一：“痕” 大抵应当被当作感觉和感觉一起训练，而不是拎出来单独训练，而我写的程序貌似可能把 “痕” 单独训练了，在下一代可能要修改，不过经过仔细的检查后，我感觉貌似不用修改的。  
  
二：由于需要信息无损，所以建议用四叉树数据结构存储权重，类似于矩阵平均分块，比如  
0 0 1 0  
0 1 0 0  
0 0 0 1  
0 0 0 0  
存储为两层四叉树（  
序1 序2  
序3 序4  
）  
1101  
0001 / 1000 / 0100 ，4 个位一节  
类似地，稀疏矩阵  
1 1 0 0 0 0 0 0  
0 1 1 0 0 0 0 0  
0 0 1 1 0 0 0 0  
0 0 0 1 1 0 0 0  
0 0 0 0 1 1 0 0  
0 0 0 0 0 1 1 0  
0 0 0 0 0 0 1 1  
0 0 0 0 0 0 0 1  
存储为  
1101  
1101 / 0010 / 1101  
1101 / 0010 / 1101 / 0010 / 1101 / 0010 / 1101  
可见，它只用了 44 个位，虽然可能还要存一些额外信息，比如层索引，但 44<64，可见对于大的稀疏矩阵它是挺有用的，而且它可以直接用来计算（如果与向量进行计算需要是类似的二叉树数据结构向量(序1，序2)），涉及的相应机制如下：  
1. 用二叉树数据结构存储向量，用四叉树数据结构存储矩阵  
2. 采用多个1父级2子级的父子级指针并行遍历树数据结构的每父子级层  
3. 用节掩码来指导剪枝，注意是节掩码，而非位掩码，每个节有1 bit 掩码，其功能与由 4 位一节改为 5 位一节并以多出来的一位用作掩码类似。