1. Title: First-order theorem proving

2. Sources:

   (a) James P Bridge, Sean B Holden and Lawrence C Paulson

   University of Cambridge
   Computer Laboratory
   William Gates Building
   15 JJ Thomson Avenue
   Cambridge CB3 0FD
   UK

   +44 (0)1223 763500
   forename.surname@cl.cam.ac.uk

   (b) Sean B Holden - details as (a).

   (c) Date: 17th April 2013

3. Past Usage:

   (a) Machine learning for first-order theorem proving: learning to select a good heuristic James P Bridge, Sean B Holden and Lawrence C Paulson Submitted for publication in the Journal of Automated Reasoning, Springer 2012/13.

   Please include a citation if you use this data.

   (b) We wish to predict which of a set of five heuristics will provide the fastest proof, given features derived from a theorem to be proved. A sixth possible prediction is to decline to attempt a proof, should the theorem be assessed as too difficult.

   (c) In prediction terms this is a challenging problem. However we can do better than any individual heuristic and obtain performance comparable to that of a hand-crafted selection mechanism employing around 75 addition heuristics. The ability to decline a proof is also beneficial.

4. Relevant Information Paragraph:

   Files:

   Expanding the tarball ml-prove.tar produces a directory ml-prove/ containing the files:

   ```
   all-data-raw.csv            - raw data used to derive training,
                                 validation and test data.
   all-data-raw-statistics.txt - min, max, mean and standard deviation
                                 for raw data. (Tabulated below.)
   train.csv                   - actual training, validation and test
   validation.csv                sets used.
   test.csv
   all-data-statistics.txt     - min, max and correlation data for
                                 combined actual data. (Tabulated below.)
   ```

   Raw data:

   Columns 1 to 14 are the static features and columns 15 to 53 are

the dynamic features. (See the paper for a description of static and dynamic features.) The final five columns denote the time in seconds taken by each of the five heuristics to prove the relevant theorem. There was a time limit of 100 seconds. An entry of -100 denotes failure to obtain a proof within the time limit. The first half of this data corresponds to the training data used. The second half was permuted and split to obtain the validation and test sets.

Training, validation and test data:

These are the sets used in the reported experiments. Two redundant features (static feature 5 and dynamic feature 21 in the raw data) were removed. The features in the training set are normalised to zero mean and unit variance. Validation/test data was normalized using the coefficients computed for the training set. Labels are in the final six columns. The first five of those correspond to the five heuristics (H1 to H5) and contain +1 if the corresponding heuristic found a proof and was the fastest to do so, and -1 otherwise. The final column contains +1 where no heuristic finds a proof within the time limit and -1 otherwise (H0 in the paper).

5. Number of Instances:

   6118 in the raw data.

   The training, validation and test sets have 3059, 1529 and 1530 respectively.

6. Number of Attributes:

   There are 13 static and 38 dynamic features for each instance. (See the paper for details regarding static/dynamic features. The raw data has two more features, which are redundant.) Columns 1 to 13 contain static features and columns 14 to 51 dynamic features.

7. Description of attributes:

   The full names for each attribute are provided in the paper, tables 2 and 3.

   Raw data: all attributes are numeric. Attributes 5, 9, 11, 13 and 35 are integer-valued. All other attributes are continuous.

   Training, validation and test data: all data are numeric and continuous on account of being normalized.

8. Missing Attribute Values:

   There are no missing values.

9. Class Distribution: number of positive instances in the sets for each heuristic (H1 to H5) and the "decline" option H0.

	H1	H2	H3	H4	H5	H0
Training set:	556	229	373	303	312	1286
Validation set:	260	133	187	146	159	644
Test set:	273	124	188	168	153	624

10. Attribute statistics:

Statistics for the raw data:

We do not include correlations as this data includes time measurements rather than actual classes.

Attribute	Min	Max	Mean	SD
1	0	1	0.36785	0.31389
2	0.0078125	1	0.83086	0.21303
3	0	1	0.26426	0.26827
4	0	1	0.30074	0.26399
5	0	0	0	0
6	0.00038153	1	0.35776	0.26964
7	0	0.98214	0.10619	0.1193
8	0	0.9966	0.53605	0.30238
9	1	244	7.0814	7.3035
10	1	39.07	2.64	2.2741
11	1	86	7.7257	7.3965
12	1	11	2.7553	0.93885
13	12	16240	159.76	318.14
14	4.4643	990.14	26.838	38.297
15	0.14141	1	0.8451	0.20275
16	0.048106	1477.7	14.238	42.392
17	0.0031504	0.97917	0.32239	0.17827
18	0.016667	7010	62.371	383.47
19	0.0040984	3	0.79035	0.3646
20	0.044484	3.7954	0.86357	0.30873
21	0.03125	5	1.0941	0.32019
22	0.21483	27.052	1.3219	0.58307
23	0.076923	7.5	0.98681	0.45421
24	0.29257	2.7525	1.0079	0.20582
25	0.23077	8.3333	1.3338	0.70145
26	0.54418	4.8414	1.1857	0.38731
27	0.0017567	13.647	0.82934	0.80024
28	0.01868	7.2062	0.86554	0.47486
29	0.051613	449.2	2.8181	14.844
30	0.22842	78.779	1.7701	2.4387
31	0	0.72727	0.04486	0.10227
32	0	0.85859	0.11004	0.16052
33	0.14141	1	0.8451	0.20275
34	0	7.7273	0.012384	0.14083
35	0	0	0	0
36	0	0.40404	0.0053444	0.0225
37	0	0.52525	0.02126	0.050914
38	0	5.404	0.029045	0.1183
39	0	142.7	4.0076	9.5463
40	0	176.01	7.0783	10.955
41	0	142.6	3.2732	8.7734
42	0	2.7778	0.024949	0.090029
43	0	142.7	3.9567	9.5333
44	0	1.8182	0.0098467	0.060589
45	0	7.8889	0.01567	0.14505
46	0	1	0.19265	0.33357
47	0	1	0.69969	0.29722
48	0	1	0.13459	0.25073
49	0	1	0.18258	0.32845
50	0	1	0.1277	0.26172
51	0	1	0.27041	0.32167
52	0	1	0.046002	0.10017
53	0	1	0.68359	0.32092

Statistics for the combined training, validation and test data:

We do not include mean and standard deviation as the data are normalized.

			Correlation with predicted attribute				
Attribute	Min	Max	H1	H2	H3	H4	H5	H0 (decline)
1	-1.1052	2.0094	0.002146	0.018808	0.053781	-0.05868	-0.055589	0.022243
2	-3.7356	0.83152	-0.0096346	-0.0083581	-0.10085	0.032679	-0.03394	0.079926
3	-0.98411	2.7381	0.064511	0.0044296	0.10732	-0.014341	-0.061767	-0.077094
4	-1.0652	2.6448	0.0035693	0.00027939	0.041497	-0.058916	-0.047575	0.034685
5	-1.2401	2.3662	0.012026	0.012048	0.15051	-0.073614	-0.069968	-0.028027
6	-0.88058	7.1945	0.012745	0.07273	0.049764	-0.0097244	-0.046281	-0.048485
7	-1.7638	1.4393	-0.015752	-0.039438	-0.15385	0.06948	0.080651	0.044121
8	-0.82637	32.27	0.0638	-0.034726	-0.016868	0.012842	0.046311	-0.055503
9	-0.70995	15.495	0.078303	0.0083392	-0.034436	0.002618	0.033225	-0.064424
10	-0.91358	10.654	-0.13423	-0.031277	-0.1219	-0.021194	-0.096237	0.27425
11	-1.8597	8.7417	-0.074573	0.02795	-0.027282	-0.072398	-0.13654	0.18864
12	-0.39657	42.424	-0.065158	-0.018158	-0.099504	-0.026572	-0.073284	0.1878
13	-0.54978	23.342	0.041242	0.014417	-0.043775	-0.020486	-0.032732	0.021781
14	-3.4736	0.75143	-0.0082202	-0.067825	-0.091643	0.069091	0.078551	0.01405
15	-0.31199	31.367	0.026761	0.0072152	0.015031	-0.015863	-0.0097697	-0.019017
16	-1.7758	3.5936	0.10717	-0.032362	-0.038655	-0.058026	0.040632	-0.029209
17	-0.16059	18.089	-0.0428	0.016927	0.040962	-0.039735	-0.043521	0.047678
18	-2.2338	6.2511	-0.070495	0.026704	0.018875	-0.13101	-0.12438	0.18383
19	-2.7772	9.7705	-0.045209	-0.0076608	0.023149	-0.10657	-0.12364	0.16485
20	-3.3938	12.535	-0.090452	0.032468	-0.081587	-0.02245	0.00090069	0.11971
21	-1.6224	37.494	0.00055005	-0.013199	-0.021748	-0.035844	-0.0087721	0.04853
22	-2.0133	14.345	0.021364	0.084079	0.10798	-0.10595	-0.084204	-0.018032
23	-3.4243	8.4063	-0.026558	0.061894	0.060129	-0.073706	-0.008226	-0.0032276
24	-1.5872	10.102	-0.057337	0.09175	0.091461	-0.059409	-0.068288	0.011591
25	-1.6551	9.4181	-0.062921	0.060763	0.026404	-0.063904	-0.079265	0.085613

```
26          -0.9819    15.033      -0.034107  0.051906   0.047598   -
0.08317        -0.076884   0.064343
27          -1.751     12.878      -0.042957  0.029734   0.059429   -
0.096057       -0.084145   0.087833
28          -0.1774    27.844      -0.03759   0.012154   -0.0078761 -
0.031427       -0.032431   0.066818
29          -0.5859    28.717      -0.044935  0.0032447  -0.011599  -
0.046033       -0.042716   0.095105
30          -0.44033   6.6755      -0.032097  0.058129   0.029079   -
0.06687        -0.077477   0.062085
31          -0.67335   4.6948      0.030832   0.048634   0.097226   -
0.044665       -0.049855   -0.057301
32          -3.4736    0.75143     -0.0082202 -0.067825  -0.091643
0.069091       0.078551    0.01405
33          -0.092482  54.739      -0.013186  -0.010443  -0.018006  -
0.010453       -0.0084866  0.039507
34          -0.22559   16.234      0.065962   0.071334   0.027832   -
0.0069337 -0.020891   -0.091715
35          -0.42105   9.9934      0.044079   0.13205          0.058075
0.00014594 -0.014078   -0.13663
36          -0.35232   68.461      0.041986   0.06243          0.044788
0.00073451 -0.014495   -0.088107
37          -0.41853   14.542      -0.059786  0.046906   0.083345   -
0.067165       -0.075797   0.052818
38          -0.65677   15.754      -0.084272  0.04263          0.09013    -
0.083737       -0.08659    0.086391
39          -0.37137   15.949      -0.055556  0.042761   0.075241   -
0.059883       -0.069622   0.048956
40          -0.26431   29.35       0.041966   0.055452   0.029756   -
0.014736       -0.014861   -0.064608
41          -0.41369   14.567      -0.059774  0.046369   0.083293   -
0.066937       -0.075561   0.052853
42          -0.16374   30.596      -0.029609  0.009295   -0.008666  -
0.025563       -0.015875   0.048978
43          -0.11229   54.273      -0.014475  -0.012921  -0.019959  -
0.011158       -0.0094043  0.044156
44          -0.57834   2.4111      -0.023614  0.055402   0.04686    -
0.02783        -0.090859   0.029562
45          -2.4037    0.99018     0.008524   -0.022984  -0.0071944
0.054038       -0.011789   -0.014992
46          -0.54541   3.4576      0.073899   -0.025784  0.068701
0.0084746      -0.039061   -0.070023
47          -0.55678   2.4793      -0.02353   0.056086   0.046787   -
0.035891       -0.091956   0.034767
48          -0.4899    3.2982      -0.030158  0.068629   0.062895   -
0.04733        -0.085002   0.025046
49          -0.78459   2.2579      -0.032389  0.090761   0.086016   -
0.05451        -0.091454   0.0076241
50          -0.46486   9.5615      0.060724   0.0042777  -0.039375
0.071214       0.009376    -0.072531
51          -2.1318    0.93219     0.013509   -0.092307  -0.073924
0.032407       0.088739    0.014998
```