Multisets: The Basics

Multisets are like sets, but the elements can be repeated: the number of times that an element appears in a multiset is its *multiplicity* (if an element does not appear, then its multiplicity is zero). For example, if you have two green marbles and one red marble, then you have two distinct elements: 'green' with multiplicity 2, and 'red' with multiplicity 1. If you only keep the distinct elements of a multiset, then you get a set, which is called the *support* of the multiset.

Figure 1: A multiset and its support.

You have an *inclusion* between two multisets if all elements of the first are also contained in the second with at least the same multiplicity.

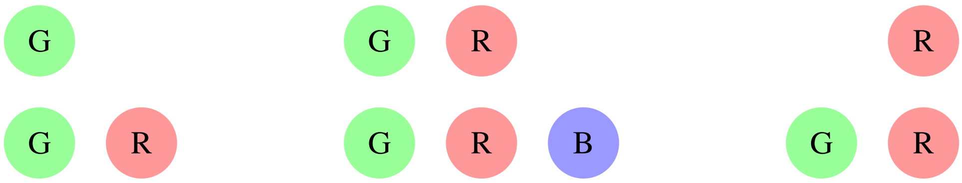

Figure 2: The multiset "2 green, 1 red" is included in "2 green, 2 red, 1 blue", but it is not included in "1 green, 2 red".

The *sum* of two multisets is the multiset obtained by taking all their elements together, which amounts to summing the multiplicities.
The *intersection* of two multisets is the largest multiset which is included in both (simply take the common elements), while the *union* is the smallest multiset which includes both (for the union one has to take each element with sufficient multiplicity). Of course we may similarly define sum, intersection, and union for several multisets (or just for one multiset, the result being the multiset itself).
One can easily show that intersection and union, considered as operations between multisets, are commutative, associative, and each of them is distributive with respect to the other. This means that the following set of formulas hold for any three multisets A, B, C:

$$\begin{aligned} A\cap B &= B\cap A & A\cup B &= B\cup A \\ (A\cap B)\cap C &= A\cap(B\cap C) & (A\cup B)\cup C &= A\cup(B\cup C) \\ (A\cap B)\cup C &= (A\cup C)\cap(B\cup C) & (A\cup B)\cap C &= (A\cap C)\cup(B\cap C)\,. \end{aligned}$$

There are also further formulas involving the sum (maybe you want to find some of them out yourself?). Notice that to prove these and several other properties for multisets it suffices to consider each element and analyze its multiplicities in the various multisets.

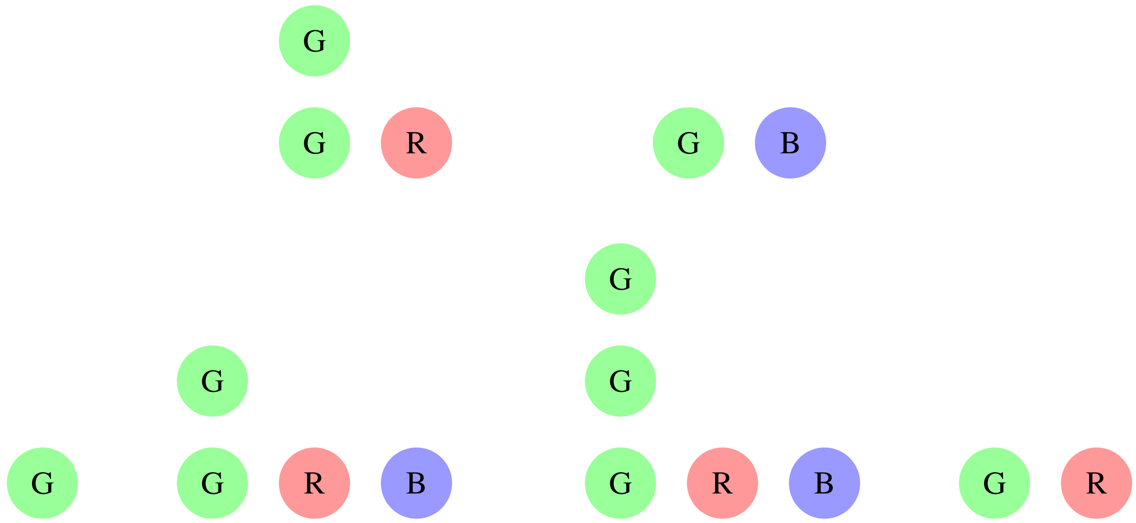

Figure 3: The four multisets below are respectively the intersection, the union, the sum, and the difference of the two multisets above.

Two or several multisets are said to be *disjoint* if their intersection is empty, which means that they have no common elements. Disjoint multisets are also *pairwise disjoint* if any two of them are disjoint, so if any two of them have no common elements (note that this is in general a stronger condition). Two disjoint multisets (or more generally several pairwise disjoint multisets) have many additional properties, for example their union equals their sum.

Finally, if a multiset is included in another one, you may consider their *difference*: you simply remove all elements of the former from the latter. There is also a way of defining the difference of any two multisets: say we want to subtract from some multiset M_1 the multiset M_2, and we need to define the difference multiset D. Consider any element: If it occurs in M_1 less often than in M_2, then its multiplicity in D is zero. Otherwise its multiplicity in D is the difference of the multiplicites in M_1 and in M_2. We now illustrate this definition with an example. If you have some piece of fruits on your shopping list, then consider what happens to the greengrocer's supplies after your purchase: if the greengrocer has what you need, you just subtract what you buy, and otherwise you buy as much as possible, leaving the shop without supplies of something (the difference between the multiset "greengrocer's supplies before the purchase" and the multiset "shopping list" is the multiset "greengrocer's supplies after the purchase").

As you may have noticed, multisets operations are a generalization of set operations. Sometimes more than one generalization is possible. For example there is a different definition for the intersection (which amounts to multiplying the multiplicities), because this also generalizes the usual intersection for sets. There is also another definition of difference, where the elements of the multiset that we subtract become so to say 'prohibited', and we remove them entirely from the other set.

EXERCISES

EXERCISE 1: In a multiset, what is the relation between the total number of elements and the multiplicities of the distinct elements?

EXERCISE 2: To make some fruity milkshake you can use either 1 kiwi and 1 banana, or 2 kiwis and 3 strawberries. What do you need to buy if you plan to make one of the two milkshakes, but you are still undecided about which one? Which operation of multisets are you doing?

EXERCISE 3: What are the multiplicities for the intersection, the union, and the sum, in terms of those of the original multisets?

EXERCISE 4: Can you describe the multiplicities of the elements for disjoint multisets and for pairwise disjoint multisets?

EXERCISE 5: If you have three disjoint multisets, does their union equal their sum?

SOLUTIONS

SOLUTION 1: The total number of elements of a multiset is the sum of the multiplicities of the distinct elements.

SOLUTION 2: You would buy 2 kiwis, 1 banana, and 3 strawberries. You are doing the union of the multisets of the ingredients.

SOLUTION 3: The multiplicity for the intersection is the minimum of the multiplicities. The multiplicity for the union is the maximum of the multiplicities. The multiplicity for the sum is the sum of the multiplicities.

SOLUTION 4: For disjoint multisets, the minimum of the multiplicities of an element is zero because there must be some multiset that does not contain that element (otherwise there would be a common element). For pairwise disjoint multisets, each element can occur in only one of the multisets, otherwise there would be two multisets with some common element. Thus all multiplicities for an element are zero, with at most one exception.

SOLUTION 5: In general the sum may be larger than the union. For two or more multisets, the union equals the sum if and only if the multisets are pairwise disjoint: this condition is necessary and sufficient. Indeed, considering the multiplicities for an element, we want that the maximum of the multiplicities equals the sum of the multiplicities: this happens if and only if the multiplicities with at most one exception are zero. Thus the multisets are pairwise disjoint.