Problem Description

Draw the DFA accepting the language:

\{ s \mid 'ab' appears in s exactly 2 times \}

Solution:

Problem Syntactic Mistake

Feedback: The correct language is \{ s \mid 'ab' appears in s exactly 2 times \}

Grade: 5/10

Solution Syntactic Mistake

Feedback: One more state should be made final

Grade: 9/10

Technique

1. Synthesize a logic (MSO) description of the student attempt and the problem solution (brute force and pruning)
2. Compute tree edit distance between the two descriptions to produce grade
3. Use the logic descriptions and the edit script to produce the feedback (highlight edits)

Technique

Compute size of the set of misclassified string:
- \( S = \text{correct solution} \)
- \( A = \text{student attempt} \)
- Difference: \( D = S \setminus A \cup A \setminus S \)
- Size(\( D,S \)) = \( \lim_{n \to \infty} \frac{D^n}{S^n} \)
- Approximate to finite \( n \)
- Feedback with counterexample in \( D \)

Experimental Results

- Compared with human graders on 800 real student attempts
- Identical solutions receive same grades and correct attempts awarded max score (unlike human)
- 90% cases consistent with human grader (+/- 3 points)
- On disagreeing cases, human grader often realized that his assigned grade was inaccurate after reading tool’s feedback
- Always assigns full score to correct solutions (unlike human)
- Tool limited to small DFAs (< 10 states) and small alphabets (< 3 symbols).
  Not a big limitation in practice

Ongoing work

- Evaluation of quality of feedback
- Test the tool on Automata Theory courses in Fall at Penn and UIUC
- Grading and Feedback for
  - Regular expressions,
  - NFAs,
  - ….