PAI 2018: KingSheep Results

Suzanne Tolmeijer

3. April 2018

Games

Of the 52 students enrolled for the course Practical Artificial Intelligence, 43 handed in agents for the game of KingSheep.

There were three maps used to test the agents: 'desert', 'heaven', and 'the corridor'. The list of all games was constructed by matching each submission against both of the predefined agent-pairs on each map plus matching each submission against every other submission on all three maps. Kingsheep is played on maps that are symmetrical and starting positions are deterministic. Therefore, the player that gets the first move (team1) potentially has an advantage. To address any fairness issues resulting from this circumstance, each match was duplicated with switched team positions of the agent-pairs. Therefore, all agent-pairs were matched against each other twice.

This resulted in the following number of games per student:

Student versus predefined agents:
2 (predefined agents * 2 (possible positions) * 3 (possible maps) = 12
Student versus other students:
42 (other students) * 2 (possible positions) * 3 (possible maps) = 252

In total, the following number of games was played:

Total number of games
43*12 + 43*126 = 5934
(preset games for 43 students + games against others in position 1)

Running all games (excluding debugging and reruns) cost approximately 30 hours.

Game Outcomes

The following game outcomes were observed in the evaluation:

- MAX_TURNS: A game of Kingsheep ends after 100 turns if no winner has been established before. The winning team is chosen based on the in-game score. Draws are possible.

- SHEEP_DIED: When the wolf of team A eats the sheep of team B, the game ends immediately and team A wins the game.

- TIME_LIMIT: When an agent of team A requires more than one second to compute the next action its team is disqualified and team B wins the game.

- ERROR: The game resulted in a runtime error and could not be concluded gracefully. Typically a Java StackOverflowError or NullPointerError when the agents returned *null* as action for a creature. With the current version of the game there is no reliable method to establish a winner under such circumstances. No points are awarded to any team.

The following bar plot shows the different outcomes of the games.
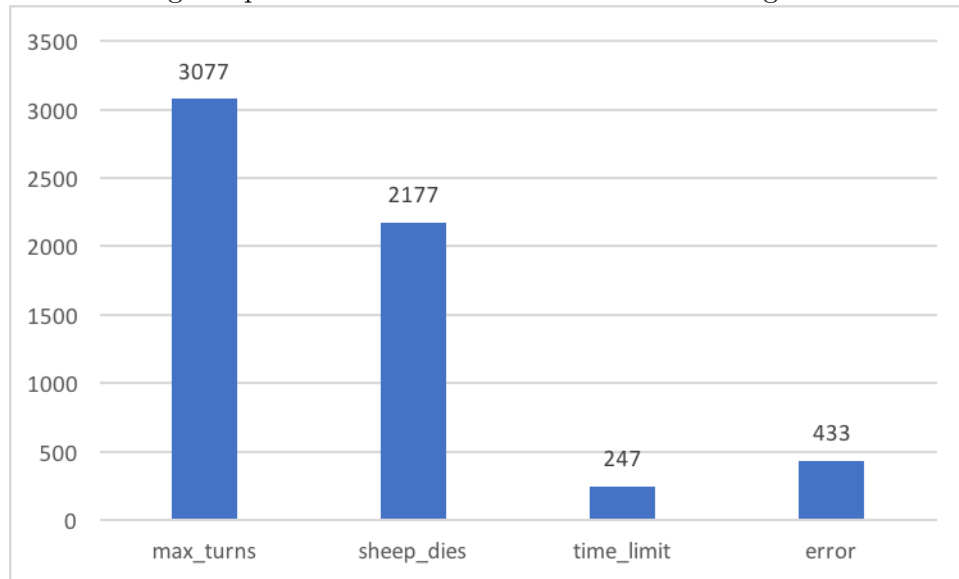
Evaluation Environment

The evaluation was run on a single machine with a 2.9 GHz Intel Core i7 processor and 16 GB RAM. The Java Virtual Machine for the evaluation was run with the following non-default parameters: -Xmx15G -Xss50M, which translates to 15 GB total heap memory size and 50 MB initialised stack memory per thread.

Results

The following table lists the cumulated points for each evaluated submission. The scores are listed separately for games against predefined agents (P) and games against other submissions (S). Entries are ordered by cumulated total points. The maximum possible points that could be achieved by a submission equals the number of games it participated in. Therefore, the maximum would be 12 and 252 for predefined and student opponents respectively and 274 in total.

nickname	P	S	total
Mormon	4.49	181.89	186.38
monsker	7.73	169.24	176.97
twitch.tv/kingsheep	12.00	155.67	167.67
robin$on	11.87	151.23	163.10
lucky	11.82	150.86	162.68
daca	8.73	147.06	155.79
Saki	10.06	146.66	156.73
HotShot	12.00	145.06	157.06
Yolo	11.69	144.78	156.48
Sheepeater69	11.50	144.77	156.27
Kingslayer Nelyo	9.18	144.72	153.90
Walti Ruedisueli	12.00	144.33	156.33
sheep_killer123	12.00	144.18	156.18
Dr.Reaper	12.00	143.95	155.95
Rick Roll	11.88	143.87	155.74
sheepyFressi	9.47	142.58	152.05
holzmano	10.44	142.33	152.77
Soap Opera	10.64	141.96	152.60
Robert Dapper	12.00	141.18	153.18
FeelsBadMan	8.41	141.00	149.41
sheepard	12.00	140.74	152.74
lollipop	11.00	140.31	151.31
Kind of disabled sheep that has no chance of winning against any opponent. Or does it? I dont know.	12.00	140.30	152.30
kingsheep007	12.00	140.10	152.10
Lynxxx	11.67	139.65	151.32
Bot Marley	11.50	139.60	151.11
RollingRou	12.00	139.31	151.31
I'M IN CHARGGEEE NOW!	8.87	138.94	147.81
A/N	8.48	138.62	147.10
carpe_diem	11.43	138.44	149.87
Notice me sheepAI	10.73	138.14	148.87
TelenovelaStar	11.08	137.46	148.54
timboloni	10.62	136.53	147.15
SgtUnicorn	12.00	135.24	147.24
sholf_weep	11.52	134.83	146.35
Dr. Propper	10.71	134.38	145.09
shaun_the_sheep	11.00	132.94	143.94
Rapperswil Chinderzoo schalalalala	11.36	126.87	138.23
dalaha	12.00	126.44	138.44
Guy Fieri	10.00	120.60	130.60
hukkepakk	9.43	120.56	129.99
Silence_of_the_Lambs	3.17	59.77	62.93
creatureone	0.00	5.75	5.75

The submissions were graded from 1 (worst) to 6 (best). For each submission i, let p_i denote its cumulated points from games against predefined agents and si are its cumulated points from games against all other submissions. A submissions grade is defined as:

$$g_i = \frac{p_i}{12} * 2.5 + \frac{s_i}{max(S)} * 2.5 + 1. \tag{1}$$

Half of the grade is assigned based on the performance against predefined agents, specifically 2.5 grade-points are distributed according to the fraction of achieved points to the maximum possible points against the predefined agents (6). The second 2.5 grade-points are assigned based on the relative performance against the best achieved of any submission score in S. The following table lists all submissions ordered by grade.

The next table lists the counts of different game outcomes for all the games each submission has played. Note that the occurrence of a particular outcome for a submission does not mean that the submission won or lost due to that reason but only that a game it was involved in ended for that reason. For example, the sheep died outcome means that a game the submission was involved in ended because a sheep died. But we can not tell if it was the sheep of that particular submission or the sheep of the opponent that died. This also means that being part of games causing errors does not automatically imply you caused the error.

As can be seen, two students caused a larger amount of errors. Because these games had to be excluded, their grades are unfortunately significantly lower. Besides these two outliers, everyone created pretty good code, that had some similar approaches in some cases (cooperation is valued in this group apparently..!) which caused quite a lot of max-turns, since many students were worthy opponents to each other.

nickname	G
twitch.tv/kingsheep	5.64
robin$on	5.55
lucky	5.54
HotShot	5.49
Walti Ruedisueli	5.48
sheep_killer123	5.48
Dr.Reaper	5.48
Rick Roll	5.45
Robert Dapper	5.44
sheepard	5.43
Kind of disabled sheep that has no chance of winning against any opponent. Or does it? I dont know.	5.43
Yolo	5.43
kingsheep007	5.43
RollingRou	5.41
Sheepeater69	5.38
SgtUnicorn	5.36
Lynxxx	5.35
Bot Marley	5.32
carpe_diem	5.28
sholf_weep	5.25
dalaha	5.24
lollipop	5.22
TelenovelaStar	5.20
Soap Opera	5.17
Notice me sheepAI	5.13
holzmano	5.13
shaun_the_sheep	5.12
Saki	5.11
Rapperswil Chinderzoo schalalalala	5.11
timboloni	5.09
Dr. Propper	5.08
monsker	4.94
sheepyFressi	4.93
Kingslayer Nelyo	4.90
daca	4.84
I'M IN CHARGGEEE NOW!	4.76
Guy Fieri	4.74
FeelsBadMan	4.69
A/N	4.67
hukkepakk	4.62
Mormon	4.43
Silence_of_the_Lambs	2.48
creatureone	1.08

!

nickname	MAX_TURNS	SHEEP_DIED	TIME_LIMIT	ERROR
FeelsBadMan	43	206	6	9
carpe_diem	123	122	7	12
A/N	125	123	6	10
Soap Opera	158	89	5	11
HotShot	175	70	8	11
hukkepakk	60	191	4	9
sheepard	164	84	7	9
Rick Roll	152	95	7	10
Robert Dapper	155	89	6	14
sholf_weep	117	113	22	11
Notice me sheepAI	219	25	6	14
creatureone	7	1	1	255
daca	95	152	7	10
Kind of disabled sheep that has no chance of winning against any opponent. Or does it? I dont know.	169	79	5	11
I'M IN CHARGGEEE NOW!	122	46	89	7
twitch.tv/kingsheep	151	98	5	10
Dr. Propper	152	97	5	10
holzmano	137	110	6	11
Dr.Reaper	129	118	6	11
TelenovelaStar	141	105	8	10
SgtUnicorn	163	83	8	10
monsker	100	123	14	27
Bot Marley	191	58	4	11
sheep_killer123	179	67	8	10
Yolo	149	96	10	9
Silence_of_the_Lambs	48	78	4	134
Rapperswil Chinderzoo schalalalala	185	67	3	9
sheepyFressi	90	160	3	11
lucky	169	51	35	9
Mormon	243	7	4	10
lollipop	135	109	9	11
RollingRou	152	89	9	14
robin$on	195	45	12	12
Saki	166	83	5	10
Lynxxx	176	71	6	11
Walti Ruedisueli	139	65	51	9
Kingslayer Nelyo	133	116	5	10
timboloni	58	191	6	9
dalaha	125	82	8	49
shaun_the_sheep	122	127	4	11
Sheepeater69	207	35	11	11
kingsheep007	169	78	5	12
Guy Fieri	153	46	55	10