

Word Break

Core Idea (DP):

Define $dp[i]$ = true iff the prefix $s[0...i]$ can be segmented.

Transition:

- For each i from $1...n$, set $dp[i]$ = true if there exists a $j < i$ such that:
 - $dp[j] == \text{true}$
 - $s[j..i]$ is in wordDict.

Answer is $dp[n]$.

How to make it fast:

- Precompute a set from wordDict for $O(1)$ membership.
- Let L_{min} = min length of words, L_{max} = max length of words.
- Only check j where $i-j$ is in $[L_{min}, L_{max}]$. This prunes a ton.
- Break early when you set $dp[i] = \text{true}$.

Complexity:

- Worst case $\sim O(n * (\# \text{ lengths checked})) = O(n * (L_{max} - L_{min} + 1))$, with $O(n)$ space.

Alternative: BFS on indices:

Think of indices as nodes. From index i , you can jump to $i + \text{len}(w)$ for any word w that matches $s[i...i + \text{len}(w)]$. Use a queue and a visited boolean array to avoid revisiting i . Return true if you ever reach n . This often performs similarly to DP.

Small walkthrough:

$S = \text{"applepenapple"}$, $\text{dict} = \{\text{"apple", "pen"}\}$.

$L_{min} = 3$, $L_{max} = 5$

$dp[5]$ becomes true via "apple", then $dp[9]$ via "pen", then $dp[13]$ via "apple".

Final $dp[n] = \text{true}$.

Extra optimization (optional):

- Group dictionary words by length; iterate only those lengths.
- If you like trees, you can traverse a tree from each i and mark $dp[end]$ when a terminal node is hit, but the simple DP is usually enough here.