Timings

- Using core features
  - Trunk Moses: 47 minutes
  - Sparse Moses: 54 minutes

- Using extended features (and a smaller model)
  - Sparse Moses: 372 minutes
Optimisations

- Pre-calculate features which only depend on phrase pair / rule
- Reduce (or eliminate!) string concatenation in names
- Don’t store full feature vectors unless needed
- Use a faster feature vector