-order graph encodings, showing (a) argument ordering and (b) variable binding.

Tips & Tricks

- Cyclic learning rate really helpful for this
- Vanilla SGD+Nesterov momentum
- Batch normalisation: a bit unclear but theoretically good.

Results

- 79.8% DeepMath unseen set

Table 3: Propositional Entailment Accuracy

model	valid	easy	hard	big	massive	exam
PossibleWorldNet	98.7	98.6	**96.7**	**93.9**	73.4	96.0
TopDownNet	95.5	95.9	83.2	81.6	83.6	96.0
Contribution	**99.4**	**99.3**	91.2	88.3	**89.2**	**97.0**

The Future

- Pooling?
- Hypergraphs?
- Weird convolutions?
- New domains?