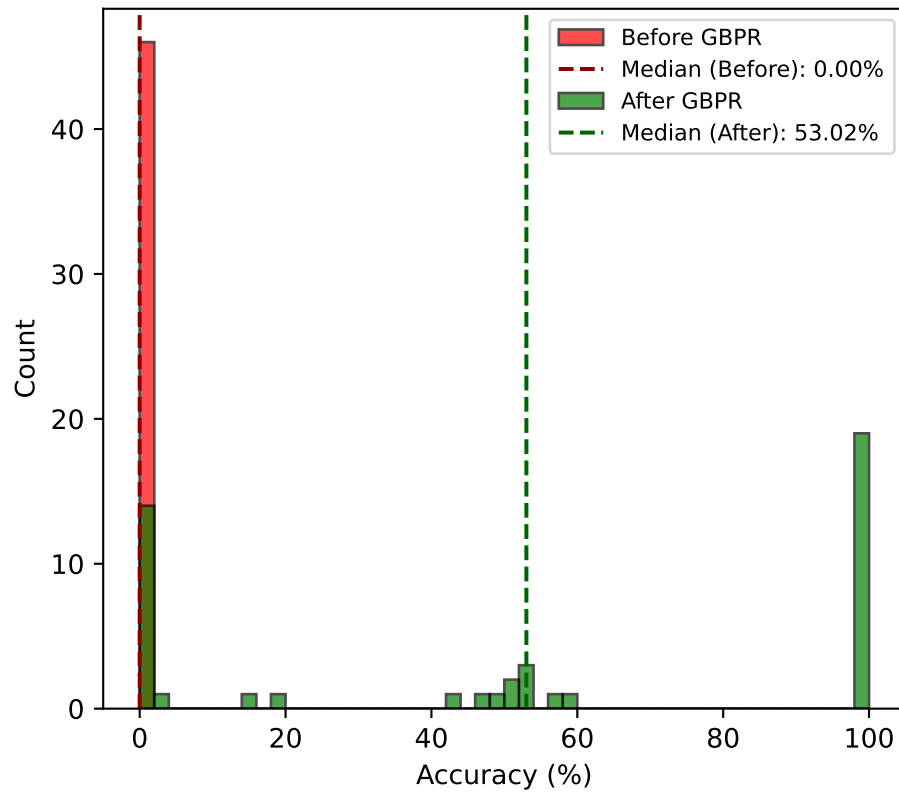
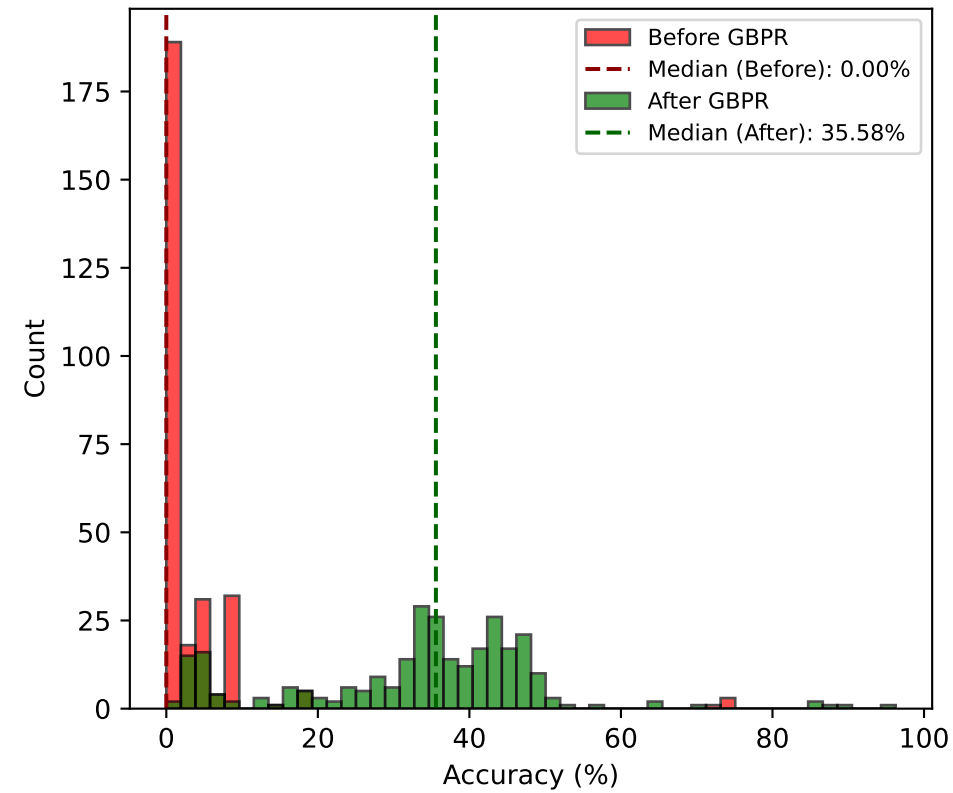


Distribution of Correctness Accuracy by Program (Before and After GBPR)

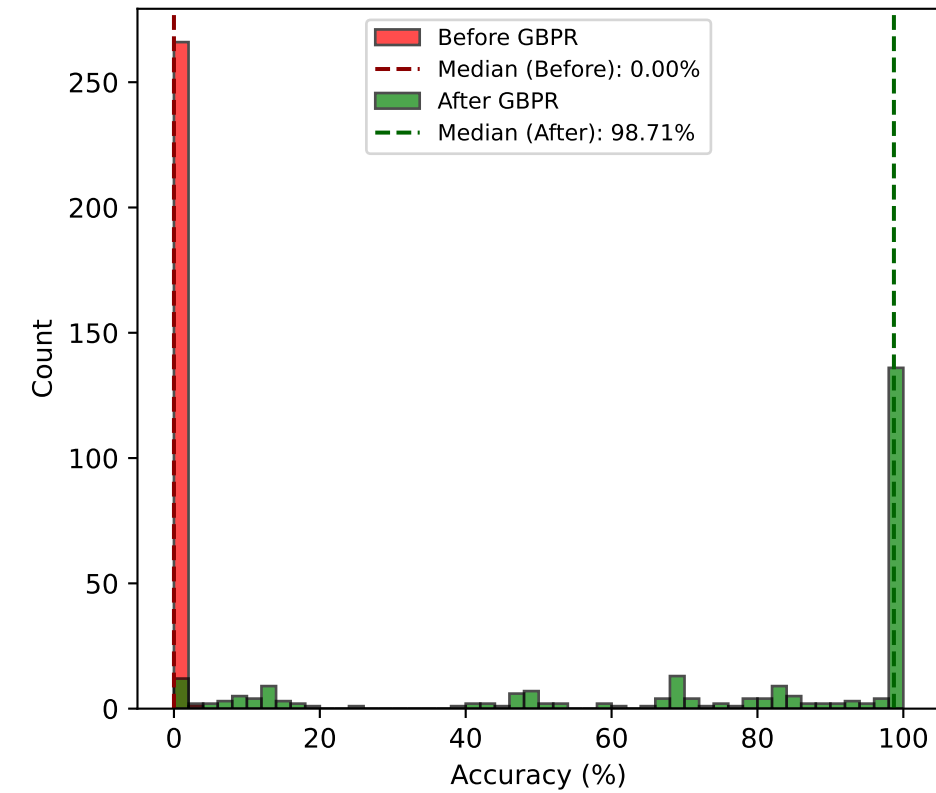
hist
(Models: Before=46, After=46)



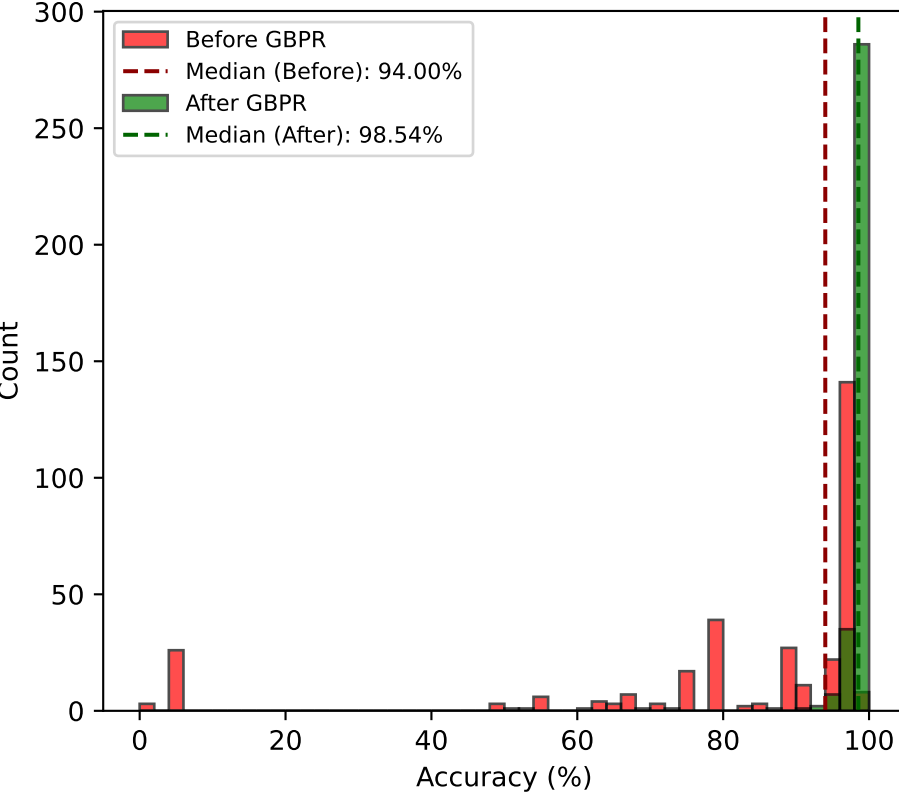
most_freq
(Models: Before=284, After=284)



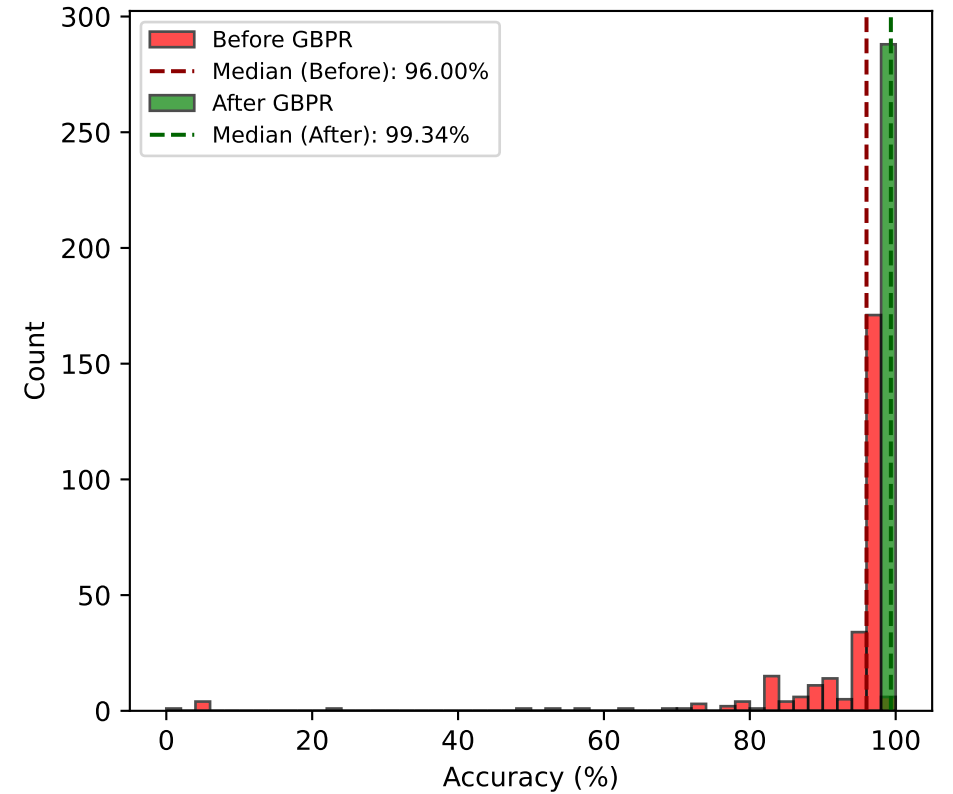
reverse
(Models: Before=268, After=268)



shuffle_dyck
(Models: Before=331, After=331)



shuffle_dyck2
(Models: Before=288, After=288)



sort
(Models: Before=249, After=249)

