★ ★ ★ ★ ★ Average Rating: 4.81 (132 votes)

Premium

Overview

This article is for advanced readers. It introduces the following ideas: Depth First Search (DFS), Memoization, Dynamic programming, Topological Sorting. It explains the relation between dynamic programming and topological sorting.

Solution

Approach #1 (Naive DFS) [Time Limit Exceeded]

Intuition

DFS can find the longest increasing path starting from any cell. We can do this for all the cells.

Algorithm

Each cell can be seen as a vertex in a graph G. If two adjacent cells have value $a < b$, i.e. increasing then we have a directed edge (a, b). The problem then becomes:

> Search the longest path in the directed graph G.

Naively, we can use DFS or BFS to visit all the cells connected starting from a root. We update the maximum length of the path during search and find the answer when it finished.

Usually, in DFS or BFS, we can employ a set `visited` to prevent the cells from duplicate visits. We will introduce a better algorithm based on this in the next section.

Java

Copy

```java
// Naive DFS Solution
// Time Limit Exceeded
public class Solution {
  private static final int[][] dirs = {{0, 1}, {1, 0}, {0, -1}, {-1, 0}};
  private int m, n;

  public int longestIncreasingPath(int[][] matrix) {
      if (matrix.length == 0) return 0;
      m = matrix.length;
      n = matrix[0].length;
      int ans = 0;
      for (int i = 0; i < m; ++i)
          for (int j = 0; j < n; ++j)
              ans = Math.max(ans, dfs(matrix, i, j));
      return ans;
  }

  private int dfs(int[][] matrix, int i, int j) {
      int ans = 0;
      for (int[] d : dirs) {
          int x = i + d[0], y = j + d[1];
          if (0 <= x && x < m && 0 <= y && y < n && matrix[x][y] > matrix[i][j])
              ans = Math.max(ans, dfs(matrix, x, y));
      }
      return ++ans;
  }
}
```

Complexity Analysis

- Time complexity : $O(2^{m+n})$. The search is repeated for each valid increasing path. In the worst case we can have $O(2^{m+n})$ calls.
- Space complexity : $O(mn)$. For each DFS we need $O(h)$ space used by the system stack, where h is the maximum depth of the recursion. In the worst case, $O(h) = O(mn)$.

Approach #2 (DFS + Memoization) [Accepted]

Intuition

Cache the results for the recursion so that any subproblem will be calculated only once.

Algorithm

From previous analysis, we know that there are many duplicate calculations in the naive approach.

One optimization is that we can use a set to prevent the repeat visit in one DFS search. This optimization will reduce the time complexity for each DFS to $O(mn)$ and the total algorithm to $O(m^2n^2)$.

Here, we will introduce more powerful optimization, Memoization.

> In computing, memoization is an optimization technique used primarily to speed up computer programs by storing the results of expensive function calls and returning the cached result when the same inputs occur again.

In our problem, we recursively call `dfs(x, y)` for many times. But if we already know all the results for the four adjacent cells, we only need constant time. During our search if the result for a cell is not calculated, we calculate and cache it; otherwise, we get it from the cache directly.

Java

Copy

```java
// DFS + Memoization Solution
// Accepted and Recommended
public class Solution {
```

Testcase Run Code Result Debugger

Accepted Runtime: 5 ms

Your input `[[9,9,4],[6,6,8],[2,1,1]]`

Output 4 Diff

Expected 4

Console Use Example Testcases

```
    private static final int[][] dirs = {{0, 1}, {1, 0}, {0, -1}, {-1, 0}};
    private int m, n;

    public int longestIncreasingPath(int[][] matrix) {
        if (matrix.length == 0) return 0;
        m = matrix.length; n = matrix[0].length;
        int[][] cache = new int[m][n];
        int ans = 0;
        for (int i = 0; i < m; ++i)
            for (int j = 0; j < n; ++j)
                ans = Math.max(ans, dfs(matrix, i, j, cache));
        return ans;
    }

    private int dfs(int[][] matrix, int i, int j, int[][] cache) {
        if (cache[i][j] != 0) return cache[i][j];
        for (int[] d : dirs) {
            int x = i + d[0], y = j + d[1];
            if (0 <= x && x < m && 0 <= y && y < n && matrix[x][y] > matrix[i][j])
                cache[i][j] = Math.max(cache[i][j], dfs(matrix, x, y, cache));
        }
        return ++cache[i][j];
    }
}
```

Complexity Analysis

- Time complexity : $O(mn)$. Each vertex/cell will be calculated once and only once, and each edge will be visited once and only once. The total time complexity is then $O(V+E)$. V is the total number of vertices and E is the total number of edges. In our problem, $O(V) = O(mn)$, $O(E) = O(4V) = O(mn)$.
- Space complexity : $O(mn)$. The cache dominates the space complexity.

Approach #3 (Peeling Onion) [Accepted]

Intuition

The result of each cell only related to the result of its neighbors. Can we use dynamic programming?

Algorithm

If we define the longest increasing path starting from cell (i, j) as a function

$$f(i, j)$$

then we have the following transition function

$$f(i,j) = max\{f(x,y) | (x,y) \text{ is a neighbor of} (i,j) \text{ and matrix}[x][y] > \text{matrix}[i][j]\} + 1$$

This formula is the same as used in the previous approaches. With such transition function, one may think that it is possible to use dynamic programming to deduce all the results without employing DFS!

That is right with one thing missing: we don't have the dependency list.

For dynamic programming to work, if problem B depends on the result of problem A, then we must make sure that problem A is calculated before problem B. Such order is natural and obvious for many problems. For example the famous Fibonacci sequence:

$$F(0) = 1, F(1) = 1, F(n) = F(n-1) + F(n-2)$$

The subproblem $F(n)$ depends on its two predecessors. Therefore, the natural order from 0 to n is the correct order. The dependent is always behind the dependee.

The terminology of such dependency order is "Topological order" or "Topological sorting":

> Topological sorting for Directed Acyclic Graph (DAG) is a linear ordering of vertices such that for every directed edge (u, v), vertex u comes before v in the ordering.

In our problem, the topological order is not natural. Without the value in the matrix, we couldn't know the dependency relation of any two neighbors A and B. We have to perform the topological sort explicitly as a preprocess. After that, we can solve the problem dynamically using our transition function following the stored topological order.

There are several ways to perform the topological sorting. Here we employ one of them called "Peeling Onion".

The idea is that in a DAG, we will have some vertex who doesn't depend on others which we call "leaves". We put these leaves in a list (their internal ordering does matter), and then we remove them from the DAG. After the removal, there will be new leaves. We do the same repeatedly as if we are peeling an onion layer by layer. In the end, the list will have a valid topological ordering of our vertices.

In out problem, since we want the longest path in the DAG, which equals to the total number of layers of the "onion". Thus, we can count the number of layers during "peeling" and return the counts in the end without invoking dynamic programming.

Java

Copy

```
// Topological Sort Based Solution
// An Alternative Solution
public class Solution {
    private static final int[][] dir = {{0, 1}, {1, 0}, {0, -1}, {-1, 0}};
    private int m, n;
    public int longestIncreasingPath(int[][] grid) {
        int m = grid.length;
        if (m == 0) return 0;
        int n = grid[0].length;
        // padding the matrix with zero as boundaries
        // assuming all positive integer, otherwise use INT_MIN as boundaries
        int[][] matrix = new int[m + 2][n + 2];
        for (int i = 0; i < m; ++i)
            System.arraycopy(grid[i], 0, matrix[i + 1], 1, n);

        // calculate outdegrees
        int[][] outdegree = new int[m + 2][n + 2];
        for (int i = 1; i <= m; ++i)
            for (int j = 1; j <= n; ++j)
                for (int[] d: dir)
                    if (matrix[i][j] < matrix[i + d[0]][j + d[1]])
                        outdegree[i][j]++;

        // find leaves who have zero out degree as the initial level
        n += 2;
        m += 2;
        List<int[]> leaves = new ArrayList<>();
```

Complexity Analysis

- Time complexity : $O(mn)$. The the topological sort is $O(V+E) = O(mn)$. Here, V is the total number of vertices and E is the total number of edges. In our problem, $O(V) = O(mn), O(E) = O(4V) = O(mn)$.
- Space complexity : $O(mn)$. We need to store the out degrees and each level of leaves.

Remarks

- Memoization: for a problem with massive duplicate calls, cache the results.
- Dynamic programming requires the subproblem solved in topological order. In many problems, it coincides the natural order. For those who doesn't, one need perform topological sorting first. Therefore, for those problems with complex topology (like this one), search with memorization is usually an easier and better choice.

Report Article Issue

Comments: 67 | Best | Most Votes | Newest to Oldest | Oldest to Newest

Type comment here... (Markdown is supported)

Post

kiljaeden ★ 514 March 21, 2021 1:42 AM

I think it is worthy to mention that for most of this kind of questions that we could not add memorization upon a DFS. This question is a special case. Normally when you could move to 4 directions, there would be cycle so you could not memorize the result. However since this question is strictly increasing, thus it is a DAG.

149 Show 8 replies Reply

jguodev ★ 102 March 14, 2018 10:45 PM

Who can help explain why the Time complexity of Approach 1 is O(2^(m+n))?
Thanks.

74 Show 26 replies Reply

ghostfacechillah ★ 110 Last Edit: August 21, 2018 1:31 PM

"Usually, in DFS or BFS, we can employ a set visited to prevent the cells from duplicate visits. We will introduce a better algorithm based on this in the next section." but I did not find which part in the article explains why we don't have to maintain such a visited set. My guess is because the path is increasing, we will never visit a node with smaller value.

34 Show 6 replies Reply

jinchuan ★ 37 December 26, 2016 1:32 PM

the time complexity of the second solution should be O(mn) in total, isn't it?

25 Show 1 reply Reply

haotianliu1801 ★ 10 Last Edit: October 24, 2018 12:53 AM

Why do we have to increment ans by 1 at the end of DFS in solution #1?

6 Show 2 replies Reply

Kushina ★ 46 August 24, 2019 12:00 PM

Appreciate the `Remarks` summarizing the insights/take-aways from the solution. This should be made a more common practice, if possible. Many thanks!

4 Reply

himani_01 ★ 132 February 9, 2021 7:48 AM

According to me the complexity for approach 1 should be (4^mn)
How are we reaching to m+n factor here .. Since we can move in any of the 4 direction. I feel the longest path could be as long as m*n.

5 Show 3 replies Reply

harleyquinn ★ 5 Last Edit: September 22, 2018 2:09 PM

can you give more examples of questions where onion peeling is used for solving the question?

3 Show 3 replies Reply

hanzhoutang ★ 315 November 6, 2018 7:56 PM

It can also be a dp problem.
The idea is that we first update the dp matrix from top to down, left to right and then update it from bottom to top, right to left. We continue this procedure until dp matrix doesn't update. Here is my code:

```
int longestIncreasingPath(vector<vector<int>>& matrix) {
    if(matrix.empty()) return 0;
    if(matrix[0].empty()) return 0;
    vector<vector<int>> dp(matrix.size(),vector<int>(matrix[0].size(),1));
    int ret = 0;
    int oldvalue = 0;
    int updating = false;
```

Read More

5 Show 2 replies Reply

losercoder ★ 15 August 29, 2020 1:21 PM

In approach 1, for each vertex, we are doing work equivalent to O(2^(mn)) since the longest snake path consists of mn vertices. Now without memoization, this will be repeated across all mn *vertices. So isn't the time complexity of approach 1 O(mn*2^(mn)) ??

3 Show 1 reply Reply