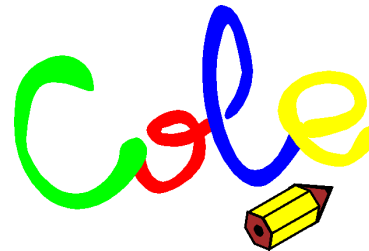


Dynamic Programming of Partial Parses

David Cabrero Souto Jesús Vilares Ferro Manuel Vilares Ferro

Grupo CoLe



Universidad de Vigo & Universidade da Coruña

Introduction

- The challenge of robust parsing:

Introduction

- The challenge of robust parsing:
 - incomplete grammars

Introduction

- The challenge of robust parsing:
 - incomplete grammars
 - ill-formed input

Introduction

- The challenge of robust parsing:
 - incomplete grammars
 - ill-formed input
- ... but there could be enough information.

Introduction

- The challenge of robust parsing:
 - incomplete grammars
 - ill-formed input

... but there could be enough information.
- Solutions:
 - Automatic error recovery
 - No parsing
 - Partial parsing

Approaches to partial parsing

Approaches to partial parsing

- Based on the idea of Chunking [Abney]

Approaches to partial parsing

- Based on the idea of Chunking [Abney]
 - n-phases (usually 2)

Approaches to partial parsing

- Based on the idea of Chunking [Abney]
 - n-phases (usually 2)
 - First phase: parse minimal constituents (chunks)

Approaches to partial parsing

- Based on the idea of Chunking [Abney]
 - n-phases (usually 2)
 - First phase: parse minimal constituents (chunks)
 - Second phase: join chunks

Approaches to partial parsing

- Based on the idea of Chunking [Abney]
 - n-phases (usually 2)
 - First phase: parse minimal constituents (chunks)
 - Second phase: join chunks
- Modify conventional parsers

Modify conventional parsers

1. First of all: Extend grammars

Modify conventional parsers

1. First of all: Extend grammars

• CFG: $\mathcal{G} = (N, \Sigma, P, S)$

Modify conventional parsers

1. First of all: Extend grammars

- CFG: $\mathcal{G} = (N, \Sigma, P, S)$
Nonterminals

Modify conventional parsers

1. First of all: Extend grammars

- CFG: $\mathcal{G} = (N, \Sigma, P, S)$
Nonterminals, Terminals

Modify conventional parsers

1. First of all: Extend grammars

• CFG: $\mathcal{G} = (N, \Sigma, P, S)$

Nonterminals, Terminals, Productions

Modify conventional parsers

1. First of all: Extend grammars

• CFG: $\mathcal{G} = (N, \Sigma, P, S)$

Nonterminals, Terminals, Productions, Initial symbol

Modify conventional parsers

1. First of all: Extend grammars

- CFG $\mathcal{G} = (N, \Sigma, P, S)$

Nonterminals, Terminals, Productions, Initial symbol

- Change to: (N, Σ, P, S)

Set of Initial symbols

Modify conventional parsers

1. Extend grammars

- CFG $\mathcal{G} = (N, \Sigma, P, S)$

Nonterminals, Terminals, Productions, Initial symbol

- Change to: $(N, \Sigma, P, \mathcal{S})$
Set of Initial symbols

2. Extend parser: Entry point [Jacobs]

- Start at any point
- Finish at any point

Descriptive Framework

Descriptive Framework

- Common descriptive framework [Sikkel,Shieber]

Descriptive Framework

- Common descriptive framework [Sikkel,Shieber]
- Key idea: Parsing as deduction

Descriptive Framework

- Common descriptive framework [Sikkel, Shieber]
- Key idea: Parsing as deduction
 - Items
 - Ex. Top-Down [$\bullet\beta, j$]

Descriptive Framework

- Common descriptive framework [Sikkel, Shieber]
- Key idea: Parsing as deduction

- Items

Ex. Top-Down $[\bullet\beta, j]$

- Deduction steps

Ex. Top-Down $\frac{[\bullet B\beta, j]}{[\bullet\gamma\beta, j]} \langle B \rightarrow \gamma \in R \rangle$

$$\frac{[\bullet w_{j+1}\beta, j]}{[\bullet\beta, j + 1]}$$

Descriptive Framework

- Common descriptive framework [Sikkel, Shieber]
- Key idea: Parsing as deduction
 - Items
 - Ex. Top-Down $[\bullet\beta, j]$
 - Deduction steps
 - Ex. Top-Down $\frac{[\bullet B\beta, j]}{[\bullet\gamma\beta, j]} \langle B \rightarrow \gamma \in R \rangle$
 - $\frac{[\bullet w_{j+1}\beta, j]}{[\bullet\beta, j+1]}$
 - Initial (axioms) and final (goals) items.
 - Ex. Top-Down $[\bullet S, 0]$ and $[\bullet, n]$

Partial Top-Down

Axioms: $[\bullet S, 0]$

Goals: $[\bullet, n]$

Partial Top-Down

Axioms: $[\bullet S, 0]$

Goals: $[\bullet, n]$

- Change Initial Symbol (S) by set of Initial Symbols (S)

Partial Top-Down

Axioms: $[\bullet A, 0], A \in \mathcal{S}$

Goals: $[\bullet, n]$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})

Partial Top-Down

Axioms: $[\bullet A, 0], A \in \mathcal{S}$

Goals: $[\bullet, n]$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})
- Start at any point

Partial Top-Down

Axioms:

$$[\bullet A, i], A \in \mathcal{S}, 0 \leq i \leq n$$

Goals: $[\bullet, n]$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})
- Start at any point

Partial Top-Down

Axioms: $[\bullet A, i], A \in \mathcal{S}, 0 \leq i \leq n$

Goals: $[\bullet, n]$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})
- Start at any point
- Finish at any point

Partial Top-Down

Axioms: $[\bullet A, i], A \in \mathcal{S}, 0 \leq i \leq n$

Goals: $[\bullet, j], 0 \leq j \leq n$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})
- Start at any point
- Finish at any point

Partial Top-Down

Axioms: $[\bullet A, i], A \in \mathcal{S}, 0 \leq i \leq n$

Goals: $[\bullet, j], 0 \leq j \leq n$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})
- Start at any point
- Finish at any point
 1. Finish after starting
 2. Remember starting point

Partial Top-Down

Axioms:

$$[\bullet A, i, i], A \in \mathcal{S}, 0 \leq i \leq n$$

Goals:

$$[\bullet, i, j], 0 \leq i \leq j \leq n$$

- Change Initial Symbol (S) by set of Initial Symbols (\mathcal{S})
- Start at any point
- Finish at any point
 1. Finish after starting
 2. Remember starting point

Partial Top-Down (2)

- Next: modify deduction steps

Partial Top-Down (2)

- Next: modify deduction steps

$$\text{Prediction} \quad \frac{[\bullet B\beta, j]}{[\bullet \gamma\beta, j]} \langle B \rightarrow \gamma \in R \rangle$$

$$\text{Scanning} \quad \frac{[\bullet w_{j+1}\beta, j]}{[\bullet \beta, j + 1]}$$

- New in items: starting point

Partial Top-Down (2)

- Next: modify deduction steps

$$\text{Prediction} \quad \frac{[\bullet B\beta, j]}{[\bullet \gamma\beta, j]} \langle B \rightarrow \gamma \in R \rangle$$

$$\text{Scanning} \quad \frac{[\bullet w_{j+1}\beta, j]}{[\bullet \beta, j + 1]}$$

- New in items: starting point
 - Add starting point
 - Pass-through starting point

Partial Top-Down (2)

- Next: modify deduction steps

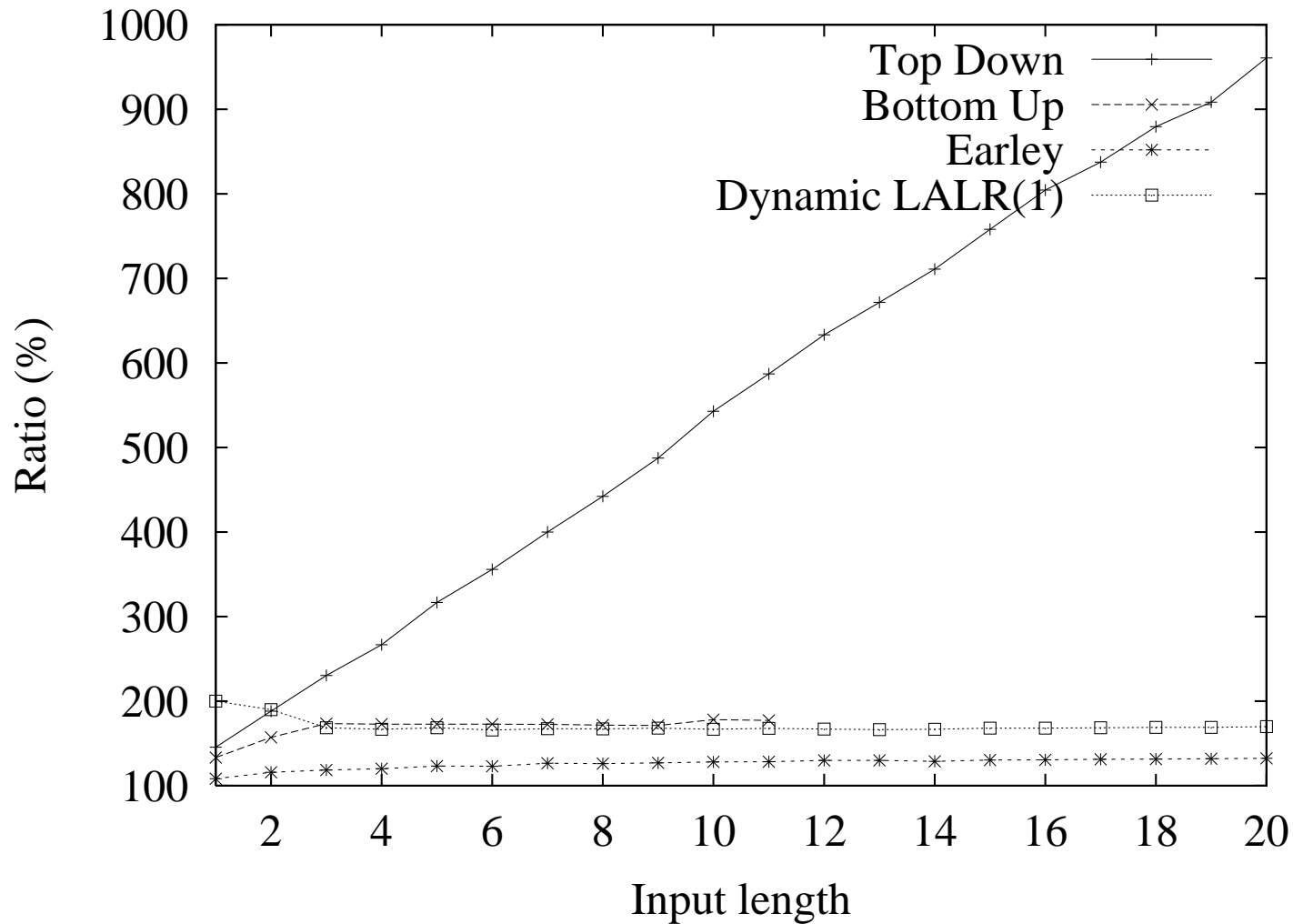
$$\text{Prediction} \quad \frac{[\bullet B\beta, i, j]}{[\bullet \gamma\beta, i, j]} \langle B \rightarrow \gamma \in R \rangle$$

$$\text{Scanning} \quad \frac{[\bullet w_{j+1}\beta, i, j]}{[\bullet \beta, i, j + 1]}$$

- New in items: starting point
 - Add starting point
 - Pass-through starting point

Experimental Results

- Simple grammar: palindromes



Experimental Results (2)

- Results as expected
 - Top-Down: Combinatorial explosion
 - Bottom-Up: No more combinatorial explosion
 - LALR & Earley (mixed strategys): Good mix

Future work

- Real tests
- Proofs of soundness and correctness (TR ?)
- “Syntax sugar”++