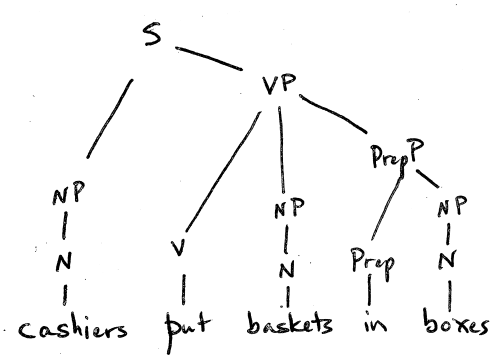
CS 6740/INFO 6300, Fall 2019
CFGs, long-distance dependencies, tree insertion grammar, Sept 5, 2019

A Recall from last lecture

We want grammars that allow us to assign "good" structures to "good sentences":



. . . but don't allow us to assign structures to "bad sentences", like "* they puts baskets from sleep".

Many required *lexically-induced* features ⇒ proliferation of sub-types:

$$\mathrm{VP}_{\text{subject:animateNP,1:puttableNP,2:locationPP}} \rightarrow \mathrm{V}_{\text{subject:animateNP,1:puttableNP,2:locationPP}}\ \mathrm{NP}_{\text{puttable,directobjectformPPlocation}}$$

B Developing CFG parsing intuitions

1. Start symbol is S; decompositions $\mathcal{R}$ is the following set[1] (using "tree notation" rather than "rule notation" $\mathrm{S} \rightarrow \mathrm{R}$), which we'll consider to induce: Terminals $\mathcal{T} = \{\text{no}\}$; nonterminals $\mathcal{N} = \{\mathrm{S}, \mathrm{L}, \mathrm{R}\}$. *Question*: how many parse trees does this CFG assign to "no no no"?

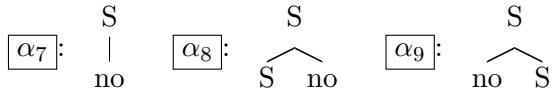
$\boxed{\alpha_1}$: S / R $\quad$ $\boxed{\alpha_2}$: S / L

$\boxed{\alpha_3}$: R / no $\quad$ $\boxed{\alpha_4}$: L / no

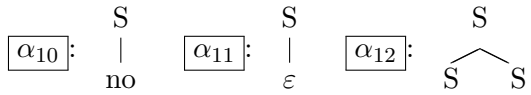
$\boxed{\alpha_5}$: L / (L, no) $\quad$ $\boxed{\alpha_6}$: R / (no, R)

[1] Tree layout using the QobiTree package.

2. *Question*: how many parse trees does the (induced) CFG below assign to "no no no"?



3. A CFG worth knowing about. "ε" is the *empty string*.



How many parses for "no no no"?

Is the set of sentences it assigns structures to the same as in the previous CFG?

C Back to the "category proliferation problem" from last time ("let's be clever engineers")

From the grammar-designer's perspective, we'd like:

4. VP $\rightarrow$ V NP PrepP

- VP agreement $\sqcup$ V agreement
- V's first argument constraints $\sqcup$ NP
- V's second argument constraints $\sqcup$ PrepP

where $\sqcup$ means "unify" ("smallest set of constraints consistent with both operands' constraint sets").

D Handling gaps in question inversion (another type of long-distance dependencies)

5. what do the cashiers put in the boxes?

6. where do the cashiers put the boxes?

7. * the cashiers put boxes

8. * what do the cashiers put the boxes

9. $S_{question} \rightarrow$ NP-WH AUX NP VP

- NP-WH $\sqcup$ VP direct object
- VP's direct object must have gap: $+$
- Subject NP must have gap: $-$
- ... (other agreement constraints)

Add to 4:

- at most one of NP, PrepP can have gap: $-$

10. NP $\rightarrow \varepsilon$

- NP gap $\sqcup$ *yes*

E Tree substitution grammars: extend the domain of locality

$\boxed{\alpha_{put}}$: