[I]

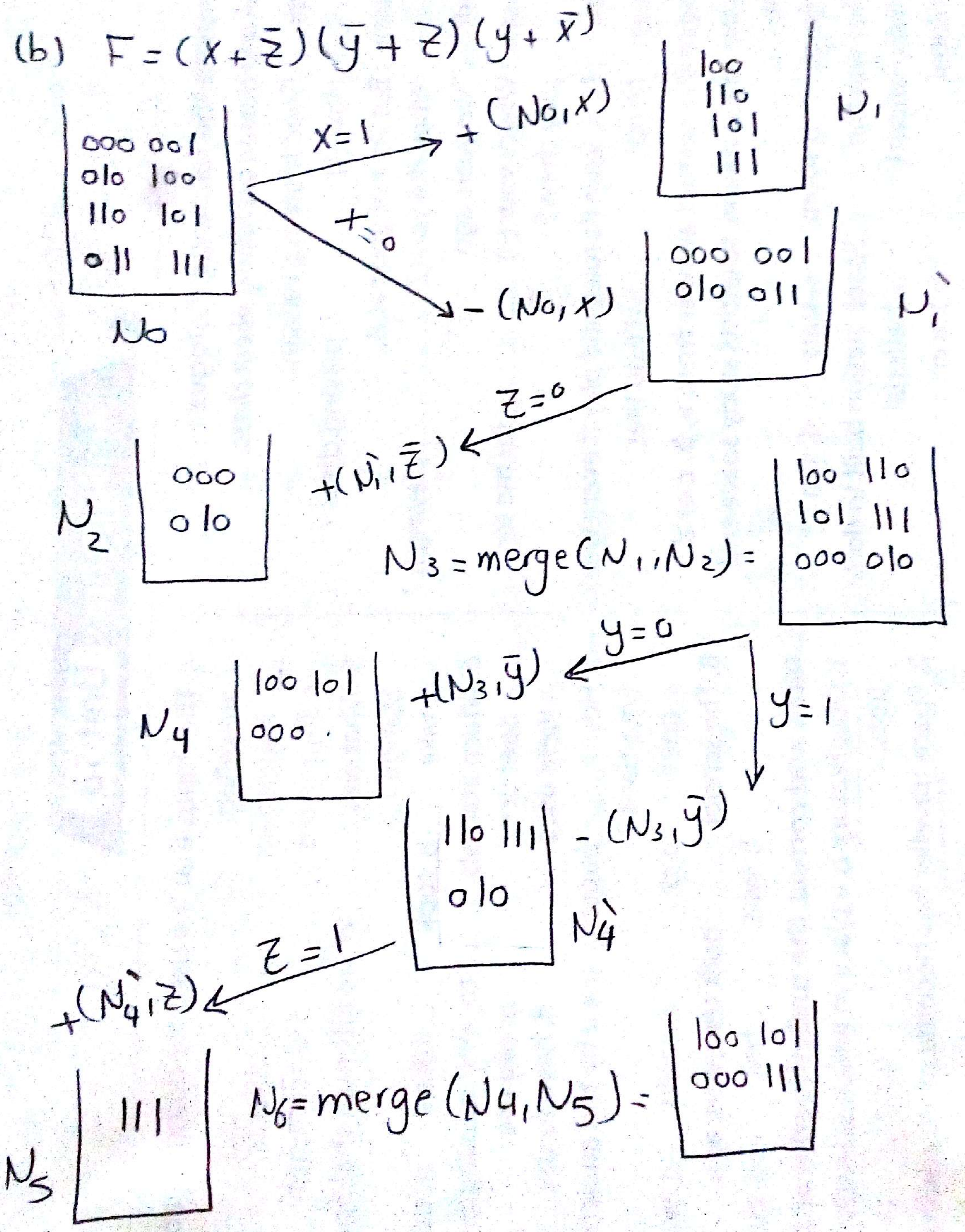
bioinformatics Exam
2014-2015

Question 1

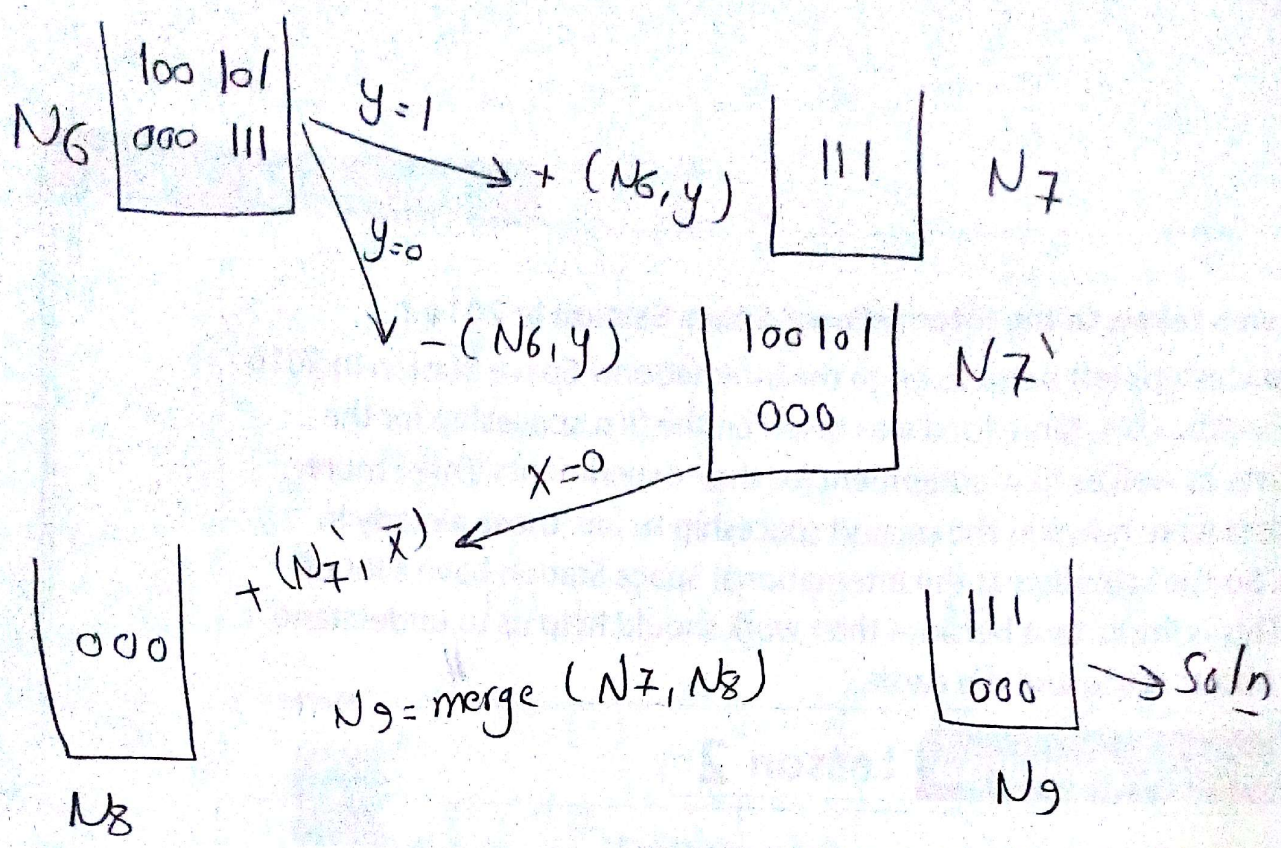
(a)

مش شايفا

(b) $F = (x + \bar{z})(\bar{y} + z)(y + \bar{x})$



2



Lipton's Algorithm in TTpL

(1) input (N0)
(2) $N_1 = +(N_0, X)$
(3) $N_1' = -(N_0, X)$
(4) $N_2 = +(N_1', Z)$
(5) $N_3 = merge(N_1, N_2)$
(6) $N_4 = +(N_3, \bar{y})$
(7) $N_4' = -(N_3, \bar{y})$
(8) $N_5 = +(N_4', Z)$
(9) $N_6 = merge(N_4, N_5)$
(10) $N_7 = +(N_6, y)$
(11) $N_7' = -(N_6, y)$
(12) $N_8 = +(N_7', \bar{X})$
(13) $N_9 = merge(N_7, N_8)$
(14) detect (N_9)

#

3

Question 2

(a) $\Pi = (1\ 6\ 2\ 3\ 4\ 5)$

$i = 1$, $\Pi_j = i$, $j = 1$

$i = 2$, $\Pi_j = i$, $j = 3$, $\rho(2,3)$

$\Pi = (1\ 2\ 6\ 3\ 4\ 5)$

$\Pi = (1\ 2\ 3\ 6\ 4\ 5)$

$i = 3$, $\Pi_j = i$, $j = 4$, $\rho(3,4)$

$i = 4$, $\Pi_j = i$, $j = 5$, $\rho = (4,5)$

$\Pi = (1\ 2\ 3\ 4\ 6\ 5)$

$\Pi = (1\ 2\ 3\ 4\ 5\ 6)$

number of reversal = 4

$i = 5$, $\Pi_j = i$, $j = 6$, $\rho = (5,6)$

$\Pi = (1\ 6\ 2\ 3\ 4\ 5)$

$\rho(2,6)$

$\Pi = (1\ 5\ 4\ 3\ 2\ 6)$

$\rho(2,5)$

$\Pi = (1\ 2\ 3\ 4\ 5\ 6)$

min number of Reversal = 2

#

4

(b)

$L = \{2, 2, 3, 3, 4, 5, 6, 7, 8, 10\}$

$X = \{0\}$ max element in $L = 10$

$X = \{0, 10\}$ $L = \{2, 2, 3, 3, 4, 5, 6, 7, 8\}$

Remove 8

$y = 2$ $y = 8$

$\Delta(y, x) = \{2, 8\}$ $\Delta(y, x) = \{8, 2\}$ → back tracking

$X = \{0, 2, 10\}$ $L = \{2, 3, 3, 4, 5, 6, 7\}$

Remove 7

~~$y = 3$~~ $y = 7$

~~$\Delta(y, x) = \{3, 7\}$~~ $\Delta(y, x) = \{7, 5, 3\}$

$X = \{0, 2, 7, 10\}$ $L = \{2, 3, 4, 6\}$

Remove 6

$y = 4$ $y = 6$

$\Delta(y, x) = \{4, 2, 3, 6\}$ $\Delta(y, x) = \{6, 4, 1, 4\}$

$X = \{0, 2, 4, 7, 10\}$ $L = \{\}$

[5]

Question 3

(a) Problem → this algorithm may work Forever

There is no guarantee that eliminating Some break points will not introduce others, leading to an endless Cycle.

Improved Breakpoint Reversal Sort Algorithm

```
Improved Breakpoint Reversal Sort (Π)
  while b(Π) > 0
    if Π has a decreasing strip
        Among all reversals, choose reversal ρ minimizing b(Π·ρ)
    else
        Choose a reversal ρ that flips an increasing strip in Π
  Π ← Π·ρ
  output Π
return
```

#

(b) مش طلبنا

(6)

Question 4

(a) Alignment

A	G	G	T	A	C	T	T
C	C	A	T	A	C	G	T
A	C	G	T	T	A	G	T
A	C	G	T	C	C	A	T
C	C	G	T	A	C	G	G

Profile matrix

A	(3)	0	1	0	(3)	1	1	0
C	2	(4)	0	0	1	(4)	0	0
G	0	1	(4)	0	0	0	(3)	1
T	0	0	0	(5)	1	0	1	(4)

Consensus AC GT ACGT

Score 3+4+4+5+3+4+3+4

Total Distance 2+1+1+0+2+1+2+1

Sum 5 5 5 5 5 5 5 5

The motif finding problem is a maximization problem while The median string problem is a minization problem

motif finding problem need to maximizing score

median string problem need to minizing total distance

#

17)

(b)

Motif finding problem	Gold bug problem
- motifs Encode for a message in the "Genetic language"	- Symbols Encode for a message in the "English language"
- analyze the Frequencies of pattern in DNA message	- analyze the Frequencies of Pattern in Gold bug message
- establish regulatory Motif	- word in the English dictionary
hints	hints
- hidden seq. is of length 8	- English message
- the pattern is not exactly the same in each array because random points mutation may occur in the seq.	- Each symbol Correspond to one letter
	- No punctuation

but the motif finding problem is harder than the Gold bug problem because

- we don't have the Complete dictionary of Motifs
- Genetic language doesn't have standard Grammer

#