

Exercise 1

One can simply use the ~~extended~~ respective definition of intersection and union in the case of multisets and thus construct a similarity metric that is analogous to the Jaccard Similarity:

$$\text{Jac-mult}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

, where

$A \cap B$ is the largest multiset that is a multisubset of A and B , i.e.

$$a \in A \text{ and } a \in B \Rightarrow a \in A \cap B,$$

and $A \cup B$ is the smallest multiset such that A and B are multisubsets of it.

Example: $A = \{a, a, b, b, b, c, d, d\}$
 $B = \{b, b, c, c, c, d, d, e\}$

$$\Rightarrow A \cup B = \{a, a, b, b, b, c, c, c, d, d, e\}$$

$$A \cap B = \{b, b, c, d, d\}$$

So in the ^{case} where both multisets are in fact sets, this similarity metric corresponds exactly to the Jaccard Similarity.