



a5在a4的基础上更加与时俱进了一些。它基于seq2seq这套encoder + decoder + attention的架构，尝试在更小粒度的character level做embedding。你需要改进的地方如下：

1. 如何利用新的char-level CNN生成原有encoder所需要的embedding
2. 如何在原有decoder输出为<UNKNOWN>时利用新的char-level RNN生成一个词

开篇强调本次作业code-heavy，因为"...we give you much less help in writing and debugging that code..."

## a5相关课程

1. Lecture 11. Convolutional Networks for NLP: 同lecture 6 & 7的性质一样，这节课也是介绍char-level NMT的预备知识Convolutional Neural Network。考虑到本课提到的CNN和Highway network都需要你在a5里徒手撸一遍，建议还是听一下...
2. Lecture 12. Information from parts of words: Subword Models: 如标题所示，本节课讲的是... subword level modeling，其中既有char-level也有wordpiece-level。听完lecture 11 & 12解锁a5。

## A5的一些坑

1. 注意padding要做的不只是padding；在word length过大时（a5中 `max_word_length = 21`）还需将其truncate，否则embedding会炸。
2. 独立生成一个tensor时（如 `torch.tensor(...)`，即不利用上游的tensor计算得出），记得设置 device，否则切到GPU时会炸。

