In addition to perturbations to strong and weak students, perturbation to intermediate students is also performed. The filter is given as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | Strong | Intermediate | Weak |
| topic\_fam |  |  |  |
| correct\_so\_far |  |  |  |
| avg\_time |  |  |  |

The feature topic\_fam represents the familiarity of previously chosen part of a topic by students; the feature correct\_so\_far is the quantized ratio of correct responses to total number of responses; and avg\_time measures the average time taken by students on each topic. Here and represent current state and next values respectively. For strong and intermediate students, we would expect them to be more familiar to the topic after some learning engagement while for weak students, the learning effect may be negligible. In addition, strong students are more likely to give a correct response and spend less time in learning as their familiarity grows.

The reward is calculated in such a way that correct responses on harder questions attain higher rewards (1 to 4) while incorrect responses on easier questions attain higher penalties (-8 to -2). Since in general students would get more reward than penalty, the penalty is scaled from [-4,-1] to [-8,-2].

The cumulative reward from initial states across 1000 steps for various MDPs is shown as below:

图形用户界面, 图示

描述已自动生成

From the graphs, we can see that the original MDP always has the best performance because it is where the policy is derived. When only perturbing three features, weaker students are affected more. As shown by the top-left graph, there is a big difference between reward of weak students and reward of intermediate students. However, as more features are included, the effect starts to diminish. The bottom-right graph shows that even with all features perturbed, the performance of policy is not affected. This means that the policies would be robust in the real-world setting where students have different learning styles. One drawback is that including more features inevitably increases the cost of computation. A policy that favors weaker students would be desired.