universität koblenz
weiter:denken

Network Theory and Dynamic Systems

11. Information Cascades

SOSE 2025

Dr. -Ing. Stefania Zourlidou

Institute for Web Science and Technologies
Universität Koblenz

WeST
Web Science and Technologies

Recap from Previous Lecture

- Ideas, Information, Influence
- Epidemic Spreading
  - Diseases
  - Rumor
- Opinion
- Search
  - Local Search
  - Searchability

Objectives of this Lecture

- what is information cascade?
- a simple herding experiment
- Bayes's Rule
  - what is it?
  - in the herding experiment
- a simple, general cascade model

1. Information Cascades

Following the Crowd (1/4)

- When people are connected by a network, it becomes possible for them to influence each other's behavior and decisions
- **Example 1**
  - Suppose that you are choosing a restaurant in an unfamiliar town, and based on your own research about restaurants, you intend to go to restaurant **A**
  - However, when you arrive you see that no one is eating in restaurant A, whereas restaurant B next door is nearly full
  - If you believe that other diners have tastes similar to yours, and that they too have some information about where to eat, it may be rational to join the crowd at **B** *rather than to follow your own information*

Following the Crowd (2/4)

- To see how this is possible…
  - Suppose that each diner has obtained independent but imperfect information about which of the two restaurants is better
  - Then if there are already many diners in restaurant B, the information that you can infer from their choices may be more powerful than your own private information, in which case it would in fact make sense for you to join them *regardless of your own private information*
  - In this case, we say that **herding**, or an **information cascade**, has occurred

Following the Crowd (3/4)

- **Information cascade**, comes from the work of Banerjee
  - the concept was also developed in other work around the same time by Bikhchandani, Hirshleifer, and Welch
- An information cascade has the potential to occur when people make decisions sequentially, with later people watching the actions of earlier people and from these actions inferring something about what the earlier people know
- In the restaurant example, when the first diners to arrive chose restaurant B, they conveyed information to later diners *about what they knew*
- A *cascade* then develops when *people abandon their own information in favor of inferences based on earlier people's actions*
- Individuals in a cascade are imitating the behavior of others, but it is *not* mindless imitation
  - Rather, it is the result of drawing rational inferences from limited information

Following the Crowd (4/4)

Example 2 (Milgram, Bickman, and Berkowitz in the 1960s)

- The experimenters had groups of people, ranging in size from just one person to as many as fifteen people, stand on a street corner and stare up into the sky
- They then observed how many passersby stopped and also looked up at the sky
- They found that with only one person looking up, very few passersby stopped
- If five people were staring up into the sky, then more passersby stopped, but most still ignored them
- Finally, with fifteen people looking up, they found that 45% of passersby stopped and also stared up into the sky
- The experimenters interpreted this result as demonstrating a social force for conformity that grows stronger as the group conforming to the activity becomes larger
- But another possible explanation – essentially, a possible mechanism giving rise to the conformity observed in this kind of situation – is rooted in the idea of information cascades

Informational vs. Direct-Benefit Effects (1/2)

universität koblenz
weiter:denken

- There is also a fundamentally different class of rational reasons why you may want to imitate what other people are doing
  - You may want to copy the behavior of others if there is **a direct benefit** to you from aligning your behavior with their behavior
- **Example**
  - Consider the first fax machines to be sold. A fax machine is useless if no one else owns one, and so in evaluating whether to buy one, it's very important to know whether there are other people who own one as well – not just because their purchase decisions convey information, but because they directly affect the fax machine's value to you as a product
  - A similar argument can be made for computer operating systems, social networking sites, and other kinds of technology where you directly benefit from choosing an option that has a large user population

Informational vs. Direct-Benefit Effects (2/2)

- **Example (cont)**
  - This type of direct-benefit effect is different from the informational effects we discussed previously: here, the actions of others affect your payoffs directly, rather than indirectly by changing your information
- Many decisions exhibit **both** information and direct-benefit effects
  - For example, in the technology-adoption decisions just discussed, you potentially learn from others' decisions and benefit from compatibility with them
- In some cases, the two effects are even in **conflict**
  - If you have to wait in a long line to get into a popular restaurant, you are choosing to let the informational benefits of imitating others outweigh the direct inconvenience (from waiting) that this imitation causes you

2. A Simple Herding Experiment

A Simple Herding Experiment (1/8)

- A simple herding experiment created by Anderson and Holt to illustrate how the mathematical models for information cascade work

(a) There is a decision to be made – for example, whether to adopt a new technology, wear a new style of clothing, eat in a new restaurant, or support a particular political position

(b) People make decisions sequentially, and each person can observe the choices made by those who acted earlier

(c) Each person has some private information that helps guide their decision

(d) A person can’t directly observe the private information that other people *know*, but he or she can make inferences about this private information from what they *do*

A Simple Herding Experiment (2/8)

- Imagine the experiment taking place in a classroom, with a large group of students
- The experimenter puts an urn at the front of the room with three marbles hidden in it; she announces that there is a 50% chance that the urn contains two red marbles and one blue marble ("majority-red" urn), and a 50% chance the urn contains two blue marbles and one red marble ("majority-blue" urn)
- One by one, each student comes to the front of the room and draws a marble from the urn
- He looks at the color and then places it back in the urn **without** showing it to the rest of the class
- The student then guesses whether the urn is majority-red or majority-blue and publicly announces this guess to the class
  - The public announcement is the key part of the setup: the students who have not yet had their turn don't get to see which colors the earlier students draw, but they do get to hear the guesses that are being made
  - This parallels the previous example with the two restaurants: one by one, each diner needs to guess which is the better restaurant, and while they don't get to see the reviews read by the earlier diners, they do get to see which restaurant these earlier diners chose

A Simple Herding Experiment (3/8)

Let's now consider what we should expect to happen when this experiment is performed

- We will assume that all the students reason *correctly* about what to do when it is their turn to guess, using everything they have heard so far

- Things are fairly straightforward for the first two students; they become interesting once we reach the third student
- The First Student
  - The first student should follow a simple decision rule for making a guess: if he sees a red marble, it is better to guess that the urn is majority-red; and if he sees a blue marble, it is better to guess that the urn is majority-blue
  - This means the first student's guess conveys perfect information about what he has seen

A Simple Herding Experiment (4/8)

universität koblenz
weiter:denken

- The Second Student
  - If the second student sees the same color that the first student announced, then her choice is simple: she should guess this color as well. However, suppose she sees the opposite color – say that she sees red while the first guess was blue. Since the first guess was exactly what the first student saw, the second student can essentially reason as though she got to draw twice from the urn, seeing blue once and red once. In this case, she is indifferent about which guess to make; we will assume in this case that she breaks the tie by guessing the color she saw. Thus, whichever color the second student draws, her guess too conveys perfect information about what she has seen
- The Third Student
  - Things start to get interesting here. If the first two students have guessed opposite colors, then the third student should just guess the color he sees, since it will effectively break the tie between the first two guesses
  - But suppose the first two guesses have been the same – say they've both been blue – and the third student draws red. Since we've decided that the first two guesses convey perfect information, the third student can reason in this case as though he saw three draws from the urn: two blue, and one red. Given this information, he should guess that the urn is majority-blue, ignoring his own private information

A Simple Herding Experiment (5/8)

- The Third Student (cont.)
  - More generally, the point is that when the first two guesses are the same, the third student should guess this color as well, regardless of which color he draws from the urn. And the rest of class will only hear his guess; they don’t get to see which color he’s drawn
  - In this case, an *information cascade has begun*
  - The third student makes the same guess as the first two, *regardless of which color he draws from the urn*, and hence regardless of his own private information

A Simple Herding Experiment (6/8)

- The Fourth Student and Onward
  - Let's consider just the "interesting" case of the third student discussed above, in which the first two guesses were the same – suppose they were both blue. In this case, we've argued that the third student will also announce a guess of blue, regardless of what he actually saw
  - Now consider the situation faced by the fourth student, getting ready to make a guess having heard three guesses of "blue" in a row. She knows that the first two guesses conveyed perfect information about what they saw. She also knows that, given this, the third student was going to guess "blue" no matter what he saw – so his guess conveys no information
  - As a result, the fourth student is in exactly the same situation – from the point of view of making a decision – as the third student. Whichever color she draws, it will be outweighed by the two draws of blue by the first two students, and so she should guess "blue" regardless of what she sees
  - This will continue with all the subsequent students: if the first two guesses are "blue," then everyone in order will guess "blue" as well (a completely symmetric thing happens if the first two guesses are "red"). An information cascade has taken hold: no one is under the illusion that every single person is drawing a blue marble, but once the first two guesses turn out "blue," the future announced guesses become worthless and so everyone's best strategy is to rely on the limited genuine information they have available

A Simple Herding Experiment (7/8)

- An information cascade can occur as long as the first two guesses are the same. While the setting of this experiment is highly stylized, it reveals important insights about information cascades

  1. First, it demonstrates how easily cascades can emerge under the right structural conditions. It also illustrates how a *strange pattern* — where a large group of students all make the same guess — can arise even when each individual is acting rationally
  2. Second, it highlights how information cascades can result in suboptimal decisions
     - Imagine an urn that contains mostly red marbles
     - Each student draws a marble in private and announces a guess about the urn's majority color
     - Suppose the chance of drawing a blue marble is 1/3. There's a $1/3 \times 1/3 = 1/9$ probability that both of the first two students independently draw blue marbles
     - If that happens, both guess "blue." Because later students observe only the previous guesses and not the actual draws, they will also guess "blue," regardless of what they draw. As a result, everyone will follow the initial guesses, and the group ends up making a wrong collective decision — even though the urn is majority-red
     - This shows that increasing the number of participants doesn't necessarily prevent error

A Simple Herding Experiment (8/8)

3. Third, it illustrates that cascades – despite their potential to produce long runs of conformity – can be fundamentally very fragile

- For example, suppose that, in a class of 100 students, the first two guesses are “blue,” and all subsequent guesses are proceeding – as predicted – to be “blue” as well.
- Now, suppose that students 50 and 51 both draw red marbles, and they each “cheat” by showing their marbles directly to the rest of the class
- In this case, the cascade has been broken: when student 52 gets up to make a guess, she has four pieces of genuine information to go on: the colors observed by students 1, 2, 50, and 51. Since two of these colors are blue and two are red, she should make the guess based on her own draw, which will break the tie
- The point is that everyone knew the initial run of 49 “blue” guesses had very little information supporting it, and so it was easy for a fresh infusion of new information to overturn it. This is the essential fragility of information cascades: even after they have persisted for a long time, they can be overturned with comparatively little effort

3. Bayes’s Rule:
A Model of Decision Making under Uncertainty

Bayes's Rule: A Model of Decision Making under Uncertainty (1/5)

What is the probability this is the better restaurant, given the reviews I've read and the crowds I see in each one?

Or

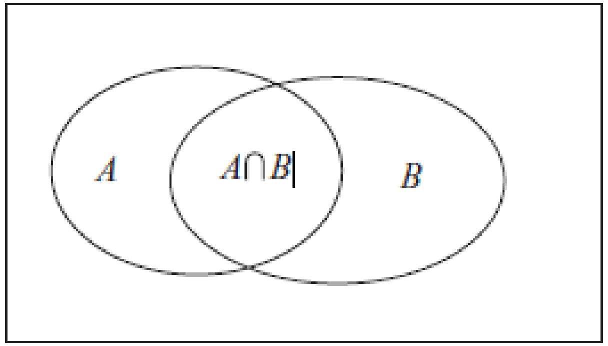
What is the probability this urn is majority-red, given the marble I just drew and the guesses I've heard?

Bayes's Rule: A Model of Decision Making under Uncertainty (2/5)

- **Conditional Probability and Bayes' Rule:** We will compute the probabilities of various *events*, and use these probabilities to reason about decision making
  - In the context of the previous experiment, an event could be "the urn is majority-blue," or "the first student draws a blue marble"
  - Given any event A, we will denote its probability of occurring by Pr[A]
  - Whether an event occurs or not is the result of certain random outcomes
  - We imagine a large sample space in which each point in the sample space consists of a particular realization for each of these random outcomes

Bayes's Rule: A Model of Decision Making under Uncertainty (3/5)

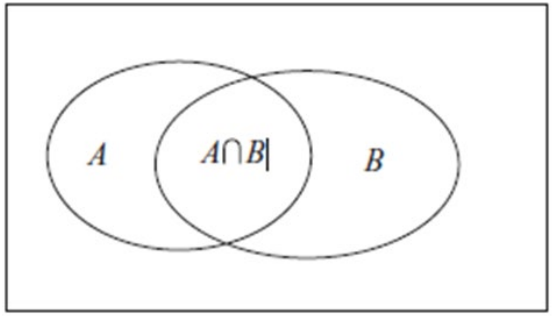
- Given a sample space, events can be pictured graphically as in Figure
  - the unit-area rectangle in the figure represents the sample space of all possible outcomes
  - the event A is then a region within this sample space: the set of all outcomes where event A occurs
  - the probability of A corresponds to the area of this region
- The relationship between two events (e.g., A and B) can be illustrated graphically as well
- The area where they overlap corresponds to the joint event when both A and B occur
  - This event is the intersection of A and B, and it's denoted by A ∩ B



Bayes's Rule: A Model of Decision Making under Uncertainty (4/5)

- **Example**: A may be the event that the urn in the previously mentioned experiment is majority-blue, and B may be the event that the ball you've drawn is blue
- We will refer to this quantity as the *conditional probability of A given B* and denote it by Pr [A | B]
- We can think of this as the fraction of the area of region B occupied by A ∩ B, and so we define

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$



- Similarly, the conditional probability of B given A is

$$\Pr[B \mid A] = \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{\Pr[A \cap B]}{\Pr[A]}$$

Bayes's Rule: A Model of Decision Making under Uncertainty (5/5)

- where the second equality follows simply because A ∩ B and B ∩ A are the same set. Rewriting previous Equations , we have

$$\Pr[A \mid B] \times \Pr[B] = \Pr[A \cap B] = \Pr[B \mid A] \times \Pr[A]$$

- and therefore, dividing through by Pr [B]

$$\Pr[A \mid B] = \frac{\Pr[A] \times \Pr[B \mid A]}{\Pr[B]}$$

- This Equation is called *Bayes' rule*
- When we want to make explicit that we're interested in the effect of event B on the probability of an event A, we refer to Pr [A] as the *prior probability* of A, since it reflects our understanding of the probability of A without knowing anything about whether B has occurred
- We refer to Pr [A | B] as the *posterior probability* of A given B, since it reflects our new understanding of the probability of A now that we know B has occurred

An Example Bayes's Rule (1/4)

We will apply Bayes' rule when a decision maker is assessing the probability that a particular choice is the best one, *given* the event that he has received certain private information, observed certain other decisions, or both

- Suppose that in some city 80% of taxi cabs are black and the remaining 20% are yellow
- A witness to a hit-and-run accident involving a taxi states that the cab involved was yellow. Suppose that eyewitness testimony is imperfect in the sense that witnesses sometimes misidentify the colors of cabs
- Suppose that if a taxi is yellow then a witness will claim it is yellow after the fact 80% of the time; if it is black, they will claim it is black 80% of the time
- Interpreting eyewitness testimony is at some level a question of conditional probability: what is the probability the cab is yellow (or black), given that the witness says it is yellow?

An Example Bayes's Rule (2/4)

- Let "true" denote the true color of the cab, and let "report" denote the reported color of the cab
- Let Y denote yellow and B denote black
- We are looking for the value of Pr [*true* = Y | *report* = Y]
- Applying Bayes´ Equation with A equal to the event *true* = Y and B equal to the event *report* = Y, we have

$$\Pr[\mathit{true} = \mathrm{Y} \mid \mathit{report} = \mathrm{Y}] = \frac{\Pr[\mathit{true} = \mathrm{Y}] \times \Pr[\mathit{report} = \mathrm{Y} \mid \mathit{true} = \mathrm{Y}]}{\Pr[\mathit{report} = \mathrm{Y}]}$$

- We've been told that
  - Pr [*report* = Y | *true* = Y] is 0.8 (the accuracy of eyewitness testimony)
  - Pr [*true* = Y] is 0.2 (the frequency of yellow taxi cabs)

An Example Bayes's Rule (3/4)

- We now need to figure out the denominator
- There are two ways for a witness to report that a cab is yellow:
  - One is for the cab to actually be yellow, and the other is for it to actually be black. The probability of getting a report of yellow via the former option is

$$\Pr[\mathit{true} = \mathrm{Y}] \times \Pr[\mathit{report} = \mathrm{Y} \mid \mathit{true} = \mathrm{Y}] = 0.2 \times 0.8 = 0.16$$

  - And the probability of getting a report of yellow via the latter option is

$$\Pr[\mathit{true} = \mathrm{B}] \times \Pr[\mathit{report} = \mathrm{Y} \mid \mathit{true} = \mathrm{B}] = 0.8 \times 0.2 = 0.16$$

- The probability of a report of yellow is the sum of these two probabilities,

$$\begin{aligned}\Pr[\mathit{report} = \mathrm{Y}] &= \Pr[\mathit{true} = \mathrm{Y}] \times \Pr[\mathit{report} = \mathrm{Y} \mid \mathit{true} = \mathrm{Y}] \\ &\quad + \Pr[\mathit{true} = \mathrm{B}] \times \Pr[\mathit{report} = \mathrm{Y} \mid \mathit{true} = \mathrm{B}] \\ &= 0.2 \times 0.8 + 0.8 \times 0.2 = 0.32.\end{aligned}$$

An Example Bayes's Rule (4/4)

- We can now put everything together

$$\begin{aligned}\Pr[\mathit{true} = \mathrm{Y} \mid \mathit{report} = \mathrm{Y}] &= \frac{\Pr[\mathit{true} = \mathrm{Y}] \times \Pr[\mathit{report} = \mathrm{Y} \mid \mathit{true} = \mathrm{Y}]}{\Pr[\mathit{report} = \mathrm{Y}]} \\ &= \frac{0.2 \times 0.8}{0.32} \\ &= 0.5.\end{aligned}$$

- If the witness says the cab was yellow, the conclusion is that it is in fact equally likely to have been yellow or black
- Because the frequency of black and yellow cabs makes black substantially more likely in the absence of any other information (0.8 versus 0.2), the witness's report had a substantial effect on our beliefs about the color of the particular cab involved
- But the report should not lead us to believe that the cab was in fact more likely to have been yellow than black

A Second Example: Spam Filtering (1/4)

- One application where Bayes' rule has been very influential is in email *spam detection* – automatically filtering unwanted e-mail out of a user's incoming e-mail stream
- Bayes' rule was a crucial conceptual ingredient in the first generation of e-mail spam filters, and it continues to form part of the foundation for many spam filters
  - Suppose that you receive a piece of e-mail whose subject line contains the phrase "check this out" (a popular phrase among spammers)
  - Based just on this (and without looking at the sender or the message content), what is the chance the message is spam?

$$\Pr\left[\textit{message is spam} \mid \textit{subject contains ``check this out''}\right]$$

  - To make this equation and the ones that follow a bit simpler to read, let's abbreviate *as*

$$\Pr\left[\textit{spam} \mid \textit{``check this out''}\right]$$

A Second Example: Spam Filtering (2/4)

- To determine this value, we need to know some facts about your e-mail and the general use of the phrase “check this out” in subject lines
- Suppose that 40% of all your e-mail is spam and the remaining 60% is e-mail you want to receive
- Also, suppose that 1% of all spam messages contain the phrase “check this out” in their subject lines, while 0.4% of all nonspam messages contain this phrase
- Writing these in terms of probabilities
  - Pr [*spam*] = 0.4; this is the prior probability that an incoming message is spam
  - also $\Pr[\text{“}\textit{check this out}\text{”} \mid \textit{spam}] = 0.01$
  - and $\Pr[\text{“}\textit{check this out}\text{”} \mid \textit{not spam}] = 0.004$

A Second Example: Spam Filtering (3/4)

- We're now in a situation completely analogous to the calculations involving eyewitness testimony: we can use Bayes' rule to write

$$\Pr[spam \mid \text{"check this out"}] = \frac{\Pr[spam] \times \Pr[\text{"check this out"} \mid spam]}{\Pr[\text{"check this out"}]}$$

- Based on what we know, we can determine that the numerator is 0.4 × 0.01 = 0.004
- For the denominator, as in the taxi cab example, we note that there are two ways for a message to contain "check this out" – either by being spam or by not being spam

$$\begin{aligned}\Pr[\text{"check this out"}] &= \Pr[spam] \times \Pr[\text{"check this out"} \mid spam] \\ &\quad + \Pr[not\ spam] \times \Pr[\text{"check this out"} \mid not\ spam] \\ &= 0.4 \times 0.01 + 0.6 \times 0.004 = 0.0064.\end{aligned}$$

- Dividing numerator by denominator, we get our answer

$$\Pr[spam \mid \text{"check this out"}] = \frac{0.004}{0.0064} = \frac{5}{8} = 0.625$$

A Second Example: Spam Filtering (4/4)

- Although spam (in this example) forms less than half of your incoming e-mail, a message whose subject line contains the phrase “check this out” is – in the absence of any other information – more likely to be spam than not
- In practice, spam filters built on Bayes’ rule look for a wide range of different signals in each message – the words in the message body, the words in the subject, properties of the sender (do you know them? what kind of an e-mail address are they using?), properties of the mail program used to compose the message, and other features
- Each signal provides its own estimate for whether the message is spam or not, and spam filters then combine these estimates to arrive at an overall guess about whether the message is spam
- For example, if we also knew that the message came from someone you send mail to every day, then presumably this competing signal – strongly indicating that the message is not spam – should outweigh the presence of the phrase “check this out” in the subject

4. Bayes's Rule in the Herding Experiment

Bayes's Rule in the Herding Experiment (1/5)

universität koblenz
weiter:denken

- Each student's decision is intrinsically based on determining a conditional probability
  - he is trying to estimate the conditional probability that the urn is majority-blue or majority-red, given what she has seen and heard
- To maximize her chance of winning the monetary reward for guessing correctly, she should guess majority-blue if

$$\Pr[\textit{majority-blue} \mid \textit{what she has seen and heard}] > \frac{1}{2}$$

- and guess majority-red otherwise
- If the two conditional probabilities are both exactly 0.5, then it doesn't matter what she guesses
- Also we know the prior probabilities of majority-blue and majority-red are each 1/ 2

$$\Pr[\textit{majority-blue}] = \Pr[\textit{majority-red}] = \frac{1}{2}$$

Bayes's Rule in the Herding Experiment (2/5)

- Also, based on the composition of the two kinds of urns

$$\Pr[\mathit{blue} \mid \mathit{majority\text{-}blue}] = \Pr[\mathit{red} \mid \mathit{majority\text{-}red}] = \frac{2}{3}$$

- Suppose that the first student draws a blue marble. He therefore wants to determine Pr [majority-blue | blue] and, just as in the examples from, he can use Bayes' rule to calculate

$$\Pr[\mathit{majority\text{-}blue} \mid \mathit{blue}] = \frac{\Pr[\mathit{majority\text{-}blue}] \times \Pr[\mathit{blue} \mid \mathit{majority\text{-}blue}]}{\Pr[\mathit{blue}]}$$

- The numerator is $\frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$. For the denominator, we reason just as we did before, by noting that there are two possible ways to get a blue marble – if the urn is majority-blue or if it is majority-red

$$\begin{aligned}\Pr[\mathit{blue}] &= \Pr[\mathit{majority\text{-}blue}] \times \Pr[\mathit{blue} \mid \mathit{majority\text{-}blue}] \\ &\quad + \Pr[\mathit{majority\text{-}red}] \times \Pr[\mathit{blue} \mid \mathit{majority\text{-}red}] \\ &= \frac{1}{2} \times \frac{2}{3} + \frac{1}{2} \times \frac{1}{3} = \frac{1}{2}.\end{aligned}$$

Bayes's Rule in the Herding Experiment (3/5)

- The answer Pr [blue] = 1 / 2 makes sense, given that the roles of blue and red in this experiment are completely symmetric. Dividing numerator by denominator, we get

$$\Pr[\textit{majority-blue} \mid \textit{blue}] = \frac{1/3}{1/2} = \frac{2}{3}.$$

- Since this conditional probability is greater than 1/2 , we get the intuitive result that the first student should guess majority-blue when he sees a blue marble. Note that in addition to providing the basis for the guess, Bayes' rule provides a probability, namely 2/3 , that the guess will be correct
- The calculation is very similar for the second student, and we skip it so as to move on to the calculation for the third student, where a cascade begins to form
- Let's suppose, that the first two students have announced guesses of blue, and the third student draws a red marble. As we discussed there, the first two guesses convey genuine information, so the third student knows that there have been three draws from the urn, consisting of the sequence of colors blue, blue. and red. What he wants to know is

$$\Pr[\textit{majority-blue} \mid \textit{blue}, \textit{blue}, \textit{red}]$$

Bayes's Rule in the Herding Experiment (4/5)

- So as to make a guess about the urn. Using Bayes' rule we get

$$\Pr[\textit{majority-blue} \mid \textit{blue, blue, red}] = \frac{\Pr[\textit{majority-blue}] \times \Pr[\textit{blue, blue, red} \mid \textit{majority-blue}]}{\Pr[\textit{blue, blue, red}]}$$

- Since the draws from the urn are independent, the probability Pr[blue, blue,red| majority-blue] is determined by multiplying the probabilities of the three respective draws together:

$$\Pr[\textit{blue, blue, red} \mid \textit{majority-blue}] = \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3} = \frac{4}{27}$$

- To determine Pr [blue, blue, red], as usual we consider the two different ways this sequence could have happened – if the urn is majority-blue or if it is majority-red

$$\begin{aligned}\Pr[\textit{blue, blue, red}] &= \Pr[\textit{majority-blue}] \times \Pr[\textit{blue, blue, red} \mid \textit{majority-blue}] \\ &\quad + \Pr[\textit{majority-red}] \times \Pr[\textit{blue, blue, red} \mid \textit{majority-red}] \\ &= \frac{1}{2} \times \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{6}{54} = \frac{1}{9}.\end{aligned}$$

Bayes's Rule in the Herding Experiment (5/5)

- Plugging all this back into Equation, we get

$$\Pr\left[\textit{majority-blue} \mid \textit{blue}, \textit{blue}, \textit{red}\right] = \frac{\frac{4}{27} \times \frac{1}{2}}{\frac{1}{9}} = \frac{2}{3}$$

- Therefore, the third student should guess majority-blue (from which he will have a 2/3 chance of being correct)
- This outcome confirms our intuitive observation that the student should ignore what he sees (red) in favor of the two guesses he's already heard (both blue)
- Finally, once these three draws from the urn have taken place, all future students will have the same information as the third student, and so they will all perform the same calculation, resulting in an information cascade of blue guesses

5. A Simple, General Cascade Model

A Simple, General Cascade Model (1/12)

Let's return to the motivation for the herding experiment

- The experiment served as a stylized metaphor for any situation in which people make decisions sequentially, basing these decisions on a combination of their own private information and observations of what earlier people have done
- We now formulate a model that covers such situations in general
  - We will see that Bayes' rule predicts in this general model that cascades will form, with probability tending to 1 as the number of people goes to infinity

A Simple, General Cascade Model (2/12)

universität koblenz
weiter:denken

- **Formulating the Model** Consider a group of people (numbered 1, 2, 3,...) who will sequentially make decisions; that is, individual 1 will decide first, then individual 2 will decide, and so on. We can describe the decision as a choice between accepting or rejecting some option
    - This decision could be about whether to adopt a new technology, wear a new fashion, eat in a new restaurant, commit a crime, vote for a particular political candidate, or choose one route to a common destination rather than an alternative route
- **First Model Ingredient:** States of the World. At the start of everything, before any individual has made a decision, we assume that the world is randomly placed into one of two possible states: it is placed in either a state in which the option is actually a good idea or a state in which the option is actually a bad idea
    - We imagine that the state of the world is determined by some initial random event that the individuals can't observe, but they will try to use what they observe to make inferences about this state
    - For example, the world is either in a state where the new restaurant is good or a state where it is bad; the individuals in the model know that it was randomly placed in one of these two states, and they're trying to figure out which

A Simple, General Cascade Model (3/12)

- We write the two possible states of the world as G, which represents the state where the option is a good idea, and B, which represents the state where the option is a bad idea
- We suppose that each individual knows the following fact: the initial random event that placed the world into state G or B placed it into state G with probability p, and into state B with probability 1 − p. This will serve as the prior probabilities of G and B; in other words, Pr [G] = p, and hence Pr [B] = 1 − Pr [G] = 1 − p
- **Second Model Ingredient:** Payoffs. Each individual receives a payoff based on her decision to accept or reject the option. If the individual chooses to reject the option, she receives a payoff of 0. The payoff for accepting depends on whether the option is a good idea or a bad idea
- Let's suppose that if the option is a good idea then the payoff obtained from accepting it is a positive number $v_g > 0$. If the option is a bad idea, then the payoff is a negative number $v_b < 0$. We will also assume that the expected payoff from accepting in the absence of other information is equal to 0; in other words, $v_g p + v_b(1 - p) = 0$
    - That is, before an individual gets any additional information, the expected payoff from accepting is the same as the payoff from rejecting

A Simple, General Cascade Model (4/12)

universität koblenz
weiter:denken

- **Third Model Ingredient:** Signals. In addition to the payoffs, we also want to model the effect of private information. We assume that, before any decisions are made, each individual gets a private signal that provides information about whether accepting is a good idea or a bad idea
- The private signal is designed to model private information that the person happens to know, beyond just the prior probability p that accepting the option is a good idea
- The private signal does not convey perfect certainty about what to do (since we want to model individual uncertainty even after the signal comes in), but it does convey useful information. Specifically, there are two possible signals: a high signal (denoted H), which suggests that accepting is a good idea, and a low signal (denoted L), which suggests that accepting is a bad idea
- We can make this precise by saying that if accepting is in fact a good idea then high signals are more frequent than low signals: $\Pr[H \mid G] = q > 1/2$, while $\Pr[L \mid G] = 1 - q < 1/2$. Similarly, if accepting the option is a bad idea, then low signals are more frequent: $\Pr[L \mid B] = q$ and $\Pr[H \mid B] = 1 - q$, for this same value of $q > 1/2$

A Simple, General Cascade Model (5/12)

- The probability of receiving a low or high signal, as a function of the two possible states of the world (G or B)

		States	
		B	G
Signals	L	q	$1-q$
	H	$1-q$	q

- The herding experiment from fits the properties of this more abstract model
  - The two possible states of the world are that the urn placed at the front of the room was majority-blue or that it was majority-red
  - We can think of “accepting” as guessing “majority-blue”; this is a good idea (G) if the true urn really is majority-blue and a bad idea (B) otherwise
  - The prior probability of accepting being a good idea is p = 1/2 . The private information in the experiment is the color of the ball the individual draws; it’s a “high” signal if it is blue, and so Pr [H | G] = Pr [blue | majority-blue] = q = 2/3

A Simple, General Cascade Model (6/12)

- Similarly, to return to the two-restaurant example from the opening section, “accepting” could correspond to choosing the first restaurant, A; it’s a good idea if restaurant A is actually better than the second restaurant, B
- The private information could be a review that you read of the first restaurant, with a high signal corresponding to a review comparing it favorably to restaurant B next door
- If choosing the first restaurant is actually good, there should be a higher number of such reviews, so $\Pr[H \mid G] = q > 1/2$

A Simple, General Cascade Model (7/12)

- **Individual Decisions.** We now want to model how people should make decisions about accepting or rejecting. First, let's consider how someone should do this based only on their own private signal, and then consider the effect of observing the earlier decisions of others
- Suppose that a person gets a high signal. This shifts their expected payoff from $v_g\Pr[G] + v_b\Pr[B] = 0$ to $v_g\Pr[G \mid H] + v_b\Pr[B \mid H]$

$$
\begin{aligned}
\Pr[G \mid H] &= \frac{\Pr[G] \times \Pr[H \mid G]}{\Pr[H]} \\
&= \frac{\Pr[G] \times \Pr[H \mid G]}{\Pr[G] \times \Pr[H \mid G] + \Pr[B] \cdot \Pr[H \mid B]} \\
&= \frac{pq}{pq + (1-p)(1-q)} \\
&> p,
\end{aligned}
$$

- where in the second line we compute the denominator Pr[H] as usual by expanding out the two possible ways of getting a high signal (if the option is a good idea or a bad idea). The final inequality follows since we have $pq + (1 - p)(1 - q) < pq + (1 - p)q = q$ in the denominator

A Simple, General Cascade Model (8/12)

This result makes sense

- A high signal is more likely to occur if the option is good than if it is bad, so if an individual observes a high signal they raise their estimate of the probability that the option is good
- As a result, the expected payoff shifts from zero to a positive number, and so they should accept the option
- A completely analogous calculation shows that if the individual receives a low signal, they should reject the option

A Simple, General Cascade Model (9/12)

universität koblenz
weiter:denken

- **Multiple Signals.** We know from the herding experiment that an important step in reasoning about how people make decisions in sequence is to understand how an individual should use the evidence of multiple signals
- Using Bayes' rule, it's not hard to reason directly about an individual's decision when they get a sequence S of independently generated signals consisting of *a* high signals and *b* low signals, interleaved in some fashion. We do this by deriving the following facts

  (i) The posterior probability Pr [G | S] is greater than the prior probability Pr [G] when a>b

  (ii) The posterior probability Pr [G | S] is less than the prior probability Pr [G] when a < b

  (iii) The two probabilities Pr [G | S] and Pr [G] are equal when a = b
- As a result, individuals should accept the option when they get more high signals than low signals, and reject it when they get more low signals than high signals
- They are indifferent when they get the same number of each
  - In this simple setting with a sequence of signals, individuals can decide according to a majority vote over the signals they receive

A Simple, General Cascade Model (10/12)

- Here we will try to justify facts (i)–(iii), using Bayes' rule and a bit of algebra. To apply Bayes' rule, we write

$$\Pr[G \mid S] = \frac{\Pr[G] \times \Pr[S \mid G]}{\Pr[S]},$$

where S is a sequence with a high signals and b low signals

- To compute Pr [S | G] in the numerator, we note that because the signals are generated independently, we can simply multiply their probabilities, which gives us a factors of q and b factors of (1 − q), and so Pr [S | G] = $q^a(1-q)^b$
- To compute Pr [S], we consider that S can arise if the option is a good idea or a bad idea, so

$$\Pr[S] = \Pr[G] \times \Pr[S \mid G] + \Pr[B] \times \Pr[S \mid B]$$
$$= pq^a(1-q)^b + (1-p)(1-q)^a q^b.$$

- Plugging this back into Equation , we get

A Simple, General Cascade Model (11/12)

$$\Pr[G \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)(1-q)^a q^b}.$$

- What we want to know is how this expression compares to p
- If we were to replace the second term in the denominator by $(1 - p)q^a(1 - q)^b$, then the denominator would become $pq^a(1 - q)^b + (1 - p)q^a(1 - q)^b = q^a(1 - q)^b$, and so the whole expression would become

$$\frac{pq^a(1-q)^b}{q^a(1-q)^b} = p$$

A Simple, General Cascade Model (12/12)

- So the question is: *does this replacement make the denominator smaller or larger?* $\frac{pq^a(1-q)^b}{q^a(1-q)^b} = p$

(i) If a>b, then this replacement makes the denominator larger, since q > 1/2 and we now have more factors of q and fewer factors of 1 − q. Since the denominator gets larger, the overall expression gets smaller as it is converted to a value of p, and therefore Pr [G | S] > p = Pr [G]

(ii) If a < b , the argument is symmetric: this replacement makes the denominator smaller, and hence the overall expression larger

- So Pr [G | S] < p = Pr [G]

(iii) Finally, if a = b, then this replacement keeps the value of the denominator the same,

- So Pr [G | S] = p = Pr [G]

6. Sequential Decision Making and Cascades

Sequential Decision Making and Cascades (1/7)

universität koblenz
weiter:denken

Let's now consider what happens when individuals make decisions in sequence

- As before, we want to capture situations in which each person can see what earlier people do, but not what they know
- In our model, this means that when a given person decides whether to accept or reject the option, they have access to their own private signal and also the accept/reject decisions of all earlier people
  - Important: they do not see the actual private signals of any of these earlier people
- The reasoning is now very similar to what we did for the sequence of students in the herding experiment
  - Person 1 will follow his own private signal
  - Person 2 will know that person 1's decision reveals their private signal, and so it is as though person 2 gets two signals. If these signals are the same, person 2's decision is easy. If they are different, person 2 will be indifferent between accepting and rejecting. Here we assume she follows her own private signal. Thus, either way, person 2 is following her own signal

Sequential Decision Making and Cascades (2/7)

universität koblenz
weiter:denken

- As a result, person 3 knows that person 1 and person 2 both acted on their private signals, so it is as though person 3 has received three independent signals (the two he infers, and his own observation). We know from the argument previously, that person 3 will follow the majority signal (high or low) in choosing whether to accept or reject
- This means that if person 1 and person 2 have made opposite decisions (i.e., they received opposite signals), then person 3 will use his own signal as the tiebreaker. Hence, future people will know that person 3's decision was based on his own signal, and so they can use this information in their own decisions
- On the other hand, if person 1 and person 2 have made the same decision (i.e., received the same signal), then person 3 will follow this regardless of what his own signal says
- Hence, future people will know that person 3's decision conveys no information about his signal, and these future people will all be in the same position as person 3
  - In this case, a cascade has begun. That is, we are in a situation where no individual's decision can be influenced by his own signal
  - No matter what they see, every individual from person 3 on will make the same decision that persons 1 and 2 made

Sequential Decision Making and Cascades (3/7)

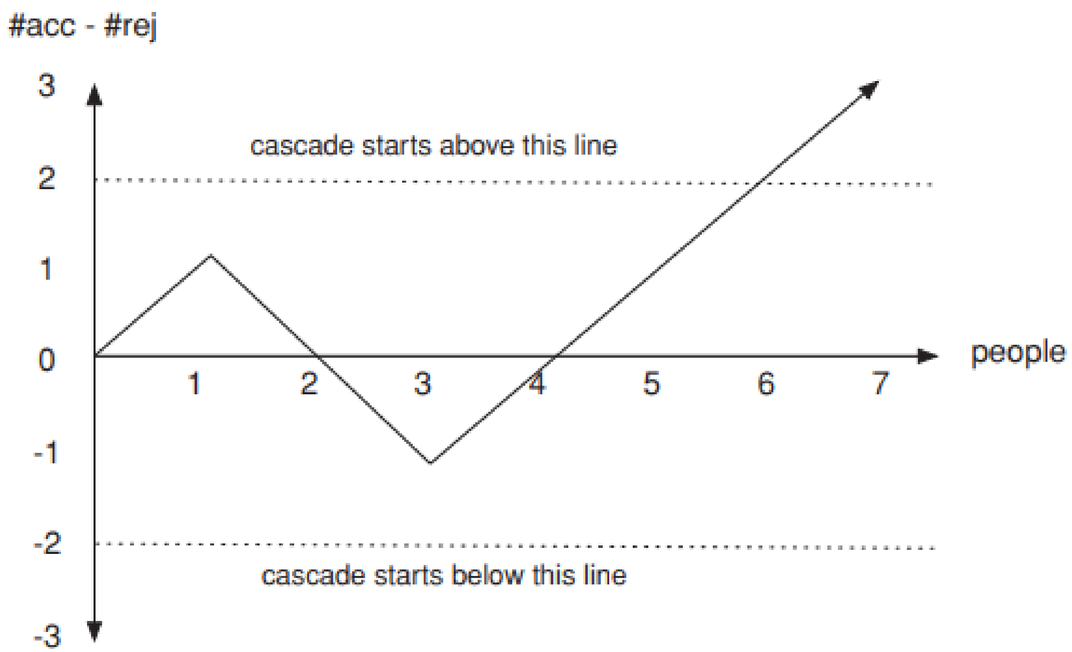
- Let's now consider how this process unfolds through future people beyond person 3
- In particular, let's consider the perspective of a person numbered N
- Suppose that person N knows that everyone before her has followed their own signal; that is, suppose the accept/reject decisions of these earlier people exactly coincide with whether they received a high or low signal, and person N knows this. There are several possible cases to consider
  - If the number of acceptances among the people before N is equal to the number of rejections, then N's signal will be the tie-breaker, and so N will follow her own signal
  - If the number of acceptances among the people before N differs from the number of rejections by one, then either N's private signal will make her indifferent or it will reinforce the majority signal. Either way, N will follow her private signal (since we assume a person follows their own signal in the case of indifference)
  - If the number of acceptances among the people before N differs from the number of rejections by two or more, then however N's private signal turns out, it won't outweigh this earlier majority. As a result, N will follow the earlier majority and ignore her own signal

Sequential Decision Making and Cascades (4/7)

- Moreover, in this case, the people numbered N + 1, N + 2, and onward will know that person N ignored her own signal (whereas we've assumed that all earlier people were known to have followed their private signals)
- So they will each be in exactly the same position as N. This means that each of them will also ignore their own signals and follow the majority; hence, a cascade has begun
- We can therefore sum up the behavior of the decision-making process as follows
  - As long as the number of acceptances differs from the number of rejections by at most one, each person in sequence is simply following their own private signal in deciding what to do
  - But once the number of acceptances differs from the number of rejections by two or more, a cascade takes over, and everyone simply follows the majority decision forever

Sequential Decision Making and Cascades (5/7)

- A cascade begins when the difference between the number of acceptances and rejections reaches two
- The figure illustrates this for a sample outcome of the process, in which we plot the difference between the number of acceptances and the number of rejections over time as people make decisions
- This plot moves up or down by one each time a new decision is made, since either the number of acceptances or the number of rejections grows by exactly one with each decision
- Once the difference between the number of acceptances and the number of rejections escapes from the narrow horizontal ribbon around zero – that is, once the plot moves at least two steps away from the x-axis – a cascade begins and runs forever



Sequential Decision Making and Cascades (6/7)

universität koblenz
weiter:denken

- Finally, it is very difficult for this difference to remain in such a narrow interval (between −1 and +1) forever
  - For example, during the period of time when people are following their own signals, if three people in a row ever happen to get the same signal, a cascade will definitely have begun
- Now, let's argue that the probability of finding three matching signals in a row converges to 1 as the number of people, N, goes to infinity
  - Suppose we divide the first N people into blocks of three consecutive people each (people 1, 2, 3; people 4, 5, 6; people 7, 8, 9; and so on)
  - Then the people in any one block will receive identical signals with probability $q^3 + (1 - q)^3$
  - The probability that none of these blocks consists of identical signals is therefore

$(1 - q^3 - (1 - q)^3)^{N/3}$

  - As N goes to infinity, this quantity goes to zero, which means as the number of people goes to infinity, the probability that a cascade begins converges to 1. Thus, in the limit, a cascade takes place in this model almost surely

Sequential Decision Making and Cascades (7/7)

- This is an extremely simple model of individual decision-making
- In more general versions, for example, it could well be the case that people don't see all the decisions made earlier but only some of them; that not all private signals convey equal information, or that not everyone receives the same payoffs
- Many of these more general variants become much more complicated to analyze, and they can differ in their specifics (for example, the condition for a cascade to begin is clearly not always as simple as having the number of acceptances differ from the number of rejections by at least two)
- *When people can see what others do but not what they know, there is an initial period when people rely on their own private information; but as time goes on, the population can tip into a situation where people – still behaving fully rationally – begin ignoring their own information and following the crowd*

7. Summary

Summary (1/6)

(i) Cascades can be wrong. If, for example, accepting the option is in fact a bad idea but the first two people happen to get high signals, a cascade of acceptances will start immediately, even though it is the wrong choice for the population

(ii) Cascades can be based on very little information. Because people ignore their private information once a cascade starts, only the pre-cascade information influences the behavior of the population. This means that if a cascade starts relatively quickly in a large population, most of the private information that is collectively available to the population (in the form of private signals to individuals) is not being used

(iii) Cascades are fragile. The previous point makes them easy to start, but it can also make them easy to stop. One manifestation of this fact is that people who receive slightly superior information can overturn even long-lived cascades

Summary (2/6)

- For example, suppose that a cascade of acceptances is under way in our model; so we know that the number of high signals exceeded the number of low signals by two at the time the cascade began
- Now suppose someone making a decision in the midst of this cascade happens to receive two private signals
  - If they are both low signals, then this person (taking into account the earlier signals he can infer) has now seen an equal number of high and low signals
  - Because he is indifferent, our assumption is that he will reject (since his own signals were low) despite the long run of acceptances that preceded him
- A single public signal can have the same effect
  - if, in the midst of a cascade, there is a public signal that everyone sees, then the next person to decide in effect receives two signals (the public one and her own private one), with similar consequences

Summary (3/6)

- The main lesson to be learned from studying cascades is
  - to be careful in drawing conclusions about the best course of action from the behavior of a crowd
  - the crowd can be wrong even if everyone is rational and everyone takes the same action
- This forms an interesting contrast with an argument made by popular general audience books such as James Surowiecki's The Wisdom of Crowds – that the aggregate behavior of many people with limited information can sometimes produce very accurate results
  - In his opening example, Surowiecki notes that, if many people are guessing independently, then the average of their guesses is often a surprisingly good estimate of whatever they are guessing about (perhaps the number of jelly beans in a jar, or the weight of a bull at a fair)
  - The key to this argument of course is that the individuals each have private information (their signals), and they guess independently, without knowing what the others have guessed. If instead they guess sequentially, and can observe the earlier guesses of others, then we are back in the cascade setting and there would be no reason to expect the average guess to be good at all

Summary (4/6)

- These observations suggest how the possibility of cascades can affect the behavior of individuals or groups in a range of different situations
    - One setting that is susceptible to cascades is the style of group decision making in which a committee of people sit around a table and discuss potential solutions to a problem
- For example, consider a hiring committee that needs to decide whether to make a job offer to candidate A or candidate B. In these kinds of situations, a common strategy is to go around the table, asking people in sequence to express their support for option A or option B
- But if the participants assume that they all have roughly equal insight into the problem, then a cascade can quickly develop: if a few people initially favor A, others may be led to conclude that they should favor A, even if they initially preferred B on their own
- The cascade principles we've seen in this lecture suggest that this may not just be an issue of social pressure to conform to the majority, but in fact a rational approach to decision making, in which you assume that the people speaking before you have information about the problem that is comparable in quality to your own

Summary (5/6)

universität koblenz
weiter:denken

- Such considerations suggest an inherent tension between getting a group of experts to work together and build on each other’s ideas, on the one hand, and giving them the opportunity to form their own opinions, on the other
- This in turn suggests strategies for balancing this tension, potentially by forcing experts to reach partial decisions independently before entering a phase of collaboration and consensus
- It also suggests that if certain people are known to have particularly good information about a problem, it can matter whether they weigh in earlier in the process or later

Summary (6/6)

- Marketers also use the idea of cascades to attempt to get a buying cascade started for a new product. If they can induce an initial set of people to adopt the new product, then those who make purchasing decisions later on may also adopt the product even it is no better than, or perhaps even worse than, competing products
- This is most effective if these later consumers are able to observe the adoption decisions, but not how satisfied the early customers actually were with their choices; this is consistent with the idea that cascades arise naturally when *people can see what others do but not what they know*
- If the payoffs (or statistics based on the payoffs) from earlier consumers are visible, this can help prevent a cascade of bad choices – another example of how changing the information available to a group of people can have an effect on their overall behavior

References

[1] Easley & Kleinberg, J. (2010). *Networks, Crowds and Markets*. Cambridge: Cambridge University Press.

- Chapter 16: Information Cascades

[2] OLAT course page: https://olat.vcrp.de/url/RepositoryEntry/4669112833