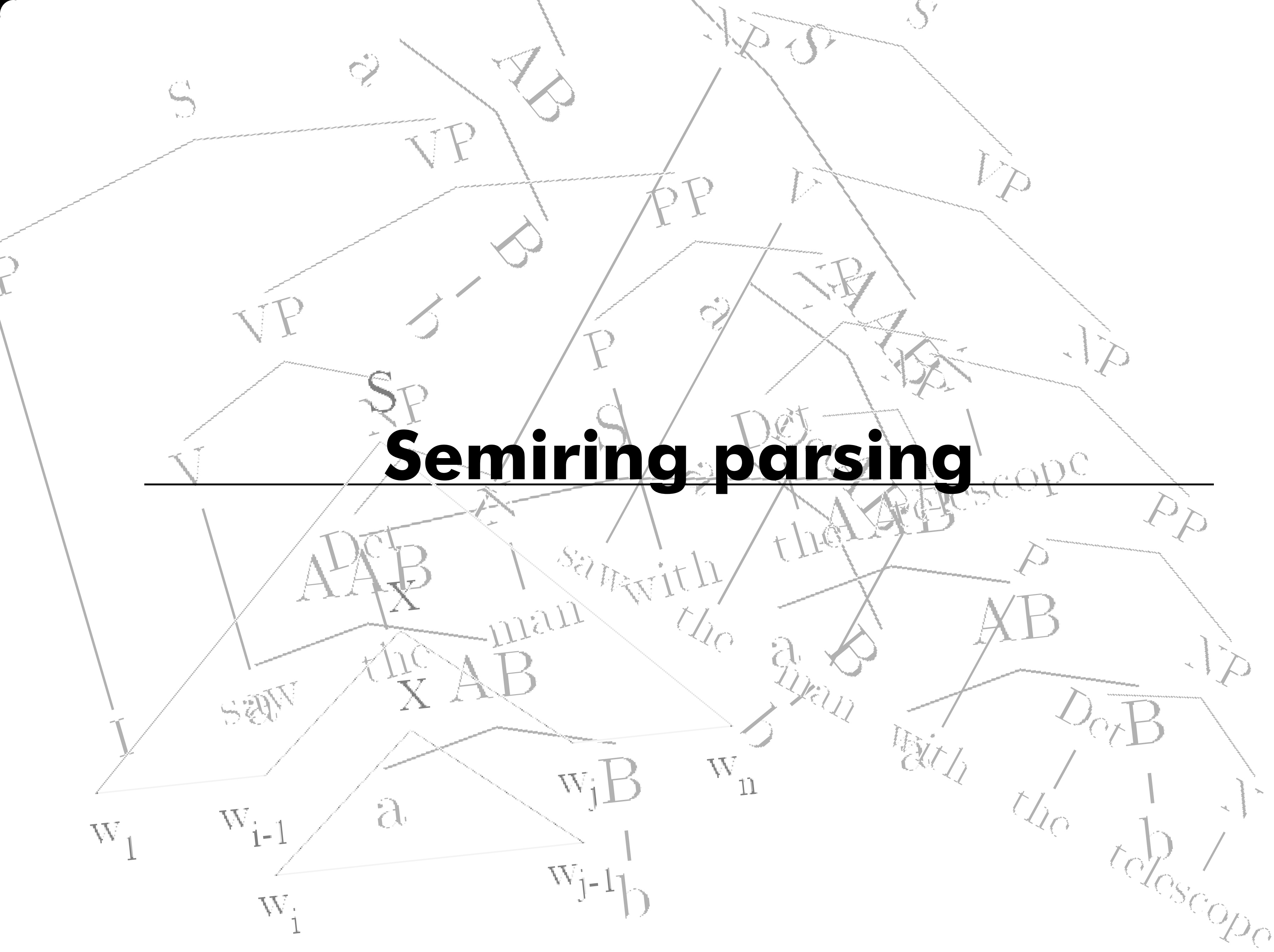


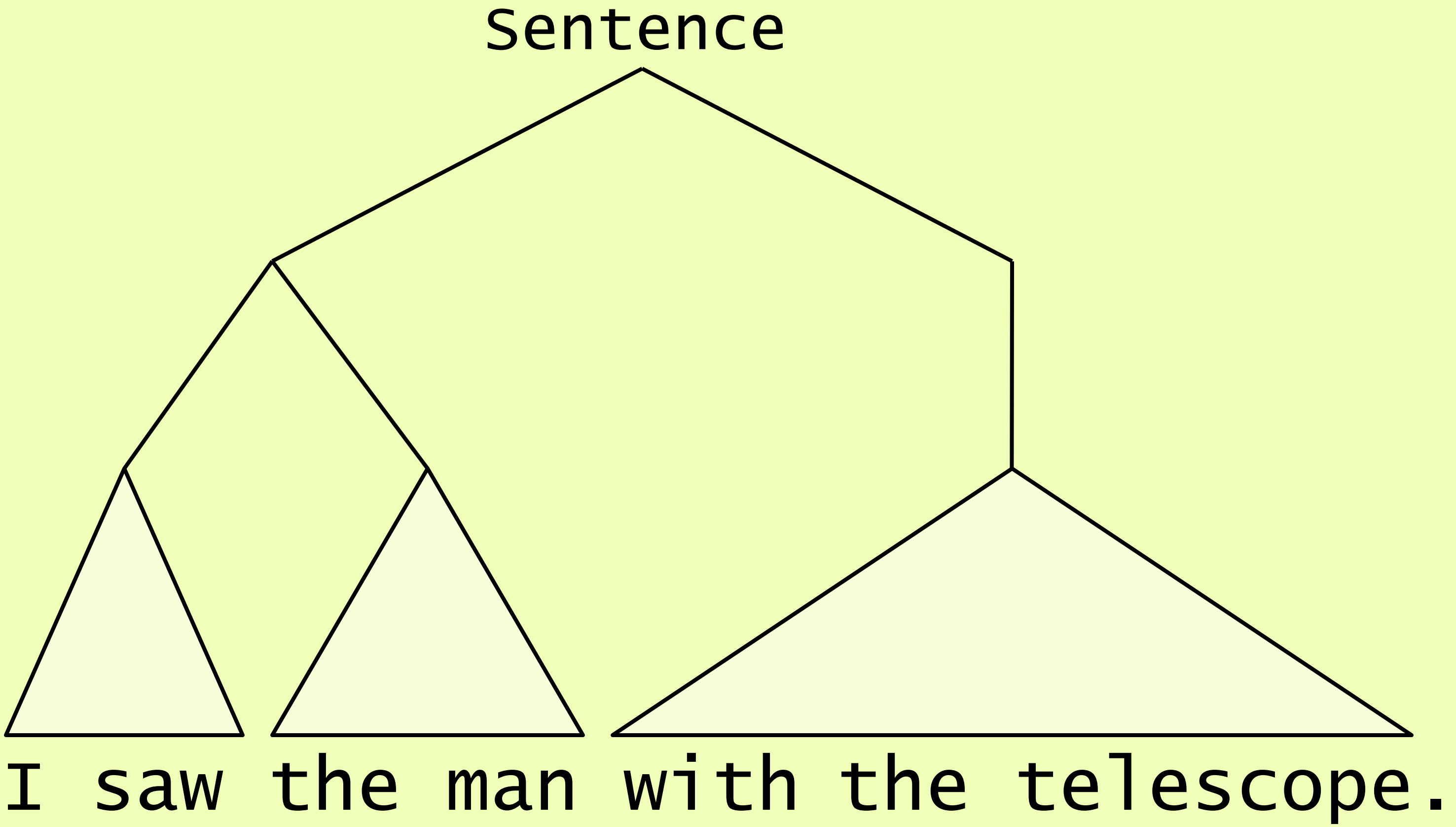
Semiring parsing

Arnd Hartmanns

Probabilistic grammars

Motivation

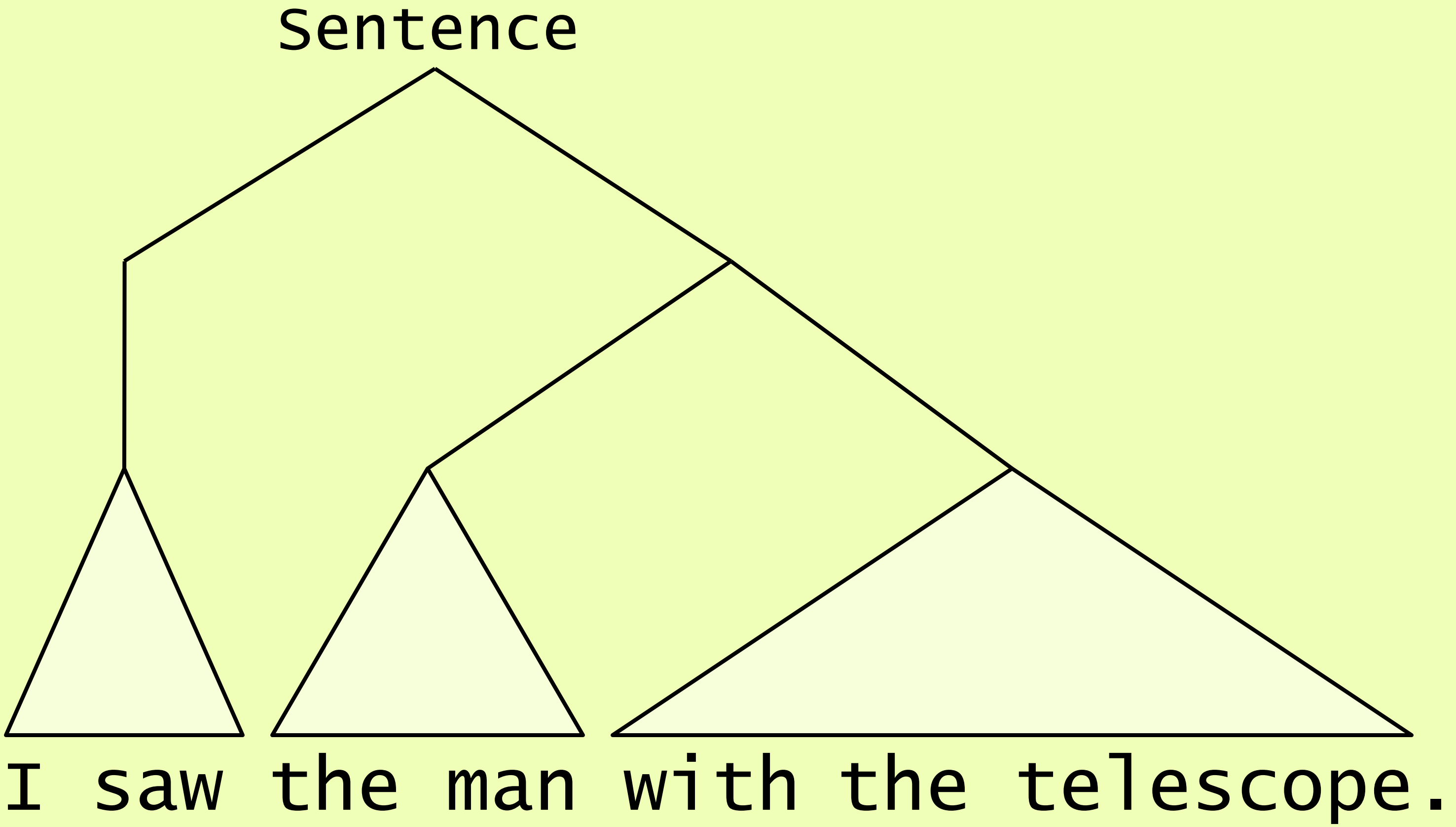
Natural language is ambiguous:



Probabilistic grammars

Motivation

Natural language is ambiguous:

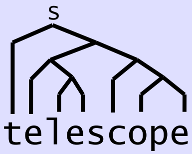
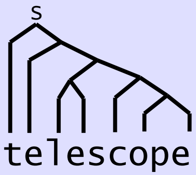


Probabilistic grammars

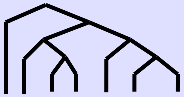
Probabilistic context-free grammars

Context-free grammar: $G = (N, T, S, R)$

+ probability distribution on derivations

e.g. $P($$) = 0.001$, but $P($$) = 0.000001$

Use $p\colon R \to [0,1]$ s.t. $\forall A \in N\colon \sum_{(A\to\alpha)\in R} p(A \to \alpha) = 1$

and get $P($$) = \prod_{(A\to\alpha)\in}$$p(A \to \alpha)$

Probabilistic grammars

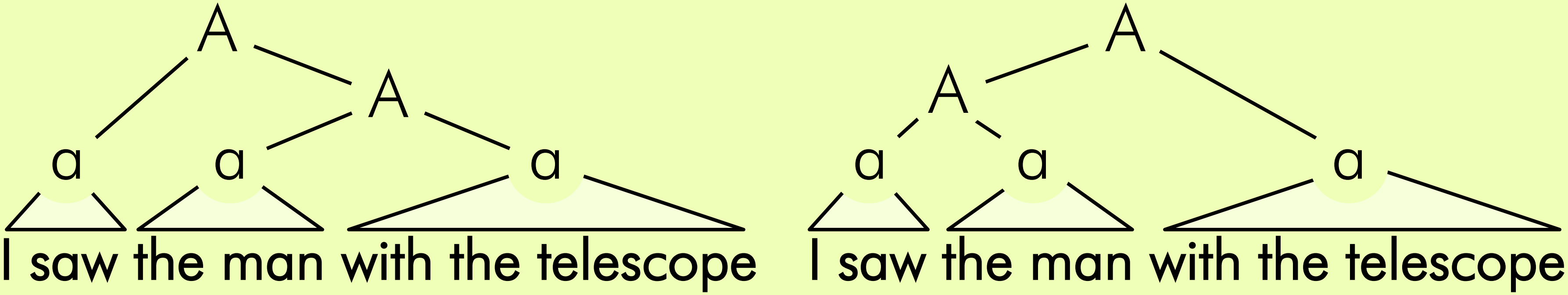
Example PCFG

A → AP

A → PA

A → PP

P → I saw | the man | with the telescope



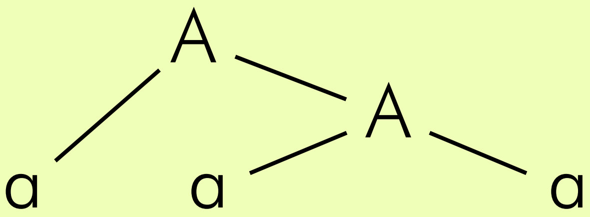
Probabilistic grammars

Example PCFG

$$A \rightarrow Aa \qquad p(A \rightarrow Aa) = 0.4$$

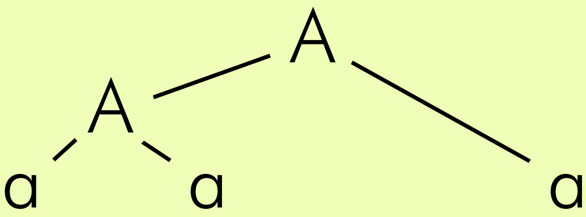
$$A \rightarrow aA \qquad p(A \rightarrow aA) = 0.1$$

$$A \rightarrow aa \qquad p(A \rightarrow aa) = 0.5$$



lower probability

$P = 0.1 \times 0.5 = \mathbf{0.05}$



higher probability

$P = 0.4 \times 0.5 = \mathbf{0.2}$

Probabilistic grammars

Interesting values

Inside probability

Viterbi

Viterbi-derivation

Viterbi-n-best

Outside probability

Telescope grammar

$$p(A \rightarrow Aa) = 0.4$$
$$p(A \rightarrow aA) = 0.1$$
$$p(A \rightarrow aa) = 0.5$$

Example calculations

Input: . a . a . a .
1 2 3 4

inside(1, A, 4)

= P() + P()

$= \quad 0.2 \quad + \quad 0.05$

$= \quad 0.25$

viterbi-derivation(1, A, 4)

= 

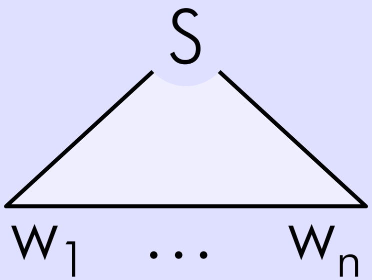
$\text{viterbi}(1, A, 4) = 0.2$

Semirings?

Extending CKY

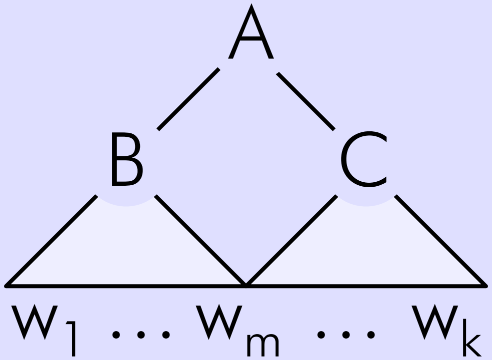
CKY parsing

Input: $w_1 \ldots w_n$; Goal item: $[1, S, n+1]$



Beyond recognition

$[i, A, k]$ provable $\Leftrightarrow V[i, A, k] = \text{true}$



Extending CKY

CKY parsing

Input: $w_1 \ldots w_n$; Goal item: $[1, S, n+1]$

Rules: $\dfrac{(A \to w_i) \in R}{[i, A, i+1]}$, $\dfrac{(A \to BC) \in R \quad [i, B, m] \quad [m, C, k]}{[i, A, k]}$

Beyond recognition

Unary rule: $(A \to w_i) \in R \Rightarrow V[i, A, i+1] = \text{true}$

Binary rule: $(A \to BC) \in R \Rightarrow$

$$V[i, A, k] = V[i, A, k] \vee (V[i, B, m] \wedge V[m, C, k])$$

$\text{success} = V[1, S, n+1]$

Extending CKY

CKY parsing

Input: $w_1 \ldots w_n$; Goal item: $[1, S, n+1]$

Rules: $\dfrac{(A \rightarrow w_i) \in R}{[i, A, i+1]}$, $\dfrac{(A \rightarrow BC) \in R \quad [i, B, m] \quad [m, C, k]}{[i, A, k]}$

Beyond recognition

Unary rule: $(A \rightarrow w_i) \in R \Rightarrow V[i, A, i+1] = p(A \rightarrow w_i)$

Binary rule: $(A \rightarrow BC) \in R \Rightarrow$

$$V[i, A, k] = V[i, A, k] \oplus (V[i, B, m] \otimes V[m, C, k]) \otimes p(A \rightarrow BC))$$

$$\text{inside} = V[1, S, n+1]$$

Semirings

Semiring definition

Recall: field → ring

→ **semiring**

$+ \quad 0 \quad \xcancel{-x} \quad \times \quad 1 \quad \xcancel{x^{-1}}$

Complete semiring: $\sum^{\infty}$ is well-defined

Some semirings

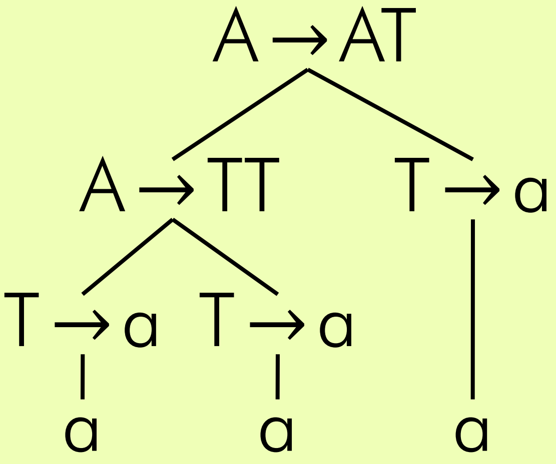
Natural numbers: $\langle \mathbb{N}[0,\infty], +, \times, 0, 1 \rangle$

Reals with max: $\langle \mathbb{R}[0,1], \max, \times, 0, 1 \rangle$

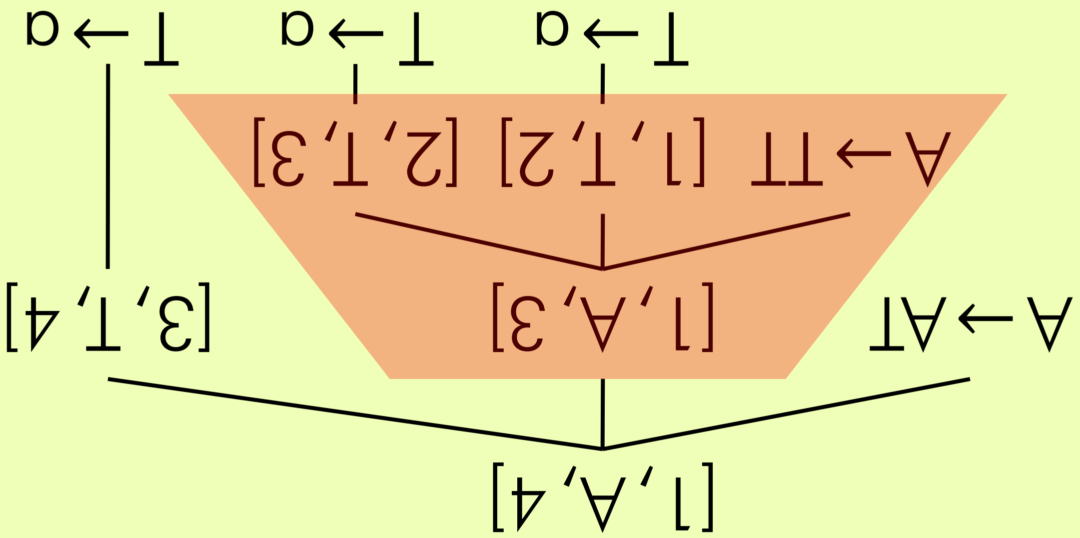
Extending CKY

Derivations

Grammar



Parser

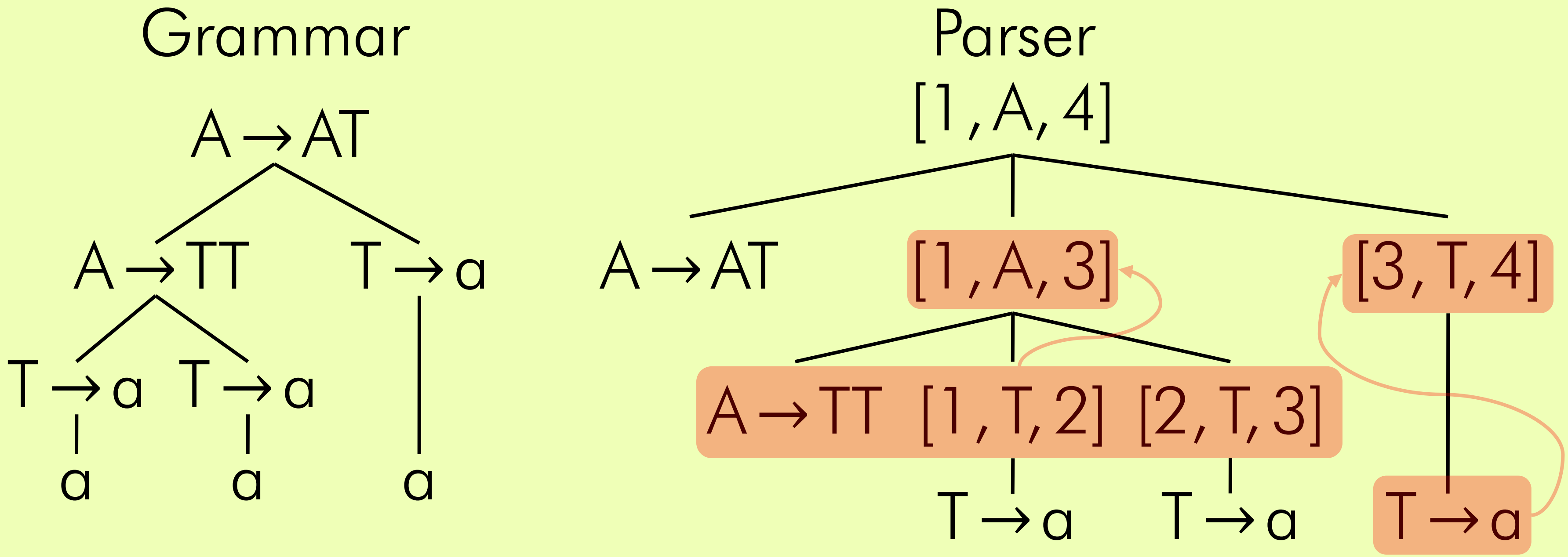


Derivation values

Grammar: Multiply all rule values

Extending CKY

Derivations



Derivation values

Grammar: Multiply all rule values

Parser: Multiply rule values recursively via item values

Semiring computations

Notations

Value of a rule	$R(A \rightarrow BC)$	– from semiring
Grammar derivation	$E = e_1 \ldots e_m$	– list of rules
Item derivation tree	$D = D_1 \ldots D_m$	– leaves are rules

Grammar

Value of a derivation:

$$V_G(E) = \otimes_{i=1}^{m} R(e_i)$$

Word, derivable by $E_1 \ldots E_k$:

$$V_G = \oplus_{i=1}^{k} V_G(E_i)$$

Parser

$$V(D) = \otimes_{d \text{ leaf}} R(d)$$

$$= \begin{cases} R(D) & \text{(leaf node)} \\ \otimes_{i=1}^{m} V(D_i) & \text{(inner node)} \end{cases}$$

Item x, heading $D_1 \ldots D_k$:

$$V(x) = \oplus_{i=1}^{k} V(D_i)$$

Semirings

Useful semirings

Recognition: $\langle\ \{\text{true}, \text{false}\}, \vee, \wedge, \text{false}, \text{true}\ \rangle$

Derivation number: $\langle\ \mathbb{N}[0,\infty], +, \times, 0, 1\ \rangle$

Derivation forest: $\langle\ 2^{\mathbb{E}}, \cup, \cdot, \emptyset, \{\langle\rangle\}\ \rangle$

Inside probability: $\langle\ \mathbb{R}[0,\infty], +, \times, 0, 1\ \rangle$

Viterbi: $\langle\ \mathbb{R}[0,1], \max, \times, 0, 1\ \rangle$

Viterbi-derivation: $\langle\ \mathbb{R}[0,1] \times 2^{\mathbb{E}}, \max_{\text{Vit}}, \times_{\text{Vit}}, \langle 0, \emptyset\rangle, \langle 1, \{\langle\rangle\}\rangle\ \rangle$

Viterbi-n-best: *way too complicated…*

Semiring computations

Derivation forest example

$$\langle 2^{\mathbb{E}}, \cup, \cdot, \emptyset, \{\langle\rangle\} \rangle$$

$$\text{Input:} \quad \underset{1}{.}\, a \,\underset{2}{.}\, a \,\underset{3}{.}\, a \,\underset{4}{.}$$

$$V([1,T,2]) = \{\langle T \rightarrow a \rangle\}$$

$$V([2,T,3]) = \{\langle T \rightarrow a \rangle\}$$

$$V([3,T,4]) = \{\langle T \rightarrow a \rangle\}$$

$$\frac{(T \rightarrow a)}{[i,T,i+1]}$$

Semiring computations

Derivation forest example

$$\langle 2^{\mathbb{E}}, \cup, \cdot, \emptyset, \{\langle\rangle\} \rangle$$

Input: $\underset{1}{.}\,a\,\underset{2}{.}\,a\,\underset{3}{.}\,a\,\underset{4}{.}$

$$V([1,T,2]) = \{\langle T \rightarrow a \rangle\}$$

$$V([2,T,3]) = \{\langle T \rightarrow a \rangle\}$$

$$V([3,T,4]) = \{\langle T \rightarrow a \rangle\}$$

$$V([1,A,3]) = \{\langle A \rightarrow TT, T \rightarrow a, T \rightarrow a \rangle\}$$

$$V([2,A,4]) = \{\langle A \rightarrow TT, T \rightarrow a, T \rightarrow a \rangle\}$$

$$\frac{\begin{array}{c}(A \rightarrow TT)\\ [i,T,m]\\ [m,T,k]\end{array}}{[i,A,k]}$$

Semiring computations

Derivation forest example

$$\langle 2^{\mathbb{E}}, \cup, \cdot, \emptyset, \{\langle\rangle\} \rangle$$

Input: $\quad . \, a \, . \, a \, . \, a \, .$

$\qquad\qquad 1 \quad 2 \quad 3 \quad 4$

$V([1,T,2]) = \{\langle T \to a \rangle\}$

$V([2,T,3]) = \{\langle T \to a \rangle\}$

$V([3,T,4]) = \{\langle T \to a \rangle\}$

$V([1,A,3]) = \{\langle A \to TT, T \to a, T \to a \rangle\}$

$V([2,A,4]) = \{\langle A \to TT, T \to a, T \to a \rangle\}$

$$V([1,A,4]) = \{\langle A \to AT, A \to TT, T \to a, T \to a, T \to a \rangle\}$$
$$\cup \; \{\langle A \to TA, A \to TT, T \to a, T \to a, T \to a \rangle\}$$

$(A \to TA)$

$$\frac{[i,T,m] \quad [m,A,k]}{[i,A,k]}$$

Semiring parsing

Beyond CKY

Works for many parsers *e.g. Earley, but also for TAGs*

Omissions

Outside values *complicated, but similar proofs*

$\sum^{\infty}$ Infinite summation *for $A \rightarrow A$, semiring-dependent*

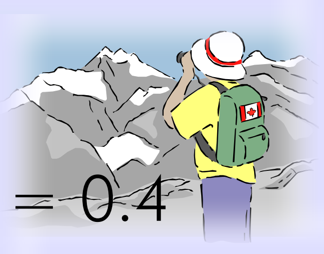
Further reading

Joshua Goodman: *Semiring parsing*

…and his Ph.D. thesis

Semiring parsing

Summary



Natural language processing problems

Probabilistic grammars

$p(A \rightarrow Aa) = 0.4$

Inside probability, Viterbi, ...

Semiring operation substitution

$\oplus \quad \otimes$

Semiring parsing

Summary

Natural language processing problems

Probabilistic grammars

$p(A \to Aa) = 0.4$

Inside probability, Viterbi, …

Semiring operation substitution $\oplus$ $\otimes$

one parser

many values

?