## **Al Assignment 3**

# Vikranth Udandarao 2022570

**Theory** 

A 1 Assignment 3

hat proved sampling > P(M) = Gount (M) strengths > coast difficult to implement. May require large samples per accurate estimates -> wood for easy perobability Zers effecture jos ments with low prebability events, eg - P(moisel) -> unbiased Estimates 11- Rejection Sampling -> Accepting sample with a perdability given by -> f Caccept) = min (1, q cn)/mp(n)) - whose -m = por posed distribution constant officieny whele Long under Target
Aone under Peroposal Areq Weaknesses Strengths Can be Turnfficient if rejecteation grate is high. > Good Jos distributions with which applies hore. May waste ramples for distribution

calculate the values for all mariables - M.P.S. and then update each variable as -M, CHA M P (M, M2 (H) , --)-N2 (+t) M P(M2/M, (++1),---) and iterate for many steps-Weaknesses Strengths -) Good and useful for estimating More Complem to implement joint distribution 
-) Efficient for high dimensional Requires superification date with conditional perdoability of conditional perdoabilities 1c using Bayes Theorem -P'CM = AUT MP: Businon = P(M= win) PCP= Burning [M=NIS) = 0,80 to,20 = 0, 60. Nor of ferains = 30 N= 100 PCP=lienouse | M= Train) Expected numbers = 30×0,4=12

P. CM=Rig N P= Runners) = P(M=atg)

P(M=Toraln NP=Levius) = P(P=Reinuse | M=Toraln)

P(M=Toraln NP=Levius) = P(P=Reinuse | M=Toraln)

= 0-4 00 to-3 = 0-12 = 127
I when we increase the sample rise our accuses dosen to true value and more data near dosen to true value and more data near the normalise intended erreas decreases of and the tory of our dataset - Race overton. In the case of precision, the rose of our dataset - Race overton inthe less perdoals 1/14 (ap = 0-2) are ostinated perdoals 1/14 (ap = 0-2) are ostinated perdoals 1/14 (ap = 0-2) are ostinated

	books
	B = People reading journate regularly
	J= People access academi Journals Regularly
_	C = Person participates in book like
a)	
	P(B=6 J=j)= 0.780
2.	P (J=j   B= D = 0+
3.	P(C=e B=6) = 0.320
4.	PC J=1/1 B= 76) = 0.227
5.	P(7(8: 16 V.J:)) - anoto 0.090
6	P(J-j/8=76) - 02880 0.716
7.	P(C=c / J=j) - maga 0.088
8	P(C=c V J-j) = acréroro 0.631
9.	P(J=11:0 = 0.400
10.	P(J=) = 0500
11.	P(C17B) = 0.0044

2 6- Checking the amons of perobability. 1-P(n) =0 for any event of n\_ Lo satisfied for all statements 2- Law of mutual Enclusivity- $P(X \cap Y) = P(X) + P(Y) - P(X \cup Y)$ normally and P(XUY) - P(A) + P(B) of nutually and nutually FO9 -P(B=6)-P (B=60 J=J) = P(J=J | B=6) + TR=6) P(B=6 UJ=1) = P(B=6) + P(J=5) PCB=6NJ=J) -1 0.11=0.5 +0.6 PCB) -> P(B=26) = 0-6833-PLUNTS + PC-BAJ) + P(BAJ)

7 Mg n, 727 fo -663-0-910-

= 0.227+0.6 P(B) + on4 P(B)\_

John all statements. hor, p(BU5) = 0-91  $\rho(7BN7) = 0.09$  hor, p(BUJ) + P(7BN)

S, P(RVJ)+P(¬BN¬J)=1\_ -D C) mutually Endusing quent\_

3. Sum of porobabilities of all possible outcomes is 1.
use god above - smallowy all stustements

6 B	5	C	Sont Pooleb. 4.9
-	26.5		, 100,000,000
Ves	Vy	Yes	PC BNS 1 () = 0-088
40	40	As.	p(8 a 5 n7c) = 0,186
40	N∞	y es	P(Bn75n()=0.132
Yes	No	No	P(Bn-5n()= 0.278
No	Yes	40	P(7B751()= 0,000
No	40	No	P(7805070)=0.25
Np	No	40	P(787757()=0,0004
Mo	No	Np	P(7B17517()=0.089

d- for ordependence of 5 and C-

P(COB) = P(CIB) PCB) = 0.2186

P((178) = P((178) P(78) = 0-004) P(C) = P((18) + P((178)\_

-) 6,2186 +0,0614 = 0.22

=- P(C)-P(J) = 0-110 to -088

=== Not ridependent.

for J and B,

P(T) XP(B) = 0.5 X0-683 + 0.4 X0-67

=-Not Indergendent.

For redefendence of Band Cz

PCBOC) = PCB) - PCC), then

P(B) = 0,22 P(B) = 0,22

P(R10) = P(B) - P(E)

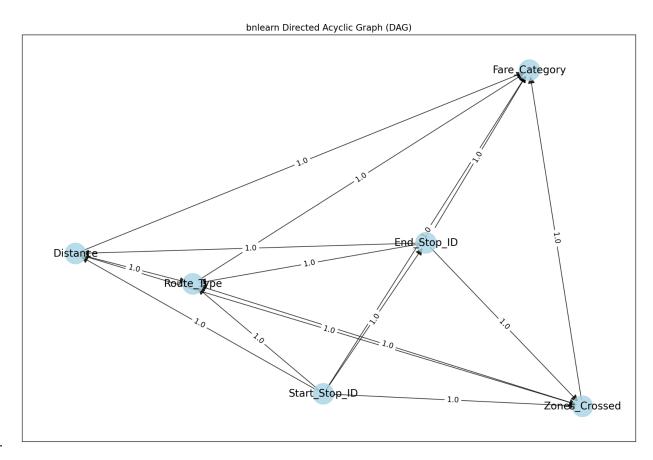
-) P(BAC) = 0.2186 \$0.3240-13

Is Not Thdesendent

pcB) = pgrobability adversiet attack is a backtone it tack ((A) advenced persons believes. pim) = Pordonbildy of a misclassification Now, by Bayes' Rule -PCA (M) = PCM (A) PCA) p(m) P(M) = P(M:IA)PCA) + P(MIB)PCB) - P(MIANB)P(ANB) P(B), P(A) & P(M) 1 Likelihoods .adversal probability of mischnification given PCM(B) - Peropability of mischention gives itakhors P(M A NB) - Probability of misotorifiation given adverted

pertubation gives abscention of M which c- Effect of conditioning on B-P(AIM) = P(MIA)P(A) PCM) p (M) is dependent on B-6 f(m) = P(M/A) P(A) + P(M/B) P(B) - P.CM (AMB) PCAMB) We see PCB) of heads to a P (M) Therene herre; PCA IM will decrease also leading PCM/A) to decrease V PCA(M)PCM) = PCM/A) P(A)

### Coding



**Pruning Methodology: Edge Pruning** 

- Edge Pruning involves iteratively removing edges from the Bayesian Network and evaluating the impact on the model's structure score (BIC - Bayesian Information Criterion).
- The objective is to retain edges that significantly contribute to the network while removing those that add complexity without improving predictive performance.

#### Implementation Steps:

- The initial Bayesian Network was defined with all possible feature dependencies (DAG\_edges).
- o For each edge in the network:
  - i. The edge was temporarily removed, creating a pruned network.
  - ii. The structure score (BIC) of the pruned network was computed after fitting it to the data.

1.

- iii. If the pruned network's BIC was higher than the previous best score, the pruned network was retained as the new best model.
- The process was repeated until no further improvement in the structure score was observed.

#### **Model Evaluation:**

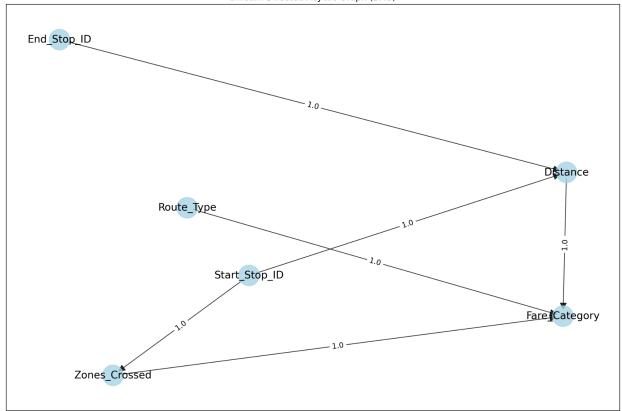
 Bayesian Information Criterion (BIC) was used as the metric to evaluate model performance. A higher BIC score indicates a better balance between model fit and complexity.

#### Visualization:

 The pruned network structure was visualized to provide a clear understanding of the optimized dependencies between features.

#### Results

- Initial Network Score (BIC): The score from the initial network.
- **Pruned Network Score (BIC):** After pruning, the best network achieved an increase in BIC score, indicating an improvement in performance.
- Performance Improvement:
  - **Efficiency:** The pruned network has fewer edges, resulting in faster model fitting and reduced computational cost.



#### **Initial Bayesian Network:**

- The base Bayesian Network was constructed using predefined relationships between the variables.
- It provided a foundational structure for fare classification but included redundancies and possibly unnecessary dependencies.

#### **Optimization Technique Applied:**

#### • Hill Climbing Algorithm:

- The Hill Climbing algorithm was applied to refine the network structure.
- This approach systematically searches for a structure with a higher Bayesian Information Criterion (BIC) score, indicating better model fit.

#### • Parameter Learning:

 The optimized structure was fitted with parameters to improve the Conditional Probability Distributions (CPDs) using the dataset.

#### **Code Implementation:**

- Structure learning with Hill Climbing was achieved using bn.structure\_learning.fit().
- The optimized DAG (Directed Acyclic Graph) was then used to define the new network.

The refined Bayesian Network was visualized to highlight structural improvements.

#### Visualization:

#### Base Network:

The initial network contains numerous edges, representing all potential relationships between variables. This design, while comprehensive, may include redundancies or relationships that are not significant. (Refer to Image: Base Bayesian Network)

#### Optimized Network:

The refined network has fewer edges, focusing on the most significant dependencies as identified through the Hill Climbing algorithm. This results in a simpler, more efficient structure. (Refer to Image: Optimized Bayesian Network)

#### **Performance Improvement:**

- The optimized network showed a higher BIC score compared to the initial network, indicating better data fit.
- Redundant relationships were pruned, which likely improves the computational efficiency of the model.

#### Discussion

#### 1. Advantages of Optimization:

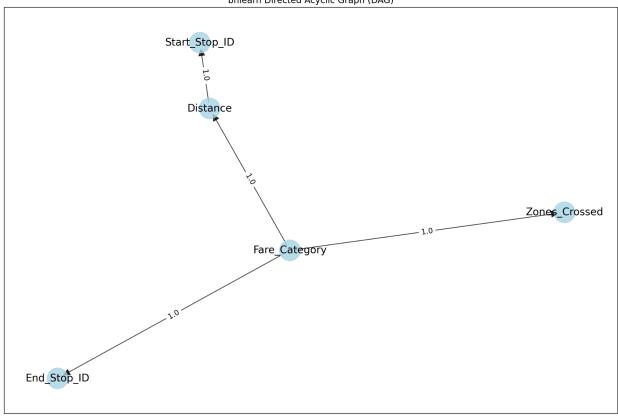
- The Hill Climbing method ensures that only the most critical relationships are retained, leading to a simplified structure.
- Fewer dependencies reduce computational overhead, enhancing the model's efficiency.

#### 2. Model Accuracy:

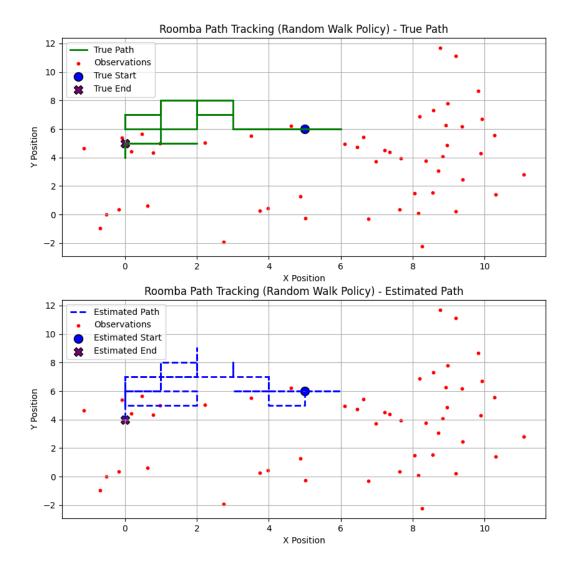
 By focusing on statistically significant relationships, the optimized network likely improves prediction accuracy. However, further validation on test data is required to quantify this improvement.

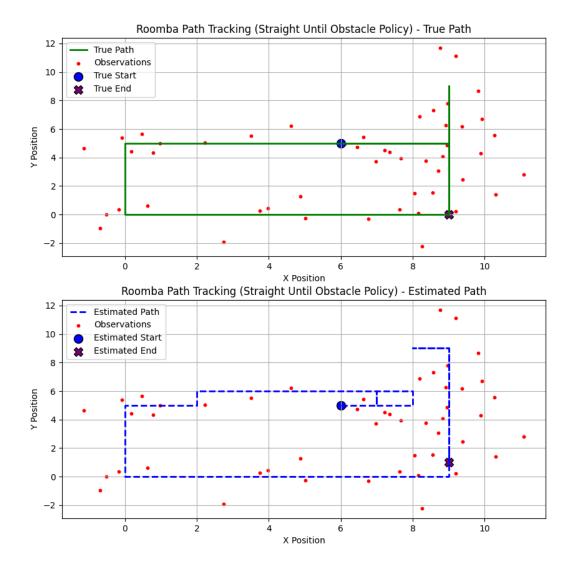
#### 3. Efficiency:

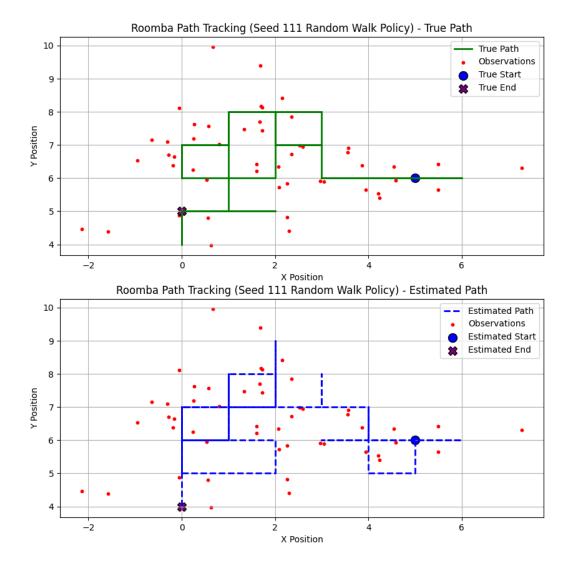
 The reduction in edges leads to faster parameter estimation and inference, making the optimized model more suitable for real-world applications.

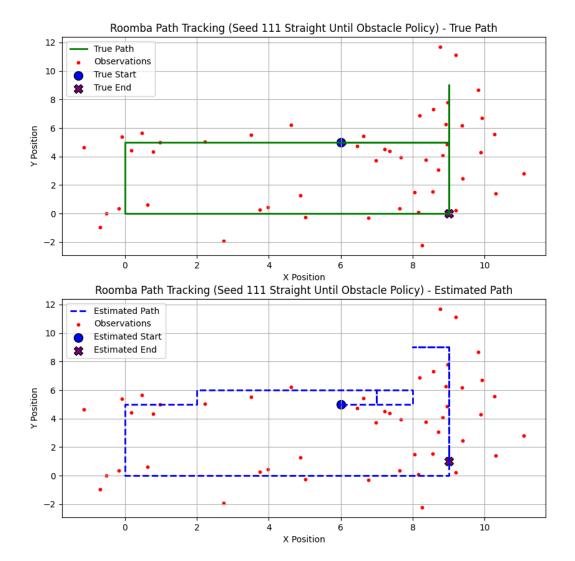


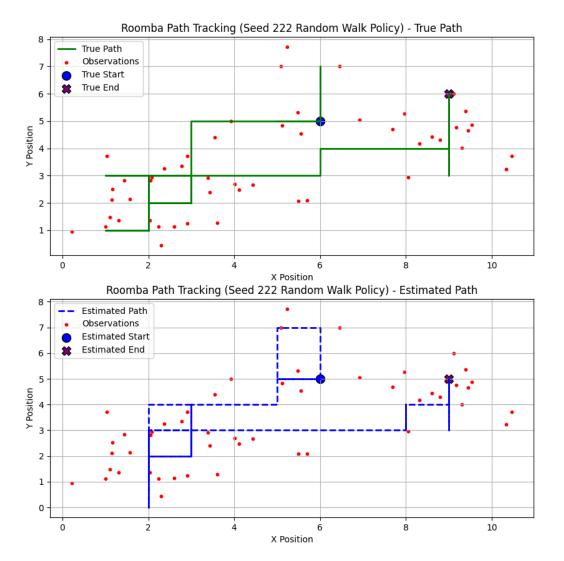
2. Seed values are 111, 222, 444, 765

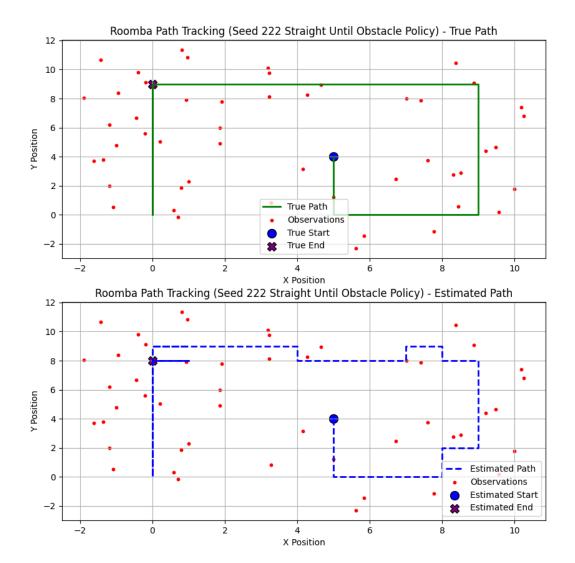


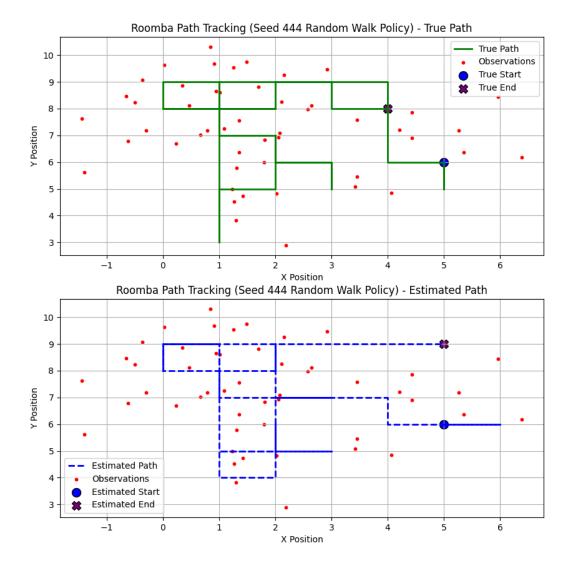


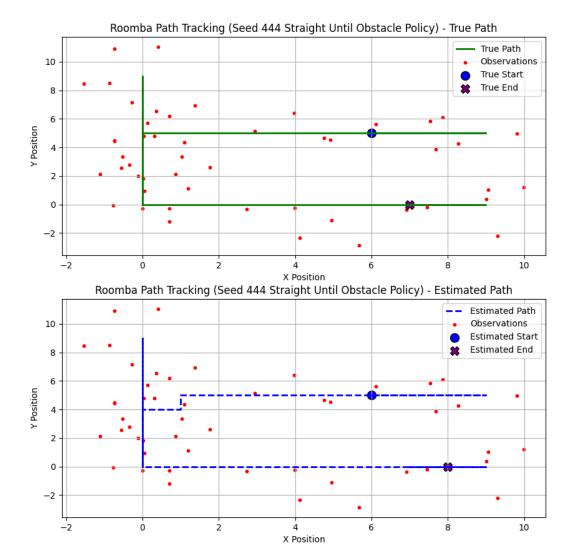


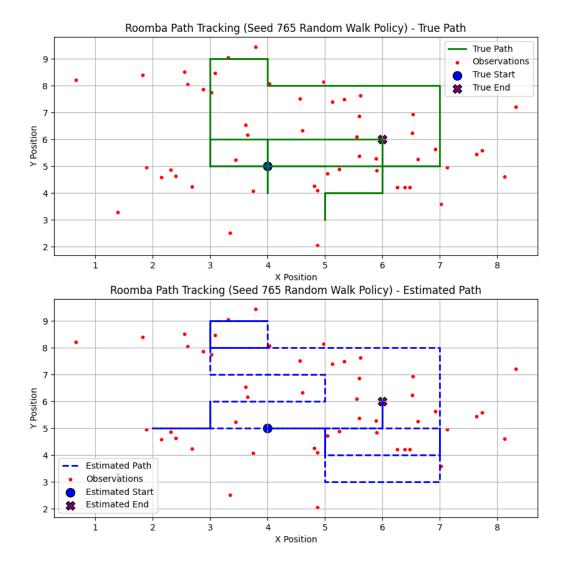


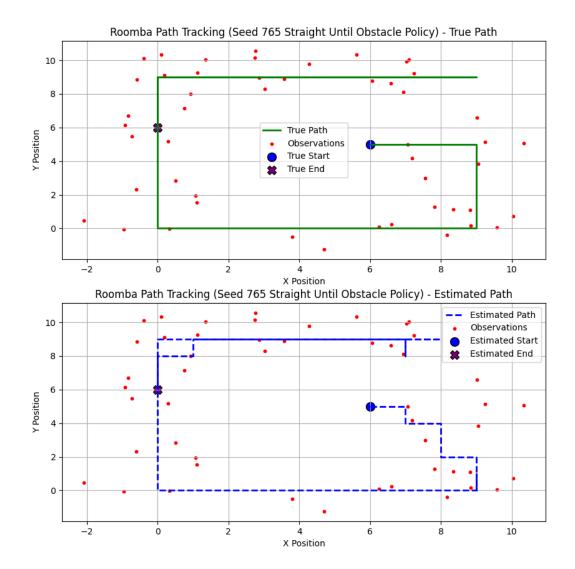












C.

Seed	Policy	Accuracy
111	Random Walk	34.00
111	Straight Until Obstacle	52.00
222	Random Walk	64.00

222	Straight Until Obstacle	72.00
444	Random Walk	54.00
444	Straight Until Obstacle	72.00
764	Random Walk	58.00
764	Straight Until Obstacle	84.00

#### **Key Observations**

#### 1. Random Walk:

- Accuracy varies significantly between seeds (34% to 64%).
- This is expected due to the high randomness in movements, making it difficult for the Viterbi algorithm to track accurately.

#### 2. Straight Until Obstacle:

- Accuracy is consistently higher compared to random\_walk, reaching as high as 84% for seed 765.
- This deterministic policy makes transitions more predictable, allowing the Viterbi algorithm to estimate the path more effectively.

#### 3. Comparison Between Policies:

- Straight Until Obstacle is more accurate overall due to its predictable movement pattern.
- Randomness in random\_walk introduces significant uncertainty, leading to lower accuracy.

```
seed, policy, estimated_path

111, random_walk, "[((5, 6), 'S'), ((5, 5), 'N'), ((4, 5), 'N'), ((4, 6), 'N'), ((3, 6), 'N'), ((4, 6), 'N'), ((4, 7), 'N'), ((4, 6), 'N'), ((5, 6), 'N'), ((6, 6), 'N'), ((5, 6), 'N'), ((4, 6), 'N'), ((4, 7), 'N'), ((3, 8), 'N'), ((3, 7), 'N'), ((2, 7), 'N'), ((2, 8), 'N'), ((2, 7), 'N'), ((1, 7), 'N'), ((1, 8), 'N'), ((2, 8), 'N'), ((2, 9), 'N'), ((2, 8), 'N'), ((2, 7), 'N'), ((1, 7), 'N'), ((1, 6), 'N'), ((1, 6), 'N'), ((1, 7), 'N'), ((1, 7), 'N'), ((1, 7), 'N'), ((1, 6), 'N'), ((1, 6), 'N'), ((1, 7), 'N'), ((1, 6), 'N'), ((1, 6), 'N'), ((1, 5), 'N'), ((0, 6), 'N'), ((1, 6), 'N'), ((2, 6), 'N'), ((2, 5), 'N'), ((1, 5), 'N'), ((0, 5), 'N'), ((0, 4), 'N')]"
```

```
111, straight_until_obstacle, "[((6, 5), 'E'), ((7, 5), 'S'), ((7, 6),
((5, 6), 'W'), ((4, 6), 'W'), ((3, 6), 'W'), ((2, 6), 'N'), ((2, 5), 'W')
((0, 1), "N"), ((0, 0), "E"), ((1, 0), "E"), ((2, 0), "E"), ((3, 0), "E")
((9, 0), 'S'), ((9, 1), 'S'), ((9, 2), 'N'), ((9, 1), 'S'), ((9, 2),
((9, 8), 'S'), ((9, 9), 'W'), ((8, 9), 'E'), ((9, 9), 'N'), ((9, 8),
                                                                      'N'),
((9, 7), 'N'), ((9, 6), 'N'), ((9, 5), 'N'), ((9, 4), 'N'), ((9, 3), 'N'),
((9, 2), 'N'), ((9, 1), 'N')]"
222, random walk, "[((6, 5), 'E'), ((6, 6), 'N'), ((6, 7), 'N'), ((5, 7),
'N'), ((3, 3), 'N'), ((3, 4), 'N'), ((2, 4), 'N'), ((2, 3), 'N'), ((3, 3), (3, 3))
'N'), ((3, 2), 'N'), ((3, 3), 'N'), ((2, 3), 'N'), ((2, 2), 'N'), ((2, 3),
'N'), ((2, 2), 'N'), ((2, 1), 'N'), ((2, 2), 'N'), ((2, 3), 'N'), ((3, 3),
'N'), ((4, 3), 'N'), ((5, 3), 'N'), ((6, 3), 'N'), ((7, 3), 'N'), ((8, 3),
'N'), ((8, 4), 'N'), ((8, 3), 'N'), ((8, 4), 'N'), ((9, 4), 'N'), ((9, 3),
'N'), ((9, 5), 'N')]"
222, straight until obstacle, "[((5, 4), 'N'), ((5, 3), 'N'), ((5, 2), 'N'),
((5, 1), "N"), ((5, 0), "E"), ((6, 0), "E"), ((7, 0), "E"), ((8, 0), "S"),
((8, 1), 'S'), ((8, 2), 'E'), ((9, 2), 'S'), ((9, 3), 'S'), ((9, 4), 'S'),
((8, 9), 'W'), ((7, 9), 'N'), ((7, 8), 'W'), ((6, 8), 'W'), ((5, 8), 'W')
((0, 9), 'E'), ((1, 9), 'W'), ((0, 9), 'N'), ((0, 8), 'N'), ((0, 7),
((0, 1), "N"), ((0, 0), "S"), ((0, 1), "S"), ((0, 2), "S"), ((0, 3), "S")
((1, 8), 'W'), ((0, 8), 'N')]"
444, random walk, "[((5, 6), 'S'), ((6, 6), 'N'), ((5, 6), 'N'), ((4, 6),
'N'), ((4, 7), 'N'), ((3, 7), 'N'), ((2, 7), 'N'), ((2, 8), 'N'), ((2, 9),
'N'), ((1, 6), 'N'), ((1, 5), 'N'), ((1, 4), 'N'), ((2, 4), 'N'), ((2, 5),
```

```
'N'), ((1, 7), 'N'), ((1, 8), 'N'), ((0, 8), 'N'), ((0, 9), 'N'), ((1, 9),
'N'), ((0, 9), 'N'), ((0, 8), 'N'), ((0, 9), 'N'), ((0, 8), 'N'), ((0, 9),
'N'), ((0, 9), 'N'), ((1, 9), 'N'), ((2, 9), 'N'), ((3, 9), 'N'), ((4, 9),
'N'), ((5, 9), 'N')]"
444, straight until obstacle, "[((6, 5), 'E'), ((7, 5), 'E'), ((8, 5), 'E'),
((9, 5), 'W'), ((8, 5), 'W'), ((7, 5), 'W'), ((6, 5), 'W'), ((5, 5),
((4, 5), 'W'), ((3, 5), 'W'), ((2, 5), 'W'), ((1, 5), 'N'), ((1, 4),
((0, 4), 'N'), ((0, 3), 'N'), ((0, 2), 'N'), ((0, 1), 'N'), ((0, 0), 'S'),
((0, 6), 'S'), ((0, 7), 'S'), ((0, 8), 'S'), ((0, 9), 'N'), ((0, 8), 'N'),
((0, 2), 'N'), ((0, 1), 'S'), ((0, 2), 'N'), ((0, 1), 'N'), ((0, 0), 'E'),
((1, 0), 'E'), ((2, 0), 'E'), ((3, 0), 'E'), ((4, 0), 'E'), ((5, 0),
((6, 0), 'E'), ((7, 0), 'E'), ((8, 0), 'E'), ((9, 0), 'W'), ((8, 0), 'W'),
765, random walk, "[(4, 5), W'), (5, 5), N'), (5, 4), N'), (5, 5),
'N'), ((6, 5), 'N'), ((6, 6), 'N'), ((6, 5), 'N'), ((5, 5), 'N'), ((5, 4),
'N'), ((5, 3), 'N'), ((6, 3), 'N'), ((7, 3), 'N'), ((7, 4), 'N'), ((7, 5),
'N'), ((7, 4), 'N'), ((6, 4), 'N'), ((5, 4), 'N'), ((5, 5), 'N'), ((4, 5),
'N'), ((3, 5), 'N'), ((3, 6), 'N'), ((3, 5), 'N'), ((2, 5), 'N'), ((3, 5),
'N'), ((3, 6), 'N'), ((4, 6), 'N'), ((5, 6), 'N'), ((5, 7), 'N'), ((4, 7),
'N'), ((7, 6), 'N'), ((7, 5), 'N'), ((7, 4), 'N'), ((7, 5), 'N'), ((6, 5),
'N'), ((6, 6), 'N')]"
765, straight until obstacle, "[((6, 5), 'E'), ((7, 5), 'N'), ((7, 4), 'E'),
((8, 4), "N"), ((8, 3), "N"), ((8, 2), "E"), ((9, 2), "N"), ((9, 1), "N")
((9, 0), 'S'), ((9, 1), 'N'), ((9, 0), 'W'), ((8, 0), 'W'), ((7, 0),
                                                                       'W'),
((6, 0), 'W'), ((5, 0), 'W'), ((4, 0), 'W'), ((3, 0), 'W'), ((2, 0),
((1, 0), \overline{)}, ((0, 0), \overline{)}, ((0, 1), \overline{)}, ((0, 2), \overline{)}, ((0, 3), \overline{)}
((0, 4), 'S'), ((0, 5), 'S'), ((0, 6), 'S'), ((0, 7), 'S'), ((0, 8),
((5, 9), 'E'), ((6, 9), 'E'), ((7, 9), 'N'), ((7, 8), 'S'), ((7, 9), 'E'),
```

```
((3, 9), 'W'), ((2, 9), 'W'), ((1, 9), 'W'), ((0, 9), 'N'), ((0, 8), 'N'), ((0, 7), 'N'), ((0, 6), 'N')]"
```