

CM224 HW 4 Solution

Ashish Kumar Singh (UID:105479019)

October 30, 2021

Problem 1. Clark's Algorithm

Answer 1. d) GGGAA

Solution 1. Using Clark's Algorithm, we get following haplotypes: AGGAA, AGAAA, GGAGG, GAGGA, AAAAA, GGGAA
So, the last haplotype added is GGGAA

Problem 2. Expectation Maximization

Answer 2. (b) $1/4$

Solution 2. Possible haplotypes phasing are:

- 1) 01111 00110
- 2) 01110 00111
- 3) 10110 00100
- 4) 10100 00110
- 5) 10111 10110

These pairs give 8 distinct haplotypes:

- i) 01111 ii) 00110 iii) 01110 iv) 00111 v) 10110 vi) 00100 vii) 10100 viii) 10111

We give them equal probability of $1/8$. For Expectation step, we get following probabilities for haplotype phasing:

- 1) 01111 00110 (0.5)
- 2) 01110 00111 (0.5)
- 3) 10110 00100 (0.5)
- 4) 10100 00110 (0.5)
- 5) 10111 10110 (1.0)

For Maximization step, we get following probabilities for haplotypes:

- i) 01111 ($1/12$) ii) 00110 ($1/4$) iii) 01110 ($1/12$) iv) 00111 ($1/12$) v) 10110 ($1/4$) vi) 00100 ($1/12$) vii) 10100 ($1/12$) viii) 10111 ($1/12$)

We see the maximum probability value is $(1/4)$.