Explanation of QA Results

The score for an individual question is the reciprocal of the rank at which the first correct response was found, or 0 if no correct response was found in the top 5 responses. The score for the run as a whole is the mean of the reciprocal ranks over the 198 test questions. Also reported is the number of questions for which no correct response was found.

The graph plots the score for a given run against the median score for each question. The median scores are computed separately over runs with a maximum answer length of 50 bytes and runs with a maximum answer length of 250 bytes. The median score is printed as a cross hatch, and the question ids along the x-axis are sorted by decreasing median score. An arrow runs from the median score to the score for the current run. If the run's score is greater than the median, the arrowhead is filled; otherwise the arrowhead is empty. If no arrow appears, then the run's score is identical to the median score for that question.

The order of the questions in the graphs is given below.

Table 1: Question number by decending median for runs with maximum answer length 50 bytes

1	5	14	44	82	112	122	144	147	170	172	178	29	39	71
78	85	90	132	160	185	18	19	24	28	48	63	80	127	155
157	163	171	180	192	3	33	45	51	70	76	84	95	110	137
149	195	35	121	151	2	4	6	7	8	9	10	11	12	13
15	16	17	20	21	22	23	25	26	27	30	31	32	34	36
37	38	40	41	42	43	46	47	49	50	52	53	54	55	56
57	58	59	60	61	62	64	65	66	67	68	69	72	73	74
75	77	79	81	83	86	87	88	89	91	92	93	94	96	97
98	99	100	101	102	103	104	105	106	107	108	109	111	113	114
115	116	117	118	119	120	123	124	125	126	128	129	130	133	134
135	136	138	139	140	141	142	143	145	146	148	150	152	153	154
156	158	159	161	162	164	165	166	167	168	169	173	174	175	176
177	179	181	182	183	186	187	188	189	190	191	193	194	196	197
198	199	200												

Table 2: Question number by decending median for runs with maximum answer length 250 bytes

1	13	14	16	18	28	29	30	39	44	45	46	48	50	58
63	66	80	82	90	110	112	117	118	122	127	128	139	147	155
163	178	180	5	10	26	35	38	60	65	71	76	78	83	85
99	103	120	124	132	134	140	144	154	160	162	170	172	3	19
24	37	40	41	42	51	57	68	70	79	84	95	100	106	137
151	157	185	190	6	52	88	102	121	145	149	171	183	192	199
33	59	62	130	138	168	188	200	2	4	7	8	9	11	12
15	17	20	21	22	23	25	27	31	32	34	36	43	47	49
53	54	55	56	61	64	67	69	72	73	74	75	77	81	86
87	89	91	92	93	94	96	97	98	101	104	105	107	108	109
111	113	114	115	116	119	123	125	126	129	133	135	136	141	142
143	146	148	150	152	153	156	158	159	161	164	165	166	167	169
173	174	175	176	177	179	181	182	186	187	189	191	193	194	195
196	197	198												