

- Prefix beam search的过程,每个时刻有如下3个动作: 1. 扩展:根据前缀串和当前时刻的输出,计算新串的概率。
- 2. 归约:将规约串相同的候选概率相加。
- 3. 裁剪:仅保留top k个最好的序列做下一时刻的拓展,绿色的表示保留,红色的表示被裁减掉,图示中k为3。