Scott Farrar CLMA, University of Washington farrar@u.washington.edu

February 1, 2010

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Today's lecture

Earley algorithm

- Earley: introduction
- Example of Earley algorithm

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

In naive search, top-down parsing is inefficient because structures are created over and over again.

- Need a way to record that a particular structure has been predicted
- Need a way to record *where* the structure was predicted wrt the input

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

 \checkmark Never explores trees that aren't potential solutions, ones with the wrong kind of root node.

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

- $\checkmark\,$ Never explores trees that aren't potential solutions, ones with the wrong kind of root node.
- X But explores trees that do not match the input sentence (predicts input before inspecting input).

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

- \checkmark Never explores trees that aren't potential solutions, ones with the wrong kind of root node.
- X But explores trees that do not match the input sentence (predicts input before inspecting input).
- X Naive top-down parsers never terminate if G contains recursive rules like $X \rightarrow X Y$ (left recursive rules).

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

- \checkmark Never explores trees that aren't potential solutions, ones with the wrong kind of root node.
- X But explores trees that do not match the input sentence (predicts input before inspecting input).
- X Naive top-down parsers never terminate if G contains recursive rules like $X \rightarrow X Y$ (left recursive rules).
- X Backtracking may discard valid constituents that have to be re-discovered later (duplication of effort).

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

- \checkmark Never explores trees that aren't potential solutions, ones with the wrong kind of root node.
- X But explores trees that do not match the input sentence (predicts input before inspecting input).
- X Naive top-down parsers never terminate if G contains recursive rules like $X \rightarrow X Y$ (left recursive rules).
- X Backtracking may discard valid constituents that have to be re-discovered later (duplication of effort).

Use a top-down strategy when you know what kind of constituent you want to end up with (e.g. NP extraction, named entity extraction). Avoid this strategy if you're stuck with a highly recursive grammar.

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

The Earley Parsing Algorithm: an efficient top-down parsing algorithm that avoids some of the inefficiency associated with purely naive search with the same top-down strategy (cf. recursive descent parser).

- Intermediate solutions are created only once and stored in a chart (dynamic programming).
- Left-recursion problem is solved by examining the input.
- Earley is not picky about what type of grammar it accepts, i.e., it accepts arbitrary CFGs (cf. CKY).

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley Parsing Algorithm (J&M, p. 444)

function EARLEY-PARSE(words, grammar) returns chart

```
 \begin{array}{l} \operatorname{Enqueue}((\gamma \rightarrow \bullet S, [0,0]), chart[0]) \\ \text{for } i \leftarrow \text{from 0 to } \operatorname{Length}(words) \text{ do} \\ & \text{for each state in } chart[i] \text{ do} \\ & \text{if } \operatorname{Incomplete}(state) \text{ and } \operatorname{Next-Cat}(state) \text{ is not POS then} \\ & \operatorname{PreDictor}(state) \\ & \text{elseif } \operatorname{Incomplete}(state) \text{ and } \operatorname{Next-Cat}(state) \text{ is POS then} \\ & \operatorname{SCanner}(state) \\ & \text{else} \\ & \operatorname{COmpleter}(state) \\ & \text{end} \\ \\ \text{end} \\ \\ \text{return}(chart) \end{array}
```

◆□▶ ◆□▶ ◆注▶ ◆注▶ 注 のへで

The rationale is to fill in a chart with the solutions to the subproblems encountered in the top-down parsing process.

- Based on an input string of length n, build a 1D array (called a chart) of length n + 1 to record the solutions to subproblems
- Chart entries are lists of **states**, or info about partial solutions.
- States represent attempts to discover constituents.

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Empty Earley chart

Chart[0]:	 astronomers saw stars with ears partial solutions
Chart[1]:	astronomers • saw stars with ears
	partial solutions
Chart[2]:	astronomers saw • stars with ears
	partial solutions
Chart[3]:	astronomers saw stars • with ears
	partial solutions
Chart[4]:	astronomers saw stars with $ullet$ ears
	partial solutions
Chart[5]:	astronomers saw stars with ears $ullet$

Assumed indexing scheme:

 \bullet_0 astronomers \bullet_1 saw \bullet_2 stars \bullet_3 with \bullet_4 ears \bullet_5

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

A state consists of:

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

A state consists of:

• a subtree corresponding to a grammar rule $S \rightarrow NP \ VP$

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

A state consists of:

- a subtree corresponding to a grammar rule $S \rightarrow NP VP$
- info about progress made towards completing this subtree

$$S \rightarrow NP \bullet VP$$

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

A state consists of:

- a subtree corresponding to a grammar rule $S \rightarrow NP VP$
- info about progress made towards completing this subtree

 $S \rightarrow NP \bullet VP$

S the position of the subtree wrt input

 $S \rightarrow NP \bullet VP, [0,3]$

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

A state consists of:

- a subtree corresponding to a grammar rule $S \rightarrow NP VP$
- info about progress made towards completing this subtree

 $S \rightarrow NP \bullet VP$

the position of the subtree wrt input

 $S \rightarrow NP \bullet VP, [0,3]$

 pointers to all contributing states in the case of a parser (cf. recognizer)

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Setting up the Earley algorithm: dotted rules

Definition

A dotted rule is a data structure used in top-down parsing to record parital solutions towards discovering a constituent.

- S → •VP, [0, 0]
 Predict an S will be found which consists of a VP; the S will begin at 0.
- NP → Det Nominal, [1, 2]
 Predict an NP starting at 1; an Det has been found; Nominal is expected next.
- VP → V NP•, [0, 3]
 A VP has been found starting at 0 and spanning to 3; the constituents of VP are V and NP.

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Dotted rules and corresponding graph



・ロト・西ト・ヨト・ヨー シック・

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley: fundamental operations

- **Predict** sub-structure (based on grammar)
- Scan partial solutions for a match
- Complete a sub-structure (i.e., build constituents)

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Sample grammar from J&M

- $S \rightarrow NP VP \qquad NP \rightarrow NP PP$
- $PP \rightarrow P NP \qquad NP \rightarrow N$
- $VP \rightarrow V NP$ $N \rightarrow astronomers$
- $VP \rightarrow VP PP \quad N \rightarrow ears$
- $P \rightarrow with \qquad N \rightarrow stars$
- $V \rightarrow saw$ $N \rightarrow telescopes$

ambiguous, PP attachment

astronomers saw stars with ears

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

How to represent progress towards finding an S node?

Add a dummy rule to grammar:

$$\gamma \rightarrow S$$

This seeds the chart as the base case for recursion.

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

How to represent progress towards finding an S node?

Add a dummy rule to grammar:

$$\gamma \rightarrow \bullet S$$

This seeds the chart as the base case for recursion.

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Chart[0]: "• astronomers saw stars with ears" state ID dotted rule position back pointer operation

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

・ロト ・四ト ・ヨト ・ヨー うへぐ

Chart[0]: "• astronomers saw stars with ears" state ID dotted rule position back pointer operation

Enqueue dummy start state

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Chart[0]:	: "• astro	nomers sa	w stars wi	th ears"
state D	dotted rule	position	back pointer	operation
S0	$TOP ightarrow \bullet S$	[0,0]	[]	Seed

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Chart[0]:	: "• astror	nomers sa	w stars wi	th ears"
state D	dotted rule	position	back pointer	operation
S0	TOP ightarrow ullet S	[0,0]	[]	Seed

Scott Farrar CLMA, University of Washington farrar@u.washington.ec

Earley algorithm Earley: introduction Example of Earley algorithm

For each state in chart[i]... processing S0

Chart[0]:	👘 👋 👋 👋	nomers sa	w stars wi	th ears"
state D	dotted rule	position	back pointer	operation
S0	$TOP ightarrow \bullet S$	[0,0]	[]	Seed

if state incomplete and $\operatorname{NextCat}$ is not POS, then $\operatorname{Predictor}$

```
procedure PREDICTOR((A \rightarrow \alpha \bullet B \beta, [i,j]))
for each (B \rightarrow \gamma) in GRAMMAR-RULES-FOR(B,grammar) do
ENQUEUE((B \rightarrow \bullet \gamma, [j,j]),chart[j])
end
```

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Chart[0]	: "• astror	nomers sa	w stars wi	th ears"
state D	dotted rule	position	back pointer	operation
S0	$TOP \rightarrow \bullet S$	[0,0]	[]	Seed

if state incomplete and $\operatorname{NextCat}$ is not POS, then $\operatorname{Predictor}$

procedure
$$PreDictor((A \rightarrow \alpha \bullet B \beta, [i, j]))$$

for each $(B \rightarrow \gamma)$ in GRAMMAR-RULES-FOR $(B, grammar)$ do
 $ENQUEUE((B \rightarrow \bullet \gamma, [j, j]), chart[j])$
end

$$S \rightarrow NP VP \qquad NP \rightarrow NP PP$$

$$PP \rightarrow P NP \qquad NP \rightarrow N$$

$$VP \rightarrow V NP$$
 $N \rightarrow astronomers$

$$VP \rightarrow VP PP \quad N \rightarrow ears$$

$$P \rightarrow with \qquad N \rightarrow stars$$

$V \rightarrow saw$ $N \rightarrow telescopes$

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

◆□▶ ◆□▶ ◆□▶ ◆□▶ ◆□ ◆ ○ ◆

C	hart[0]	: "• astron	omers sa	w stars wi	th ears"
	state ID	dotted rule	position	back pointer	operation
	S0	$TOP ightarrow \bullet S$	[0,0]	[]	Seed
	S1	$S ightarrow \bullet NP VP$	[0,0]	[]	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

C	hart[0]	: "• astrono	omers sa	w stars wi	th ears"
	state ID	dotted rule	position	back pointer	operation
	S0	$TOP \rightarrow \bullet S$	[0,0]	[]	Seed
	S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

processing S1

Chart[0]	: "• astrono	omers sa	w stars wi	th ears"
state ID	dotted rule	position	back pointer	operation
S0	$TOP ightarrow \bullet S$	[0,0]	[]	Seed
S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor

if state incomplete and $\operatorname{NextCat}$ is not POS, then $\operatorname{Predictor}$

procedure PREDICTOR(
$$(A \rightarrow \alpha \bullet B \beta, [i,j])$$
)
for each $(B \rightarrow \gamma)$ in GRAMMAR-RULES-FOR(*B*,grammar) do
ENQUEUE($(B \rightarrow \bullet \gamma, [j,j])$,chart[j])
end

- $S \rightarrow NP VP \qquad NP \rightarrow NP PP$
- $PP \rightarrow P NP \qquad NP \rightarrow N$
- $VP \rightarrow V NP$ $N \rightarrow astronomers$
- $VP \rightarrow VP PP \quad N \rightarrow ears$
- $P \rightarrow with \qquad N \rightarrow stars$
- $V \rightarrow saw$ $N \rightarrow telescopes$

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

・ロト ・日 ・ ・ ヨ ・ ・ 日 ・ うへぐ

Chart[0]:	 "• astrono 	mers sa	iw stars wi	th ears"
state D	dotted rule	position	back pointer	operation
S0	$TOP \rightarrow \bullet S$	[0,0]	[]	Seed
S1	$S ightarrow \bullet NP VP$	[0,0]	[]	Predictor
S2	$NP \rightarrow \bullet NP PP$	[0,0]	[]	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

С	hart[0]:	 "• astrono 	mers sa	iw stars wi	th ears"
	state ID	dotted rule	position	back pointer	operation
	S0	TOP ightarrow ullet S	[0,0]	[]	Seed
	S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor
	S2	$NP \rightarrow \bullet NP PP$	[0,0]	[]	Predictor
	S3	NP ightarrow ullet N	[0,0]	[]	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Cl	nart[0]:	"• astrono	mers sa	w stars wi	th ears"
	state D	dotted rule	position	back pointer	operation
	S0	$TOP \rightarrow \bullet S$	[0,0]	[]	Seed
	S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor
	S2	$NP \rightarrow \bullet NP PP$	[0,0]	[]	Predictor
	S3	NP ightarrow ullet N	[0,0]	Ū	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

processing S2... but NP at position [0,0] has already been considered 233

Chart[0]: "• astronomers saw stars with ears"

state ID	dotted rule	position	back pointer	operation
S0	$TOP \rightarrow \bullet S$	[0,0]	[]	Seed
S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor
S2	NP ightarrow ullet NP PP	[0,0]	Ū	Predictor
S3	NP ightarrow ullet N	[0,0]	[]	Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed
state ID	dotted rule	position	back pointer	operation
S0	TOP ightarrow ullet S	[0,0]	[]	Seed
S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor
S2	$NP \rightarrow \bullet NP PP$	[0,0]	[]	Predictor
S 3	NP ightarrow ullet N	[0,0]	0	Predictor

processing S3 ...

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S0	$TOP \rightarrow \bullet S$	[0,0]	[]	Seed
S1	$S \to \bullet NP VP$	[0,0]	[]	Predictor
S2	NP ightarrow ullet NP PP	[0,0]	Ō	Predictor
S3	NP ightarrow ullet N	[0,0]	Ō	Predictor
if state inco	omplete and Nevt(at is a PO	S then Scanner	

if state incomplete and NextCat is a POS, then Scanner

procedure SCANNER($(A \rightarrow \alpha \bullet B \ \beta, [i,j])$) if $B \subset POS(word[j])$ then ENQUEUE($(B \rightarrow word[j], [j,j+1])$, chart[j+1])

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S0	TOP ightarrow ullet S	[0,0]	[]	Seed
S1	$S ightarrow \bullet NP VP$	[0,0]	[]	Predictor
S2	$NP \rightarrow \bullet NP PP$	[0,0]	[]	Predictor
S3	$NP \rightarrow \bullet N$	[0,0]	[]	Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

Scanner: Add $N \rightarrow astronomers \bullet [0,1]$ at Chart[0+1]

Chart[1]: state ID	"astronomers dotted rule	• SaW position	stars with back pointer	ears" operation	Earley algorithm Scott Farrar
state ID S4	dotted rule $N \rightarrow astronomers \bullet$	position [0,1]	l	operation Scanner	Scott Farrar CLMA, University of Washington far- rar@u.washington.ed Earley algorithm Earley: introduction Example of Earley algorithm

Chart[1]: state ID	"astronomers dotted rule	• SaW position	stars with back pointer	ears" operation	Earley algorithm Scott Farrar
state ID S4	dotted rule $N \rightarrow astronomers \bullet$	position [0,1]	l	operation Scanner	Scott Farrar CLMA, University of Washington far- rar@u.washington.ed Earley algorithm Earley: introduction Example of Earley algorithm





Which states from Chart[0] require the current constit. to be complete?

C	hart[1]:	"astronomers	• saw	stars with	ears"
	state ID	dotted rule	position	back pointer	operation
	S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
	S5	NP ightarrow N ullet	[0,1]	[S4]	Completer

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

С	hart[1]:	"astronomers	• saw	stars with	ears"
	state ID	dotted rule	position	back pointer	operation
	S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
	S5	NP ightarrow N ullet	[0,1]	[S4]	Completer

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

processing S5...else COMPLETER



Which states from Chart[0] require the current constit. to be complete?

Chart[1]: "astronomers	• saw	stars with	ears"
state ID	dotted rule	position	back pointer	operation
S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
S6	S ightarrow NP ullet VP	[0,1]	[S5]	Completer

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

"astronomers • saw stars with ears" Chart[1]: state D dotted rule position back pointer operation S4 Scanner $N \rightarrow astronomers \bullet$ [0,1] [] S5 $NP \rightarrow N \bullet$ [S4] Completer [0,1]S6 $S \rightarrow NP \bullet VP$ [S5] [0,1]Completer S7 $NP \rightarrow NP \bullet PP$ [0,1][S5] Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

C	hart[1]:	"astronomers	• saw	stars with	ears"
	state ID	dotted rule	position	back pointer	operation
	S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
	S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
	S6	$S \rightarrow NP \bullet VP$	[0,1]	[S5]	Completer
	S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

processing S6...PREDICTOR

Chart[1]:	"astronomers	• saw	stars with	ears"
state D	dotted rule	position	back pointer	operation
S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
S6	$S \rightarrow NP \bullet VP$	[0,1]	[S5]	Completer
S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer
S8	$VP \rightarrow \bullet V NP$	[1,1]	0	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Chart[1]:	"astronomers	• saw	stars with	ears"
state D	dotted rule	position	back pointer	operation
S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
S6	S ightarrow NP ullet VP	[0,1]	[S5]	Completer
S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer
S8	$VP \rightarrow \bullet V NP$	[1,1]	0	Predictor
S9	$VP \rightarrow \bullet VP PP$	[1,1]	Ō	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Chart[1]:	"astronomers	• saw	stars with	ears"
state D	dotted rule	position	back pointer	operation
S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
S6	S ightarrow NP ullet VP	[0,1]	[S5]	Completer
S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer
S8	$VP \rightarrow \bullet V NP$	[1,1]	0	Predictor
S9	$VP \rightarrow \bullet VP PP$	[1,1]	Ō	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

C	hart[1]:	"astronomers	• saw	stars with	ears"
	state ID	dotted rule	position	back pointer	operation
	S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
	S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
	S6	$S \rightarrow NP \bullet VP$	[0,1]	[S5]	Completer
	S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer
	S8	$VP \rightarrow \bullet V NP$	[1,1]	0	Predictor
	S9	$VP \rightarrow \bullet VP PP$	[1,1]	Ō	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

processing S7...PREDICTOR

Chart[1]: "astronomers • saw stars with ears" state ID dotted rule position back pointer operation S4 $N \rightarrow astronomers \bullet$ [0,1][] Scanner S5 [S4] $NP \rightarrow N \bullet$ [0,1]Completer S6 $S \rightarrow NP \bullet VP$ [S5] [0,1]Completer S7 $NP \rightarrow NP \bullet PP$ [0,1][S5] Completer S8 $VP \rightarrow \bullet V NP$ [1,1][] Predictor [] S9 $VP \rightarrow \bullet VP PP$ [1,1]Predictor Π S10 $PP \rightarrow \bullet P NP$ [1,1]Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

С	hart[1]:	"astronomers	• saw	stars with	ears"
	state D	dotted rule	position	back pointer	operation
	S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
	S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
	S6	$S \rightarrow NP \bullet VP$	[0,1]	[S5]	Completer
	S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer
	S8	$VP \rightarrow \bullet V NP$	[1,1]	0	Predictor
	S9	$VP \rightarrow \bullet VP PP$	[1,1]	Ō	Predictor
	S10	$PP \rightarrow \bullet P NP$	[1,1]	Π	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ee

Earley algorithm Earley: introduction Example of Earley algorithm

Processing S8...SCANNER: Add $V \rightarrow saw \bullet [1,2]$ at Chart[1+1]

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ○ □ ○ ○ ○ ○

С	hart[1]:	"astronomers	• saw	stars with	ears"
	state D	dotted rule	position	back pointer	operation
	S4	N ightarrow astronomers ullet	[0,1]	[]	Scanner
	S5	NP ightarrow N ullet	[0,1]	[S4]	Completer
	S6	$S \rightarrow NP \bullet VP$	[0,1]	[S5]	Completer
	S7	$NP \rightarrow NP \bullet PP$	[0,1]	[S5]	Completer
	S8	$VP \rightarrow \bullet V NP$	[1,1]	0	Predictor
	S9	$VP \rightarrow \bullet VP PP$	[1,1]	Ō	Predictor
	S10	$PP \rightarrow \bullet P NP$	[1,1]	Π	Predictor

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

Processing S9...PREDICTOR, but VP already expanded in Chart[1]

Chart[1]: "astronomers • saw stars with ears" state ID dotted rule position back pointer operation S4 $N \rightarrow astronomers \bullet$ [0,1][] Scanner S5 [S4] $NP \rightarrow N \bullet$ [0,1]Completer S6 $S \rightarrow NP \bullet VP$ [S5] [0,1]Completer S7 $NP \rightarrow NP \bullet PP$ [0,1][S5] Completer S8 $VP \rightarrow \bullet V NP$ [1,1][] Predictor [] S9 $VP \rightarrow \bullet VP PP$ [1,1]Predictor $PP \rightarrow \bullet P NP$ Π S10 [1,1]Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

Processing S10...SCANNER, but no P in input at words[1]

state ID dotted rule position back pointer operation

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Example of Earley algorithm

state ID	dotted rule	position	back pointer	operation
S11	V ightarrow saw ullet	[1,2]	[]	Scanner

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S11	V ightarrow saw ullet	[1,2]	[]	Scanner
S12	$VP \rightarrow V \bullet NP$	[1,2]	[S11]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ee

state ID	dotted rule	position	back pointer	operation
S11	V ightarrow saw ullet	[1,2]	[]	Scanner
S12	VP o V ullet NP	[1,2]	[S11]	Completer
S13	$NP \rightarrow \bullet NP PP$	[2,2]	[]	Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ee

state ID	dotted rule	position	back pointer	operation
S11	V ightarrow saw ullet	[1,2]	[]	Scanner
S12	VP o V ullet NP	[1,2]	[S11]	Completer
S13	$NP \rightarrow \bullet NP PP$	[2,2]	[]	Predictor
S14	NP ightarrow ullet N	[2,2]	0	Predictor

state ID	dotted rule	position	back pointer	operation
S15	N o stars ullet	[2,3]	[]	Scanner

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S15	N ightarrow stars ullet	[2,3]	[]	Scanner
S16	NP ightarrow N ullet	[2,3]	[S15]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S15	N ightarrow stars ullet	[2,3]	[]	Scanner
S16	NP ightarrow N ullet	[2,3]	[S15]	Completer
S17	VP o V NP ullet	[1,3]	[S11,S16]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S15	N ightarrow stars ullet	[2,3]	[]	Scanner
S16	NP ightarrow N ullet	[2,3]	[S15]	Completer
S17	VP ightarrow V $NP ightarrow$	[1,3]	[S11,S16]	Completer
S18	$NP \rightarrow NP \bullet PP$	[2,3]	[S16]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ee

state ID	dotted rule	position	back pointer	operation
S15	N ightarrow stars ullet	[2,3]	[]	Scanner
S16	NP ightarrow N ullet	[2,3]	[S15]	Completer
S17	VP ightarrow V NP ullet	[1,3]	[S11,S16]	Completer
S18	NP ightarrow NP ullet PP	[2,3]	[S16]	Completer
S19	S o NP VP ullet	[0,3]	[S5,S17]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S15	N ightarrow stars ullet	[2,3]	[]	Scanner
S16	NP ightarrow N ullet	[2,3]	[S15]	Completer
S17	VP ightarrow V NP ullet	[1,3]	[S11,S16]	Completer
S18	$NP \rightarrow NP \bullet PP$	[2,3]	[S16]	Completer
S19	S o NP VP ullet	[0,3]	[S5,S17]	Completer
S20	VP ightarrow VP ullet PP	[1,3]	[S17]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S15	N ightarrow stars ullet	[2,3]	[]	Scanner
S16	NP ightarrow N ullet	[2,3]	[S15]	Completer
S17	VP ightarrow V $NP ightarrow$	[1,3]	[S11,S16]	Completer
S18	NP ightarrow NP ullet PP	[2,3]	[S16]	Completer
S19	S o NP VP ullet	[0,3]	[S5,S17]	Completer
S20	VP ightarrow VP ullet PP	[1,3]	[S17]	Completer
S21	$PP \rightarrow \bullet P NP$	[3,3]	[]	Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S22	P ightarrow with ullet	[3,4]	[]	Scanner

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S22	P ightarrow with ullet	[3,4]	[]	Scanner
S23	$PP \rightarrow P \bullet NP$	[3,4]	[S19]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S22	P ightarrow with ullet	[3,4]	[]	Scanner
S23	$PP \rightarrow P \bullet NP$	[3,4]	[S19]	Completer
S24	NP ightarrow ullet NP PP	[4,4]	0	Predictor
Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S22	P ightarrow with ullet	[3,4]	[]	Scanner
S23	$PP \rightarrow P \bullet NP$	[3,4]	[S19]	Completer
S24	$NP \rightarrow \bullet NP PP$	[4,4]	[]	Predictor
S25	NP ightarrow ullet N	[4,4]	[]	Predictor

state ID dotted rule	position	back pointer	operation	
----------------------	----------	--------------	-----------	--

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP o P NP \bullet$	[3,5]	[S19,S27]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	PP ightarrow P NP ullet	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	PP ightarrow P NP ullet	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V $NP ightarrow$	[1,5]	[S11,S29]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V NP ullet	[1,5]	[S11,S29]	Completer
S32	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V NP ullet	[1,5]	[S11,S29]	Completer
S32	NP ightarrow NP ullet PP	[2,5]	[S29]	Completer
S33	VP ightarrow V NP ullet	[2,5]	[S11,S29]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V NP ullet	[1,5]	[S11,S29]	Completer
S32	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer
S33	VP ightarrow V NP ullet	[2,5]	[S11,S29]	Completer
S34	NP ightarrow NP ullet PP	[2,5]	[S29]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	$N \rightarrow ears \bullet$	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V NP ullet	[1,5]	[S11,S29]	Completer
S32	NP ightarrow NP ullet PP	[2,5]	[S29]	Completer
S33	VP ightarrow V NP ullet	[2,5]	[S11,S29]	Completer
S34	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer
S35	S ightarrow NP VP ullet	[0,5]	[S5,S30]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V NP ullet	[1,5]	[S11,S29]	Completer
S32	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer
S33	VP ightarrow V NP ullet	[2,5]	[S11,S29]	Completer
S34	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer
S35	S o NP VP ullet	[0,5]	[S5,S30]	Completer
S36	S o NP VP ullet	[0,5]	[S5,S31]	Completer

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

state ID	dotted rule	position	back pointer	operation
S26	N ightarrow ears ullet	[4,5]	[]	Scanner
S27	NP ightarrow N ullet	[4,5]	[S26]	Completer
S28	$PP \rightarrow P NP \bullet$	[3,5]	[S19,S27]	Completer
S29	NP ightarrow NP PP ullet	[2,5]	[S16,S28]	Completer
S30	VP ightarrow VP PP ullet	[1,5]	[S17,S28]	Completer
S31	VP ightarrow V NP ullet	[1,5]	[S11,S29]	Completer
S32	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer
S33	VP ightarrow V NP ullet	[2,5]	[S11,S29]	Completer
S34	$NP \rightarrow NP \bullet PP$	[2,5]	[S29]	Completer
S35	S ightarrow NP VP ullet	[0,5]	[S5,S30]	Completer
S36	S ightarrow NP VP ullet	[0,5]	[S5,S31]	Completer
S37	$PP \rightarrow \bullet P NP$	[5,5]	[]	Predictor

Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Parse 1



Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Parse 2



Earley algorithm

Scott Farrar CLMA, University of Washington farrar@u.washington.ed

Earley algorithm Earley: introduction Example of Earley algorithm

・ロ・・聞・・思・・思・・ しゃくの