

Parsing with Context-Free Grammars



Berlin Chen

Graduate Institute of Computer Science & Information Engineering
National Taiwan Normal University

References:

1. *Natural Language Understanding*, chapter 3 (3.1~3.4, 3.6)
2. *Speech and Language Processing*, chapters 9, 10

Grammars and Sentence Structures

- Describe the structure of sentences and explore ways of characterizing all the legal structures in a language
 - How a sentence is broken into its major subparts/constituents ?
 - E.g., *John ate the cat*

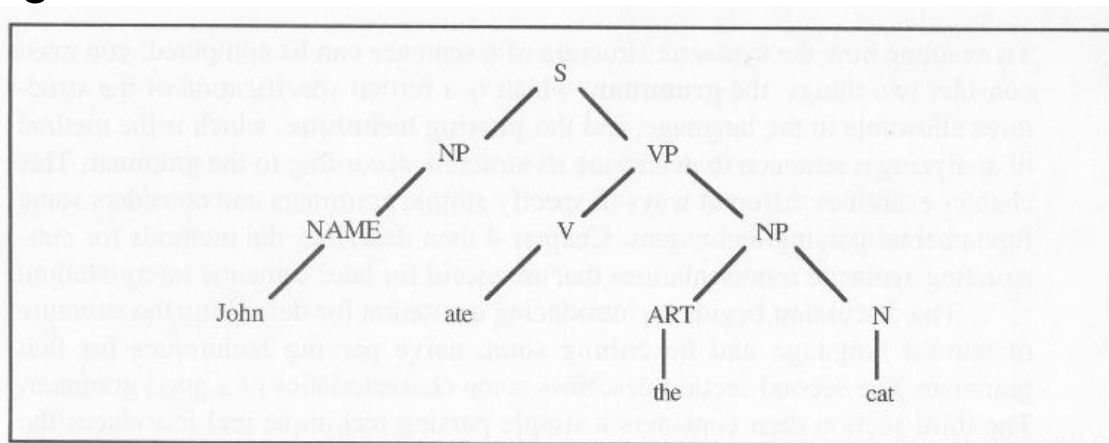


Figure 3.1 A tree representation of *John ate the cat*

(S (NP (NAME John)
(VP (V ate)
(NP (ART the)
(N cat))))

| | |
|---------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 5. $NAME \rightarrow John$ |
| 2. $VP \rightarrow V NP$ | 6. $V \rightarrow ate$ |
| 3. $NP \rightarrow NAME$ | 7. $ART \rightarrow the$ |
| 4. $NP \rightarrow ART N$ | 8. $N \rightarrow cat$ |

Grammar 3.2 A simple grammar

Rewrite rules

Context-Free Grammars (CFGs)

- Grammars consist entirely of rules with a single symbol on the left-hand side
- Formalized by Chomsky(1956), and Backus (1959)
 - Also called Backus-Naur Form (BNF)
- Also called phrase-structure grammars
- The most commonly used mathematical system for modeling the constituent structure in natural languages
 - Ordering
 - What are the rules that govern the ordering of words and bigger units in the language
 - Constituency
 - How do words group into units and what do we say about how the various kinds of units behave

Major Characteristics of CFGs (1/3)

- CFG examples
 - Consist of a set of rules (productions)

NP → *Det Nominal*

NP → *ProperNoun*

Nominal → *Noun | Noun Nominal*

Det → *a*

Det → *the*

Noun → *flight*

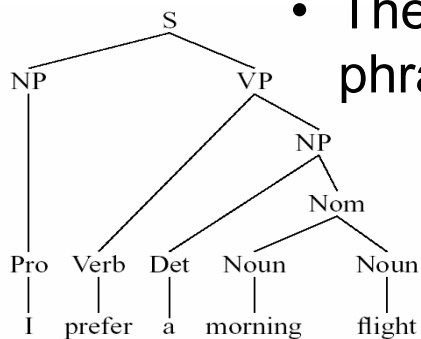
Rewrite the symbol on the left with
a string of symbols on the right

mother

a book club (*NP* → *Det Nominal*,
Nominal → *Noun Nominal*,
Nominal → *Noun*, *Det* → *a*,
Noun → *book*, *Noun* → *club*)

Major Characteristics of CFGs (2/3)

- Symbols used are divided into two classes: terminal and non-terminal symbols
 - A single **non-terminal symbols** on the left side of the arrow (\rightarrow) while one or more terminal or non-terminal symbols on the right side
 - The **terminal symbol** is a word, while in the lexicon, the **non-terminal symbol** associated with each word is its lexical category, or part-of-speech
 - The non-terminal symbol can be a larger constituent (e.g. a phrasal unit) in addition to the lexical category



Major Characteristics of CFGs (3/3)

- The notion of context in CFGs has nothing to do with the ordinary meaning of the **word context** in language
- All it really means that the non-terminal on the left-hand side of a rule is out of there all by itself

$$A \rightarrow B C$$

- We can rewrite an A as a B followed by a C regardless of the context in which A is found

Generation and Parsing (1/2)

- CFGs can be thought of as a device for *generating sentences* or a device for *assigning a structure to a given sentence* (i.e. parsing)
 - **Sentence generation**
 - Start from the S symbol, randomly choose and apply rewrite rules (or productions), until a sequence of words is generated
 - **Parsing**
 - Identify the structure of a sentence given a grammar
 - Top-down or bottom-up strategies

Generation and Parsing (2/2)

- Generation

| | |
|---------------------|------------------|
| S | |
| => NP VP | (rewriting S) |
| => NAME VP | (rewriting NP) |
| => John VP | (rewriting NAME) |
| => John V NP | (rewriting VP) |
| => John ate NP | (rewriting V) |
| => John ate ART N | (rewriting NP) |
| => John ate the N | (rewriting ART) |
| => John ate the cat | (rewriting NP) |

| | |
|---------------|----------------|
| 1. S → NP VP | 5. NAME → John |
| 2. VP → V NP | 6. V → ate |
| 3. NP → NAME | 7. ART → the |
| 4. NP → ART N | 8. N → cat |

- Parsing (with a top-down strategy)

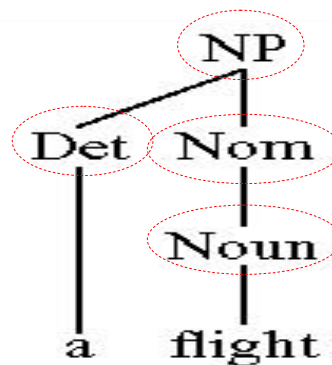
| | |
|---------------------|-------------------|
| John ate the cat | |
| => NAME ate the cat | (rewriting John) |
| => NAME V the cat | (rewriting ate) |
| => NAME V ART cat | (rewriting the) |
| => NAME V ART N | (rewriting cat) |
| => NP V ART N | (rewriting NAME) |
| => NP V NP | (rewriting ART N) |
| => NP VP | (rewriting V NP) |
| => S | (rewriting NP VP) |

Major Constituents of a CFG

- A CFG G has four parameters (“4-tuple”)
 1. A set of non-terminal symbols (or “variables”) N
 2. A set of terminal symbols Σ (disjoint from N)
 3. A set of productions P , each of the form $A \rightarrow \alpha$, where A is a non-terminal symbol and α is a string of symbols from the infinite set of strings $(\Sigma \cup N)^*$
 4. A designated start symbol S (or N^1)
- CFG is a generative grammar
 - The language is defined by the set of possible sentences “generated” by the grammar
 - Via the concept of “derivation”

Derivations

- Derivation: a sequence of rules applied to a string that accounts for that string
 - Can be represented by a parse tree
 - E.g. a parse tree for “a flight”



A derivation represented by a parse tree

- But, usually languages are derivable from the designated start symbol (S)
 - The “sentence” node
 - The set of strings derivable from S called sentences

Derivations as Trees

- Directly Derive

- Directly derive:

$$A \rightarrow \beta, \alpha A \gamma \Rightarrow \alpha \beta \gamma$$

- Derive *directly derives*

where $\alpha, \beta, r, \alpha_i \in (\Sigma \cup N), m \geq 1$

$$\alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{m-1} \Rightarrow \alpha_m \quad \therefore \alpha_1 \overset{*}{\underset{\text{derives}}{\Rightarrow}} \alpha_m$$

- **Parsing:** the process of taking a string and a grammar and returning a (or many) parse tree(s) for that string
 - Analogous to running a finite-state transducer with a tape
 - But, is it more powerful (?)

More Complex Derivations (1/3)

- $S \rightarrow NP VP$
 - Units S , NP , and VP are in the language
 - S consists of an NP followed immediately by a VP
 - There may be many kinds of S
 - NPs and VPs can occur at other places (on the left sides) of the set of rules

- E.g.

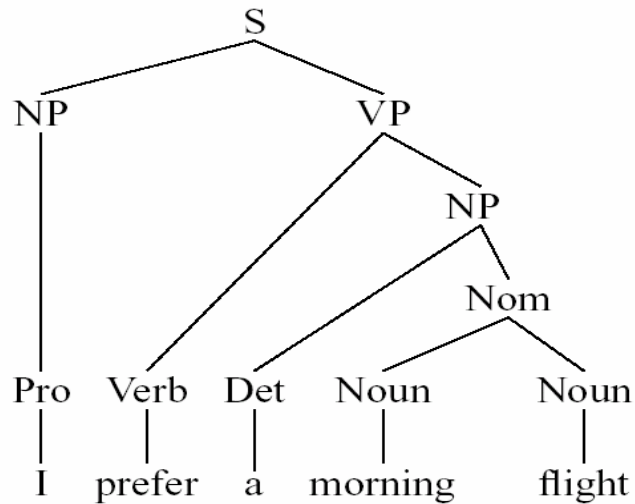
$NP \rightarrow Det Nominal$

$NP \rightarrow ProperNoun$

$VP \rightarrow Verb NP$

$VP \rightarrow Verb NP PP$

More Complex Derivations (2/3)



Bracketed notation/List notation

$[S [NP[Pro I]][VP[V prefer][NP[Det a][Nom [N morning][N flight]]]]]$

| | |
|------------------------------------|---------------------------------|
| $S \rightarrow NP VP$ | I + want a morning flight |
| $NP \rightarrow Pronoun$ | I |
| <i>Proper-Noun</i> | Los Angeles |
| <i>Det Nominal</i> | a + flight |
| $Nominal \rightarrow Noun Nominal$ | morning + flight |
| <i>Noun</i> | flights |
| $VP \rightarrow Verb$ | do |
| <i>Verb NP</i> | want + a flight |
| <i>Verb NP PP</i> | leave + Boston + in the morning |
| <i>Verb PP</i> | leaving + on Thursday |
| $PP \rightarrow Preposition NP$ | from + Los Angeles |

| | |
|--------------------|---|
| <i>Noun</i> | \rightarrow flights breeze trip morning ... |
| <i>Verb</i> | \rightarrow is prefer like need want fly |
| <i>Adjective</i> | \rightarrow cheapest non-stop first latest other direct ... |
| <i>Pronoun</i> | \rightarrow me I you it ... |
| <i>Proper-Noun</i> | \rightarrow Alaska Baltimore Los Angeles Chicago United American ... |
| <i>Determiner</i> | \rightarrow the a an this these that ... |
| <i>Preposition</i> | \rightarrow from to on near ... |
| <i>Conjunction</i> | \rightarrow and or but ... |

More Complex Derivations (3/3)

- Recursion

- The non-terminal on the left also appears somewhere on the right (directly or indirectly)

NP → NP PP [[The flight] [to Boston]]

VP → VP PP [[departed Miami] [at noon]]

- E.g.

- flights from Denver
- Flights from Denver to Miami
- Flights from Denver to Miami in February
- Flights from Denver to Miami in February on Friday

Sentence-level Construction of English

- Declaratives: A plane left.
 $S \rightarrow NP VP$
- Imperatives: Show the lowest fare.
 $S \rightarrow VP$
- Yes-No Questions: Did the plane leave?
 $S \rightarrow Aux NP VP$
- WH Questions: When did the plane leave?
 $S \rightarrow WH Aux NP VP$

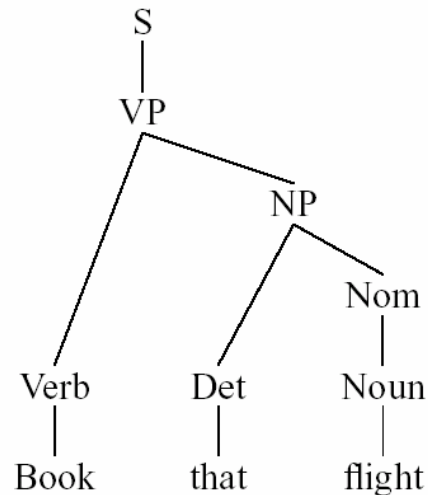
Parsing Strategies

- Top-Down Parsing
 - Start with the S symbol and search through different ways to rewrite the symbols until the input sentence is generated, or until all possibilities have been explored
- Bottom-Up Parsing
 - Start with the words in the input sentence and use the rewrite rules backward to **reduce the sequence of symbols** until it consists solely of S
 - **The left side of each rule is used to rewrite the symbols on the right side**
 - Take **a sequence of symbols** and match it to the right side of the rule

Parsing Strategies (1/2)

Parsing as Search

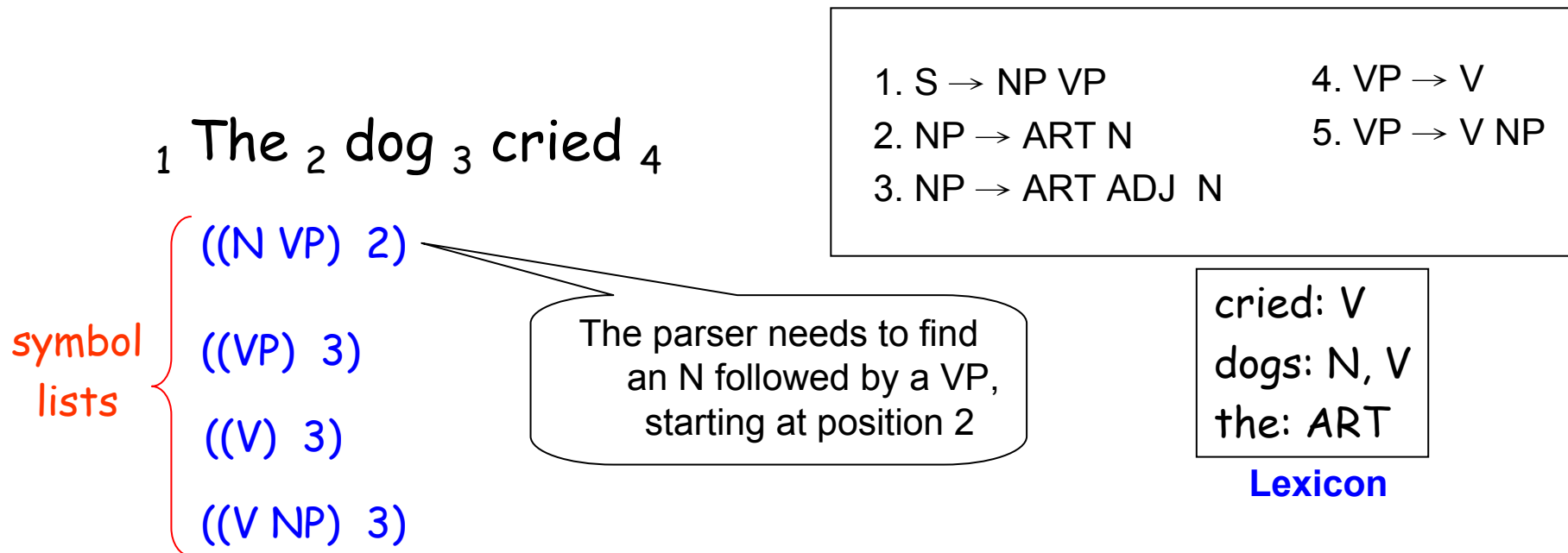
- Different search algorithms, such as depth-first search or breadth-first search algorithms, can be applied



The record of the parsing process, either in top-down or bottom-up manners, can be used to generate the parse tree representation

The Top-Down Parser (2/2)

- Start with the S symbol and rewrite it into a sequence of terminal symbols that matches the classes of the words in the input sentence
 - The state of the parse at any given time can be represented as a list of symbols that are the results of operations applied so far



The Simple Top-Down Parser (1/3)

the possibilities list

| Step | Current State | Backup States | Comment |
|------|----------------|--|--|
| 1. | ((S) 1) | | initial position |
| 2. | ((NP VP) 1) | | rewriting S by rule 1 |
| 3. | ((ART N VP) 1) | | rewriting NP by rules 2 & 3 |
| 4. | ((N VP) 2) | ((ART ADJ N VP) 1) | matching ART with <i>the</i> |
| 5. | ((VP) 3) | ((ART ADJ N VP) 1) | matching N with <i>dogs</i> |
| 6. | ((V) 3) | ((ART ADJ N VP) 1) ((V NP) 3) ((ART ADJ N VP) 1) | rewriting VP by rules 5–8 |
| 7. | (() 4) | | the parse succeeds as V is matched to <i>cried</i> , leaving an empty grammatical symbol list with an empty sentence |

Check the input sentence to see if match ASAP

new states are put onto the front of the possibilities list

Figure 3.5 Top-down depth-first parse of ₁ *The* ₂ *dogs* ₃ *cried* ₄

- | | |
|-------------------------------|--------------------------|
| 1. $S \rightarrow NP VP$ | 4. $VP \rightarrow V$ |
| 2. $NP \rightarrow ART N$ | 5. $VP \rightarrow V NP$ |
| 3. $NP \rightarrow ART ADJ N$ | |

cried: V
dogs: N, V
the: ART

The Simple Top-Down Parser (2/3)

- **Algorithm**

1. Select the current state: take the first state off the possibilities list and call it C
 - If the possibilities list is empty, then the algorithm fails
2. If C consists of an empty symbol list and is at the sentence end position, the algorithm succeeds
3. Otherwise, generate the next possible states
 - If the first symbol on the symbol list is a **lexical symbol** (part-of-speech tag), and the next word in the sentence can be in that class, then create a new state by removing the first symbol from the symbol list and update the word position, and **add it to the possibilities list**
 - Otherwise, if the first symbol on the symbol list of C is a non-terminal, generate a new state for each rule in the grammar that can rewrite that non-terminal symbol and **add them all to the possibilities list**

where to put the new states depends on the search strategies

The Simple Top-Down Parser (3/3)

- One more example

1 The 2 old 3 man 4 cried 5

- | | |
|-------------------------------|--------------------------|
| 1. $S \rightarrow NP VP$ | 4. $VP \rightarrow V$ |
| 2. $NP \rightarrow ART N$ | 5. $VP \rightarrow V NP$ |
| 3. $NP \rightarrow ART ADJ N$ | |

cried: V
old: ADJ, N
man: N, V
the: ART

new states are put onto the front of the possibilities list

| Step | Current State | Backup States | Comment |
|------|--------------------|---------------------------------------|--|
| 1. | ((S) 1) | | |
| 2. | ((NP VP) 1) | | S rewritten to NP VP |
| 3. | ((ART N VP) 1) | ((ART ADJ N VP) 1) | NP rewritten producing two new states |
| 4. | ((N VP) 2) | ((ART ADJ N VP) 1) | |
| 5. | ((VP) 3) | ((ART ADJ N VP) 1) | the backup state remains |
| 6. | ((V) 3) | ((V NP) 3) ((ART ADJ N VP) 1) | |
| 7. | (() 4) | ((V NP) 3) ((ART ADJ N VP) 1) | |
| 8. | ((V NP) 3) | ((ART ADJ N VP) 1) | the first backup is chosen |
| 9. | ((NP) 4) | ((ART ADJ N VP) 1) | |
| 10. | ((ART N) 4) | ((ART ADJ N) 4) ((ART ADJ N VP) 1) | looking for ART at 4 fails |
| 11. | ((ART ADJ N) 4) | ((ART ADJ N VP) 1) | fails again |
| 12. | ((ART ADJ N VP) 1) | | now exploring backup state saved in step 3 |
| 13. | ((ADJ N VP) 2) | | |
| 14. | ((N VP) 3) | | |
| 15. | ((VP) 4) | | |
| 16. | ((V) 4) | ((V NP) 4) | |
| 17. | (() 5) | | success! |

Figure 3.6 A top-down parse of 1 The 2 old 3 man 4 cried 5

Search strategies for the Top-Down Parser (1/3)

- Depth-first search: DFS (LIFO: last-in first-out)
 - The possibilities list is a stack
 - Step 1 always take the first element off the list
 - Step 3 always puts (adds) the new states **on the front of the list**
- Breadth-first search: BFS (FIFO: first-in first-out)
 - The possibilities list is a queue
 - Step 1 always take the first element off the list
 - Step 3 always puts (adds) the new states **on the end of the list**

Search strategies for the Top-Down Parser (2/3)

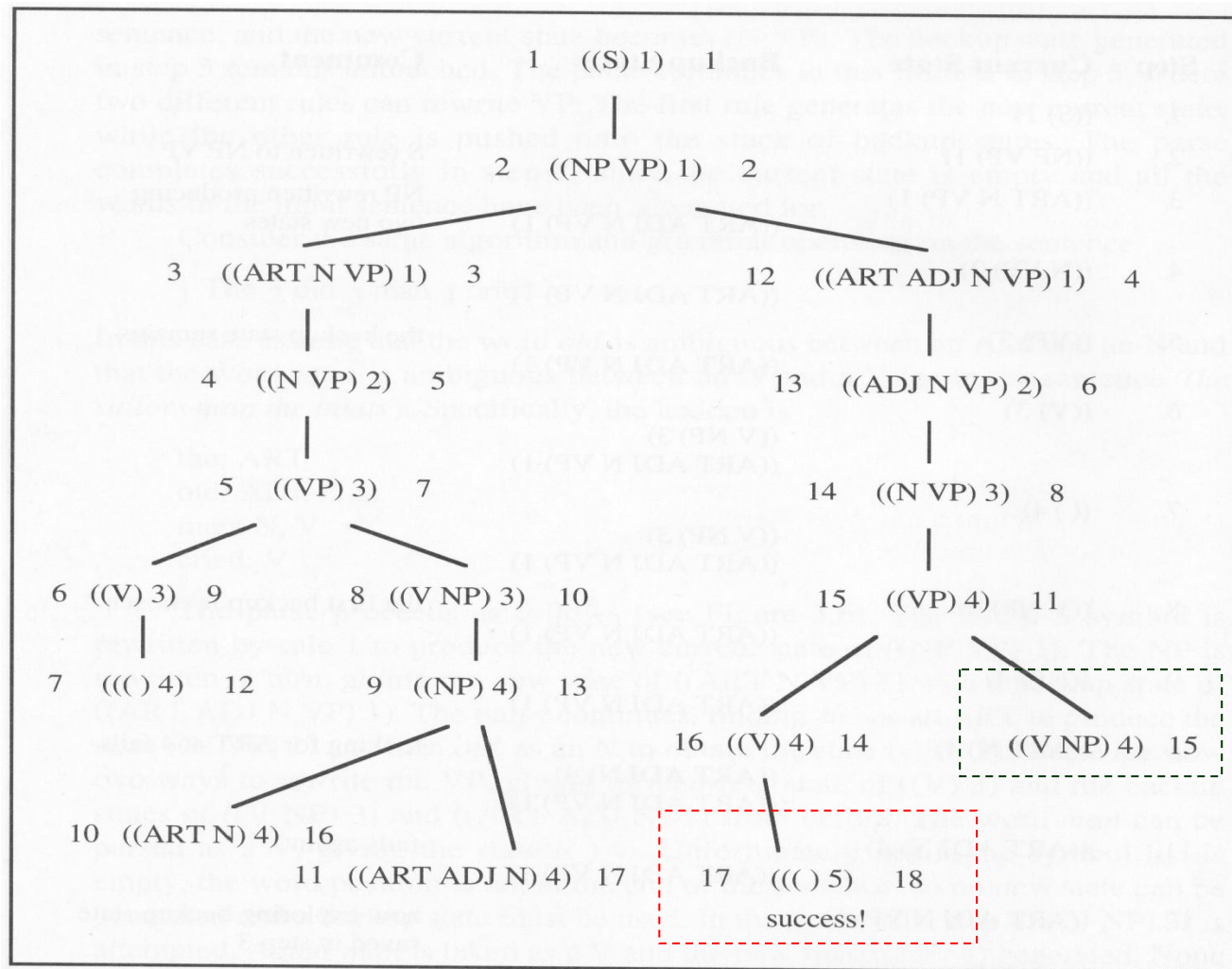


Figure 3.7 Search tree for two parse strategies (depth-first strategy on left; breadth-first on right)

Not examined
by DFS

Search strategies for the Top-Down Parser (3/3)

- Comparison of DFS and BFS
 - DFS
 - **One interpretation** is considered and expanded until fails; only then is the second one considered
 - Often moves quickly to the a solution but in other cases may spend considerable time pursuing futile paths
 - BFS
 - **All interpretations** are considered alternatively, each being expanded one step at a time
 - Explore each possible solution to a certain depth before moving on

Many parsers built today use the DFS strategy because it tends to minimize the no. of backup states needed and thus uses less memory and requires less bookkeeping

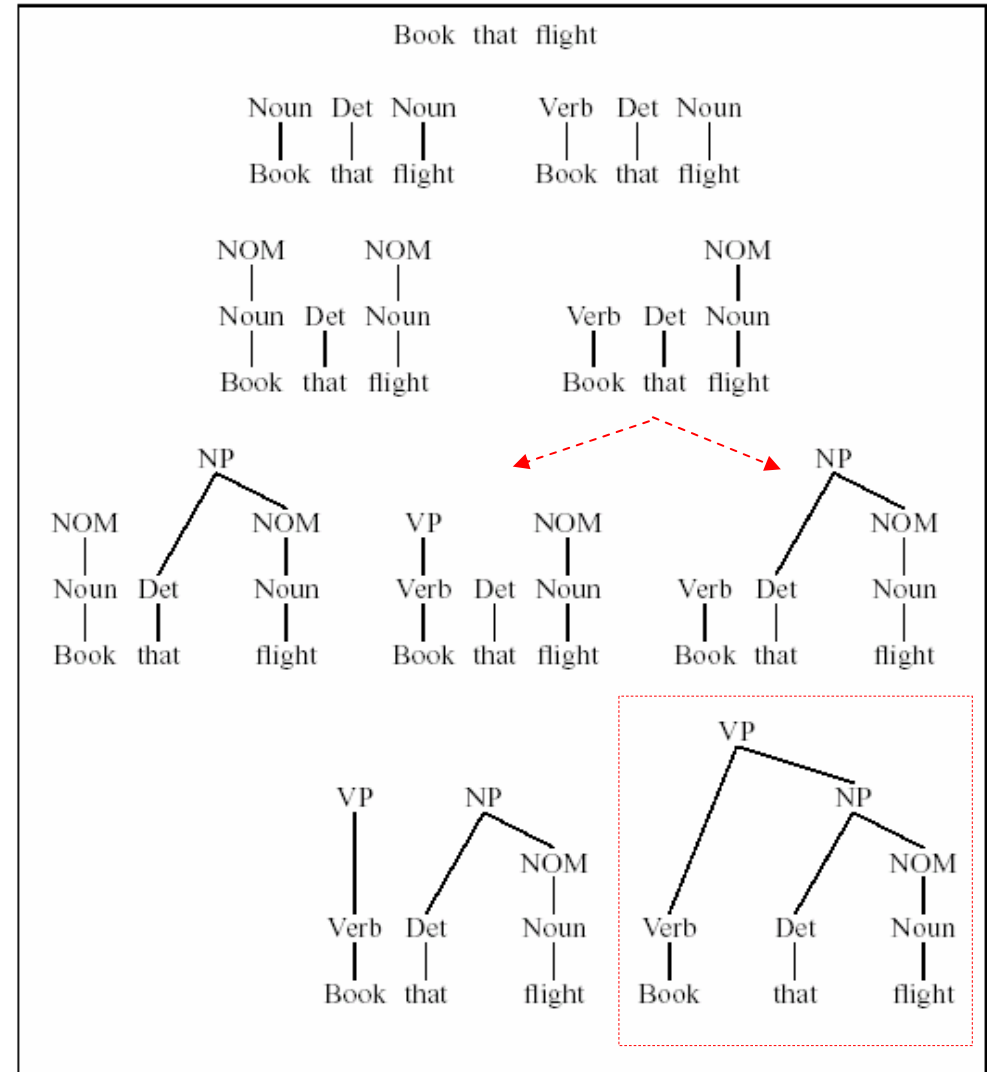
The Bottom-Up Parser (1/2)

- Start with the words of the input, and try to build tree from the words up, by applying rules from the grammar one at a time
 - The right hand side of some rules might fit
 - Successful if the parser succeeds in building a tree rooted in the start symbol (or a symbol list with S and positioned at the end of the input sentence) that covers all the input

The Bottom-Up Parser (2/2)

| | |
|------------------------------------|--|
| $S \rightarrow NP VP$ | $Det \rightarrow that \mid this \mid a$ |
| $S \rightarrow Aux NP VP$ | $Noun \rightarrow book \mid flight \mid meal \mid money$ |
| $S \rightarrow VP$ | $Verb \rightarrow book \mid include \mid prefer$ |
| $NP \rightarrow Det Nominal$ | $Aux \rightarrow does$ |
| $Nominal \rightarrow Noun$ | |
| $Nominal \rightarrow Noun Nominal$ | $Prep \rightarrow from \mid to \mid on$ |
| $NP \rightarrow Proper-Noun$ | $Proper-Noun \rightarrow Houston \mid TWA$ |
| $VP \rightarrow Verb$ | |
| $VP \rightarrow Verb NP$ | $Nominal \rightarrow Nominal PP$ |

1 Book 2 that 3 flight 4



Comparing Top-Down and Bottom-UP Parsing

- Top-Down
 - Never wastes time exploring trees that can't result in an S
 - But spends considerable effort on S trees that are not consistent with the input
- Bottom-Up
 - Never suggest trees that are not least locally grounded in the actual input
 - Trees that have no hope of leading to an S , or fitting in with any of their neighbors, are generated with wild abandon
 - Only check the input once

Problems with Parsing (1/4)

- **Left-recursion**

- A non-terminal category that has a derivation that includes itself anywhere along its leftmost branch

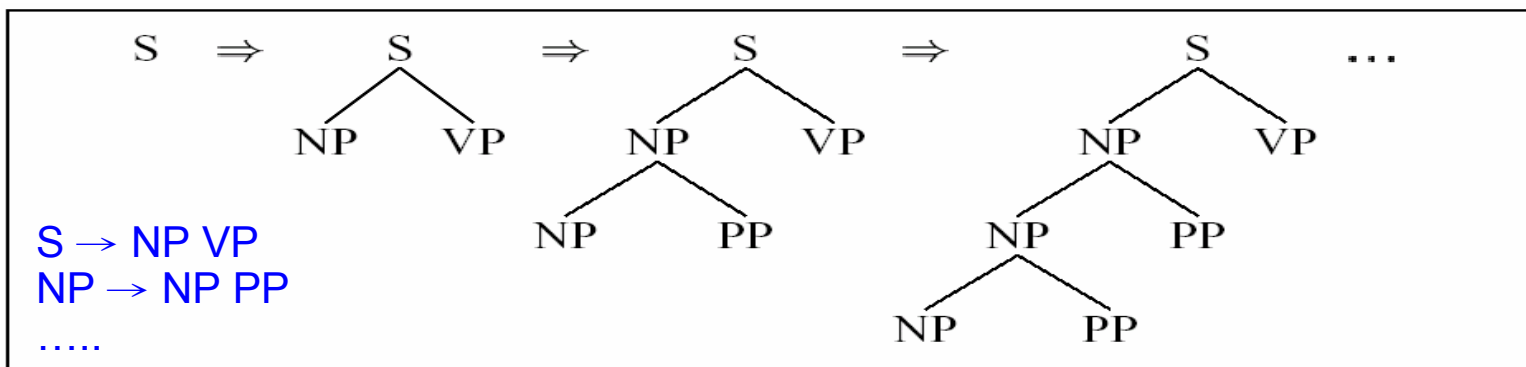
NP → Det Nominal

Det → NP 's

- Especially, the **immediately** left-recursive rule

NP → NP 's N

- E.g. causing an infinite loop in top-down parsing with DFS search strategy



Problems with Parsing (2/4)

- **Ambiguity**

- Structural ambiguity: arises in the syntactic structures used in parsing
 - The grammar assigns more than one possible parse to a sentence
 - Attachment ambiguity
 - » Most frequently seen for adverbial phrases
- Coordination ambiguity

I shot an elephant in my pajamas.

old men and women

Parsers which do not incorporate disambiguators must simply return all the possible parse trees for a given input.

Problems with Parsing (3/4)

- **Ambiguity**

- Basic ways to alleviate the ambiguity problem

- Dynamic programming

- Used to exploit the regularities in the search space so that **the common subpart are derived only once**

- Reduce some of the costs associated with ambiguity

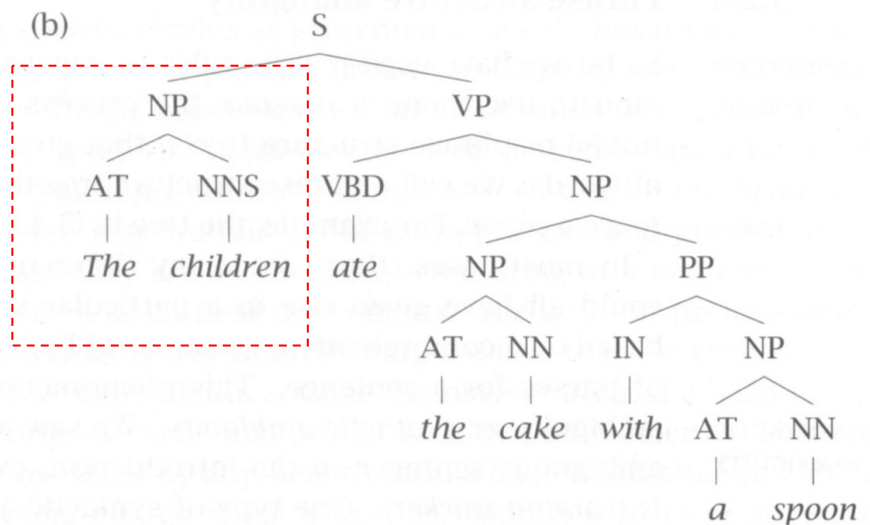
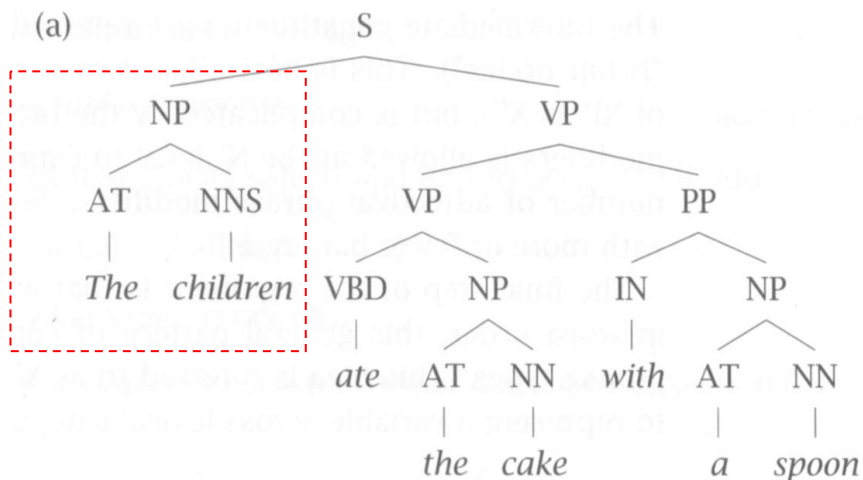
- Implicitly store all possible parses by storing all the constituents with links that enable the parses to be reconstructed

- Heuristic search

- Augment the parser's search strategy with heuristics that guide it towards likely parses first

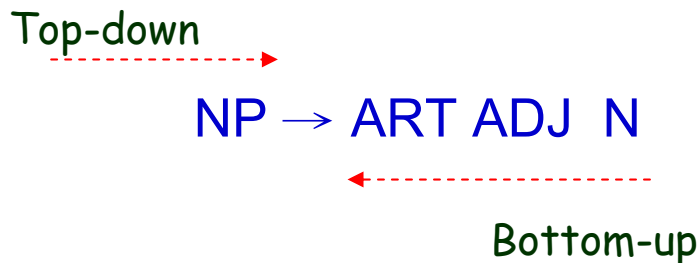
Problems with Parsing (4/4)

- Repeated Parsing of Subtrees
 - The parser often builds valid subtrees for portions of the input, then discards them during the backtracking
 - It has to rebuild these subtrees again
 - Some constituents are constructed more than once



The Bottom-Up Chart Parser (1/29)

- A data structure called chart is introduced
 - Allow the parser to store the partial results of matching as it done so far
 - Such that the work would not be reduplicated
- The basic operation of a chart parser involves combining an active arc with a complete constituents (keys)
 - Three kinds of data structures
 - The agenda (to store new complete constituents)
 - The active arcs (i.e. partial parse trees)
 - The chart



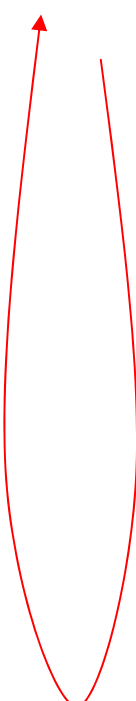
- A subtree corresponding to a single grammar rule
- Information about the progress made in completing the subtree
- Position of the subtree has to do with the input

Space Complexity vs. Time Complexity

The Bottom-Up Chart Parser (2/29)

- The Bottom-Up Chart Parsing Algorithm

1. If the agenda is empty, look up the **interpretations** for the next word in the input and add them to the agenda
2. Select a constituent from the agenda (Call it constituent C from position p_1 to p_2)
3. For each rule in the grammar of form $X \rightarrow CX_1 \dots X_n$, add an **active arc** of form $X \rightarrow \circ CX_1 \dots X_n$ from position p_1 to p_1
4. Add C to the chart using the following **arc extension algorithm**
 - 4.1 Insert C **into the chart** from position p_1 to p_2
 - 4.2 For any active arc of the form $X \rightarrow X_1 \dots \circ C \dots X_n$ from position p_0 to p_1 add a new active arc $X \rightarrow X_1 \dots C \circ \dots X_n$ from p_0 to p_2
 - 4.3 For any active arc of the form $X \rightarrow X_1 \dots X_n \circ C$ from position p_0 to p_1 , then add a new constituent of type X from p_0 to p_2 **to the agenda**



Loop until
no input left

The Bottom-Up Chart Parser (3/29)

- Example

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
large: ADJ
can: N, AUX
holds: N, V
Water: N

Initialization

Chart:



Input: 1 The 2 large 3 can 4 holds 5 the 6 water 7

Note that depth-first strategy is used here
=> The agenda is implemented as a stack.

The Bottom-Up Chart Parser (4/29)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
large: ADJ
can: N, AUX
hold: N, V
Water: N

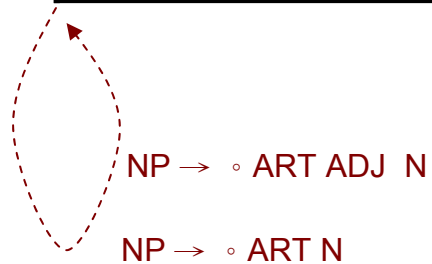
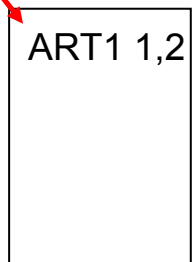
Loop 1

Enter ART1: (*the* from 1 to 2) Look at next word

Chart:



Agenda:



The Bottom-Up Chart Parser (5/29)

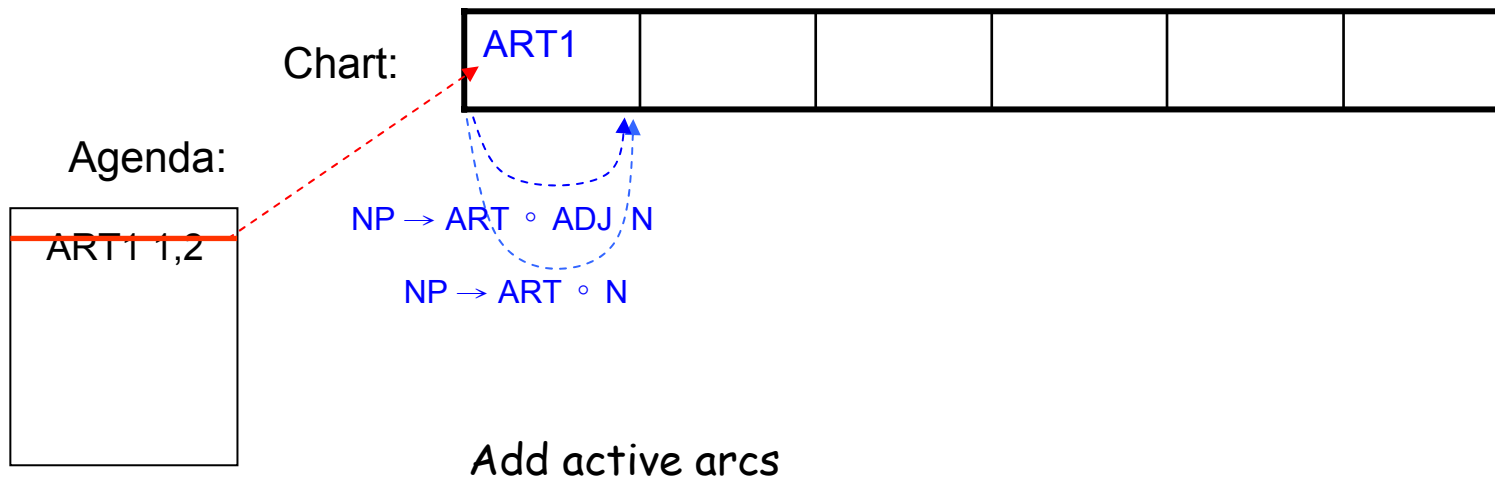
- Example ₁ The ₂ large ₃ can ₄ holds ₅ the ₆ water ₇

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 1

Enter ART1: (*the* from 1 to 2) (using the arc extension algorithm)



The Bottom-Up Chart Parser (6/29)

- Example ₁ The ₂ large ₃ can ₄ holds ₅ the ₆ water ₇

1. $S \rightarrow NP VP$

2. $NP \rightarrow ART ADJ N$

3. $NP \rightarrow ART N$

4. $NP \rightarrow ADJ N$

5. $VP \rightarrow AUX VP$

6. $VP \rightarrow V NP$

the: ART

large: ADJ

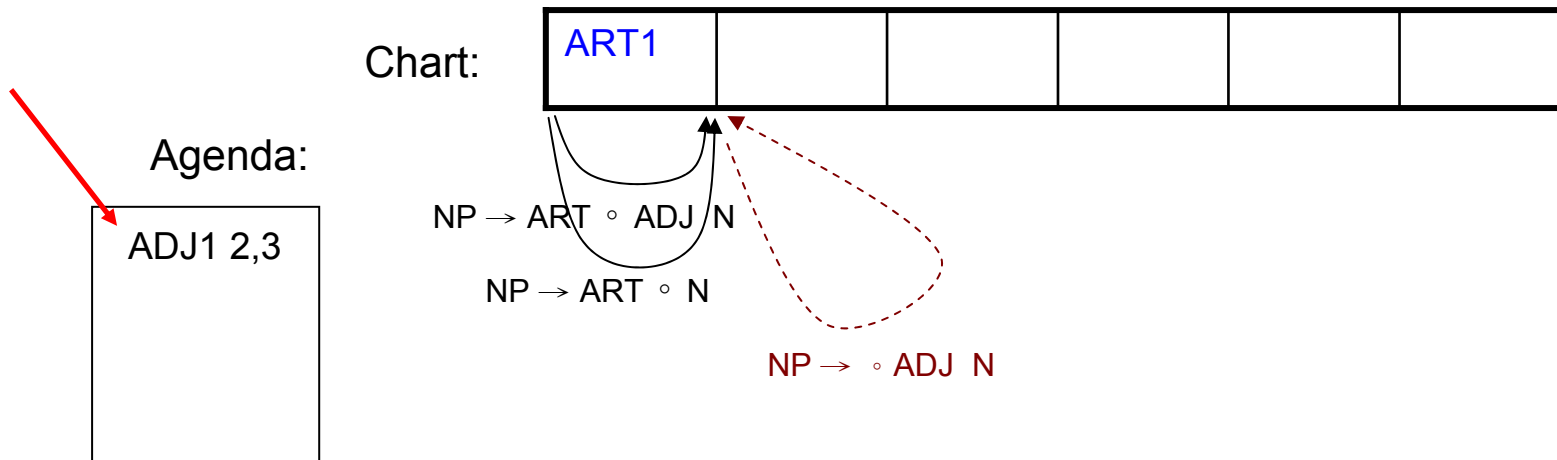
can: N, AUX

hold: N, V

Water: N

Loop 2 Look at next word

Enter ADJ1: ("large" from 2 to 3)



The Bottom-Up Chart Parser (7/29)

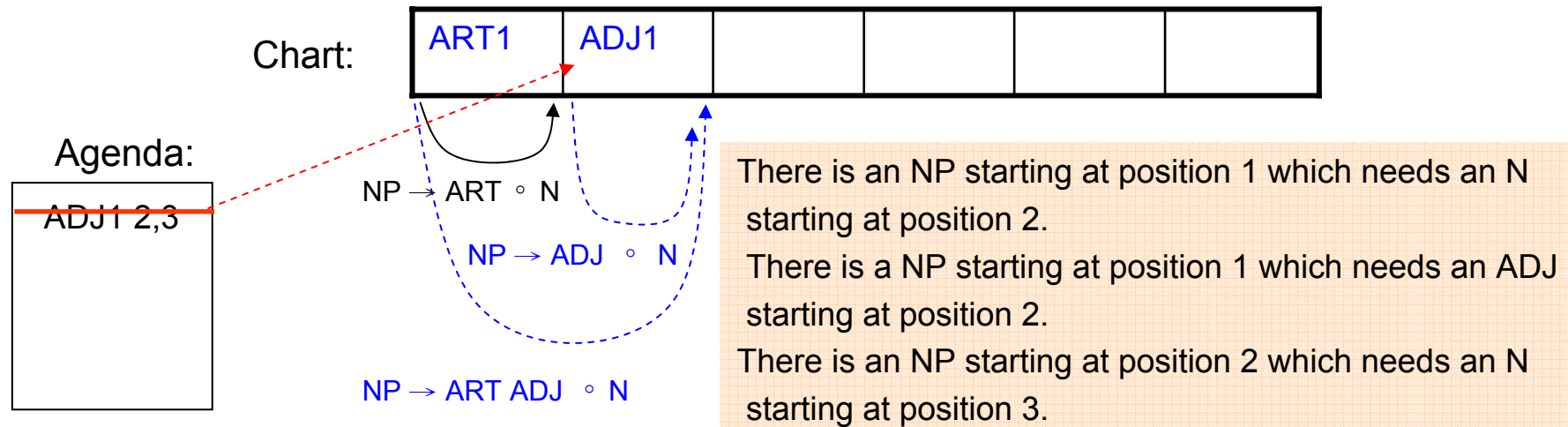
- Example ₁ The ₂ large ₃ can ₄ holds ₅ the ₆ water ₇

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 2 (using the arc extension algorithm)

Enter ADJ1: (“large” from 2 to 3)



The Bottom-Up Chart Parser (8/29)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

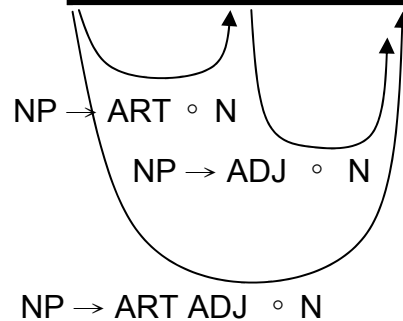
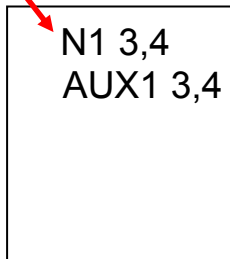
the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 3 Look at next word

Enter N1: ("can" from 3 to 4)



Agenda:



No active arcs are added here!

The Bottom-Up Chart Parser (9/29)

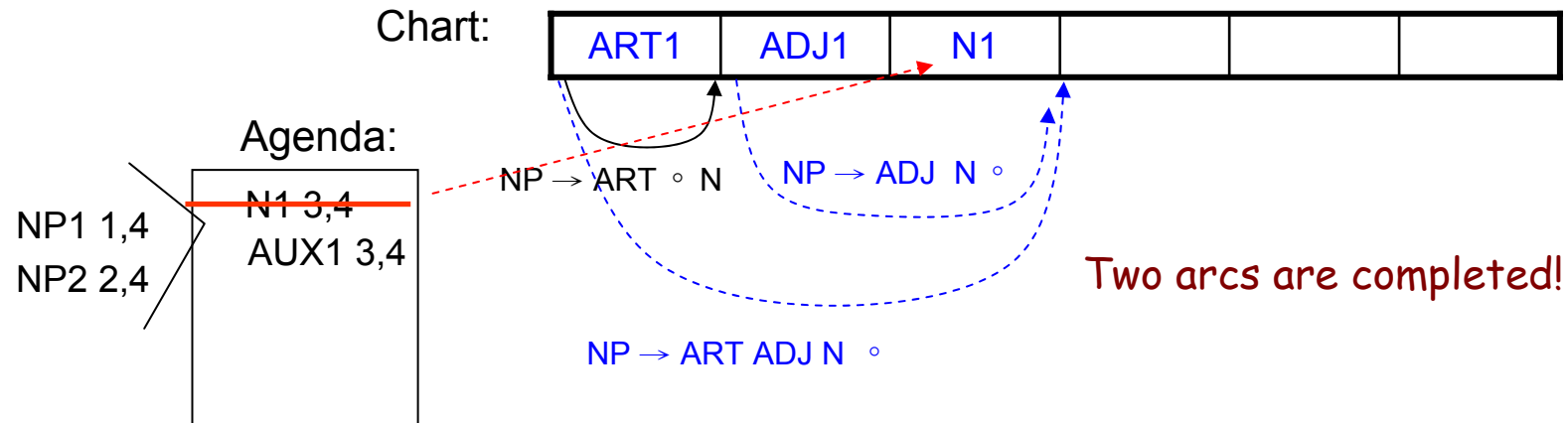
- Example 1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 3

Enter N1: (“can” from 3 to 4) (using the arc extension algorithm)



The Bottom-Up Chart Parser (10/29)

- Example

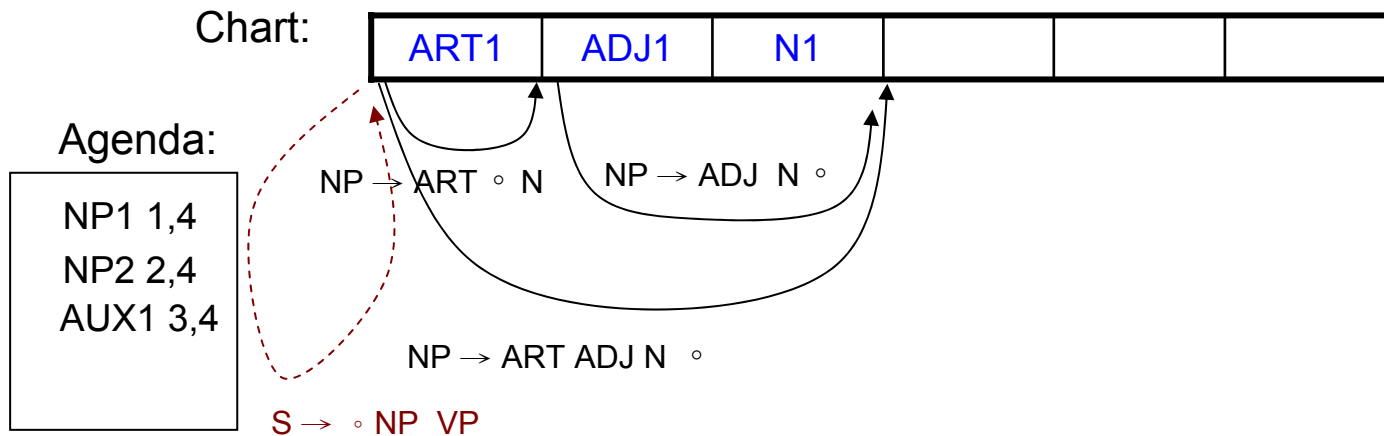
1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 4

Enter NP1: (“the large can” from 1 to 4)



The Bottom-Up Chart Parser (11/29)

- Example

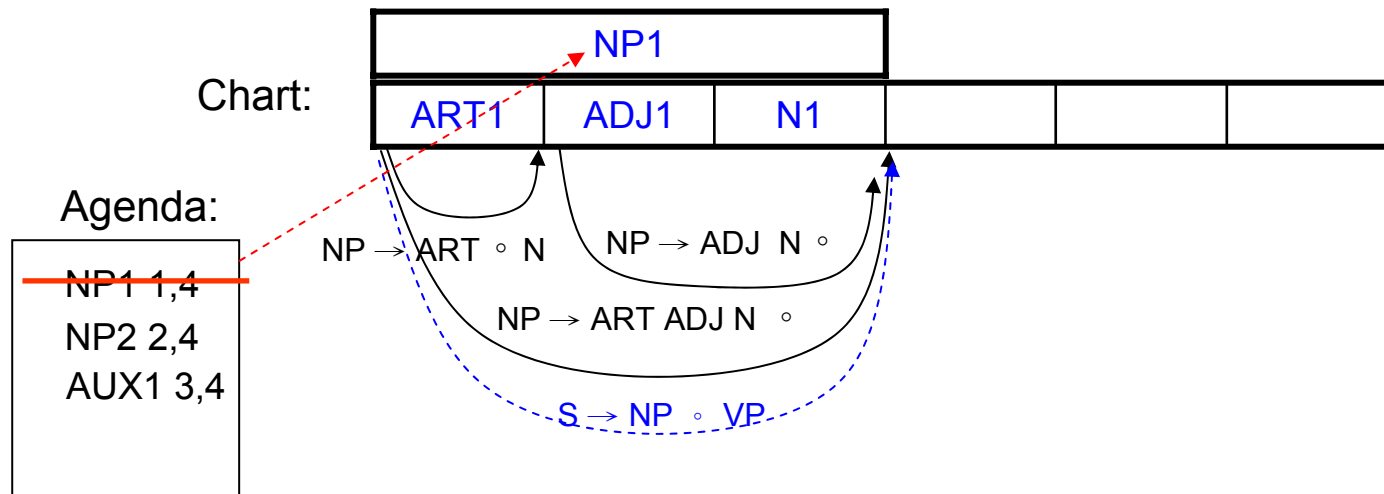
₁ The ₂ large ₃ can ₄ holds ₅ the ₆ water ₇

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 4

Enter NP1: (“the large can” from 1 to 4) (using the arc extension algorithm)



The Bottom-Up Chart Parser (12/29)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

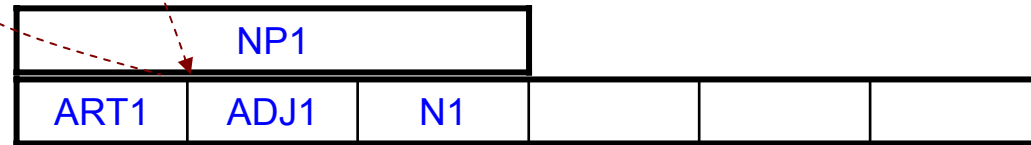
the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 5

Enter NP2: ("*large can*" from 2 to 4)

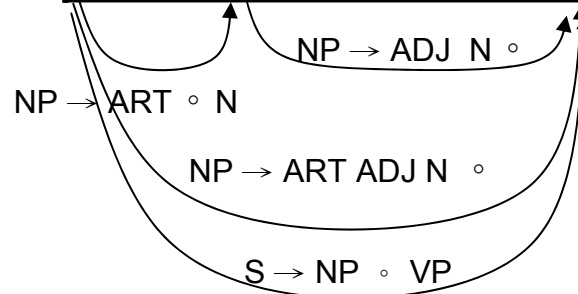
$S \rightarrow \circ NP VP$

Chart:



Agenda:

NP2 2,4
 AUX1 3,4



The Bottom-Up Chart Parser (13/29)

- Example

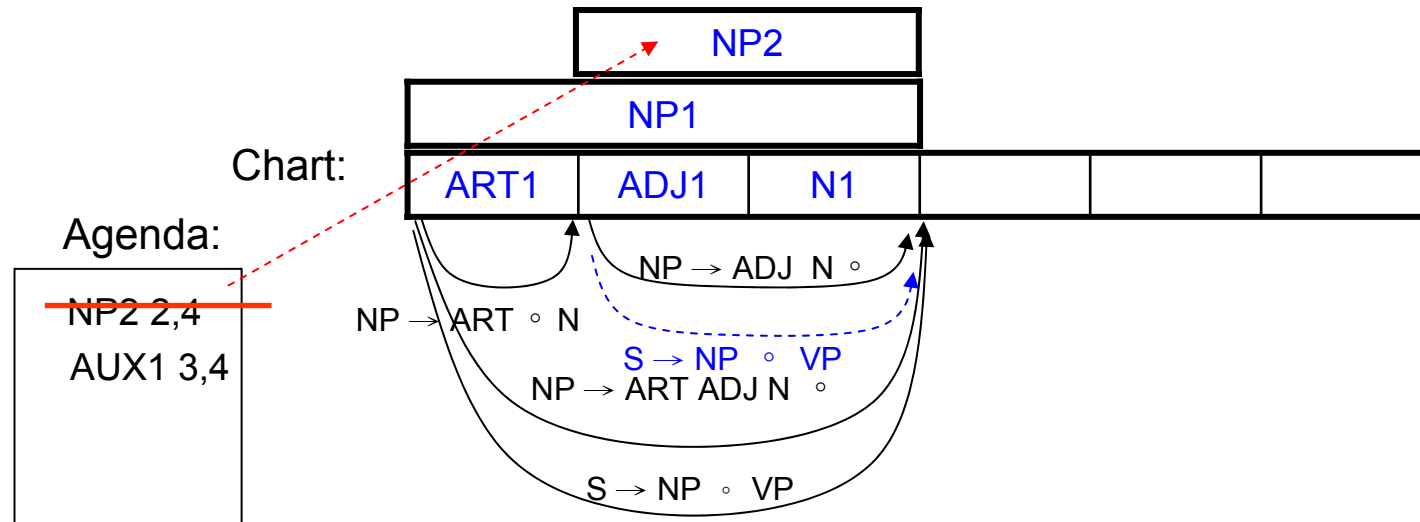
1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 5 (using the arc extension algorithm)

Enter NP2: ("large can" from 2 to 4)



The Bottom-Up Chart Parser (14/29)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

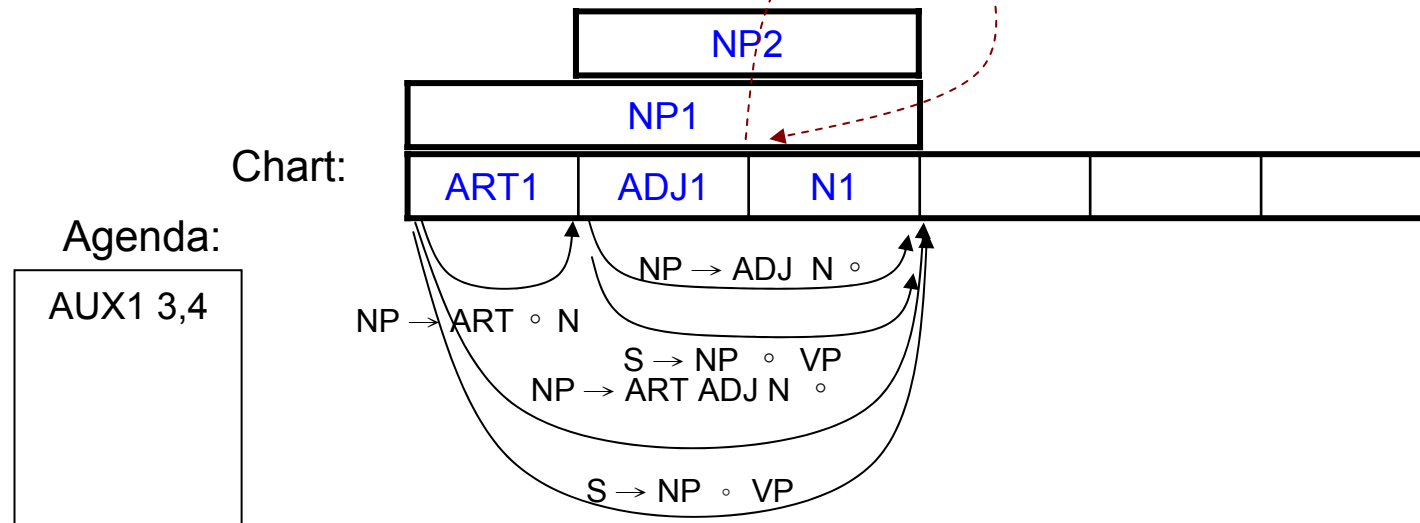
| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 6

Enter AUX1: ("can" from 3 to 4)

$VP \rightarrow \circ AUX VP$



The Bottom-Up Chart Parser (15/29)

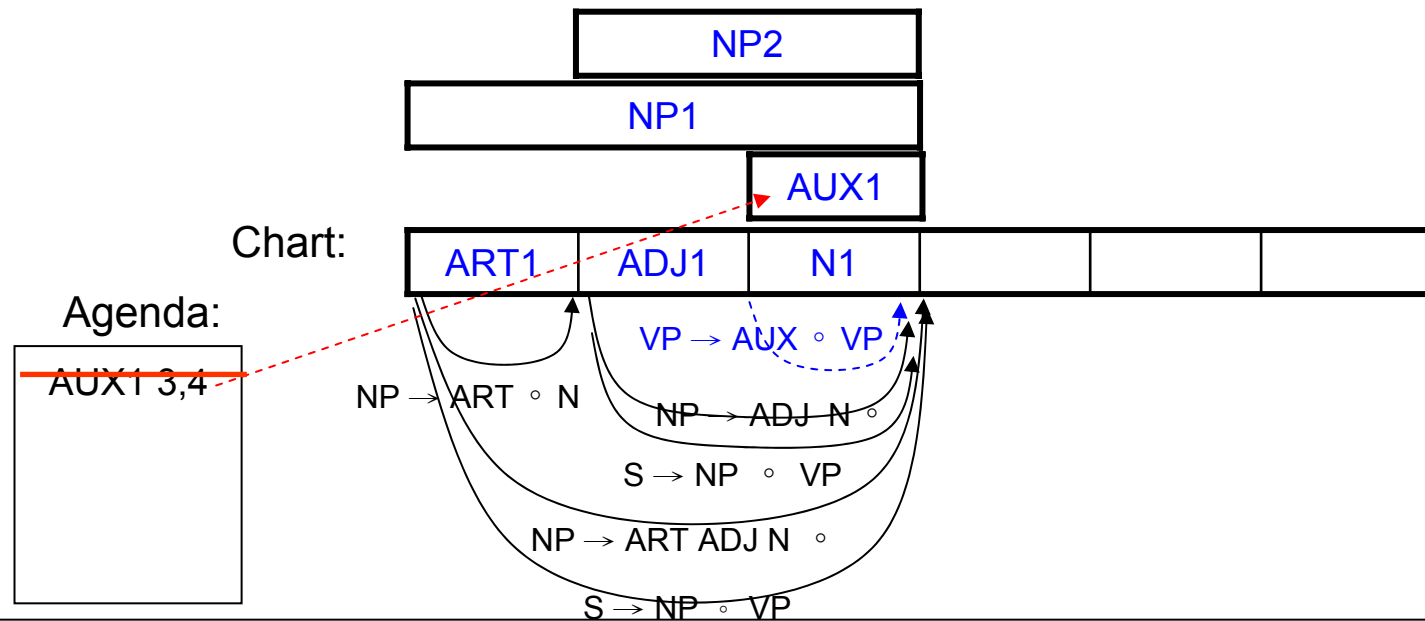
- Example 1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 6 (using the arc extension algorithm)

Enter AUX1: (“can” from 3 to 4)



The Bottom-Up Chart Parser (16/29)

- Example

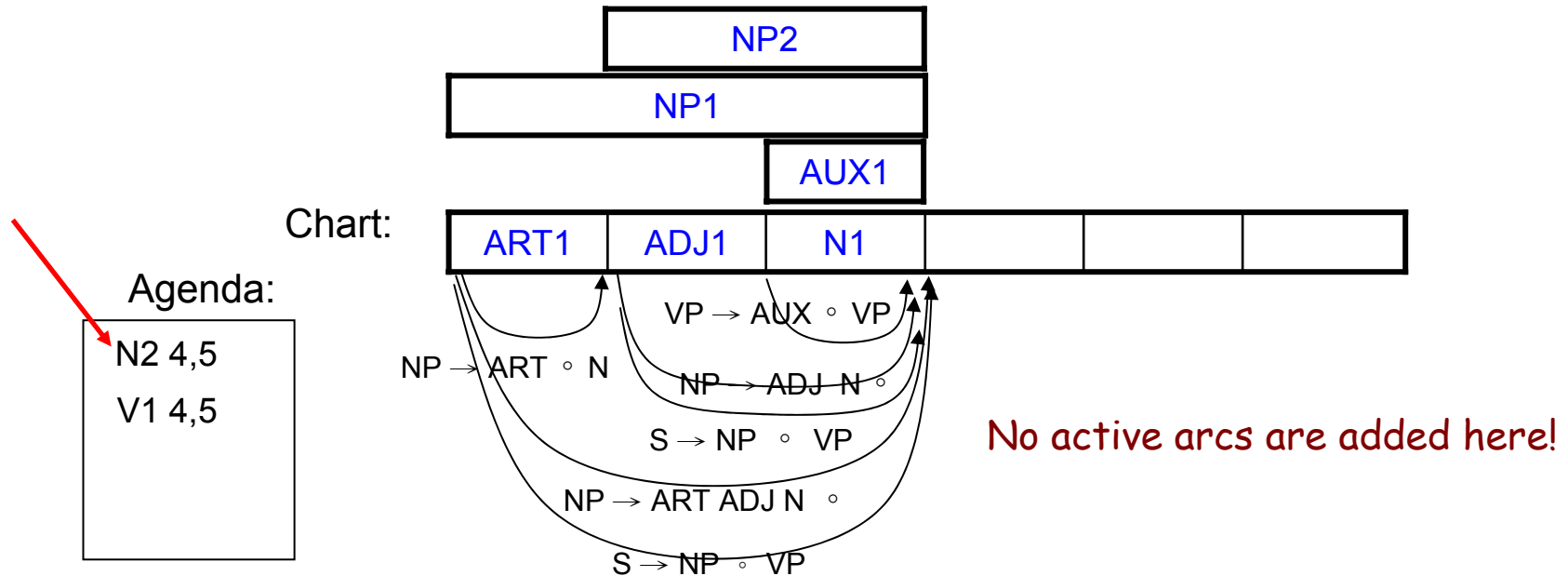
1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------|----------------|
| 1. S → NP VP | 4. NP → ADJ N |
| 2. NP → ART ADJ N | 5. VP → AUX VP |
| 3. NP → ART N | 6. VP → V NP |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 7 Look at next word

Enter N2: ("hold" from 4 to 5)



The Bottom-Up Chart Parser (17/29)

- Example

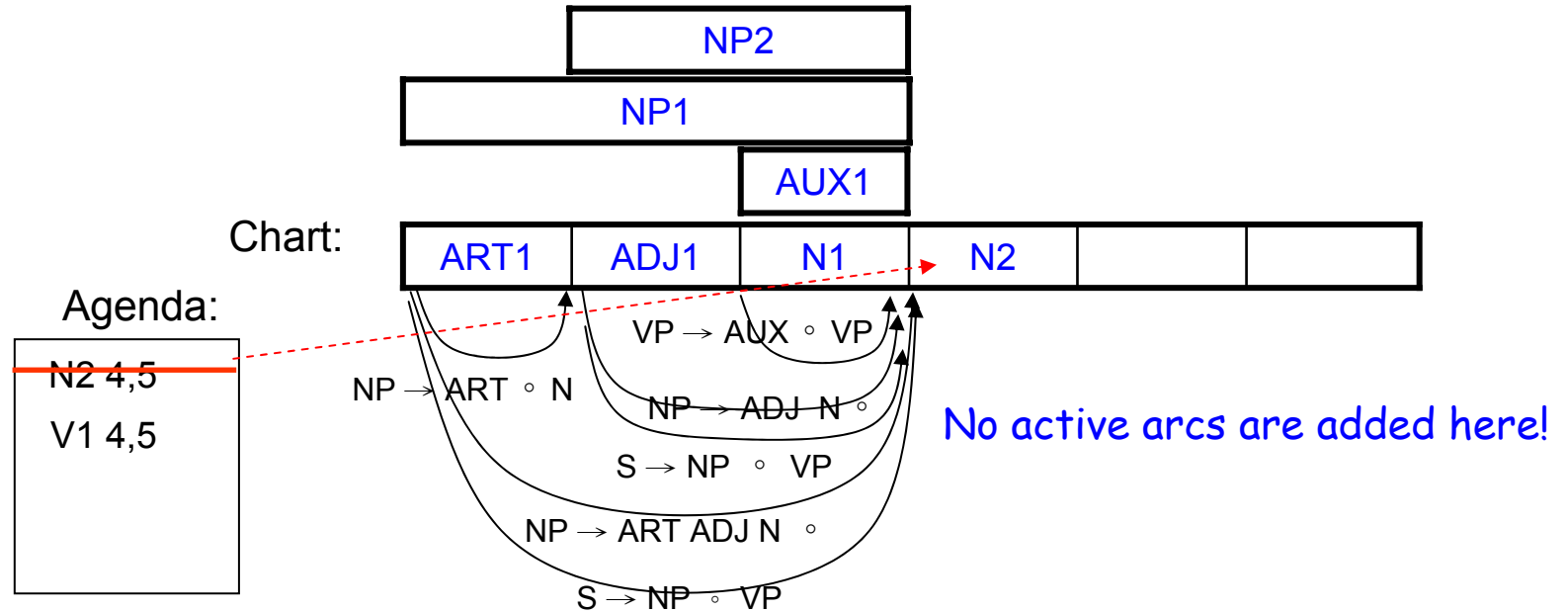
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 7 (using the arc extension algorithm)

Enter N2: ("hold" from 4 to 5)



The Bottom-Up Chart Parser (18/29)

- Example

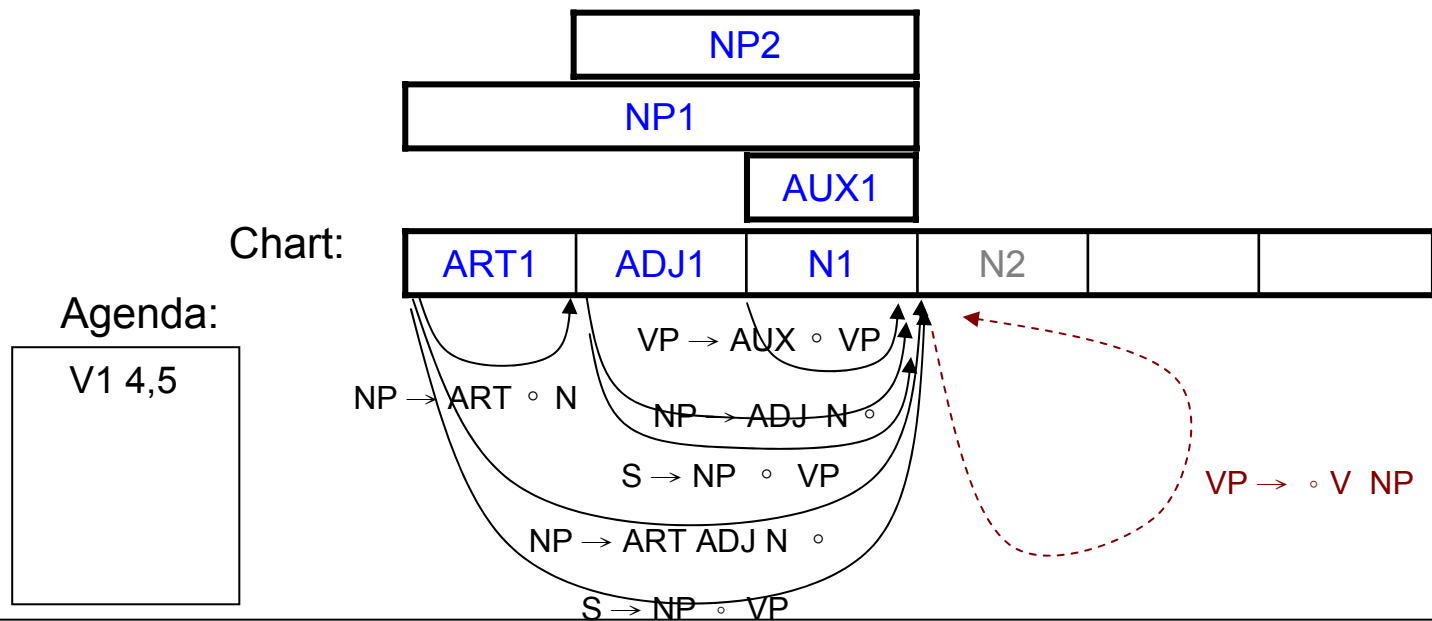
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 8

Enter V1: ("hold" from 4 to 5)



The Bottom-Up Chart Parser (19/29)

- Example

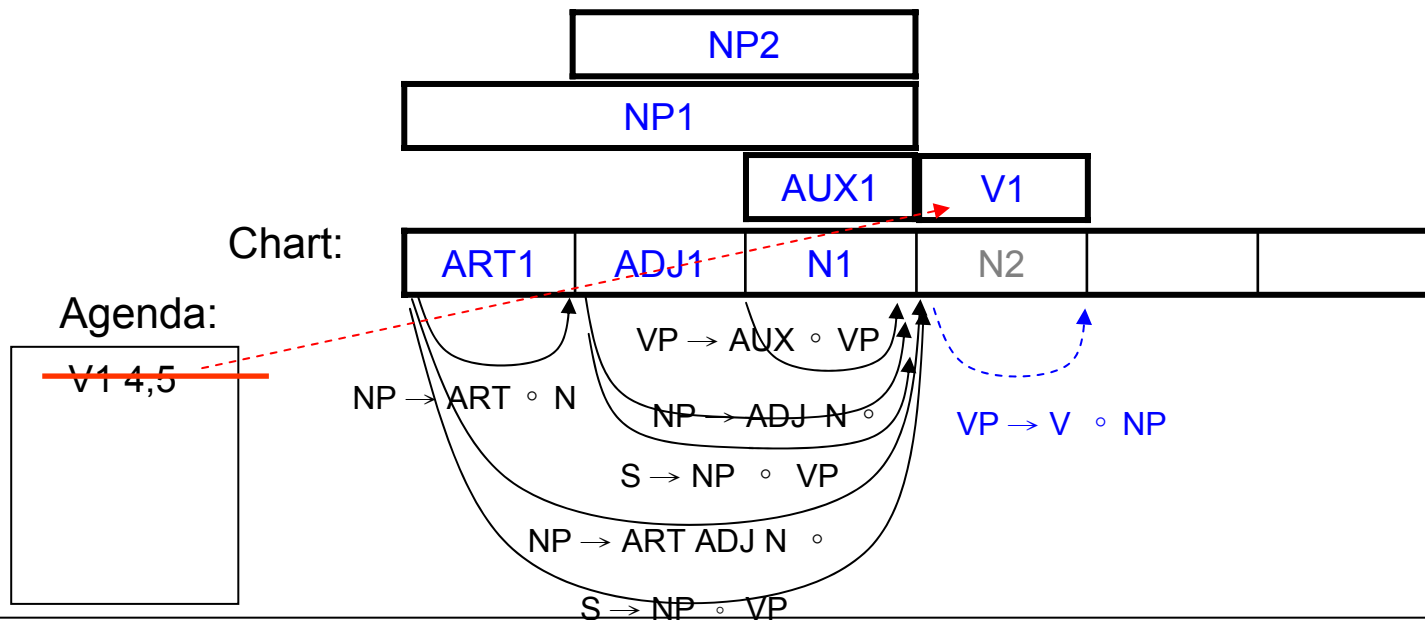
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 8 (using the arc extension algorithm)

Enter V1: (“hold” from 4 to 5)



The Bottom-Up Chart Parser (20/29)

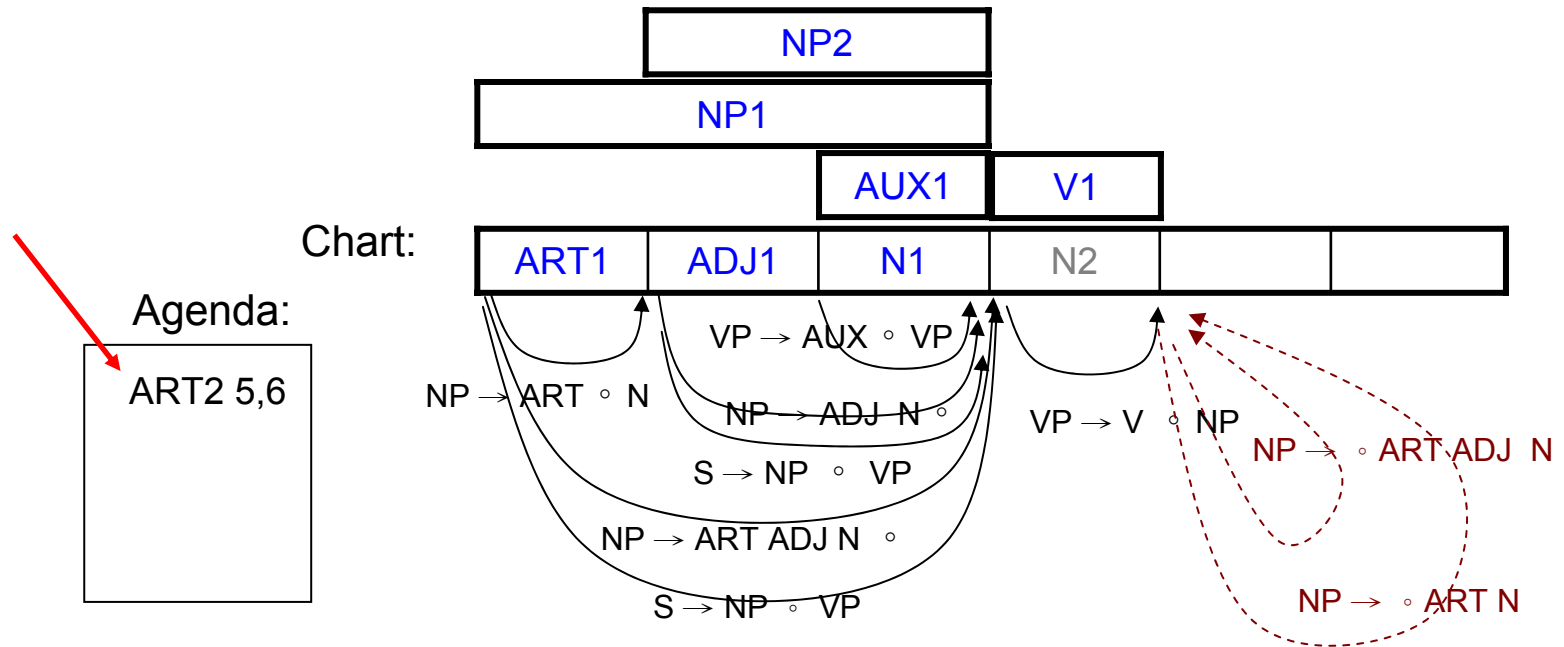
- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 9 Look at next word
 Enter **ART2**: ("*the*" from 5 to 6)



The Bottom-Up Chart Parser (21/29)

- Example

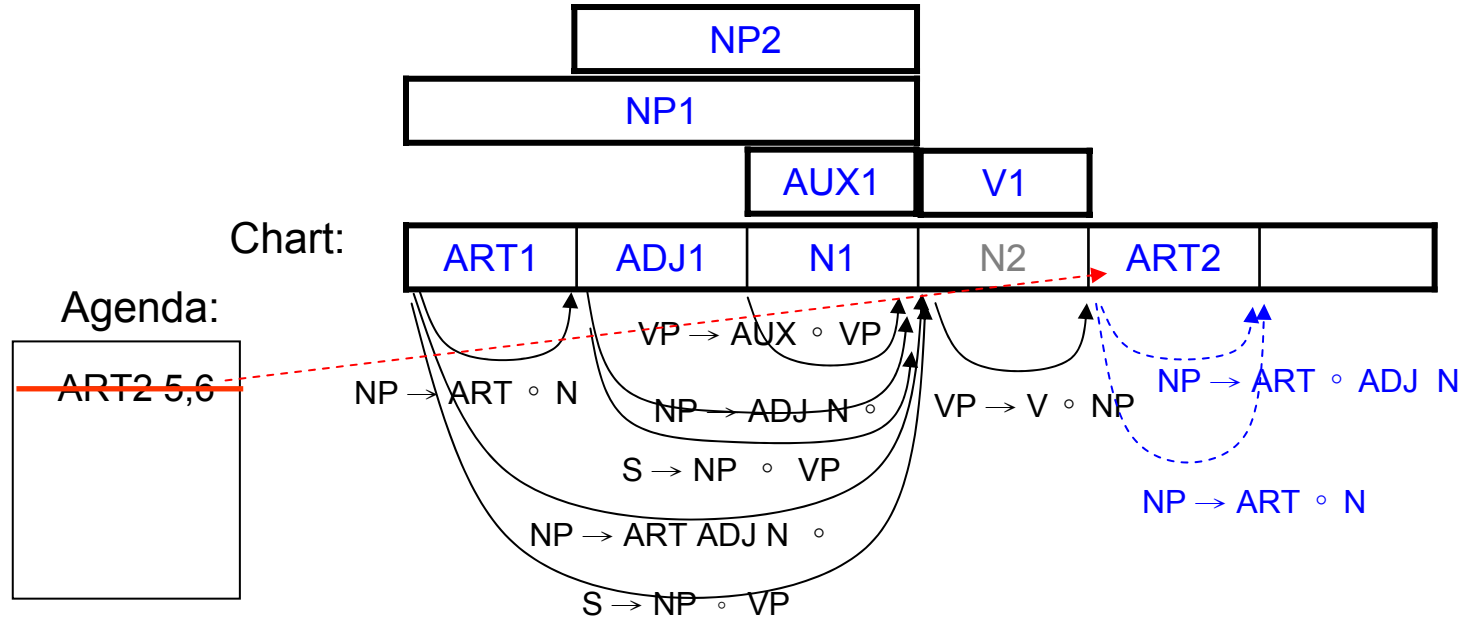
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 9 (using the arc extension algorithm)

Enter ART2: ("the" from 5 to 6)



The Bottom-Up Chart Parser (22/29)

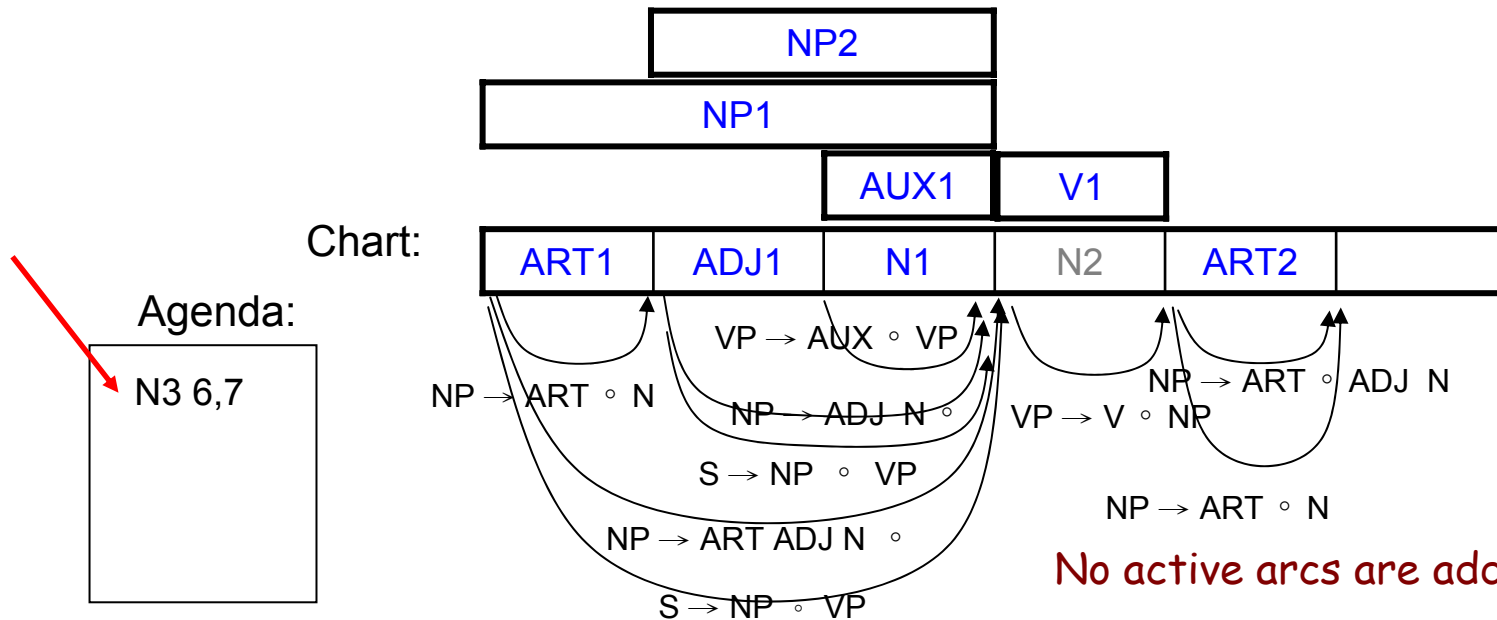
- Example $_1$ The $_2$ large $_3$ can $_4$ holds $_5$ the $_6$ water $_7$

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 10 Look at next word

Enter N3: ("water" from 6 to 7)



The Bottom-Up Chart Parser (23/29)

- Example

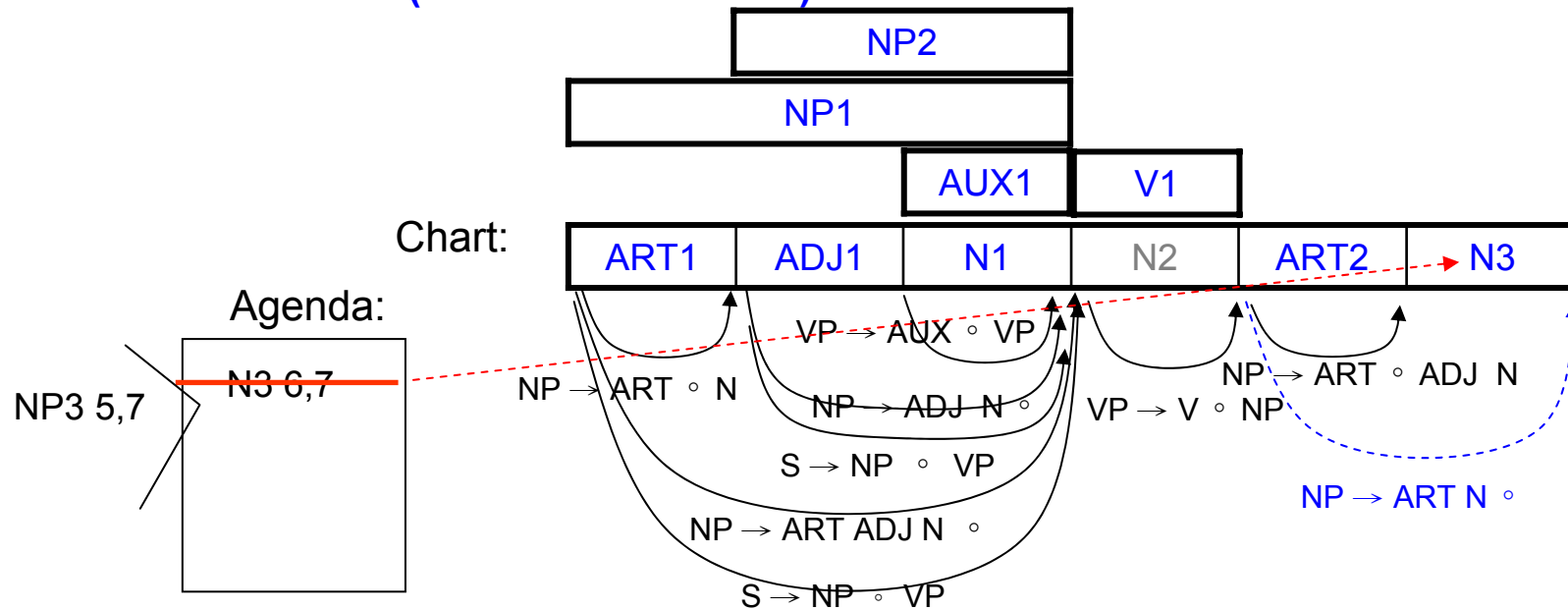
1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 10 (using the arc extension algorithm)

Enter N3: ("water" from 6 to 7)



The Bottom-Up Chart Parser (24/29)

- Example

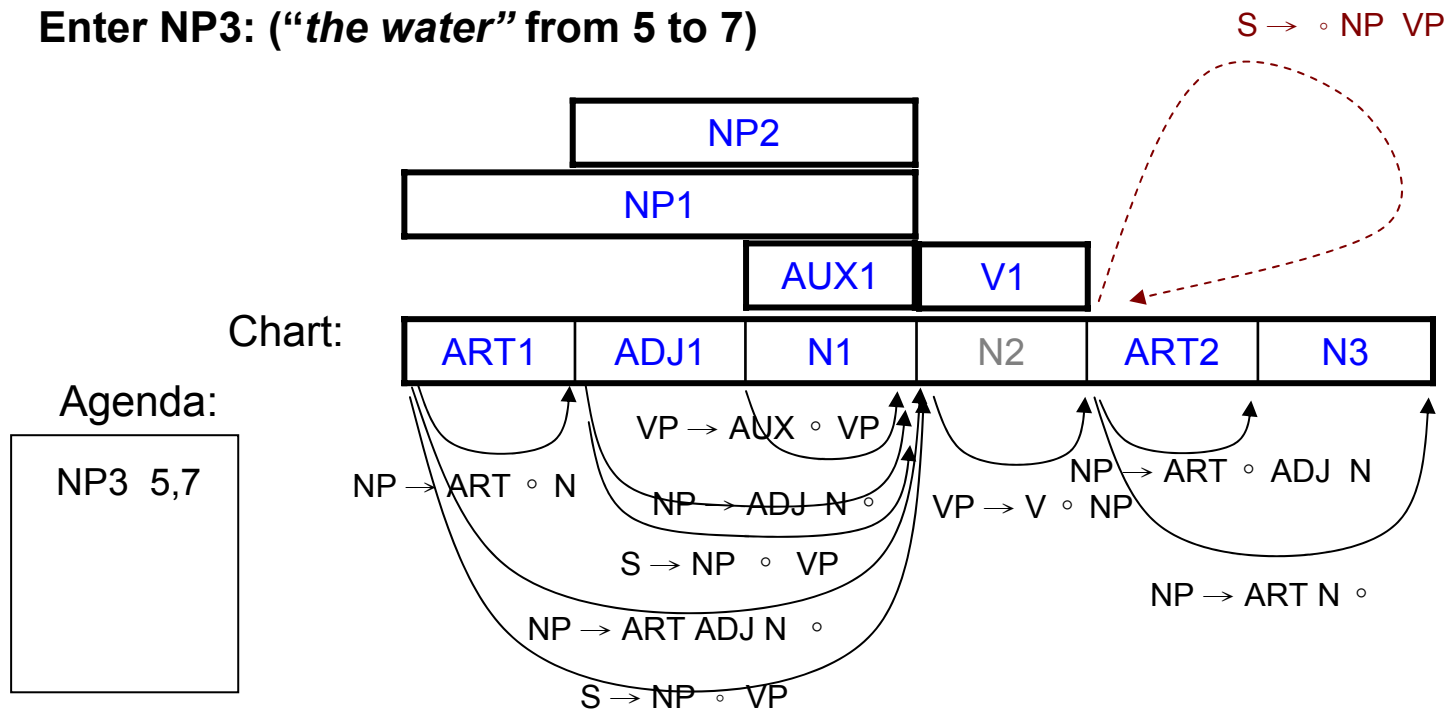
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 11

Enter NP3: ("the water" from 5 to 7)



The Bottom-Up Chart Parser (25/29)

- Example

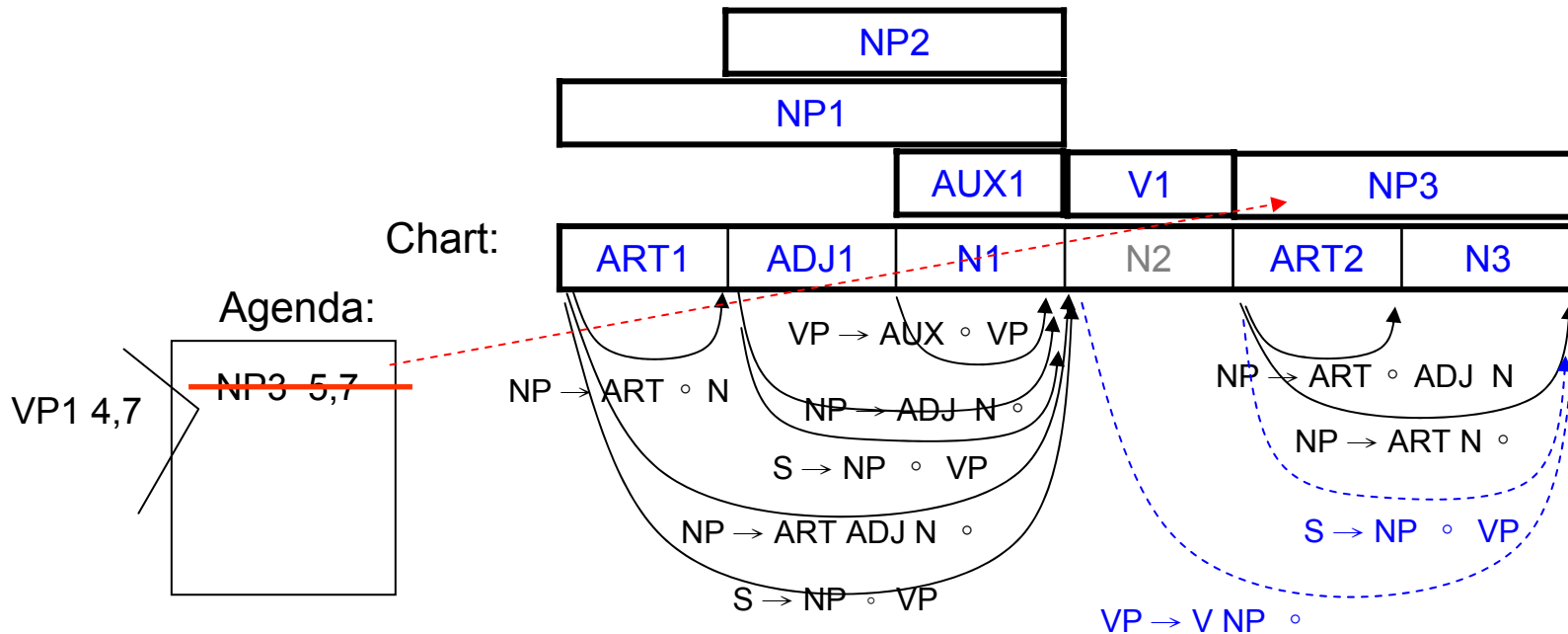
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 11 (using the arc extension algorithm)

Enter NP3: ("the water" from 5 to 7)



The Bottom-Up Chart Parser (26/29)

- Example

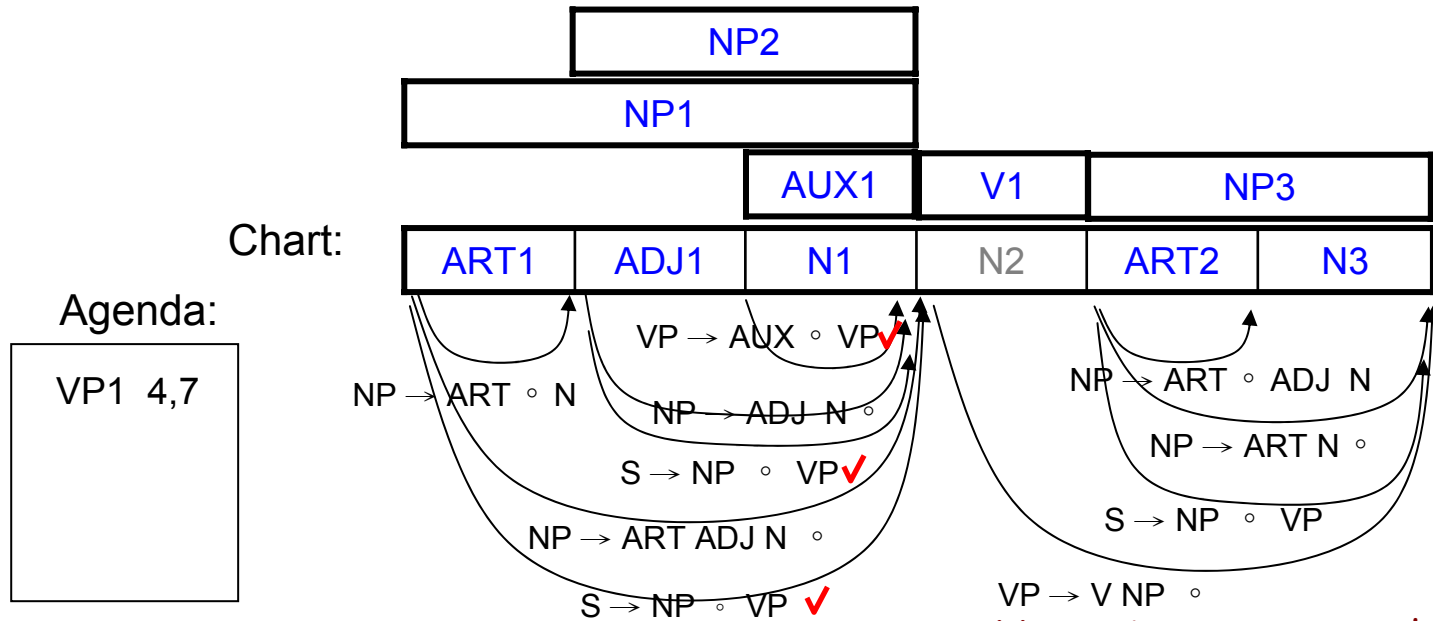
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 12

Enter VP1: (“hold the water” from 4 to 7)



No active arcs are added here!

The Bottom-Up Chart Parser (27/29)

- Example

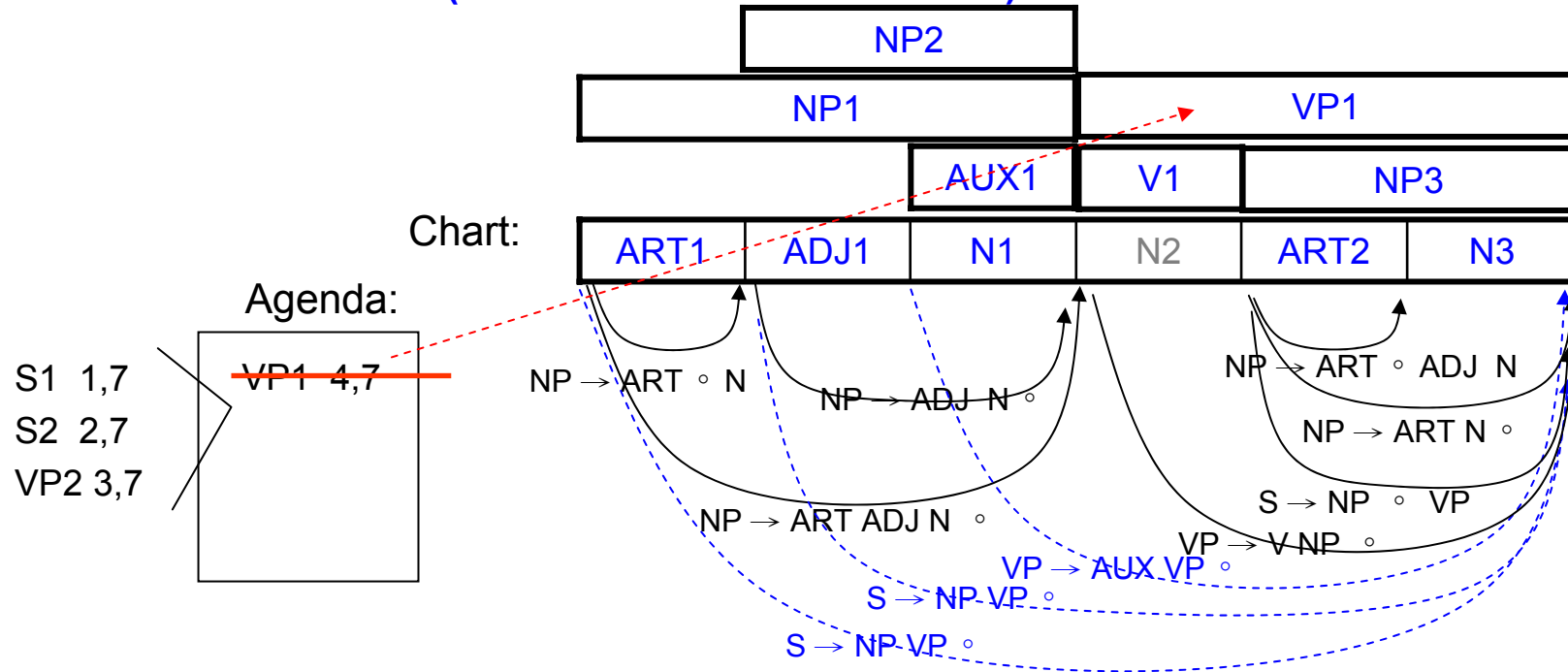
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 12 (using the arc extension algorithm)

Enter VP1: ("hold the water" from 4 to 7)



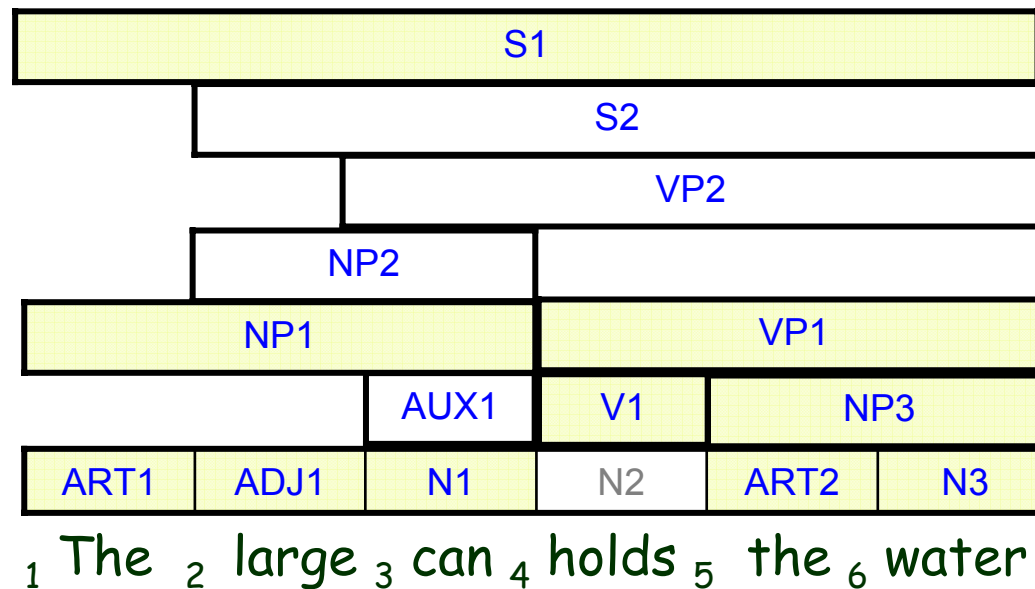
The Bottom-Up Chart Parser (28/29)

- Example ₁ The ₂ large ₃ can ₄ holds ₅ the ₆ water ₇

| | |
|-------------------|----------------|
| 1. S → NP VP | 4. NP → ADJ N |
| 2. NP → ART ADJ N | 5. VP → AUX VP |
| 3. NP → ART N | 6. VP → V NP |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

The final Chart



The Bottom-Up Chart Parser (29/29)

- Characteristics
 - The algorithm always moves forward through the chart making additions as it goes
 - Arcs are never removed and the algorithm never backtracks to a previous chart entry once it has moved on
 - Different S structures might share the common subparts represented in the chart only once

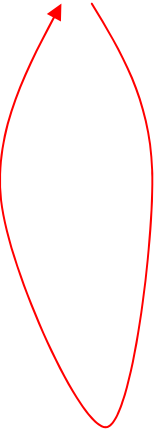
The Top-Down Chart Parser (1/27)

- The Top-Down Chart Parsing Algorithm

Initialization: For every rule in the grammar of form $S \rightarrow X_1 \dots X_k$, add an arc labeled $S \rightarrow \circ X_1 \dots X_k$ using **the arc introduction algorithm**

1. If the agenda is empty, look up the interpretations of the next word and add them to the agenda
2. Select a constituent C from the agenda
3. Using **the arc extension algorithm**, combine C with every active arc on the chart. Any new constituents are added to the agenda
4. For any active arcs created in step 3, add them to the chart using the following **top-down arc introduction algorithm**

Loop until
no input left



- To add an arc $S \rightarrow C_1 \dots \circ C_i \dots C_n$ ending at position j , do the following:

For each rule in the grammar of form $C_i \rightarrow X_1 \dots X_k$, recursively add the new arc $C_i \rightarrow \circ X_1 \dots X_k$ from position j to j

The Top-Down Chart Parser (2/27)

- Recall “**the arc extension algorithm**”
 - Insert **C into the chart** from position p_1 to p_2
 - For any active arc of the form $X \rightarrow X_1 \dots \circ C \dots X_n$ from position p_0 to p_1 add a new active arc $X \rightarrow X_1 \dots C \circ \dots X_n$ from p_0 to p_2
 - For any active arc of the form $X \rightarrow X_1 \dots X_n \circ C$ from position p_0 to p_1 , then add a new constituent of type X from p_0 to p_2 **to the agenda**

The Top-Down Chart Parser (3/27)

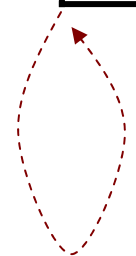
- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Initialization: (using the arc introduction algorithm)



Agenda:



- $S \rightarrow \circ NP VP$
- $NP \rightarrow \circ ART ADJ N$
- $NP \rightarrow \circ ART N$
- $NP \rightarrow \circ ADJ N$

} using the arc introduction algorithm

Note that no checking of 3rd-person-sg or non-3rd-person-sg verbs was applied here.

The Top-Down Chart Parser (4/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

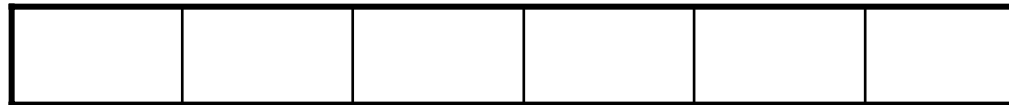
the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 1

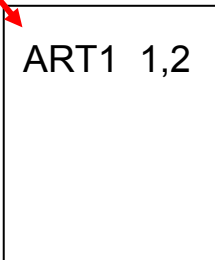
| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Enter ART1 (“the” from 1 to 2): Look at next word

Chart:



Agenda:



- $S \rightarrow \circ NP VP$
- $NP \rightarrow \circ ART ADJ N$
- $NP \rightarrow \circ ART N$
- $NP \rightarrow \circ ADJ N$

The Top-Down Chart Parser (5/27)

- Example

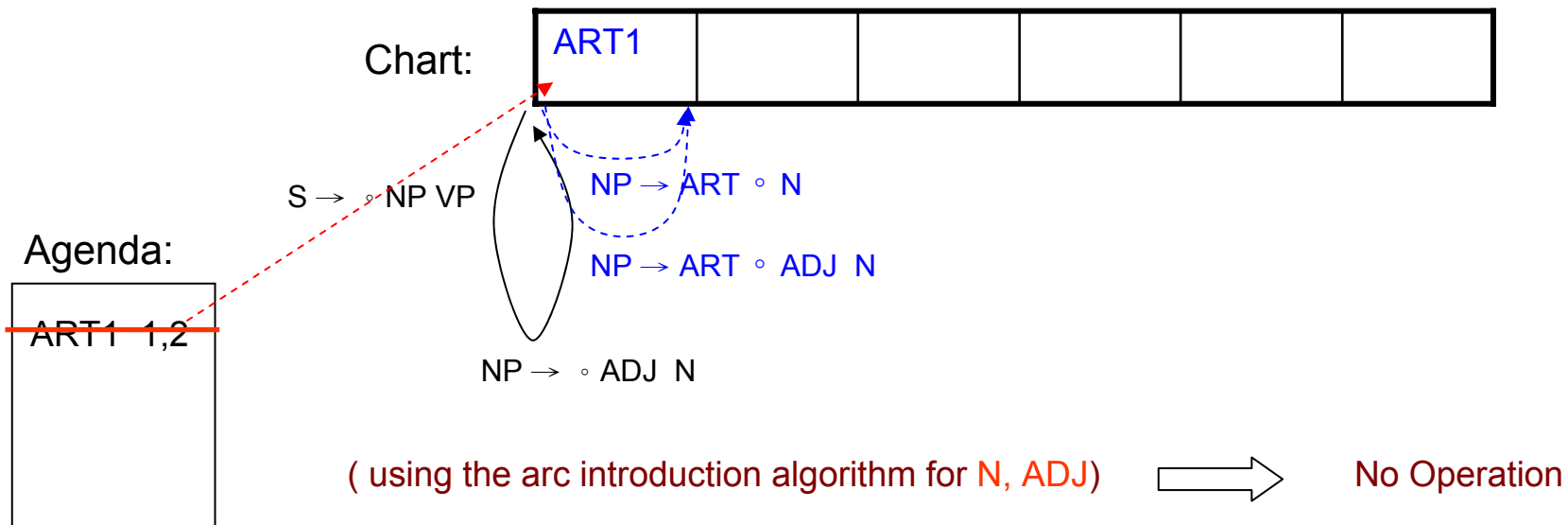
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 1

Enter ART1 (“the” from 1 to 2): (using the arc extension algorithm)



The Top-Down Chart Parser (6/27)

- Example

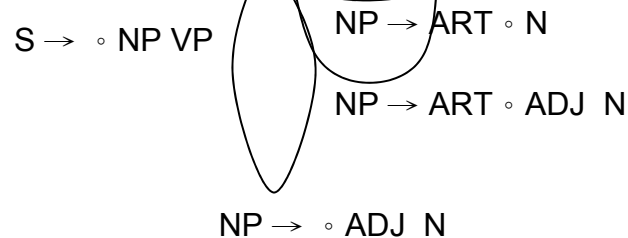
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

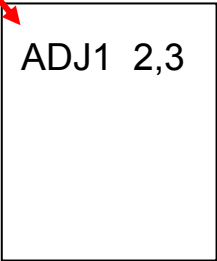
the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 2

Enter ART1 ("*large*" from 2 to 3): Look at next word



Agenda:



The Top-Down Chart Parser (7/27)

- Example

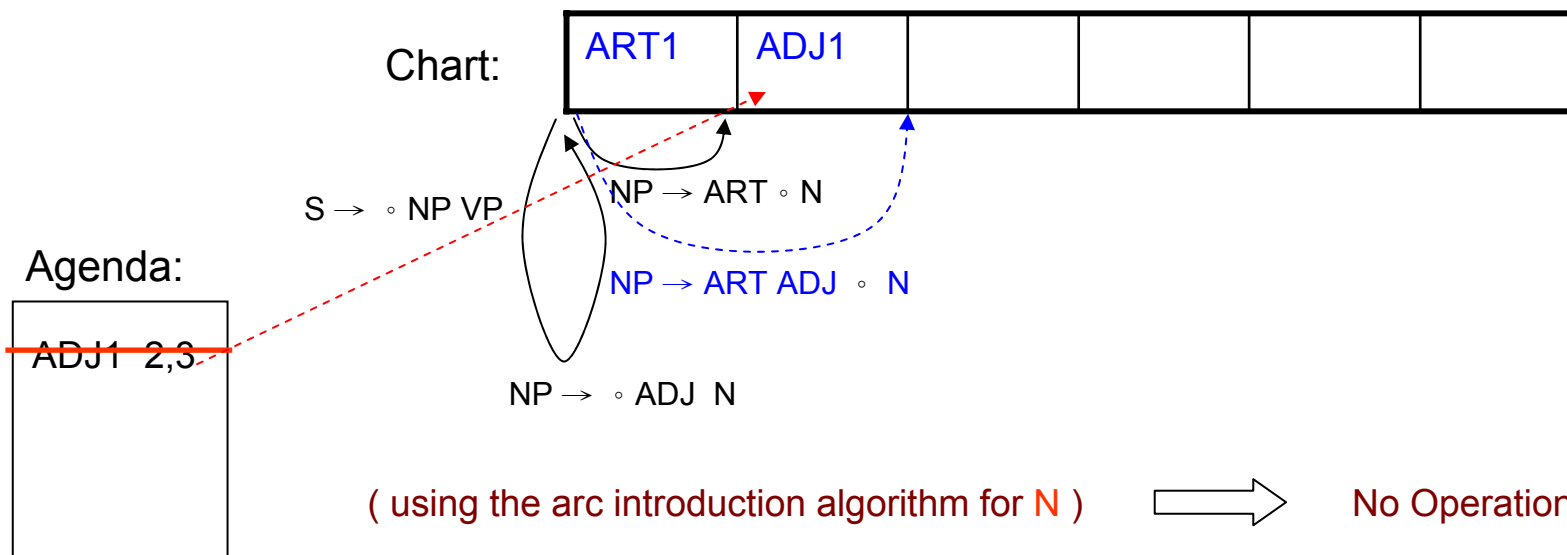
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 2

Enter ADJ1 (“large” from 2 to 3): (using the arc extension algorithm)



The Top-Down Chart Parser (8/27)

- Example

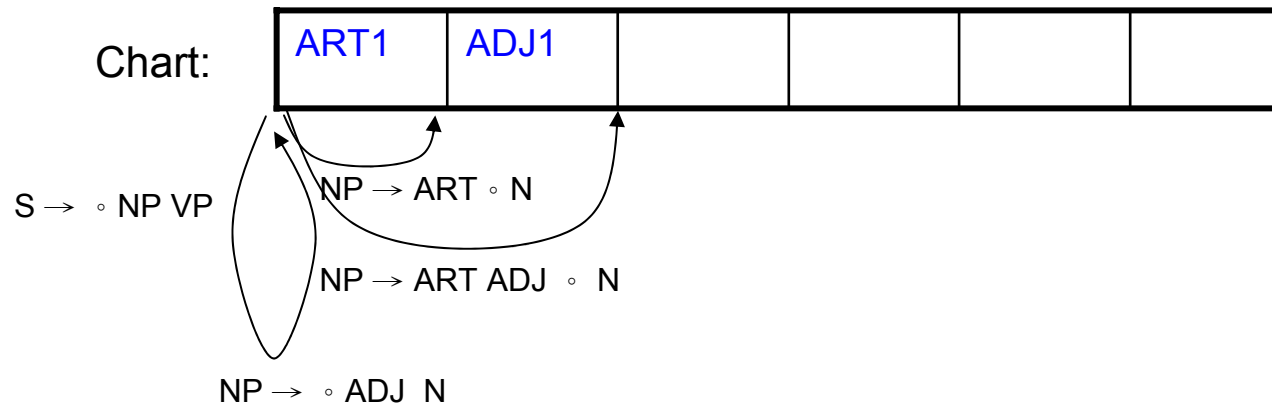
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 3

Enter N1 (“can” from 3 to 4): Look at next word



Agenda:

| |
|----------|
| N1 3,4 |
| AUX1 3,4 |

The Top-Down Chart Parser (9/27)

- Example

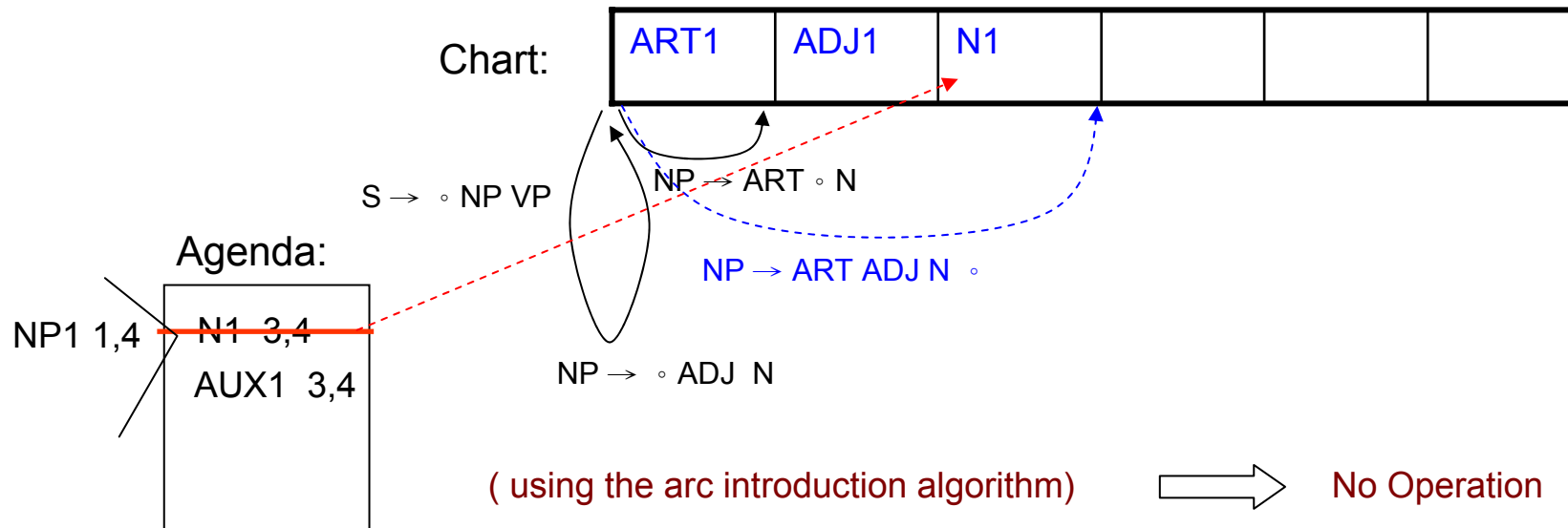
1 The 2 large 3 can 4 holds 5 the 6 water 7

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 3

Enter N1 ("can" from 3 to 4): (using the arc extension algorithm)



The Top-Down Chart Parser (10/27)

- Example

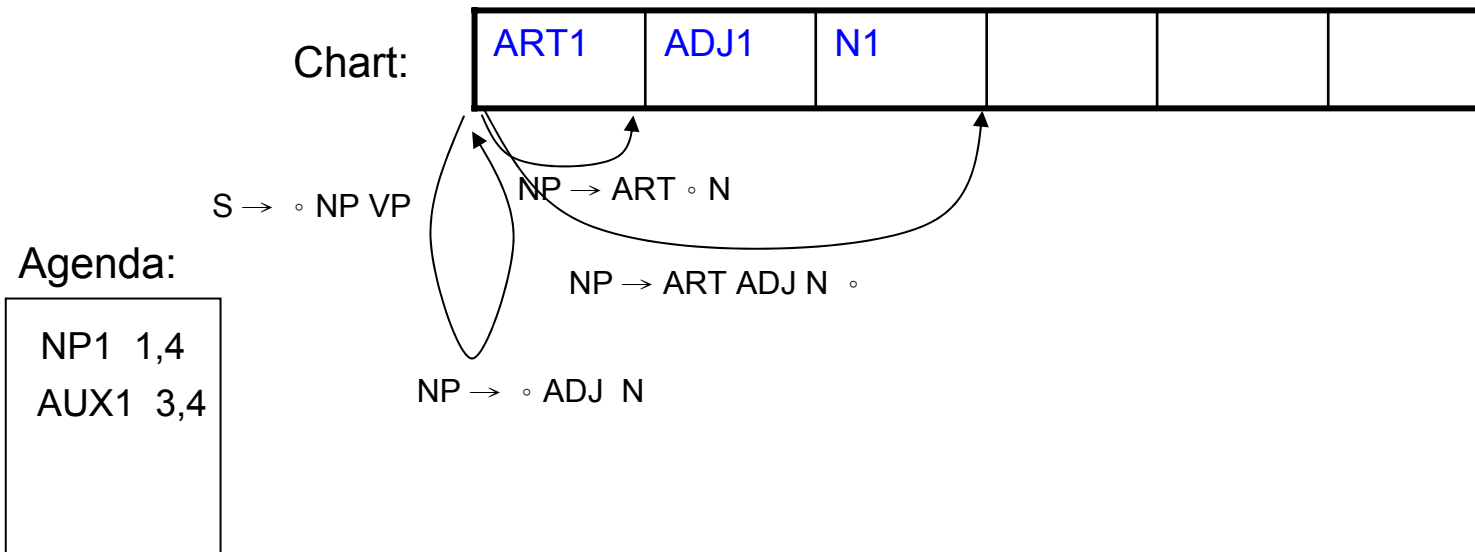
1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 4

Enter NP1 (“the large can” from 1 to 4):



The Top-Down Chart Parser (11/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

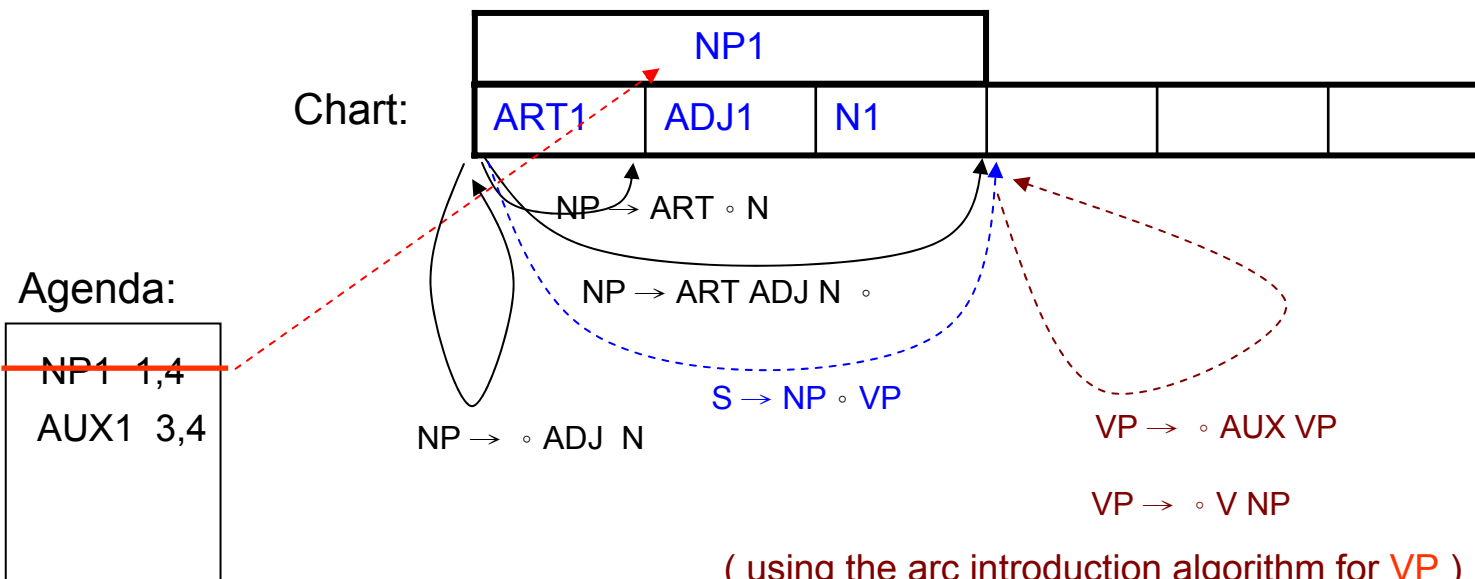
| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 4

Enter NP1 (“the large can” from 1 to 4):

(using the arc extension algorithm)



(using the arc introduction algorithm for VP)

The Top-Down Chart Parser (12/27)

- Example

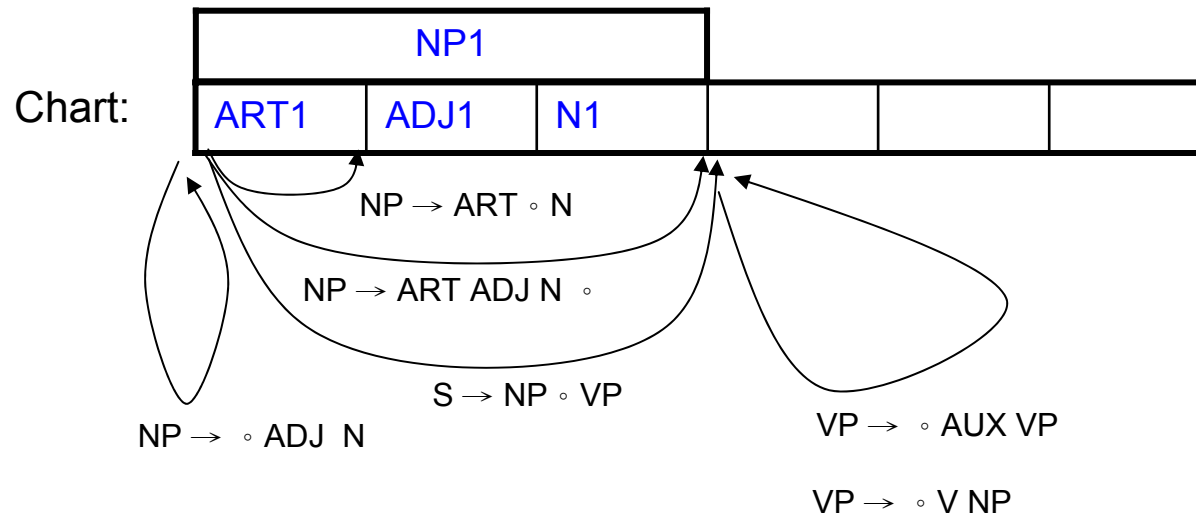
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 5

Enter AUX1 (“can” from 3 to 4):



Agenda:

| |
|----------|
| AUX1 3,4 |
|----------|

The Top-Down Chart Parser (13/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

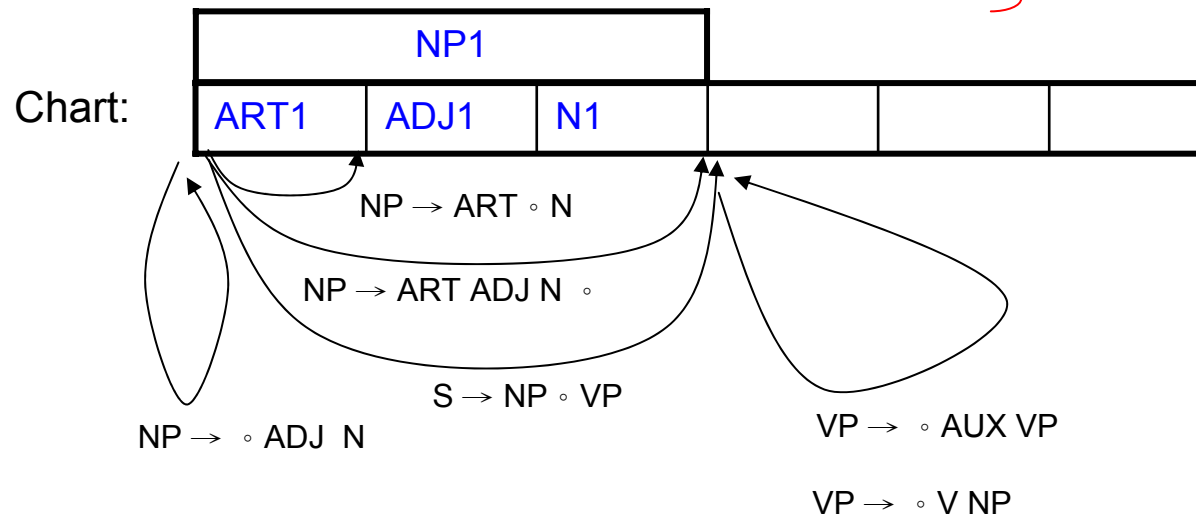
| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 5

Enter AUX1 (“can” from 3 to 4):

(using the arc extension algorithm)
 (using the arc introduction algorithm)

} No Operation



Agenda:

| |
|---------------------|
| AUX1 3,4 |
|---------------------|

The Top-Down Chart Parser (14/27)

- Example

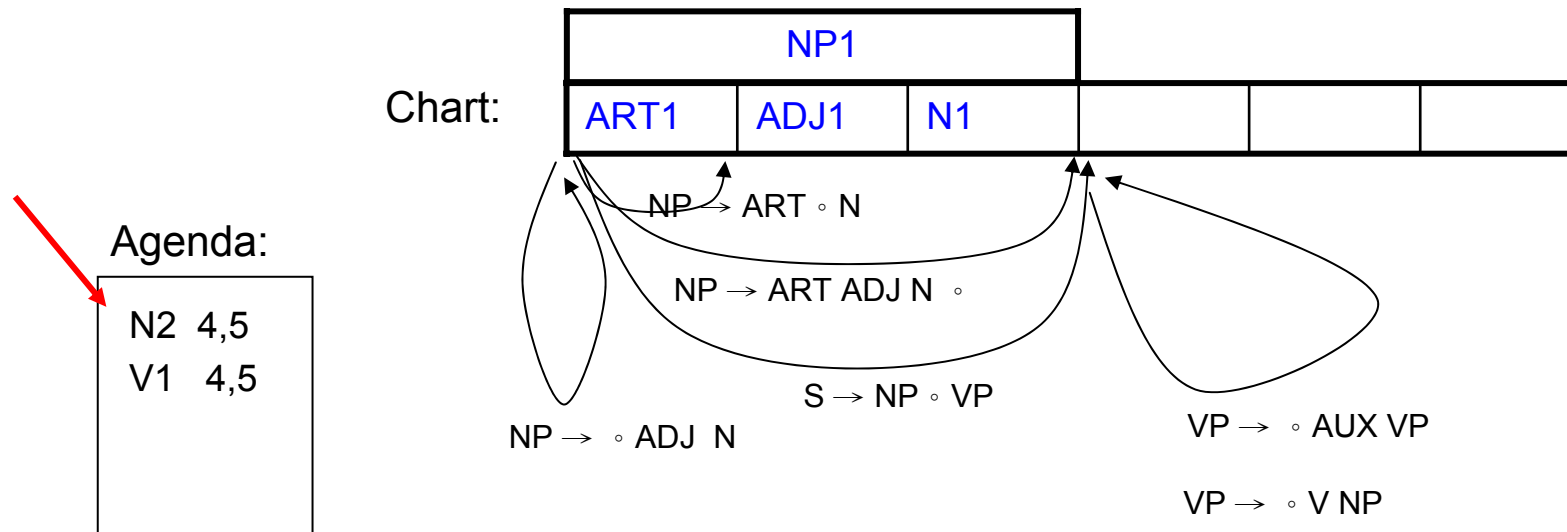
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------|----------------|
| 1. S → NP VP | 4. NP → ADJ N |
| 2. NP → ART ADJ N | 5. VP → AUX VP |
| 3. NP → ART N | 6. VP → V NP |

Loop 6

Enter N2 (“holds” from 4 to 5): Look at next word



The Top-Down Chart Parser (15/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

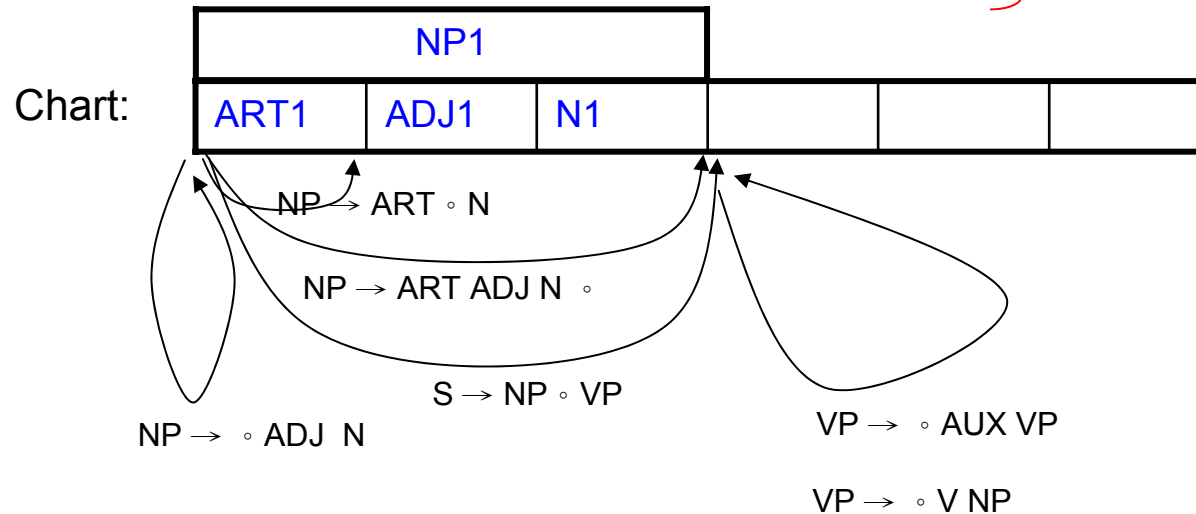
Loop 6

Enter N2 ("holds" from 4 to 5):

(using the arc extension algorithm)

(using the arc introduction algorithm)

} **No Operation**



Agenda:

| |
|-------------------|
| N2 4,5 |
| V1 4,5 |

The Top-Down Chart Parser (16/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

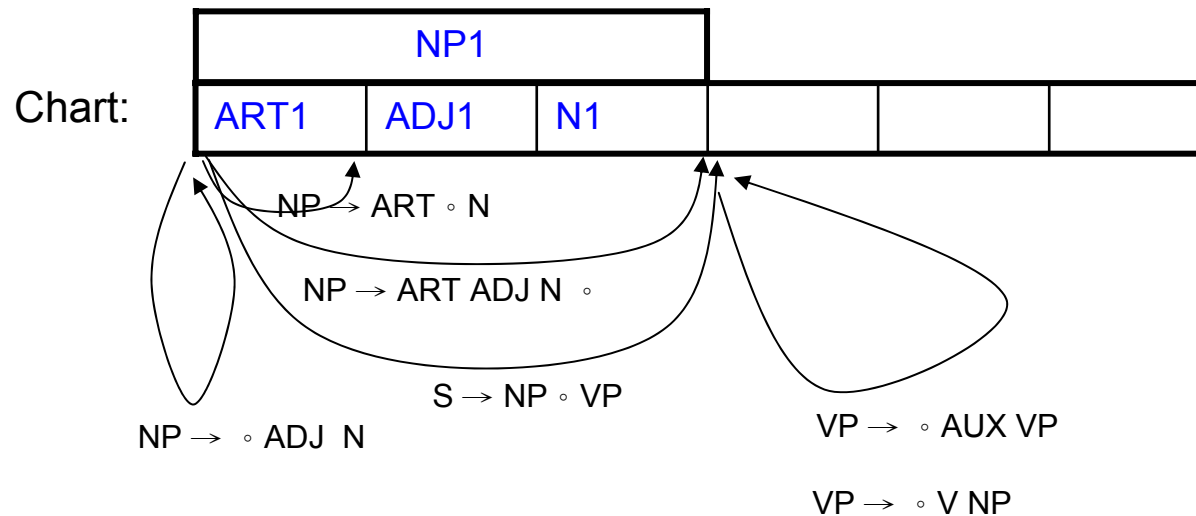
| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 7

Enter V1 (“holds” from 4 to 5):

Agenda:

| |
|--------|
| V1 4,5 |
|--------|



The Top-Down Chart Parser (17/27)

- Example

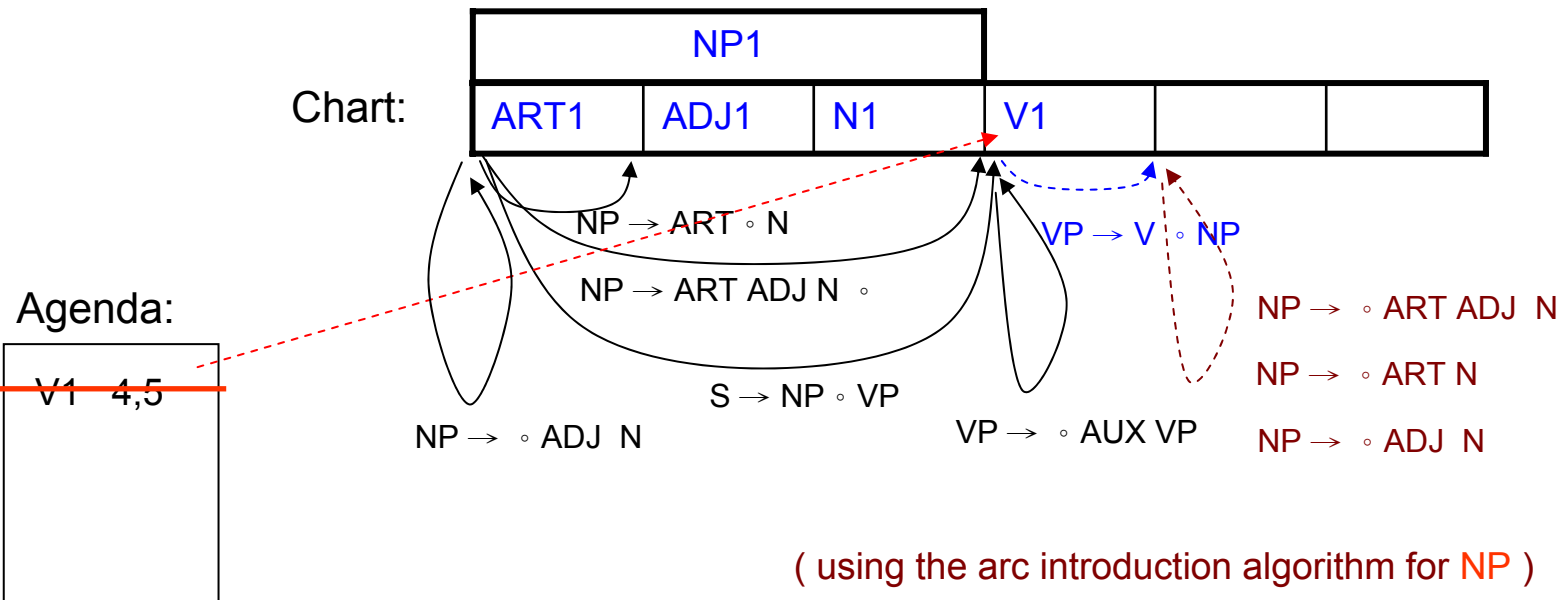
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
large: ADJ
can: N, AUX
hold: N, V
Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 7
Enter V1 ("holds" from 4 to 5):

(using the arc extension algorithm)



The Top-Down Chart Parser (18/27)

- Example

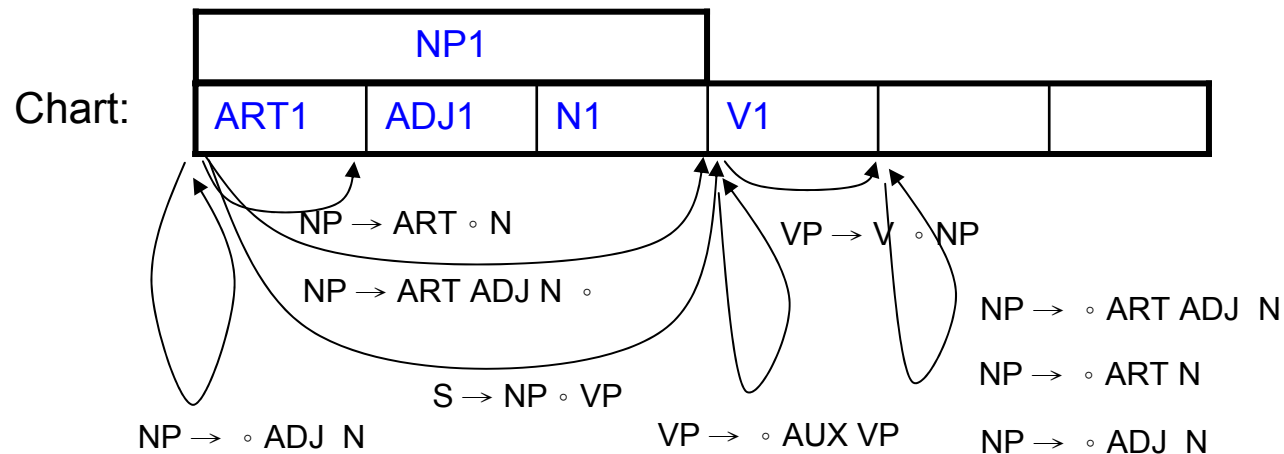
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 8

Enter **ART2** ("*the*" from 5 to 6): Look at next word



The Top-Down Chart Parser (19/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

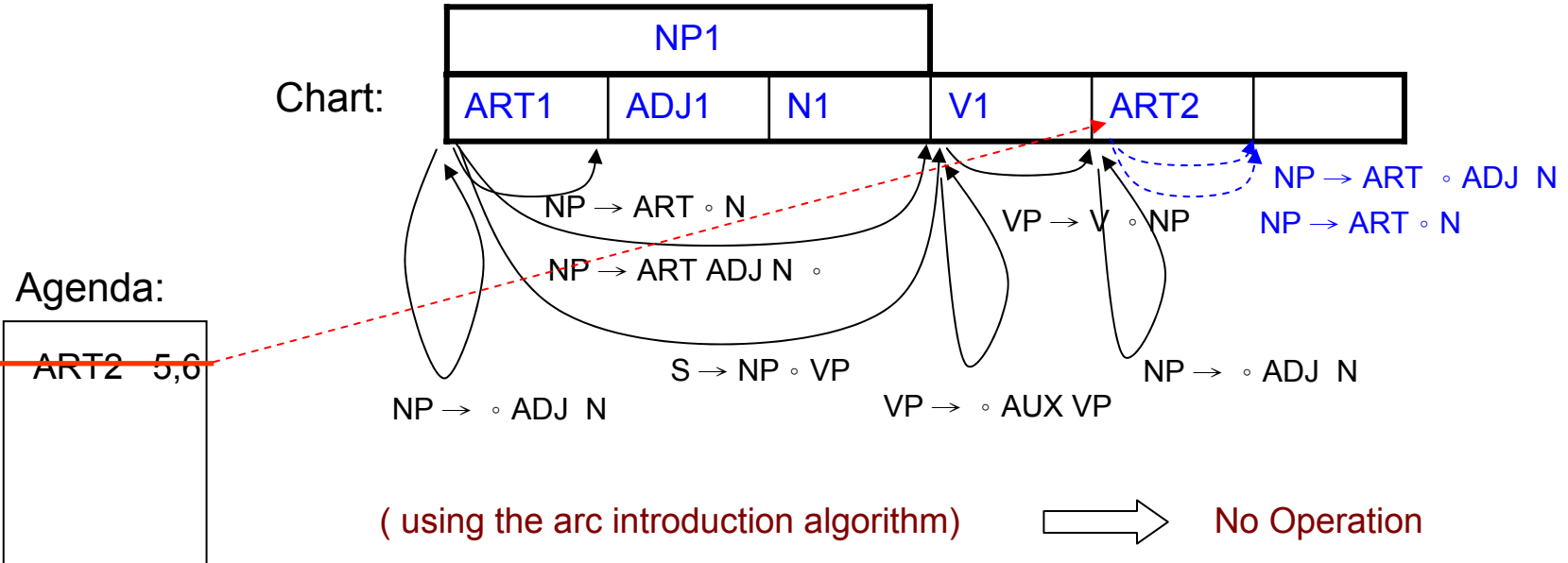
the: ART
large: ADJ
can: N, AUX
hold: N, V
Water: N

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 8

Enter ART2 (“the” from 5 to 6):

(using the arc extension algorithm)



The Top-Down Chart Parser (20/27)

- Example

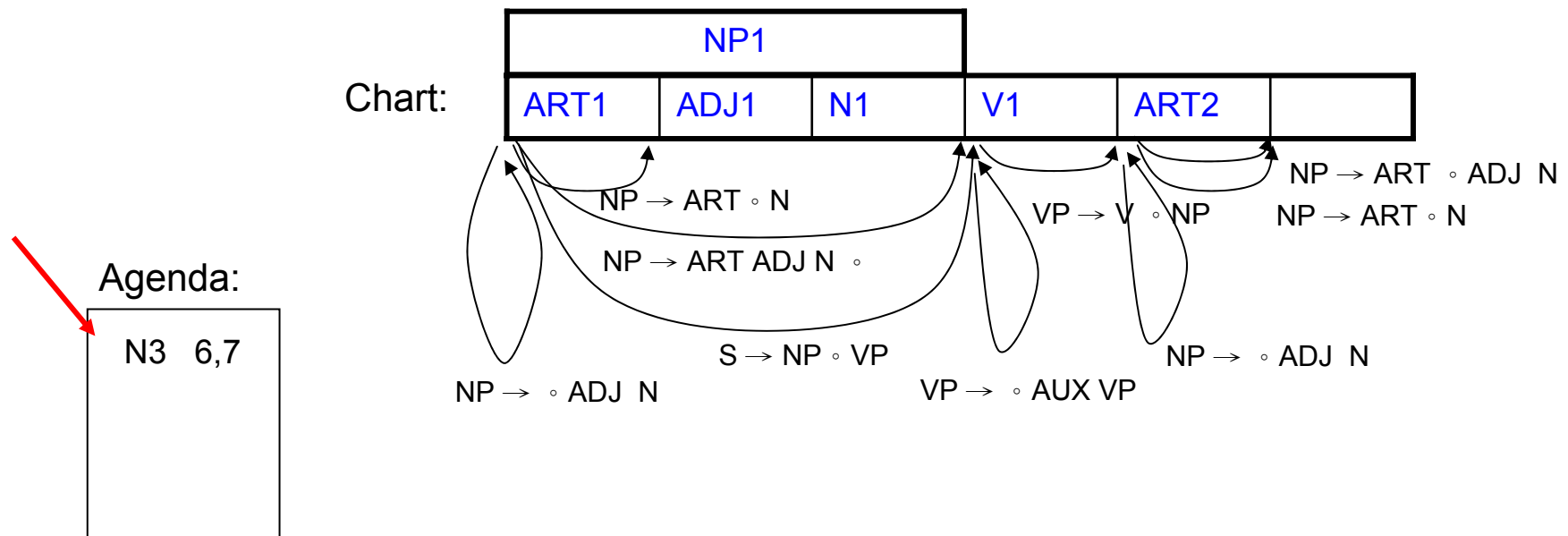
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 9

Enter N3 (“water” from 6 to 7): Look at next word



The Top-Down Chart Parser (21/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

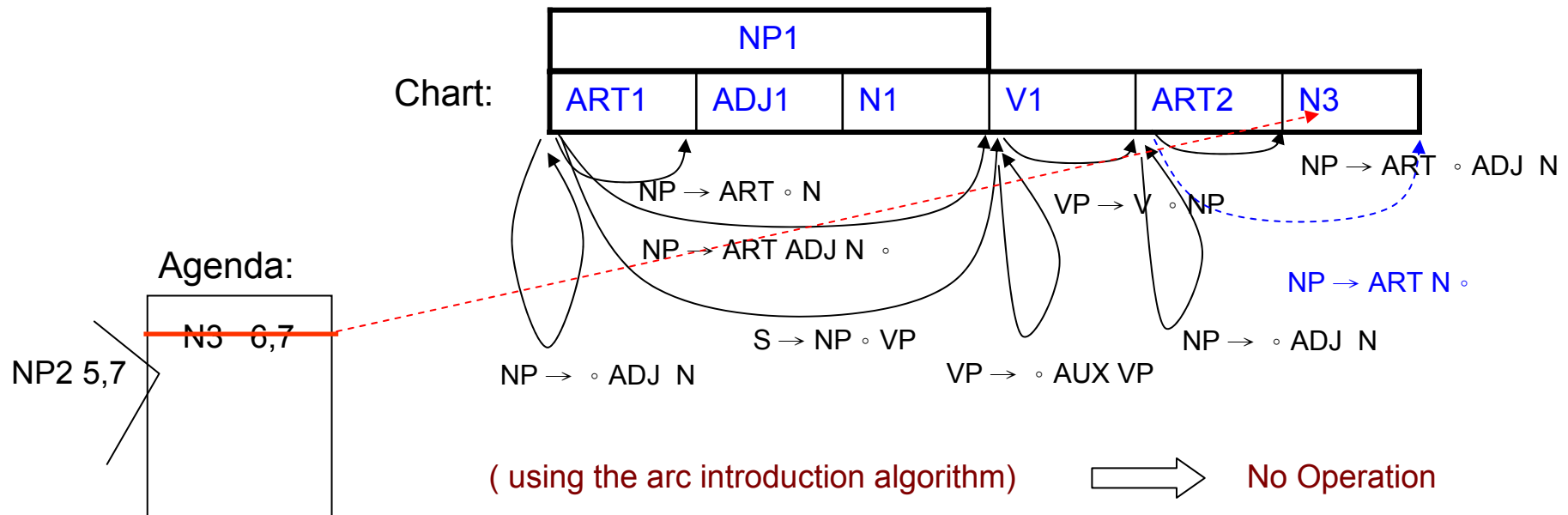
the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 9

Enter N3 ("water" from 6 to 7):

(using the arc extension algorithm)



The Top-Down Chart Parser (22/27)

- Example

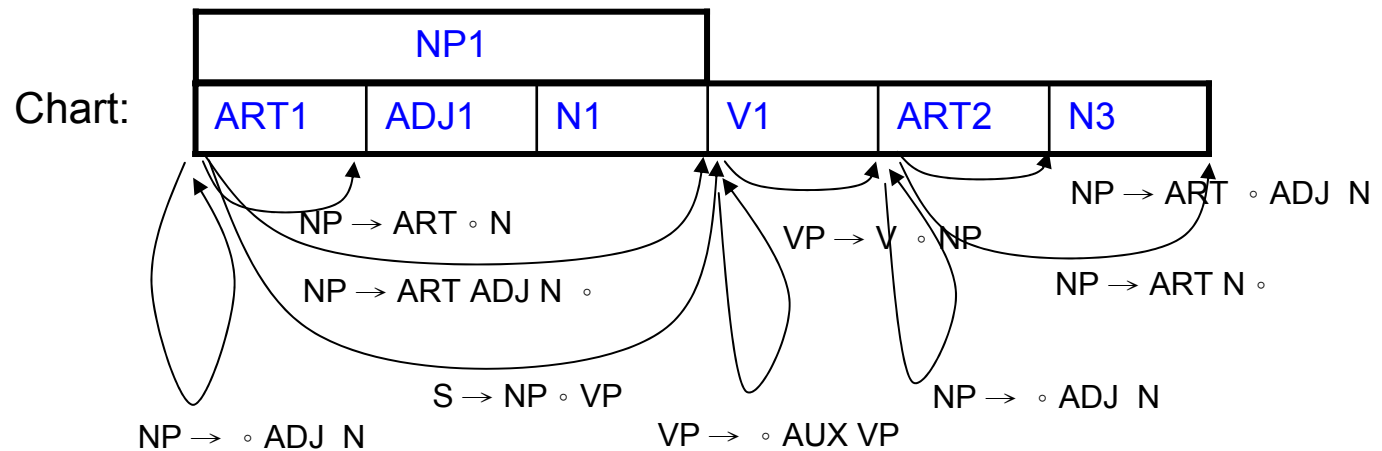
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 10

Enter NP2 (“the water” from 5 to 7):



Agenda:

| |
|---------|
| NP2 5,7 |
|---------|

The Top-Down Chart Parser (23/27)

- Example

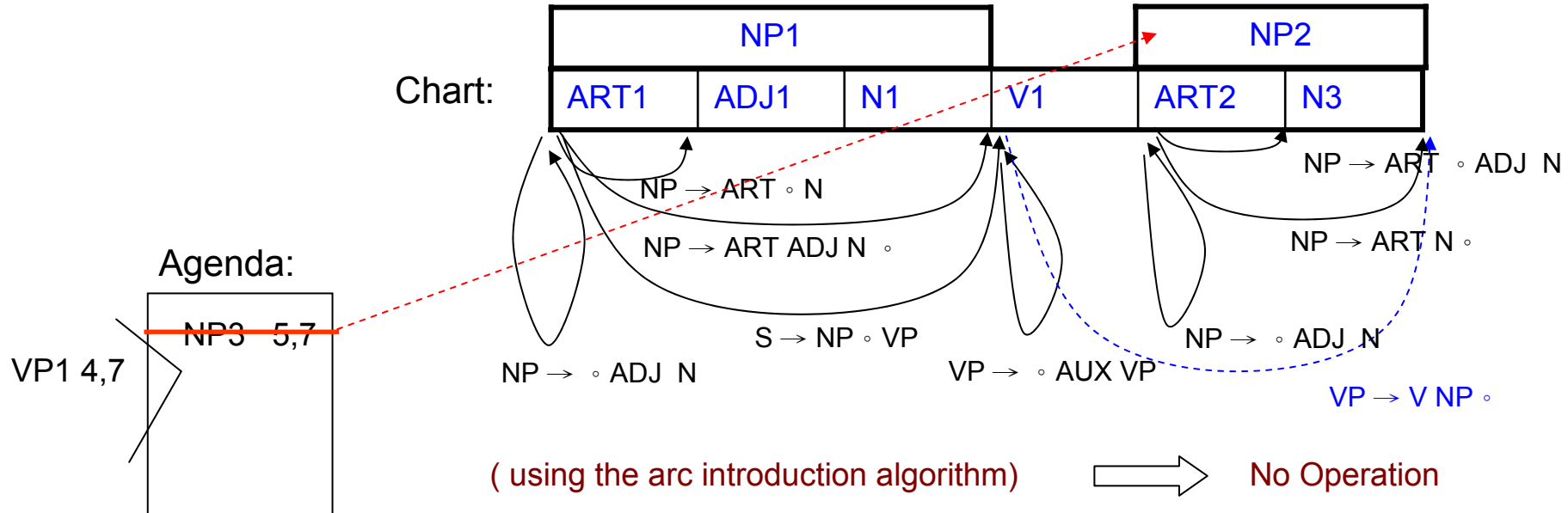
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 10

Enter NP2 (“the water” from 5 to 7):
 (using the arc extension algorithm)



The Top-Down Chart Parser (24/27)

- Example

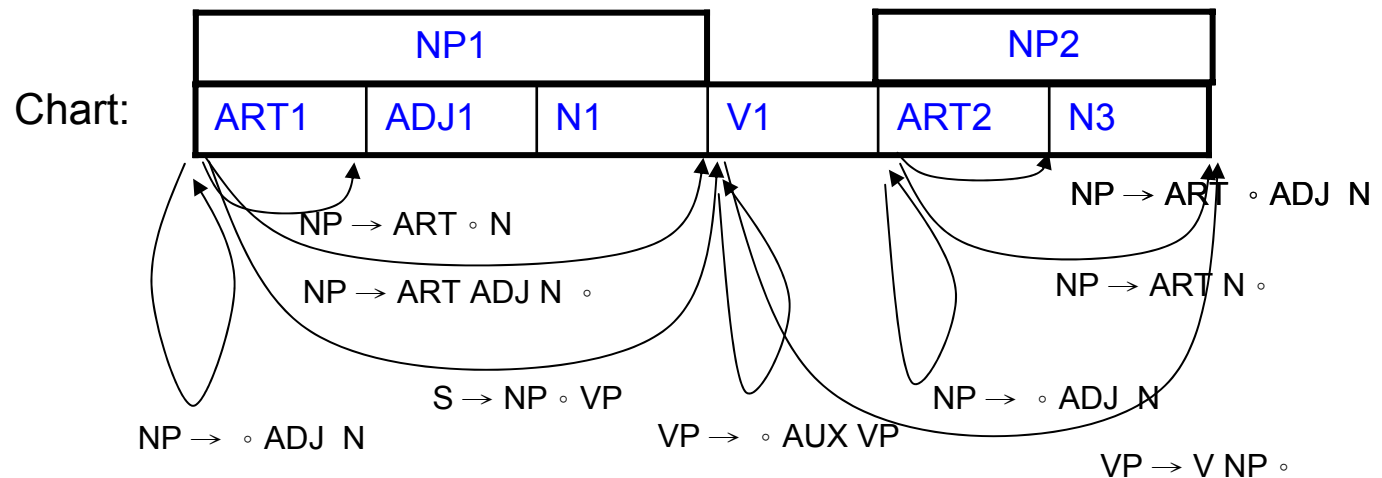
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 11

Enter VP1 (“holds the water” from 4 to 7):



Agenda:

VP1 4,7

The Top-Down Chart Parser (25/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

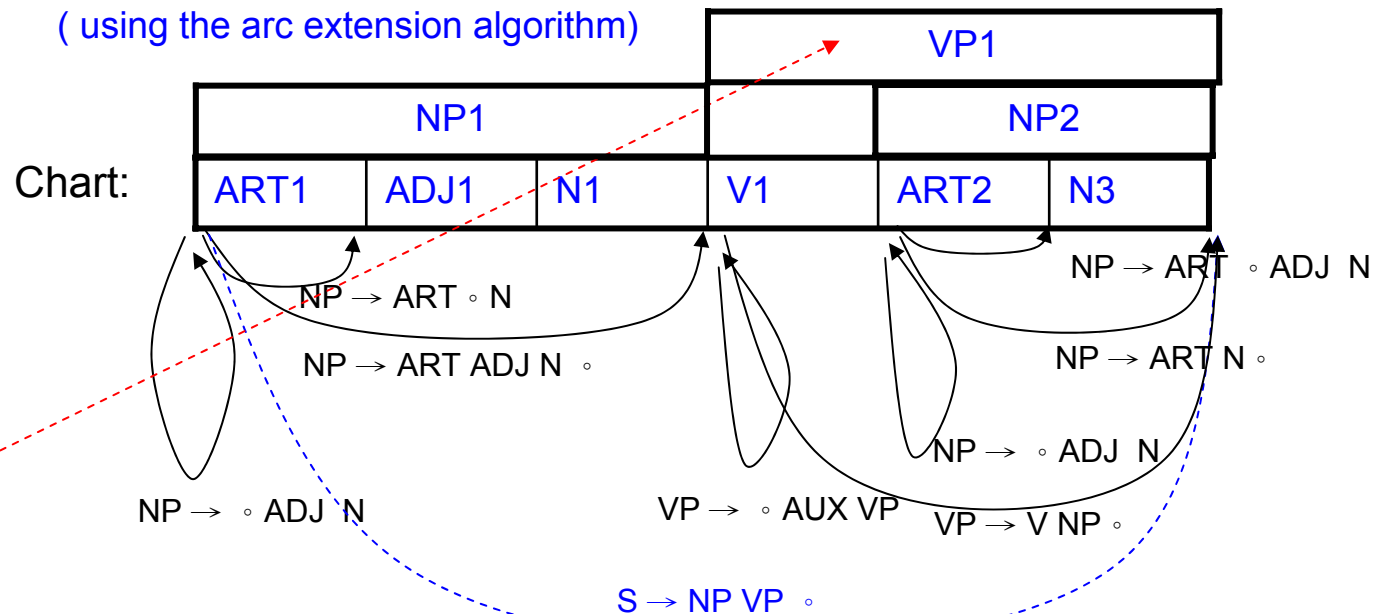
the: ART
large: ADJ
can: N, AUX
hold: N, V
Water: N

- | | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

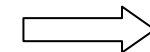
Loop 11

Enter VP1 (“holds the water” from 4 to 7):

(using the arc extension algorithm)



(using the arc introduction algorithm)



No Operation

The Top-Down Chart Parser (26/27)

- Example

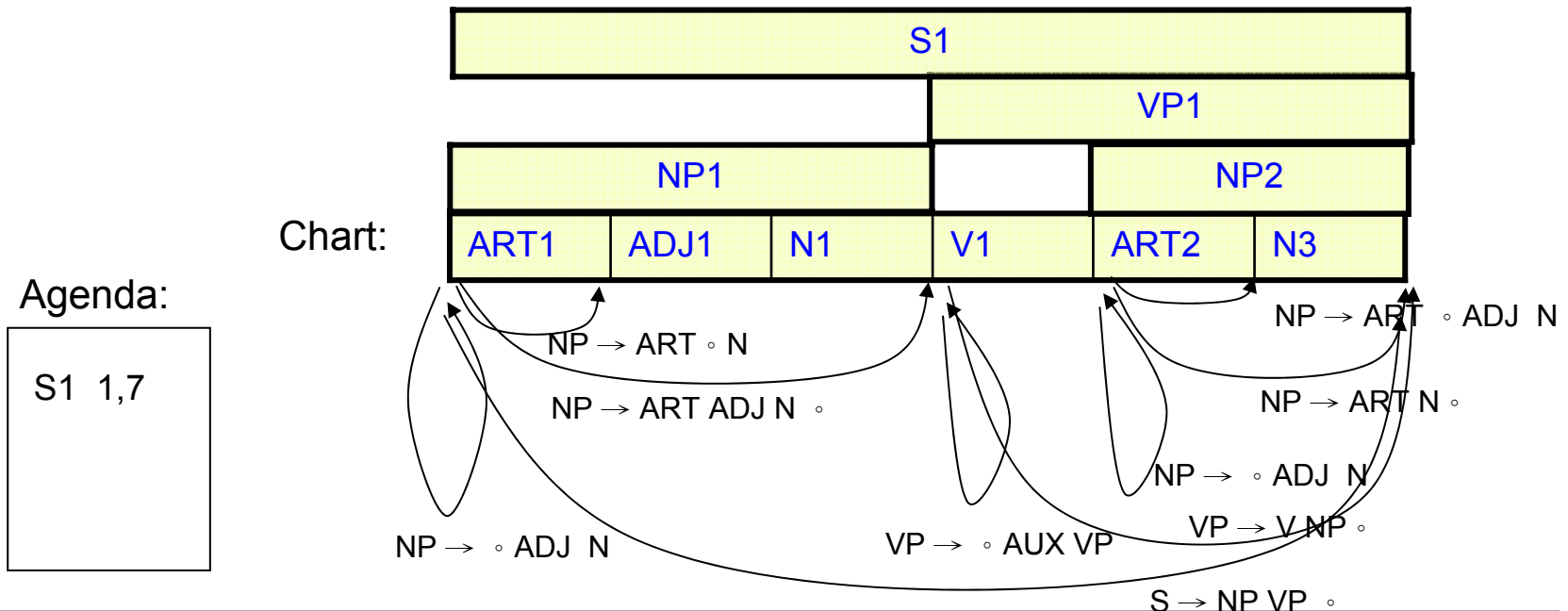
1 The 2 large 3 can 4 holds 5 the 6 water 7

the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

Loop 12

Enter S1 (“the large can holds the water” from 1 to 7):



The Top-Down Chart Parser (27/27)

- Example

1 The 2 large 3 can 4 holds 5 the 6 water 7

| | |
|-------------------------------|----------------------------|
| 1. $S \rightarrow NP VP$ | 4. $NP \rightarrow ADJ N$ |
| 2. $NP \rightarrow ART ADJ N$ | 5. $VP \rightarrow AUX VP$ |
| 3. $NP \rightarrow ART N$ | 6. $VP \rightarrow V NP$ |

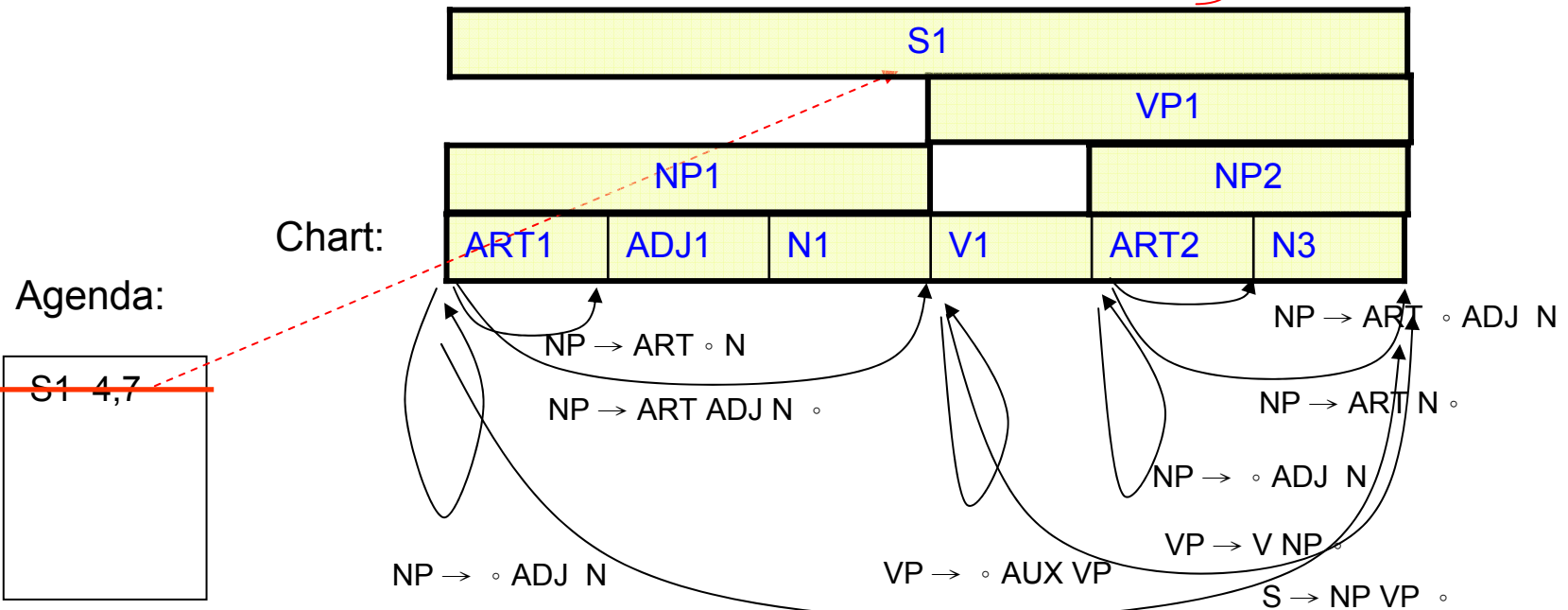
the: ART
 large: ADJ
 can: N, AUX
 hold: N, V
 Water: N

Loop 12

Enter S1 ("the large can holds the water" from 1 to 7):

(using the arc extension algorithm)
 (using the arc introduction algorithm)

No Operation



Comparisons

- The number of constituents generated by the top-down chart parser has dropped from 15 to 10
- In practice, the top-down method is considerably more efficient for any reasonable grammar

Homework-3

1. Natural Language Understanding, chapter 3
 - Exercises 8 and 9