Mingjia Yao, 905302291
STATS 232C, Homework 3

Some explanations about my results:
-For the posterior probabilities of the goals, my results are not between 0 and 1, but the results should be proportional to the actual posterior probabilities (refer to the homework description).
-For the calculation of pi (at | st, g), the equation is shown as:
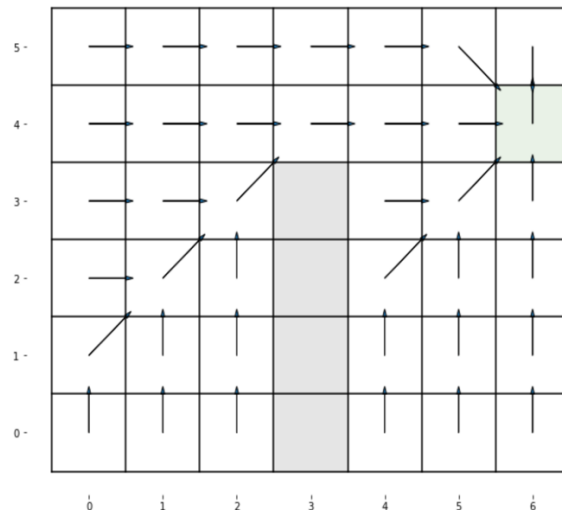
$$\pi(a_t|s_t, g) \propto e^{\beta Q_g^\pi(s_t, a_t)}$$

in the homework description.

However, if we use this equation to calculate pi, the result would be extremely large. Therefore, I used log instead of exp to calculate pi.
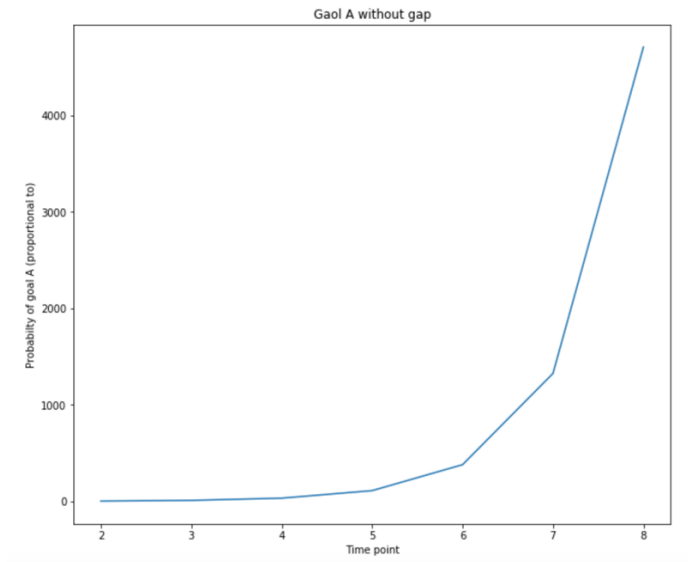
Goal A, without gap

Visualization of value of goal A without gap.   Visualization of policy of goal A without gap

The graph plotting the posterior probability of goal A without gap at each time point for the given trajectory sequences.
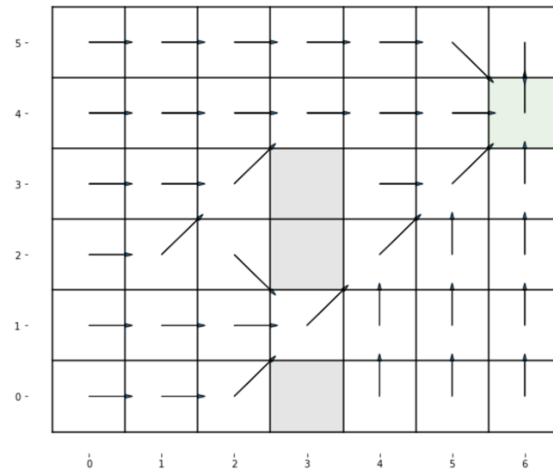The probabilities are proportional to:
[3.1379882404511474, 10.083985411991591, 33.08324092956145, 110.93174740658924, 379.8297746257766, 1324.9363289419246, 4705.747517484912]
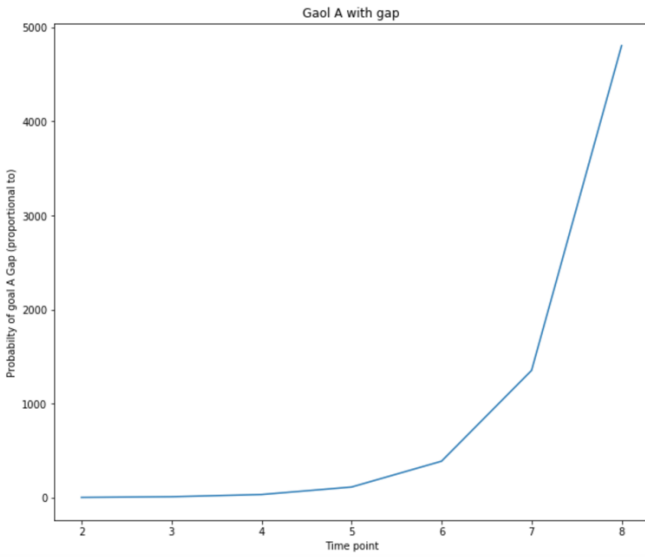
Goal A, with gap

	0	1	2	3	4	5	6
5	62.613	66.961	71.538	76.356	81.427	86.765	87.179
4	62.933	67.298	71.893	76.729	81.821	87.179	92.821
3	62.56	66.905	71.479	0	81.427	86.765	87.179
2	62.166	66.491	70.356	0	81.013	81.427	81.821
1	61.92	66.232	70.771	75.548	75.962	76.356	76.729
0	61.547	65.839	70.356	0	71.164	71.538	71.893

Visualization of value of goal A with gap.   Visualization of policy of goal A with gap

Gaol A with gap

The graph plotting the posterior probability of goal A with gap at each time point for the given trajectory sequences.
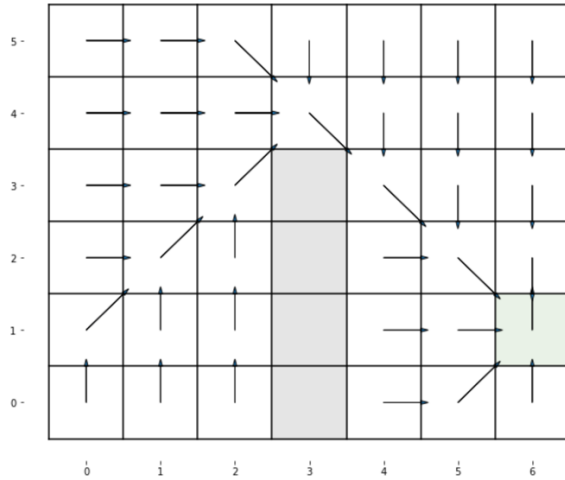
The probabilities are proportional to:
[3.202847205891906, 10.292411021839404, 33.76703786432827, 113.22459451024147, 387.6
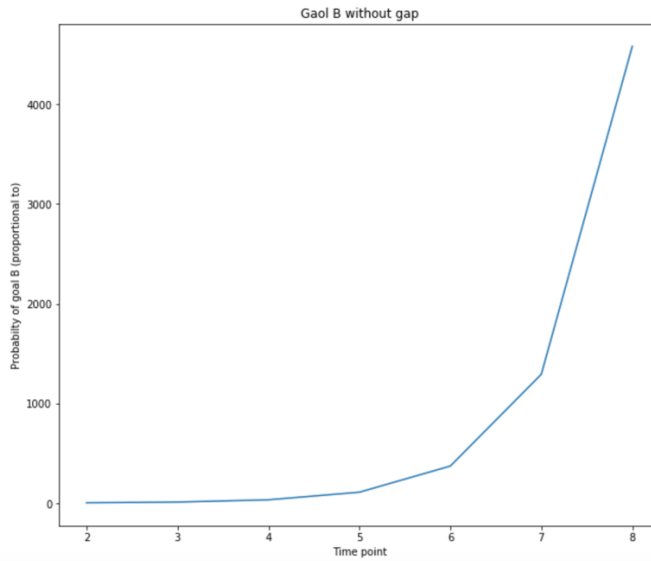804756342923, 1352.321443187641, 4803.010667872254]

Goal B, without gap

Visualization of value of goal B without gap.
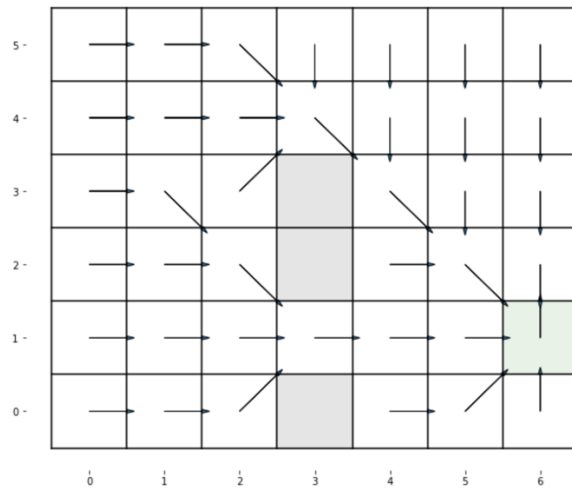

Visualization of policy of goal B without gap

The graph plotting the posterior probability of goal B without gap at each time point for the given trajectory sequences.

The probabilities are proportional to:
[3.121152038274025, 9.977545074384132, 32.63569334856427, 108.91450966094524, 371.23 255919461366, 1291.2642534958634, 4580.00537153257]
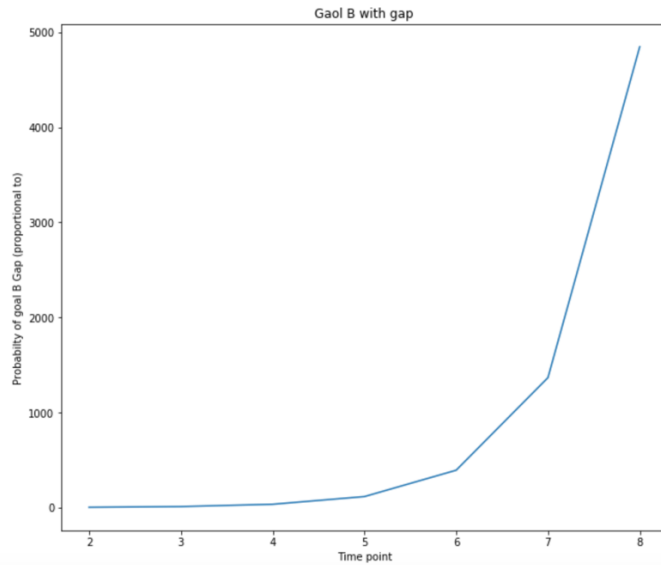
Goal B, with gap

Visualization of value of goal B with gap.

61.547	65.839	70.356	70.771	71.164	71.538	71.893
61.92	66.232	70.771	75.548	75.962	76.356	76.729
62.166	66.491	70.356	0	81.013	81.427	81.821
62.56	66.905	71.479	0	81.427	86.765	87.179
62.933	67.298	71.893	76.729	81.821	87.179	92.821
62.56	66.905	71.479	0	81.427	86.765	87.179

Visualization of policy of goal B with gap

The graph plotting the posterior probability of goal B with gap at each time point for the given trajectory sequences.
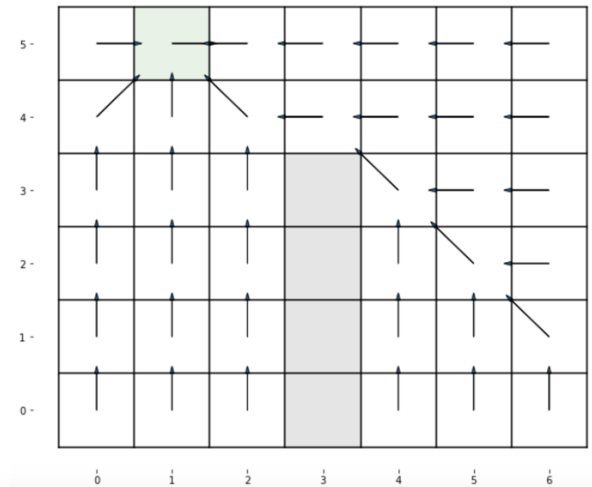
The probabilities are proportional to:
[3.2191833585636265, 10.561401585077677, 34.54543812744511, 115.28786758588984, 392.95600308938873, 1366.8252737102025, 4848.013960926132]
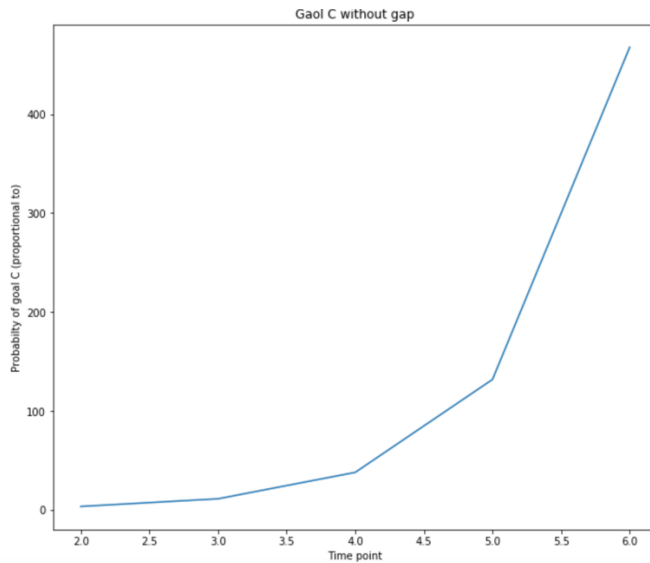
Goal C, without gap

Visualization of value of goal C without gap.

Visualization of policy of goal C without gap

The graph plotting the posterior probability of goal C without gap at each time point for the given trajectory sequences.
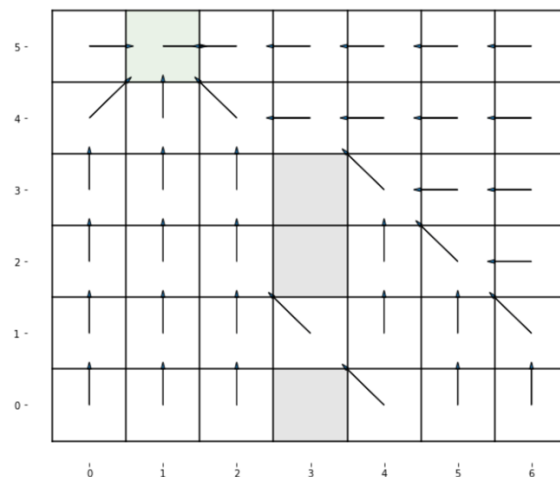The probabilities are proportional to:
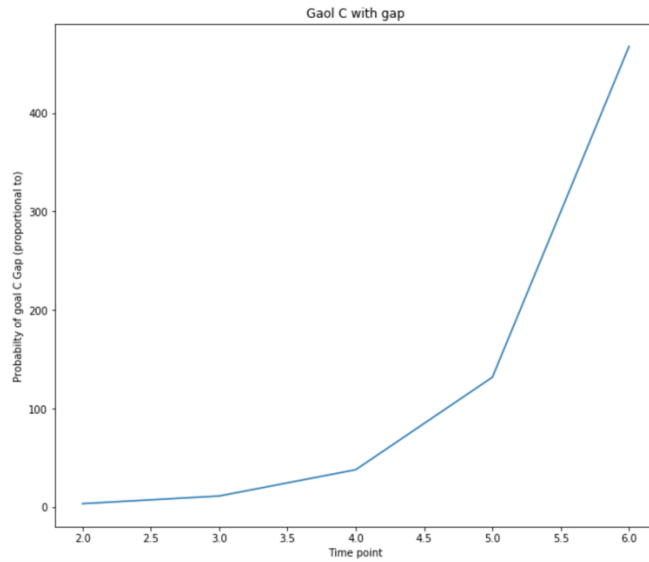[3.287819241259324, 11.024419686967674, 37.74756052223363, 131.67244248182197, 467.6581480148082]

Goal C, without gap

Visualization of value of goal C with gap.

Visualization of policy of goal C with gap

The graph plotting the posterior probability of goal C with gap at each time point for the given trajectory sequences.

The probabilities are proportional to:

[3.287819253742368, 11.024419805821687, 37.7475610492872, 131.6724450925809, 467.658 158549859]