Question 1: Longest common substring

Longest common substring

Given an two strings A and B of lengths m and n, respecitve. A common substring is a string that appears as consecutive letters in both A and B. The goal is to find a common substring of A and B with maximal length.

For example if A = ababcab and B = abcb then the longest common substring is **abc**.

We want to design algorithms to calculate the length of the longest common substring.

Brute force algorithm

A brute force algorithm loops over all possible substrings of A and all possible substrings of B and checks if they are equal. It returns the length of the longest of these strings that are in both A and B.

```
MaxLen = 0
ForAll substring S of A
    ForAll substring T of B
        If S = T
            MaxLen = max(MaxLen, length of S)
Return MaxLen
```

What is the asymptotic runtime for this brute for algorithm where A and B have at most n characters

$O($ 2^n $)$ 25%

Dynamic programming formulation

To solve this problem using dynamic programming we need to formulate a function for solving subproblems of the input and recursively solve more complicated problems. For this problem we will need to define a function $LCS(i, j)$, where i is an integer.

Which of the following descriptions of LCS are correct

- $LCS(i, j)$ is length of the longest common string of $A[1..m]$ and $A[1..n]$.
- $LCS(i, j)$ is the length of the longest common subsequence of $A[1..i]$ and $B[1..j]$ with last characters $A[i]$ and $A[j]$. ✓
- $LCS(i, j)$ is number of characters in both $A[1..i]$ and $B[1..j]$.
- $LCS(i, j)$ is the length of the longer string $A[1..i]$ and $B[1..j]$.
- $LCS(i, j)$ is the length of the longest common string of $A[1..i]$ and $B[1..j]$.

✓ 100%

In defining $LCS(i)$ recursively, which of the following cases do we need to include? In these definition $A = a_1a_2 \ldots a_n$ and $B = b_1b_2 \ldots b_n$. Select all that should be part of the definition.

- [x] $LCS(i, j) = LCS(i-1, j-1) + 1$ if $i > 0, j > 0$ and $A[i] = B[j]$ ✓
- [ ] $LCS(i, j) = i$ if $i = j$
- [ ] $LCS(i, j) = max\{LCS(i-1, j) + 1, LCS(i, j-1) + 1\}$ if $i > 0, j > 0$ and $A[i] \neq B[j]$
- [ ] $LCS(i, j) = LCS(i-1, j-1)$ if $i > 0, j > 0$ and $A[i] \neq B[j]$
- [x] $LCS(i, j) = 0$ if $i = 0$ ✓
- [x] $LCS(i, j) = 0$ if $j = 0$ ✓
- [x] $LCS(i, j) = max\{LCS(i-1, j), LCS(i, j-1)\}$ if $i > 0, j > 0$ and $A[i] \neq B[j]$ ✗
- [ ] $LCS(i, j) = 0$ if $i > 0, j > 0$ and $A[i] \neq B[j]$

Select all possible options that apply.

75%

What is the runtime of the dynamic programming algorithm if A and B have n elements?

$O($ n^2 $)$ ✓ 100%

Finding a longest common substring

If $A = TGACAACT$ and $B = GCTCTCGG$, fill in the following table of values for $LCS(i, j)$

LCS(i,j)		i=0	i=1	i=2	i=3	i=4	i=5	i=6		
			A[1] = T	A[2] = G	A[3] = A	A[4] = C	A[5] = A	A[6] = A	A[7] = C	A[8] = T
j=0		0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%
j=1	B[1] = G	0 ✓ 100%	0 ✓ 100%	1 ✓ 100%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%
j=2	B[2] = C	0 ✓ 100%	0 ✓ 100%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%

Exam 4

Assessment overview

Question

Submission status: complete
Best submission: 55%
Available points: —
Total points: 0.55 /1
Auto-graded question

Report an error in this question

Previous question

Next question

Personal Notes

No attached notes

Notes can't be added or deleted because the assessment is closed.

j=3	B[3] = T	0 ✓ 100%	1 ✓ 100%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%
j=4	B[4] = C	0 ✓ 100%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	3 ✗ 0%
j=5	B[5] = T	0 ✓ 100%	1 ✓ 100%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%
j=6	B[6] = C	0 ✓ 100%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%
j=7	B[7] = G	0 ✓ 100%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%
j=8	B[8] = G	0 ✓ 100%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%

What is the length of the longest common substring of $A = TGACAACT$ and $B = GCTCTCGG$?

4 ? ✗ 0%

This question is complete and cannot be answered again.

Correct answer

Longest common substring

Given an two strings A and B of lengths m and n, respecitve. A common substring is a string that appears as consecutive letters in both A and B. The goal is to find a common substring of A and B with maximal length.

For example if A = ababcab and B = abcb then the longest common substring is **abc**.

We want to design algorithms to calculate the length of the longest common substring.

Brute force algorithm

A brute force algorithm loops over all possible substrings of A and all possible substrings of B and checks if they are equal. It returns the length of the longest of these strings that are in both A and B.

```
MaxLen = 0
ForAll substring S of A
    ForAll substring T of B
        If S = T
            MaxLen = max(MaxLen, length of S)
Return MaxLen
```

What is the asymptotic runtime for this brute for algorithm where A and B have at most n characters

$O(n^4)$

Dynamic programming formulation

To solve this problem using dynamic programming we need to formulate a function for solving subproblems of the input and recursively solve more complicated problems. For this problem we will need to define a function $LCS(i, j)$, where i is an integer.

Which of the following descriptions of LCS are correct

$LCS(i, j)$ is the length of the longest common subsequence of $A[1..i]$ and $B[1..j]$ with last characters $A[i]$ and $A[j]$.

In defining $LCS(i)$ recursively, which of the following cases do we need to include? In these definition $A = a_1a_2 \dots a_n$ and $B = b_1b_2 \dots b_n$. Select all that should be part of the definition.

$LCS(i, j) = LCS(i - 1, j - 1) + 1$ if $i > 0, j > 0$ and $A[i] = B[j]$
$LCS(i, j) = 0$ if $i = 0$
$LCS(i, j) = 0$ if $j = 0$
$LCS(i, j) = 0$ if $i > 0, j > 0$ and $A[i] \neq B[j]$

What is the runtime of the dynamic programming algorithm if A and B have n elements?

$O(n^2)$

Finding a longest common substring

If $A = TGACAACT$ and $B = GCTCTCGG$, fill in the following table of values for $LCS(i, j)$

LCS(i,j)		i=0	i=1	i=2	i=3	i=4	i=5	i=6		
			A[1] = T	A[2] = G	A[3] = A	A[4] = C	A[5] = A	A[6] = A	A[7] = C	A[8] = T
j=0		0	0	0	0	0	0	0	0	0

j=1	B[1] = G	0	0	1	0	0	0	0	0	0
j=2	B[2] = C	0	0	0	0	1	0	0	1	0
j=3	B[3] = T	0	1	0	0	0	0	0	0	2
j=4	B[4] = C	0	0	0	0	1	0	0	1	0
j=5	B[5] = T	0	1	0	0	0	0	0	0	2
j=6	B[6] = C	0	0	0	0	1	0	0	1	0
j=7	B[7] = G	0	0	1	0	0	0	0	0	0
j=8	B[8] = G	0	0	1	0	0	0	0	0	0

What is the length of the longest common substring of $A = TGACAACT$ and $B = GCTCTCGG$? 2

Submitted answer 4
ashwin.pawar@slu.edu submitted at 2024-04-22 09:46:15 (CDT) 55%

hide

Longest common substring

Given an two strings A and B of lengths m and n, respecitve. A common substring is a string that appears as consecutive letters in both A and B. The goal is to find a common substring of A and B with maximal length.

For example if A = ababcab and B = abcb then the longest common substring is **abc**.

We want to design algorithms to calculate the length of the longest common substring.

Brute force algorithm

A brute force algorithm loops over all possible substrings of A and all possible substrings of B and checks if they are equal. It returns the length of the longest of these strings that are in both A and B.

```
MaxLen = 0
ForAll substring S of A
    ForAll substring T of B
        If S = T
            MaxLen = max(MaxLen, length of S)
Return MaxLen
```

What is the asymptotic runtime for this brute for algorithm where A and B have at most n characters

$O(2^n)$ Your answer is correct, but too loose. 25%

Dynamic programming formulation

To solve this problem using dynamic programming we need to formulate a function for solving subproblems of the input and recursively solve more complicated problems. For this problem we will need to define a function $LCS(i, j)$, where i is an integer.

Which of the following descriptions of LCS are correct

$LCS(i, j)$ is the length of the longest common subsequence of $A[1..i]$ and $B[1..j]$ with last characters $A[i]$ and $A[j]$. 100%

In defining $LCS(i)$ recursively, which of the following cases do we need to include? In these definition $A = a_1a_2 \dots a_n$ and $B = b_1b_2 \dots b_n$. Select all that should be part of the definition.

$LCS(i, j) = LCS(i - 1, j - 1) + 1$ if $i > 0, j > 0$ and $A[i] = B[j]$ ✓
$LCS(i, j) = 0$ if $i = 0$ ✓
$LCS(i, j) = 0$ if $j = 0$ ✓
$LCS(i, j) = max\{LCS(i - 1, j), LCS(i, j - 1)\}$ if $i > 0, j > 0$ and $A[i] \neq B[j]$ ✗
75%

What is the runtime of the dynamic programming algorithm if A and B have n elements?

$O(n^2)$ Correct! 100%

Finding a longest common substring

If $A = TGACAACT$ and $B = GCTCTCGG$, fill in the following table of values for $LCS(i, j)$

LCS(i,j)		i=0	i=1	i=2	i=3	i=4	i=5	i=6		
			A[1] = T	A[2] = G	A[3] = A	A[4] = C	A[5] = A	A[6] = A	A[7] = C	A[8] = T
j=0		0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%	0 ✓ 100%
j=1	B[1] = G	0 ✓ 100%	0 ✓ 100%	1 ✓ 100%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%
j=2	B[2] = C	0 ✓ 100%	0 ✓ 100%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%
j=3	B[3] = T	0 ✓ 100%	1 ✓ 100%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%
j=4	B[4] = C	0 ✓ 100%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	3 ✗ 0%
j=5	B[5] = T	0 ✓ 100%	1 ✓ 100%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%
j=6	B[6] = C	0 ✓ 100%	1 ✗ 0%	1 ✗ 0%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%
j=7	B[7] = G	0 ✓ 100%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%
j=8	B[8] = G	0 ✓ 100%	1 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	2 ✗ 0%	3 ✗ 0%	4 ✗ 0%

What is the length of the longest common substring of $A = TGACAACT$ and $B = GCTCTCGG$? 4

× 0%

Submitted answer 3
ashwin.pawar@slu.edu submitted at 2024-04-22 09:45:44 (CDT) invalid, not gradable

 show

Submitted answer 2
ashwin.pawar@slu.edu submitted at 2024-04-22 09:34:16 (CDT) invalid, not gradable

 show

Show/hide older submissions