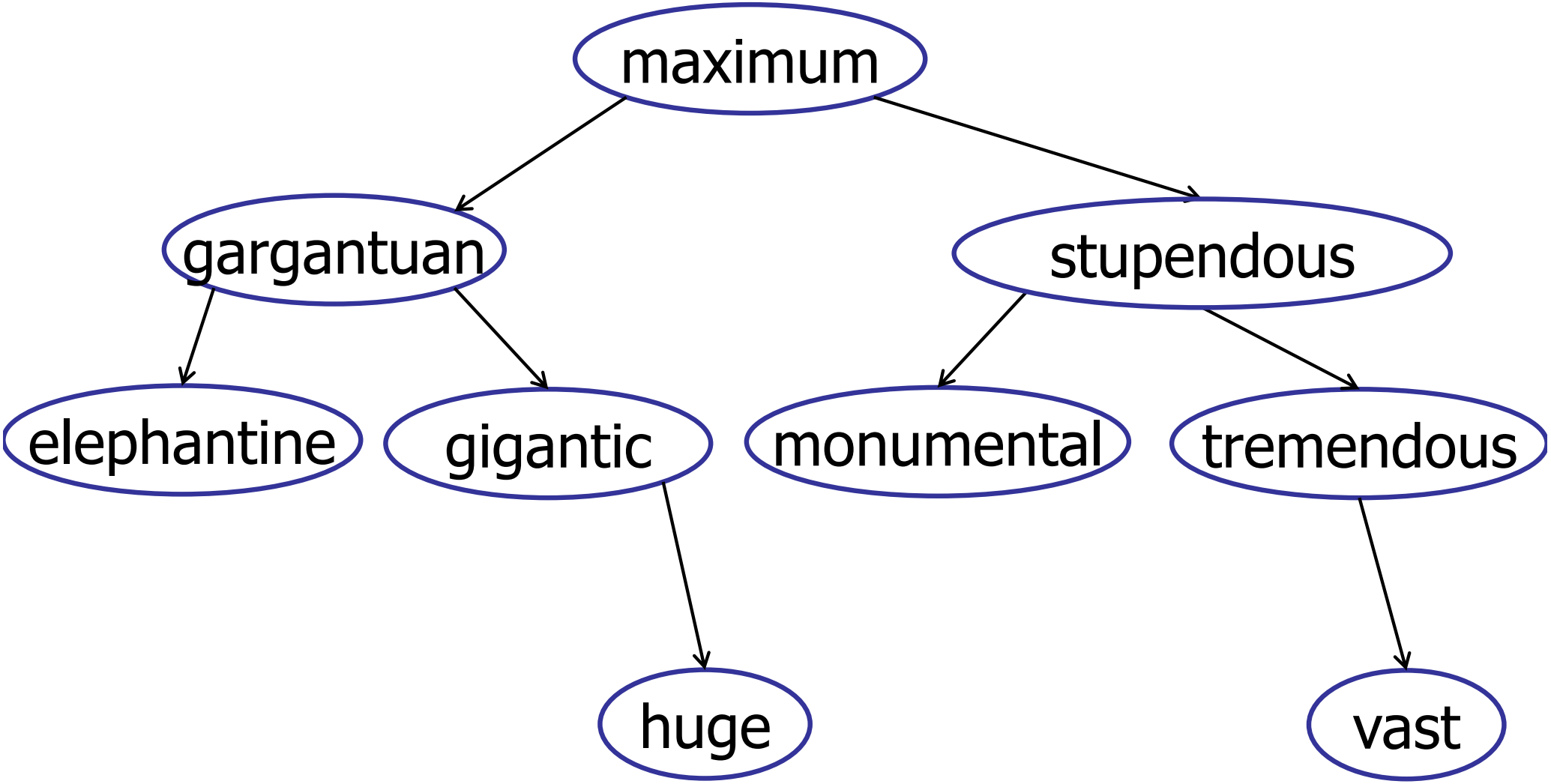
CS2040S
Data Structures and Algorithms
(e-learning edition)

Tries!

What about text strings?



Implement a searchable dictionary!

What about text strings?

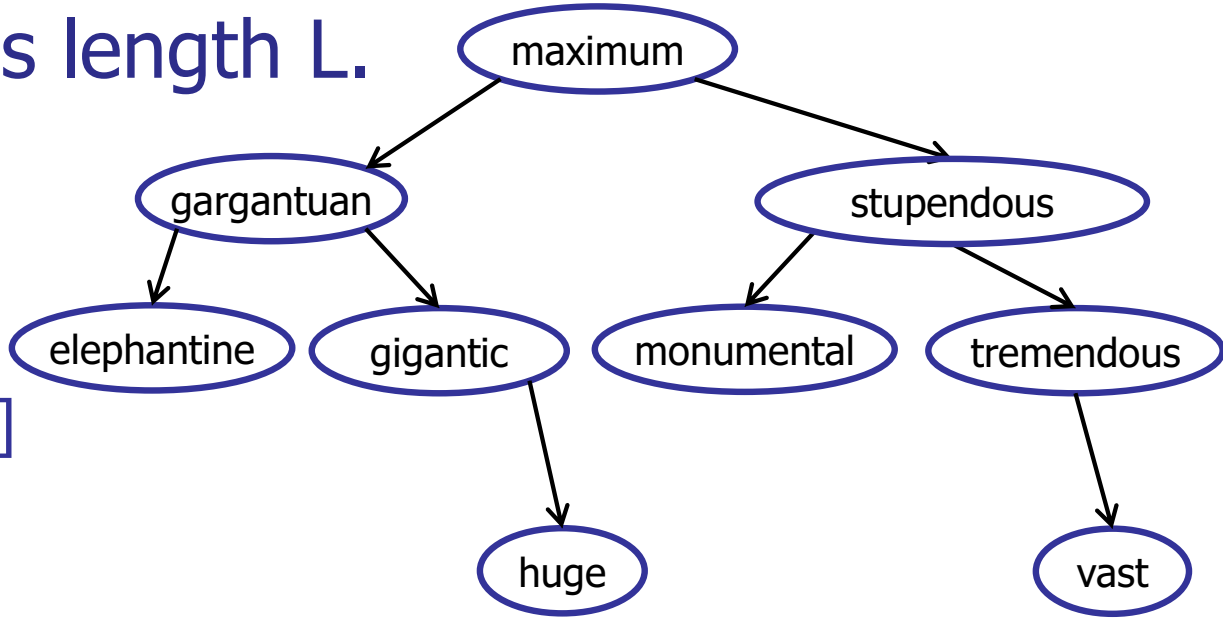
Cost of comparing two strings:

- Cost[A ?? B] = min(A.length, B.length)
- Compare strings letter by letter (?)

Cost of tree operation:

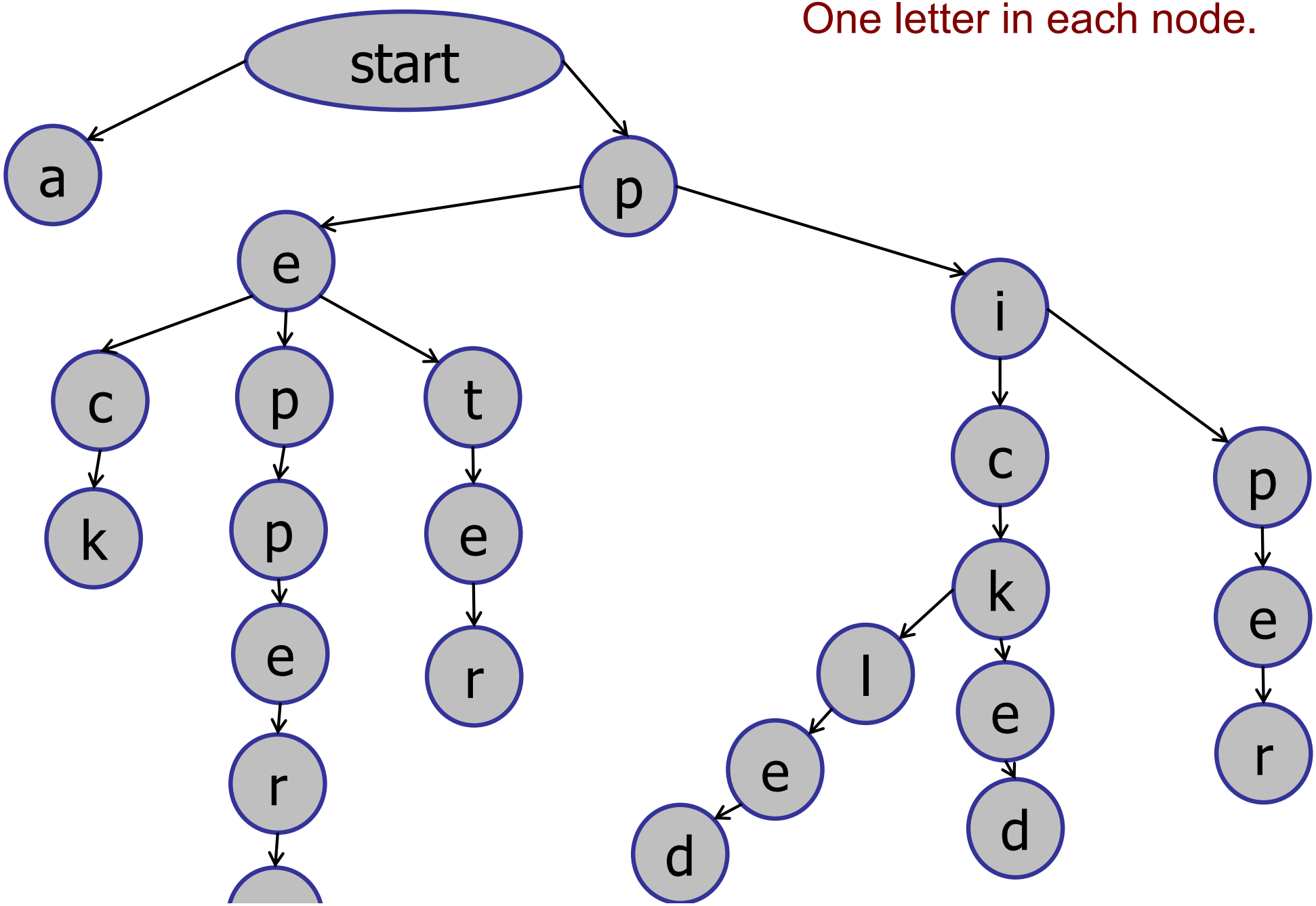
- Assume string has length L.
- Cost: O(hL)

[Optimizations are possible.]



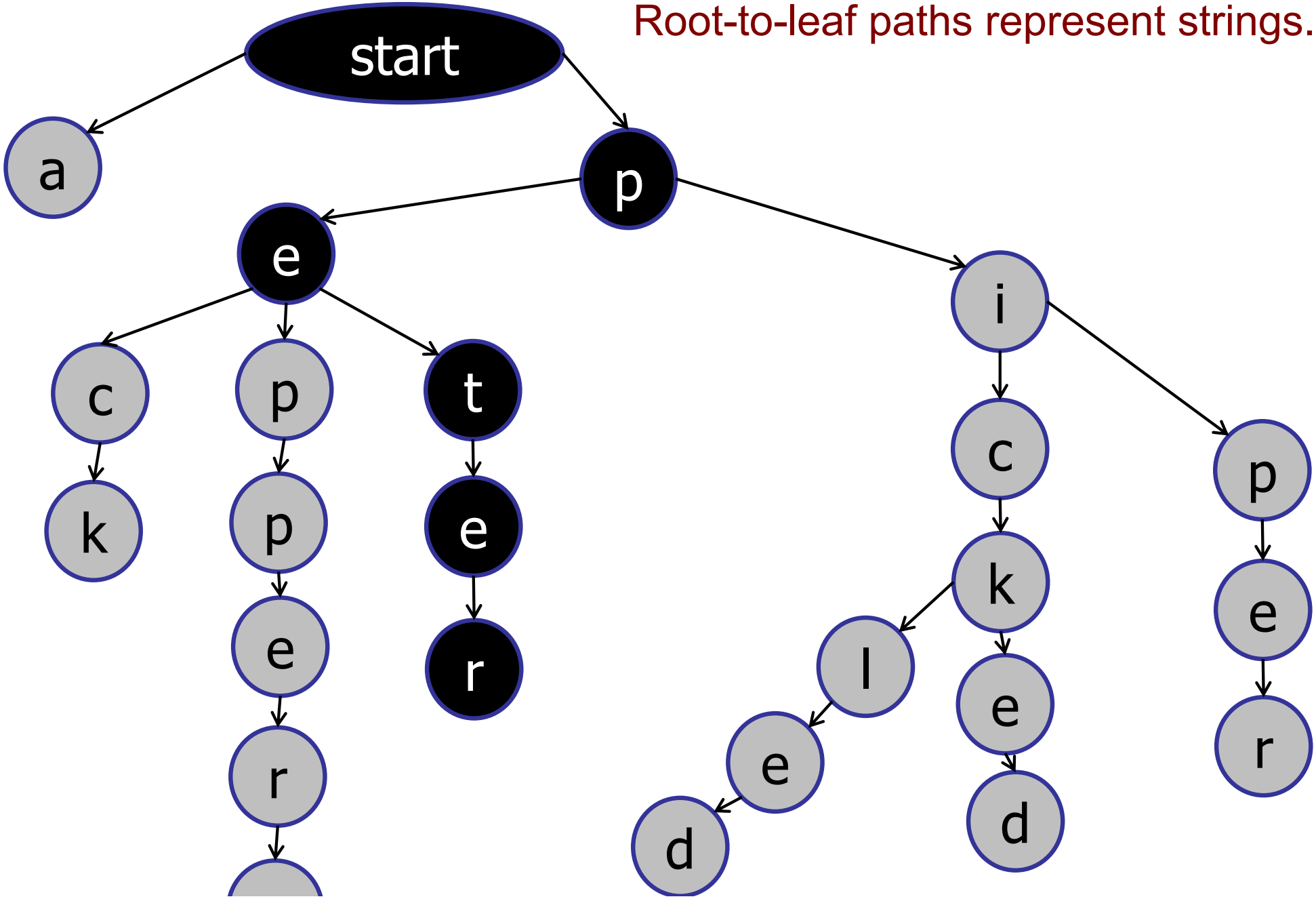
Trie [prounounced: try]

One letter in each node.

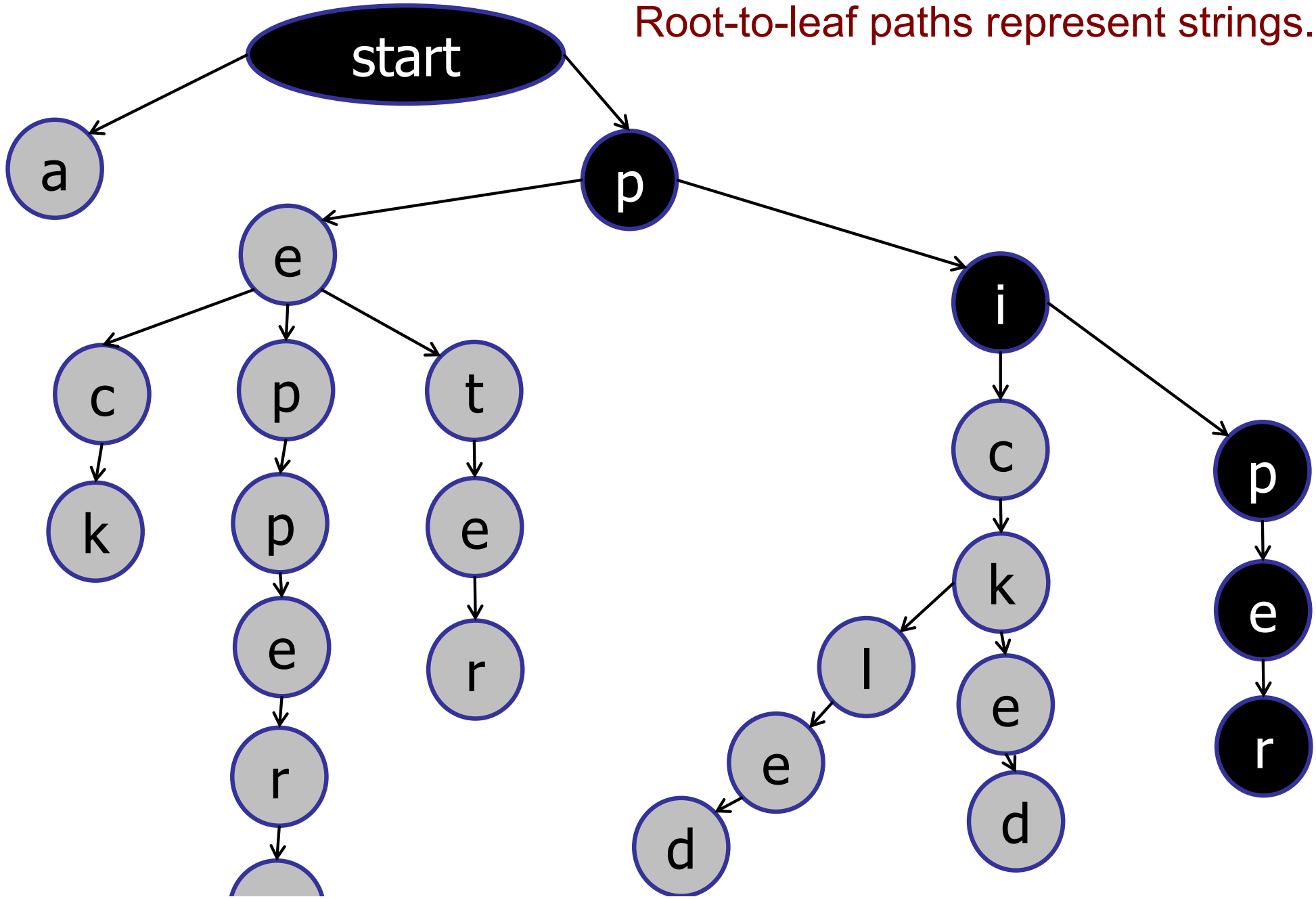


Trie [prounounced: try]

Root-to-leaf paths represent strings.

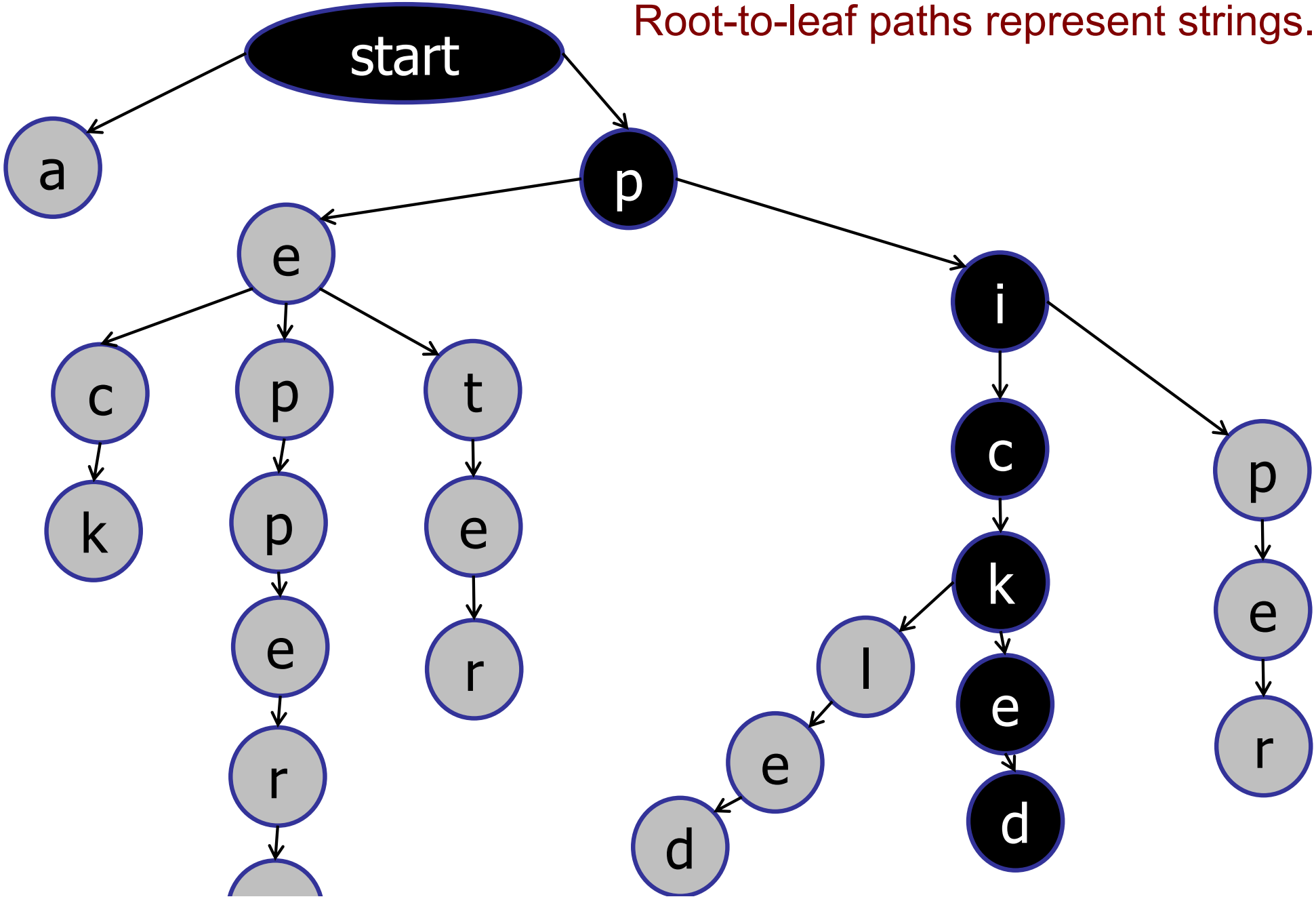


Trie [prounounced: try]



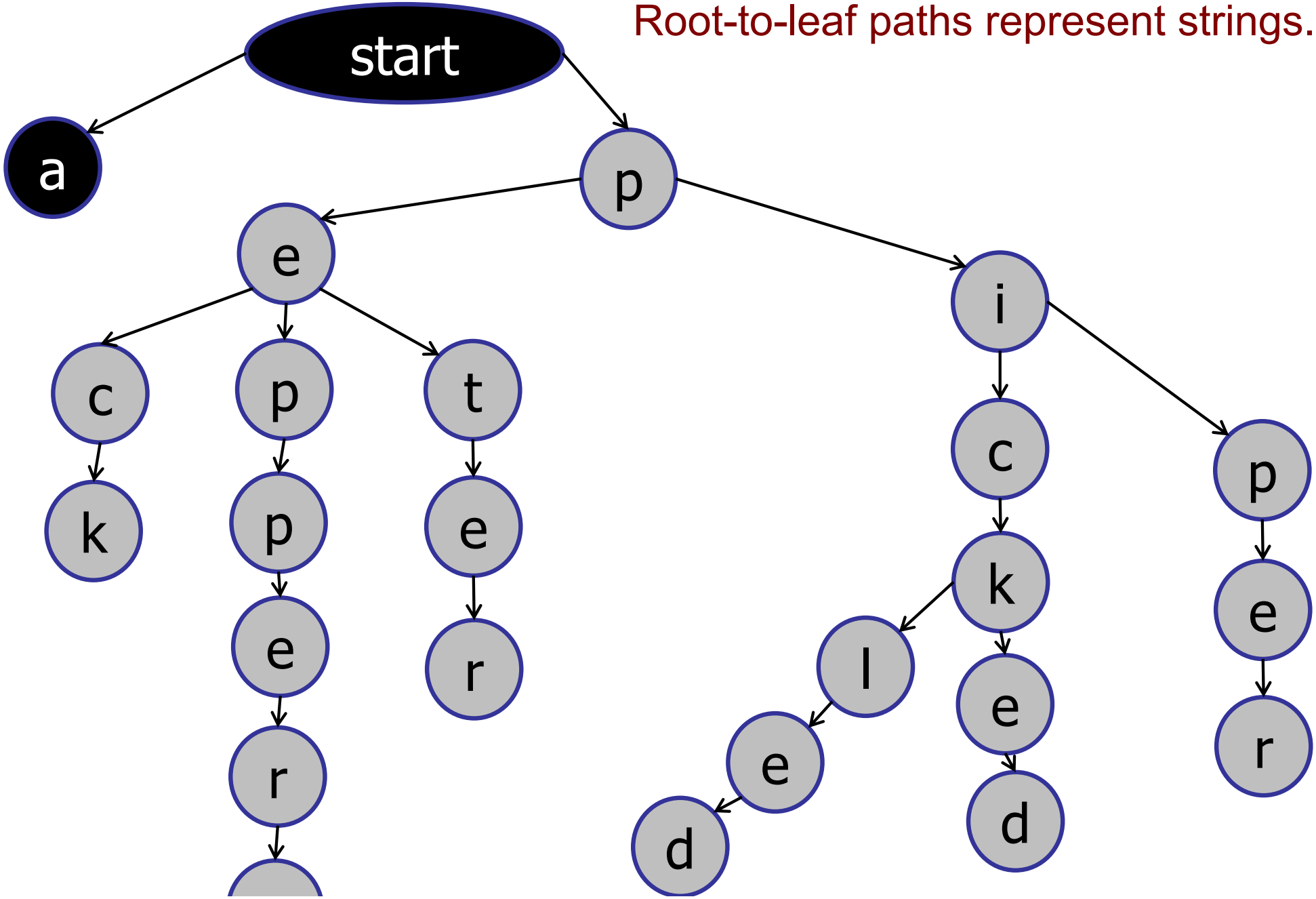
Trie [prounounced: try]

Root-to-leaf paths represent strings.



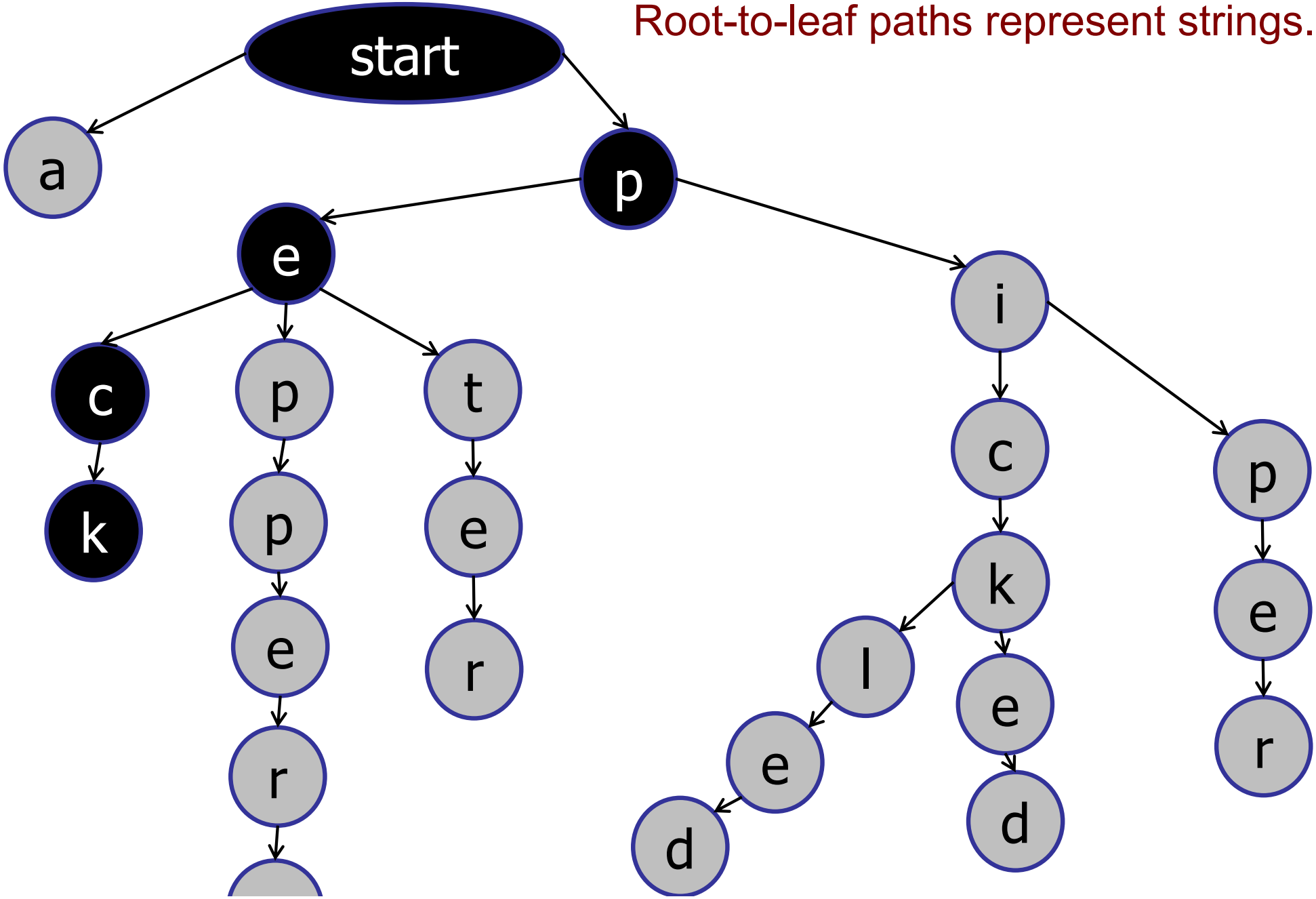
Trie [prounounced: try]

Root-to-leaf paths represent strings.



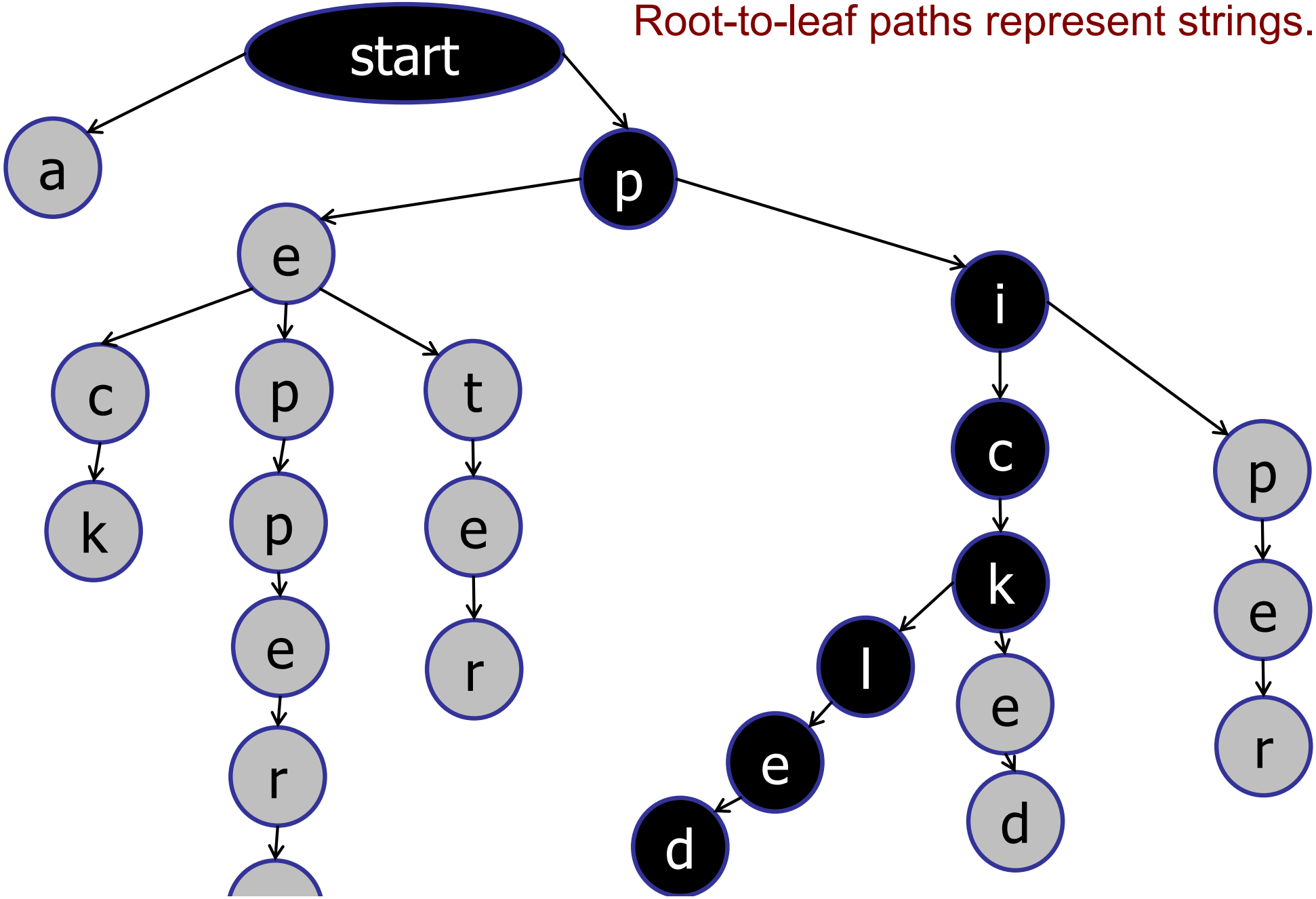
Trie [prounounced: try]

Root-to-leaf paths represent strings.



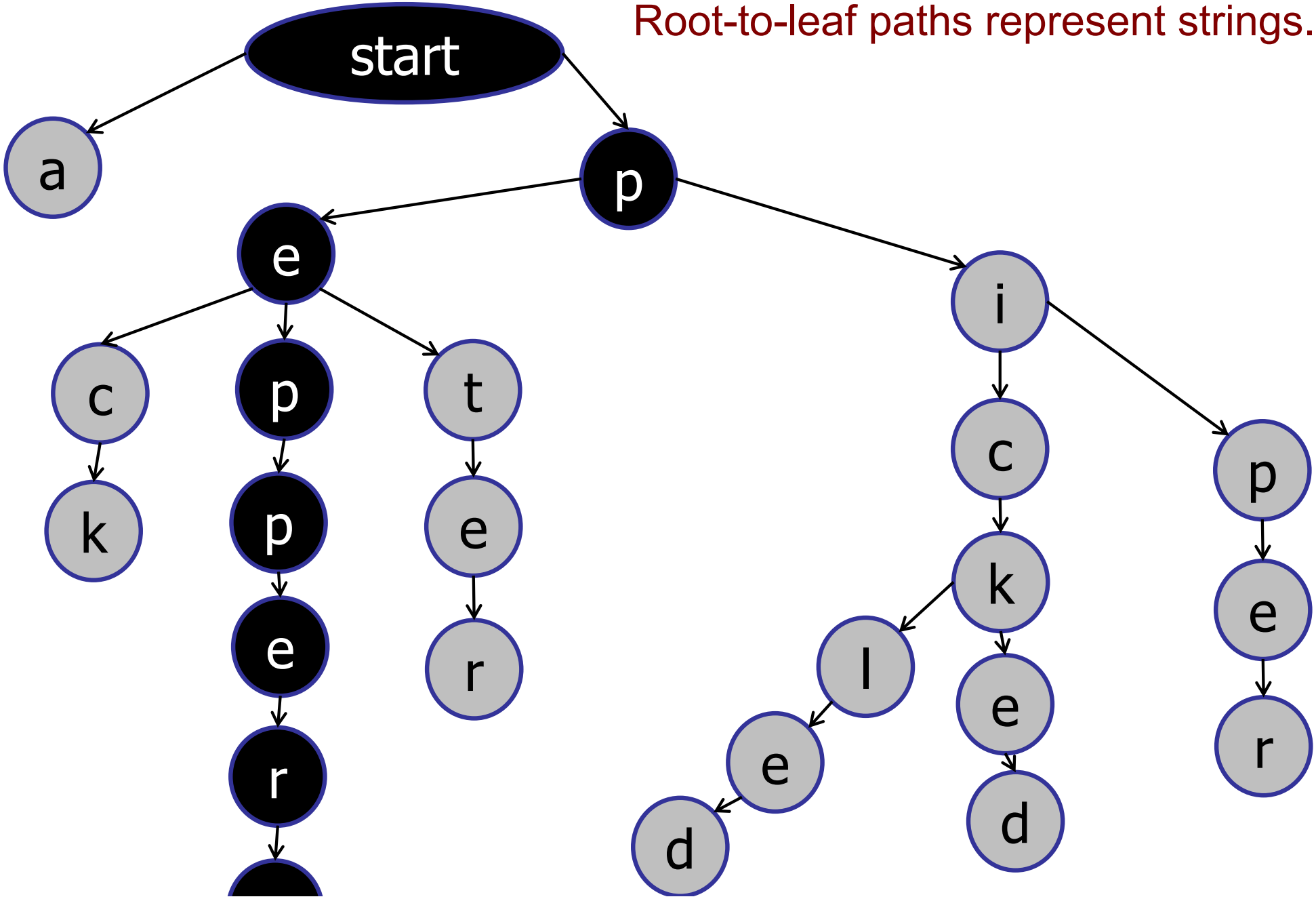
Trie [prounounced: try]

Root-to-leaf paths represent strings.

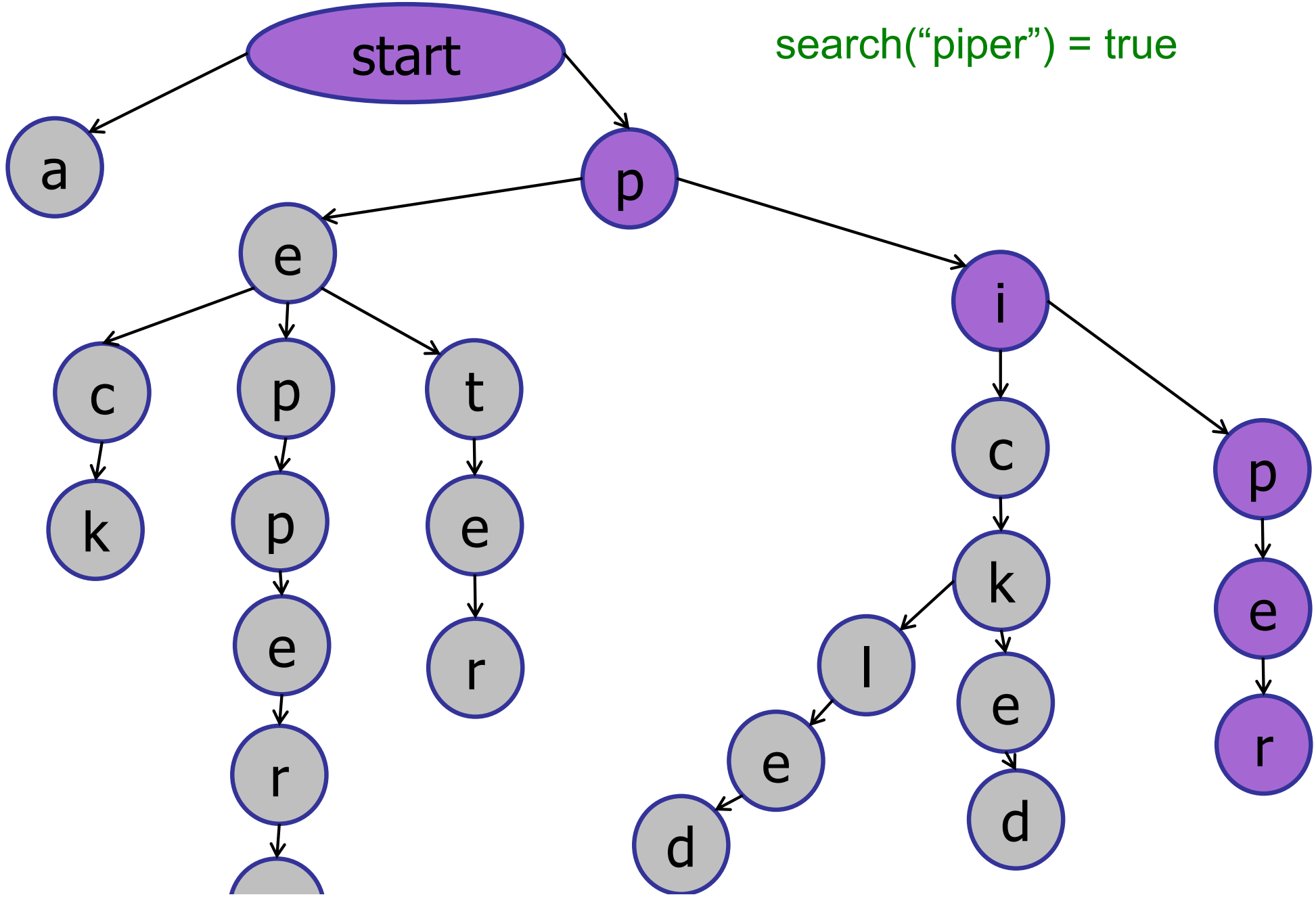


Trie [prounounced: try]

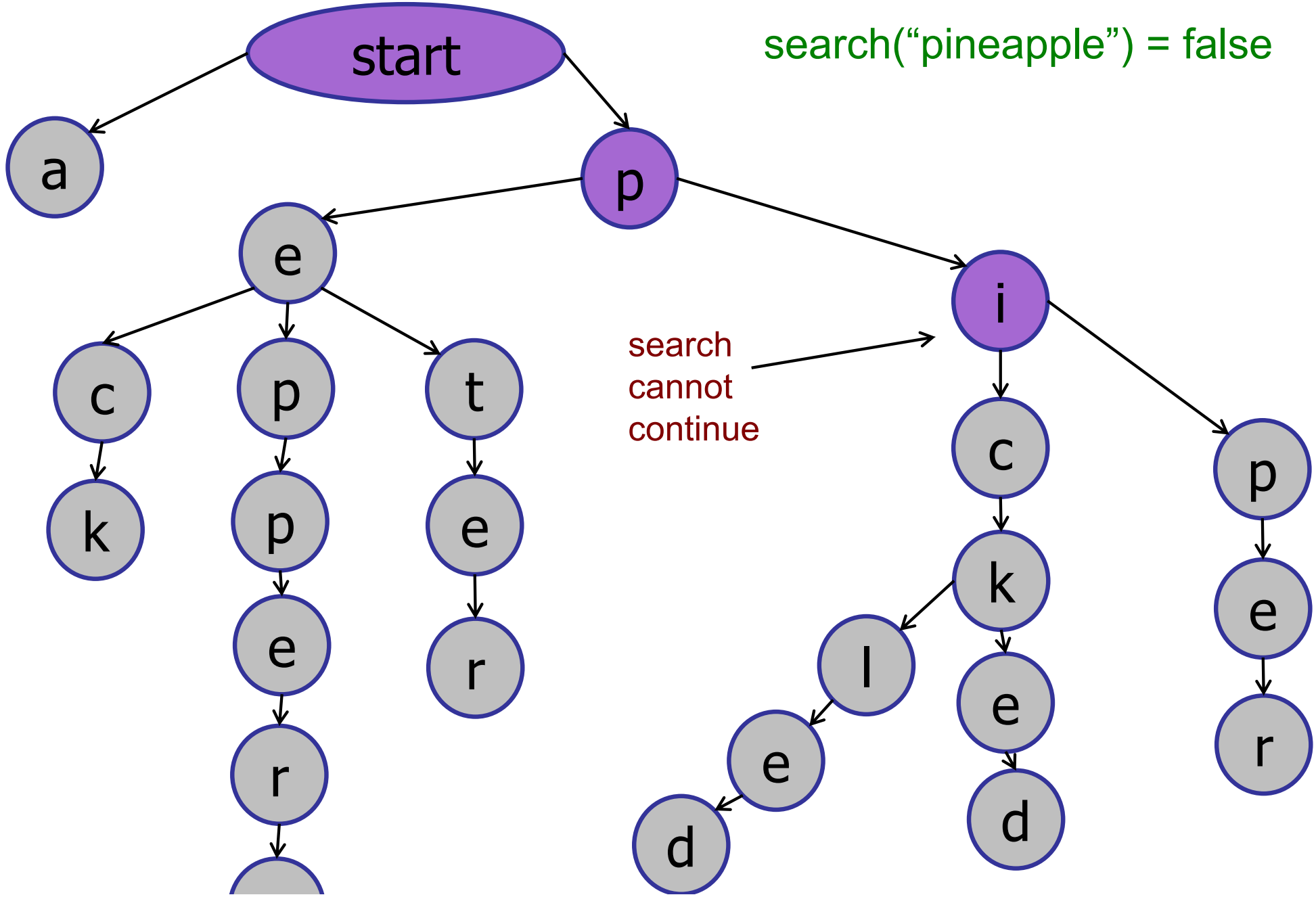
Root-to-leaf paths represent strings.



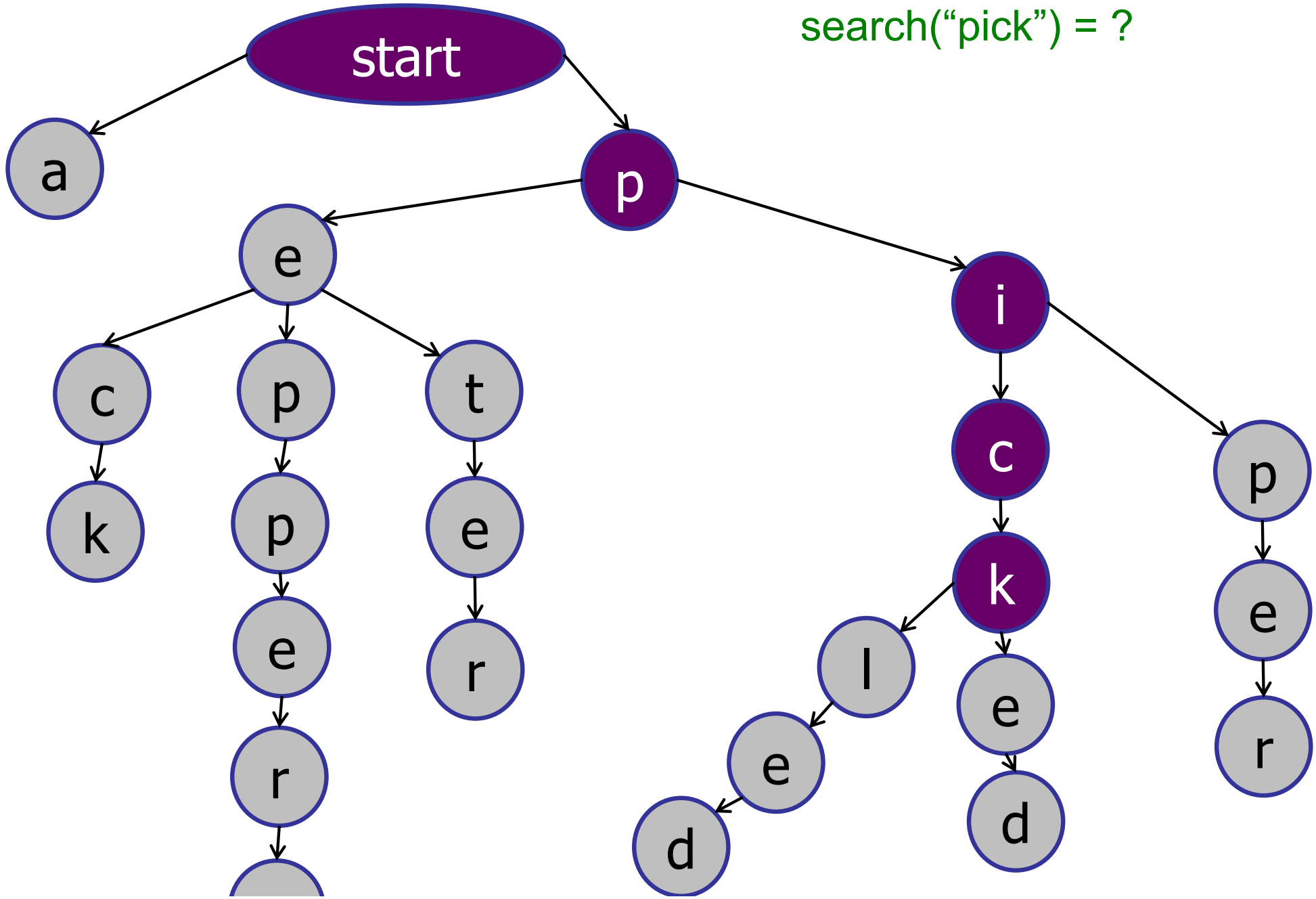
Searching a Trie



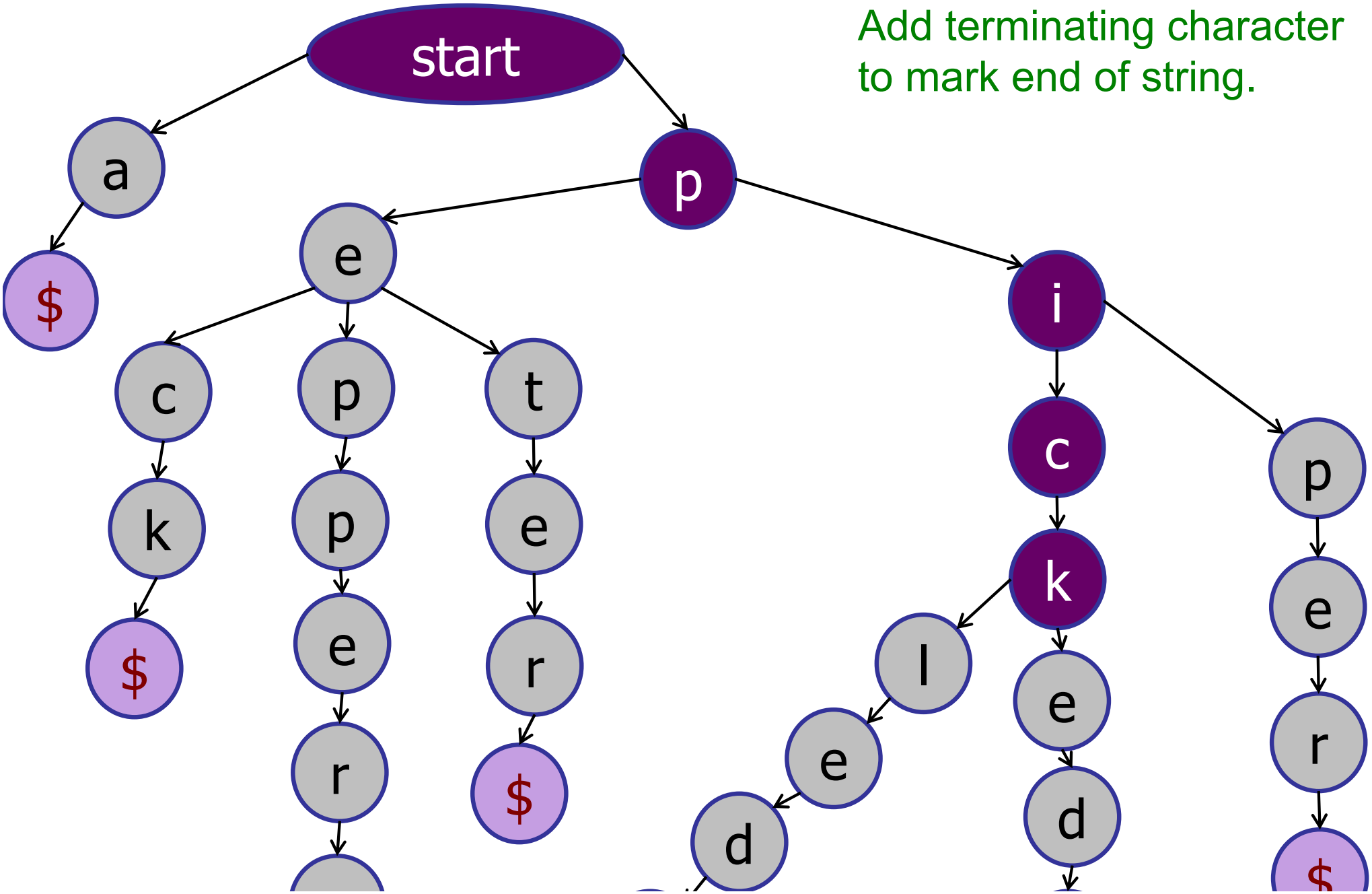
Searching a Trie



Trie Details

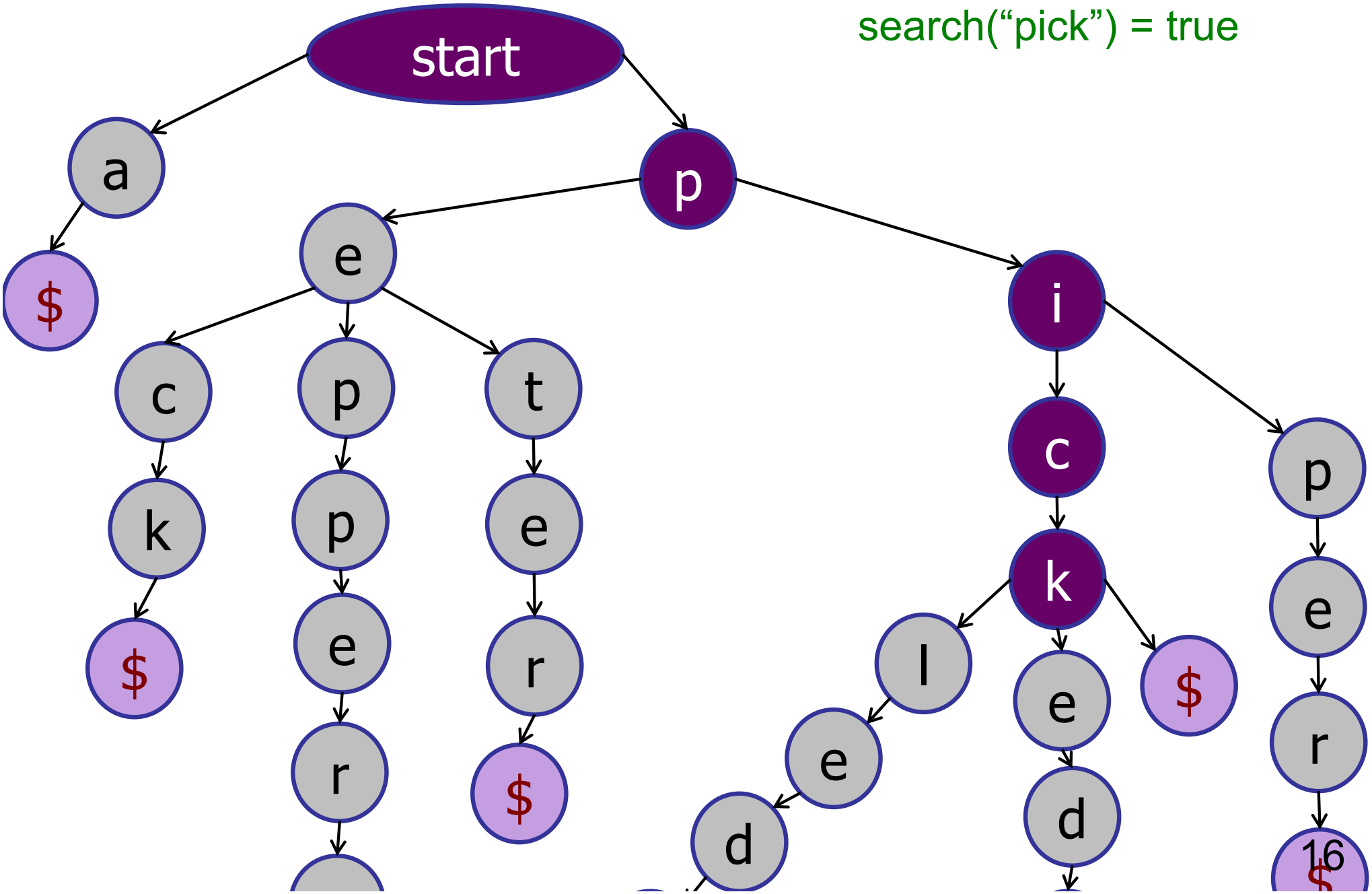


Trie Details



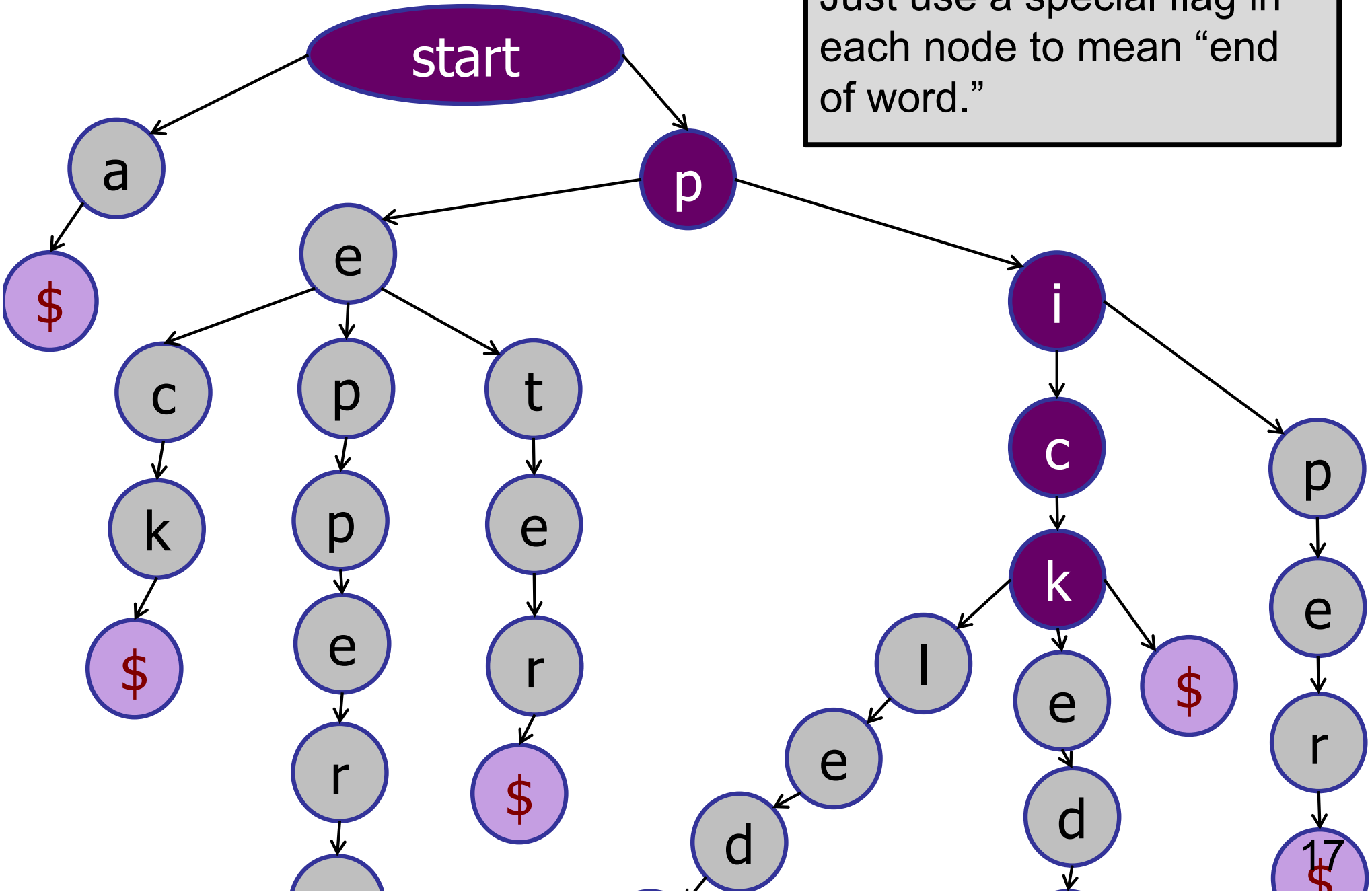
Add terminating character to mark end of string.

Trie Details



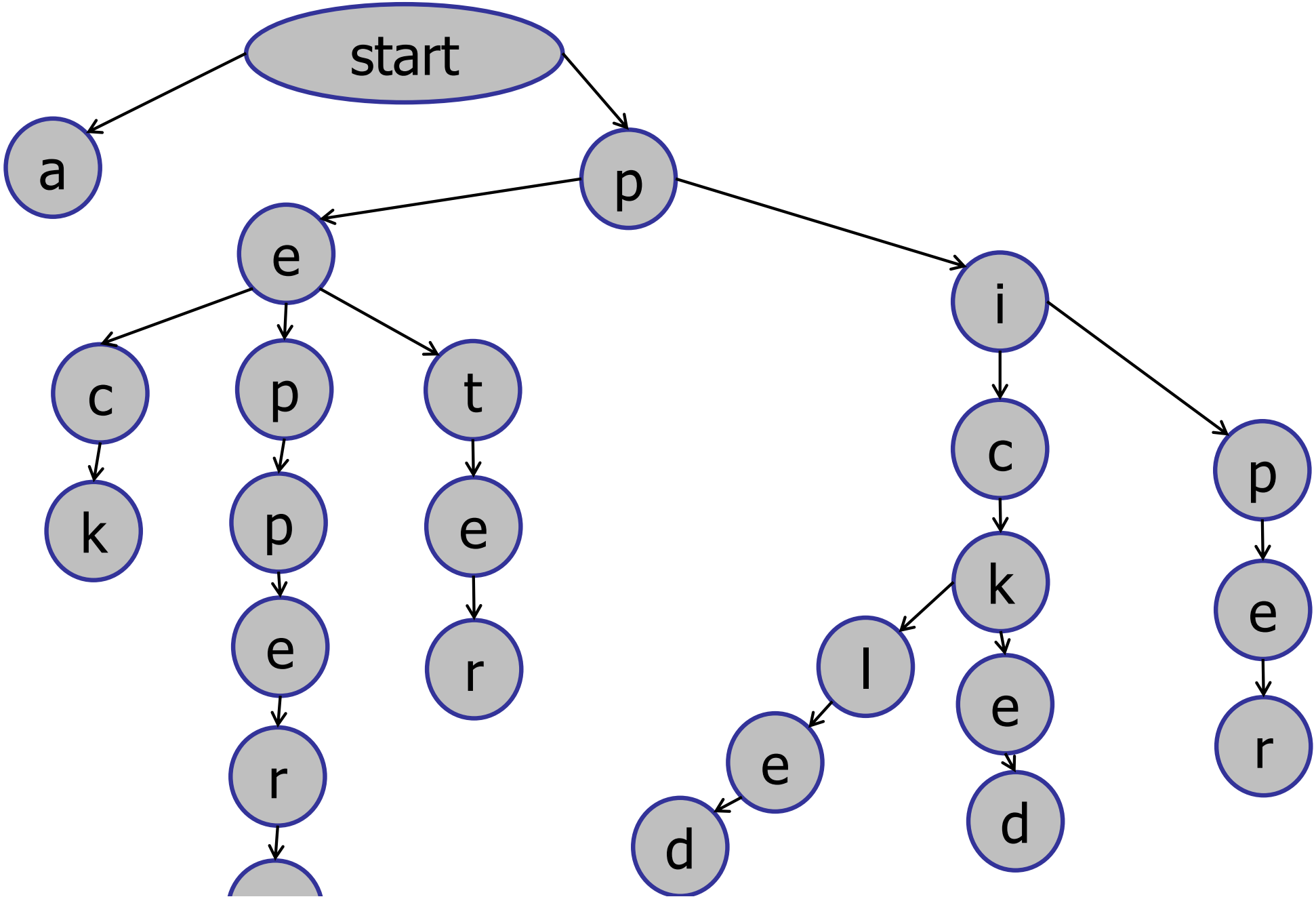
Trie Details

Or:
Just use a special flag in each node to mean “end of word.”



Trie

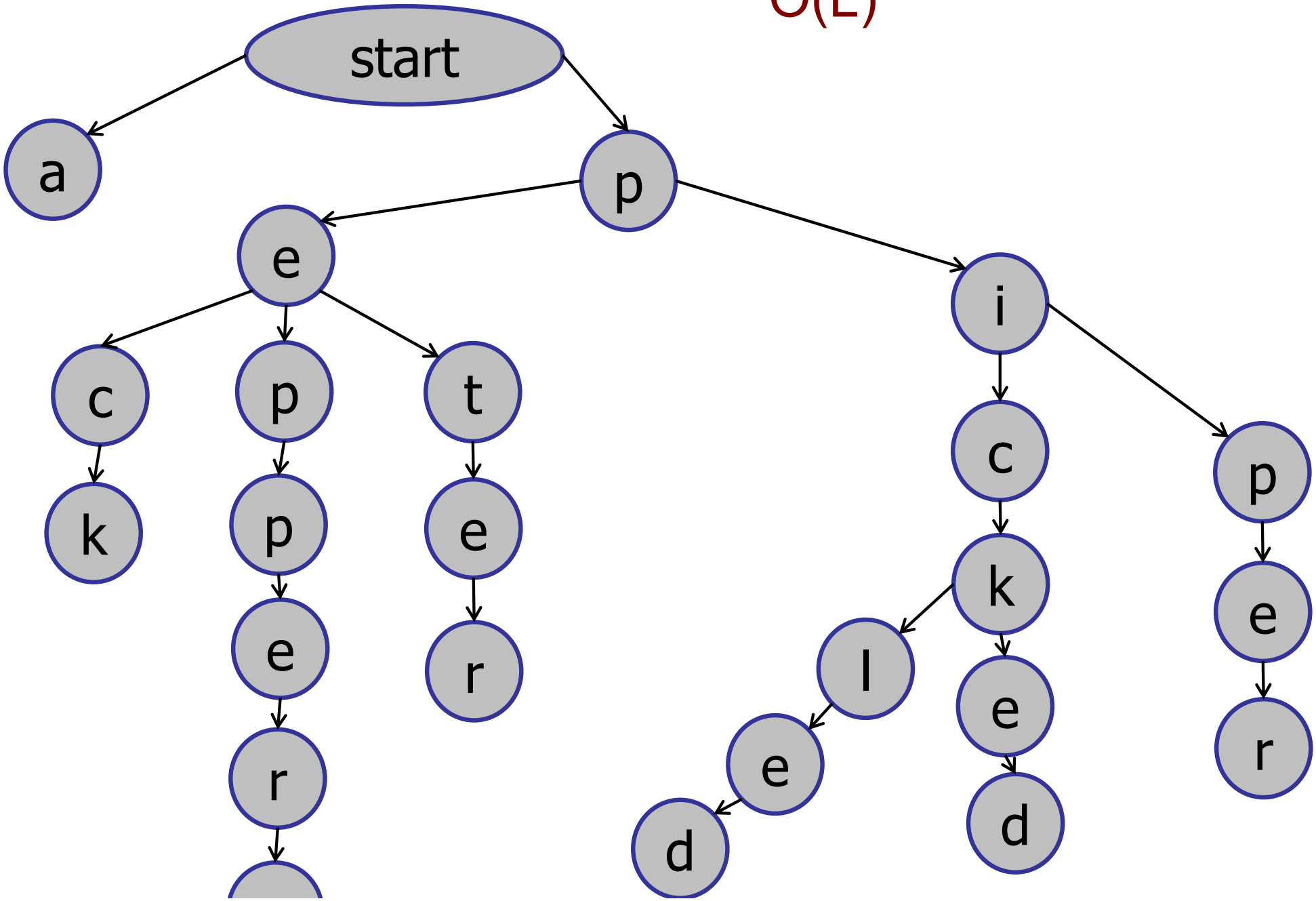
Cost to search for a string of length L?



Trie

Cost to search for a string of length L?

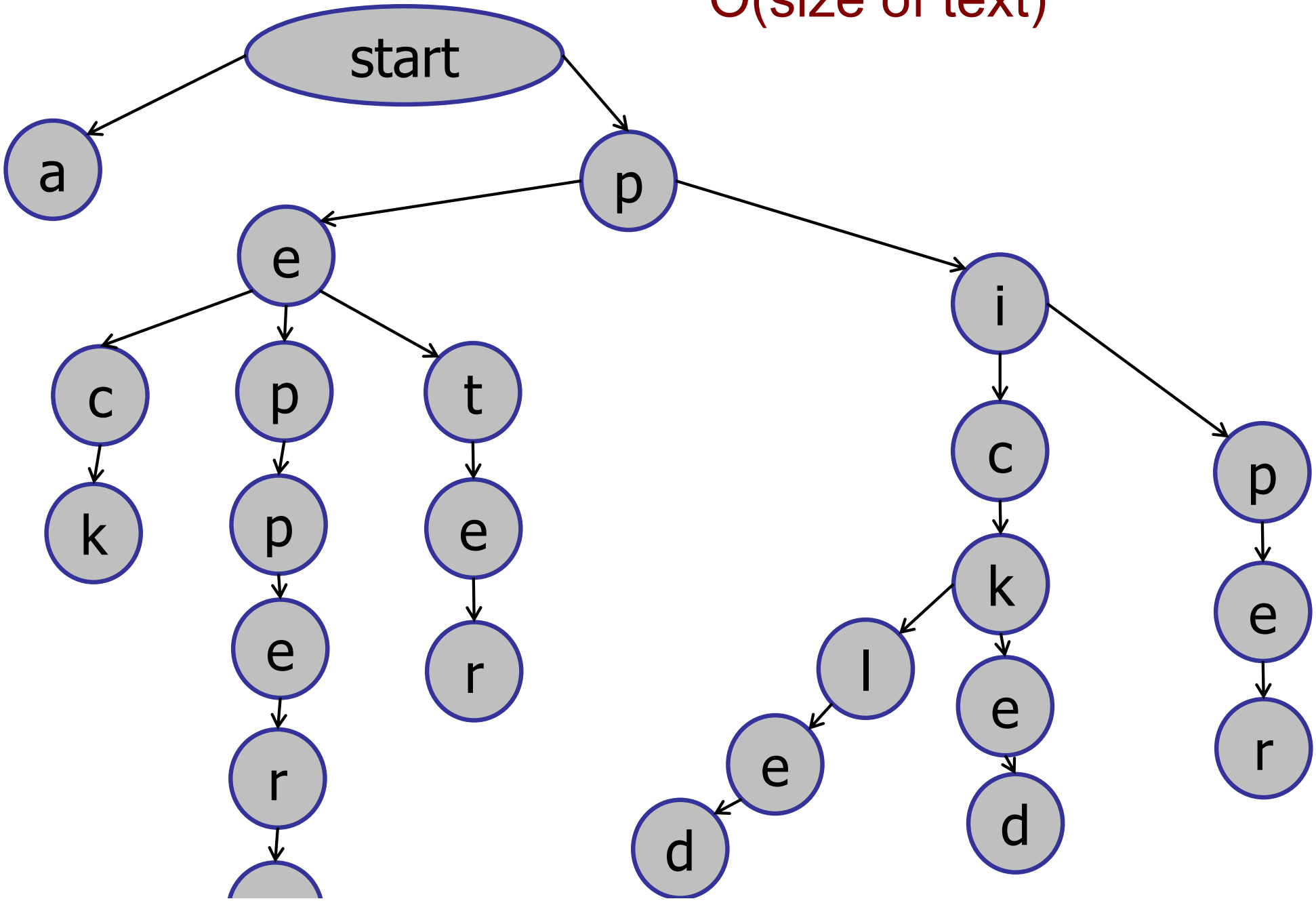
O(L)



Trie

Space for storing a try?

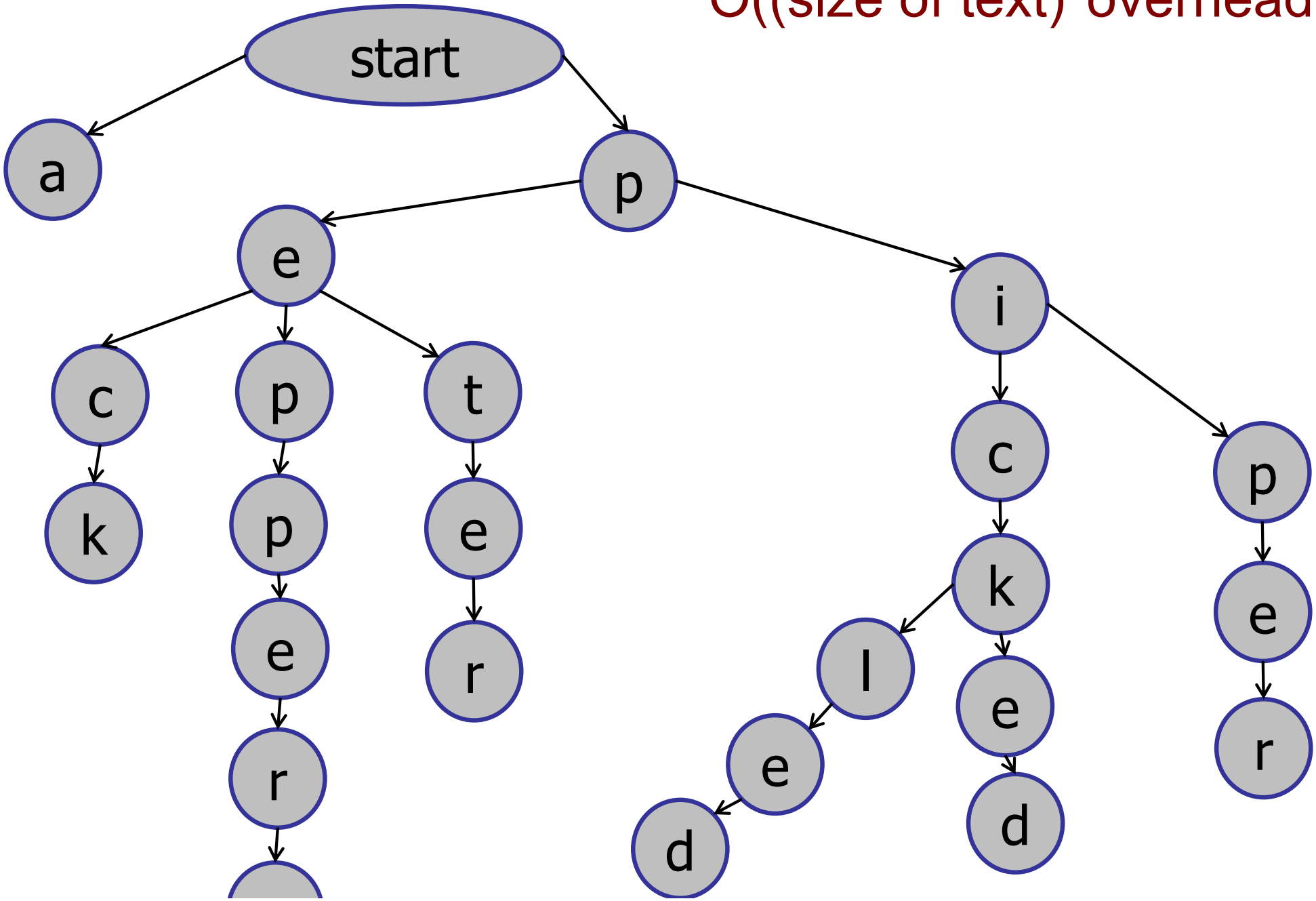
O(size of text)



Trie

Space for storing a try?

O((size of text)*overhead)



Trie Tradeoffs

Time:

- Trie tends to be faster: O(L).
- Does not depend on size of total text.
- Does not depend on number of strings.

Even faster if string is not in trie!

Trie Tradeoffs

Time:

- Trie tends to be faster: O(L).
- Does not depend on size of total text.
- Does not depend on number of strings.

Space:

- Trie tends to use more space.
- BST and Trie use O(text size) space.
- But Trie has more nodes and more overhead.

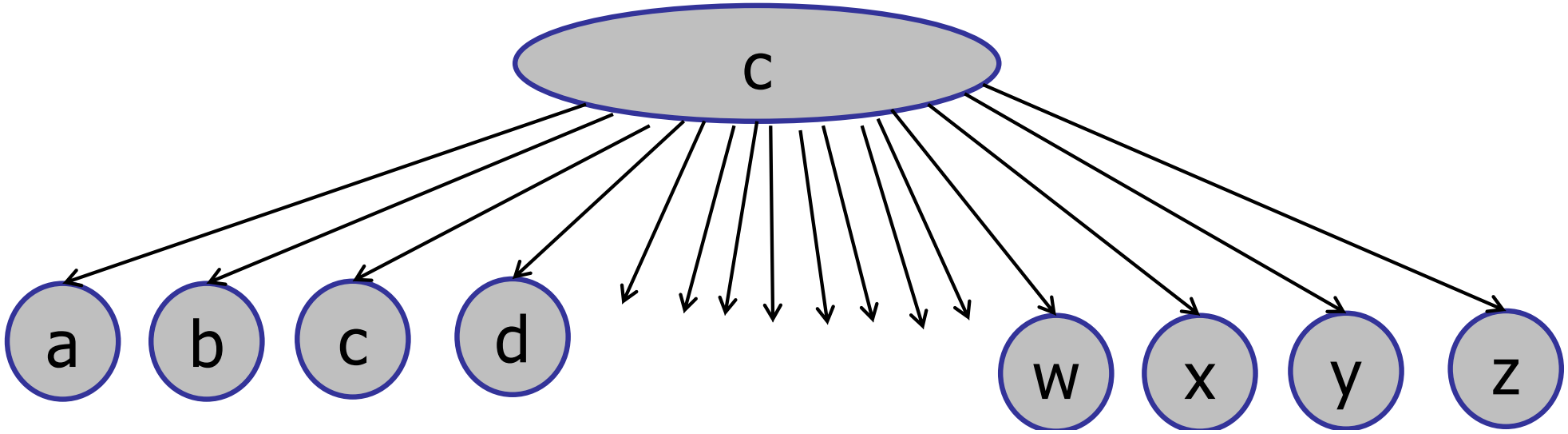
Trie Space

Trie node:

- Has many children.
- For strings: fixed degree.
- Ascii character set: 256

wasted space?

```
TrieNode children[] = new TrieNode[256];
```



Trie Applications

String dictionaries

- Searching
- Sorting / enumerating strings

Partial string operations:

- Prefix queries: find all the strings that start with pi.
- Long prefix: what is the longest prefix of "pickling" in the trie?
- Wildcards: find a string of the form "pi??le" in the trie.