

Misleading Information Reward in Presence of Particle Impoverishment

Mittwoch, 18. November 2020 09:19

Evaluationsergebnisse der C++ Implementierung des Solvers:

Compiled in debug mode
PSOPII/EUROPE/TEST/EN/1 - 200

Round	#Steps	disc. Reward	undisc. Reward
0	12	48.256	89.000
1	13	33.151	66.362
2	10	55.630	91.606
3	9	59.438	92.000
4	15	33.984	76.697
5	13	41.734	82.346
6	20	-12.830	28.000
7	17	26.445	69.527
8	14	41.601	87.000
9	13	35.096	69.795
10	10	55.630	91.000
11	17	32.815	84.000
12	5	-85.166	-104.000
13	14	41.601	87.000
14	15	38.521	86.000
15	13	44.843	88.000
16	17	32.815	84.000
17	11	51.840	90.000
18	11	44.364	77.500
19	11	-67.899	-110.000
20	6	72.854	95.000
21	15	38.521	86.000
22	16	35.595	85.000
23	8	63.800	93.000
24	12	48.256	89.000
25	14	41.601	87.000
26	14	-36.100	-52.691
27	11	29.378	52.457
28	11	43.918	76.754
29	14	41.601	87.000

Average total discounted reward (stdDev) = 31.245796 (.35_288949)
 Average total undiscounted reward (stdDev) = -62.512572 (.56_032346)

Ergebnis: (Bei 100 rounds ähnliche Ergebnisse)

Der Solver funktioniert in den meisten Fällen wie er soll (der final reward wird gefunden).

Manchmal jedoch wird der Reward nicht gefunden und das Ergebnis ist ein hoher negativer Reward.

Das führt zusammen zu einem geringeren Mittelwert des IPFT Paper. (Unterschied hier zur Tabelle ist, dass der gesampelt wurde, also wie bei sample_trajectory.)

Baptism

Der Solver wählt oft die Aktion 0, obwohl er noch nicht (mit ausreichender Sicherheit) in der Nähe der Optimalen Größe ist.

Digitized by srujanika@gmail.com

Beispielfehlerszenario A:

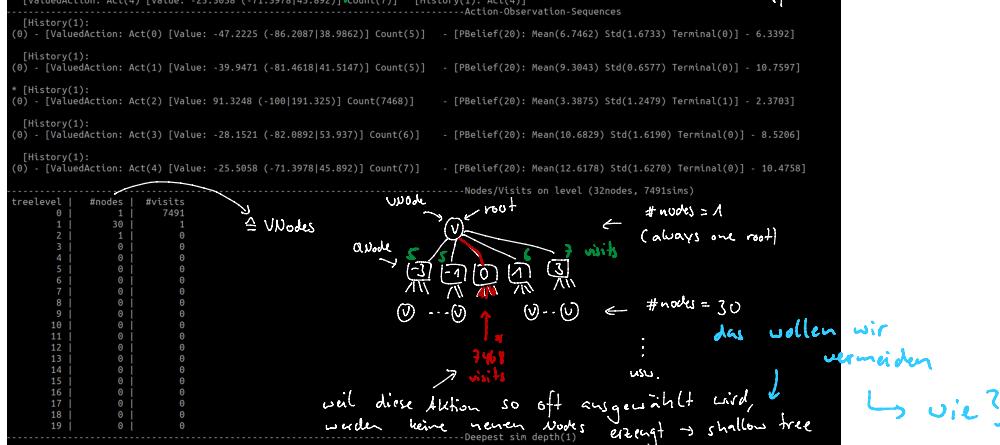
```

Search(1000,11ms, 7491 sims); [Value:Action: Act(2) [Value: 91.3248 (-108|191...325)] Count(7468) ] 
-----Root belief
PBelief(20): Mean(9.6092) Std(1.6324) Terminal(0)

Index | weight | value
-----+
 0 | 0.050000 | 1.7003
 1 | 0.050000 | 9.6376
 2 | 0.050000 | 10.1384
 3 | 0.050000 | 10.2247
 4 | 0.050000 | 10.7598
 5 | 0.050000 | 10.9999
 6 | 0.050000 | 10.4726
 7 | 0.050000 | 8.9347
 8 | 0.050000 | 9.2815
 9 | 0.050000 | 9.5353
10 | 0.050000 | 0.0242
11 | 0.050000 | 9.6946
12 | 0.050000 | 10.7702
13 | 0.050000 | 9.3272
14 | 0.050000 | 9.1568
15 | 0.050000 | 9.6167
16 | 0.050000 | 9.6757
17 | 0.050000 | 10.6167
18 | 0.050000 | 9.6881
19 | 0.050000 | 9.5846
-----+
  
```

nur ein "Outlier"

state reward: -100 !
information ras: 191
oft ausgewählt!
(+468 mal!)



Grund für die Auswahl von Aktion 0:

- Information reward überwiegt State Reward!

Grund für hohen Information Reward:

Nach Resampling sind alle Particles gleich!

Das fñhrt zu sehr hohem Information Reward, da Standardabweichung null ist und die Bandbreite nach Rule-of-Thumb gleich dem sqrt(Machine-Epsilon) = 1.49012e-08 (wegen max operator)

In den nächsten Schritten wird die Aktion, die zu dem "Single-state" Particle set geführt hat immer wieder ausgeführt, da hoher reward erhalten (in UCB Action selection)

```
max@max-ThinkPad-P1:~/ma/workspace/wsipft-devel/lib/ipft$ ./ipft-gtest-test_ipft_solver
Running main() from gtest_main.cc
[=====] Running 1 test from 1 test case.
[-----] Global test environment set-up.
[----] 1 test from InfrSolverTest
```

```

RUN : 1 IppSolverTest.Test1r15SolverSearchStep8
T20201118 09:37:24.801137 11783 {ipft.cpp:333} [Value: -47.6098 (-88.59149,98.9814)]
T20201118 09:37:24.803285 11783 {ipft.cpp:333} [Value: -5.92706 (-55.210149,2831)]
T20201118 09:37:24.803320 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.0933
T20201118 09:37:24.803457 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.805485 11783 {ipft.cpp:333} [Value: -31.9656 (-73.073641,108)]
T20201118 09:37:24.806794 11783 {ipft.cpp:333} [Value: -51.4699 (-100|46.5129)]
T20201118 09:37:24.808115 11783 {ipft.cpp:333} [Value: -58.8489 (-100|41.1511)]
T20201118 09:37:24.808717 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.2198
T20201118 09:37:24.808721 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.808746 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 2.9628
T20201118 09:37:24.808752 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.808762 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.0557
T20201118 09:37:24.808766 11783 {ipft.cpp:333} [Value: -57.6655 (-100|21.1346)]
T20201118 09:37:24.808768 11783 {ipft.cpp:333} [Value: -110.399 (-100|10.399)]
T20201118 09:37:24.808803 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.0253
T20201118 09:37:24.808817 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)] ←
T20201118 09:37:24.808817 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 2.6771
T20201118 09:37:24.808824 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.808843 11783 {ipft.cpp:333} [Value: -57.1655 (-100|42.8345)] ←
T20201118 09:37:24.808852 11783 {ipft.cpp:333} [Value: -61.9607 (-100|38.0393)]
T20201118 09:37:24.808856 11783 {ipft.cpp:333} [Value: -61.9607 (-100|38.0393)]
T20201118 09:37:24.808875 11783 {ipft.cpp:333} [Value: -78.0954 (-100|18.046)] ←
T20201118 09:37:24.808883 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.1118
T20201118 09:37:24.808894 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.809126 11783 {ipft.cpp:333} [Value: -66.252 (-100|33.748)]
T20201118 09:37:24.809188 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.0333
T20201118 09:37:24.809293 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)] ←
T20201118 09:37:24.809432 11783 {ipft.cpp:333} [Value: -61.1561 (-100|38.8439)] ←
T20201118 09:37:24.809478 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 2.9721
T20201118 09:37:24.809564 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.809571 11783 {ipft.cpp:333} [Value: -68.8094 (-100|31.422)]
T20201118 09:37:24.809704 11783 {ipft.cpp:333} [Value: -66.4944 (-100|31.422)]
T20201118 09:37:24.809947 11783 {ipft.cpp:333} [Value: -67.1359 (-100|32.49426)]
T20201118 09:37:24.810111 11783 {ipft.cpp:333} [Value: -66.6916 (-100|33.3084)]
T20201118 09:37:24.810254 11783 {ipft.cpp:333} [Value: -111.183 (-100|10.1826)]
T20201118 09:37:24.810384 11783 {ipft.cpp:333} [Value: -131.235 (-100|31.2349)]
T20201118 09:37:24.810498 11783 {ipft.cpp:333} [Value: -44.3649 (-100|55.6351)]
T20201118 09:37:24.810614 11783 {ipft.cpp:333} [Value: -122.093 (-100|22.0926)]
T20201118 09:37:24.810729 11783 {ipft.cpp:333} [Value: -102.768 (-100|27.76848)]
T20201118 09:37:24.810831 11783 {ipft.cpp:333} [Value: -100.494 (-100|27.4316)]
T20201118 09:37:24.811003 11783 {ipft.cpp:333} [Value: -59.2616 (-100|40.744)]
T20201118 09:37:24.818061 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.0898
T20201118 09:37:24.812128 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.813144 11783 {ipft.cpp:333} [Value: -54.8223 (-100|35.1777)]
T20201118 09:37:24.813174 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.1557
T20201118 09:37:24.814166 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.815170 11783 {ipft.cpp:333} [Value: -130.007 (-100|30.0074)]
T20201118 09:37:24.815199 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.1201
T20201118 09:37:24.815203 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.817053 11783 {ipft.cpp:333} [Value: -72.042 (-100|27.742)]
T20201118 09:37:24.818233 11783 {particle_belfef.cpp:198} After resampling due to particle deletion: All particles equal to 3.0466
T20201118 09:37:24.819062 11783 {ipft.cpp:333} [Value: 1009.34 (-100|1109.34)]
T20201118 09:37:24.820334 11783 {ipft.cpp:333} [Value: -53.8296 (-100|46.1704)]
T20201118 09:37:24.821247 11783 {ipft.cpp:333} [Value: -64.1773 (-100|35.8229)]
T20201118 09:37:24.821279 11783 {ipft.cpp:333} [Value: -57.6517 (-100|42.3483)]
T20201118 09:37:24.824099 11783 {ipft.cpp:333} [Value: -59.5594 (-100|40.4406)]

```

Das Problem tritt auch auf, wenn nicht die Action null gewählt wird:

Fällt, dann aber nicht so auf!

Frage:

Gibt es im IPFT Julia code irgendein Mechanismus der diesen Fehler auffängt? Oder warum tritt der Fehler in der julia implementierung nicht so häufig auf, wenn man das Beispielefehlerszenario A mehrfach in dem Julia code ausführt? Siehe test_step.jl file.