

RECITATION 11, 10/21/09

①

TOPICS

- STABILITY
- SORTING STRINGS

STABLE SORTING ALGORITHM PRESERVES
THE INITIAL ORDERING OF IDENTICAL ELEMENTS

EXAMPLE:

	MEDALS	SECONDARY KEY ↓	PRIMARY KEY ↓	
ALPHABETICAL LIST OF COUNTRIES	AUSTRALIA	3	SLOVENIA	13
	AUSTRIA	2	SLOVAKIA	12
	CAMBODIA	7	U. A. E	12
	CAMEROON	7	UNITED STATES	12
	CHILE	10	CHILE	10
	CHINA	10	CHINA	10
	IRAN	10	U. K.	10
	IRAQ	4	...	
	SLOVAKIA	12		
	SLOVENIA	13		
	UNITED ARAB EMIRATES	12		
	UNITED KINGDOM	10		
	UNITED STATES	12		

STABLE SORT COMPARES JUST NUMBERS

NO NEED FOR EXPENSIVE COMPARISON OF STRINGS

RADIX SORT

2

WANT TO SORT N STRINGS OF LENGTH L
OVER ALPHABET Σ

INPUT:
ABDC
BADA
ACAA
BCAB

STABLE
COUNTING
SORT
ACCORDING
TO

$O(N + |\Sigma|)$
TIME

~~ABDC~~
SORTED
BADA
ACAA
BCAB
ABDC

STABLE
SORTING
ACCORDING
TO

$O(N + |\Sigma|)$
TIME

SORTED
ACAA
BCAB
BADA
ABDC

STABLE
SORTING
ACCORDING
TO

~~ACAA~~
SORTED
BADA
ABDC
ACAA
BCAB

SORT
ACCORDING
TO

ABDC
ACAA
BADA
BCAB

ALL SORTED

TIME COMPLEXITY:

L TIMES STABLE COUNTING SORT =

$$= L \cdot O(N + |\Sigma|) = O(L(N + |\Sigma|))$$

DIFFERENT LENGTHS

3

$S = \text{SUM OF LENGTHS OF ALL STRINGS}$
(NONE OF THEM EMPTY)

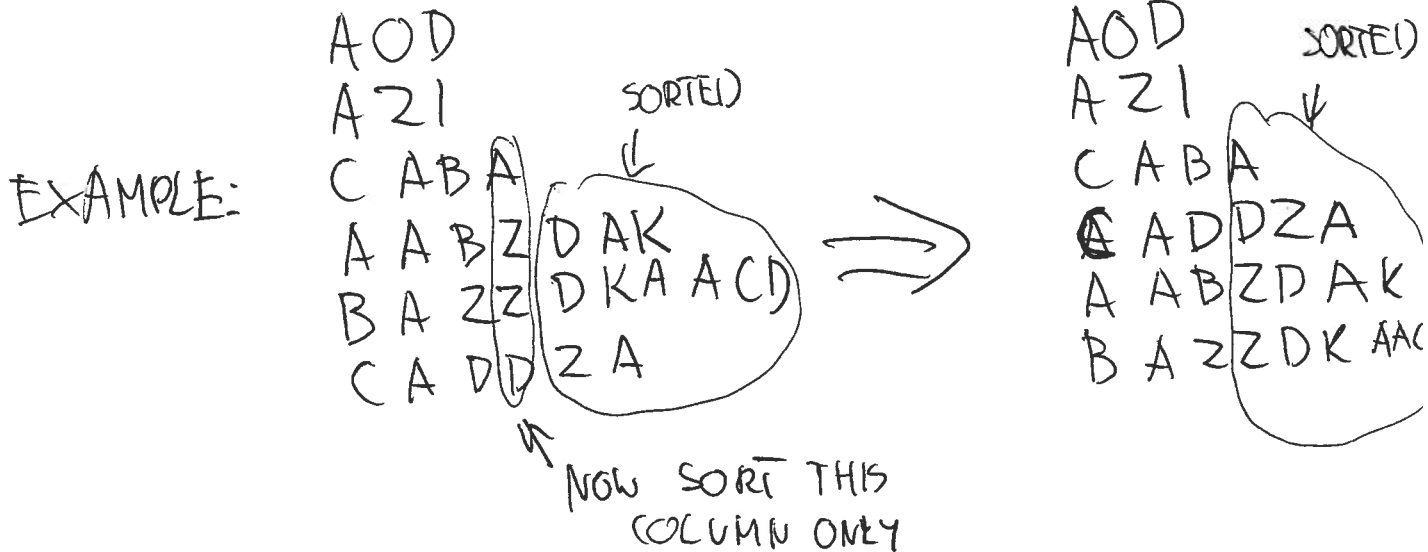
PADDING STRINGS WITH SPECIAL CHARACTERS CAN RESULT
IN $\Theta(S \cdot (S + |\Sigma|))$ TIME IN WORST-CASE

DIFFERENT SOLUTION:

1. USE COUNTING SORT TO SORT BY LENGTH
IN ~~$\Theta(S \cdot (S + |\Sigma|))$~~ $O(S)$ TIME:

AOD
AZI
CABA
CADDZA
AABZDAK
BAZZDKAACD

2. PROCEED AS BEFORE FROM LEAST SIGNIFICANT
COLUMN BUT SORT ONLY STRINGS THAT
HAVE A CHARACTER IN A GIVEN COLUMN



COMPLEXITY:

FOR EACH COLUMN:

$$O(|\Sigma| + \underbrace{\text{\#CHARACTERS IN THE COLUMN}})$$

SUMMING OVER ALL COLUMNS

$$O(\underbrace{\text{\#COLUMNS}}_{\leq S} \cdot |\Sigma| + S)$$

$$O(|\Sigma| \cdot S)$$

NOTE:

IF Σ BIG (SAY, UNICODE), THE $|\Sigma|$ FACTOR CAN BE PROHIBITIVE

BUT THE ALGORITHM CAN BE

IMPROVED TO $O(|\Sigma| + S)$

↑
ADDITION
INSTEAD
OF MULTIPLICATION