

### C Final Problem Set and Struggling-Level Classification

The final set of 37 Java programming problems used in our experiments is listed below. Problems were classified into Low, Medium, and High struggling levels based on the total number of human student submissions in the original dataset, serving as a proxy for student struggle. The problems were sorted by submission count and then evenly divided into three tiers.

Table 3. List of the 37 Finalized Problems and the Students' Struggling Levels

Problem Name	Struggling Level
countevens	Low
sortasum	Low
luckysum	Low
squirrelplay	Low
cigarparty	Low
centeredaverage	Low
fizzbuzz	Low
noteensum	Low
datefashion	Low
repeatend	Low
tenrun	Low
answercell	Low
love6	Low
zeromax	Medium
has77	Medium
endother	Medium
sum67	Medium
withoutten	Medium
in1to10	Medium
countcode	Medium
xybalance	Medium
mixstring	Medium
roundsum	Medium
countclumps	Medium
fix45	Medium
bobthere	Medium
alarmclock	High
evenlyspaced	High
greenticket	High
zipzap	High
getsandwich	High
xyzthere	High
caughtspeeding	High
plusout	High
without2	High
makechocolate	High
blackjack	High