

DOMjudge Jury Manual

by the DOMjudge team

Sat, 11 Jun 2016 14:38:14 -0300

This document provides information about DOMjudge aimed at a jury member operating the system during the contest: viewing and checking submissions and working with clarification requests. A separate manual is available for teams and administrators. Document version: a717e98

Contents

1	DOMjudge overview	3
1.1	Features	3
1.2	Copyright and licencing	3
1.3	Contact	4
2	General	5
2.1	Judges and Administrators	5
2.2	Scoreboard	5
3	Before the contest	6
3.1	Problems and languages	6
3.2	Verifying testdata	6
3.3	Testing jury solutions	6
3.4	Practice Session	7
4	During the contest	8
4.1	Monitor teams	8
4.2	Judging Submissions	8
4.3	Clarification Requests	10
5	After the contest	12
A	Problem package format	13

1 DOMjudge overview

DOMjudge is a system for running programming contests, like the ACM ICPC regional and world championship programming contests.

This means that teams are on-site and have a fixed time period (mostly 5 hours) and one computer to solve a number of problems (mostly 8-11). Problems are solved by writing a program in one of the allowed languages, that reads input according to the problem input specification and writes the correct, corresponding output.

The judging is done by submitting the source code of the solution to the jury. There the jury system automatically compiles and runs the program and compares the program output with the expected output.

This software can be used to handle the submission and judging during such contests. It also handles feedback to the teams and communication on problems (clarification requests). It has web interfaces for the jury, the teams (their submissions and clarification requests) and the public (scoreboard).

1.1 Features

A global overview of the features that DOMjudge provides:

- Automatic judging with distributed (scalable) judge hosts
- Web interface for portability and simplicity
- Modular system for plugging in languages/compiler and validators
- Detailed jury information (submissions, judgings, diffs) and options (rejudge, clarifications, resubmit)
- Designed with security in mind
- Has been used in many live contests
- Open Source, Free Software

1.2 Copyright and licencing

DOMjudge is developed by Jaap Eldering, Nicky Gerritsen, Keith Johnson, Thijs Kinkhorst and Tobias Werth; Peter van de Werken has retired as developer. Many other people have contributed (apologies for any oversights): Jeroen Bransen, Stijn van Drongelen, Rob Franken, Jacob Kleerekoper, Ruud Koot, Jan Kuipers, Richard Lobb, Alex Muntada, Dominik Paulus, Jeroen Schot, Matt Steele, Shuhei Takahashi, and Jeroen van Wolffelaar. Some code has been ported from the ETH Zurich fork by Christoph Krautz, Thomas Rast et al.

DOMjudge is Copyright (c) 2004 - 2016 by the DOMjudge developers and its contributors.

DOMjudge, including its documentation, is free software; you can redistribute it and/or modify it under the terms of the *GNU General Public License* <<http://www.gnu.org/copyleft/gpl.html>> as published by the Free Software Foundation; either version 2, or (at your option) any later version. See the file COPYING.

Additionally, parts of this system are based on other programs, which are covered by other copyrights. See the administrator's manual for details.

1.2.1 About the name and logo

The name of this judging system is inspired by a very important and well known landmark in the city of Utrecht: the dome tower, called the ‘Dom’ in Dutch. The logo of the 2004 Dutch Programming Championships (for which this system was originally developed) depicts a representation of the Dom in zeros and ones. We based the name and logo of DOMjudge on that.

We would like to thank Erik van Sebille, the original creator of the logo. The logo is under a GPL licence, although Erik first suggested a "free as in beer" licence first: you’re allowed to use it, but you owe Erik a free beer in case might you encounter him.

1.3 Contact

The DOMjudge homepage can be found at: <https://www.domjudge.org/>

We have a low volume [mailing list for announcements](#) of new releases.

The authors can be reached through the development mailing list: domjudge-devel@domjudge.org . You need to be subscribed before you can post. See [the list information page](#) for subscription and more details.

Some developers and users of DOMjudge linger on the IRC channel dedicated to DOMjudge on the Freenode network: server `irc.freenode.net`, channel `#domjudge`. Feel free to drop by with your questions and comments, but note that it may sometimes take a bit longer than a few minutes to get a response.

2 General

The jury interface is accessed through a web browser. The main page shows a list of various overviews, and the most important of those are also included in the menu bar at the top. The menu bar will refresh occasionally to allow for new information to be presented. It also has the current ‘official’ contest time in the top-right corner.

Most pieces of information are clickable and bring up a new page with details. Many items also have tooltips that reveal extra information when the mouse is hovered over them. Problem, language and team pages have lists with corresponding submissions for that problem, language or team. Tables can be sorted by clicking on the column headers.

The most important pages are ‘Submissions’: the list of submitted solutions made by teams, sorted by newest first, and ‘Scoreboard’: the canonical overview of current standings.

2.1 Judges and Administrators

The DOMjudge system discerns between *judges* and *administrators* (admins). An administrator is responsible for the technical side of DOMjudge: installation and keeping it running. The jury web interface may be used by both, but depending on your assigned role you may have more options.

2.2 Scoreboard

The scoreboard is the most important view on the contest.

The scoreboard will display an upcoming contest from the given ‘activatetime’; the contest name and a countdown timer is shown. Only at the first second of the real start of the contest it will show the problems to the teams and public, however. The jury always has a full view on the scoreboard.

It is possible to freeze the scoreboard at a given time, commonly one hour before the contest ends, to keep that last hour interesting for all. From that time on, the public and team scoreboard will not be updated anymore (the jury scoreboard will) and indicate that they are frozen. It will be unfrozen at a specified time, or by a button click in the jury interface. Note that the way freezing works, a submission from before the freeze and judged after may still update the scoreboard even when frozen.

The problem headings can display the colours of balloons associated with them, when set.

Nearly everything on the scoreboard can be clicked to reveal more detailed information about the item in question: team names, specific submissions and problem headers. Many cells will show additional ‘title text’ information when hovering over them. The score column lists the number of solved problems and the total penalty time for each team. Each cell in a problem column lists the number of submissions, and if the problem was solved, this is followed by the time of the first correct submission in minutes since contest start. Any penalty time incurred for previous incorrect submissions is included in the team’s total time.

3 Before the contest

Before the contest starts, a number of things will need to be configured by the administrator. You can check that information, such as the problem set(s), test data and time limits, contest start- and end time, the time at which the scoreboard will be frozen and unfrozen, all from the links from the front page.

Note that multiple contests can be defined, with corresponding problem sets, for example a practice session and the real contest.

3.1 Problems and languages

The problem sets are listed under ‘Problems’. It is possible to change whether teams can submit solutions for that problem (using the toggle switch ‘allow submit’). If disallowed, submissions for that problem will be rejected, but more importantly, teams will not see that problem on the scoreboard. Disallow judge will make DOMjudge accept submissions, but leave them queued; this is useful in case an unexpected problem shows up with one of the problems. Timelimit is the maximum number of seconds a submission for this problem is allowed to run before a ‘TIMELIMIT’ response is given (to be multiplied possibly by a language factor). Note that a ‘timelimit overshoot’ can be configured to let submissions run a bit longer. Although DOMjudge will use the actual limit to determine the verdict, this allows judges to see if a submission is close to the timelimit.

Problems can be imported and exported into and from DOMjudge using zip-files that contain the problem metadata and testdata files, based on the `problemarchive.org` format. See appendix A (Problem package format -specification) for details. Problems can have special *compare* and *run* scripts associated to them, to deal with problem statements that require non-standard evaluation. For more details see the administrator’s manual.

The ‘Languages’ overview is quite the same. It has a timefactor column; submissions in a language that has time factor 2 will be allowed to run twice the time that has been specified under Problems. This can be used to compensate for the execution speed of a language, e.g. Java.

3.2 Verifying testdata

For checking whether the your testdata conforms to the specifications of your problem statement, we recommend the `checktestdata` program, which is available from a [separate repository](#) . It allows you to not only check on simple (spacing) layout errors, but a simple grammar file must be specified for the testdata, according to which the testdata is checked. This allows e.g. for bounds checking.

This program is built upon the separate library `libchecktestdata.h` that can be used to write the syntax checking part of special compare scripts: it can easily handle the tedious task of verifying that a team’s submission output is syntactically valid, leaving just the task of semantic validation to another program.

3.3 Testing jury solutions

Before a contest, you will want to have tested your reference solutions on the system to see whether those are judged as expected and maybe use their runtimes to set timelimits for the problems.

The simplest way to do this is to include the jury solutions in a problem zip file and upload this. You can also upload a zip file containing just solutions to an existing problem. Note that the zip archive has to adhere

to the [Kattis problem package format](#) . For this to work, the jury/admin who uploads the problem has to have an associated team to which the solutions will be assigned. The solutions will automatically be judged if the contest is active (but it need not have started yet). You can verify whether the submissions gave the expected answer from the link on the jury/admin index page.

3.4 Practice Session

If your contest has a test session or practice contest, use it also as a general rehearsal of the jury system: judge test submissions as you would do during the real contest and answer incoming clarification requests.

4 During the contest

4.1 Monitor teams

Under the Teams menu option, you can get a general impression of the status of each team: a traffic light will show either of the following:

gray

the team has not (yet) connected to the web interface at all;

red

the team has connected but not submitted anything yet;

yellow

one or more submissions have been made, but none correct;

green

the team has made at least one submission that has been judged as correct.

This is especially useful during the practice session, where it is expected that every team can make at least one correct submission. A team with any other colour than green near the end of the session might be having difficulties.

4.2 Judging Submissions

4.2.1 Flow of submitted solutions

The flow of an incoming submission is as follows.

1. Team submits solution. It will either be rejected after basic checks, or accepted and stored as a *submission*.
2. The first available *judgehost* compiles, runs and checks the submission. The outcome and outputs are stored as a *judging* of this submission. Note that judgehosts may be restricted to certain contests, languages and problems, so that it can be the case that a judgehost is available, but not judging an available submission.
3. If verification is not required, the result is automatically recorded and the team can view the result and the scoreboard is updated (unless after the scoreboard freeze). A judge can optionally inspect the submission and judging and mark it verified.
4. If verification is required, a judge inspects the judging. Only after it has been approved (marked as *verified*) will the result be visible outside the jury interface. This option can be enabled by setting `verification_required` on the *configuration settings* admin page.

4.2.2 Submission judging status codes

The interface for jury and teams shows the status of a submission with a code.

QUEUED/PENDING

submission received and awaiting a judgehost to process it *;

JUDGING

a judgehost is currently compiling/running/testing the submission *;

TOO-LATE

submission submitted after the contest ended;

CORRECT

submission correct, problem solved;

COMPILER-ERROR

the compiler gave an error while compiling the program;

TIMELIMIT

program execution time exceeded the time defined for the problem;

RUN-ERROR

a kind of problem while running the program occurred, for example segmentation fault, division by zero or exitcode unequal to 0;

NO-OUTPUT

there was no output at all from the program;

WRONG-ANSWER

the output of the program did not exactly match the expected output;

* in the team interface a submission will only show as PENDING to prevent leaking information of problem time limits. The jury can see whether a submission is QUEUED or JUDGING. In case of required verification, a submission will show as PENDING to the team until the judging has been verified.

Under the Submissions menu, you can see submitted solutions, with the newest one at the top. Click on a submission line for more details about the submission (team name, submittime etc), a list of judgments and the details for the most recent judging (runtime, outputs, diff with testdata). There is also a switch available between newest 50, only unverified, only unjudged or all submissions. The default (coloured) diff output shows differences on numbered lines side by side separated by a character indicating how the lines differ: ! for different contents, \$ for different or missing end-of-line characters, and one of <> when there are no more lines at the end of the other file.

Under the submission details the ‘view source code’ link can be clicked to inspect the source code. If the team has submitted code in the same language for this problem before, a diff output between the current and previous submission is also available there.

It is possible to edit the source code and resubmit it if you have a team associated to your user. This does not have any effect for the teams, but allows a judge to perform a ‘what if this was changed’-analysis.

A submission can have multiple judgments, but only one valid judging at any time. Multiple judgments occur when rejudging, see 4.2.3 (Rejudging).

4.2.3 Rejudging

In some situations it is necessary to rejudge one or more submissions. This means that the submission will re-enter the flow as if it had not been judged before. The submittime will be the original time, but the program will be compiled, run and tested again.

This can be useful when there was some kind of problem: a compiler that was broken and later fixed, or testdata that was incorrect and later changed. When a submission is rejudged, the old judging data is kept but marked as ‘invalid’.

You can rejudge a single submission by pressing the ‘Rejudge’ button when viewing the submission details. It is also possible to rejudge all submissions for a given language, problem, team or judgehost; to do so, go to the page of the respective language, problem, team or judgehost, press the ‘Rejudge all’ button and confirm.

There are two different ways to run a rejudging, depending on whether the `create rejudging` button toggled:

1. Without this button toggled, an "old-style" rejudging is performed where the results are directly made effective.
2. When toggled, a "rejudging" set is created, and all affected submissions are rejudged, but the new judgments are not made effective yet. Instead, the jury can inspect the results of the rejudging (under the rejudging tab). Based on that the whole rejudging, as a set, can be applied or cancelled, keeping the old judgments as is.

Submissions that have been marked as ‘CORRECT’ will not be rejudged. Only DOMjudge admins can override this restriction using a tickbox.

Teams will not get explicit notifications of rejudgings, other than a potentially changed outcome of their submissions. It might be desirable to combine rejudging with a clarification to the team or all teams explaining what has been rejudged and why.

4.2.4 Ignored submissions

Finally, there is the option to *ignore* specific submissions using the button on the submission page. When a submission is being ignored, it is as if was never submitted: it will show strike-through in the jury’s and affected team’s submission list, and it is not visible on the scoreboard. This can be used to effectively delete a submission for some reason, e.g. when a team erroneously sent it for the wrong problem. The submission can also be unignored again.

4.3 Clarification Requests

Communication between teams and jury happens through Clarification Requests. Everything related to that is handled under the Clarifications menu item.

Teams can use their web interface to send a clarification request to the jury. The jury can send a response to that team specifically, or send it to all teams. The latter is done to ensure that all teams have the same information about the problem set. The jury can also send a clarification that does not correspond to a specific request. These will be termed ‘general clarifications’.

Under Clarifications, three lists are shown: new clarifications, answered clarifications and general clarifications. It lists the team login, the problem concerned, the time and an excerpt. Click the excerpt for details about that clarification request.

Every incoming clarification request will initially be marked as unanswered. The menu bar shows the number of unanswered requests. A request will be marked as answered when a response has been sent. Additionally it's possible to mark a clarification request as answered with the button that can be found when viewing the request. The latter can be used when the request has been dealt with in some other way, for example by sending a general message to all teams.

An answer to a clarification request is made by putting the text in the input box under the request text. The original text is quoted. You can choose to either send it to the team that requested the clarification, or to all teams. In the latter case, make sure you phrase it in such a way that the message is self-contained (e.g. by keeping the quoted text), since the other teams cannot view the original request.

The menu on every page of the jury interface will mention the number of unanswered clarification requests: "(1 new)". This number is automatically updated, even without reloading the page.

5 After the contest

Once the contest is over, the system will still accept submissions but these will receive the verdict 'TOO-LATE'. These submissions will still be judged and can be inspected by the jury, but they will not affect scoring and none of the judging details will be visible to the teams.

If the scoreboard was frozen, it will remain frozen until the time set as unfreeze time, as seen under Contests. It is possible to publish the final standings at any given moment by pressing the 'unfreeze now' button under contests.

There's not much more to be done after the contest has ended. The administrator will need to take care of backing up all system data and submissions, and the awards ceremony can start.

A Problem package format

DOMjudge supports the import and export of problems in a zip-bundle format.

The base of the format is Problem Format specification at problemformat.org . Please refer to that for the base specifications.

On top, DOMjudge defines a few extensions:

- `domjudge-problem.ini` (required): metadata file, see below.
- `problem.{pdf,html,txt}` (optional): problem statements as distributed to participants. The file extension determines any of three supported formats. If multiple files matching this pattern are available, any one of those will be used.

The file `domjudge-problem.ini` contains key-value pairs, one pair per line, of the form `key = value`. The `=` can optionally be surrounded by whitespace and the value may be quoted, which allows it to contain newlines. The following keys are supported (these correspond directly to the problem settings in the jury web interface):

- `probid` - the problem short name (e.g. "A")
- `name` - the problem displayed name
- `allow_submit` - allow submissions to this problem, disabling this also makes the problem invisible to teams and public
- `allow_judge` - allow judging of this problem
- `timelimit` - time limit in seconds per test case
- `special_run` - executable id of a special run script
- `special_compare` - executable id of a special compare script
- `points` - number of points for this problem (defaults to 1)
- `color` - CSS color specification for this problem

The `probid` key is required when importing a new problem from the `jury/problems.php` overview page, while it is ignored when uploading into an existing problem. All other keys are optional. If they are present, the respective value will be overwritten; if not present, then the value will not be changed or a default chosen when creating a new problem. Test data files are added to set of test cases already present. Thus, one can easily add test cases to a configured problem by uploading a zip file that contains only testcase files. Any jury solutions present will be automatically submitted when `allow_submit` is 1.