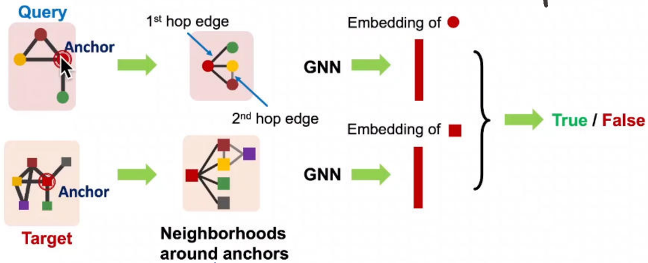
Neural Subgraph Matching

我们使用嵌入空间的形状捕获子图，而非进行穷举。



(围绕锚点的 k-hop 邻域)

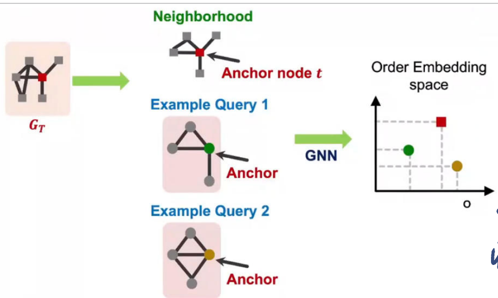
Q: 如何查找给定节点的 k-hop 邻域？

A: 对给定节点使用BFS

[Order Embedding Space]

将节点嵌入至一个有序的高维非负空间中，$v_1 \preccurlyeq v_2$ 表示 v_1 嵌入的所有维度的分量均不大于 v_2。

这揭示了一种有序性/传递性，即若 $v_1 \preccurlyeq v_2$，$v_2 \preccurlyeq v_3$，则 $v_1 \preccurlyeq v_3$。



在这个空间中，我们可以利用这种传递性很好地捕获子图关系，因为子图关系的传递性是类似的

日期: /

除此之外，有序嵌入空间还可以捕获：

1) 反对称性：若G_1是G_2子图，G_2是G_1子图，则二者同构

2) 交集下的闭包概念：单节点图为任何节点的子图（对应原点）

那么使用怎样的损失函数才能捕获这种顺序约束呢？

Objective: $\forall_{i=1}^{D} z_q[i] \le z_u[i]$ iff $G_D \subseteq G_T$

Loss (max-margin loss)：定义 $E(G_q, G_t) = \sum_{i=1}^{D} (\max(0, z_q[i] - z_t[i]))^2$

为G_q，G_t之间的"边界"(margin)

我们的目标是使：$E(G_q, G_t) = 0$，$G_q \subseteq G_T$

$E(G_q, G_t) > 0$，$G_q \not\subseteq G_T$

因此，$loss = \begin{cases} E(G_q, G_t), & G_q \subseteq G_T \\ \max(0, \alpha - E(G_q, G_T)), & G_q \not\subseteq G_T \end{cases}$

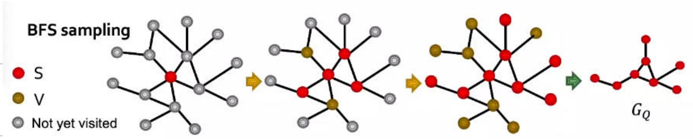
如何获取样例？

正例：通过BFS采样G_T的诱导子图G_Q

- 初始化：$S = \{v\}$，$V = \phi$
- 在每一步，从$N(S) \setminus V$中采样部分节点，归入S，其余归入V
- K步之后，获取S诱导的子图G_Q，锚点为v

通常3~5

日期:　　/



负例：通过对 G_Q 施加扰动，对其中的边或节点进行破坏即可。

那么在推理时，确定 G_Q 是否为 G_T 子图，只需对 G_Q、G_T 的锚点 q，t 分别进行嵌入（需进行遍历），当 $E(G_q, G_t) < \epsilon$（ϵ 为超参数）（软化条件）时，认为子图关系存在。

锚点t的邻域

锚点q的邻域（应为 G_Q 全图）