11-442 / 11-642:
Search Engines

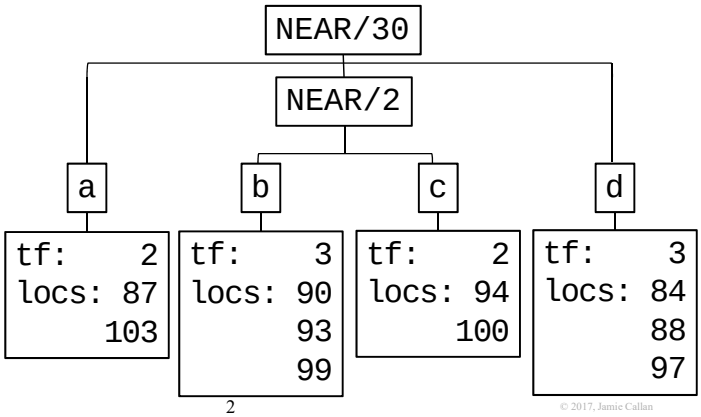
NEAR Operator Implementation

Jamie Callan
Carnegie Mellon University
callan@cs.cmu.edu

NEAR Implementation

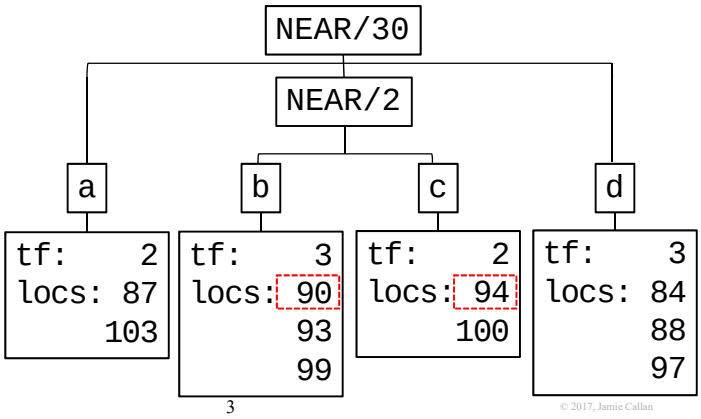
Jamie's implementation of #NEAR uses the QryIop iterators

- Sometimes the behavior is a slightly different than the greedy algorithm discussed earlier



NEAR Implementation

Initialization of the inner NEAR iterators



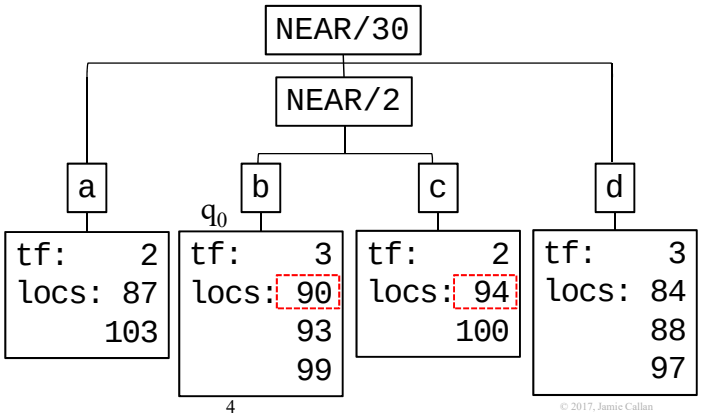
3

NEAR Implementation

Build a match left-to-right

currentLoc: 90

- Set currentLoc to location of query term q_0

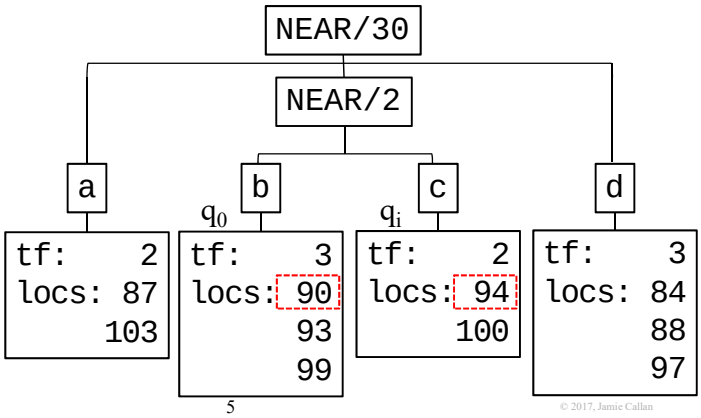


4

NEAR Implementation

Build a match left-to-right currentLoc: 90

- Advance the i'th iterator past currentLoc (if needed)

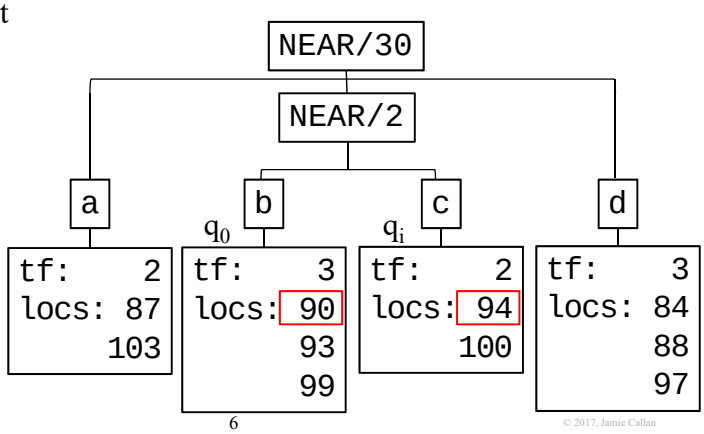


5

NEAR Implementation

Build a match left-to-right currentLoc: 90

- Compare locations of currentLoc and q_i
- Too far apart

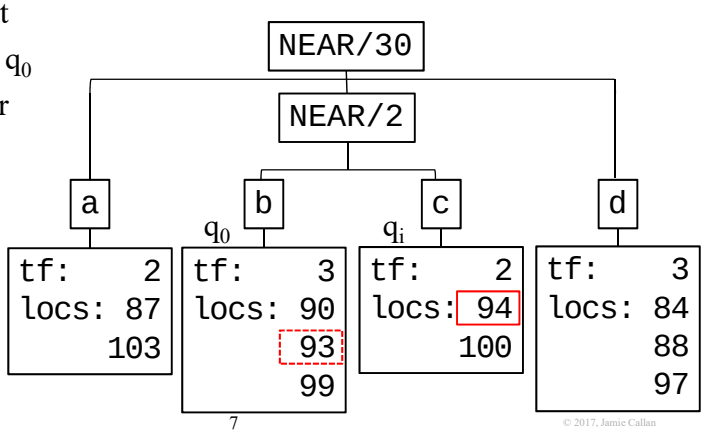


6

NEAR Implementation

Build a match left-to-right currentLoc: 90

- Compare locations of currentLoc and q_i
- Too far apart
  - Advance q_0
  - Start over

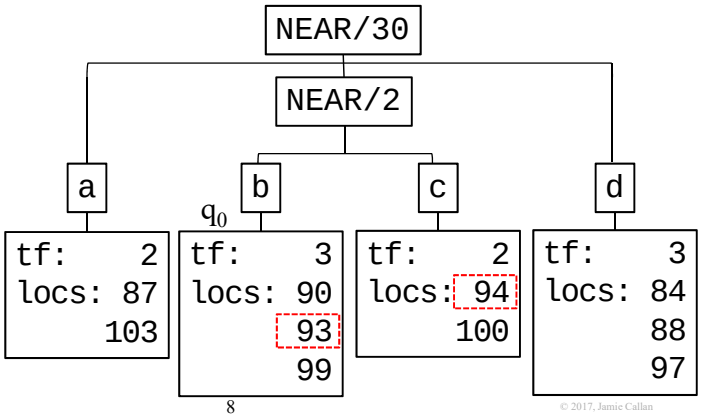


7

NEAR Implementation

Build a match left-to-right currentLoc: 93

- Set currentLoc to location of query term q_0

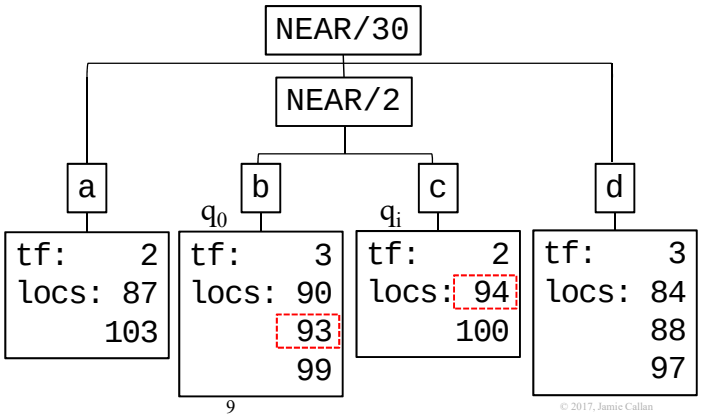


8

NEAR Implementation

Build a match left-to-right currentLoc: 93

- Advance the i'th iterator past currentLoc (if needed)

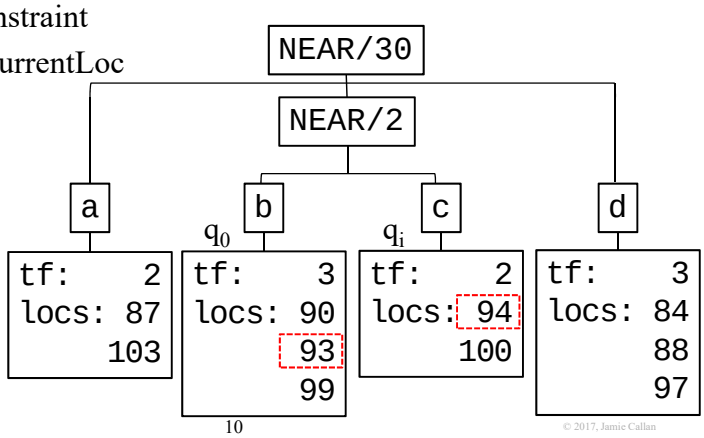


9

NEAR Implementation

Build a match left-to-right currentLoc: 94

- Compare locations of currentLoc and q_i
- Satisfied constraint
  - Update currentLoc



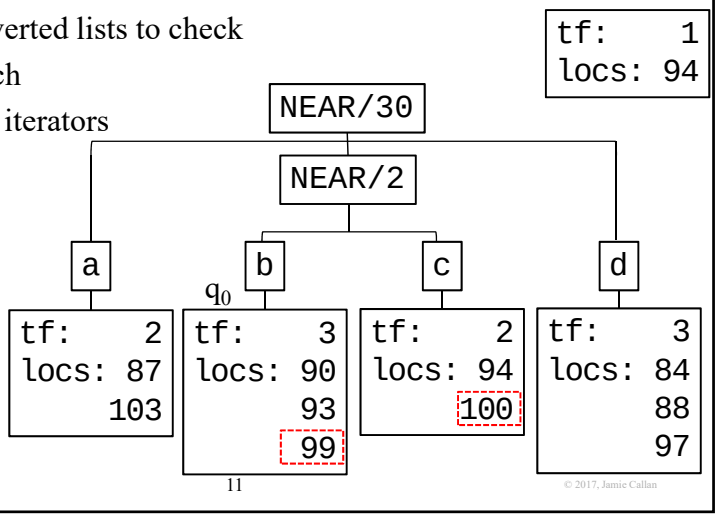
10

NEAR Implementation

Build a match left-to-right

- No more inverted lists to check
- Record match
- Advance all iterators

currentLoc: 94



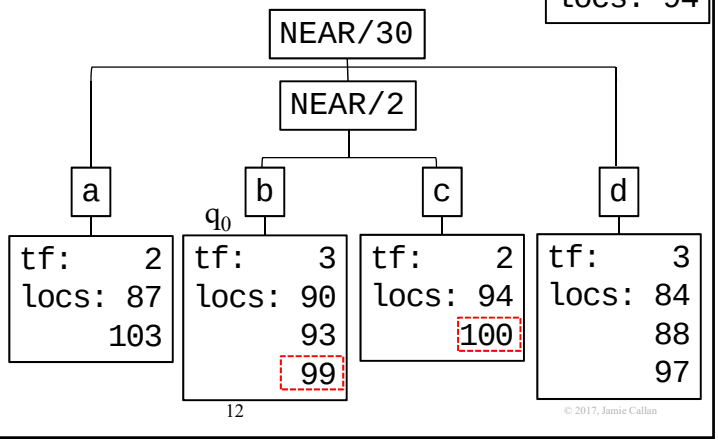
11

NEAR Implementation

Build a match left-to-right

- Set currentLoc to location of query term q_0

currentLoc: 99



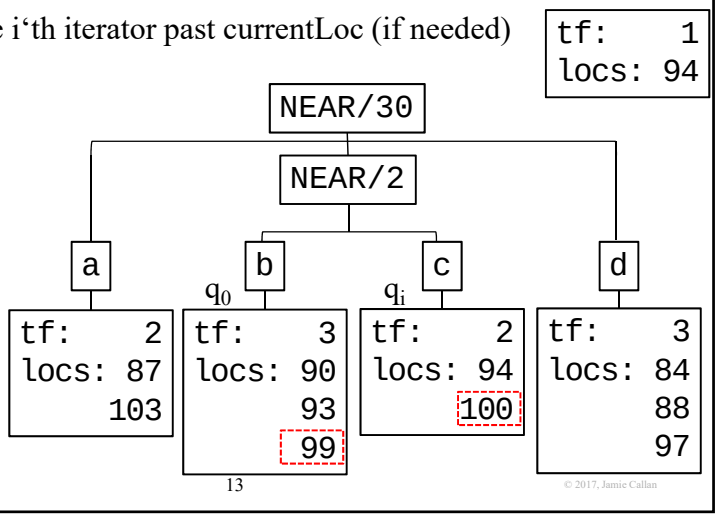
12

NEAR Implementation

Build a match left-to-right

currentLoc: 99

- Advance the i'th iterator past currentLoc (if needed)

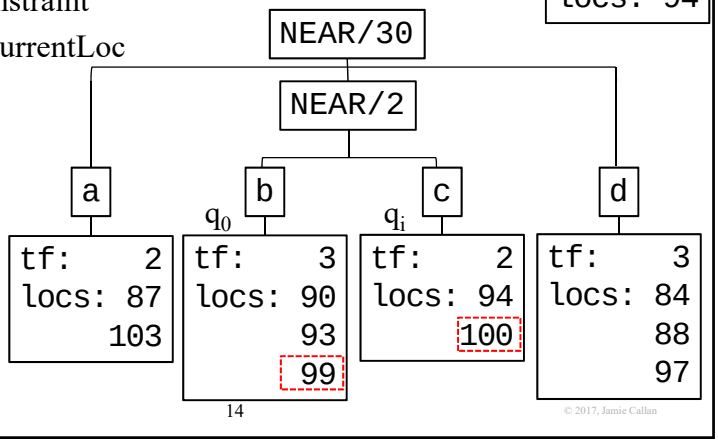


NEAR Implementation

Build a match left-to-right

currentLoc: 100

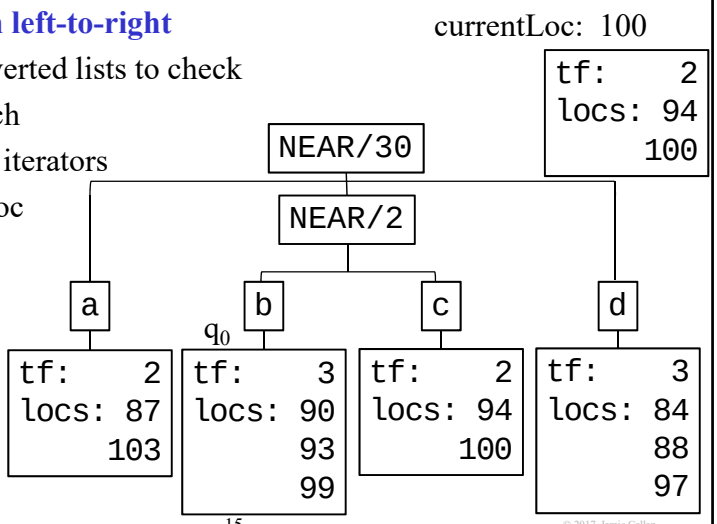
- Compare locations of currentLoc and q_i
- Satisfied constraint
  - Update currentLoc



NEAR Implementation

Build a match left-to-right

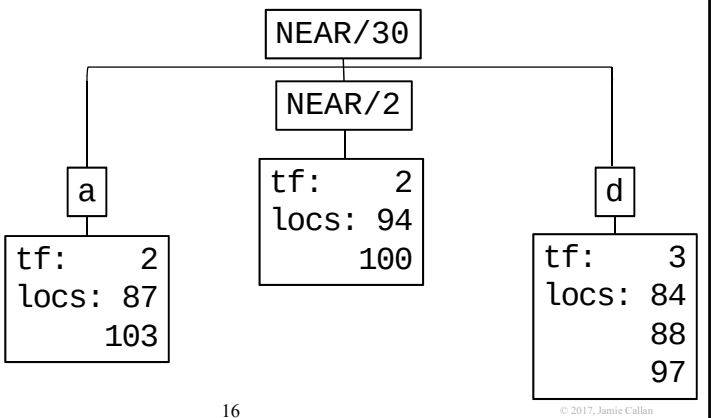
- No more inverted lists to check
- Record match
- Advance all iterators
  - End of doc



15

NEAR Implementation

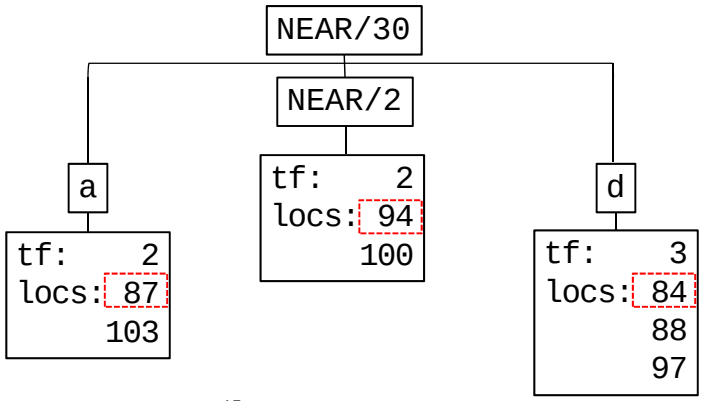
Initialization of inner NEAR operator is complete



16

NEAR Implementation

Initialization of the outer NEAR iterators



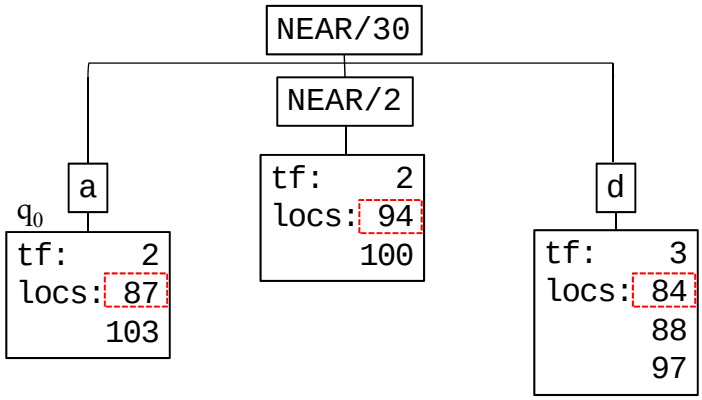
17

NEAR Implementation

Build a match left-to-right

currentLoc: 87

- Set currentLoc to location of query term q_0

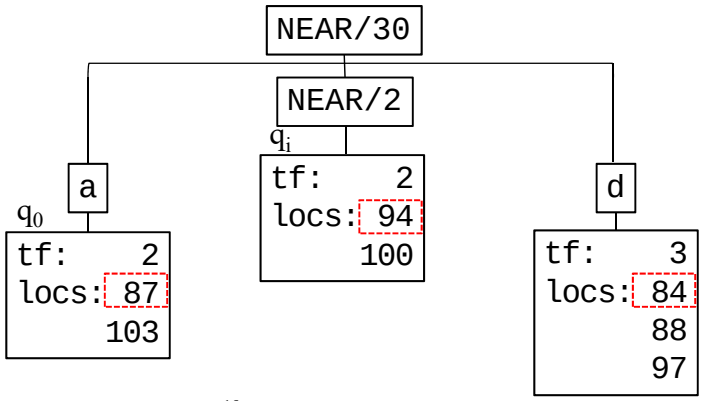


18

NEAR Implementation

Build a match left-to-right currentLoc: 87

- Advance the i'th iterator past currentLoc (if needed)

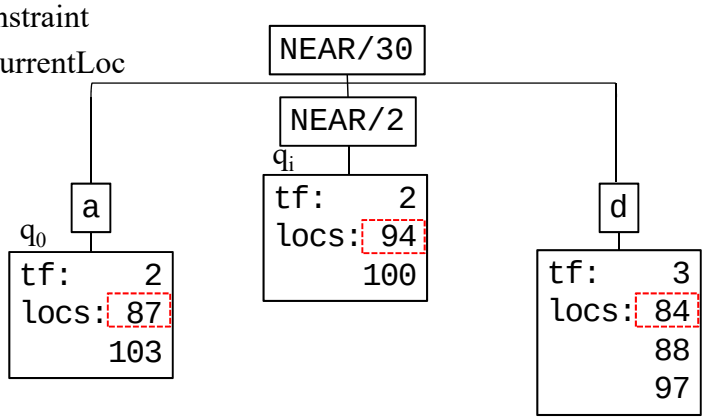


19

NEAR Implementation

Build a match left-to-right currentLoc: 94

- Compare locations of currentLoc and q_i
- Satisfied constraint
  - Update currentLoc

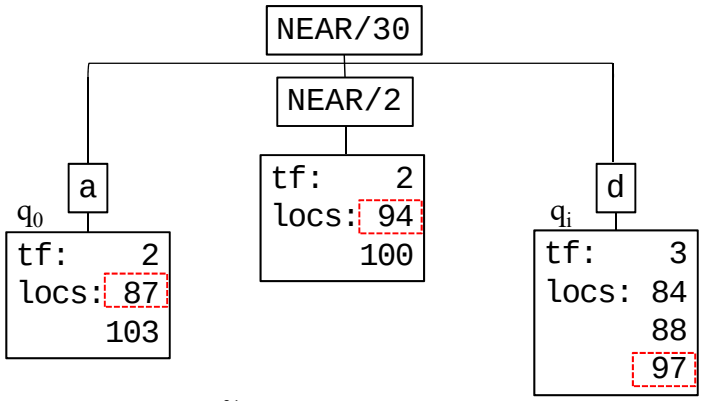


20

NEAR Implementation

Build a match left-to-right currentLoc: 94

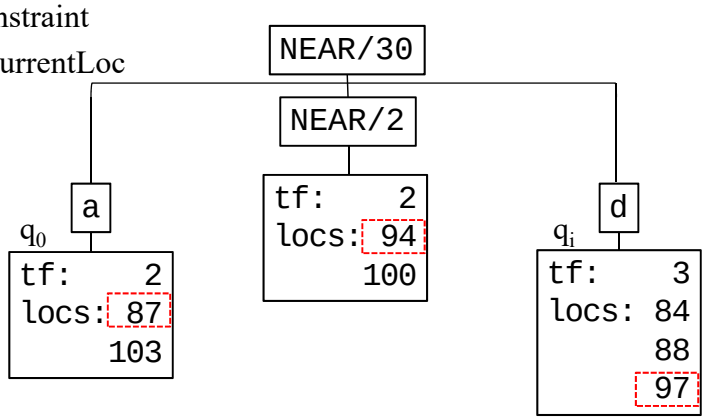
- Advance the i'th iterator past currentLoc (if needed)



NEAR Implementation

Build a match left-to-right currentLoc: 97

- Compare locations of currentLoc and q_i
- Satisfied constraint
  - Update currentLoc

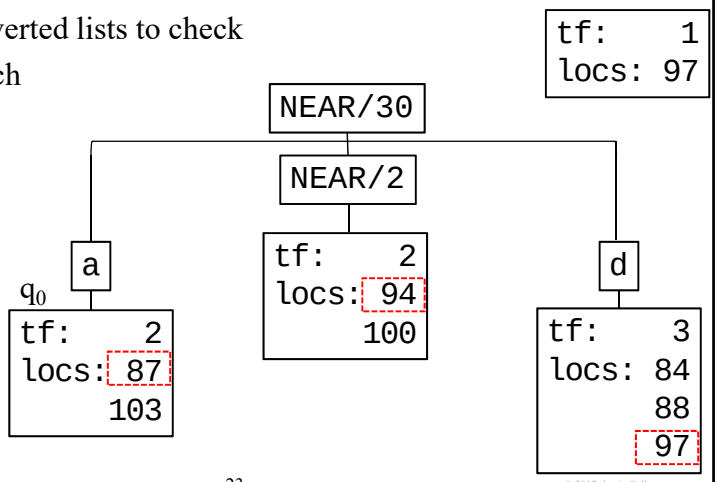


NEAR Implementation

Build a match left-to-right

- No more inverted lists to check
- Record match

currentLoc: 97

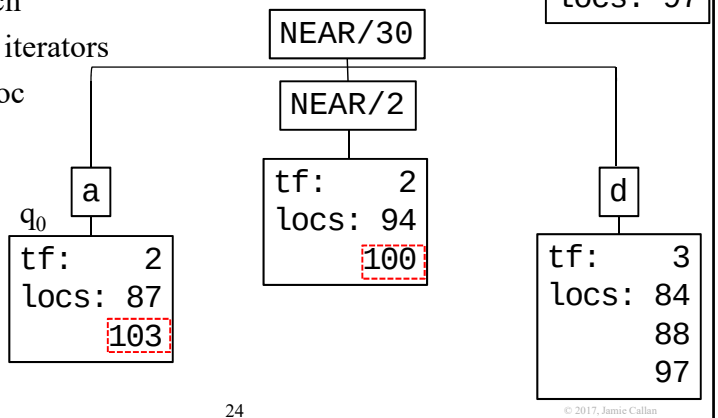


NEAR Implementation

Build a match left-to-right

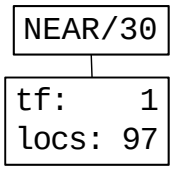
- No more inverted lists to check
- Record match
- Advance all iterators
  - End of doc

currentLoc: 97



NEAR Implementation

Initialization of outer NEAR operator is complete



NEAR Implementation

The simple greedy algorithm discussed earlier in Lecture 2 would not find a match in this document

- Use the iterators
- They make your code simpler and faster