Group Exercise 1 Question E

Philipson.S, Melin.O, Palm.J

Algorithm Description

Our algorithm samples $\lceil log(n) \rceil$ words from Alices wordlist and sends them one by one to Bob, who checks if he's got the same word in his list. If Bob can match at least one of the words he receives to his own wordlist he believes that it's a fair representation of Alices wordlist as well, and reports the Hypothesis to be true.

Pseudo Code

```
A = List of Alices words
B = List of Bobs words
  where |A| == |B|

n = The length of the lists (|A|)

for i -> 0 to log n
  random sample elem from A
  if elem exists in B
    return true
  end
end
return false
```

Theoretical Approach

The algorithm builds upon the idea that we are free to report anything if the number of common words x is $0 < x < \frac{1}{10} \cdot n$. As such, we will always report a true if we get at least 1 match, as

$$P(X = True) = 1 - P(X = False) = 1 - (1 - p)^{k \cdot log_2(n)}$$