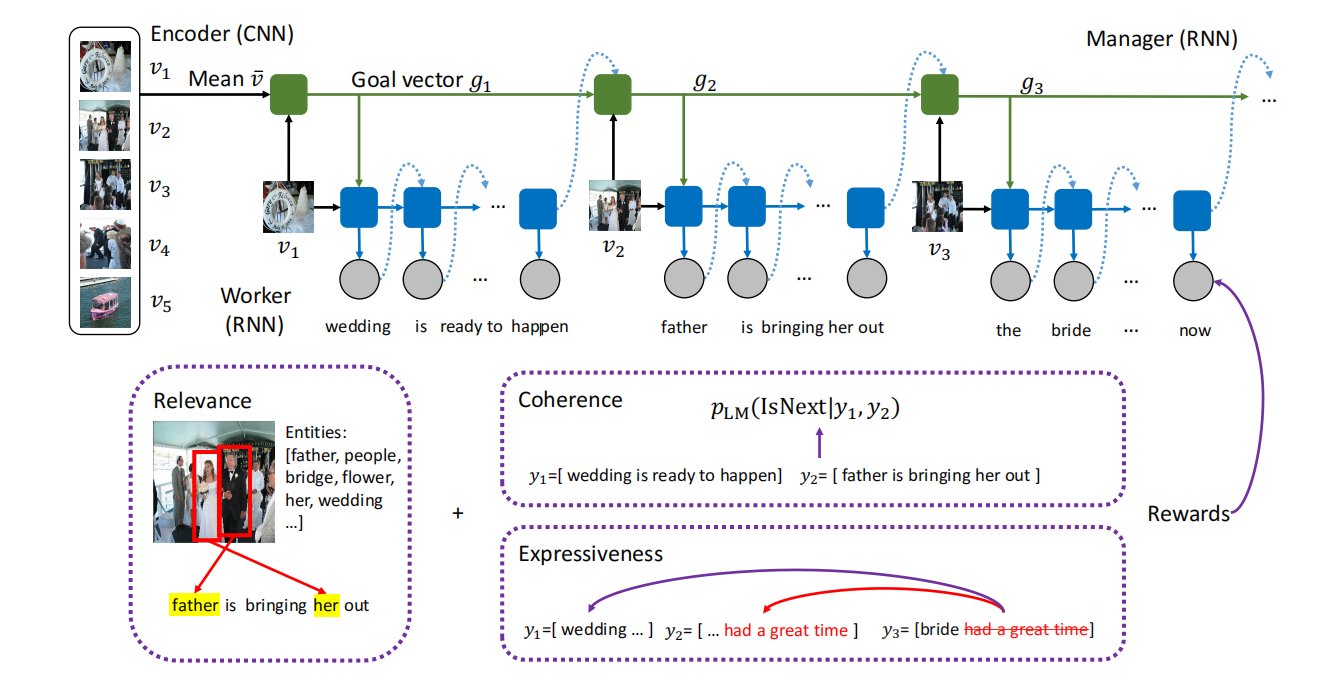
在本文中，作者的关键思想是重新审视了what makes a good story这个问题，然后提出了三个新的评价指标，它们分别是relevance(关联性),coherence(连贯性)，expressiveness（表现力），然后将这个三个指标做成一个组合的reward，接下来用强化学习来训练它，作者提出的模型名称即是Reco-RL。

作者提出的模型主要由两个部分组成，一个encoder部分，这个模块主要是由一个预训练的CNN组成（文中说的是resnet-101）,它需要从每个图像中提取深层视觉特征，将这个特征作为一个向量传递给下面的decoder部分。接下来是decoder部分，decoder部分也分成两部分，一部分是high-level decoder（manager），另一部分是low-level decoder（worker），

具体来说，manager需要考虑整个故事流程，当前图片的视觉信息，前一张图片生成的句子，根据以上三个信息生成一个目标向量(goal vector),然后将这个目标向量传递给下面的worker，worker根据这个目标向量来生成句子描述。也可以这样理解，manager是作用在句子上面的，而worker是作用在单词上面的。

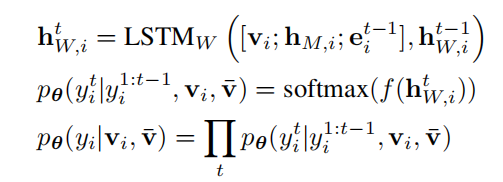


关于high-level和low-level部分，文中有专门的具体章节介绍。

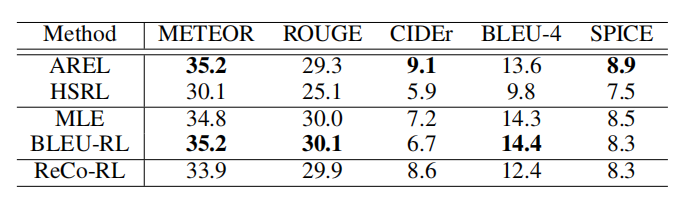
**High-level（manager）**: 它是一个LSTM网络，它是从句子层次上来把握生成故事的连贯性。在描述一张图片的时候，它需要考虑三个部分，①一组照片的总体流程；②当前图片的内容信息；③前一张图片生成的信息。换成公式来表示就是

QQ截图20201203154425

**low-level（worker）**：它也是一个LSTM网络，它需要每次预测一个单词，它是从单词层次来控制句子的连贯性。从上文可以看出，worker是由manager生成的目标向量来控制的。具体来说，当worker生成一个单词时，它需要考虑的是当前图片的信息，manager生成的目标向量以及前一个生成的单词，换成公式可以表达为：



在文章中，作者还提到了自己的三个新指标时如何计算的，这里不做具体阐述。训练过程的话，作者首先用极大似然值来训练自己的模型，然后再用强化学习的方法来训练模型，作者将自己的模型与两个baseline进行比较，一个是AREL，一个是HSRL，对比结果如下（注意这里的比较指标还是传统的句子生成指标）：



接下来，作者用他提出的三个指标来将自己的模型与其他的模型进行比较，对比结果如下：

