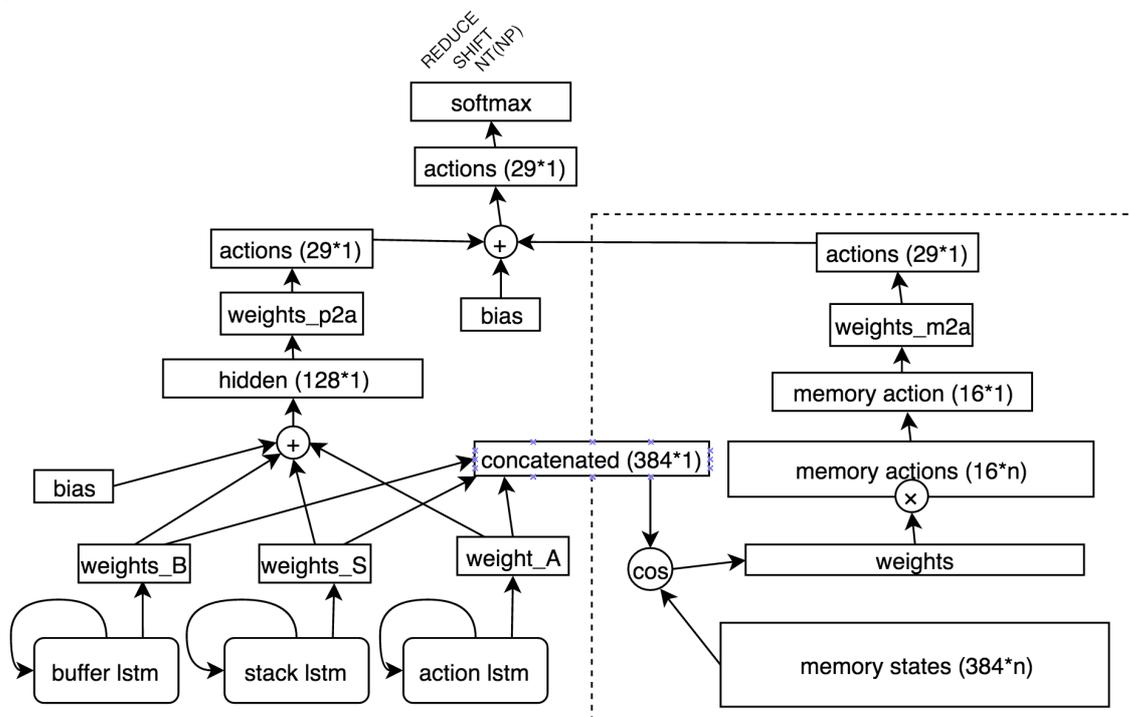


# RNNG + Memory 实验报告

张诗悦

## 2016.11.1

尝试在 RNNG 中增加 memory，现在只考虑结构如图下图所示。虚线框内部分是增加的模型结构。



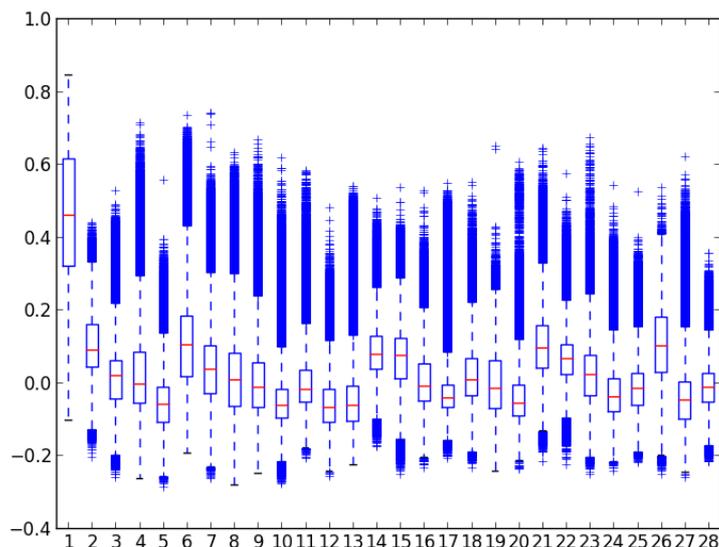
现在的实现比较简单，具体过程如下：

1. 取训练好的模型，重新输入训练集作为测试数据，获取每一步输出的 {stack, buffer, action} 384 维隐状态，共有 2599716 个隐状态。
2. 按照隐状态对应的 29 种 action 分类，分布如下：

| action      | 隐状态数    | action                 | 隐状态数     |
|-------------|---------|------------------------|----------|
| 0,REDUCE    | 792794  | 15,NT(NAC)             | 411      |
| 1,SHIFT     | 1014128 | 16,NT(WHADVP)          | 2638     |
| 2,NT(S)     | 100096  | 17,NT(PRT)             | 2641     |
| 3,NT(PP)    | 95581   | 18,NT(NX)              | 1344     |
| 4,NT(NP)    | 350432  | 19,NT(WHPP)            | 392      |
| 5,NT(PRN)   | 2420    | 20,NT(SQ)              | 350      |
| 6,NT(VP)    | 146319  | 21,NT(SBARQ)           | 222      |
| 7,NT(ADVP)  | 22447   | 22,NT(CONJP)           | 302      |
| 8,NT(SBAR)  | 30532   | 23,NT(WHADJP)          | 59       |
| 9,NT(ADJP)  | 14607   | 24,NT(INTJ)            | 127      |
| 10,NT(QP)   | 9444    | 25,NT(X)               | 176      |
| 11,NT(UCP)  | 497     | 26,NT(RRC)             | 47       |
| 12,NT(WHNP) | 9115    | 27,NT(LST)             | 38       |
| 13,NT(SINV) | 2042    | <b>28,NT(PRT ADVP)</b> | <b>1</b> |
| 14,NT(FRAG) | 514     |                        |          |

因为有些action经常出现，例如：reduce, shift。因此大部分的隐状态也集中在这些action上。也发现了一个异常第28个action应该被归到第17个action中，后续会来处理）。

3. 因为想利用模型的隐状态和memory states之间的cos距离来增加memory的影响，因此先做了一个统计工作，验证同一个action对应的states之间的距离要比不同action对应的states之间的距离近。首先，求得0-27个action对应的states的center(平均)，其次，对每个action下的states，求其与这28个center的cos距离。



上图表示“REDUCE”对应的states与28个center的cos距离。发现还是和REDUCE本身的center距离相对最近。同理，其他的action下的states都有类似的性质。也就是简单证实了做法的合理性。

4. 将28个action下的centers作为模型图中的memory states, 对应的action的embedding后的结果作为memory actions。
5. 重新训练模型，将每次stack, buffer, action三个lstm输出的hidden vector连接在一起，和memory中的28个states做cos，得到权重；按照权重组合对应的action向量；将得到的向量组合上原始模型输出的向量，具体请见模型结构图。

结果：

还在跑，暂时从模型收敛速度上看不出什么特别的提升。按理说应该会收敛的快一下才对，但是并没有。

改进方向：

需要改进的地方很多，最容易想到的是memory太少了，简化的太多，可以不只取center，应该还要在每类action中多选择一些memory加进去。