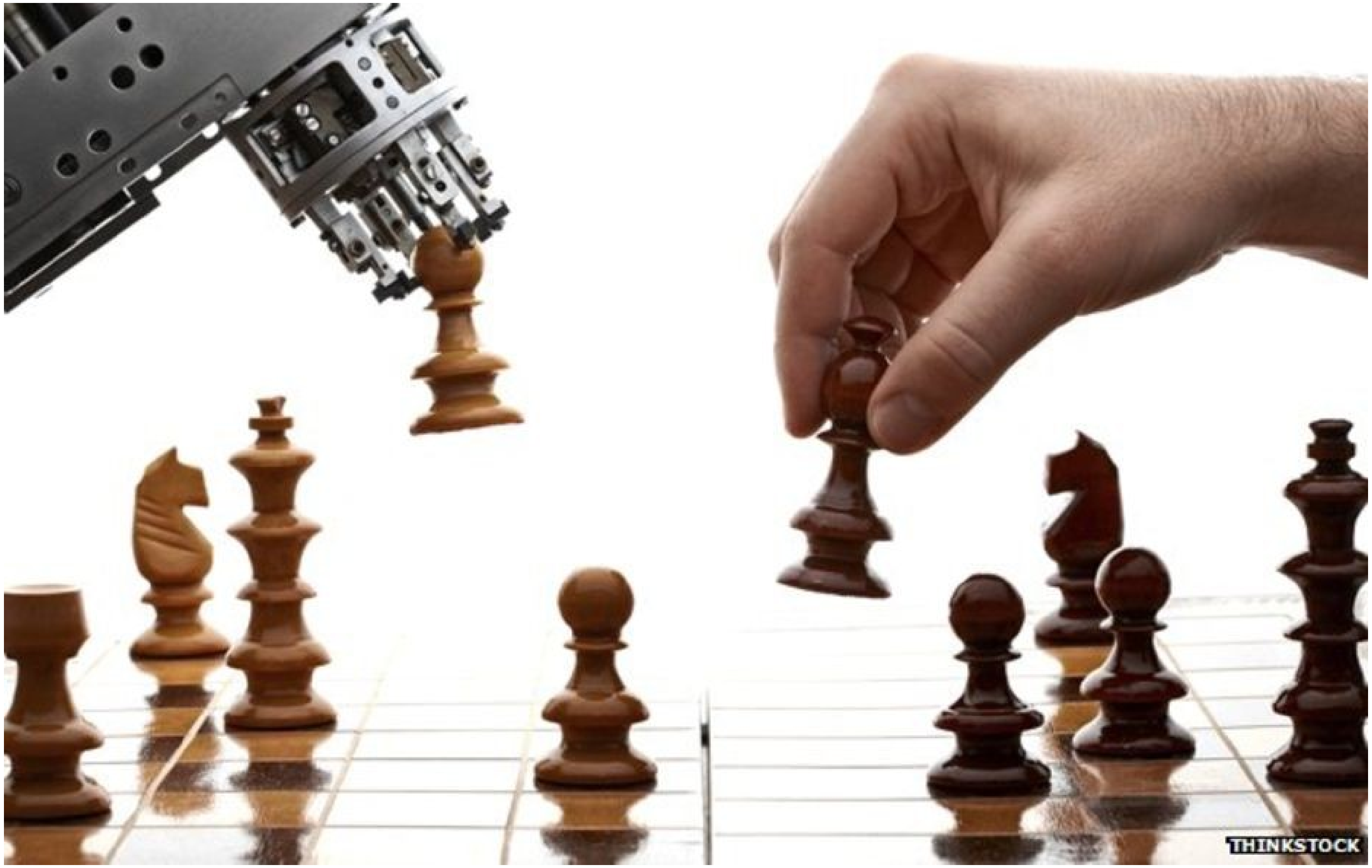
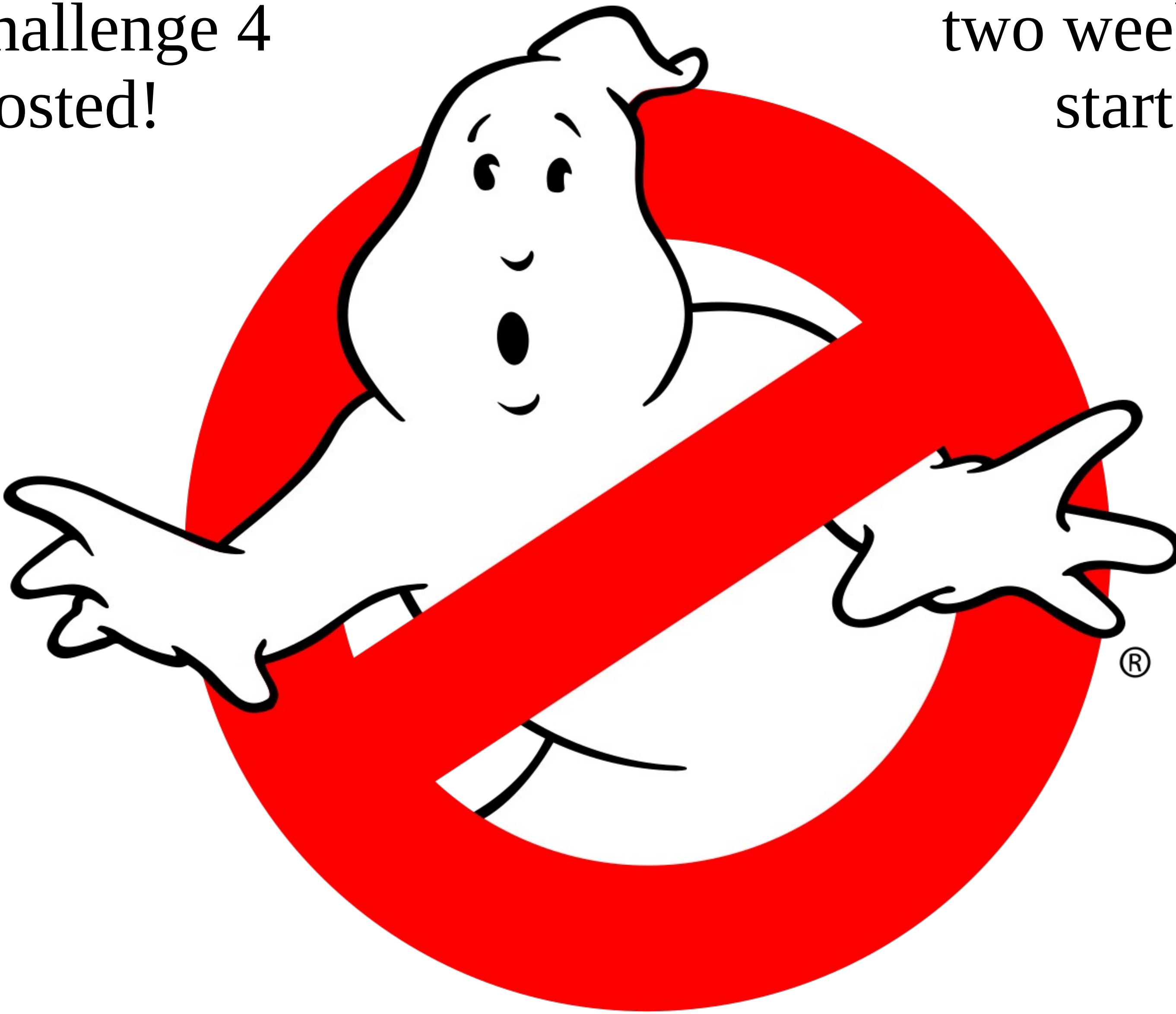


Introduction to Artificial Intelligence
COSC 4550 / COSC 5550

Professor Cheney
9/25/17

AI Challenge 4
posted!

two weeks this time,
start early!!!



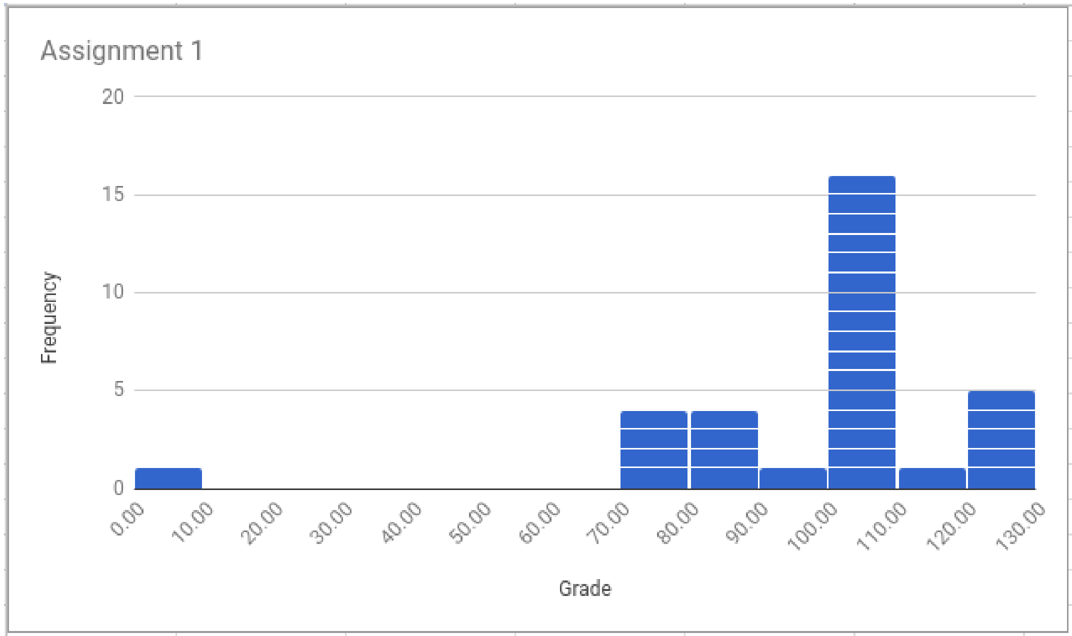
updated office hours:

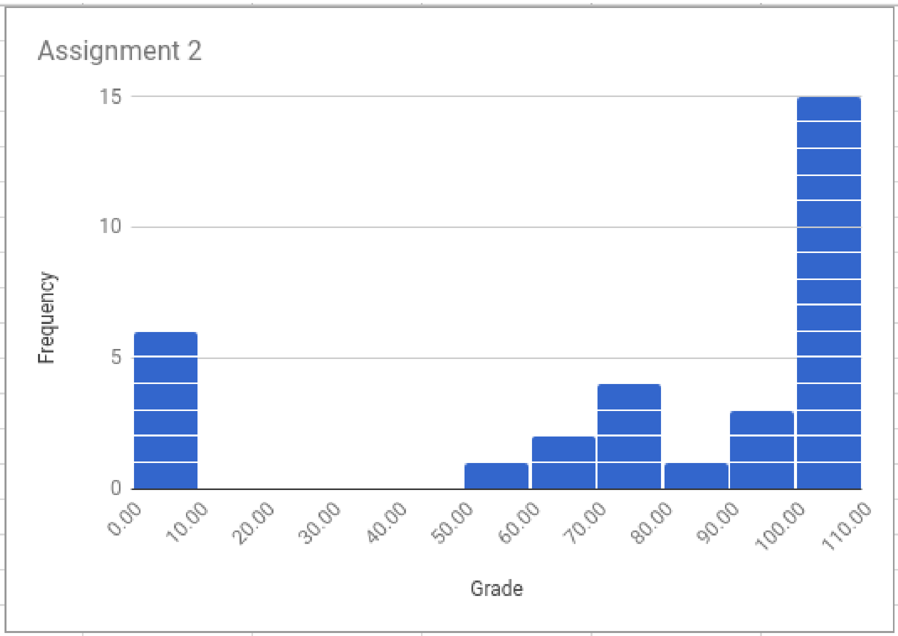
Monday 2-3pm has been moved to Wednesday 2-3pm

Film Crew in EvolvingAI Lab, Wednesday 10-11:15

(help setup Tuesday 11am)

so… how's it going?





	Mean	Median	Mode		Number Grade	Letter Grade
Assigment 1	96	100	100		93+	A
Assigment 2	73.84375	93	100		90+	A-
Assigment 3					87+	B+
Assigment 4					83+	B
Assigment 5					80+	B-
					77+	C+
					73+	C
					70+	C-
					67+	D+
					63+	D
					60+	D-
					< 60	F

my teaching grade…

mid-term teaching feedback linked on Piazza
(https://goo.gl/forms/W3MbDeQYH5xWyeU13)

what do you want to see more of?
what do you want to see less of?
what has been working for you?
what have you been struggling with?
what's been boring and repetitive?
etc.

course outline/timeline revisited

Part I: Artificial Intelligence
- Introduction
- Intelligent Agents

what are agents?
why would they need to do search?

Part II: Problem Solving
- Search
- Optimization
- Games

how do we even do search?
what's the best way to do it?

what if we have a simple setting, where we perfectly know the rules (i.e. model)?

~~**Part III: Knowledge, Reasoning, & Planning**~~

if we know things about the problem already, can we tell them to the agent, instead of making it learn them?

Part IV: Uncertainty and Reasoning
- Probability
- Bayesian Statistics
- Markov Models

if we're not sure of the model, but have a guess at how it should work, how can we update our causal understanding when new information comes?

Part V: Learning
- Unsupervised Learning
- Supervised Learning
- Reinforcement Learning

what if we have no idea (or prior assumptions) about how the world works – can we get the agent to learn correlations from the ground up?

Part VI: Communicating, Perceiving, & Acting
- Natural Language Processing
- Object Recognition
- Robotics

what additional tricks/techniques do we need to be able to apply these ideas to a variety of applications?

Bayes' rule

$$P(a \wedge b) = P(b \mid a) * P(a) = P(a \mid b) * P(b)$$

$$\frac{P(b \mid a) * P(a)}{P(a)} = \frac{P(a \mid b) * P(b)}{P(a)}$$

Bayes' rule

$$\boxed{P(b \mid a) = \frac{P(a \mid b) * P(b)}{P(a)}}$$

"This simple equation underlies most modern AI systems for probabilistic inference"

-R&N

it's just some math… what's the big deal???

$$P(b \mid a) = \frac{P(a \mid b) * P(b)}{P(a)}$$

$$P(\text{cause} \mid \text{effect}) = \frac{P(\text{effect} \mid \text{cause}) * P(\text{cause})}{P(\text{effect})}$$

$$P(\text{hypothesis} \mid \text{data}) = \frac{P(\text{data} \mid \text{hypothesis}) * P(\text{hypothesis})}{P(\text{data})}$$

$$P(\text{disease} \mid \text{symptoms}) = \frac{P(\text{symptoms} \mid \text{disease}) * P(\text{disease})}{P(\text{symptoms})}$$

for each $disease_i$ (in $disease_1$, $disease_2$, … , $disease_N$),
which one is most likley?

i.e. maximizes:

$$P(disease_i \mid symptoms) = \frac{P(symptoms \mid disease_i) * P(disease_i)}{P(symptoms)}$$

if you want to know the **cause** behind an observation...

you can figure this out if you know:

what causes tend to lead to what observations

new data

and

how likely you believed each of those causes was

(before you saw the observation)

prior belief

$$P(\text{disease}_i \mid \text{symptoms}) = \frac{P(\text{symptoms} \mid \text{disease}_i) * P(\text{disease}_i)}{P(\text{symptoms})}$$

example

a laptop manufacturer buys computer chips from two companies:

Company A sold them 100 chips, of which 5 were defective
Company B sold them 300 chips, of which 21 were defective

if I buy a laptop, what is the likelihood
that my chip came from each company?

$$P(\text{company}_A) = \frac{100}{300+100} = 0.25$$

$$P(\text{company}_B) = \frac{300}{300+100} = 0.75$$

a laptop manufacturer buys computer chips from two companies:

Company A sold them 100 chips, of which 5 were defective
Company B sold them 300 chips, of which 21 were defective

if I then observe that my processor is defective,
what is the new likelihood that my chip came from each company?

$$P(\text{company}_A \mid \text{defective}) = \frac{P(\text{defective} \mid \text{company}_A) * P(\text{company}_A)}{P(\text{defective})}$$

$$P(\text{company}_A \mid \text{defective}) = \frac{(5/100) * (100/400)}{P(\text{defective})}$$

$$P(\text{company}_B \mid \text{defective}) = \frac{(21/300) * (300/400)}{P(\text{defective})}$$

$$P(\text{company}_A \mid \text{defective}) = \frac{(5/100) * (100/400)}{P(\text{defective})} = \frac{0.0125}{P(\text{defective})}$$

$$P(\text{company}_B \mid \text{defective}) = \frac{(21/300) * (300/400)}{P(\text{defective})} = \frac{0.0525}{P(\text{defective})}$$

let's normalize it, so our probabilities add up to 1…

$$\text{normalizing constant } (\alpha) = \frac{1}{\dfrac{0.0125}{P(\text{defective})} + \dfrac{0.0525}{P(\text{defective})}} = \frac{P(\text{defective})}{0.065}$$

$$\alpha * P(\text{company}_A \mid \text{defective}) = \frac{0.0125}{\cancel{P(\text{defective})}} * \frac{\cancel{P(\text{defective})}}{0.065} = 0.192$$

$$\alpha * P(\text{company}_B \mid \text{defective}) = \frac{0.0525}{\cancel{P(\text{defective})}} * \frac{\cancel{P(\text{defective})}}{0.065} = 0.808$$

$$P(\text{hypothesis} \mid \text{data}) = \frac{P(\text{data} \mid \text{hypothesis}) * P(\text{hypothesis})}{\cancel{P(\text{data})}}$$

$$P(\text{company}_A \mid \text{defective}) = \frac{(5/100) * (100/400)}{\cancel{P(\text{defective})}}$$

$$P(\text{company}_B \mid \text{defective}) = \frac{(21/300) * (300/400)}{\cancel{P(\text{defective})}}$$

as long as we are normalizing in the end,
it's fine to ignore the likelihood of observing that data

(you know that it's equally likely in each scenario…
since it's a given that it has already happened)

based on our new observation
(that the chip was defective)

new data

we were able to update our prior belief
(25% sure chip came from A, 75% sure it came from B)

prior belief
distribution

to produce our post-observation belief
(19% sure chip came from A, 81% sure it came from B)

posterior belief
distribution

(i.e. we updated our beliefs based on new data!)

e.g. localization

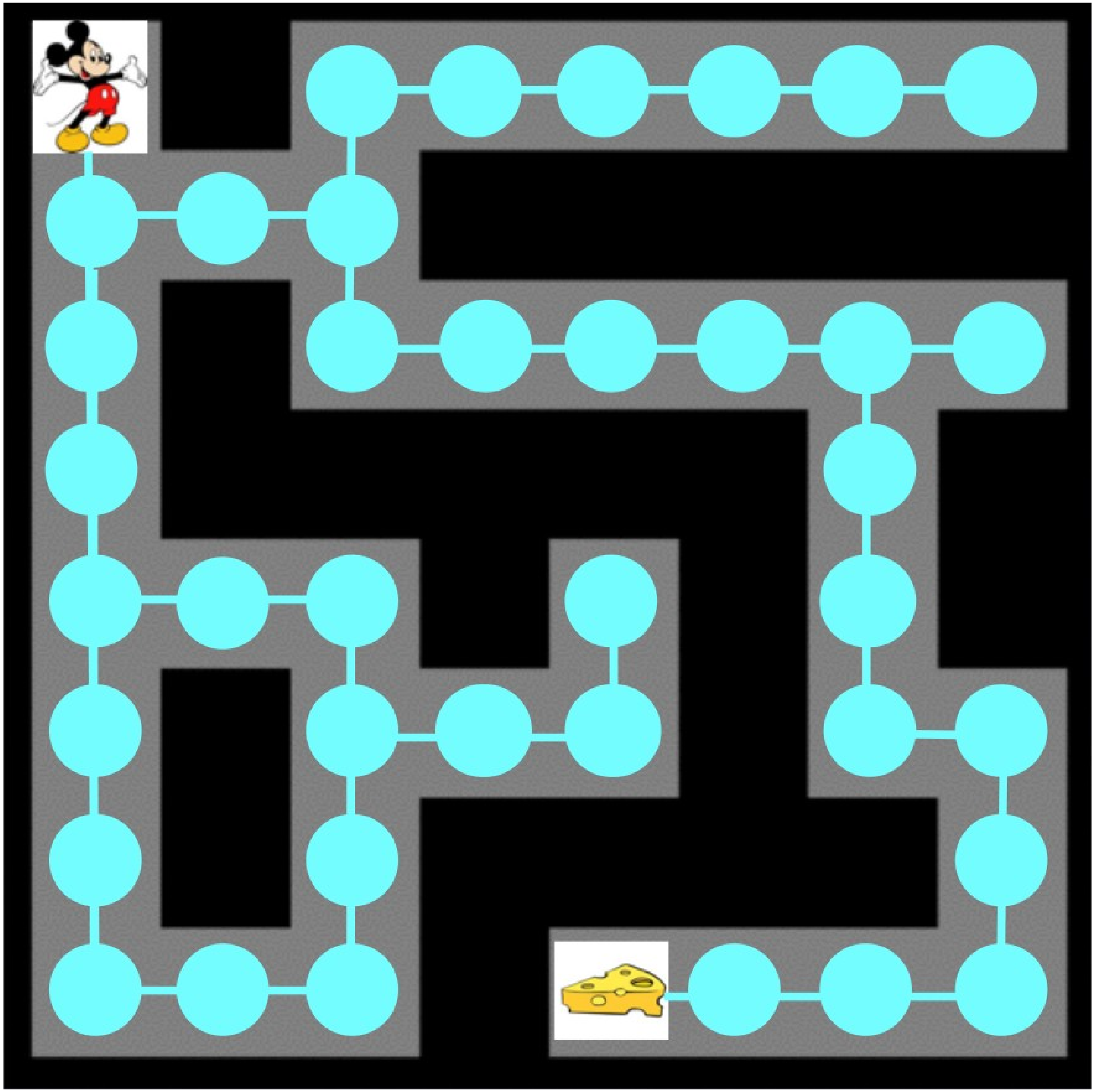
let's say you are a robot in a maze

but don't know where you are

and you can only look at one sensor at a time cause (hypothesis): I'm currently located at tile y

effect (observed data): I have x sensor readings

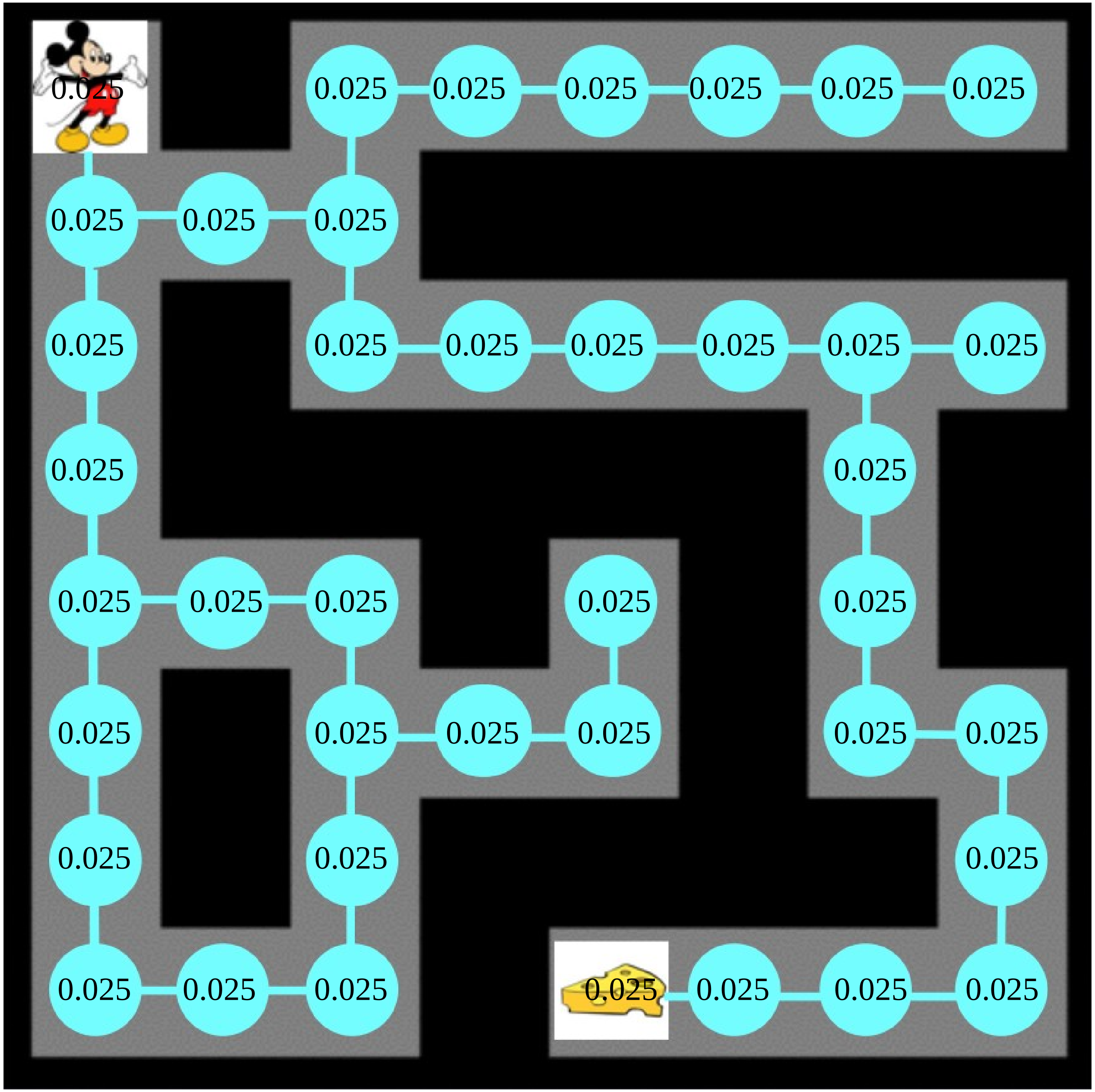
question:
where am I?



question:
where am I?

prior distribution:
who knows???
(all equally likely)

new data:
my sensors say there is
no wall above me
(with 90% accuracy)

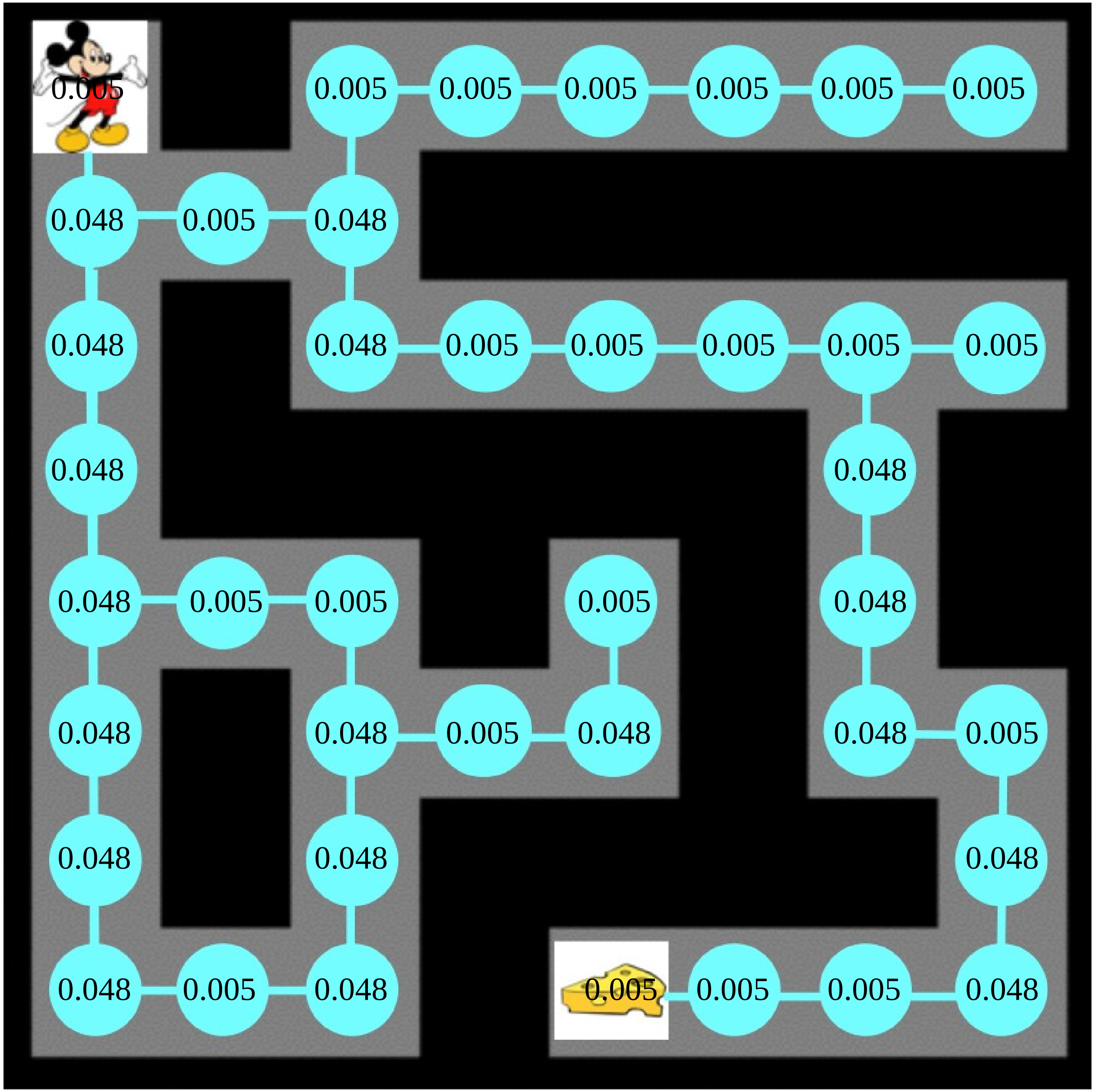


question:
where am I?

prior distribution:
who knows???
(all equally likely)

new data:
my sensors say there is no wall above me (with 90% accuracy)

posterior distribution:
I'm more likely to be in the tiles with nothing above them

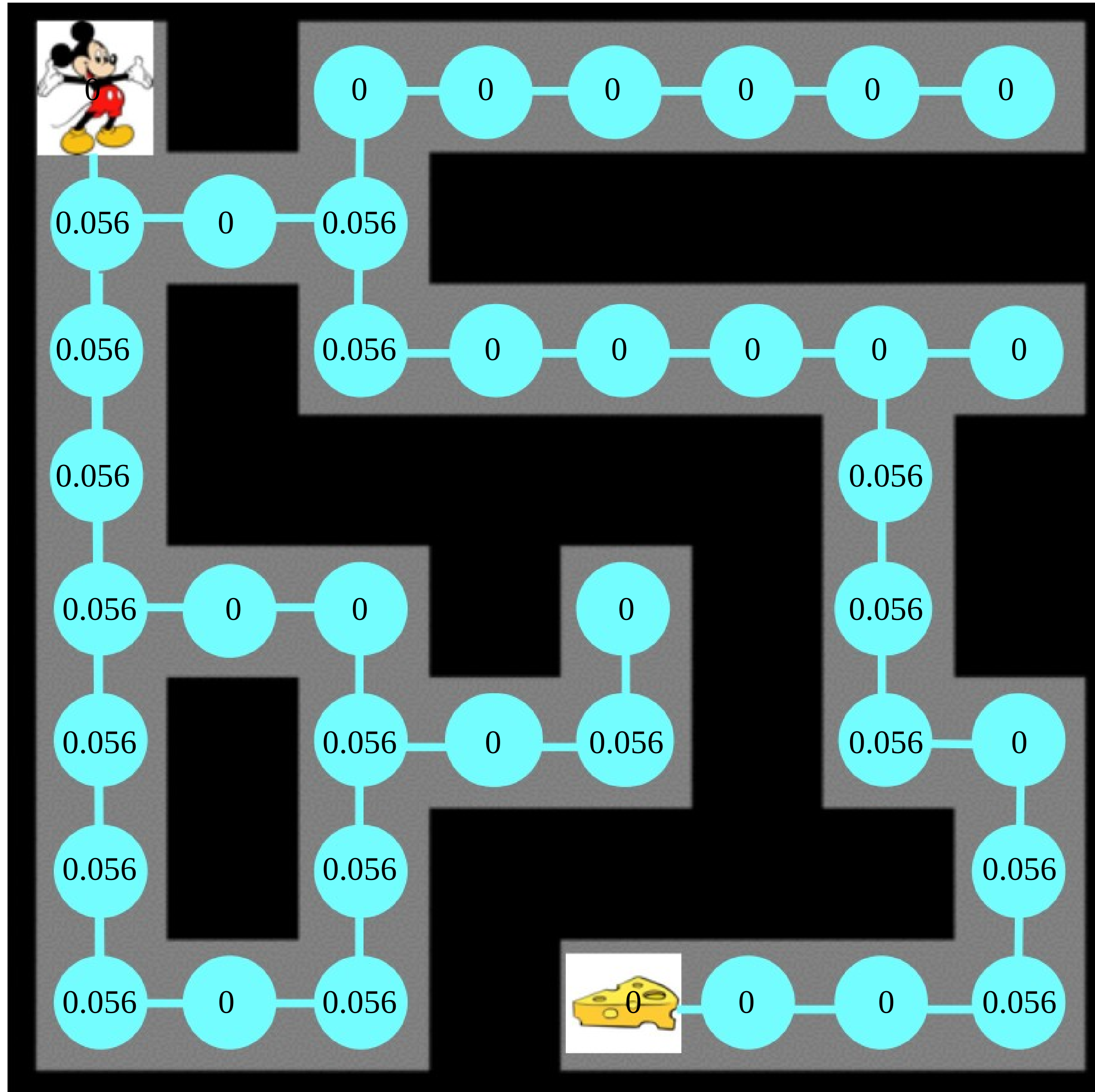


question:
where am I?

prior distribution:
who knows???
(all equally likely)

new data:
my sensors say there is
no wall above me
(with 100% accuracy)

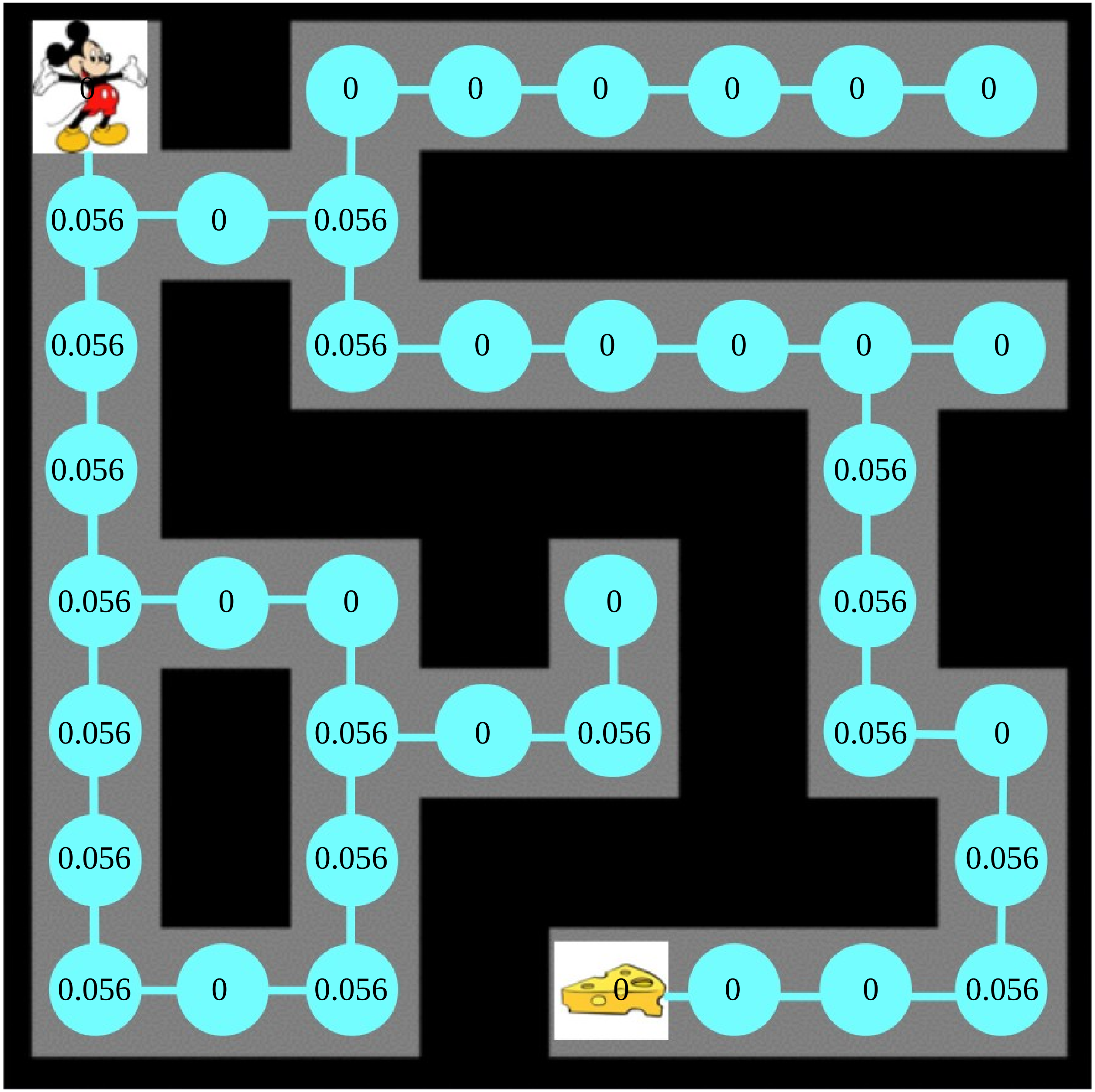
posterior distribution:
I'm more likely to be
in the tiles with nothing
above them



question:
where am I?

prior distribution:
I'm more likely to be in the tiles with nothing above them

new data:
my sensors say there is no wall to the left of me (with 100% accuracy)

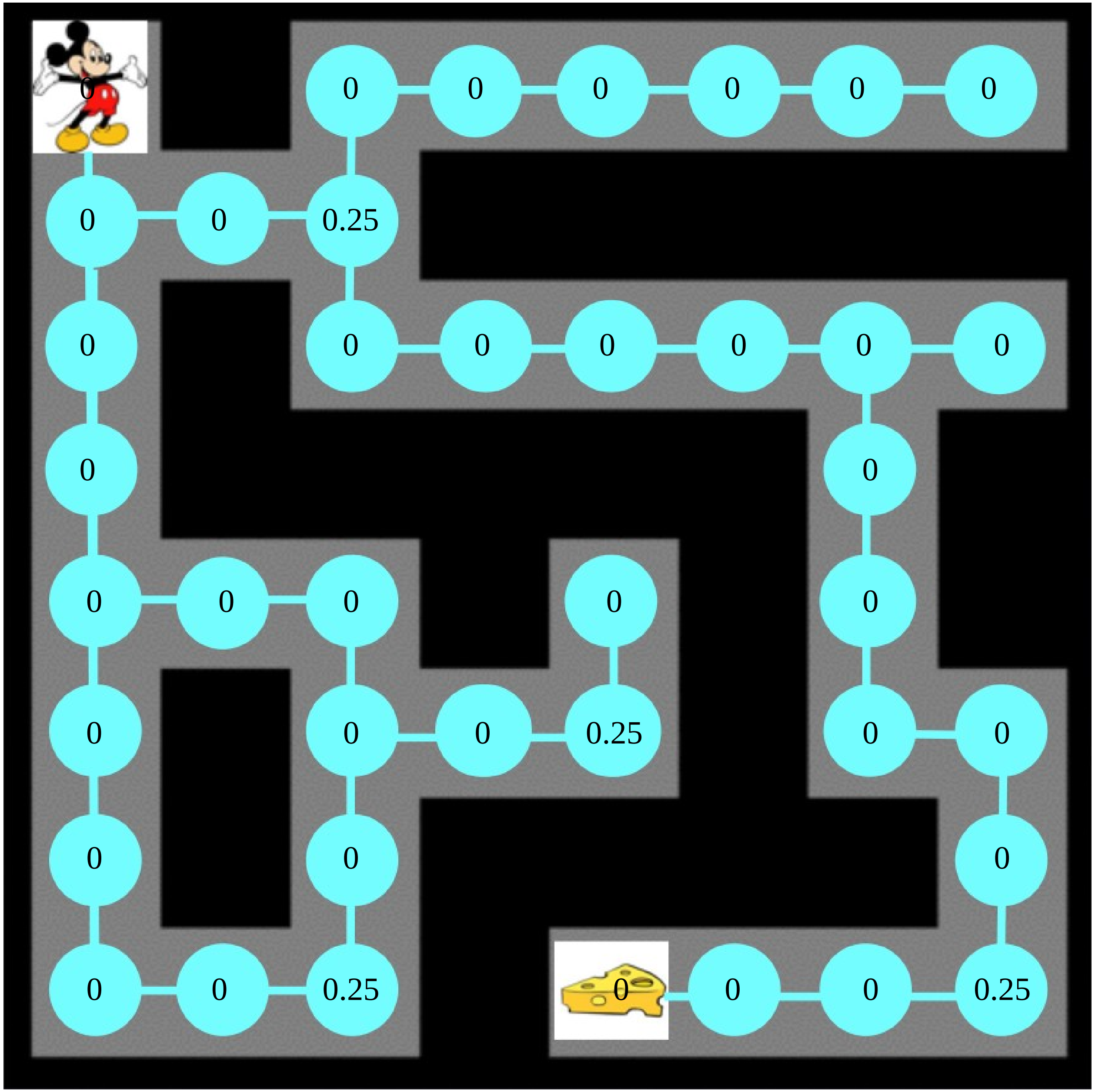


question:
where am I?

prior distribution:
I'm more likely to be in the tiles with nothing above them

new data:
my sensors say there is no wall to the left of me (with 100% accuracy)

posterior distribution:
now here is where I believe I am

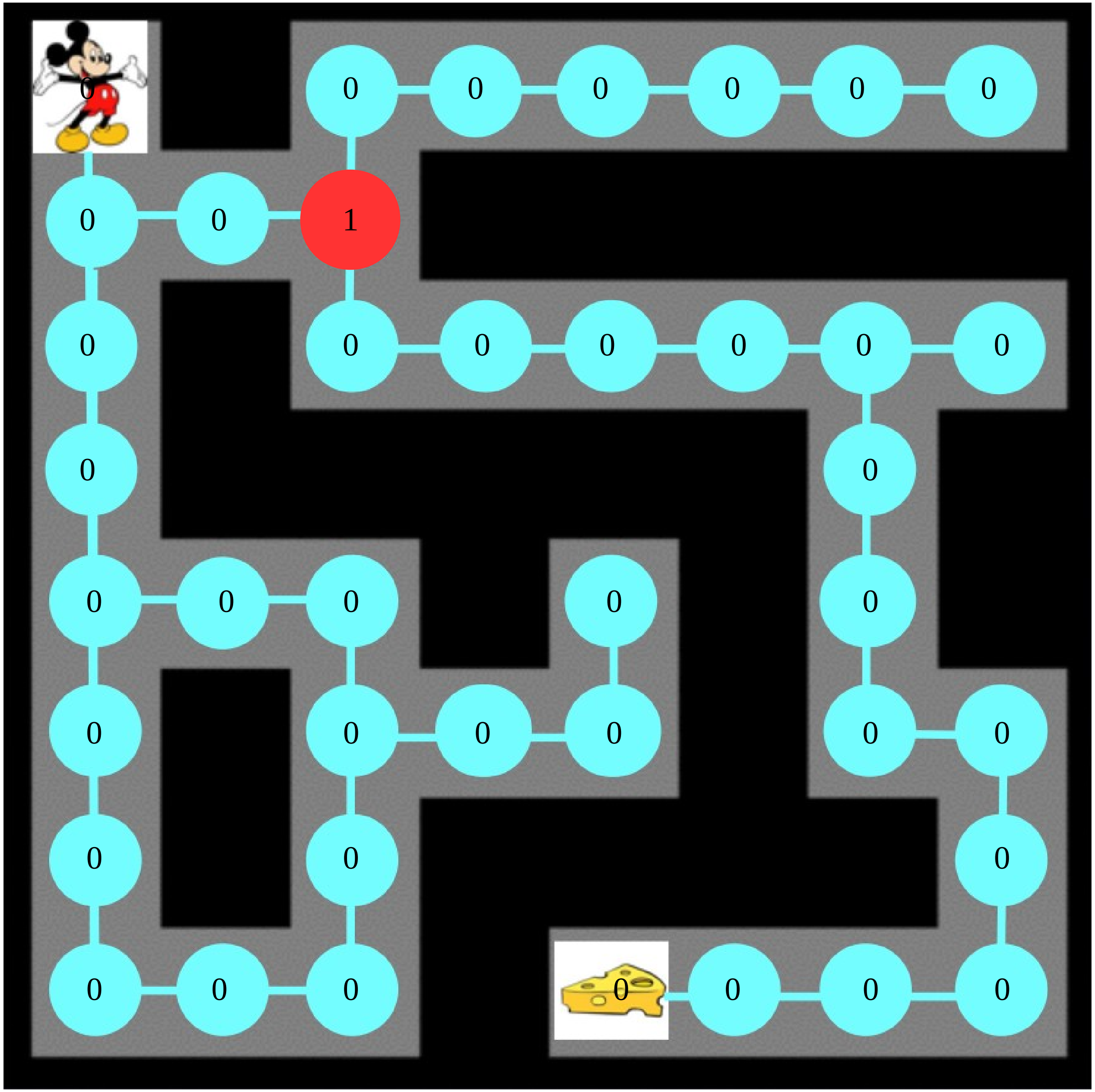


question:
where am I?

prior distribution:
now here is where
I believe I am

new data:
my sensors say there is
no wall below me
(with 100% accuracy)

posterior distribution:
I know I'm here!



Bayes' rule allows you to iteratively update your beliefs!