Allow more derivations
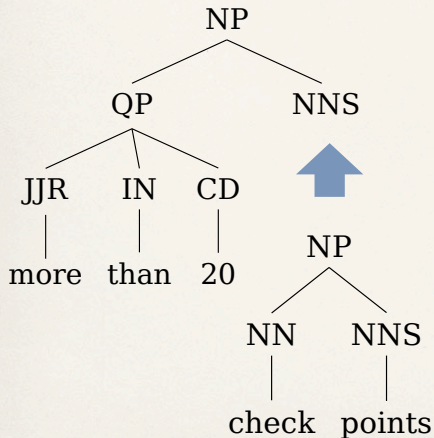

* STSG: allow only matching substitutions

* Hiero-like: allow any substitutions

* Let the model learn to choose:

    * matching substitutions

    * mismatching substitutions

    * monotone phrase-based