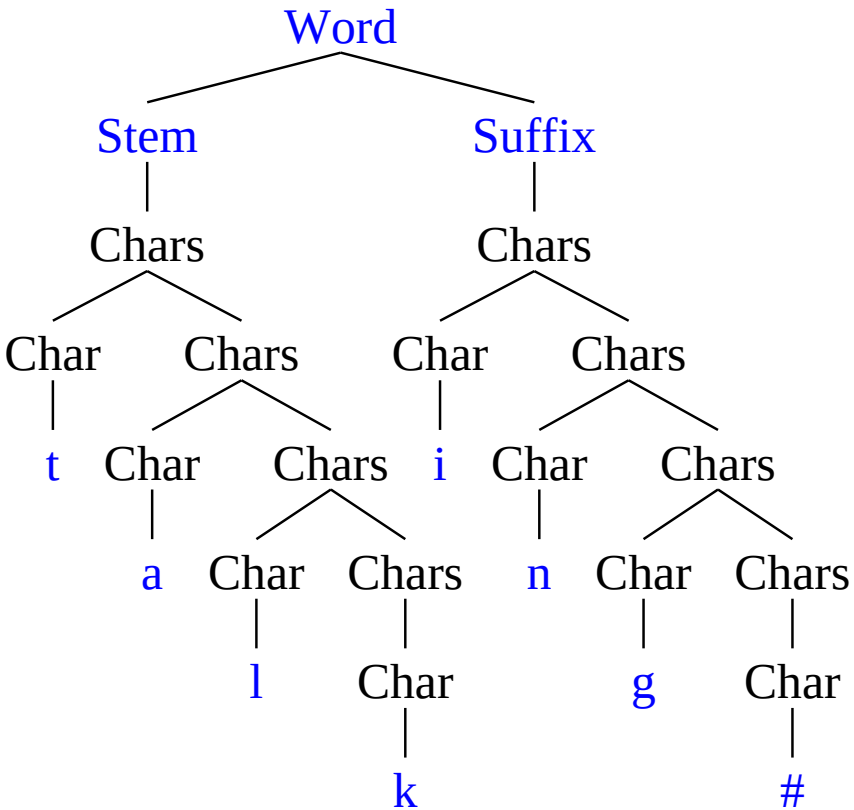
Adaptor grammars – a framework for NPB models

$$
\begin{aligned}
\text{Word} &\rightarrow \text{Stem Suffix} \\
\text{Stem} &\rightarrow \text{Chars} \\
\text{Suffix} &\rightarrow \text{Chars}
\end{aligned}
\qquad
\begin{aligned}
\text{Chars} &\rightarrow \text{Char} \\
\text{Chars} &\rightarrow \text{Char Chars} \\
\text{Char} &\rightarrow \text{a} \mid \text{b} \mid \text{c} \mid \ldots
\end{aligned}
$$



- CFG rules define possible trees
- To generate an *adapted* nonterminal
  - return a previously generated tree t with prob. $\propto n(t)$
  - generate a tree t using CFG rules with prob. $\propto \alpha_X P'(t)$

$$
\begin{aligned}
P_{\text{Word}} &\sim \text{DP}(\alpha_{\text{Word}}, P_{\text{Word}'}) \\
P_{\text{Word}'}(w) &= \sum_{s,t:s\cdot t=w} P_{\text{Stem}}(s) P_{\text{Suffix}}(t) \\
P_{\text{Stem}} &\sim \text{DP}(\alpha_{\text{Stem}}, P_{\text{Stem}'}) \\
P_{\text{Stem}'}(w) &\propto p^{|w|}
\end{aligned}
$$

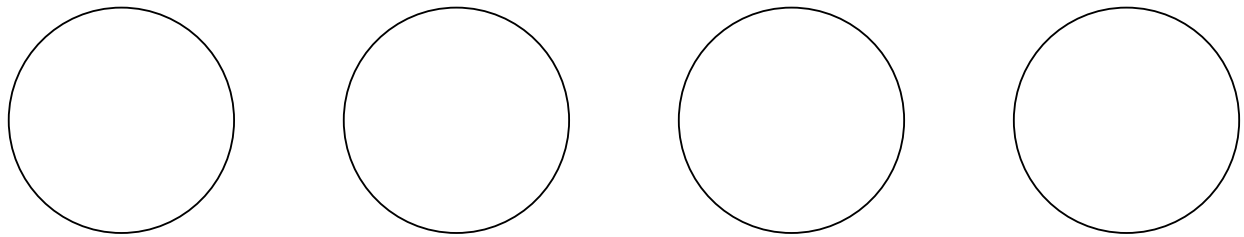
Software for using adaptor grammars

- Intended to ease exploration of Goldwater-style NPB models
  - general enough to express e.g., LDA-ish topic models
- Metropolis-within-Gibbs sampler
- Proposal distribution is *PCFG approximation* based on sample parses of other sentences
  - one production for each cached subtree
  - dynamic programming algorithm for producing proposals
- Software implements Pitman-Yor processes (not just CRPs)
- Sampling PY hyperparameters sometimes improves model accuracy
- *Table label resampling* vastly speeds convergence
- Nonidentifiability of latent states can be a potential problem

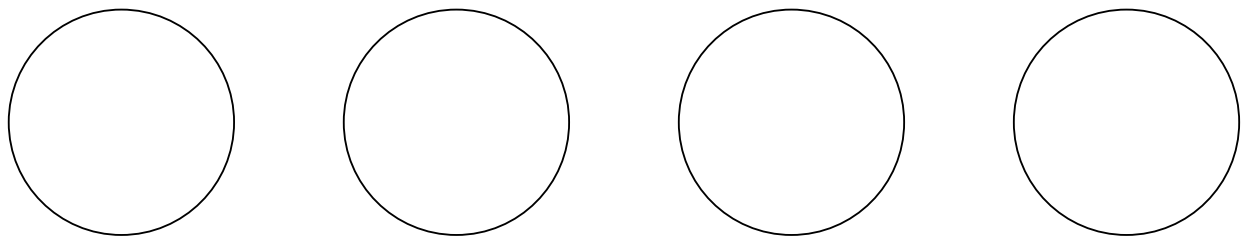
  $\Rightarrow$ evaluate a single sample of parses for all sentences

Adaptor grammar derivation (0)

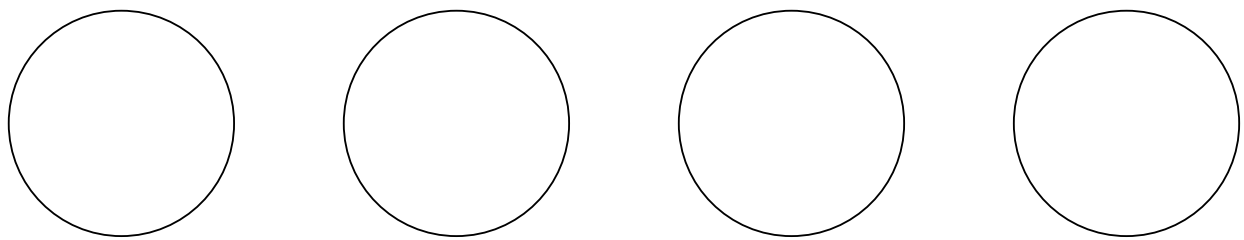
Word restaurant
Word → Stem Suffix



Stem restaurant
Stem → Phoneme*

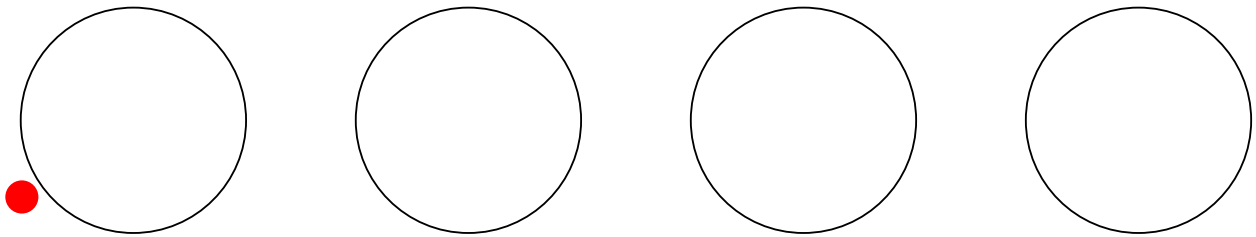


Suffix restaurant
Suffix → Phoneme$^+$

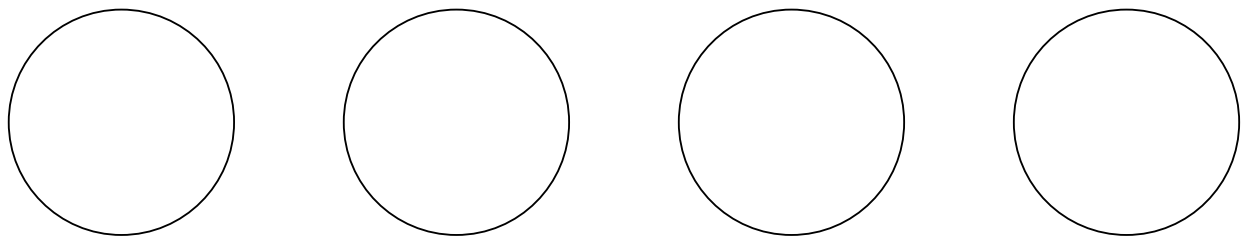


Adaptor grammar derivation (1a)

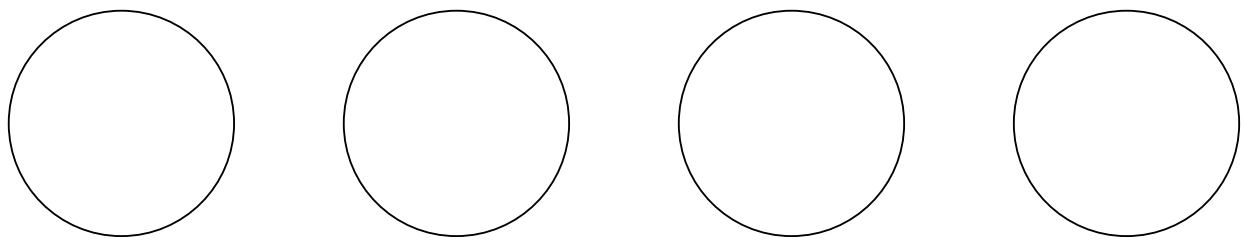
Word restaurant
Word → Stem Suffix



Stem restaurant
Stem → Phoneme*

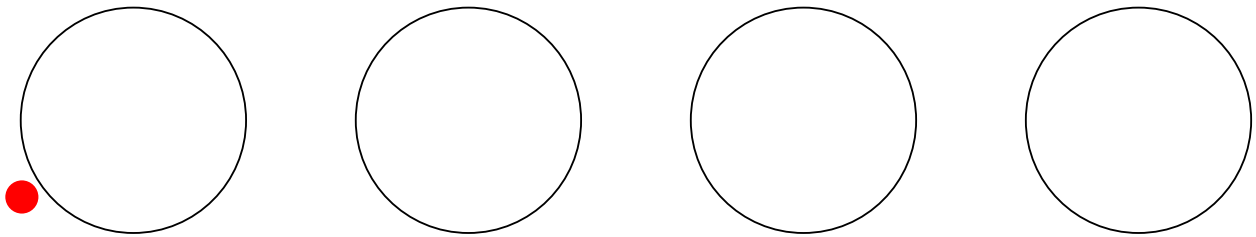


Suffix restaurant
Suffix → Phoneme$^+$

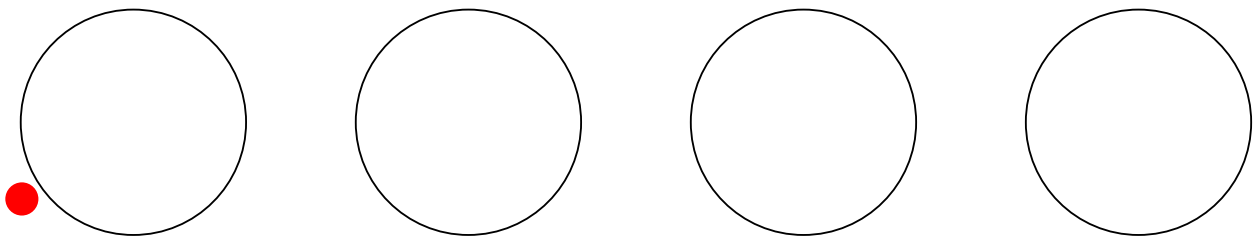


Adaptor grammar derivation (1b)

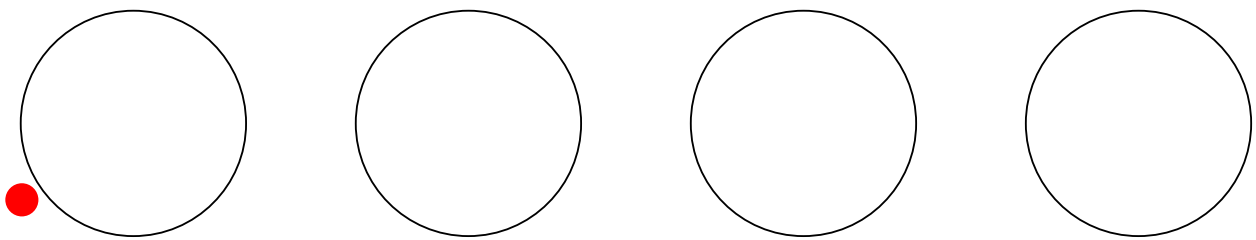
Word restaurant
Word → Stem Suffix



Stem restaurant
Stem → Phoneme*

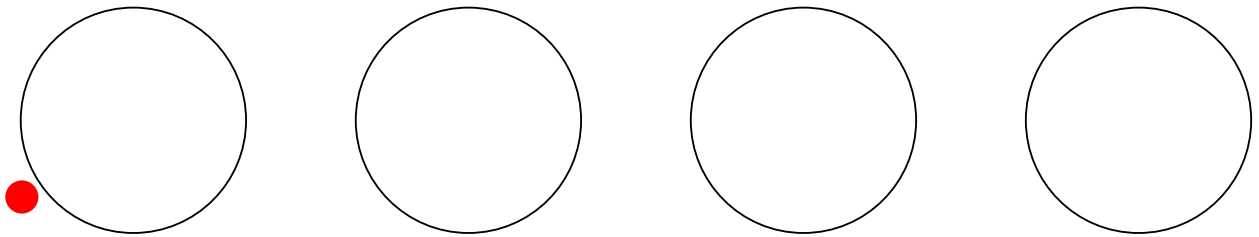


Suffix restaurant
Suffix → Phoneme$^+$

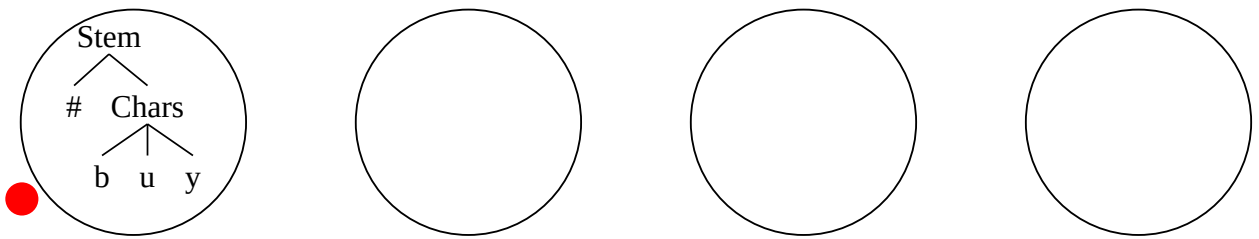


Adaptor grammar derivation (1c)

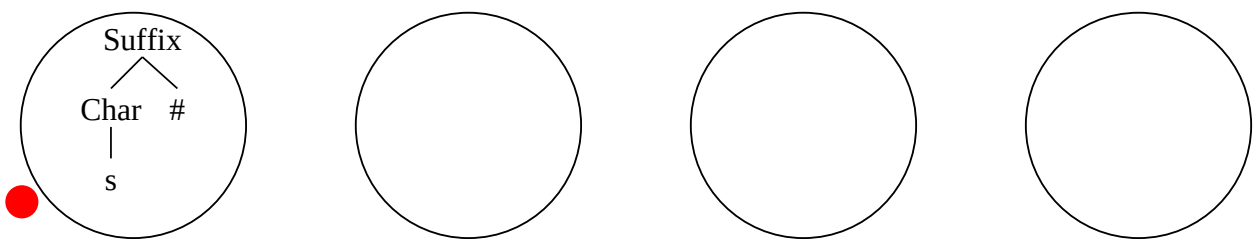
Word restaurant
Word → Stem Suffix



Stem restaurant
Stem → Phoneme⋆

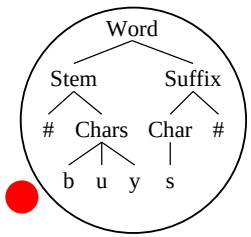


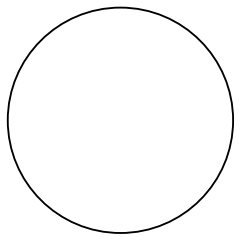
Suffix restaurant
Suffix → Phoneme⁺

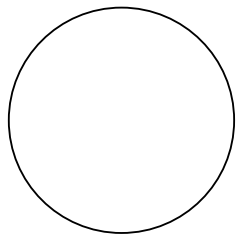


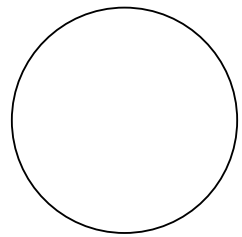
Adaptor grammar derivation (1d)

Word restaurant
Word → Stem Suffix



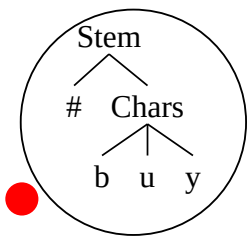


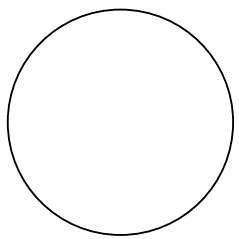


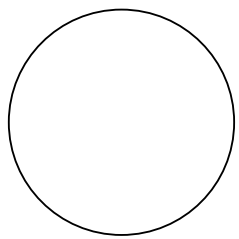


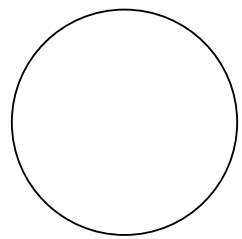
. . .

Stem restaurant
Stem → Phoneme⋆



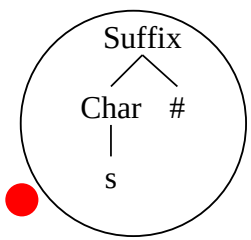


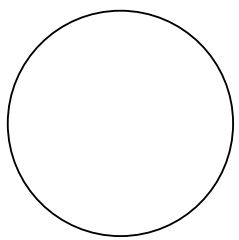


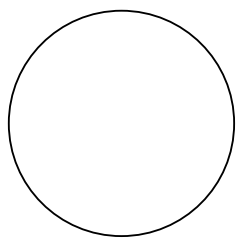


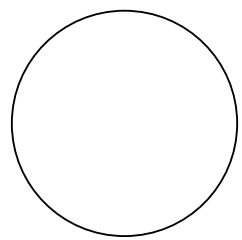
. . .

Suffix restaurant
Suffix → Phoneme+





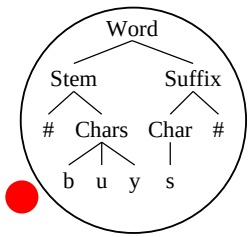


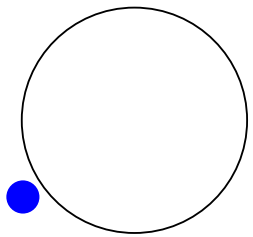


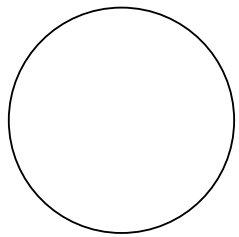
. . .

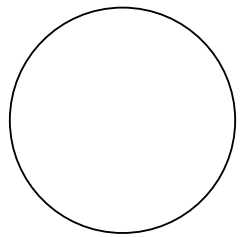
Adaptor grammar derivation (2a)

Word restaurant
Word → Stem Suffix



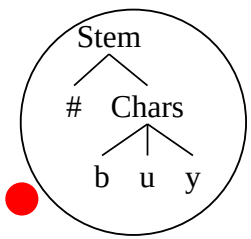


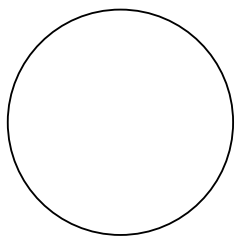


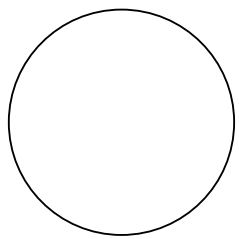


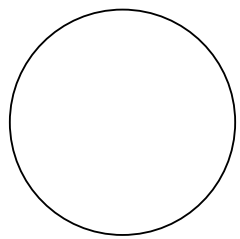
. . .

Stem restaurant
Stem → Phoneme⋆



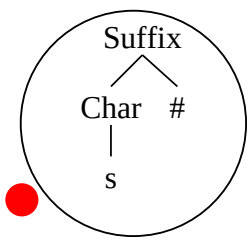


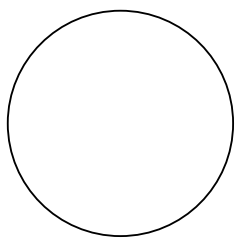


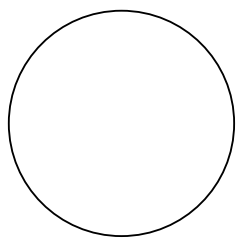


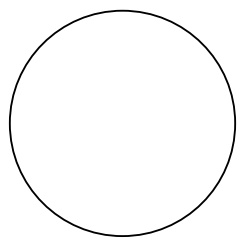
. . .

Suffix restaurant
Suffix → Phoneme+





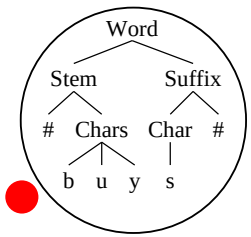


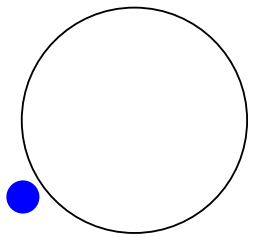


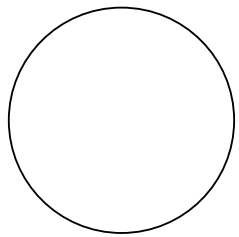
. . .

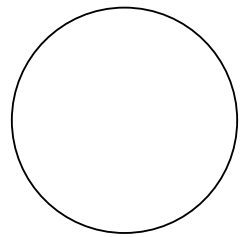
Adaptor grammar derivation (2b)

Word restaurant
Word → Stem Suffix



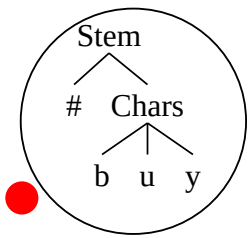


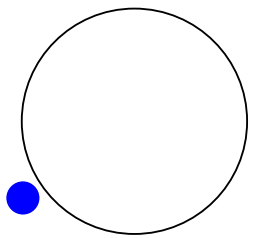


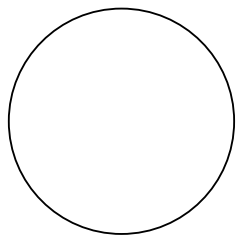


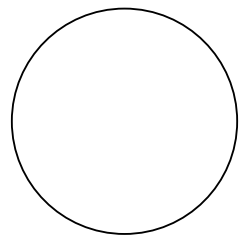
. . .

Stem restaurant
Stem → Phoneme⋆



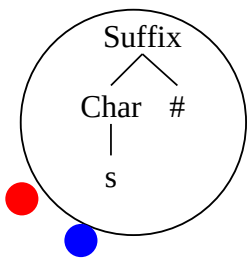


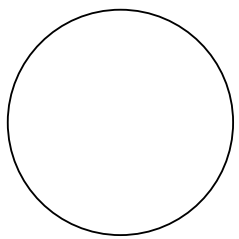


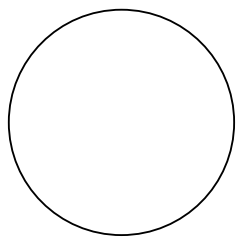


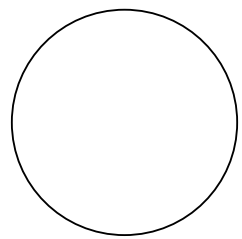
. . .

Suffix restaurant
Suffix → Phoneme+





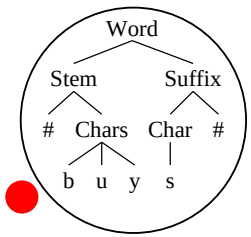


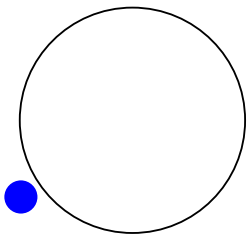


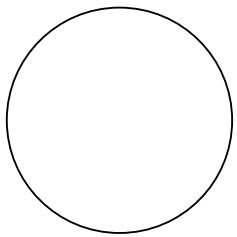
. . .

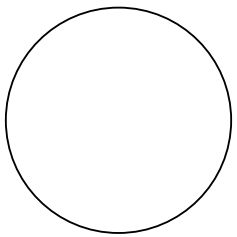
Adaptor grammar derivation (2c)

Word restaurant
Word → Stem Suffix



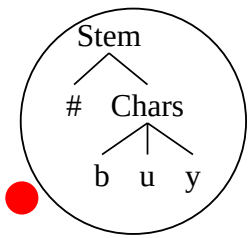


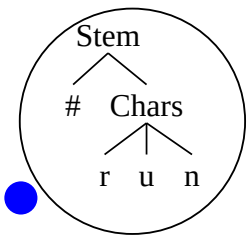


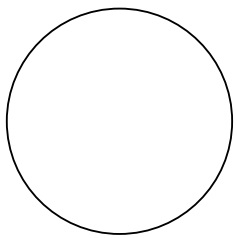


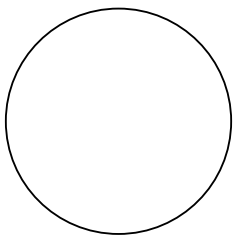
. . .

Stem restaurant
Stem → Phoneme⋆



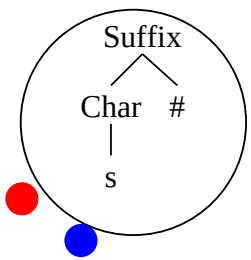


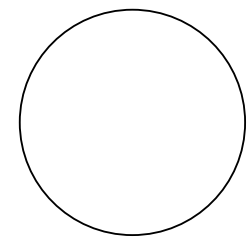


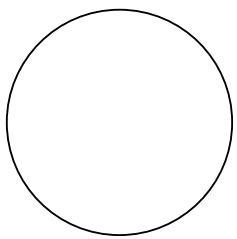


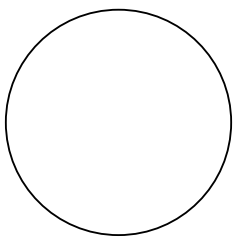
. . .

Suffix restaurant
Suffix → Phoneme⁺





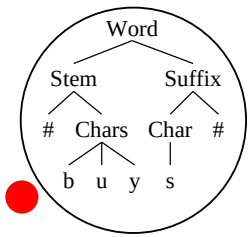


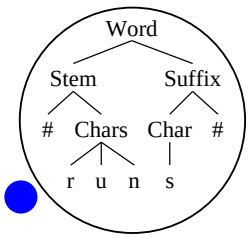


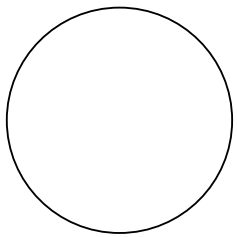
. . .

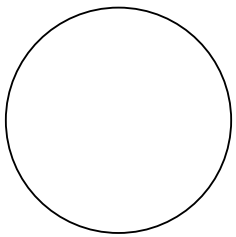
Adaptor grammar derivation (2d)

Word restaurant
Word → Stem Suffix



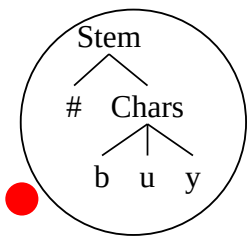


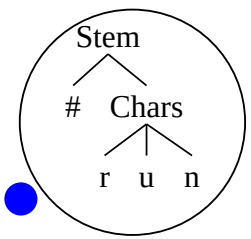


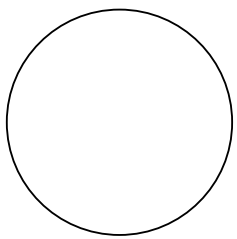


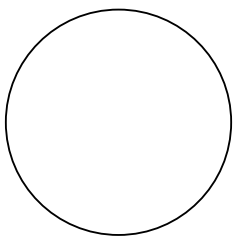
. . .

Stem restaurant
Stem → Phoneme⋆



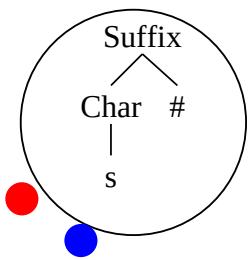


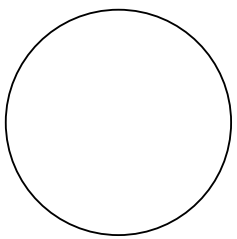


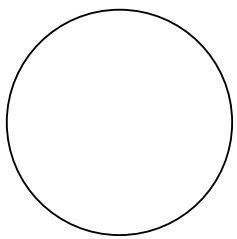


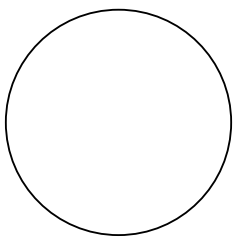
. . .

Suffix restaurant
Suffix → Phoneme⁺





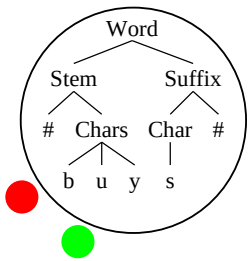


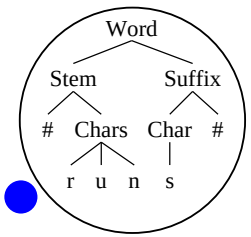


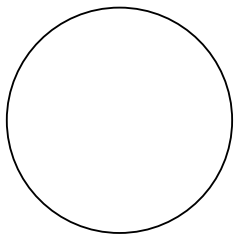
. . .

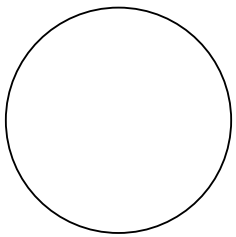
Adaptor grammar derivation (3)

Word restaurant
Word → Stem Suffix



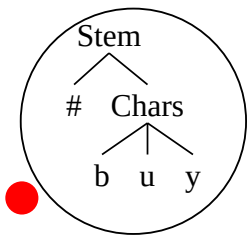


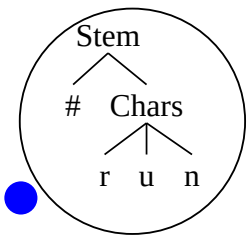


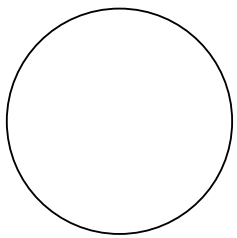


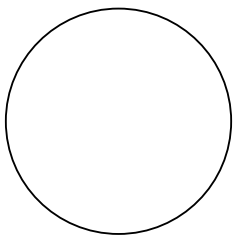
. . .

Stem restaurant
Stem → Phoneme⋆



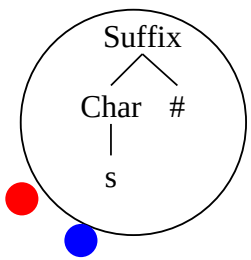


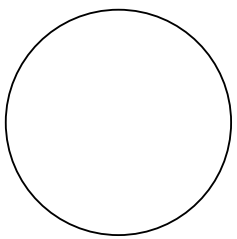


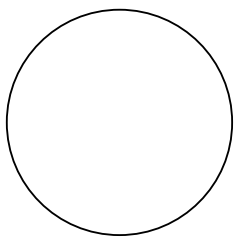


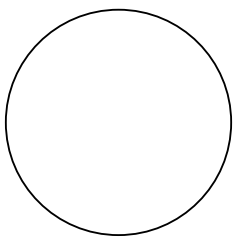
. . .

Suffix restaurant
Suffix → Phoneme⁺





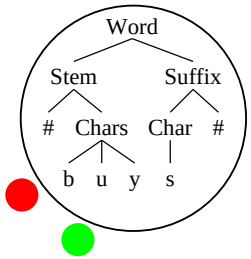


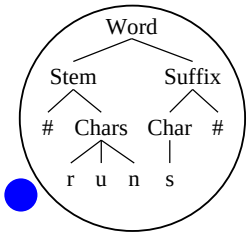


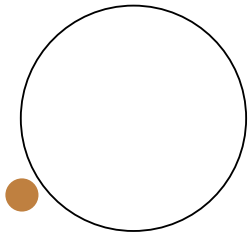
. . .

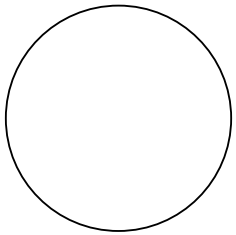
Adaptor grammar derivation (4a)

Word restaurant
Word → Stem Suffix



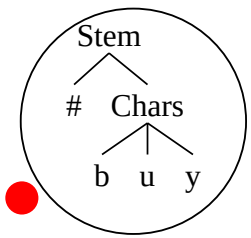


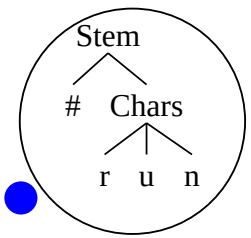


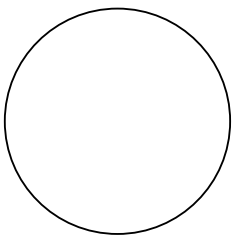


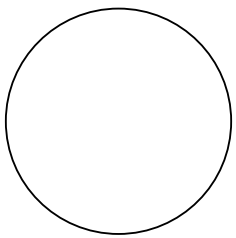
. . .

Stem restaurant
Stem → Phoneme⋆



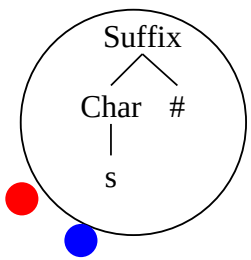


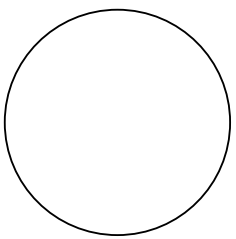


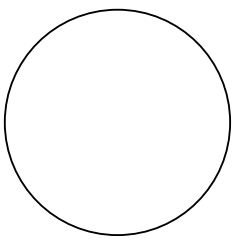


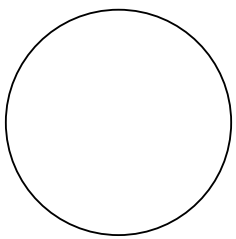
. . .

Suffix restaurant
Suffix → Phoneme⁺





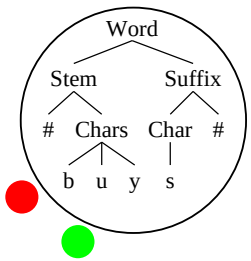


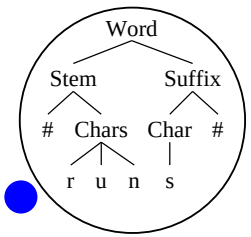


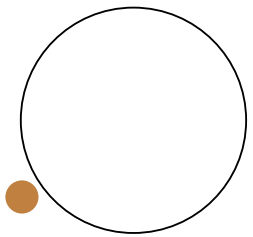
. . .

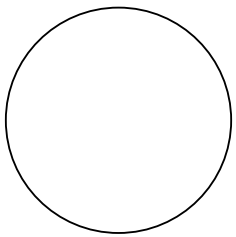
Adaptor grammar derivation (4b)

Word restaurant
Word → Stem Suffix



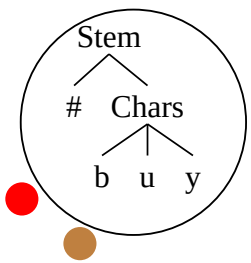


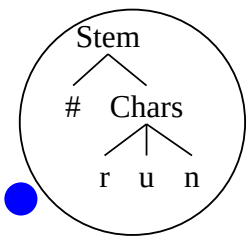


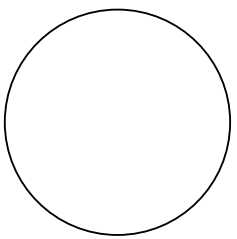


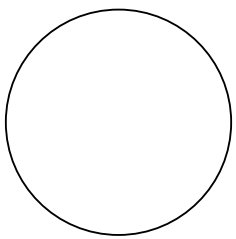
. . .

Stem restaurant
Stem → Phoneme⋆



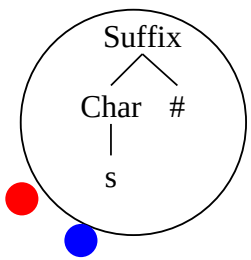


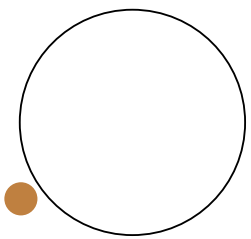


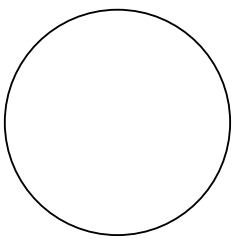


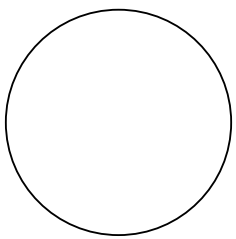
. . .

Suffix restaurant
Suffix → Phoneme+





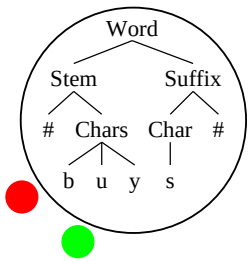


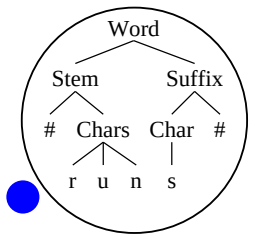


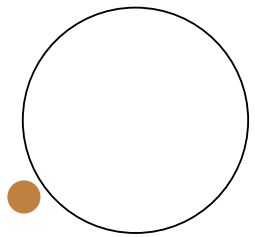
. . .

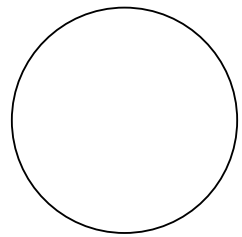
Adaptor grammar derivation (4c)

Word restaurant
Word → Stem Suffix



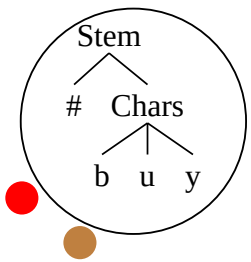


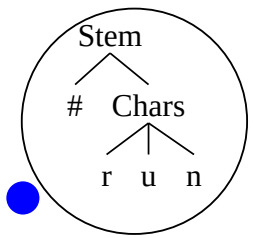


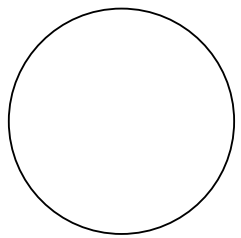


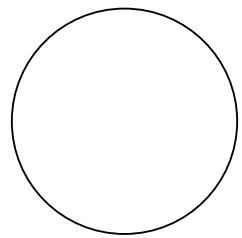
. . .

Stem restaurant
Stem → Phoneme⋆



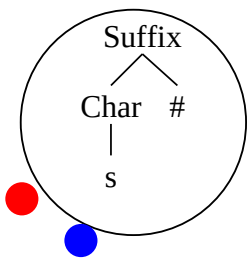


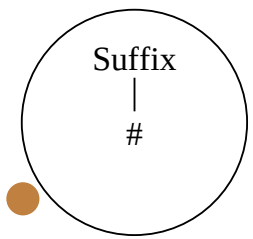


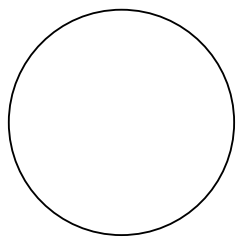


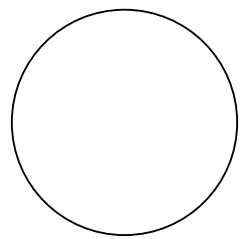
. . .

Suffix restaurant
Suffix → Phoneme+





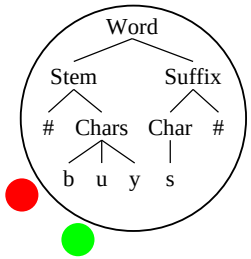


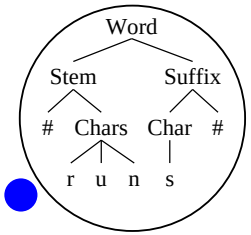


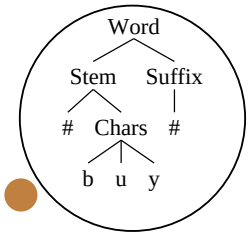
. . .

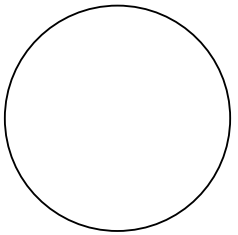
Adaptor grammar derivation (4d)

Word restaurant
Word → Stem Suffix



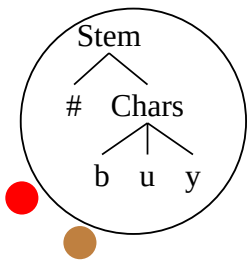


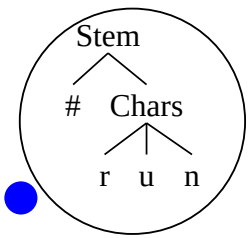


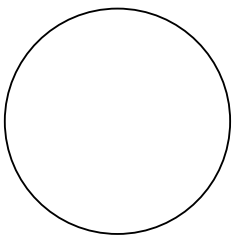


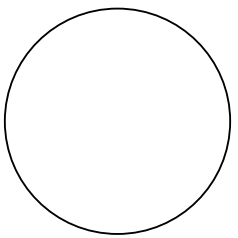
. . .

Stem restaurant
Stem → Phoneme⋆



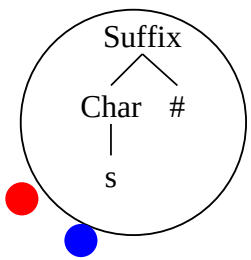


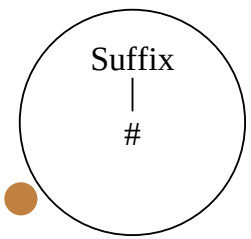


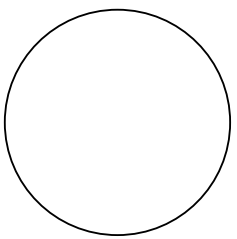


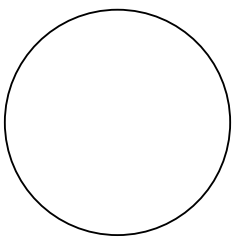
. . .

Suffix restaurant
Suffix → Phoneme⁺







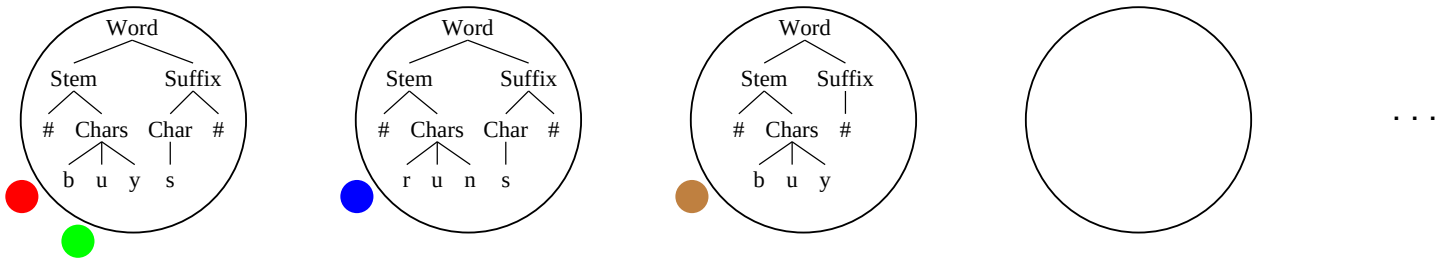


. . .

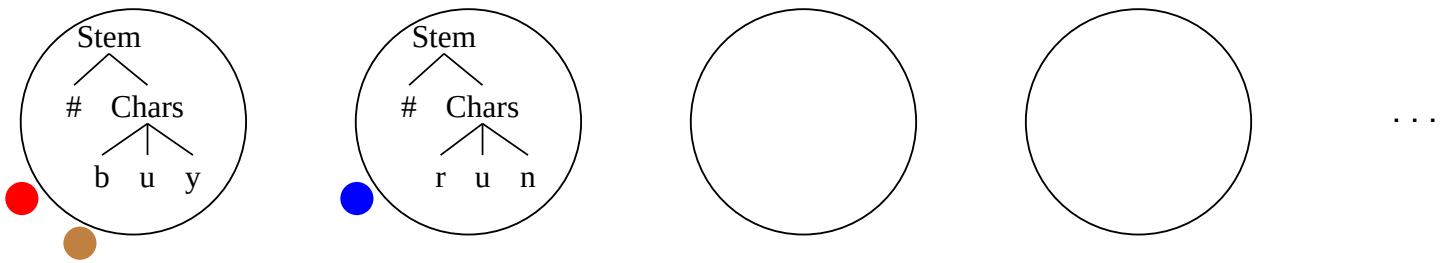
Table label resampling improves mobility

- Metropolis-within-Gibbs: resample T_i given W_i and $\vec{T}_{-i}$
- Table label resampling resamples the labels on each table
  - can change parses for many sentences at once

Word restaurant
Word → Stem Suffix



Stem restaurant
Stem → Phoneme*



Suffix restaurant
Suffix → Phoneme$^+$

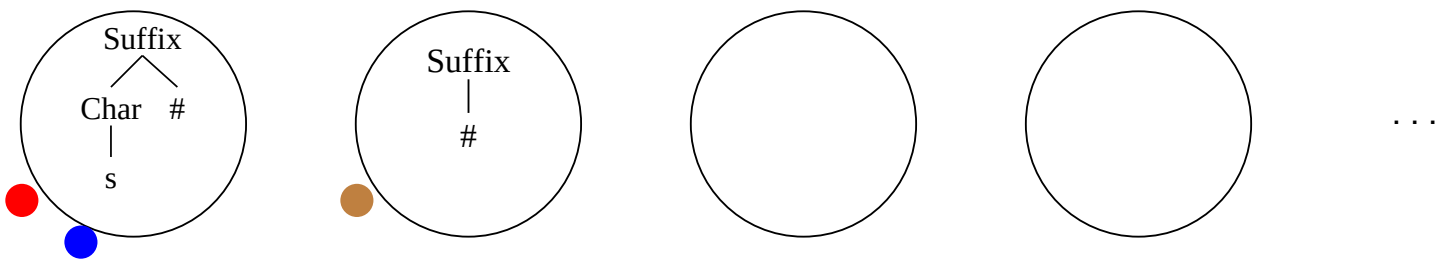


Table label resampling with Colloc grammar