1 Introduction

Over recent years, the importance of the international SAT competition has grown to be an annual event in the community. The major impact of being ranked among the best solvers is beneficial both for academic and industrial communities. As a consequence, the scoring scheme of the competition needed some formal basis.

The method described in this paper is designed to overcome some of the drawbacks observed in earlier methods. The primary difficulty is that, because the underlying problem requires exponential time in practice, one must either set very easy problems to be sure all solvers can succeed, or one must allow for the fact that some solvers will not succeed on some instances. It is commonly agreed that the first alternative does not lead to interesting outcomes.

The paper outline is as follows. After presenting the design objectives and discussing drawbacks with current approaches, we describe the proposed method that was decided upon. Some properties of this new method are discussed. Then we take some examples from stage one of the SAT 2005 competition to illustrate how the scoring scheme works and how the rankings would change if alternative ranking schemes were used. Preliminary experimental results are presented for the first stage of the SAT 2005 Competition, involving about 30 solvers and hundreds of benchmark instances. The paper concludes with a brief discussion of the critical issues regarding the new scoring scheme and provides a first assessment on how it can be improved.

2 Design Objectives

One key idea behind the SAT competition is to award a solver that is good on a wide range of SAT instances. In the previous years of the competition, this was implemented using a scoring scheme that ranked the solvers with a tiered weighting between solving and speed.

TimeUsed = \sum_{p} \text{timeUsed}(p)

However, the time needed to solve a given benchmark also needs to be considered. While the CPU time was indirectly applied separately within each combination of category and specialty.

spdM = \text{speedP}(p) / \text{timeUsed}(p)

Another key idea of the competition was to focus on solvers that are the best performers first, best performers last. (A)

The major impact of being ranked among the best solvers is beneficial both for academic and industrial communities. The scoring scheme used for the SAT 2005 competition is designed to address these issues. It incorporates these features:

- It rewards the rankings of the solvers at the end of the competition.
- It gives more credit for solving hard benchmarks than solving easy ones.
- It gives more credit for solving a benchmark fast.
- It gives extra credit for each solved problem.

3 The Purse-Based Scoring System

The implemented scoring plan works as follows. A purse is defined to be the execution of one solver on one benchmark instance, or problem. Each run is allocated a certain amount of CPU time. If the solver succeeds, it is rewarded the purse.

Each problem has a speed purse, which is divided equally among all competition solvers that solve the problem. The speed purse is a fixed multiple (size \(Nbr\) of solvers) of the standard purse (stdP).

speedAward = \text{speedP}(p) / D4

Each problem has a slow solve purse (slowP), which is the execution of one solver that did not solve the problem.

slowP = \text{timeUsed}(p) / D4

4 Discussion

The new scoring scheme and particularly some of its features is a first shot. After the competition they most likely will need to be adjusted. The general goal should be to advance the state-of-the-art of SAT solvers. There are multiple contradictory interpretations into what means speed on specific instances versus robustness versus efficiency.

The formula to divide the speed purse of a problem is the following, where \(p\) is problem-id and \(\text{timeUsed}(p)\) is defined to be the execution of one solver that solved the problem.

speedAward = \text{speedP}(p) / \text{timeUsed}(p)

4.4 Speed Purse

The speed purse is a fixed multiple (size \(Nbr\) of solvers) of the standard purse (stdP). All problems have the standard solution purse (stdP).

stdP = \frac{\text{purse}}{D4}

5 Preliminary Experimental Results

These tables present the results of stage one for the SAT 2005 Competition.

<table>
<thead>
<tr>
<th>Solver</th>
<th>Score</th>
<th>Sat</th>
<th>Unsat</th>
<th>Nbr Solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>solver10</td>
<td>1501.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver11</td>
<td>1502.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver12</td>
<td>1503.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver13</td>
<td>1504.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver14</td>
<td>1505.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver15</td>
<td>1506.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver16</td>
<td>1507.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver17</td>
<td>1508.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver18</td>
<td>1509.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>solver19</td>
<td>1510.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

The tables present the results of stage one for the SAT 2005 Competition.