FCFG Modeling of English Syntax: Joining Tense Projections with the Internal Structure of Noun Phrases

Ehva Saur and Sean Bethard

May 9, 2010

1 Towards a feature-based CFG of English

Our primary focus was on Tense Projections. For reasons of simplicity and coherence we have restricted our grammar fragment to the present tense. Past tense constructions can be captured with a simple augmentation, however. By introducing a TENSE feature and assigning it the respective [pres] and [past] values we can accommodate tense-sensitive verbs, including main verbs, auxiliaries, and do-support.

We expanded our grammar fragment by incorporating Group 3's code for the internal structure of noun phrases. After this augmentation our grammar is now able to accommodate complex determiner phrases. Joining the two grammars forced us to simplify certain agreement rules. Other rules, such as those for recursive adjectival phrases and PP-modification, remained intact. These and other adjustments are examined more closely in the sections to follow.

2 Tense Projections

In English, tense can be carried by different kinds of verbs.

- main verbs agree in person and number with the subject, show present/past alternation
- do-support agrees in person and number with the subjet, shows present/past alternation
- auxilaries agree in person and number with the subject, show present/past alternation
  - BE [progressive]: "I am leaving / You are leaving / He was leaving"
  - HAVE [perfective]: "John has left / We have left / They had left"

- modals show no subject agreement in person and number and no present/past alternation since they themselves are "tensed".
  - WILL: "I will leave."
  - MIGHT "John might eat apples."

Main verbs, auxiliaries and modals are differentiated into three different Tense Projections: VP, AuxP, and ModP.

- TP[AGR=?a] -> VP[AGR=?a] | AuxP[AGR=?a] | ModP

ModP shows no agreement. Since do-support is only obligatory for the negation of main verbs, it is added as a sub-rule under the VP. Our grammar does not extend to emphatic *do* constructions.

In the early stages of our project, we did not distinguish main verbs, auxilaries, modals and do-support. Because each of them, either overtly (main verbs, auxiliaries, DO-support) or covertly (modals), carry tense, we simply grouped them beneath a single projection. In distinguishing between these categories, however, we are now in a better position to expand our system to other aspects of English syntax, including inverted yes-no questions, non-embedded wh-questions, and inversion cases involving direct quotes and certain adverbs.

2.1 VP Structure

The verb inside the VP is specified as either transitive or intransitive. Agreement is free to move about the structure.

- TP[AGR=?a] → VP[AGR=?a]
- VP[AGR=?a] → V_trans[AGR=?a]DP | V_intrans[AGR=?a]
- TP[AGR=?a] → DO[AGR=?a] Neg InfP

On the surface, the infinitive form of English main verbs is identical to all present tense forms except in the third person singular. An infinitive category InfP is introduced in order to prevent incorrect parses like the following.

- *TP[AGR=?a] -> DO[AGR=?a] Neg VP[AGR=[PER=3, NUM=pl]]

The InfP, just like the VP, can be broken down in transitive and intransitve verbs, which are specfied in the lexicon.

- InfP → Inf_trans DP | Inf_intrans

2.2 AuxP Structure

- AuxP[AGR=?a] → Aux[ASP=perf, AGR=?a] ParP[ASP=perf]
  - John has left
- AuxP[AGR=?a] → Aux[ASP=prog, AGR=?a] ParP[ASP=prog]
  - John is leaving
- AuxP[AGR=?a] → Aux[ASP=perf, AGR=?a] ParP[ASP=perfprog]
  - John has been sleeping
- Aux[ASP=perf] → HAVE[agreement]
- Aux[ASP=prog] → BE[agreement]

2.3 ParP Structure

- ParP[ASP=perf] → Par_intrans[ASP=perf] | Par_trans[ASP=perf] DP
  - Par_intrans[ASP=perf] → 'eaten' |walked'
- ParP[ASP=prog] → Par_intrans[ASP=prog] | Par_trans[ASP=prog] DP
  - Par_trans[ASP=prog] → 'eating' | 'walking'
- ParP[ASP=perfprog] → Par[ASP=perf] ParP[ASP=prog]

It is now important to specify where Negation and Adverb Phrases can be inserted in ParP.

- ParP[ASP=prog/pref] → Neg ParP[ASP=prog/perf]
- ParP[ASP=prog/perf] → AdvP ParP[ASP=prog/perf]
- ParP[ASP=prog/perf] → AdvP Neg ParP[ASP=prog/perf]
- S → DP AdvP TP
- S → AdvP S | S

AdvP handles the adverbs outside TP that are a subcategory of AuxP.

2.4 ModP Structure

- ProgP -> Prog ParP[ASP=prog] "be doing"
- PerfP -> Perf ParP[ASP=perf] "have done"
- PerfprogP -> Perf ParP[ASP=perfprog] "have been doing"
- Prog -> BE[infinitive]
- Perf -> HAVE[infinitive]
- ModP → Mod InfP
  - John might eat apples
- ModP → Mod ProgP
  - John might be eating apples
- ModP → Mod PerfP
  - John might have eaten apples
- ModP → Mod PerfprogP
  - John might have been eating apples

3 Adverbs and Negation

The following rules manage adverbs and negation inside the ModP.

- ModP → Mod Neg InfP | Mod Neg ProgP | Mod Neg PerfP | Mod Neg PerfprogP
- ModP → Mod AdvP InfP | Mod AdvP ProgP | Mod AdvP PerfP | Mod AdvP PerfprogP
- ModP → Mod Neg AdvP InfP | Mod Neg AdvP ProgP | Mod Neg AdvP PerfP | Mod Neg AdvP PerfprogP

In English, negation is closer to the verb than the adverb. Together with the known rules for the participle phrase, the positions for adverbs and negation are specified with the following rules.

- ParP[ASP=prog/pref] → Neg ParP[ASP=prog/perf]
- ParP[ASP=prog/perf] → AdvP ParP[ASP=prog/perf]
- ParP[ASP=prog/perf] → AdvP Neg ParP[ASP=prog/perf]

Abstracting away from past tense and passive constructions, the TP is now fully specified. So far we have the following nodes within the TP.

- TP = VP
- TP = AuxP > ParP (where '>' = higher than)
- TP = ModP > ProgP /PerfP > ParP[prog/perf]
- TP = ModP > PrerfprogfP > ParP[perfprog] > ParP[prog]

4 Coordination

4.1 Conjunction rules in clauses

- S → S Join S
  - John comes and Mary leaves
- ModP → ModP Join ModP
  - John will come and might bring a cake
- AuxP → AuxP Join AuxP
  - John is walking and is singing
- InfP → InfP Join InfP
  - John will come and bring a cake
- PerfP → PerfP Join PerfP
  - John might have sang and might have won
- ProgP → ProgP Join ProgP
  - John will be singing and will be winning
- PerfprogP → PerfprogP Join PerfprogP
  - John should have been singing and should have been winning
- ParP[ASP=perf] → ParP[ASP=perf] Join ParP[ASP=perf]
  - John has won and left
- ParP[ASP=prog] → ParP[ASP=prog] Join ParP[ASP=prog]
  - John is singing and dancing

- ParP[ASP=prog] → ParP[ASP=prog] Join ParP[ASP=prog]
  - John has been singing and dancing
- ParP[ASP=perfprog] → ParP[ASP=perfprog] Join ParP[ASP=perfprog]
  - John has been singing and been dancing

4.2 Conjunction rules at the noun phrase level

- DP → DP join DP
- AdjP → AdjP Join AdjP
- AdvP → AdvP Join AdvP

The DP must be specified for agreement since "DP and DP" is always plural.

- DP[AGR=[PER=3, NUM=pl]] → DP Join DP

The following is an example for joining pronouns, for example 'you and I' corresponds to 'we' =[PER=1, NUM=pl].

- PRO[AGR=[PER=1, NUM=pl]] → PRO[AGR=[PER=2, NUM=sg]] Join PRO[AGR=[PER=1, NUM=sg]]

In this way, 'first person and second/third person' corresponds to 'we', whereas 'second person and second/third person' corresponds to 'you *pl*'.

5 Clausal and prepositional adjuncts

Unlike adverbs, adjunct clauses and prepositional adjuncts can only occur sentence-initially or sentence-finally.

- S → Sub S | S Sub (where Sub = subordinate clause)
- S → PP S | S PP

6 Comments on internal negation

Our grammar does not capture sentences like the following.

- Mary will bring a not tasty cake
- Mary will bring a not extremely tasty cake

This problem is solved by adding the following rules.

- AdjP → AdvP AdjP
- AdjP → Neg AdjP
- AdvP → Neg AdvP

Conclusion

The greatest challange we faced was proper subcategorization of the Tense Projections VP, AuxP, and ModP as well as developing the correct aspectual projections PropP, PerfP, PerfProgP, and their respective substructures. Joining these projections and expanding the grammar to cover adverbs, adjuncts, and negation was relatively painless once these foundations were established.

Lastly, we wish to note that our grammar can be further extended to include copula *be*, as in 'John is tall', main verb *be*, as in 'John is at the station', as well as main verb *have*, as in 'John has a cat'. This can be done by adding a copula *be,* a 'full' *be*, and a 'full' *have* to the lexicon and introducing the following rules.

- VP $\rightarrow$ BE_cop AdjP | BE_cop DP
- VP $\rightarrow$ BE_full PP
- VP $\rightarrow$ HAVE_full DP

The next step is to distinguish between the infinitive (for the ModP) and the other tensed forms. This is where things become somewhat tedious and for reasons of time we have omitted this part of the grammar.