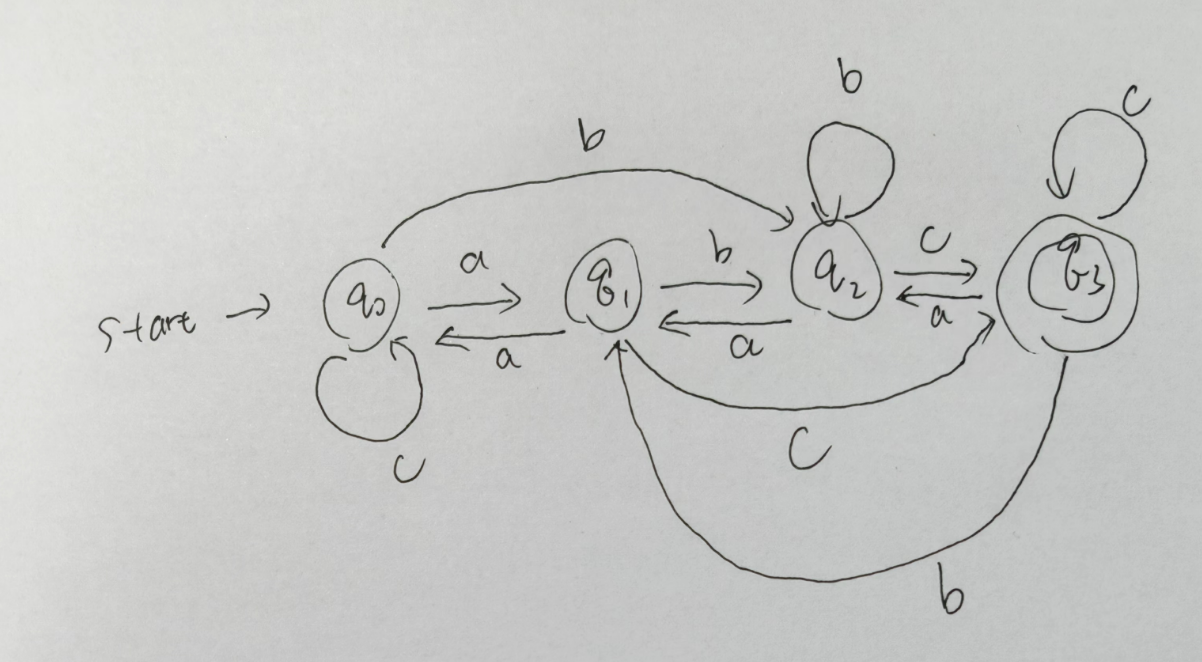
Please answer in Chinese or English.

Question 1: Common Regular Expressions

Question:  
Please list and explain the following common regular expression patterns:  
1. A regular expression that matches email addresses from 163 or 126, where the part before @ can be a digit or an uppercase or lowercase letter.  
2. A regular expression that matches at least one digit.  
3. A regular expression that matches words that start with a letter and are 5 characters long.

Question 2: DFA Diagram and its Definition

Question:  
Given the following DFA diagram, please write the DFA definition (five-tuple):  


Question 3: Differences between Pushdown Automata and DFA

Question:  
Please briefly explain the main differences between Pushdown Automata (PDA) and Deterministic Finite Automata (DFA).

Question 4: N-gram Analysis and Probability Calculation

Question:  
Given the following complete corpus:  
1. "I love natural language processing"  
2. "language processing is fun"  
3. "natural language understanding is important"  
  
Please:  
1. List all the 2-grams and 3-grams in the corpus.  
2. Calculate the probability of the 2-gram P(is | processing) and the 3-gram P(processing | natural language).

Question 5: Good-Turing Smoothing Probability Calculation

Question:  
In a given corpus, the frequency of words is as follows:  
- 100 words appear exactly once.  
- 60 words appear exactly twice.  
- 30 words appear exactly three times.  
  
Please:  
1. Calculate the probability of words that have never been seen.  
2. Calculate the smoothed probability for words that appear once and twice.

Question 6: Zero Probability Problem and Unseen Problem

Question:  
Please explain the difference between the "Zero Probability Problem" and the "Unseen Problem" in N-gram models, and briefly describe the solutions to each.