AIML: Some Elementary Combinatorics

v1.1

Dr. Max A. Little

February 3, 2024

This course studies many kinds of symbolic AI problems which can be solved using combinatorial optimization. This requires some basic knowledge from the subject of *combinatorics*: the mathematical study of arranging, organizing, comparing, selecting, generating and counting *collections* of elements (items). First, we introduce some of the elementary types of collections. For some collection X, it has *size* (number of items in the collection) denoted by $|X|$.

Sets *Unordered* collections of *distinct* items, usually written using curly braces e.g. $X = \{5, 1, 3\}$ or $Z = \{\text{apple}, \text{banana}, \text{strawberry}, \text{melon}\}$. The empty set is denoted $\emptyset$. Sets are joined with the set *union* e.g. for $Y = \{7, -1, 3, 5\}$, $X \cup Y = \{5, 1, 3\} \cup \{7, -1, 3, 5\} = \{5, 1, 3, 7, -1\}$ with $X \cup \emptyset = \emptyset \cup X = X$. Set union is *idempotent* so $X \cup X = X$. Another operation on sets is *intersection*, which gives the common elements, so $X \cap Y = \{3, 5\}$ and $Y \cap Z = \emptyset$. The difference between two sets, $Y \backslash X$ also written $Y - X$, is all the elements in Y not in X, so $Y \backslash X = \{7, -1, 3, 5\} \setminus \{5, 1, 3\} = \{7, -1\}$. A *subset* of a set, is a selection of some elements from a set, so $\{3, 5\}$ is a subset of both X and Y. The empty set $\emptyset$ is a subset of any set and a set is also a subset of itself. All the subsets of X are $\emptyset$, $\{5\}$, $\{1\}$, $\{3\}$, $\{5, 1\}$, $\{5, 3\}$, $\{1, 3\}$ and $\{5, 1, 3\}$ itself. The number of possible subsets of a set of size N is 2^N, which is a very large number even for small N, for instance for $N = 20$, $2^N = 1,048,576$.

Lists *Ordered* collections of items, not necessarily distinct, written using square braces e.g. $X = [5, 1, 1]$ or $Z = [z, u, q, z]$. When joining two lists the ordering matters, so for $Y = [2, -1]]$, $X \cup Y = [5, 1, 1, 2, -1]$ which is not the same as $Y \cup X = [2, -1, 5, 1, 1]$. The empty list $[\,]$ is the unit for list join so $X \cup [\,] = [\,] \cup X = X$. Lists are the same as *tuples* such as $(2, 1, -3)$ except that we can index them, e.g. $Z_1 = z$, $Z_2 = u$,$Z_3 = q$ and $Z_4 = z$. A *sublist* of a list, is a list which contains some elements of a list V in the same order as they appear in V, e.g. for $[-1, 7, 8, 3]$ then one sublist is $[-1, 3]$ and another is $[-1, 8]$. The empty list is a sublist of any list, and a list is a sublist of itself. The number of possible lists of size M containing elements of the set X of size N, is N^M.

Composites Collections can be *nested*, or contained inside each other. For instance, $X = \{[5,1,6],[2,-1],[\,]\}$ is a set of lists, whereas, $Y = [\{5,1,6\},\{2,-1\},\emptyset,\{2,-1\}]$ is a list of sets, and $Z = \{\{5,1,6\},\emptyset,\{2,-1\}\}$ is a set of sets. In X, while the ordering in which the lists appear in the set does not matter, there can be no duplicates lists. So, $\{[2,-1],[5,1,6],[\,]\}$ is the same as X but $\{[5,1,6],[-1,2],[\,]\}$ is not. Whereas, in Y, the ordering in which the sets appears in the list matters and there can be duplicate sets, but the ordering of the elements in each of the lists does not matter. As examples, $[\{6,5,1\},\{2,-1\},\emptyset,\{-1,2\}]$ is the same as Y but $[\{5,1,6\},\emptyset,\{2,-1\}]$ is not.

There are many ways to order, select or arrange one of the simple collection of items above. We list some of the most commonly encountered ones below.

Permutations A collection of elements *arranged in a particular order*, is called a *permutation* (of a set or a list). So, for the set $X = \{5,1,3\}$, there are six distinct permutations which we can give as lists, $[3,1,5]$, $[1,3,5]$, $[1,5,3]$, $[3,5,1]$, $[5,3,1]$ and $[5,1,3]$ (permutations cannot be given as sets since in sets the ordering does not matter). In general, for a collection of size N, there are $N! = N\times(N-1)\times\cdots\times2\times1$ permutations, which is a huge number for even small N. For instance, $3! = 6$, $4! = 24$ but $10! = 3,628,800$.

Combinations For a collection $X$of size $|X| = N$, a particular subset of $M \leq N$ items selected from X not arranged in any order, is called a *combination* of size M. For example, all the combinations of size $M = 2$ from $X = \{5,1,3\}$ are the subset $\{1,5\}$, $\{3,5\}$ and $\{1,3\}$. The number of combinations of size M from a collection of size N is given by the *binomial coefficient*, $N!/(M!(N-M)!)$. For instance, for the previous example, $N = 3$ and $M = 2$ so we have $(3\times2\times1)/2 = 3$ possible combinations. The number of combinations of size M from a collection of size N, is approximately equal to N^M.

Partitions For any set or list, we can split it up into a collection of subsets/sublists in such a way that when these subsets/sublists are joined, the result is equal to the original set/list. This is called a *partition* of a list/set. For instance, the set $X = \{5,1,3\}$ has the following five (set) partitions, $\{\{1\},\{5\},\{3\}\}$, $\{\{3\},\{5,1\}\}$, $\{\{3,5\},\{1\}\}$, $\{\{3,1\},\{5\}\}$ and $\{\{1,3,5\}\}$. We can check these are partitions since e.g. $\{3\}\cup\{5,1\} = \{1\}\cup\{5\}\cup\{3\}=\{1,3,5\} = X$. The subsets in the partition are called *blocks* or *parts* of the partition. The number of set partitions B_N of set of size N, is $B_0 = 1$, $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$ and so on. This can actually be computed using the so-called *Bell number* recurrence. For a list $X = [5,1,3]$ there are four *list partitions*, $[[5],[1],[3]]$, $[[5,1],[3]]$, $[[5],[1,3]]$ and $[[5,1,3]]$. The sublists making up each partition are called *segments* of the partition (see below). For a list of size N, there are 2^{N-1} possible partitions such as this.

Compositions Take a list of length N and form a list partition of it. Then, replace each segment in the partition with its size. Clearly, the total sum of these segment sizes must add up to N. For instance, for the example above, the list partitions of

$X = [5, 1, 3]$ are replaced by the lists $[1, 1, 1]$, $[2, 1]$, $[1, 2]$ and $[3]$. These are called *compositions* of N. In this way, the whole number N can be expressed as a sum of whole numbers in 2^{N-1} different ways. For instance, $N = 4$ has eight different compositions.

Segments/subsequences List *segments* or *subsequences* are sublists which are in consecutive order. For instance, for the list $X = [u, u, v, j, z, q]$, one segment is $[v, j, z]$ and another is $[u]$. Empty lists are segments of any list. For a list of size N, there are $O\left(N^2\right)$ such segments. For example, for the list $[1, 2, 3]$ there are seven segments $[1]$, $[1, 2]$, $[1, 2, 3]$, $[2]$, $[2, 3]$, $[3]$ and $[\,]$.

Head/tail subsequences Given a list, a *head subsequence* is a segment which starts at the beginning of the list. As an example, the list $X = [a, b, c, d]$ has head subsequences $[\,]$, $[a]$, $[a, b]$, $[a, b, c]$ and $[a, b, c, d]$. Similarly, *tail subsequences* are segments which end at the end of the list. For the list X previously, the tail subsequences are $[a, b, c, d]$, $[b, c, d]$, $[c, d]$, $[d]$ and $[\,]$. For a list of size N, there are $N + 1$ head/tail subsequences.