Algorithms, Summer 2019 at CIS

Homework 3 **Due: 7/26/19 before class**

1. Let $U = 2^u$ and $M = 2^m$. Prove that the family of hash functions $A \cdot x + b \bmod 2$, where $A \in \{0,1\}^{m \times u}$ is a random binary matrix, and $b \in \{0,1\}^m$ is a random binary vector, is a 2-universal family.

2. Recall that in class we came up with an algorithm which solved the ϵ-heavy hittes problem in a single pass, namely, in a stream of length t it output a set of size at most $1/\epsilon$ containing all items e for which $\text{count}(e) > \epsilon t$. Also, the algorithm used $O(1/\epsilon)$ words of memory. However, the algorithm was also allowed to output false positives, that is, elements e for which $\text{count}(e) \leq \epsilon t$. Suppose now you are in the data stream model but you are allowed two passes over your input stream $a_1, \ldots, a_m$. Show how you can use two passes and $O(1/\epsilon)$ words of memory to output a set containing all items e for which $\text{count}(e) > \epsilon t$, and no items e for which $\text{count}(e) \leq \epsilon t$, that is, your algorithm should not output any false positives.

3. Consider the linear program: maximize $8x + 6y$ subject to $x + 2y \leq 6$, $5x + 2y \leq 20$ and $x, y \geq 0$. Calculate the optimal solution (x, y).