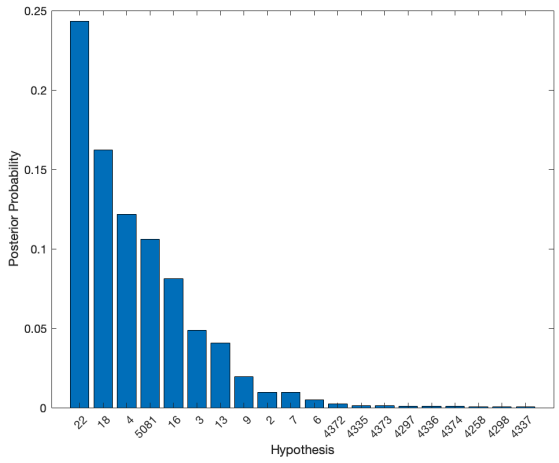
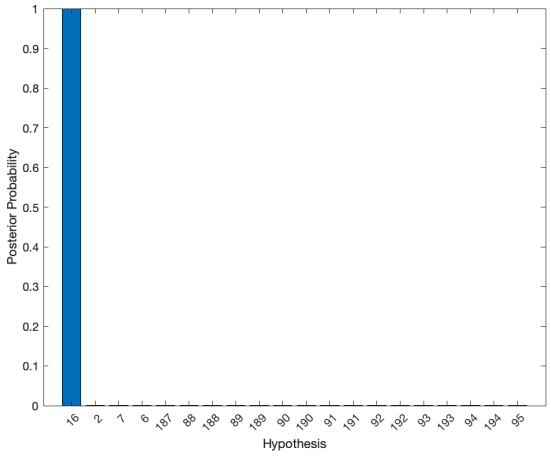
1. **Concept Learning Problem**
    a. Solution:
        i. Trial 1: [64]



Because the model is only fed one example, there is a possible range of concepts that the model infers. In this case, it learns all possible concepts of the number 64 (i.e., power of 8, power of 4, cube of 4, etc.) ranked by the size of each hypothesis. For example, the size principle makes hypothesis 22 (power of 8) most probable and makes hypothesis 7 (multiple of 2) less likely to be chosen.

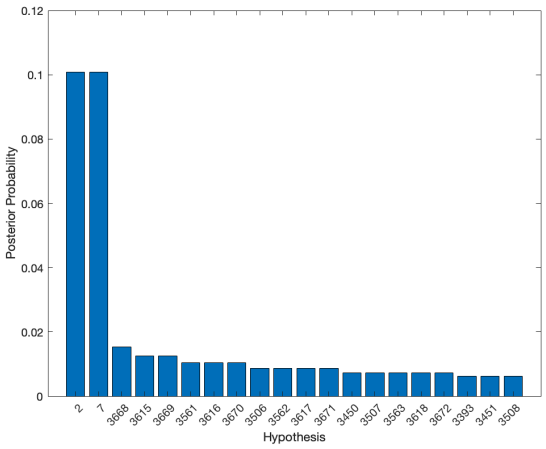
        ii. Trial 2: [8, 2, 16, 64]



For trial 2, the inferred concept is hypothesis 16 (power of 2). The high model certainty captures the underlying idea that a hypothesis with a large size is much less likely to be inferred when the sample increases in length.

In other words, although it is possible to assume that the sequence of numbers is even, there is a higher degree of specificity by inferring that they are all powers of 2 (i.e., smaller hypothesis space).

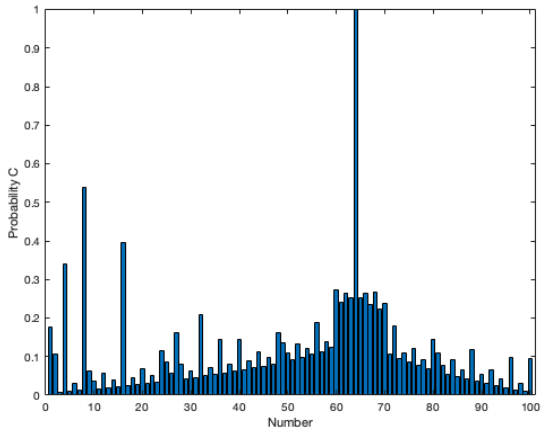
iii. Trial 3: [48, 64, 60, 72, 66]



Here, the model is indecisive between hypotheses 2 (even numbers) and 7 (multiples of 2), which have the same priors and likelihoods. Other plausible concepts include raw magnitudes that include intervals that contain the sequence of numbers demonstrated above. The raw magnitude hypotheses are approximately uniformly distributed, suggesting that each one is equally likely.

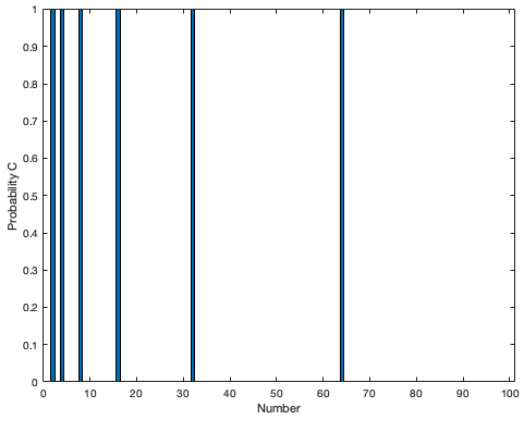
b. Solution:

i. Trial 1: [64]



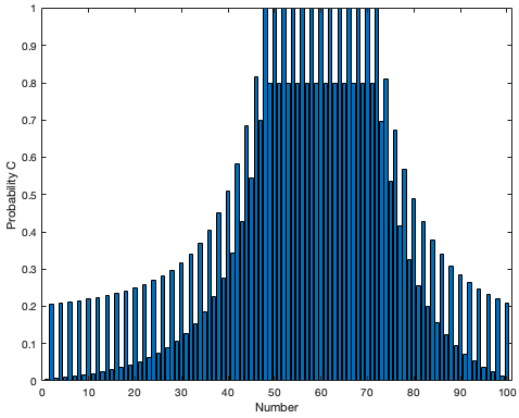
On seeing just the number 64 during trial 1, the model is unsure whether the learned concept concerns just numbers that are close to the positive example (graded) or rule-based properties (e.g., power of 4). It is more likely that the model generalizes to numbers that are closer in rule-based similarity (e.g., 4, 8, and 16), but the indecision reflects the limited sample size learned during the first trial. Thus, the model learns a diffuse similarity.

ii. Trial 2: [8, 2, 16, 64]



In trial 2, seeing more numbers that share the same properties of the number displayed in trial 1, the generalization appears ruled-based (all-or-none). Every number up to 100 that is a power of 2 (2, 4, 8, 16, 32, and 64) has the same deterministic probability (p = 1) of the learned concept in trial 2.

iii. Trial 3: [48, 64, 60, 72, 66]



Here, the model achieves a similarity-based (graded) generalization judgment. Numbers that are even (also multiples of 2) and closer in distance to the sequence of numbers have a higher probability of instantiating the same concepts learned during trial 3.