

1 HMMs and CRFs

A Hidden Markov Model (HMM) can be used a generative model of a system of linear symbols, like language.

Model	Corpus	Training accuracy	Testing accuracy	OOV-only	Time
HMM	Atis	0.888	0.865	0.180	2.987
CRF	Atis	0.999	0.928	0.273	84.971
CRF+	Atis	0.999	0.940	0.440	84.570
HMM	WSJ	0.862	0.785	0.379	51.304
CRF	WSJ	0.951	0.752	0.431	1604.749
CRF+	WSJ	0.992	0.862	0.626	5608.262

Table 1: ‘CRF+’ refers to the CRF with additional word-property features; ‘OOV-only’ refers to accuracy on out-of-vocabulary items in the held-out test data.

```
git clone https://github.com/chbrown/nlp.git
```