



Gerência de Configuração: Cálculo de Diferenças

Leonardo Gresta Paulino Murta

leomurta@ic.uff.br

Exercício motivacional

- Defina um algoritmo para o cálculo das diferenças entre dois arquivos sem ancestral comum (*diff 2*)

Linha 1
Linha 2
Linha 3
Linha 4
Linha 5
Linha 6

Linha 5
Linha 3
Linha 8
Linha 4
Linha 2
Linha 9
Linha 6
Linha 1
Linha 7

Solução

- Identificar a maior subsequência comum entre os dois arquivos

Linha 1
Linha 2
Linha 3
Linha 4
Linha 5
Linha 6

Linha 5
Linha 3
Linha 8
Linha 4
Linha 2
Linha 9
Linha 6
Linha 1
Linha 7

Solução

- Verificar o que é adicionado ou removido a essa subsequência em cada arquivo

Linha 1		Linha 5
Linha 2		
	Linha 3	
		Linha 8
	Linha 4	
Linha 5		Linha 2
		Linha 9
	Linha 6	
		Linha 1
		Linha 7

Mas... como detectar a parte comum?

- Nesse algoritmo de *diff* 2, um ponto ainda está muito abstrato:
 - Como seria possível detectar a maior sequência comum?
- Uma solução:
 - Gerar todas as subsequências de um dos elementos
 - Verificar, para cada subsequência, se ela também é subsequência do outro elemento
- Problema:
 - Complexidade $O(2^n n)$

LCS

- Características do problema
 - Pode ser decomposto em subproblemas
 - Os subproblemas se repetem em diferentes decomposições
- Outra solução, mais eficiente, pode ser obtida com o algoritmo LCS
 - Longest Common Subsequence
 - Utilizado em bioinformática e no programa *diff*
 - Utiliza programação dinâmica
 - Complexidade $O(n^2)$

LCS

- Assuma as sequências
 - $X_i = (x_1, x_2, \dots, x_i)$
 - $Y_j = (y_1, y_2, \dots, y_j)$
- Algoritmo base de decomposição

$$LCS(X_i, Y_j) = \begin{cases} \emptyset & \text{se } i = 0 \vee j = 0 \\ (LCS(X_{i-1}, Y_{j-1}), x_i) & \text{se } x_i = y_j \\ \text{maisLongo}(LCS(X_i, Y_{j-1}), LCS(X_{i-1}, Y_j)) & \text{se } x_i \neq y_j \end{cases}$$

- Para cada parte calculada
 - Armazenar em *cache*
 - Consultar o *cache* antes de decompor o cálculo de outra subsequência

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset							
L_5							
L_3							
L_8							
L_4							
L_2							
L_9							
L_6							
L_1							
L_7							

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0						
L_3	0						
L_8	0						
L_4	0						
L_2	0						
L_9	0						
L_6	0						
L_1	0						
L_7	0						

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$					
L_8	0	$\leftarrow 0 \uparrow$					
L_4	0	$\leftarrow 0 \uparrow$					
L_2	0	$\leftarrow 0 \uparrow$					
L_9	0	$\leftarrow 0 \uparrow$					
L_6	0	$\leftarrow 0 \uparrow$					
L_1	0	$\nwarrow 1$					
L_7	0	$\uparrow 1$					

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1$
L_8	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$				
L_4	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$				
L_2	0	$\leftarrow 0 \uparrow$	$\nwarrow 1$				
L_9	0	$\leftarrow 0 \uparrow$	$\uparrow 1$				
L_6	0	$\leftarrow 0 \uparrow$	$\uparrow 1$				
L_1	0	$\nwarrow 1$	$\leftarrow 1 \uparrow$				
L_7	0	$\uparrow 1$	$\leftarrow 1 \uparrow$				

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1$
L_8	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$
L_4	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$			
L_2	0	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1 \uparrow$			
L_9	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$			
L_6	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$			
L_1	0	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$			
L_7	0	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$			

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1$
L_8	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$
L_4	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\nwarrow 2$	$\leftarrow 2$	$\leftarrow 2$
L_2	0	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$		
L_9	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$		
L_6	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$		
L_1	0	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$		
L_7	0	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$		

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1$
L_8	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$
L_4	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\nwarrow 2$	$\leftarrow 2$	$\leftarrow 2$
L_2	0	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\leftarrow 2 \uparrow$
L_9	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	
L_6	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	
L_1	0	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	
L_7	0	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1$
L_8	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$
L_4	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\nwarrow 2$	$\leftarrow 2$	$\leftarrow 2$
L_2	0	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\leftarrow 2 \uparrow$
L_9	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\leftarrow 2 \uparrow$
L_6	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\nwarrow 3$
L_1	0	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\uparrow 3$
L_7	0	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\uparrow 3$

LCS (voltando ao exemplo)

	\emptyset	L_1	L_2	L_3	L_4	L_5	L_6
\emptyset	0	0	0	0	0	0	0
L_5	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$
L_3	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1$
L_8	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$
L_4	0	$\leftarrow 0 \uparrow$	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\nwarrow 2$	$\leftarrow 2$	$\leftarrow 2$
L_2	0	$\leftarrow 0 \uparrow$	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\leftarrow 2 \uparrow$
L_9	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\leftarrow 2 \uparrow$
L_6	0	$\leftarrow 0 \uparrow$	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\nwarrow 3$
L_1	0	$\nwarrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\uparrow 3$
L_7	0	$\uparrow 1$	$\leftarrow 1 \uparrow$	$\leftarrow 1 \uparrow$	$\uparrow 2$	$\leftarrow 2 \uparrow$	$\uparrow 3$

Referências

- Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C., 2001. Introduction to Algorithms, 2nd ed., MIT Press.



Gerência de Configuração: Cálculo de Diferenças

Leonardo Gresta Paulino Murta

leomurta@ic.uff.br