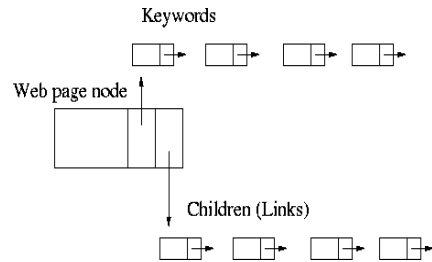


## What the Specifications Say

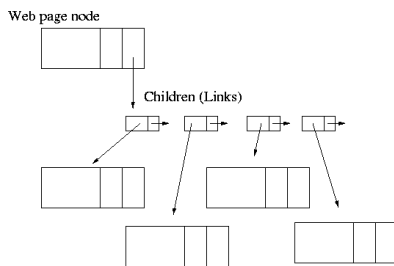
- Dynamic implementation of the tree structure
- Assume no limits on web pages and keywords
- Linked list of keywords per node
- Linked list of children per node
- Array-based heap to process web page ranks

## Keywords and Children



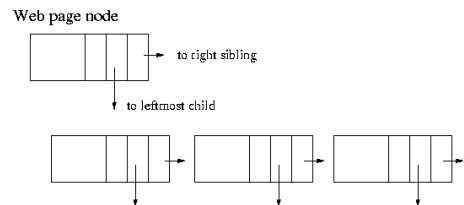
## Suggestion 1:

A child node points to a web page node



## Suggestion 2:

Child and sibling pointers on web page node



## Other Considerations

- Adding a parent pointer to the tree node
- Contents of keyword list for a web page
  - Just the keywords that occur in the page (easier to set up the tree, harder to process)
  - Or, include keywords that occur in descendant pages (involves more space but facilitates searching)

## Notes on the Min-Heap

- Since no limits will be given, the heap should be created after the file is read or the tree is created (number of web pages is known)
- To handle “ties” (web pages with the same rank), consider associating a sequence number to a web page, to ensure pages that occur earlier in the file are removed first