1 Solution to Problem 1

1a) To begin it is important to note what an inconsistent Hypothesis would mean, an inconsistent hypothesis leads to two different conclusions, more specifically for clarity it leads to contradictory conclusions and we do not want that we want reliability. This means for a hypothesis to be consistent it has to give the correct classification based on the current hypothesis. Leaving no room for errors on the training set.

While the **Least General Hypothesis** (h) which is also known as the most specific hypothesis in the hypothesis space, is such that it covers all positive examples without covering any negative examples. The least general hypothesis is such that it fits all positive instances in the training set. The consistency of the Least general Hypothesis relies on its ability to correctly classify all examples according to the target concept. In a case where all tuples or instances in the training sample are unique and correctly labelled, the Least General Hypothesis designed to fit these examples will be consistent with the training data because it matches all provided labels. In the context of what we have been taught in class, it will appear as:

$$\langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$$

The least general hypothesis by nature is **always consistent**. It is usually consistent with the training data it was derived from and covers all positive examples while excluding negatives. Certain factors that could influence its consistency in the real world are: the correctness of labels, noise in the data, missing attributes etc. But otherwise, all things being equal it always stays consistent.

1b) A hypothesis $h \in H$ that does not cover negative examples, is one that h does not incorrectly classify any negative example as positive. But the part of consistency also depends for a hypothesis h also heavily depends on its performance with respect to the positive examples in the dataset too.

A consistent hypothesis h is one that correctly classifies all the examples in the dataset, both positive and negative. Meaning it does not lead to contradictory conclusions and is very reliable. More specifically:

1. **Case of Negative examples:** The hypothesis must not classify any negative example as positive, which is satisfied in this case since h does not cover any negative examples.

2. **Case of Positive examples:** The hypothesis also must classify all positive examples as positive. If h fails to cover some or all positive examples, then it would not be consistent, because this would mean it is incorrectly classifying these positive examples (either by not classifying them at all or by classifying them as negative, depending on the context).

Consistency of a hypothesis, h is not only about only covering positive examples it must correctly take negative examples into account and vice versa. If h is able to cover all positive examples without including any negative ones, then h is indeed consistent. But if, h misses some positive examples, it is not fully consistent even if it does not cover any negative examples.

2 Solution to Problem 2

2) To prove that the post in the diagram has no dimension bigger than 3 we have to use the size of its smallest realizer or since the question says at MOST, if we have a realizer of size 2 or3 that should be sufficient for the proof. Else we provide a counter example, meaning trying to prove that the smallest realizer involves at least 4 posets.

A realizer of a partially ordered set consists of some linear extensions whose intersection give the original poset.

Smallest realizer means it uses the minimum number of linear extensions to give back original poset.

Based on this question, to prove that the Hasse diagram has no dimension ¿ 3(the smallest realizer isn't bigger than 3). We have to provide a realizer of size 3 or less.

These are the linear extensions that make up my realizers, and all their intersections result in the original poset as described by the Hasse diagram in the Homework problem.

The original poset must maintain this order: where each (a,b) means a must come before b. $< * =$ (10,9), (11,9), (12,9), (10,8), (8,4), (9,4), (6,4), (13,4), (6,7), (8,1), (13,2), (2,3), (1,3), (1,5), (4,5)

Here, $< *$ represents the immediate predecessor of each node in the poset. This specifies the items that must maintain a specific order.

To construct my 3 linear orderings I began by constructing the first one and going from left to right to create the linear extension to maintain all the orderings as defined in $< *$. I then go on to create the second linear extension by going right to left while maintaining all orderings as defined in $< *$. Then finally I create the third linear extension by paying attention to the intersection of the first two and noting the items in the intersection that are not a part of the orderings specified in $< *$. I use my third poset to correct that.

This process helped me ensure that for items whose order are not specified in $< *$ their orderings knock out each other. For examples: (2,6) if 2 comes before 6 in any of the other 2 linear extensions it must have a corresponding (6,2) such that 6 comes before 2. This is important so that when we take the intersections of all the 3 intersections, all orderings not relevant to us can knock out each other and lead to only the orderings in $< *$ being left.

My final step is to construct a table where the left side is 1-13 and each number lists out the nodes associated with it(meaning nodes that are comparable to it, whether they be the immediate predecessor or not) because those are the items whose ordering must stay the same.

I check for each of the numbers 1-13 and to avoid doing extra work when I check node 1 for example against 2 - 13 to make sure important orders as listed in the table are maintained and knockouts are also implemented, I then go and check node 2 but for node 2 I check it against 3-13. For each node, I only check against node + 1 value up until 13. By so doing I am able to correct any observed mistakes in my linear extensions if any and I am left with the following as my linear extension:

$L1 = \{10, 13, 2, 8, 1, 3, 6, 12, 11, 9, 4, 7, 5\}$
$L2 = \{6, 7, 12, 11, 10, 9, 8, 1, 13, 4, 5, 2, 3\}$
$L3 = \{13, 11, 12, 2, 10, 9, 8, 6, 4, 1, 5, 7, 3\}$

3 Solution to 3

3a) The coefficient of x^{12} in the expansion of $(2x^2+3)^{20}$ can be determined using the binomial theorem. Given the expansion $(a+b)^n$, the general term is given by $\binom{n}{k}a^{n-k}b^k$, where k is the term number.

$$(x+y)^n = \binom{n}{0}y^n x^0 + \binom{n}{1}y^{n-1}x^1 + \binom{n}{2}y^{n-2}x^2 + \ldots + \binom{n}{n}y^0 x^n$$

$$(x+y)^n = \sum_{i=0}^{n}\binom{n}{i}y^{n-i}x^i$$

For the expression $(2x^2+3)^{20}$, where $a = 2x^2$ and $b = 3$, and $n = 20$, we are interested in the term that contains x^{12}. This corresponds to raising $2x^2$ to the 6th power, as $2 \times 6 = 12$, which means $k = 6$.

Thus, the term that contributes to x^{12} is:

$$\binom{20}{6}(2x^2)^6(3)^{14}$$

To find the coefficient of x^{12}, we calculate:

$$\binom{20}{6} \cdot (2)^6 \cdot (3)^{14}$$

Substituting the values and simplifying gives us the coefficient of x^{12}.

The final answer is: $38760 \times 64 \times 4782969 = 1.1864824 \times 10^{13}$

3b) From the question we know that all the possibilities vary from 1 to 10 to the power of 20. This information allows us refer to a fixed-width representation of 20 digits for each number. But, when we are counting distinct integers with a specified range, leading zeros do not create unique values meaning '090', '0090' all represent '90'

To find out if there are more numbers containing the digit 9 or more with no 9s. We have to figure that combinatorially aka by counting.

- **Without the Digit 9**: The calculation of 9^{20} takes into account all possible ways to fill each of 20 positions with digits from 0 to 8. However, this method doesn't account for leading zeros because relies on the position's digit choice, not the representation of the number in decimal notation.
- **With At Least One Digit 9**: Therefore we can find the number that must have at least one digit 9 by subtracting the total number of possible distinct digits from the number of ways of filling the 20 positions without the digit 9. We have this to be 10^{20} - 9^{20}. Thus this gives us total number that must have at least one 9 and it considers all the possible combinations where 9 appears at least once.

Therefore, the quantity of items without the digit 9 is given by 8.7842335×10^{19}. Conversely, the quantity of items containing at least one digit 9 is represented as 9^{20}.

It can be deduced that the set containing at least one digit 9 not only has a higher exponent but also a larger base. Hence, there are more items that do not contain the digit 9 compared to those that do.

3c)To determine the size of the relation R on $X = 2^S$, where $S = \{1, 2, 3, \ldots, n\}$ and X is the power set of S, and R is defined such that $(A, B) \in R$ if and only if A is a subset of B, we can approach this problem by considering the nature of the power set and the subset relation.

Given a set S with n elements, the power set X of S contains 2^n elements, because each element of S can either be included or not included in any subset, leading to 2^n possible subsets (including the empty set and the set S itself).

For any given element A in the power set X, the number of subsets B (including A itself) for which $A \subseteq B$ depends on the number of elements not in A but in S. If A has k elements, then there are $n - k$ elements not in A. Each of these elements can either be included or not included in B, leading to 2^{n-k} possibilities for B.

To find the size of R, we sum over all subsets A of S, considering all possible sizes of A from 0 to n (inclusive):

- For a subset of size 0 (the empty set), there are 2^n choices for B (since any subset, including S itself, is a superset of the empty set).

- For a subset of size 1, there are n such subsets, and for each, there are 2^{n-1} choices for B.
- For a subset of size 2, there are $\binom{n}{2}$ such subsets, and for each, there are 2^{n-2} choices for B.
- This pattern continues up to a subset of size n (the set S itself), for which there is 1 choice for A and $2^{n-n} = 1$ choice for B (only S itself).

To add a more detailed example for the set $\{1, 2, 3\}$, the sum of each of the subsets that match the criteria in the question is $\binom{n}{0} \cdot 2^n + \binom{n}{1} \cdot 2^{n-1} + \binom{n}{2} \cdot 2^{n-2} + \binom{n}{3} \cdot 2^{n-3}$.

This pattern of $\binom{n}{0} \cdot 2^n + \binom{n}{1} \cdot 2^{n-1} + \cdots + \binom{n}{n} \cdot 2^{n-n}$.

Leads us to $\sum_{i=0}^{n} \binom{n}{i} \cdot 2^{n-i}$.

From our knowledge of binomial theorem's this is exactly equivalent to the binomial general term and summation function for binomial items.

$$\sum_{i=0}^{n} \binom{n}{i} y^{n-i} x^i = (x+y)^n \tag{1}$$

and

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} y^{n-i} x^i$$

specifically our x = 1 and our y = 2,

we plug this back in and we get:

$$(1+2)^n = \sum_{i=0}^{n} \binom{n}{i} 2^{n-i} 1^i$$

This results in $\mathbf{3^n}$, which is the closed form.
The size of the relation R on X, given the condition that $(A, B) \in R$ if and only if A is a subset of B, simplifies to 3^n under the condition that n is a real number. This result stems from considering all possible subsets A and B in the power set X of S and the fact that for each element in S, there are three possibilities regarding its presence in A and B: it can be in both A and B, in B but not in A, or in neither (which still counts towards the total because we're considering the relation of A being a subset of B

Therefore, the size of R is 3^n, reflecting the combinatorial explosion of possibilities when considering the subset relation across the power set of a set with n elements. This result illustrates the vast number of subset-superset pairs within the structure of a power set, showcasing the combinatorial richness of such relations.

4 Solution to 4

4a)

Combinatorial Proof

The identity can be understood combinatorially by interpreting both sides as counting the same quantity in two different ways.

Left-hand Side: $n \cdot 2^{n-1}$

It can be interpreted as the total number of ways to select a committee from n people and choose a leader, with the restriction that each person has a 50% chance of being on the committee but the leader must be on the committee. This is because:

- 2^{n-1} comes from considering each of the $n-1$ other members not designated as the leader, they can either be on the committee or not, independently, resulting in 2^{n-1} combinations.
- Multiplying by n allows for each person to potentially be the leader, ensuring the leader is always on the committee.

This means we are counting all the possible committees where one person is the leader and the rest of the $n-1$ people can either join or not join the committee.

Right-hand Side: $\sum_{k=1}^{n} k\binom{n}{k}$

The right-hand side counts the same scenario but structures the counting differently:

- $\binom{n}{k}$ counts the number of ways to choose a k-member committee from n people.
- Multiplying by k then accounts for choosing one of these k members as the leader.

This sum, therefore, counts the number of ways to form a committee of any size from 1 to n and then select one of the committee members as the leader. It sort of reduces the problem into smaller subgroups and by the addition principle we know we can choose a final leader from either of the smaller subgroups which is what the + sign there means.

Final Proof Statement

The left-hand side and the right-hand side count the same things differently, specifically, they count the total number of ways to form a committee from n people and appoint one of them to be the leader. The left-hand side does this by choosing the leader and then forming the committee, while the right side does the calculation by forming the committee and choosing the leader. Both methods end up having the same sum and end up being equal.

4b) Combinatorial proof for:

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$$

like we were taught, involves interpreting both sides to mean counting the number of ways to do one thing using two different combinatorial processes.

Left-hand Side: $\binom{2n}{n}$

The left-hand side, $\binom{2n}{n}$, counts the number of ways to choose n items out of a set of $2n$ items. Imagine you have two groups of n items each (for a total of $2n$ items), and you want to select n items without caring about which group they come from. A more vivid example is imagine we have 12 pairs of the same pattern of socks so 2n, and we want to pick 12 of them, since whichever group we choose from does not matter we can do this by saying out of 2n choose n.

Right-hand Side: $\sum_{k=0}^{n} \binom{n}{k}^2$

The right-hand side sums the squares of binomial coefficients, $\binom{n}{k}^2$, for k ranging from 0 to n. This involves counting the number of ways to select k items from the first group of n items and $n-k$ items from the second group of n items, ensuring that the total number of items selected is n. The square comes from the fact that you are independently choosing k items from each of two groups of n items, and you sum these counts for all possible values of k. I also think og it as a way of partitioning total elements into two cases or subsets, which means choosing k elements from first n, elements and then n-k elements from second group and application of multiplication principle here leads to the squaring up.

Final Proof Statement

Imagine dividing $2n$ items into two distinct groups, A and B, each containing n items. You want to select a total of n items from these $2n$ items.

- On the left-hand side, you are not distinguishing between the groups when selecting n items from the total $2n$ items. Thus, $\binom{2n}{n}$ directly counts the number of ways to do this.

- On the right-hand side, you consider the contributions from all possible ways of splitting your selection between the two groups. For any given k, where $0 \leq k \leq n$, you choose k items from group A (which can be done in $\binom{n}{k}$ ways) and $n-k$ items from group B (which can also be done in $\binom{n}{n-k}$ ways, or equivalently $\binom{n}{k}$ ways, due to symmetry). Since the selections from groups A and B are independent, you multiply the counts, resulting in $\binom{n}{k}^2$ for each k. The idea involves counting differently.

- Summing over all possible values of k from 0 to n accounts for all ways of distributing the selection of n items between the two groups, which must equal the total number of ways of selecting n items from the combined $2n$ items without regard to the group.

The two sides count the same thing, which is the number of ways to select n items from a set of 2n.

4c) Combinatorial proof for the identity

$$\binom{r}{r} + \binom{r+1}{r} + \binom{r+2}{r} + \ldots + \binom{n}{r} = \binom{n+1}{r+1},$$

we interpret both sides as counting the number of ways to select a subset of size $r+1$ from a set of $n+1$ elements, but from different perspectives.

Combinatorial Interpretation

In a set S of $n+1$ distinct elements. We want to count the number of subsets of S that contain exactly $r+1$ elements.

Right-hand Side: $\binom{n+1}{r+1}$

The right-hand side counts the number of ways to choose a subset of size $r+1$ from $n+1$ elements, which defines the binomial coefficient $\binom{n+1}{r+1}$.

Left-hand Side: $\sum_{k=r}^{n} \binom{k}{r}$

The left-hand side counts the same subsets but with a specific element in mind. Let's designate one element of S as a special element, say x. For any subset of size $r+1$ that includes x, we can think of removing x to be left with a subset of size r. The question then becomes: From which subset of the original n elements (excluding x) was this r-sized subset chosen?

- $\binom{r}{r}$ counts the number of ways to choose r elements from the first r elements, ensuring that the subset includes x.
- $\binom{r+1}{r}$ counts the number of ways to choose r elements from the first $r+1$ elements, and so on, up to
- $\binom{n}{r}$ which counts the ways to choose r elements from all n elements (excluding x).

Each term $\binom{k}{r}$ (for $k = r, r+1, \ldots, n$) represents the number of $r+1$-sized subsets of S that include x and r elements from the first k elements of S, where x is considered to be the $(k+1)$th element.

Final Proof Statement

The sequence of choices on the left-hand side exhaustively and uniquely covers all possible $r+1$-sized subsets of S that include x, by considering all possible subsets of the remaining n elements that can be paired with x. Since every $r+1$-sized subset of S must include x and r elements from the remaining n elements, the left-hand side counts the same set of subsets as the right-hand side, but it does so by partitioning the count based on the subsets of the first k elements for each k from r to n.

Since both sides count the same thing, which is the total number of $r+1$-sized subsets of a set with $n+1$ elements, but in different ways. We can deduce they are equal and this ends the combinatorial proof of the given identity.