

Part 1: Monte Carlo First Visit:

Epoch 0:

inital values of N(s):

```
[[0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]]
```

inital values of S(s):

```
[[0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]]
```

inital values of V(s):

```
[[0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]]
```

Epoch 1:

Episode data array:

```
[[1, 2, -1, 0.9, -1.0]]
```

N(s):

```
[[0. 1. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0.]]
```

S(s):

```
[[ 0. -1.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]]
```

V(s):

```
[[ 0. -1.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]  
 [ 0.  0.  0.  0.  0.]]
```

Epoch 10:

Episode data array:

```
[[1, 24, -1, 0.9, -1.0], [2, 24, -1, 0.9, -1.9], [3, 24, -1, 0.9, -2.71], [4, 24, -1, 0.9, -3.439], [5, 19, -1, 0.9, -4.0951], [6, 14, -1, 0.9, -4.68559], [7, 19, -1, 0.9, -5.217031], [8, 14, -1, 0.9, -5.6953279000000006], [9, 15, -1, 0.9, -6.12579511], [10, 15, -1, 0.9, -6.5132155990000005], [11, 20, -1, 0.9, -6.861894039100001], [12, 20, -1, 0.9, -7.175704635190001], [13, 15, -1, 0.9, -7.458134171671], [14, 10, -1, 0.9, -7.7123207545039], [15, 15, -1, 0.9, -7.94108867905351], [16, 10, -1, 0.9, -8.14697981114816]]
```

N(s):

```
[[1. 4. 2. 2. 2.]
 [4. 4. 2. 3. 3.]
 [4. 4. 3. 5. 3.]
 [4. 4. 5. 6. 3.]
 [3. 4. 5. 5. 0.]]
```

S(s):

```
[[ 0. -12.86697205 -19.84592538 -16.05509861 -15.29744905]
 [-19.85399605 -23.3712433 -18.74157808 -23.76375604 -22.60034614]
 [-21.70440677 -25.21383531 -28.11485107 -36.08884906 -20.44582332]
 [-25.21695353 -24.87860259 -32.41958473 -30.22132363 -23.37394858]
 [-24.51709612 -29.59278361 -32.54656665 -5. 0. ]]
```

V(s):

```
[[ 0. -3.21674301 -9.92296269 -8.0275493 -7.64872453]
 [-4.96349901 -5.84281083 -9.37078904 -7.92125201 -7.53344871]
 [-5.42610169 -6.30345883 -9.37161702 -7.21776981 -6.81527444]
 [-6.30423838 -6.21965065 -6.48391695 -5.03688727 -7.79131619]
 [-8.17236537 -7.3981959 -6.50931333 -1. 0. ]]
```

Converged after 957 episodes.

Epoch 957:

Episode data array:

```
[[1, 1, 0, 0.9, 0.0]]
```

N(s):

```
[[ 33. 361. 382. 367. 296.]
 [354. 442. 475. 443. 355.]
 [377. 459. 487. 478. 396.]
 [347. 440. 481. 496. 409.]
 [292. 370. 389. 396. 40.]]
```

S(s):

```
[[ 0. -1382.45355488 -2360.42931603 -2643.0087013
 -2316.81272605]
```

```

[-1278.92924611 -2391.32725643 -3105.16030309 -3039.47411992
 -2572.94776718]
[-2254.83122409 -3020.40811748 -3386.60805837 -3077.38041156
 -2465.307045 ]
[-2475.14086258 -3045.44092872 -3026.56346806 -2626.43185568
 -1678.15827242]
[-2292.66032624 -2700.10139996 -2360.93807199 -1565.76551186
 0. ]]
```

V(s):

```

[[ 0. -3.82951123 -6.17913434 -7.20165859 -7.82707002]
 [-3.61279448 -5.41024266 -6.53717959 -6.86111539 -7.24774019]
 [-5.98098468 -6.58040984 -6.95402065 -6.43803433 -6.22552284]
 [-7.13297079 -6.92145666 -6.29223174 -5.29522552 -4.10307646]
 [-7.85157646 -7.29757135 -6.06924954 -3.95395331 0. ]]
```

Part 2: Monte Carlo Every Visit:

Epoch 0:

inital values of N(s):

```

[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

inital values of S(s):

```

[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

inital values of V(s):

```

[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

Epoch 1:

Episode data array:

```

[[1, 1, 0, 0.9, 0.0]]
```

N(s):

```

[[1. 0. 0. 0. 0.]
```

```
[0. 0. 0. 0. 0.]
[0. 0. 0. 0. 0.]
[0. 0. 0. 0. 0.]
[0. 0. 0. 0. 0.]
```

S(s):

```
[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

V(s):

```
[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

Epoch 10:

Episode data array:

```
[[1, 24, -1, 0.9, -1.0], [2, 23, -1, 0.9, -1.9], [3, 23, -1, 0.9,
-2.71], [4, 22, -1, 0.9, -3.439], [5, 21, -1, 0.9, -4.0951], [6, 21,
-1, 0.9, -4.68559], [7, 16, -1, 0.9, -5.217031], [8, 21, -1, 0.9,
-5.6953279000000006], [9, 21, -1, 0.9, -6.12579511], [10, 22, -1, 0.9,
-6.5132155990000005], [11, 17, -1, 0.9, -6.861894039100001], [12, 12,
-1, 0.9, -7.175704635190001], [13, 17, -1, 0.9, -7.458134171671], [14,
18, -1, 0.9, -7.7123207545039], [15, 13, -1, 0.9, -7.94108867905351],
[16, 8, -1, 0.9, -8.14697981114816], [17, 9, -1, 0.9,
-8.332281830033345], [18, 14, -1, 0.9, -8.49905364703001], [19, 19,
-1, 0.9, -8.649148282327008], [20, 24, -1, 0.9, -8.784233454094307],
[21, 23, -1, 0.9, -8.905810108684877], [22, 24, -1, 0.9,
-9.01522909781639], [23, 23, -1, 0.9, -9.113706188034751], [24, 18,
-1, 0.9, -9.202335569231277]]
```

N(s):

```
[[ 1. 12. 19. 24. 25.]
 [13. 13. 15. 20. 24.]
 [19. 14. 18. 17. 18.]
 [24. 23. 22. 16. 10.]
 [24. 21. 17. 11.  0.]]
```

S(s):

```
[[ 0. -110.76440183 -181.22836336 -223.5529773
-210.62278945]
 [-77.70410823 -115.34738792 -139.77662462 -184.8229641
-198.00716414]
 [-141.17752228 -134.28132774 -171.63720913 -150.09994086
-140.40189237]
```

```
[-200.57714869 -211.38986277 -202.01241037 -114.54959889
-34.3273871 ]
[-203.59910848 -187.81444287 -135.64842651 -59.44026622
0.      ]]
```

V(s):

```
[[ 0.      -9.23036682 -9.53833491 -9.31470739 -8.42491158]
 [-5.97723909 -8.87287599 -9.31844164 -9.2411482  -8.25029851]
 [-7.43039591 -9.59152341 -9.53540051 -8.82940829 -7.80010513]
 [-8.3573812  -9.1908636  -9.18238229 -7.15934993 -3.43273871]
 [-8.48329619 -8.9435449  -7.97931921 -5.40366057  0.      ]]
```

Converged after 1432 episodes.
Epoch 1432:

Episode data array:

```
[[1, 1, 0, 0.9, 0.0]]
```

N(s):

```
[[ 66. 1275. 1897. 2110. 2249.]
 [1308. 1711. 2009. 2108. 2253.]
 [2083. 2048. 2084. 1973. 1969.]
 [2410. 2254. 2131. 1675. 1306.]
 [2534. 2461. 1983. 1263.   69.]]
```

S(s):

```
[[ 0.      -7166.42097173 -14669.92312904 -18215.92252493
-19995.68462217]
 [-7599.79350894 -12605.9246969  -16674.75960073 -18110.87756867
-19427.47420444]
 [-16696.87330189 -16986.29307699 -17605.79958142 -16191.03371051
-15449.40359373]
 [-20912.95553813 -19288.87422378 -17518.03521713 -12135.34878914
-7017.21388898]
 [-22563.45882019 -21245.50694218 -15394.61341849  -7239.76086021
0.      ]]
```

V(s):

```
[[ 0.      -5.62072233 -7.73322252 -8.63313864 -8.89092246]
 [-5.81023969 -7.36757726 -8.30002967 -8.5914979  -8.62293573]
 [-8.01578171 -8.29408842 -8.44808041 -8.20630193 -7.84631975]
 [-8.67757491 -8.55761944 -8.22057026 -7.24498435 -5.37305811]
 [-8.90428525 -8.63287564 -7.76329471 -5.73219387  0.      ]]
```

Part 3: Monte Carlo On Policy Every Visit:

Epoch 0:

inital values of N(s):

```
[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

inital values of S(s):

```
[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

inital values of V(s):

```
[[0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

Epoch 1:

Episode data array:

```
[[1, 2, -1, 0.9, -1.0], [2, 3, -1, 0.9, -1.9], [3, 8, -1, 0.9, -2.71],
 [4, 3, -1, 0.9, -3.439], [5, 8, -1, 0.9, -4.0951], [6, 3, -1, 0.9,
 -4.68559], [7, 8, -1, 0.9, -5.217031], [8, 13, -1, 0.9,
 -5.69532790000000006], [9, 8, -1, 0.9, -6.12579511], [10, 7, -1, 0.9,
 -6.51321559900000005], [11, 12, -1, 0.9, -6.861894039100001], [12, 7,
 -1, 0.9, -7.175704635190001], [13, 2, -1, 0.9, -7.458134171671], [14,
 3, -1, 0.9, -7.7123207545039], [15, 4, -1, 0.9, -7.94108867905351],
 [16, 5, -1, 0.9, -8.14697981114816], [17, 5, -1, 0.9,
 -8.332281830033345], [18, 10, -1, 0.9, -8.49905364703001], [19, 5, -1,
 0.9, -8.649148282327008]]
```

N(s):

```
[[0. 2. 4. 1. 3.]
 [0. 2. 4. 0. 1.]
 [0. 1. 1. 0. 0.]
 [0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0.]]
```

S(s):

```
[[ 0.          -8.45813417 -17.73691075  -7.94108868 -25.12840992]
 [ 0.          -13.68892023 -18.14792611   0.          -8.49905365]
 [ 0.          -6.86189404  -5.6953279   0.           0.          ]
 [ 0.           0.           0.           0.           0.          ]
 [ 0.           0.           0.           0.           0.          ]]
```

V(s):

```
[ [ 0.          -4.22906709 -4.43422769 -7.94108868 -8.37613664]
  [ 0.          -6.84446012 -4.53698153  0.          -8.49905365]
  [ 0.          -6.86189404 -5.6953279  0.          0.          ]
  [ 0.           0.           0.           0.           0.          ]
  [ 0.           0.           0.           0.           0.          ]]
```

Epoch 10:

Episode data array:

```
[ [1, 15, -1, 0.9, -1.0], [2, 15, -1, 0.9, -1.9], [3, 15, -1, 0.9,
-2.71], [4, 15, -1, 0.9, -3.439], [5, 15, -1, 0.9, -4.0951], [6, 15,
-1, 0.9, -4.68559], [7, 15, -1, 0.9, -5.217031], [8, 15, -1, 0.9,
-5.6953279000000006], [9, 15, -1, 0.9, -6.12579511], [10, 15, -1, 0.9,
-6.5132155990000005], [11, 15, -1, 0.9, -6.861894039100001], [12, 15,
-1, 0.9, -7.175704635190001], [13, 15, -1, 0.9, -7.458134171671], [14,
15, -1, 0.9, -7.7123207545039], [15, 15, -1, 0.9, -7.94108867905351],
[16, 15, -1, 0.9, -8.14697981114816], [17, 15, -1, 0.9,
-8.332281830033345], [18, 15, -1, 0.9, -8.49905364703001], [19, 15,
-1, 0.9, -8.649148282327008], [20, 15, -1, 0.9, -8.784233454094307],
[21, 15, -1, 0.9, -8.905810108684877], [22, 15, -1, 0.9,
-9.01522909781639], [23, 15, -1, 0.9, -9.113706188034751], [24, 15,
-1, 0.9, -9.202335569231277], [25, 15, -1, 0.9, -9.28210201230815],
[26, 15, -1, 0.9, -9.353891811077334], [27, 15, -1, 0.9,
-9.4185026299696], [28, 15, -1, 0.9, -9.47665236697264], [29, 15, -1,
0.9, -9.528987130275377], [30, 15, -1, 0.9, -9.57608841724784], [31,
15, -1, 0.9, -9.618479575523056], [32, 15, -1, 0.9,
-9.656631617970751], [33, 15, -1, 0.9, -9.690968456173676], [34, 15,
-1, 0.9, -9.72187161055631], [35, 15, -1, 0.9, -9.749684449500679],
[36, 15, -1, 0.9, -9.77471600455061], [37, 15, -1, 0.9,
-9.797244404095549], [38, 15, -1, 0.9, -9.817519963685994], [39, 15,
-1, 0.9, -9.835767967317395], [40, 15, -1, 0.9, -9.852191170585655],
[41, 15, -1, 0.9, -9.86697205352709], [42, 15, -1, 0.9,
-9.880274848174382], [43, 15, -1, 0.9, -9.892247363356944], [44, 15,
-1, 0.9, -9.90302262702125], [45, 15, -1, 0.9, -9.912720364319124],
[46, 15, -1, 0.9, -9.921448327887212], [47, 15, -1, 0.9,
-9.929303495098491], [48, 15, -1, 0.9, -9.936373145588643], [49, 15,
-1, 0.9, -9.942735831029779], [50, 15, -1, 0.9, -9.948462247926802],
[51, 15, -1, 0.9, -9.953616023134122], [52, 15, -1, 0.9,
-9.95825442082071], [53, 15, -1, 0.9, -9.96242897873864], [54, 15, -1,
0.9, -9.966186080864777], [55, 15, -1, 0.9, -9.9695674727783], [56,
15, -1, 0.9, -9.97261072550047], [57, 15, -1, 0.9,
-9.975349652950424], [58, 15, -1, 0.9, -9.977814687655382], [59, 15,
-1, 0.9, -9.980033218889844], [60, 15, -1, 0.9, -9.98202989700086],
[61, 15, -1, 0.9, -9.983826907300774], [62, 15, -1, 0.9,
-9.985444216570697], [63, 15, -1, 0.9, -9.986899794913628], [64, 15,
-1, 0.9, -9.988209815422264], [65, 15, -1, 0.9, -9.989388833880039],
[66, 15, -1, 0.9, -9.990449950492035], [67, 15, -1, 0.9,
-9.991404955442832], [68, 15, -1, 0.9, -9.992264459898548], [69, 15,
-1, 0.9, -9.993038013908693], [70, 15, -1, 0.9, -9.993734212517824],
[71, 15, -1, 0.9, -9.994360791266041], [72, 15, -1, 0.9,
```

```
-9.994924712139438], [73, 15, -1, 0.9, -9.995432240925494], [74, 15,
-1, 0.9, -9.995889016832944], [75, 15, -1, 0.9, -9.99630011514965],
[76, 15, -1, 0.9, -9.996670103634685], [77, 15, -1, 0.9,
-9.997003093271216], [78, 15, -1, 0.9, -9.997302783944095], [79, 15,
-1, 0.9, -9.997572505549686], [80, 15, -1, 0.9, -9.997815254994718],
[81, 15, -1, 0.9, -9.998033729495246], [82, 15, -1, 0.9,
-9.998230356545722], [83, 15, -1, 0.9, -9.99840732089115], [84, 15,
-1, 0.9, -9.998566588802035], [85, 15, -1, 0.9, -9.998709929921832],
[86, 15, -1, 0.9, -9.99883893692965], [87, 15, -1, 0.9,
-9.998955043236684], [88, 15, -1, 0.9, -9.999059538913016], [89, 15,
-1, 0.9, -9.999153585021714], [90, 15, -1, 0.9, -9.999238226519543],
[91, 15, -1, 0.9, -9.999314403867588], [92, 15, -1, 0.9,
-9.99938296348083], [93, 15, -1, 0.9, -9.999444667132748], [94, 15,
-1, 0.9, -9.999500200419472], [95, 15, -1, 0.9, -9.999550180377526],
[96, 15, -1, 0.9, -9.999595162339773], [97, 15, -1, 0.9,
-9.999635646105796], [98, 15, -1, 0.9, -9.999672081495216], [99, 15,
-1, 0.9, -9.999704873345696], [100, 15, -1, 0.9, -9.999734386011127]]
```

N(s):

```
[ [ 1. 4. 4. 16. 3.]
[ 1. 4. 18. 50. 13.]
[ 1. 3. 3. 12. 104.]
[ 4. 4. 11. 9. 6.]
[ 4. 6. 13. 13. 0.]]
```

S(s):

```
[ [ 0. -10.45813417 -17.73691075 -146.19910913
-25.12840992]
[ -9.95825442 -17.48892023 -139.53954278 -452.63283712
-112.67009584]
[ -9.95361602 -12.28189404 -20.58335329 -109.61455277
-923.92452153]
[ -33.72961984 -31.76262503 -84.24751853 -47.61172285
-23.15340484]
[ -39.56820366 -57.87269502 -116.51011412 -79.73390984
0. ]]
```

V(s):

```
[ [ 0. -2.61453354 -4.43422769 -9.13744432 -8.37613664]
[ -9.95825442 -4.37223006 -7.75219682 -9.05265674 -8.66693045]
[ -9.95361602 -4.09396468 -6.86111776 -9.13454606 -8.88388963]
[ -8.43240496 -7.94065626 -7.65886532 -5.29019143 -3.85890081]
[ -9.89205091 -9.64544917 -8.96231647 -6.13337768 0. ]]
```

Converged after 418 episodes.

Epoch 418:

Episode data array:

```
[[1, 25, 0, 0.9, 0.0]]
```


N(s):

```
[ [ 15. 228. 66. 63. 120.]  
  [ 49. 152. 74. 85. 129.]  
  [ 45. 76. 19. 28. 119.]  
  [ 38. 36. 27. 53. 84.]  
  [ 19. 24. 26. 43. 12.]]
```

S(s):

```
[ [ 0. -234.45813417 -135.53691075 -273.56910913  
    -993.59380045]  
  [ -57.95825442 -298.68892023 -291.29954278 -572.99783712  
    -1086.80898637]  
  [ -108.94361602 -210.11189404 -75.60735329 -152.97455277  
    -952.42452153]  
  [ -135.56531984 -141.81062503 -127.60751853 -131.21172285  
    -101.15340484]  
  [ -94.89297366 -131.58449502 -141.21011412 -109.73390984  
    0. ]]
```

V(s):

```
[ [ 0. -1.02832515 -2.05358956 -4.34236681 -8.27994834]  
  [-1.18282152 -1.96505869 -3.93648031 -6.74115102 -8.42487586]  
  [-2.42096924 -2.76463018 -3.97933438 -5.46337688 -8.00356741]  
  [-3.56750842 -3.93918403 -4.72620439 -2.47569288 -1.2042072 ]  
  [-4.99436703 -5.48268729 -5.43115824 -2.55195139 0. ]]
```