Chomsky-Schützenberger Characterisation for Multiple Context-Free Languages

Tobias Denkinger

Chair of Foundations of Programming
Institute of Theoretical Computer Science
Technische Universität Dresden

2014–11–27

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R) .$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ , \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

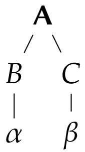
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ ,\quad B \xrightarrow{r_2} \alpha,\quad C \xrightarrow{r_3} \beta$$



Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

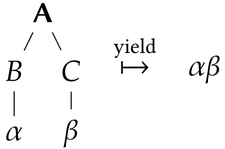
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ ,\quad B \xrightarrow{r_2} \alpha,\quad C \xrightarrow{r_3} \beta$$



Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

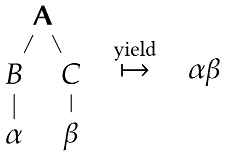
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ , \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$



$$D\colon\ A \to \alpha\ , \quad A \to \beta\ , \quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ , \quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A\}^1_{r_2}\ , \quad A \to \{^1_{r_3} A\}^1_{r_3}$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

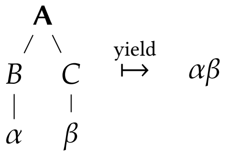
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

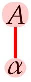
$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ , \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$



$$D\colon A \to \alpha,\ A \to \beta,\ A \to AA\ ,$$

$$A \to \{^1_{r_1} A \}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A \}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A \}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A \}^1_{r_3}$$



Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

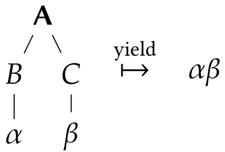
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R) \; .$$

Example (Hulden 2011)

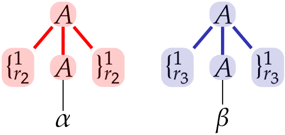
$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C \;, \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$



$$D\colon \; A \to \alpha \;, \quad A \to \beta \;, \quad A \to AA \;,$$

$$A \to \{^1_{r_1} A \}^1_{r_1} \;, \quad A \to \{^2_{r_1} A \}^2_{r_1} \;,$$

$$A \to \{^1_{r_2} A \}^1_{r_2} \;, \quad A \to \{^1_{r_3} A \}^1_{r_3}$$



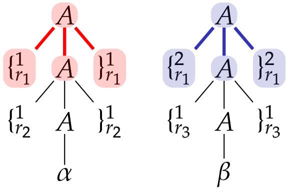
Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

For every CFL L there are

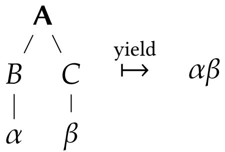
- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$



Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ , \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$



$$D\colon\ A \to \alpha\ ,\quad A \to \beta\ ,\quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A\}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A\}^1_{r_3}$$

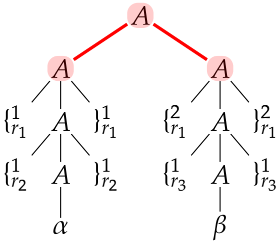
Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

For every CFL L there are

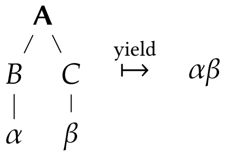
- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$



Example (Hulden 2011)

$$L\colon\ \mathbf{A} \xrightarrow{r_1} B\ C\ ,\quad B \xrightarrow{r_2} \alpha,\quad C \xrightarrow{r_3} \beta$$



$$D\colon\ A \to \alpha\ ,\quad A \to \beta\ ,\quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A\}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A\}^1_{r_3}$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

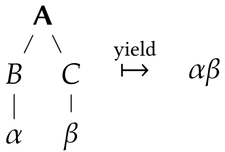
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R) \;.$$

Example (Hulden 2011)

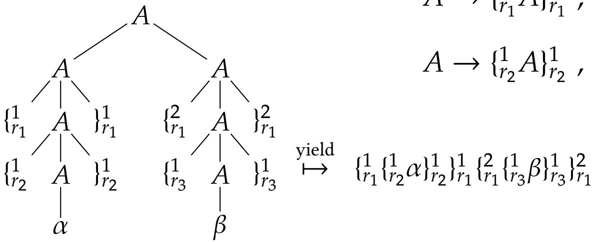
$$L\colon\; \mathbf{A} \xrightarrow{r_1} B\; C\;,\quad B \xrightarrow{r_2} \alpha,\quad C \xrightarrow{r_3} \beta$$



$$D\colon\; A \to \alpha\;,\quad A \to \beta\;,\quad A \to AA\;,$$

$$A \to \{^1_{r_1} A \}^1_{r_1}\;,\quad A \to \{^2_{r_1} A \}^2_{r_1}\;,$$

$$A \to \{^1_{r_2} A \}^1_{r_2}\;,\quad A \to \{^1_{r_3} A \}^1_{r_3}$$



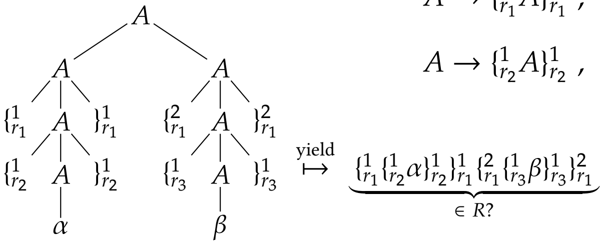
Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

For every CFL L there are

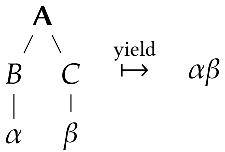
- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$



Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ ,\quad B \xrightarrow{r_2} \alpha,\quad C \xrightarrow{r_3} \beta$$



$$D\colon\ A \to \alpha\ ,\quad A \to \beta\ ,\quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A\}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A\}^1_{r_3}$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

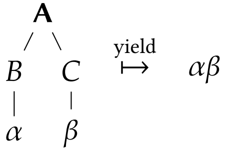
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

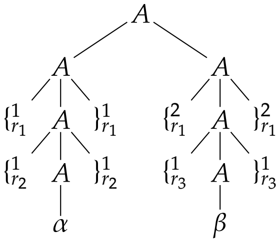
$$L\colon \mathbf{A} \xrightarrow{r_1} B\,C\,, \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$



$$D\colon\ A \to \alpha\ ,\quad A \to \beta\ ,\quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A\}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A\}^1_{r_3}$$



$$\overset{\text{yield}}{\mapsto}\ \underbrace{\{^1_{r_1}\{^1_{r_2}\alpha\}^1_{r_2}\}^1_{r_1}\{^2_{r_1}\{^1_{r_3}\beta\}^1_{r_3}\}^2_{r_1}}_{\in\ R?}$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

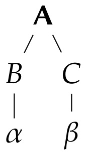
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

$$L\colon \mathbf{A} \xrightarrow{r_1} B\ C\ , \quad B \xrightarrow{r_2} \alpha, \quad C \xrightarrow{r_3} \beta$$

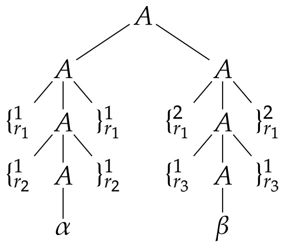


$$\overset{\text{yield}}{\mapsto}\ \alpha\beta$$

$$D\colon\ A \to \alpha\ ,\quad A \to \beta\ ,\quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

$$A \to \{^1_{r_2} A\}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A\}^1_{r_3}$$



$$\overset{\text{yield}}{\mapsto}\ \underbrace{\{^1_{r_1}\{^1_{r_2}\alpha\}^1_{r_2}\}^1_{r_1}\{^2_{r_1}\{^1_{r_3}\beta\}^1_{r_3}\}^2_{r_1}}_{\in R?}\ \overset{h}{\mapsto}\ \alpha\beta$$

Classical result

Theorem (Chomsky and Schützenberger 1963, Prop. 2)

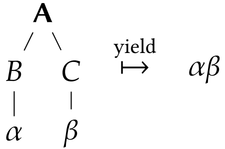
For every CFL L there are

- a Dyck language D,
- a regular language R, and
- a homomorphism h s.t.

$$L = h(D \cap R)\ .$$

Example (Hulden 2011)

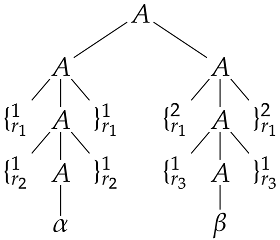
$$L\colon\ \mathbf{A} \xrightarrow{r_1} B\ C\ ,\quad B \xrightarrow{r_2} \alpha,\quad C \xrightarrow{r_3} \beta$$



$$D\colon\ A \to \alpha\ ,\quad A \to \beta\ ,\quad A \to AA\ ,$$

$$A \to \{^1_{r_1} A\}^1_{r_1}\ ,\quad A \to \{^2_{r_1} A\}^2_{r_1}\ ,$$

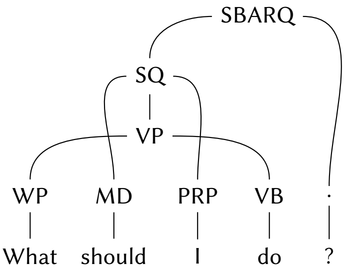
$$A \to \{^1_{r_2} A\}^1_{r_2}\ ,\quad A \to \{^1_{r_3} A\}^1_{r_3}$$



$$\overset{\text{yield}}{\mapsto}\ \underbrace{\{^1_{r_1}\{^1_{r_2}\alpha\}^1_{r_2}\}^1_{r_1}\{^2_{r_1}\{^1_{r_3}\beta\}^1_{r_3}\}^2_{r_1}}_{\in R?}\ \overset{h}{\mapsto}\ \alpha\beta$$

Non-projective trees

Example (Evang and Kallmeyer 2011)



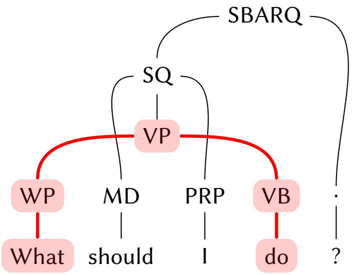
1

2

Non-projective trees

gaps / crossing edges

Example (Evang and Kallmeyer 2011)

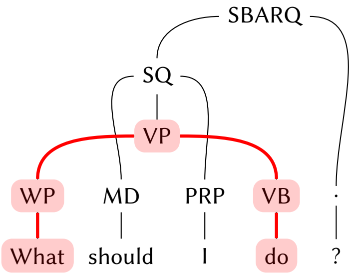


1

2

Non-projective trees

Example (Evang and Kallmeyer 2011)



gaps / crossing edges

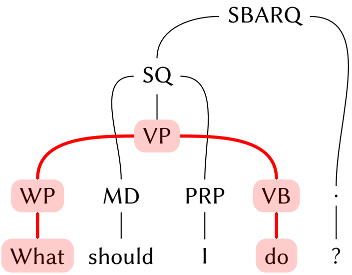
- not representable with CFG

1

2

Non-projective trees

Example (Evang and Kallmeyer 2011)



gaps / crossing edges

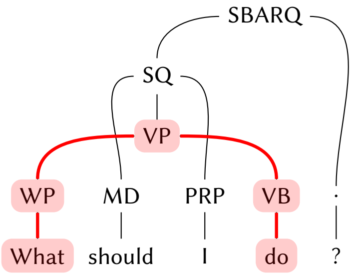
- not representable with CFG
- occur in natural language tree banks (Maier and Søgaard 2008)

1

2

Non-projective trees

Example (Evang and Kallmeyer 2011)



gaps / crossing edges

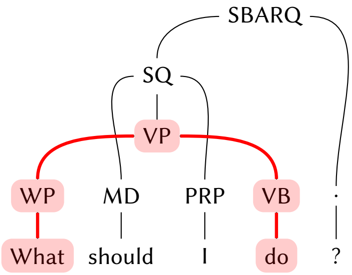
- not representable with CFG
- occur in natural language tree banks (Maier and Søgaard 2008)

	proj.	non-proj.
NeGra[1]	72.44%	27.56%
TIGER[2]	72.46%	27.54%

[1]approx. 20 000 trees
[2]approx. 50 000 trees

Non-projective trees

Example (Evang and Kallmeyer 2011)



gaps / crossing edges

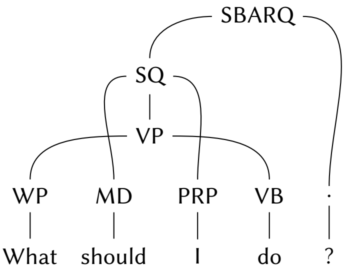
- not representable with CFG
- occur in natural language tree banks (Maier and Søgaard 2008)

	proj.	non-proj.
NeGra[1]	72.44%	27.56%
TIGER[2]	72.46%	27.54%

[1]approx. 20 000 trees
[2]approx. 50 000 trees

Non-projective trees

Example (Evang and Kallmeyer 2011)



gaps / crossing edges

- not representable with CFG
- occur in natural language tree banks (Maier and Søgaard 2008)

	proj.	non-proj.
NeGra[1]	72.44%	27.56%
TIGER[2]	72.46%	27.54%

[1] approx. 20 000 trees

[2] approx. 50 000 trees

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^{m}$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^{m}$$

Example

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1 \pi_2^1, \beta \pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^{m}$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\big((\alpha\gamma), (\alpha, \beta)\big)$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\big((\alpha\gamma), (\alpha, \beta)\big) =$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\Big((\underbrace{\alpha\gamma}_{\pi_1^1}), (\underbrace{\alpha}_{\pi_2^1}, \underbrace{\beta}_{\pi_2^2})\Big) =$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\Big((\underbrace{\alpha\gamma}_{\pi_1^1}), (\underbrace{\alpha}_{\pi_2^1}, \underbrace{\beta}_{\pi_2^2})\Big) = (\qquad , \quad)$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1 \pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1 \pi_2^1, \beta\pi_2^2\right]\Big(\big(\underbrace{\alpha\gamma}_{\pi_1^1}\big), \big(\underbrace{\alpha}_{\pi_2^1}, \underbrace{\beta}_{\pi_2^2}\big)\Big) = (\qquad, \beta\ \ )$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\Big((\underbrace{\alpha\gamma}_{\pi_1^1}), (\underbrace{\alpha}_{\pi_2^1}, \underbrace{\beta}_{\pi_2^2})\Big) = (\alpha\gamma\ \ , \beta\ \ )$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

$$\left[\pi_1^1 \pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1 \pi_2^1, \beta\pi_2^2\right]\Big((\underbrace{\alpha\gamma}_{\pi_1^1}), (\underbrace{\alpha}_{\pi_2^1}, \underbrace{\beta}_{\pi_2^2}) \Big) = (\alpha\gamma\alpha, \beta\beta)$$

Word-tuple functions

a k-ary word-tuple function

$$f\colon (\Sigma^*)^{m_1} \times \cdots \times (\Sigma^*)^{m_k} \to (\Sigma^*)^m$$

Example

1st component of the 2nd argument

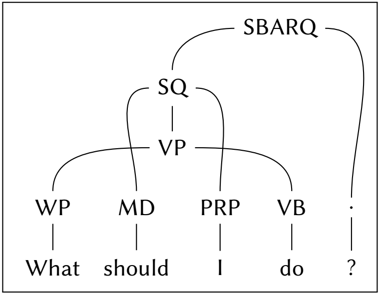
$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\colon (\Sigma^*)^1 \times (\Sigma^*)^2 \to (\Sigma^*)^2$$

$$\left[\pi_1^1\pi_2^1, \beta\pi_2^2\right]\Big((\underbrace{\alpha\gamma}_{\pi_1^1}), (\underbrace{\alpha}_{\pi_2^1}, \underbrace{\beta}_{\pi_2^2})\Big) = (\alpha\gamma\alpha, \beta\beta)$$

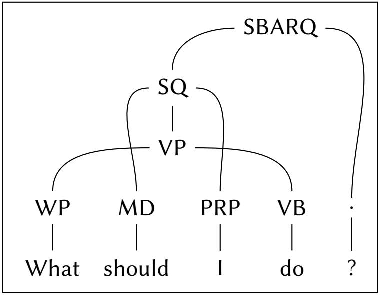
For MCFG: only *linear* word-tuple functions

does not copy argument components

Multiple context-free grammars (MCFGs)

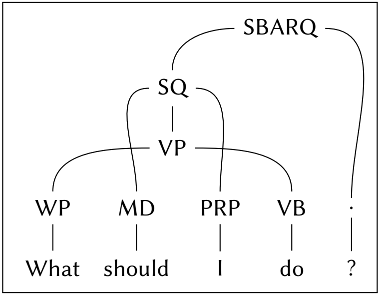


Multiple context-free grammars (MCFGs)



- gaps $\rightsquigarrow$ commas

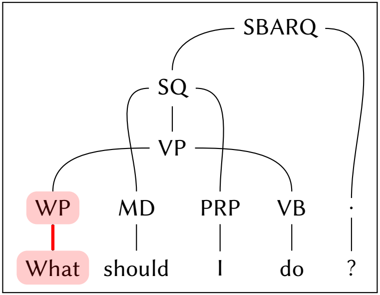
Multiple context-free grammars (MCFGs)



- gaps ⇝ commas

MCFG generating the tree:

Multiple context-free grammars (MCFGs)

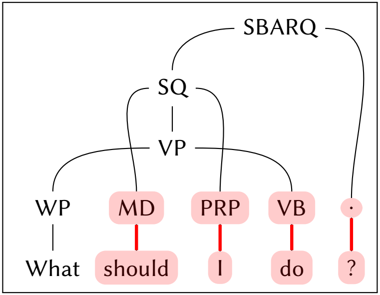


- gaps ⇝ commas

MCFG generating the tree:

WP → [What]

Multiple context-free grammars (MCFGs)



MCFG generating the tree:

WP → [What]

MD → [should]

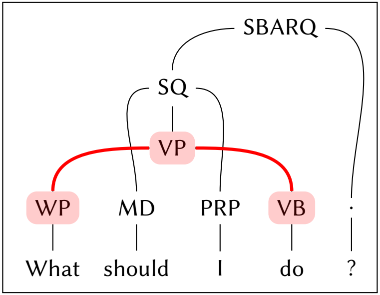
PRP → [I]

VB → [do]

. → [?]

- gaps ⇝ commas

Multiple context-free grammars (MCFGs)

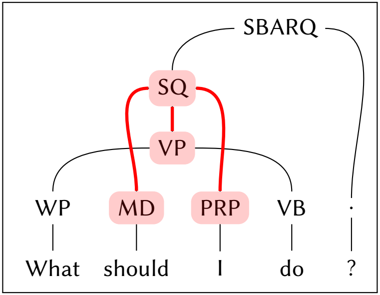


- gaps ⇝ commas

MCFG generating the tree:

$$\mathrm{WP} \rightarrow [\mathrm{What}]$$
$$\mathrm{MD} \rightarrow [\mathrm{should}]$$
$$\mathrm{PRP} \rightarrow [\mathrm{I}]$$
$$\mathrm{VB} \rightarrow [\mathrm{do}]$$
$$. \rightarrow [?]$$
$$\mathrm{VP} \rightarrow [\pi_1^1, \pi_2^1](\mathrm{WP}, \mathrm{VB})$$

Multiple context-free grammars (MCFGs)

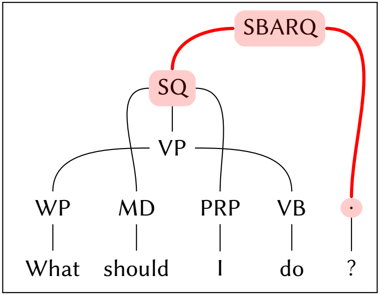


MCFG generating the tree:

$$\mathrm{WP} \rightarrow [\mathrm{What}]$$
$$\mathrm{MD} \rightarrow [\mathrm{should}]$$
$$\mathrm{PRP} \rightarrow [\mathrm{I}]$$
$$\mathrm{VB} \rightarrow [\mathrm{do}]$$
$$. \rightarrow [?]$$
$$\mathrm{VP} \rightarrow [\pi_1^1, \pi_2^1](\mathrm{WP}, \mathrm{VB})$$
$$\mathrm{SQ} \rightarrow [\pi_2^1 \pi_1^1 \pi_3^1 \pi_2^2](\mathrm{MD}, \mathrm{VP}, \mathrm{PRP})$$

- gaps $\rightsquigarrow$ commas

Multiple context-free grammars (MCFGs)

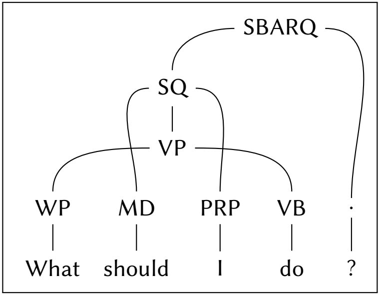


- gaps $\rightsquigarrow$ commas

MCFG generating the tree:

$$\begin{aligned}
\mathrm{WP} &\rightarrow [\mathrm{What}] \\
\mathrm{MD} &\rightarrow [\mathrm{should}] \\
\mathrm{PRP} &\rightarrow [\mathrm{I}] \\
\mathrm{VB} &\rightarrow [\mathrm{do}] \\
. &\rightarrow [?] \\
\mathrm{VP} &\rightarrow [\pi_1^1, \pi_2^1](\mathrm{WP}, \mathrm{VB}) \\
\mathrm{SQ} &\rightarrow [\pi_2^1 \pi_1^1 \pi_3^1 \pi_2^2](\mathrm{MD}, \mathrm{VP}, \mathrm{PRP}) \\
\mathbf{SBARQ} &\rightarrow [\pi_1^1 \pi_2^1](\mathrm{SQ}, .)
\end{aligned}$$

Multiple context-free grammars (MCFGs)



- gaps $\rightsquigarrow$ commas

MCFG generating the tree:

$$\mathrm{WP} \to [\text{What}]$$
$$\mathrm{MD} \to [\text{should}]$$
$$\mathrm{PRP} \to [\text{I}]$$
$$\mathrm{VB} \to [\text{do}]$$
$$. \to [?]$$
$$\mathrm{VP} \to [\pi_1^1, \pi_2^1](\mathrm{WP}, \mathrm{VB})$$
$$\mathrm{SQ} \to [\pi_2^1 \pi_1^1 \pi_3^1 \pi_2^2](\mathrm{MD}, \mathrm{VP}, \mathrm{PRP})$$
$$\mathbf{SBARQ} \to [\pi_1^1 \pi_2^1](\mathrm{SQ}, .)$$

Chomsky-Schützenberger for MCFG

Theorem

For every multiple context-free language $\mathscr{L}$ there are

Chomsky-Schützenberger for MCFG

Theorem

For every multiple context-free language $\mathscr{L}$ there are

- a multiple Dyck language $\mathscr{D}$,

Chomsky-Schützenberger for MCFG

Theorem

For every multiple context-free language $\mathscr{L}$ there are

- a multiple Dyck language $\mathscr{D}$,
- a regular language R, and

Chomsky-Schützenberger for MCFG

Theorem

For every multiple context-free language $\mathscr{L}$ there are

- a multiple Dyck language $\mathscr{D}$,
- a regular language R, and
- a homomorphism h such that

Chomsky-Schützenberger for MCFG

Theorem

For every multiple context-free language $\mathscr{L}$ there are

- a multiple Dyck language $\mathscr{D}$,
- a regular language R, and
- a homomorphism h such that

$$\mathscr{L} = h(\mathscr{D} \cap R) .$$

Bracket alphabet

Two types of brackets: rule $r\colon A \to [u_1, \ldots, u_m](A_1, \ldots, A_k)$

Bracket alphabet

Two types of brackets: rule $r\colon A \to [u_1, \ldots, u_m](A_1, \ldots, A_k)$

(i) $\{_r^i w\}_r^i$ with $1 \le i \le m$

Bracket alphabet

Two types of brackets: rule $r\colon A \to [u_1, \dots, u_m](A_1, \dots, A_k)$

(i) $\{_r^i w\}_r^i$ with $1 \leq i \leq m$
$\rightsquigarrow h(w)$ is derived in u_i

Bracket alphabet

Two types of brackets: rule $r\colon A \to [u_1, \dots, u_m](A_1, \dots, A_k)$

(i) ${}^{i}_{r}\{ w \}^{i}_{r}$ with $1 \le i \le m$
$\rightsquigarrow h(w)$ is derived in u_i

(ii) ${}^{i,j}_{r}\{ w \}^{i,j}_{r}$ with $1 \le i \le m$ and $1 \le j \le |u_i|$

Bracket alphabet

Two types of brackets: rule $r\colon A \to [u_1, ..., u_m](A_1, ..., A_k)$

(i) $\{^i_r w\}^i_r$ with $1 \leq i \leq m$
$\rightsquigarrow h(w)$ is derived in u_i

(ii) $\{^{i,j}_r w\}^{i,j}_r$ with $1 \leq i \leq m$ and $1 \leq j \leq |u_i|$
$\rightsquigarrow h(w)$ is derived in the j-th symbol of u_i

Multiple Dyck language

Multiple Dyck language

- given by an MCFG

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{^i_r A\}^i_r$

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{^i_r A\}^i_r$

multiple Dyck language

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{^i_r A\}^i_r$

multiple Dyck language

- $A_1 \to [\sigma]$

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{_r^i A\}_r^i$

multiple Dyck language

- $A_1 \to [\sigma]$
- $A_k \to f(A_\ell, A_m)$

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{^i_r A\}^i_r$

multiple Dyck language

- $A_1 \to [\sigma]$
- $A_k \to f(A_\ell, A_m)$
- $A_k \to [\{^1_r \pi^1_1\}^1_r, \ldots, \{^k_r \pi^k_1\}^k_r](A_k)$

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{^i_r A\}^i_r$

multiple Dyck language

- $A_1 \to [\sigma]$
- $A_k \to f(A_\ell, A_m)$
- $A_k \to [\{^1_r \pi^1_1\}^1_r, \ldots, \{^k_r \pi^k_1\}^k_r](A_k)$
- $A_k \to [\{^{i_1,j_1}_r \pi^1_1\}^{i_1,j_1}_r, \ldots, \{^{i_k,j_k}_r \pi^k_1\}^{i_k,j_k}_r](A_k)$

Multiple Dyck language

- given by an MCFG
- similar to classical *Dyck language*

Dyck language

- $A \to \sigma$
- $A \to AA$
- $A \to \{^i_r A\}^i_r$

multiple Dyck language

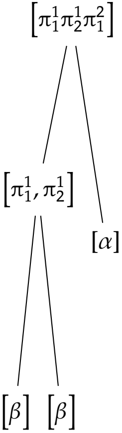
- $A_1 \to [\sigma]$
- $A_k \to f(A_\ell, A_m)$
- $A_k \to [\{^1_r \pi^1_1\}^1_r, \ldots, \{^k_r \pi^k_1\}^k_r](A_k)$
- $A_k \to [\{^{i_1,j_1}_r \pi^1_1\}^{i_1,j_1}_r, \ldots, \{^{i_k,j_k}_r \pi^k_1\}^{i_k,j_k}_r](A_k)$

Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi^1_1\pi^1_2\pi^2_1](A,B), \quad A \xrightarrow{r_2} [\pi^1_1,\pi^1_2](C,C), \quad B \xrightarrow{r_3} [\alpha], \quad C \xrightarrow{r_4} [\beta]$$

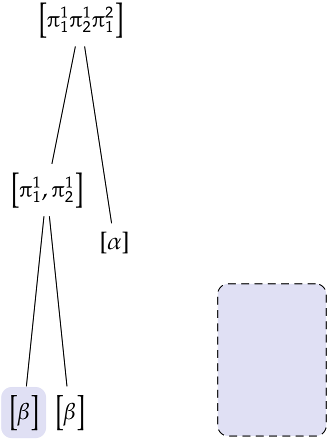
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$



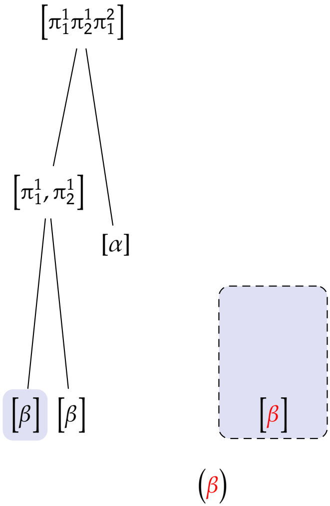
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$



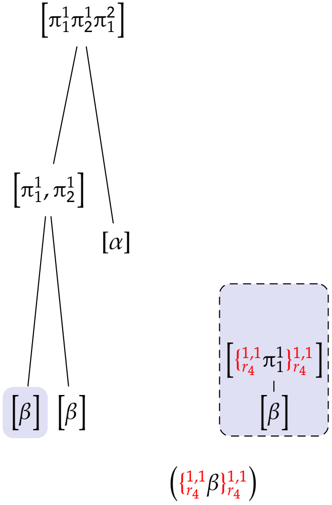
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$



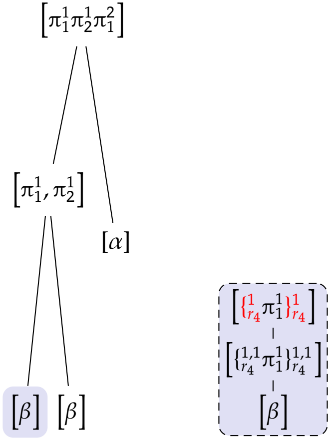
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$



Multiple Dyck derivation

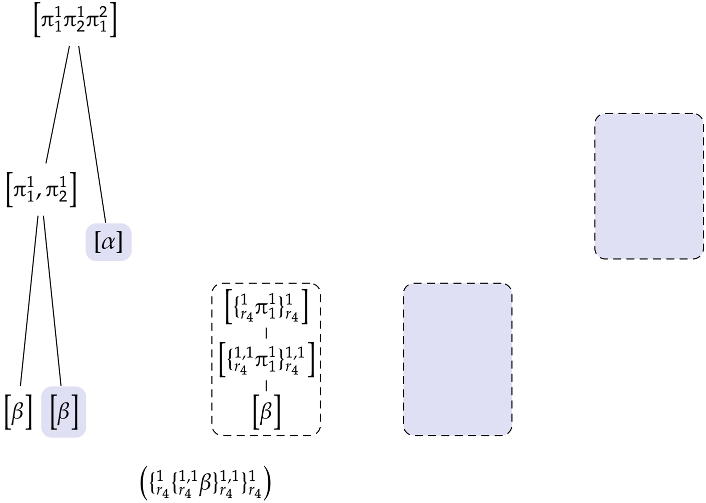
$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$



$$\left(\{_{r_4}^{1}\{_{r_4}^{1,1}\beta\}_{r_4}^{1,1}\}_{r_4}^{1}\right)$$

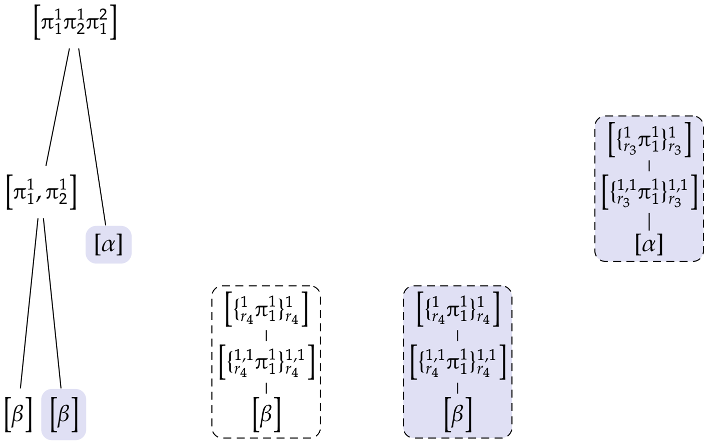
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$



Multiple Dyck derivation

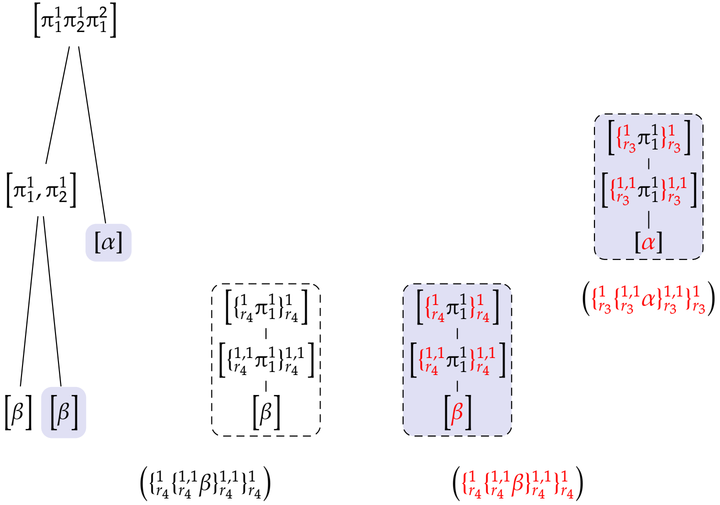
$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$



$$\left(\{_{r_4}^1\{_{r_4}^{1,1}\beta\}_{r_4}^{1,1}\}_{r_4}^1\right)$$

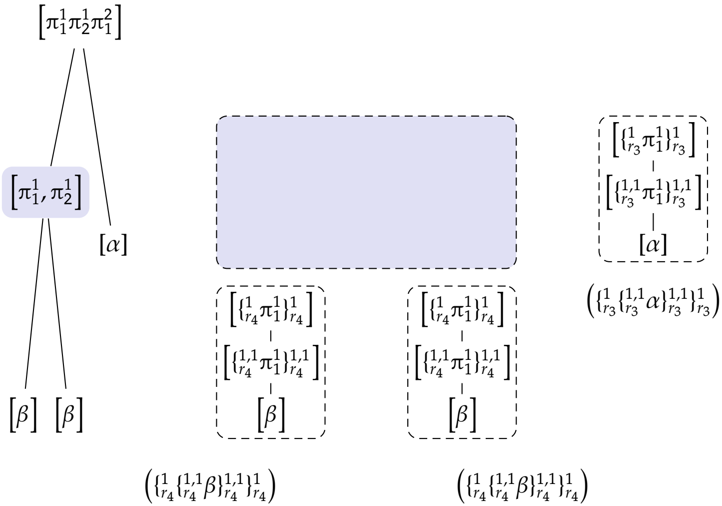
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$



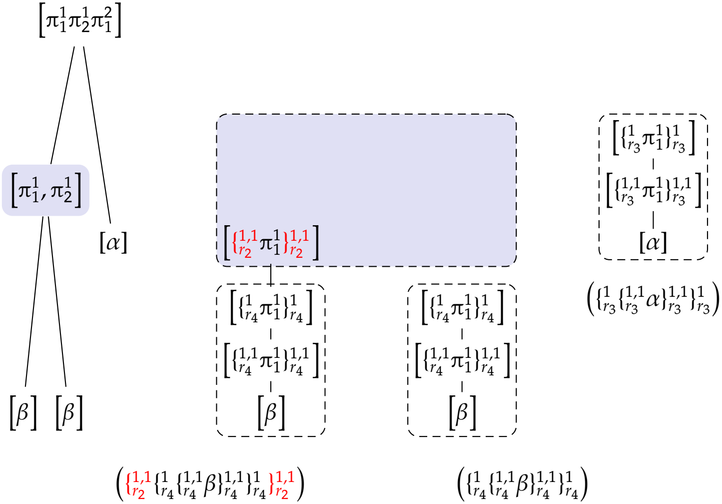
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$



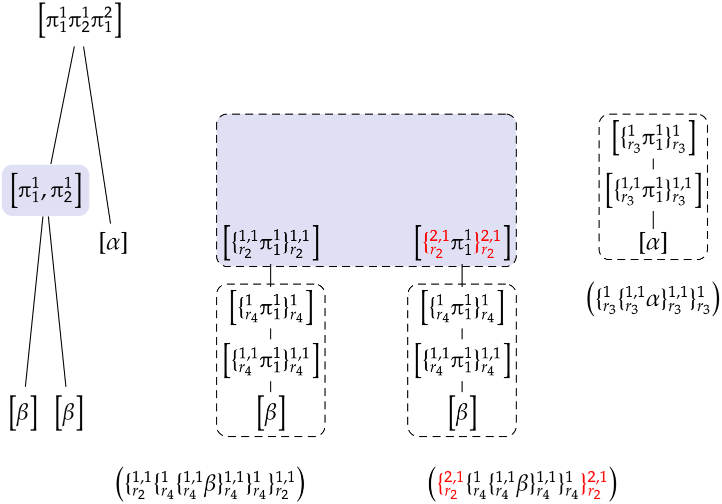
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\quad A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\quad B \xrightarrow{r_3} [\alpha],\quad C \xrightarrow{r_4} [\beta]$$



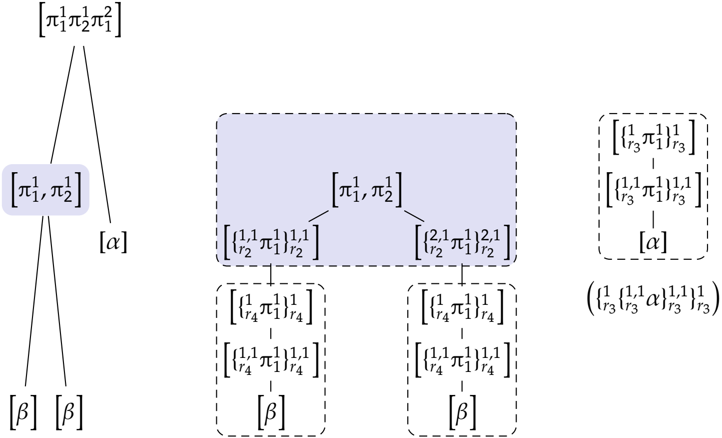
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\quad A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\quad B \xrightarrow{r_3} [\alpha],\quad C \xrightarrow{r_4} [\beta]$$



Multiple Dyck derivation

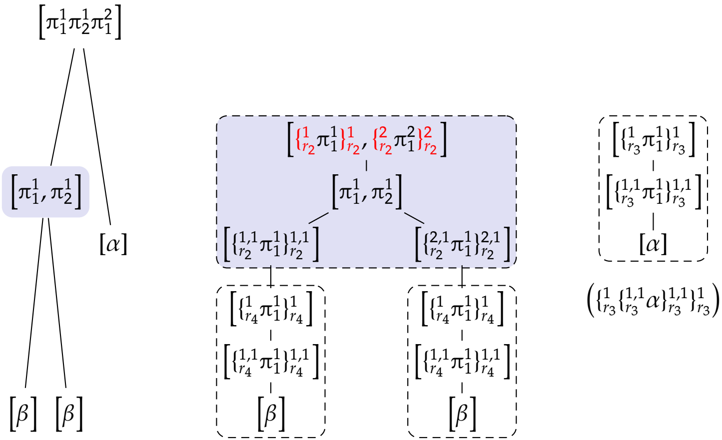
$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$



$$\left(\{^{1,1}_{r_2}\{^{1}_{r_4}\{^{1,1}_{r_4}\beta\}^{1,1}_{r_4}\}^{1}_{r_4}\}^{1,1}_{r_2} \quad , \quad \{^{2,1}_{r_2}\{^{1}_{r_4}\{^{1,1}_{r_4}\beta\}^{1,1}_{r_4}\}^{1}_{r_4}\}^{2,1}_{r_2}\right)$$

Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$

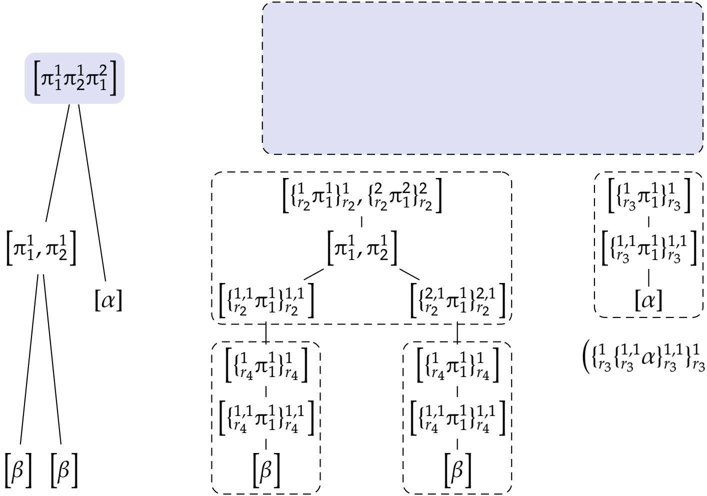


$$\left(\{^{1,1}_{r_3}\{^{1,1}_{r_3}\alpha\}^{1,1}_{r_3}\}^{1}_{r_3}\right)$$

$$\left(\{^{1}_{r_2}\{^{1,1}_{r_2}\{^{1}_{r_4}\{^{1,1}_{r_4}\beta\}^{1,1}_{r_4}\}^{1}_{r_4}\}^{1,1}_{r_2}\}^{1}_{r_2} \quad , \quad \{^{2}_{r_2}\{^{2,1}_{r_2}\{^{1}_{r_4}\{^{1,1}_{r_4}\beta\}^{1,1}_{r_4}\}^{1}_{r_4}\}^{2,1}_{r_2}\}^{2}_{r_2}\right)$$

Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\quad A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\quad B \xrightarrow{r_3} [\alpha],\quad C \xrightarrow{r_4} [\beta]$$

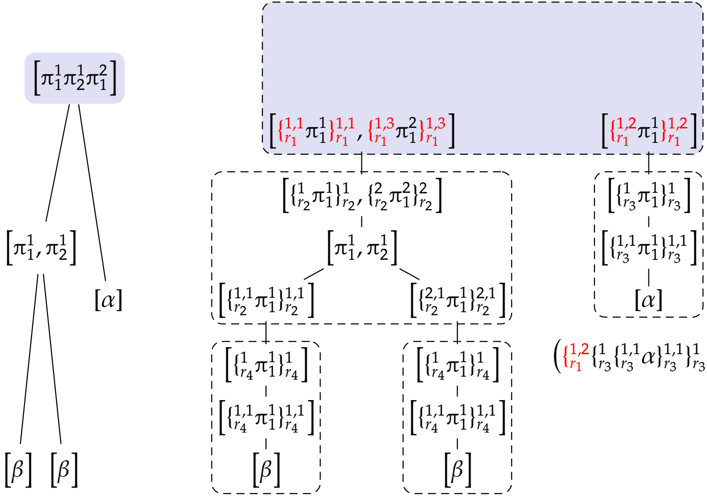


$$\left(\{_{r_3}^1\{_{r_3}^{1,1}\alpha\}_{r_3}^{1,1}\}_{r_3}^1\right)$$

$$\left(\{_{r_2}^1\{_{r_2}^{1,1}\{_{r_4}^1\{_{r_4}^{1,1}\beta\}_{r_4}^{1,1}\}_{r_4}^1\}_{r_2}^{1,1}\}_{r_2}^1 \quad , \quad \{_{r_2}^2\{_{r_2}^{2,1}\{_{r_4}^1\{_{r_4}^{1,1}\beta\}_{r_4}^{1,1}\}_{r_4}^1\}_{r_2}^{2,1}\}_{r_2}^2\right)$$

Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$

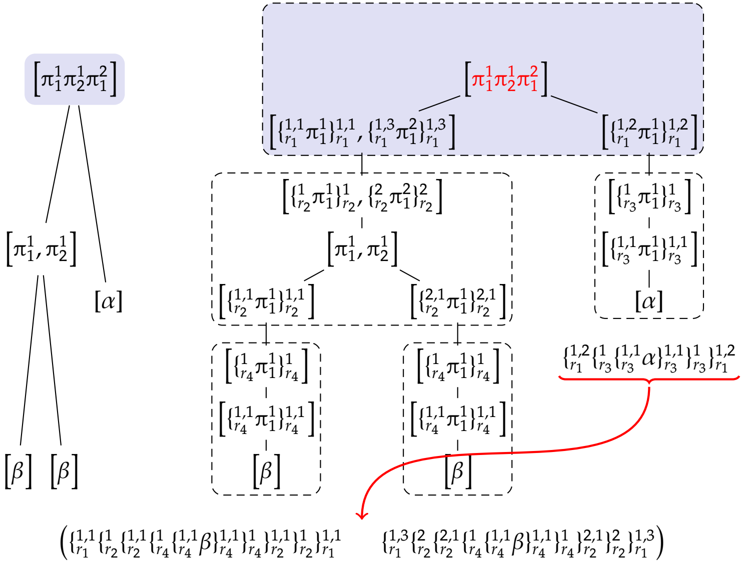


$$\left(\{^{1,2}_{r_1}\{^{1}_{r_3}\{^{1,1}_{r_3}\alpha\}^{1,1}_{r_3}\}^{1}_{r_3}\}^{1,2}_{r_1}\right)$$

$$\left(\{^{1,1}_{r_1}\{^{1}_{r_2}\{^{1,1}_{r_2}\{^{1}_{r_4}\{^{1,1}_{r_4}\beta\}^{1,1}_{r_4}\}^{1}_{r_4}\}^{1,1}_{r_2}\}^{1}_{r_2}\}^{1,1}_{r_1}\ ,\ \{^{1,3}_{r_1}\{^{2}_{r_2}\{^{2,1}_{r_2}\{^{1}_{r_4}\{^{1,1}_{r_4}\beta\}^{1,1}_{r_4}\}^{1}_{r_4}\}^{2,1}_{r_2}\}^{2}_{r_2}\}^{1,3}_{r_1}\right)$$

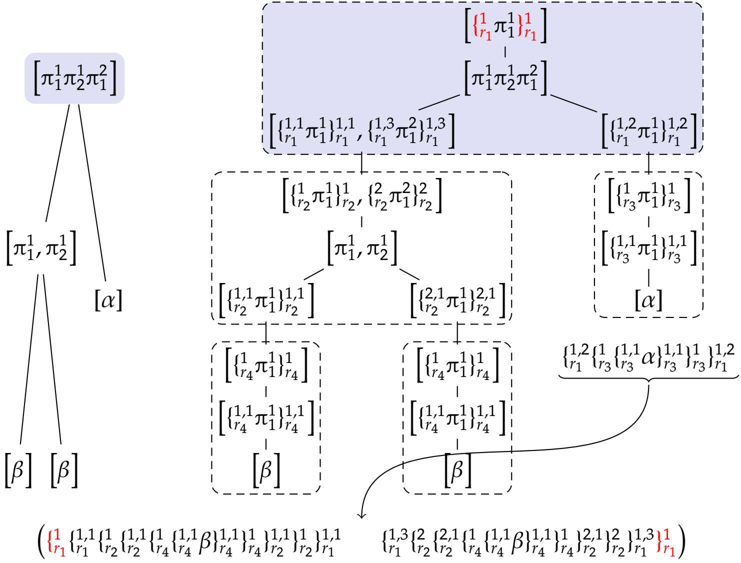
Multiple Dyck derivation

$$R:\ \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\ \ A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\ \ B \xrightarrow{r_3} [\alpha],\ \ C \xrightarrow{r_4} [\beta]$$



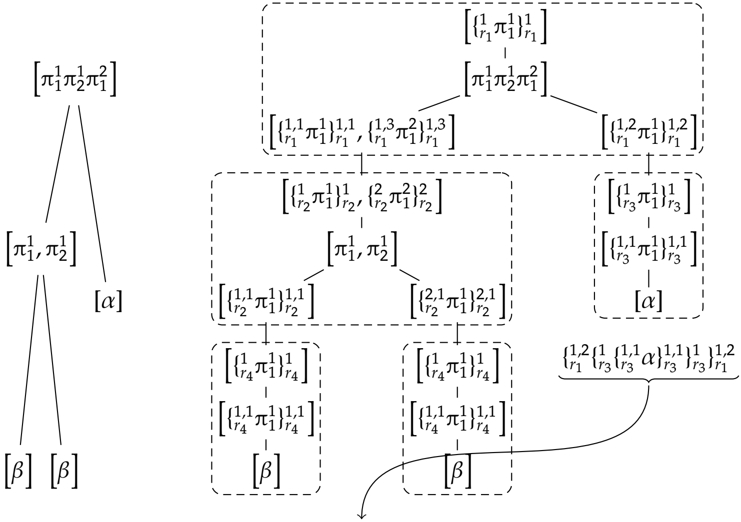
Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\quad A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\quad B \xrightarrow{r_3} [\alpha],\quad C \xrightarrow{r_4} [\beta]$$



Multiple Dyck derivation

$$R\colon \mathbf{S} \xrightarrow{r_1} [\pi_1^1\pi_2^1\pi_1^2](A,B),\;\; A \xrightarrow{r_2} [\pi_1^1,\pi_2^1](C,C),\;\; B \xrightarrow{r_3} [\alpha],\;\; C \xrightarrow{r_4} [\beta]$$



$$\left(\{^{1}_{r_1} \{^{1,1}_{r_1} \{^{1}_{r_2} \{^{1,1}_{r_2} \{^{1}_{r_4} \{^{1,1}_{r_4} \beta \}^{1,1}_{r_4} \}^{1}_{r_4} \}^{1,1}_{r_2} \}^{1}_{r_2} \}^{1,1}_{r_1} \quad \{^{1,3}_{r_1} \{^{2}_{r_2} \{^{2,1}_{r_2} \{^{1}_{r_4} \{^{1,1}_{r_4} \beta \}^{1,1}_{r_4} \}^{1}_{r_4} \}^{2,1}_{r_2} \}^{2}_{r_2} \}^{1,3}_{r_1} \}^{1}_{r_1} \right)$$

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and
- a homomorphism h such that

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and
- a homomorphism h such that

$$\mathscr{L}^{\mathrm{w}}(u) = \sum_{\substack{v \in \mathscr{D} \\ h(v) = u}} R^{\mathrm{w}}(v) \qquad \text{(finiteness conditions)}$$

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and
- a homomorphism h such that

$$\mathscr{L}^{\mathrm{w}}(u) = \sum_{\substack{v \in \mathscr{D} \\ h(v)=u}} R^{\mathrm{w}}(v) \qquad \text{(finiteness conditions)}$$

Idea

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and
- a homomorphism h such that

$$\mathscr{L}^{\mathrm{w}}(u) = \sum_{\substack{v \in \mathscr{D} \\ h(v) = u}} R^{\mathrm{w}}(v) \qquad \text{(finiteness conditions)}$$

Idea

1. weighted MCFG $\rightsquigarrow$ unweighted MCFG (support)

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and
- a homomorphism h such that

$$\mathscr{L}^{\mathrm{w}}(u) = \sum_{\substack{v \in \mathscr{D} \\ h(v) = u}} R^{\mathrm{w}}(v) \qquad \text{(finiteness conditions)}$$

Idea

1. weighted MCFG $\rightsquigarrow$ unweighted MCFG (support)
2. use unweighted construction

Chomsky-Schützenberger for weighted MCFL

Theorem

For every zero-sum and zero-product free semiring S and every S-weighted MCFL $\mathscr{L}^{\mathrm{w}}$ there are

- a multiple Dyck language $\mathscr{D}$,
- an S-weighted regular language R^{w}, and
- a homomorphism h such that

$$\mathscr{L}^{\mathrm{w}}(u) = \sum_{\substack{v \in \mathscr{D} \\ h(v) = u}} R^{\mathrm{w}}(v) \qquad \text{(finiteness conditions)}$$

Idea

1. weighted MCFG $\rightsquigarrow$ unweighted MCFG (support)
2. use unweighted construction
3. original MCFG weights in R when $\{^1_r$ is read

Summary and outlook

Summary

Outlook

Summary and outlook

Summary

- MCFG are interesting for *natural language processing*.

Outlook

Summary and outlook

Summary

- MCFG are interesting for *natural language processing*.
- A MCFL can be expressed with a *multiple Dyck language*, a *regular language* and a *homomorphism*.

Outlook

Summary and outlook

Summary

- MCFG are interesting for *natural language processing*.
- A (weighted) MCFL can be expressed with a *multiple Dyck language*, a *(weighted) regular language* and a *homomorphism*.

Outlook

Summary and outlook

Summary

- MCFG are interesting for *natural language processing*.
- A (weighted) MCFL can be expressed with a *multiple Dyck language*, a *(weighted) regular language* and a *homomorphism*.

Outlook

- Define a *parsing algorithm* for MCFG based on their Chomsky-Schützenberger characterisation (similar to Hulden 2011).

References

[1] Noam Chomsky and Marcel Paul Schützenberger. "The algebraic theory of context-free languages". 1963.

[2] Kilian Evang and Laura Kallmeyer. "PLCFRS parsing of English discontinuous constituents". 2011.

[3] Mans Hulden. "Parsing CFGs and PCFGs with a Chomsky-Schützenberger Representation". 2011.

[4] Wolfgang Maier and Anders Søgaard. "Treebanks and mild context-sensitivity". 2008.