Test Cases

I split up my tests into 3 different experiments for analysis of the scheduler in response to different inputs.

1) Initial State Test

   This experiment was designed to see how the scheduler responds to a variety of different initial state situations for the country being optimized. I kept constant parameters (frontier_size=15, depth=5, num_output_states=15, gamma=0.99, k=0.3, c=-1) for the scheduler throughout these tests.

   The input state files are all located in the 'input-states/' directory of this project. Each are described by filename in the following bullet list:

   - test_state_1: I used this mostly for testing for functionality during development. There is nothing interesting about this initial state.
   - test_state_2: The scheduler's country starts with very low food and housing, but has low to normal levels of most other resources. I wanted to see if the scheduler deliberately tries to go after one or the other, or both.
   - test_state_3: The scheduler's country has high housing, water, and farm, but low food levels. I wanted to see if the scheduler optimizes for procuring or creating food.
   - test_state_4: The scheduler's country has a gross excess of water, potential renewable energy, and potential fossil fuels, and has low levels of everything else. I was curious as to how the scheduler might optimize for trading resources it has too much of, in exchange for those it desperately needs.
   - test_state_5: Lots of every resource available, but very high levels of waste. I wanted to see how the scheduler prioritizes waste reduction.

2) Depth Test:

   Testing for the effect of max schedule depth on runtime, best utility found, and diversity of output states. I kept all other parameters constant: (test_state_5, frontier_size=15, gamma=0.99, k=0.3, c=-1). Depth was tested from the range of 1-8.

3) Frontier Size Test:

   Testing for the effect of frontier size on runtime, best utility found, and diversity of output states. All other parameters are constant: (test_state_5, depth=5, gamma=0.99, k=0.3, c=-1). Frontier size was tested with these values: [5, 10, 15, 30, 50]

My general choices for default parameters were based on quick intuition from my state quality function, resource levels in test cases, a desire to see longer, trade-heavy schedules, and a desire for the test cases to not take too long to run. Given more time, I would run some further experiments and statistical analysis on each parameter to determine what would give the best expected utility results, and what models real life more closely.

Results

1) Initial State Test
   Top 5 schedule results for each state file:

Test_State_2:

SCHEDULE 1

Final EU: 24.266427729024404
Actions:
TRANSFORM:
Country: X1
Input: [Population: 10,MetallicElements: 2,Timber: 10,MetallicAlloys: 6]
Output: [Housing: 2,HousingWaste: 2,Population: 10]
Statistics: (Reward: 14.433911610924959, Likelihood: 1, Exp_Utility: 14.433911610924959)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 15
Country2: X3, Resource2: HousingWaste, Amount2: -20
Statistics: (Reward: 19.191407412065708, Likelihood: 0.7801791549161124, Exp_Utility: 14.752915171312353)
TRANSFER:
Country1: X1, Resource1: AvailableLand, Amount1: 8
Country2: X3, Resource2: Food, Amount2: 3
Statistics: (Reward: 41.068834717762385, Likelihood: 0.6005972805887196, Exp_Utility: 24.266427729024404)

SCHEDULE 2

Final EU: 23.824356040641867
Actions:
TRANSFORM:
Country: X1
Input: [Population: 10,MetallicElements: 2,Timber: 10,MetallicAlloys: 6]
Output: [Housing: 2,HousingWaste: 2,Population: 10]
Statistics: (Reward: 14.433911610924959, Likelihood: 1, Exp_Utility: 14.433911610924959)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 15
Country2: X3, Resource2: HousingWaste, Amount2: -20

Statistics: (Reward: 19.191407412065708, Likelihood: 0.7801791549161124, Exp_Utility: 14.752915171312353)
TRANSFER:
Country1: X1, Resource1: AvailableLand, Amount1: 8
Country2: X3, Resource2: Food, Amount2: 3
Statistics: (Reward: 41.068834717762385, Likelihood: 0.6005972805887196, Exp_Utility: 24.266427729024404)

TRANSFORM:
Country: X1
Input: [Population: 1,MetallicElements: 2]
Output: [Population: 1,MetallicAlloys: 1,MetallicAlloysWaste: 1]
Statistics: (Reward: 40.33278128783475, Likelihood: 0.6005972805887196, Exp_Utility: 23.824356040641867)

SCHEDULE 3

Final EU: 23.824356040641867
Actions:
TRANSFORM:
Country: X1
Input: [Population: 10,MetallicElements: 2,Timber: 10,MetallicAlloys: 6]
Output: [Housing: 2,HousingWaste: 2,Population: 10]
Statistics: (Reward: 14.433911610924959, Likelihood: 1, Exp_Utility: 14.433911610924959)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 15
Country2: X3, Resource2: HousingWaste, Amount2: -20
Statistics: (Reward: 19.191407412065708, Likelihood: 0.7801791549161124, Exp_Utility: 14.752915171312353)
TRANSFER:
Country1: X1, Resource1: AvailableLand, Amount1: 8
Country2: X3, Resource2: Food, Amount2: 3
Statistics: (Reward: 41.068834717762385, Likelihood: 0.6005972805887196, Exp_Utility: 24.266427729024404)
TRANSFORM:
Country: X1
Input: [Population: 2,MetallicElements: 4]
Output: [Population: 2,MetallicAlloys: 2,MetallicAlloysWaste: 2]
Statistics: (Reward: 40.33278128783475, Likelihood: 0.6005972805887196, Exp_Utility: 23.824356040641867)

SCHEDULE 4

Final EU: 23.824356040641867
Actions:
TRANSFORM:
Country: X1
Input: [Population: 10,MetallicElements: 2,Timber: 10,MetallicAlloys: 6]
Output: [Housing: 2,HousingWaste: 2,Population: 10]
Statistics: (Reward: 14.433911610924959, Likelihood: 1, Exp_Utility: 14.433911610924959
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 15
Country2: X3, Resource2: HousingWaste, Amount2: -20
Statistics: (Reward: 19.191407412065708, Likelihood: 0.7801791549161124, Exp_Utility: 14.752915171312353)
TRANSFER:
Country1: X1, Resource1: AvailableLand, Amount1: 8
Country2: X3, Resource2: Food, Amount2: 3
Statistics: (Reward: 41.068834717762385, Likelihood: 0.6005972805887196, Exp_Utility: 24.266427729024404)
TRANSFORM:
Country: X1
Input: [Population: 3,MetallicElements: 6]
Output: [Population: 3,MetallicAlloys: 3,MetallicAlloysWaste: 3]
Statistics: (Reward: 40.33278128783475, Likelihood: 0.6005972805887196, Exp_Utility: 23.824356040641867)

SCHEDULE 5

Final EU: 23.824356040641867
Actions:
TRANSFORM:
Country: X1
Input: [Population: 10,MetallicElements: 2,Timber: 10,MetallicAlloys: 6]
Output: [Housing: 2,HousingWaste: 2,Population: 10]
Statistics: (Reward: 14.433911610924959, Likelihood: 1, Exp_Utility: 14.433911610924959)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 15
Country2: X3, Resource2: HousingWaste, Amount2: -20
Statistics: (Reward: 19.191407412065708, Likelihood: 0.7801791549161124, Exp_Utility: 14.752915171312353)

TRANSFER:
Country1: X1, Resource1: AvailableLand, Amount1: 8
Country2: X3, Resource2: Food, Amount2: 3
Statistics: (Reward: 41.068834717762385, Likelihood: 0.6005972805887196, Exp_Utility: 24.266427729024404)
TRANSFORM:
Country: X1
Input: [Population: 4,MetallicElements: 8]
Output: [Population: 4,MetallicAlloys: 4,MetallicAlloysWaste: 4]
Statistics: (Reward: 40.33278128783475, Likelihood: 0.6005972805887196, Exp_Utility: 23.824356040641867)

Test_State_3:

SCHEDULE 1

Final EU: 18.112528578699127
Actions:
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 18.112528578699127, Likelihood: 1, Exp_Utility: 18.112528578699127)

SCHEDULE 2

Final EU: 17.256732379405562
Actions:
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 18.112528578699127, Likelihood: 1, Exp_Utility: 18.112528578699127)
TRANSFORM:
Country: X1
Input: [Population: 1,MetallicElements: 2]
Output: [Population: 1,MetallicAlloys: 1,MetallicAlloysWaste: 1]
Statistics: (Reward: 17.256732379405562, Likelihood: 1, Exp_Utility: 17.256732379405562)

SCHEDULE 3

Final EU: 17.256732379405562
Actions:
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 18.112528578699127, Likelihood: 1, Exp_Utility: 18.112528578699127)
TRANSFORM:
Country: X1
Input: [Population: 2,MetallicElements: 4]
Output: [Population: 2,MetallicAlloys: 2,MetallicAlloysWaste: 2]
Statistics: (Reward: 17.256732379405562, Likelihood: 1, Exp_Utility: 17.256732379405562)

SCHEDULE 4

Final EU: 17.256732379405562
Actions:
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 18.112528578699127, Likelihood: 1, Exp_Utility: 18.112528578699127)
TRANSFORM:
Country: X1
Input: [Population: 3,MetallicElements: 6]
Output: [Population: 3,MetallicAlloys: 3,MetallicAlloysWaste: 3]
Statistics: (Reward: 17.256732379405562, Likelihood: 1, Exp_Utility: 17.256732379405562)

SCHEDULE 5

Final EU: 17.256732379405562
Actions:
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 18.112528578699127, Likelihood: 1, Exp_Utility: 18.112528578699127)

TRANSFORM:
Country: X1
Input: [Population: 4,MetallicElements: 8]
Output: [Population: 4,MetallicAlloys: 4,MetallicAlloysWaste: 4]
Statistics: (Reward: 17.256732379405562, Likelihood: 1, Exp_Utility: 17.256732379405562)

Test_State_4:

SCHEDULE 1

Final EU: 15.199974401039462
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 23.0
Country2: X3, Resource2: Food, Amount2: 23.0
Statistics: (Reward: 22.689551251656532, Likelihood: 0.5559501923447205, Exp_Utility: 12.170210574918565)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 12
Country2: X3, Resource2: Electronics, Amount2: 32
Statistics: (Reward: 36.94281604593867, Likelihood: 0.42695761910306285, Exp_Utility: 15.199974401039462)

SCHEDULE 2

Final EU: 14.947741564797877
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 23.0
Country2: X3, Resource2: Food, Amount2: 23.0
Statistics: (Reward: 22.689551251656532, Likelihood: 0.5559501923447205, Exp_Utility: 12.170210574918565)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 12
Country2: X3, Resource2: Electronics, Amount2: 32
Statistics: (Reward: 36.94281604593867, Likelihood: 0.42695761910306285, Exp_Utility: 15.199974401039462)
TRANSFORM:
Country: X1
Input: [Population: 1,MetallicElements: 2]
Output: [Population: 1,MetallicAlloys: 1,MetallicAlloysWaste: 1]

Statistics: (Reward: 36.352048192278, Likelihood: 0.42695761910306285, Exp_Utility: 14.947741564797877)

SCHEDULE 3

Final EU: 14.947741564797877
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 23.0
Country2: X3, Resource2: Food, Amount2: 23.0
Statistics: (Reward: 22.689551251656532, Likelihood: 0.5559501923447205, Exp_Utility: 12.170210574918565)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 12
Country2: X3, Resource2: Electronics, Amount2: 32
Statistics: (Reward: 36.94281604593867, Likelihood: 0.42695761910306285, Exp_Utility: 15.199974401039462)

TRANSFORM:
Country: X1
Input: [Population: 1,Farm: 1,Water: 1]
Output: [Population: 1,Farm: 1,Food: 3]
Statistics: (Reward: 36.352048192278, Likelihood: 0.42695761910306285, Exp_Utility: 14.947741564797877)

SCHEDULE 4

Final EU: 14.947741564797877
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 23.0
Country2: X3, Resource2: Food, Amount2: 23.0
Statistics: (Reward: 22.689551251656532, Likelihood: 0.5559501923447205, Exp_Utility: 12.170210574918565)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 12
Country2: X3, Resource2: Electronics, Amount2: 32
Statistics: (Reward: 36.94281604593867, Likelihood: 0.42695761910306285, Exp_Utility: 15.199974401039462)
TRANSFORM:
Country: X1
Input: [Population: 2,Farm: 2,Water: 2]
Output: [Population: 2,Farm: 2,Food: 6]

Statistics: (Reward: 36.352048192278, Likelihood: 0.42695761910306285, Exp_Utility: 14.947741564797877)

SCHEDULE 5

Final EU: 14.947741564797877
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 23.0
Country2: X3, Resource2: Food, Amount2: 23.0
Statistics: (Reward: 22.689551251656532, Likelihood: 0.5559501923447205, Exp_Utility: 12.170210574918565)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 12
Country2: X3, Resource2: Electronics, Amount2: 32
Statistics: (Reward: 36.94281604593867, Likelihood: 0.42695761910306285, Exp_Utility: 15.199974401039462)
TRANSFORM:
Country: X1
Input: [Population: 1,PotentialFossilEnergy: 1,Electronics: 1]
Output: [Population: 1,PotentialFossilEnergyUsable: 1,Electronics: 1,PotentialFossilEnergyUsableWaste: 1]
Statistics: (Reward: 36.352048192278, Likelihood: 0.42695761910306285, Exp_Utility: 14.947741564797877)

Test_State_5:

SCHEDULE 1

Final EU: 58.244642614231
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 24
Country2: X3, Resource2: AvailableLand, Amount2: 64
Statistics: (Reward: 84.3986874136265, Likelihood: 0.589506263422938, Exp_Utility: 49.34306111843045)
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 95.64761777096751, Likelihood: 0.589506263422938, Exp_Utility: 55.9743760208914)
TRANSFER:

Country1: X1, Resource1: Water, Amount1: 107.0
Country2: X3, Resource2: PotentialRenewableEnergyUsableWaste, Amount2: -107.0
Statistics: (Reward: 103.53824383090539, Likelihood: 0.5667269741977058, Exp_Utility: 58.244642614231)

SCHEDULE 2

Final EU: 58.2441956228842
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 24
Country2: X3, Resource2: AvailableLand, Amount2: 64
Statistics: (Reward: 84.3986874136265, Likelihood: 0.589506263422938, Exp_Utility: 49.34306111843045)
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 95.64761777096751, Likelihood: 0.589506263422938, Exp_Utility: 55.9743760208914)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 106.0
Country2: X3, Resource2: PotentialRenewableEnergyUsableWaste, Amount2: -106.0
Statistics: (Reward: 103.45624673231384, Likelihood: 0.5671675699272166, Exp_Utility: 58.2441956228842)

SCHEDULE 3

Final EU: 58.244145428503415
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 24
Country2: X3, Resource2: AvailableLand, Amount2: 64
Statistics: (Reward: 84.3986874136265, Likelihood: 0.589506263422938, Exp_Utility: 49.34306111843045)
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 95.64761777096751, Likelihood: 0.589506263422938, Exp_Utility: 55.9743760208914)
TRANSFER:

Country1: X1, Resource1: Water, Amount1: 108.0
Country2: X3, Resource2: PotentialRenewableEnergyUsableWaste, Amount2: -108.0
Statistics: (Reward: 103.62024092949696, Likelihood: 0.5662780443071981, Exp_Utility: 58.244145428503415)

SCHEDULE 4

Final EU: 58.24282128316588
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 24
Country2: X3, Resource2: AvailableLand, Amount2: 64
Statistics: (Reward: 84.3986874136265, Likelihood: 0.589506263422938, Exp_Utility: 49.34306111843045)
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 95.64761777096751, Likelihood: 0.589506263422938, Exp_Utility: 55.9743760208914)
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 105.0
Country2: X3, Resource2: PotentialRenewableEnergyUsableWaste, Amount2: -105.0
Statistics: (Reward: 103.3742496337223, Likelihood: 0.5675999730878554, Exp_Utility: 58.24282128316588)

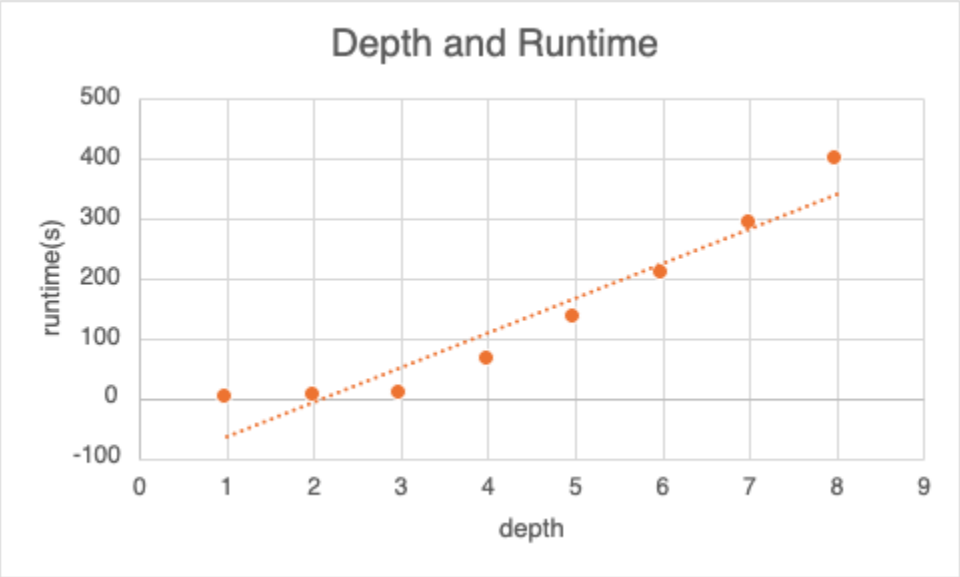
SCHEDULE 5

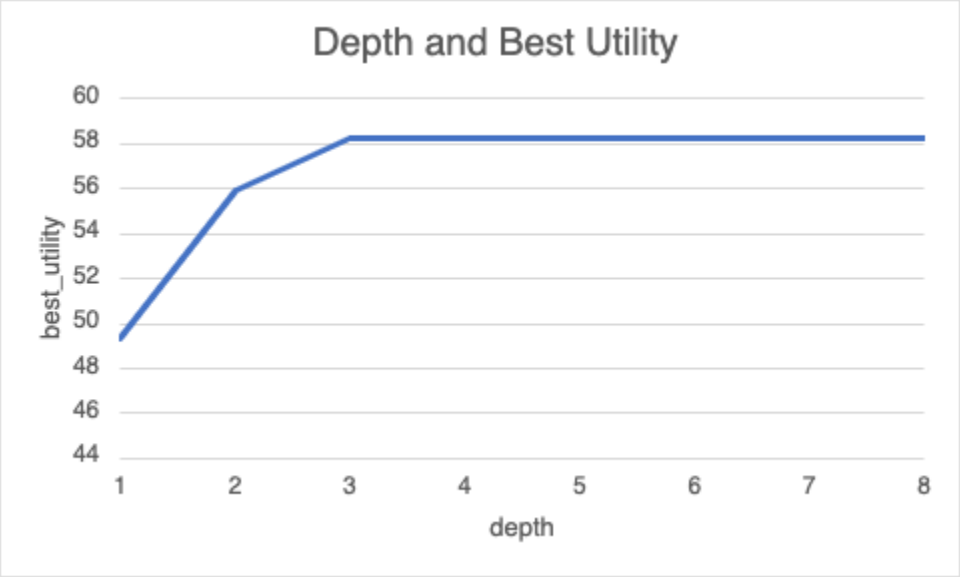
Final EU: 58.24268699111756
Actions:
TRANSFER:
Country1: X1, Resource1: Water, Amount1: 24
Country2: X3, Resource2: AvailableLand, Amount2: 64
Statistics: (Reward: 84.3986874136265, Likelihood: 0.589506263422938, Exp_Utility: 49.34306111843045)
TRANSFORM:
Country: X1
Input: [Population: 20,MetallicElements: 4,Timber: 20,MetallicAlloys: 12]
Output: [Housing: 4,HousingWaste: 4,Population: 20]
Statistics: (Reward: 95.64761777096751, Likelihood: 0.589506263422938, Exp_Utility: 55.9743760208914)
TRANSFER:

Country1: X1, Resource1: Water, Amount1: 109.0
Country2: X3, Resource2: PotentialRenewableEnergyUsableWaste, Amount2: -109.0
Statistics: (Reward: 103.7022380280885, Likelihood: 0.5658206367587338, Exp_Utility: 58.24268699111756)

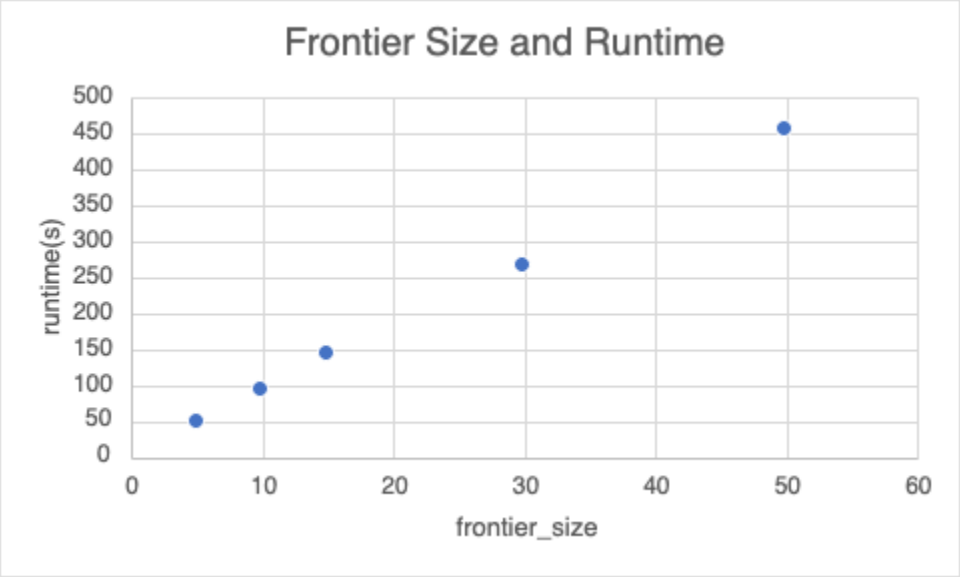
For the next 2 experiments, due to the large amount of schedule data collected, I provide charts that summarize the quantitative relationships between variables that were tested, rather than full output schedules.

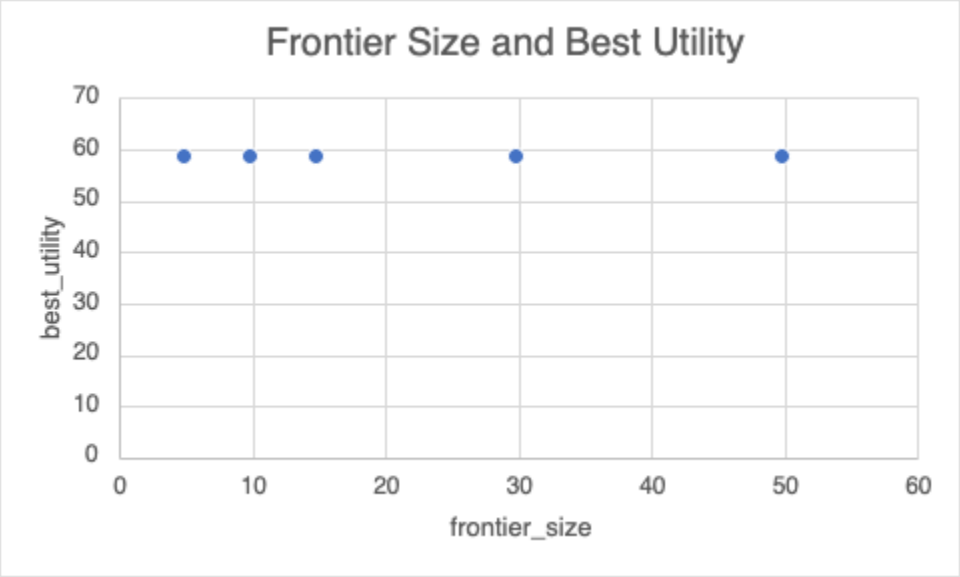
2) Depth Test





3) Frontier Size Test





Analysis

The data collected during my experiment provide several insights into how my algorithm performs in different scenarios.

The relationship between Frontier Size and Runtime appears very close to linear. This is as I expected, since we maintain a constant-size frontier, and generate successors only from those state nodes that make it onto the frontier.

The relationship between Depth and Runtime seems to be linear as well, although it changes slope at approximately depth=3. This slope change is likely due to the cost of repeating generate_successors taking over the initial overhead, and I predict that if we were to continue testing deeper depths, the runtime would remain linear.

Both approximately linear runtimes result from the fact that we keep a constant-sized frontier. This is a very desirable outcome, as it allows us to test very deep into the search space, and look for schedules far in advance. However, there are some caveats that come with the purely ‘best-fit’ beam approach that we have taken.

From the results of both experiments, it is apparent that we never find a better Expected Utility than from a depth 3 search. At first glance, this could be mistaken for having found a global maximum early on, which would be an ideal situation for our algorithm. However, from closer analysis of the output schedules from test 1, as well as all results of my experiments, it is apparent that our search becomes very concentrated in a small part of the search space right at the start. The frontier pops the best move from the initial state first, and then in almost all cases replaces the rest of the frontier with ‘second moves’ that are generated from that best depth 1 move, which are likely to be better than anything else from depth 1. This is an issue because we completely avoid schedules that might have very high delayed rewards, and could possibly beat schedules that start with the greediest move. This problem is apparent in many of the output files where the best schedule is one of depth one, and the rest of the best schedule list are just immediate successors of this schedule (see ‘output/experiment_initstate_results/state3.out’ for a clear example). In fact, throughout the results files, the first choice of action for all best schedules in any test case are almost always the exact same action! Our scheduler finds the very best first action, but doesn’t necessarily make the best choice at depth 2 onwards. Better results later on might be seen if we replaced the double-ended priority queue frontier with a non-ordered data structure, and used more of a ‘random walk’-based algorithm for choosing new nodes to explore.