

Synchronous Context-Free Grammars and Optimal Linear Parsing Strategies

Daniel Gildea

University of Rochester

Giorgio Satta

Università di Padova

Synchronous CFG

Context-free Grammar:

$$X \rightarrow AB$$

Synchronous Context-free Grammar (SCFG)

$$X \rightarrow A^{\boxed{1}} B^{\boxed{2}} C^{\boxed{3}} D^{\boxed{4}}, C^{\boxed{3}} A^{\boxed{1}} D^{\boxed{4}} B^{\boxed{2}}$$

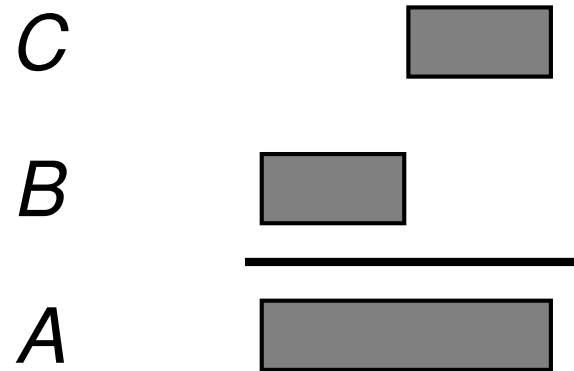
$C \rightarrow$ Powell, 鲍威尔

Synchronous CFG

- Synchronous parsing: find tree from two strings
 - used to learn grammar from parallel text
- This talk: parsing strategies for long rules
- Results also apply to translation with n-gram language model

Context-Free Grammar

$$A \rightarrow BC$$



Binary SCFG

$$A \rightarrow B^{[1]} C^{[2]}, C^{[2]} B^{[1]}$$

C



B

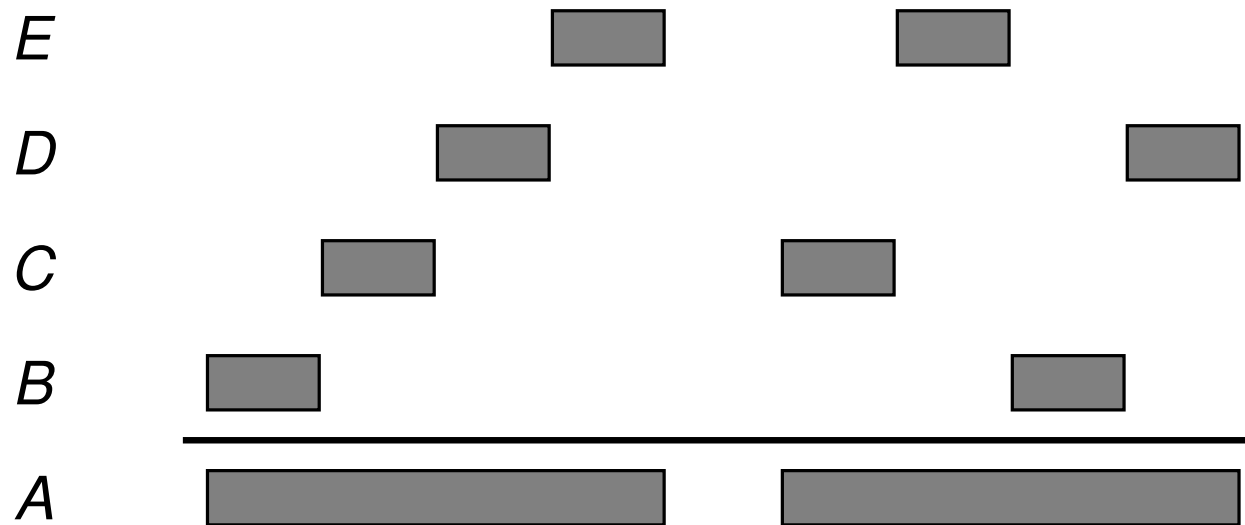


A



SCFG with 4 nonterminals

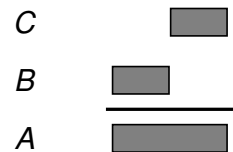
$$A \rightarrow B^{[1]} C^{[2]} D^{[3]} E^{[4]}, C^{[2]} E^{[4]} B^{[1]} D^{[3]}$$



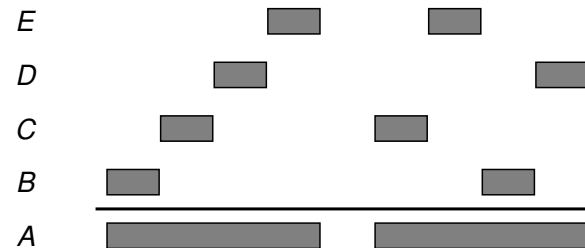
Fan-Out

Number of spans in nonterminal.

CFG: fan-out 1



SCFG: fan-out 2



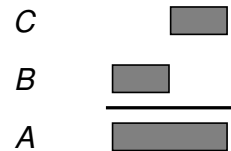
$$\varphi(G) = \max_{N \in G} \varphi(N)$$

(Rambow & Satta, 1999)

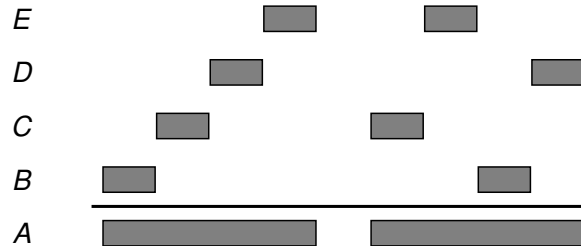
Rank

Number of nonterminals on righthand side of rule.

CFG: rank 2



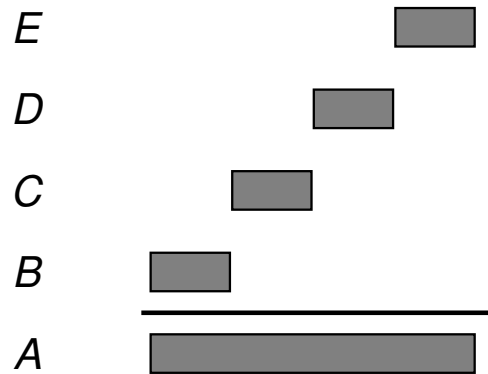
SCFG: rank r



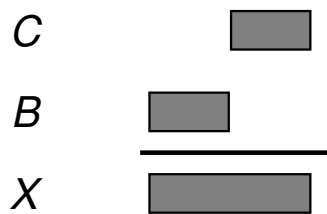
$$\rho(G) = \max_{P \in G} \rho(P)$$

Parsing Strategies

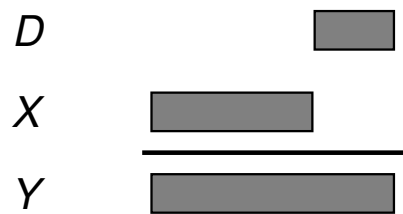
Reduce rank



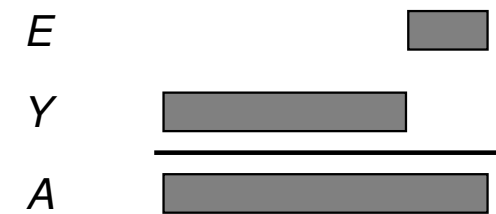
$A \rightarrow BCDE$



$X \rightarrow BC$



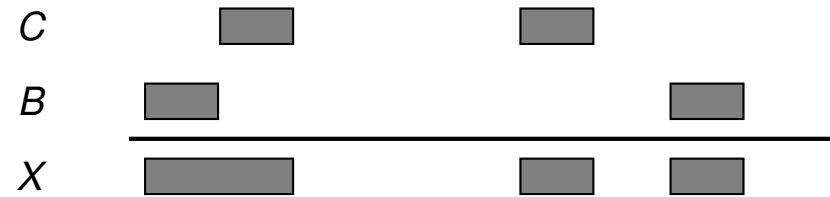
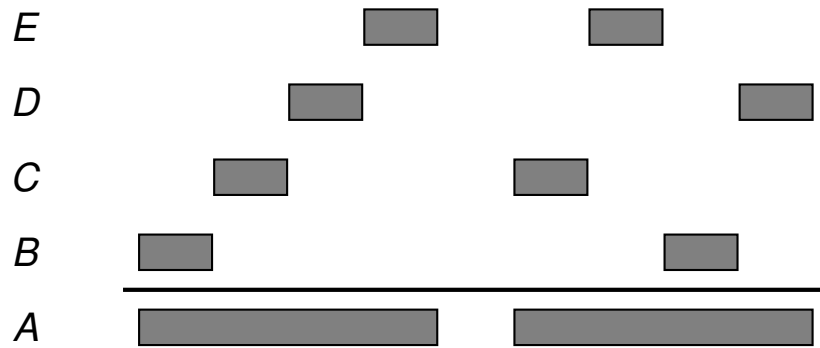
$Y \rightarrow XD$



$A \rightarrow YE$

Parsing Strategies

Reduce rank, may increase fan-out



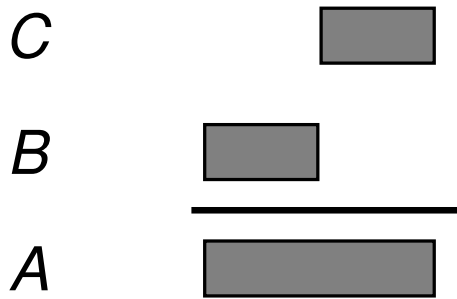
Rule Length in Synchronous CFG

- Binary grammar (ITG): parsing is $O(n^6)$ (Wu, 1997)
 - Works in real MT (Zhang et al. 2006)
- Many rules cannot be binarized without increasing fan-out (Aho and Ullman, 1972)
- Fan-out affects space and time complexity

Parsing Complexity

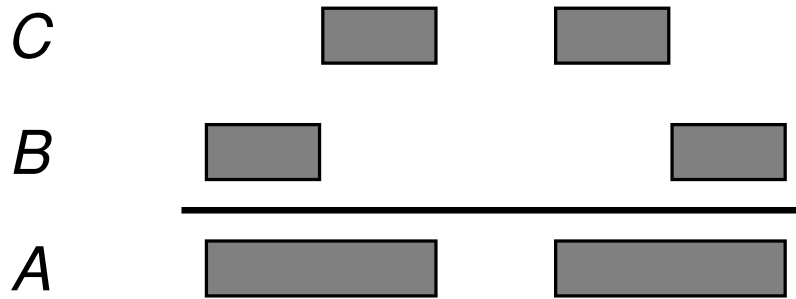
Space complexity: $O(n^{2\varphi(A)})$

Time complexity: $O(n^{\varphi(A)+\varphi(B)+\varphi(C)})$



$O(n^2)$ space

$O(n^3)$ time

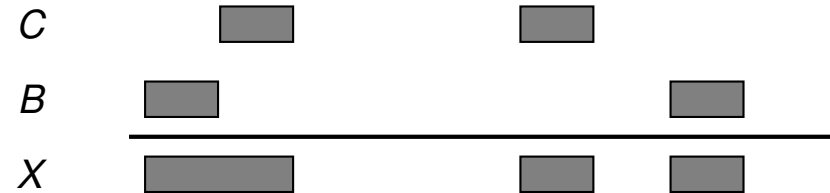
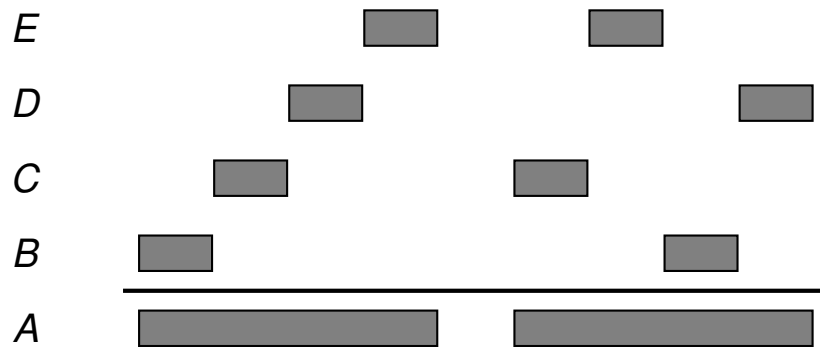


$O(n^4)$ space

$O(n^6)$ time

(Seki et al. 1991)

SCFG Parsing Strategies



naïve strategy: $O(n^{2r+2})$ time

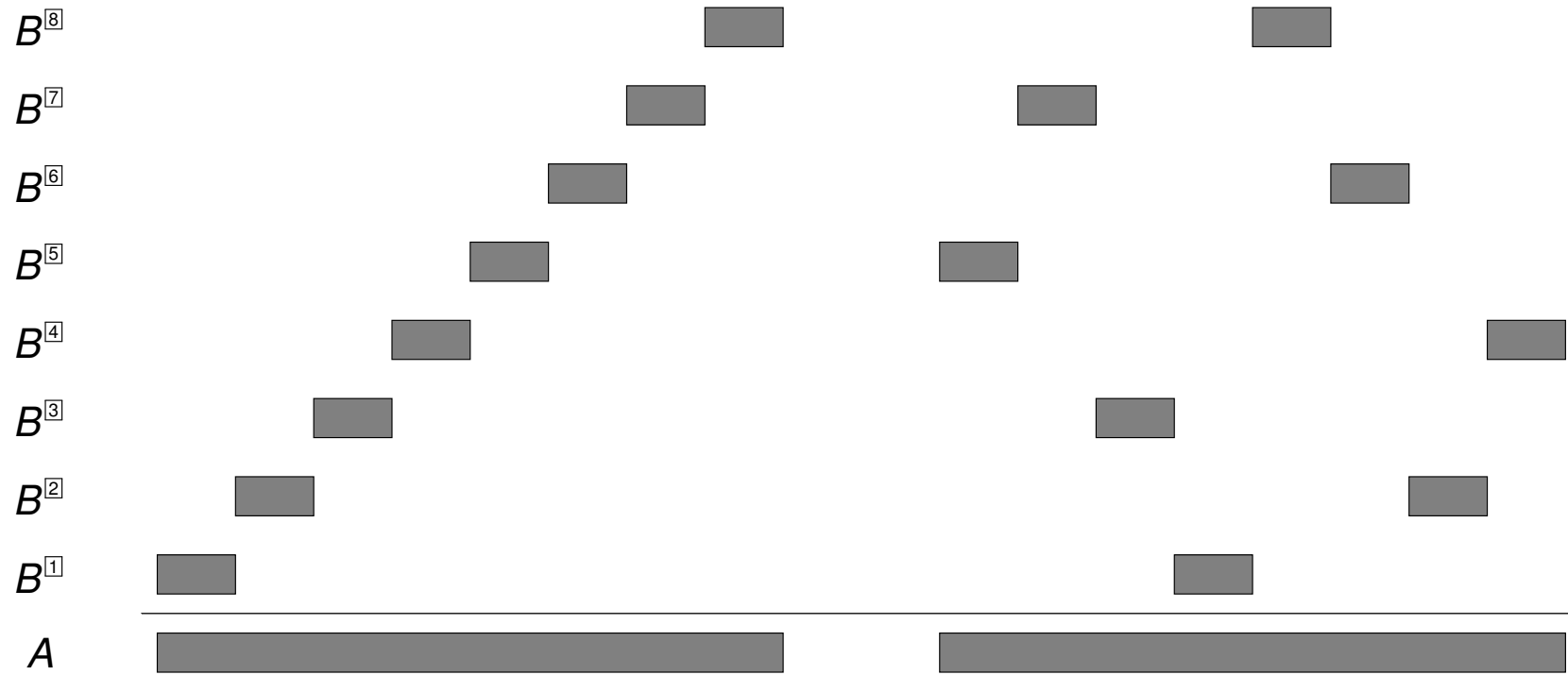
best strategy: $\Omega(n^{cr})$ for some c

(Gildea and Štefanković 2007)

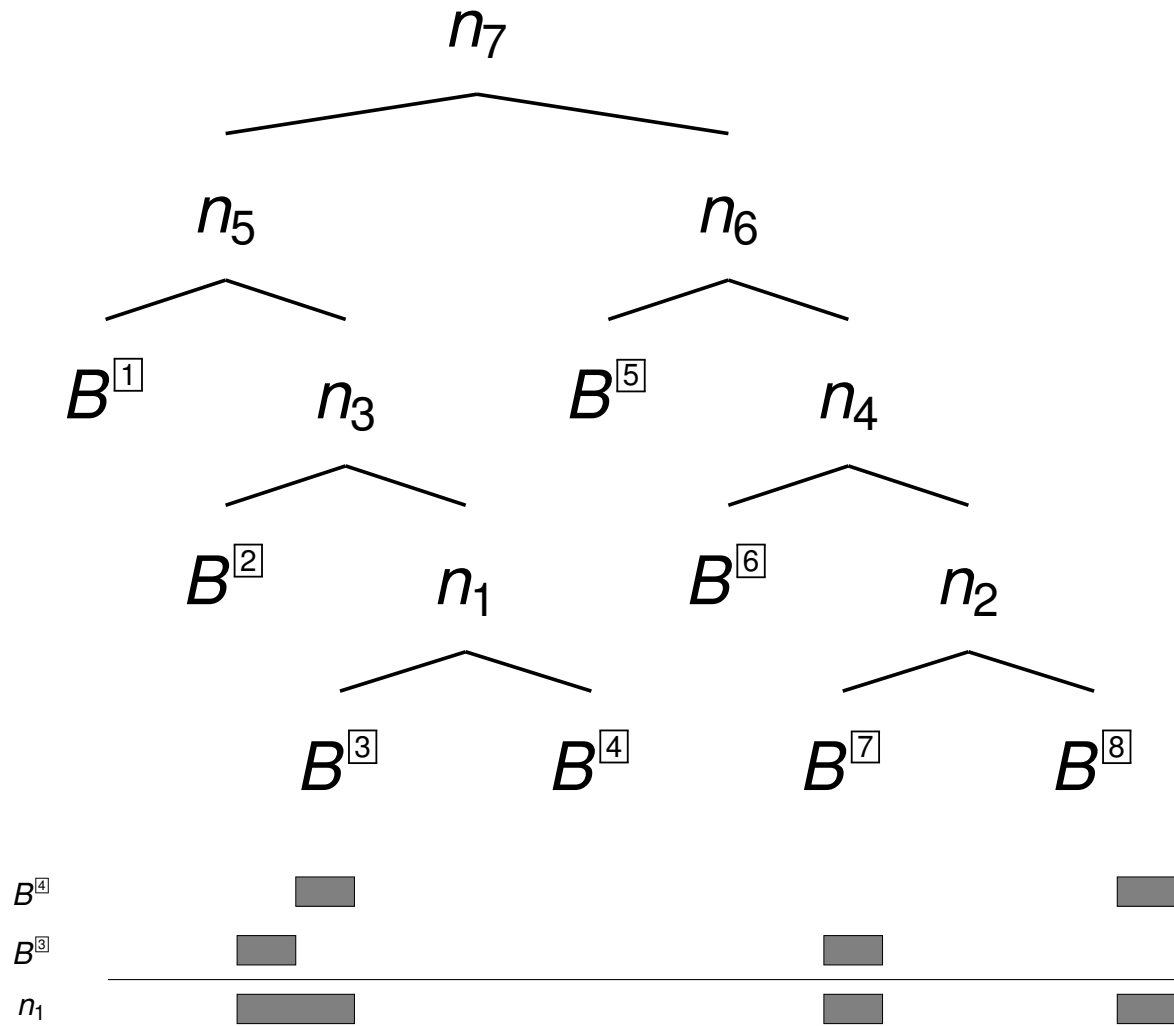
This Talk

- Finding optimal space complexity is NP-complete
- Finding optimal time complexity \Rightarrow better algs for treewidth

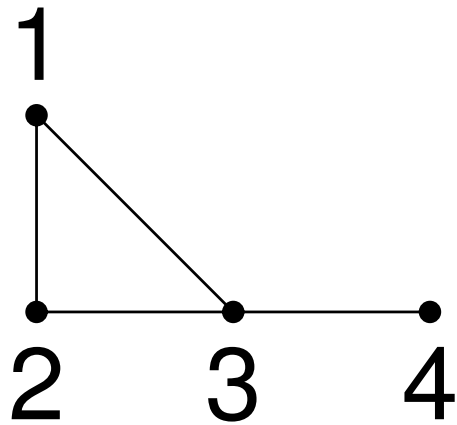
Example Rule



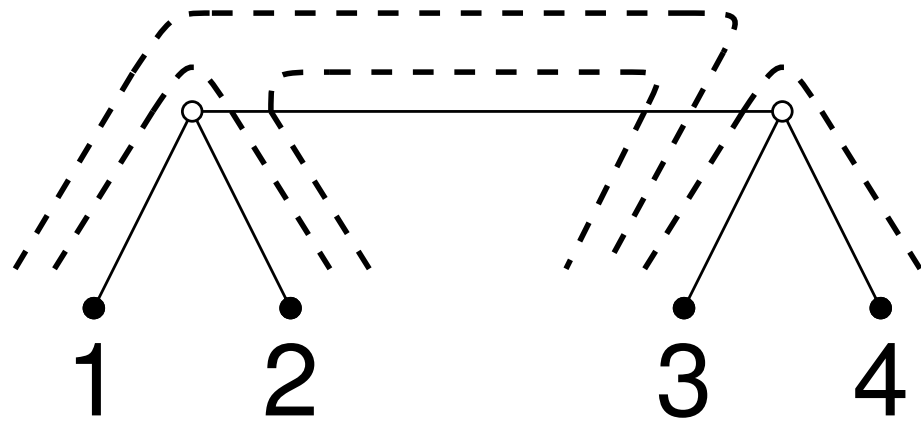
Optimal Parsing Strategy



Carving Width



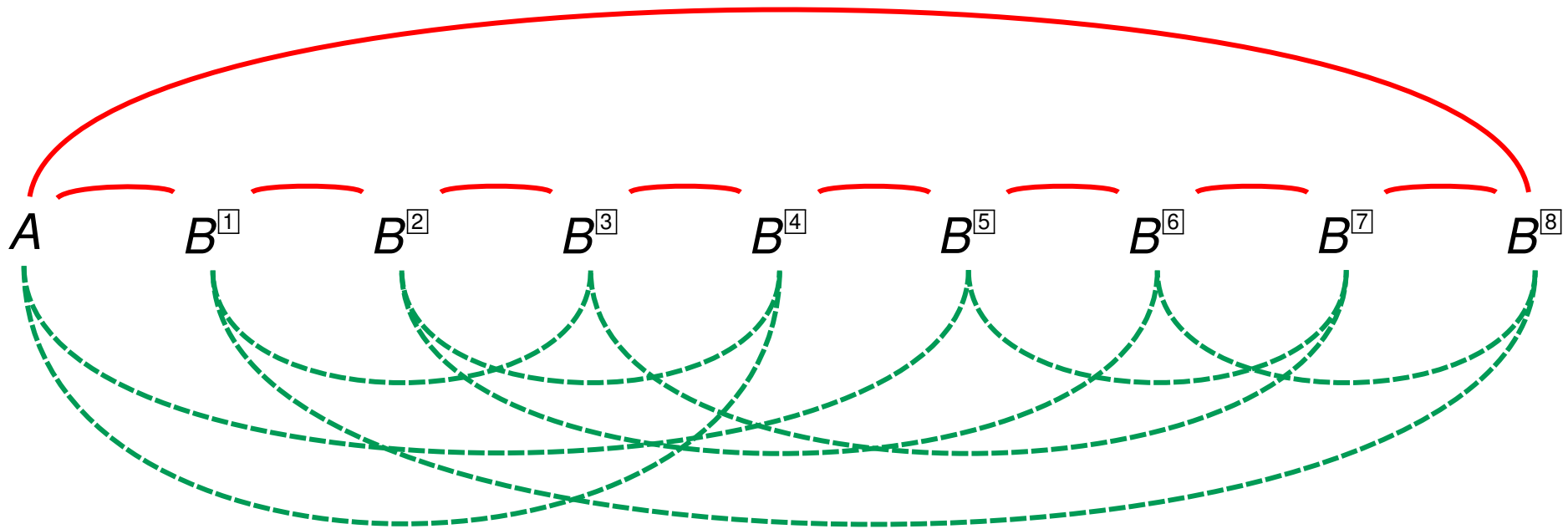
G



tree layout of G

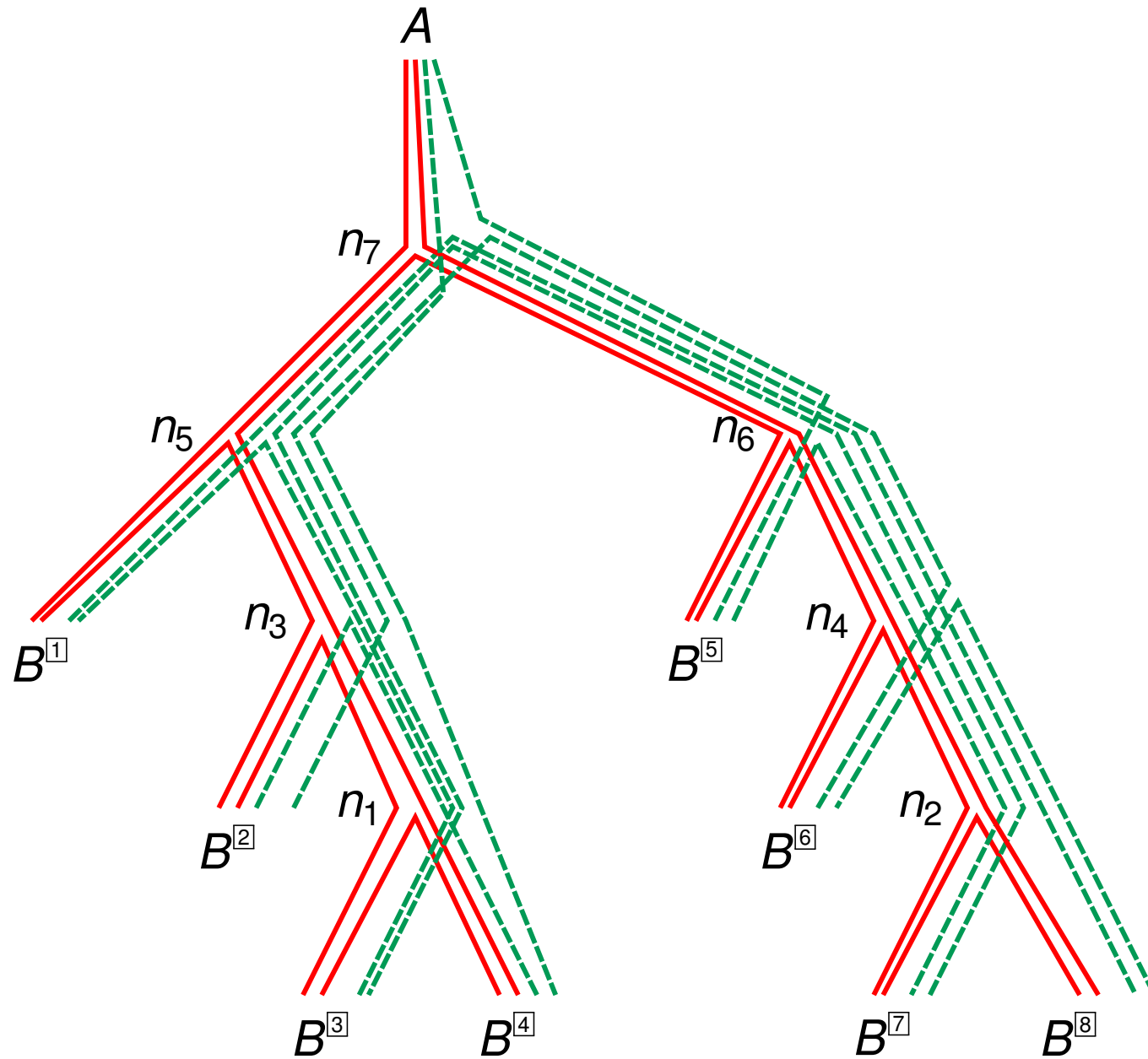
Carving width: max number edges of G routed through tree layout

Cyclic Permutation Multigraph



$$A \rightarrow B^1 B^2 B^3 B^4 B^5 B^6 B^7 B^8, \\ B^5 B^7 B^3 B^1 B^8 B^6 B^2 B^4$$

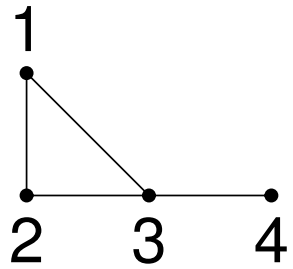
Carving Width = Space Complexity



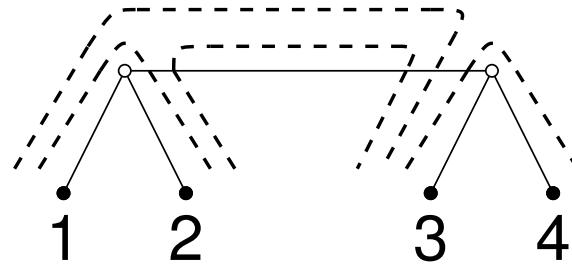
Our Reduction

- Carving width instance: (G, k)
- Construct permutation multigraph G' , integer k'
- Carving width of $G \iff$ Carving width of $G' \iff$
optimal parsing for SCFG

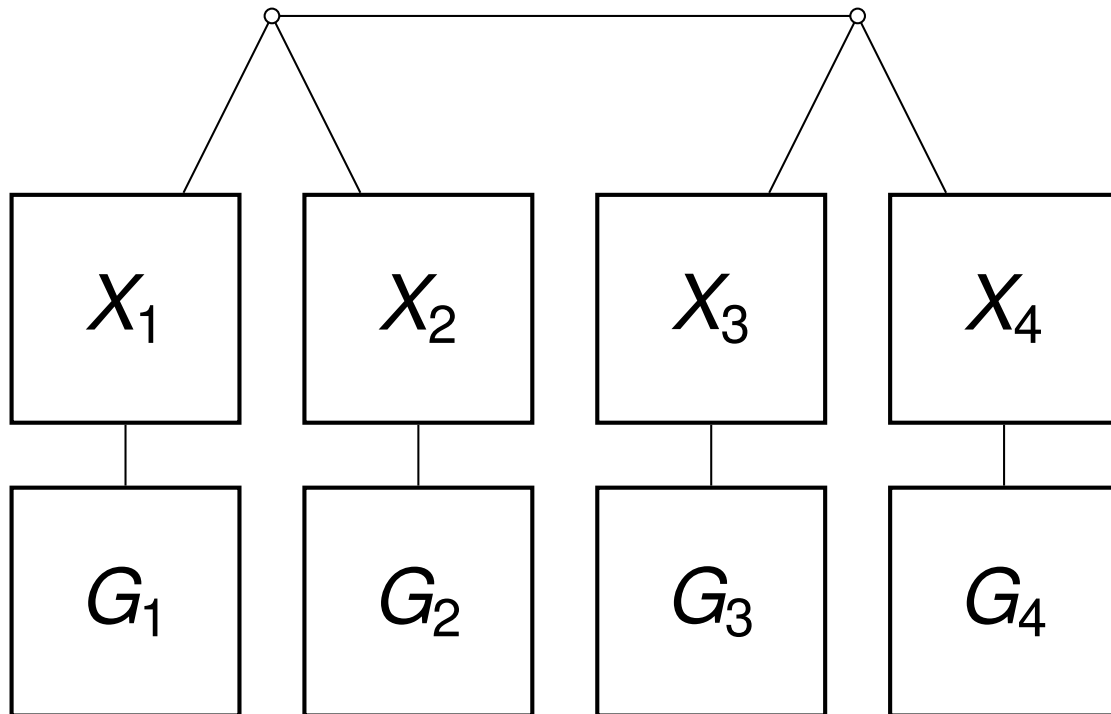
Our Construction

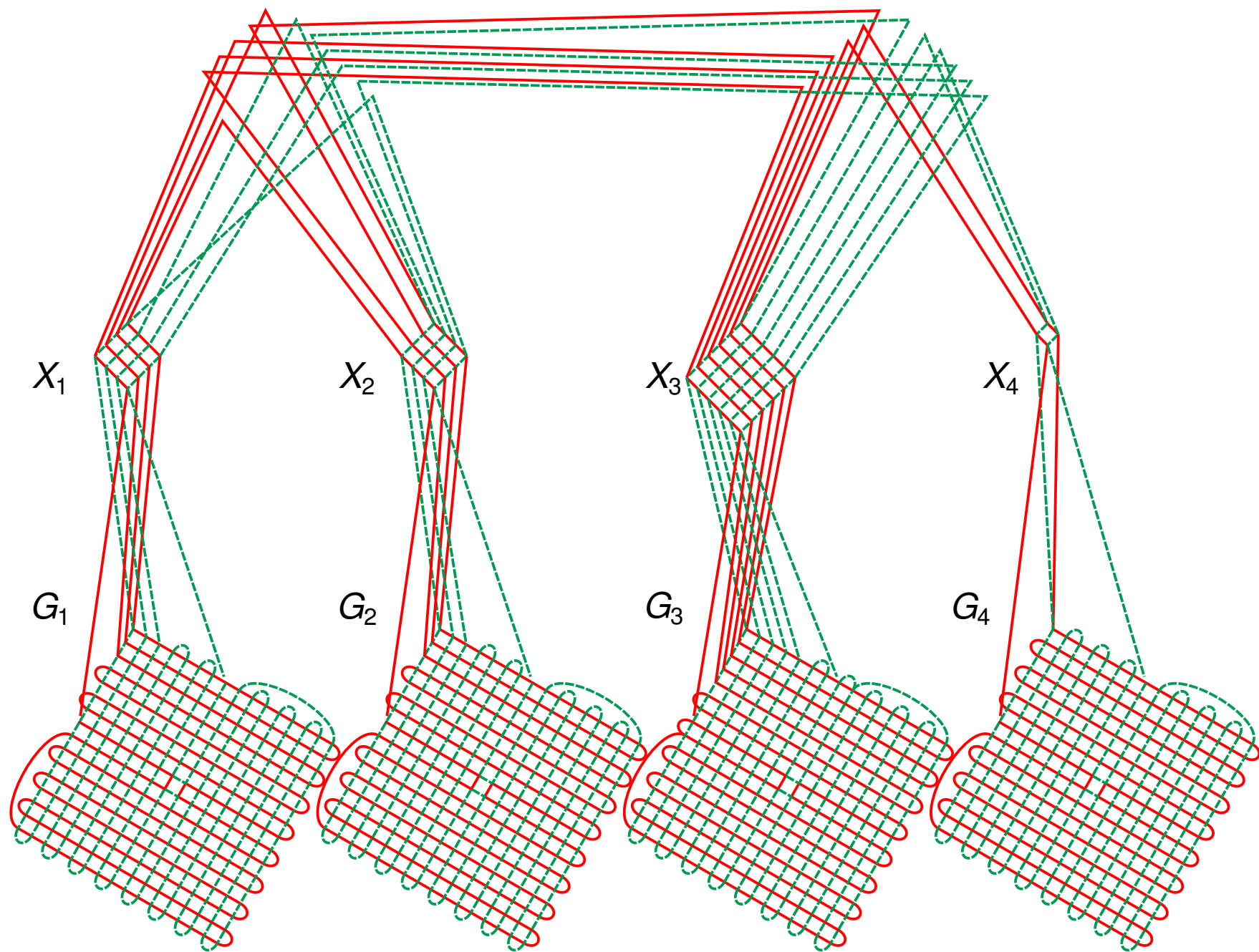


G



tree layout of G

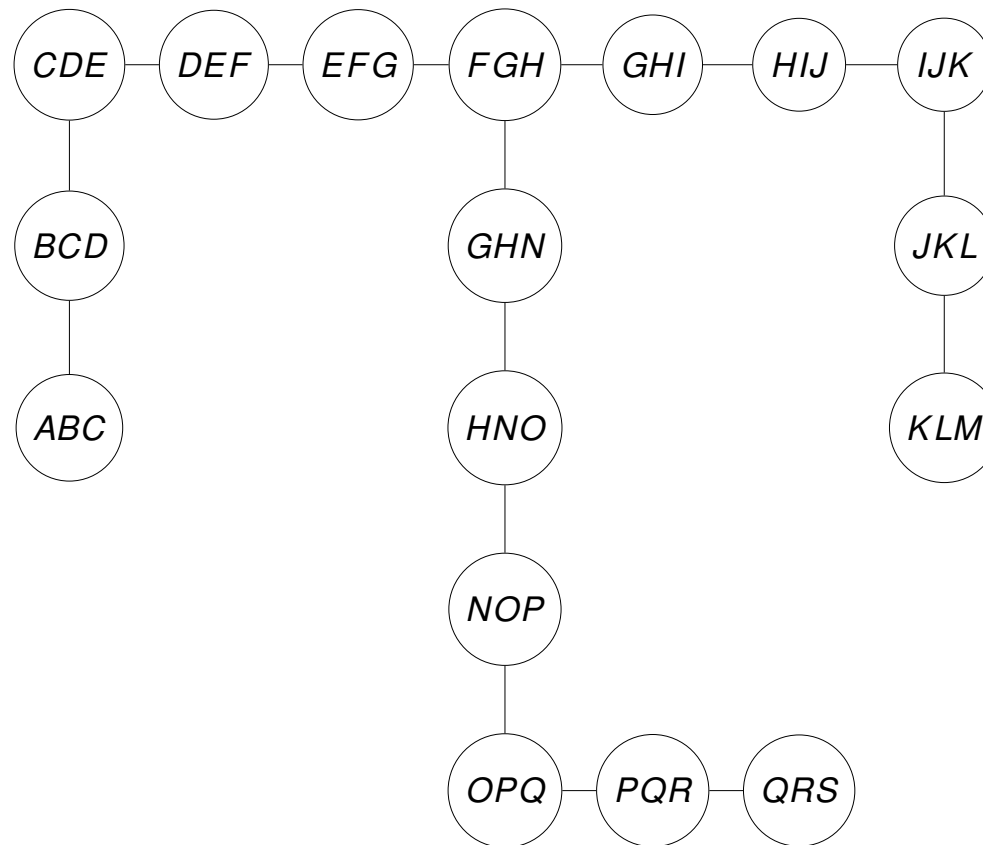
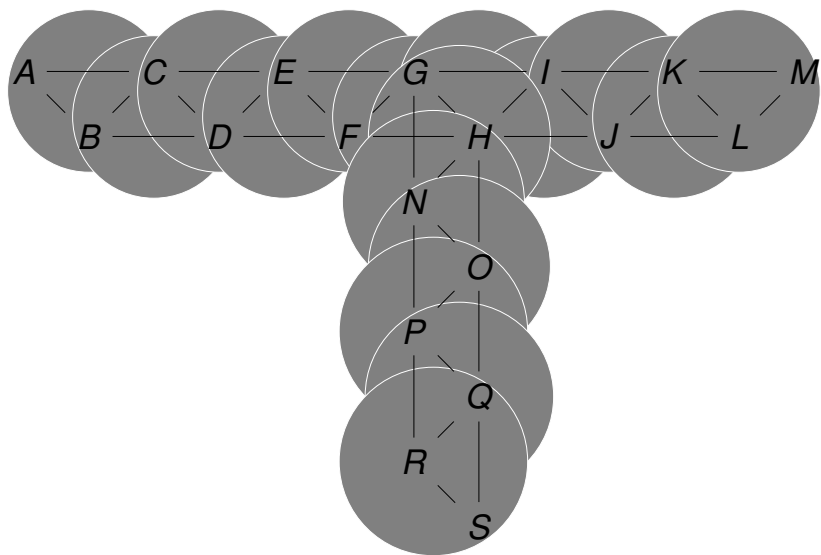




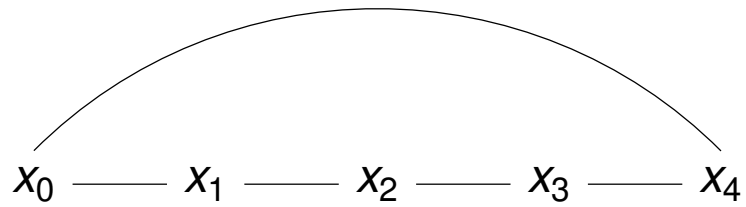
Space Complexity

Theorem 1: Finding the parsing strategy with optimal space complexity for an SCFG rule is NP-complete

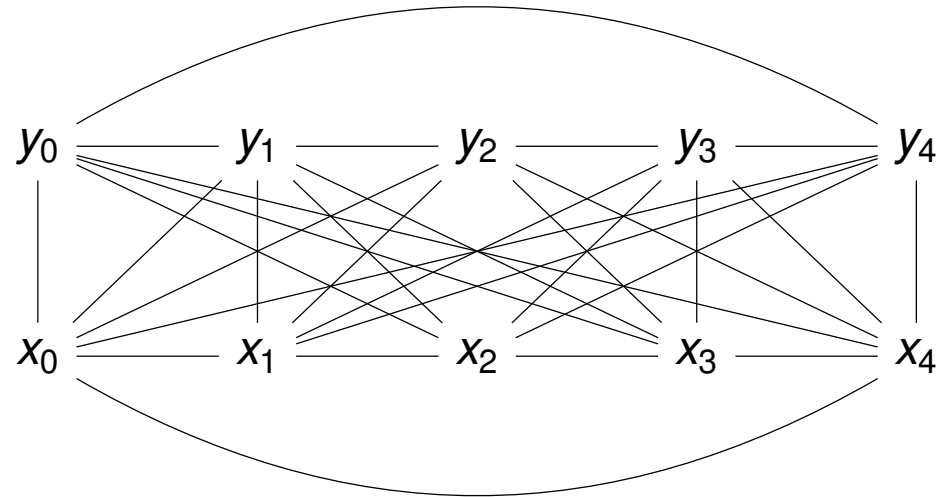
Treewidth



Dependency Graph

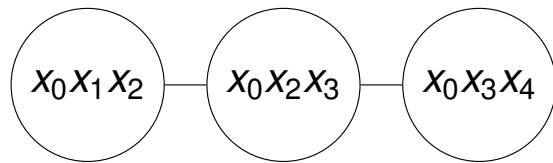
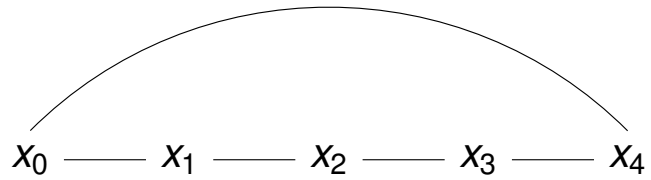


$A \rightarrow BCDE$

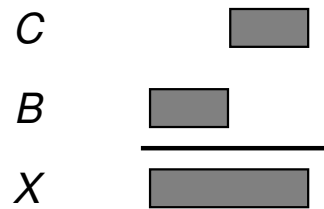


$S \rightarrow A^1 B^2 C^3 D^4, B^2 D^4 A^1 C^3$

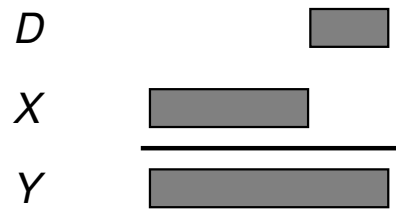
Treewidth = Time Complexity



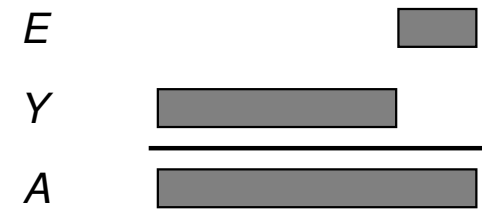
$A \rightarrow BCDE$



$X \rightarrow BC$



$Y \rightarrow XD$

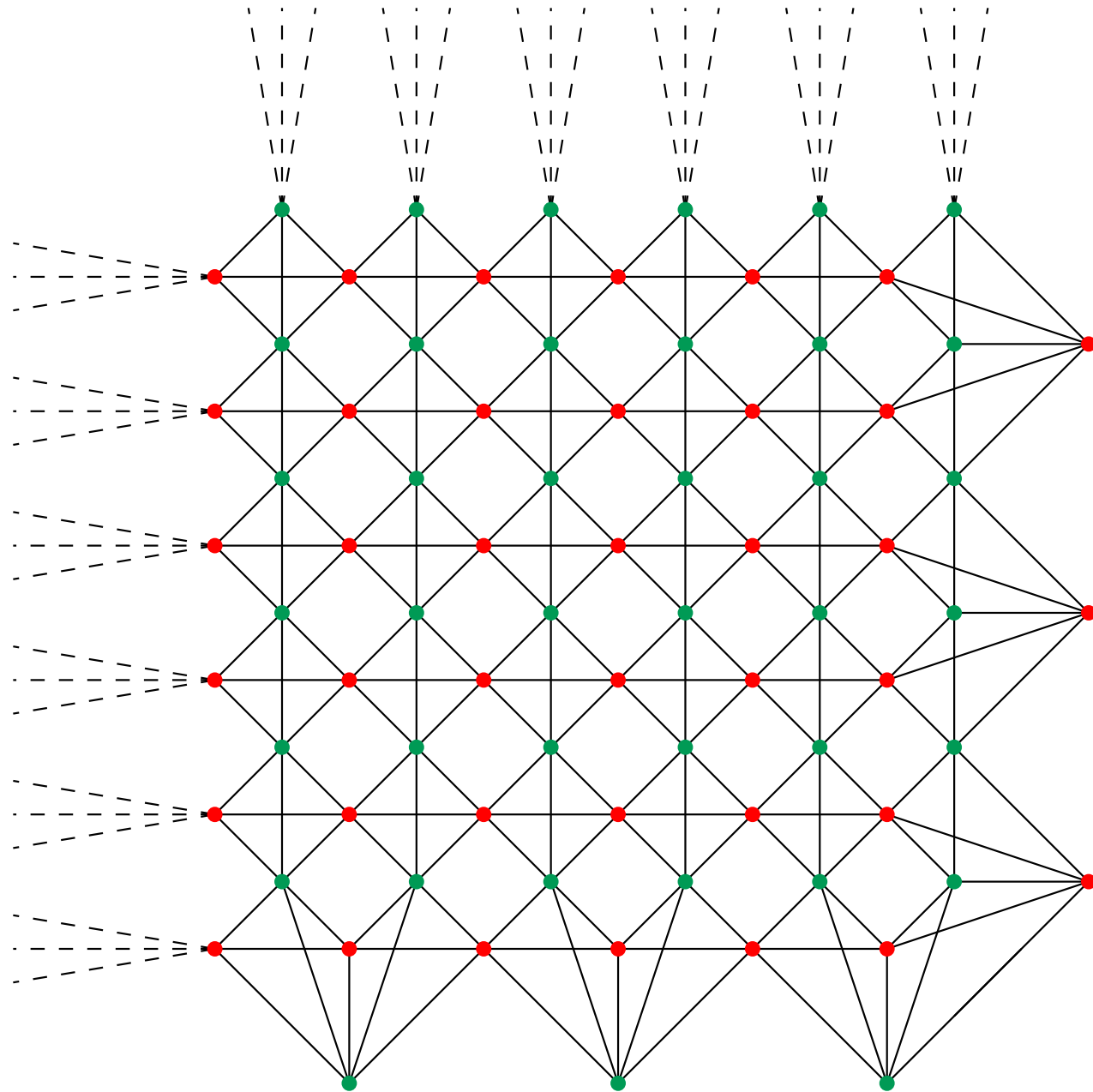


$A \rightarrow YE$

Our Reduction

- Treewidth instance: (G, k)
- Construct dependency graph G' , integer k'
- Approx of treewidth of $G \Leftrightarrow$ Treewidth of $G' \Leftrightarrow$
optimal time complexity for SCFG

Dependency Graph Construction



Approximation Algorithm for Treewidth

$$SOL < 8\Delta(G)(OPT + 1) .$$

SOL: solution using SCFG parsing strategy

OPT: optimal treewidth of input graph G

$\Delta(G)$ = degree (max num edges touching one vertex)

Time Complexity

Theorem 2: Finding the parsing strategy with optimal time complexity for an SCFG rule implies a $\Delta(G)$ -factor approximation algorithm for treewidth.

Time Complexity

Theorem 3: If finding the parsing strategy with optimal time complexity for an SCFG rule is NP-complete, then treewidth for graphs of degree 6 is NP-complete.

Conclusion

- Finding parsing strategy with best space complexity is NP-hard.
- P-time alg for finding parsing strategy with best time complexity implies better approximation algs for treewidth
- NP-hardness for time complexity implies NP-hardness for treewidth of graphs of degree six