



15-382 COLLECTIVE INTELLIGENCE – S19

LECTURE 33:

WISDOM OF THE CROWD

TEACHER:

GIANNI A. DI CARO

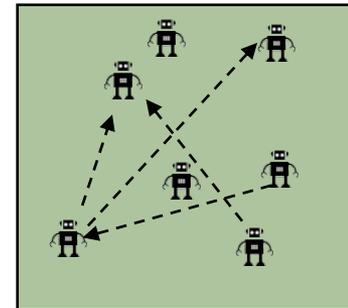
جامعة كارنيجي ميلون في قطر
Carnegie Mellon University Qatar

SO FAR...

- In Game Theory we have considered multi-agent systems with potentially *conflictual utilities*. Solution concept: equilibrium



- In PSO and ACO agents do *cooperate online* by continual information sharing (agent-to-agent in PSO, mediated by the environment in ACO)

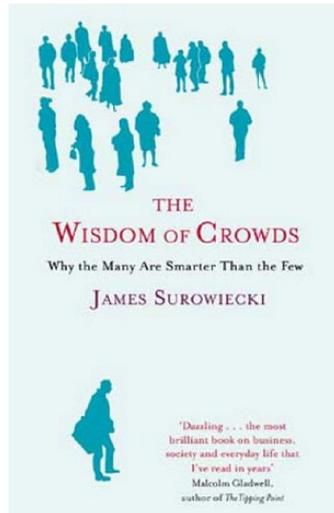


- In Auctioning and Task Allocation, agents can *compete or cooperate*, depending on the context



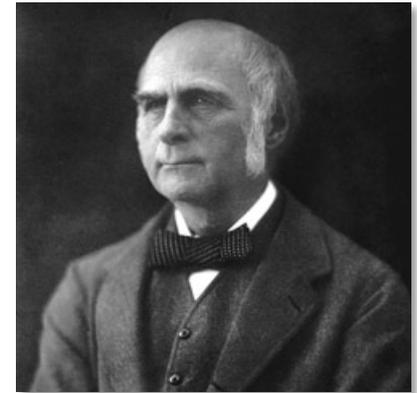
WISDOM OF THE CROWD

- Let's use a multi-agent system, a *crowd* to solve problems, like making estimates of values, taking decisions, ...
- The basic idea is that the **collective opinion of a group of individuals *can* be better than a single expert opinion**:
 - If the individuals in the crowd are experts (or, most of them are), the intuition is quite obvious
 - What if the majority is far from being an expert for the problem domain?
 - A few conditions need to be in place, to make a crowd “wise”, otherwise it may fail miserably (e.g., a number of examples from Reddit)

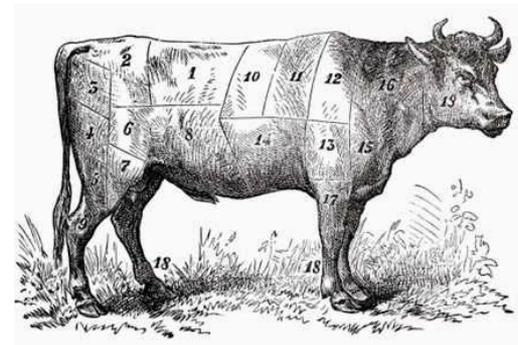


FRANCIS GALTON AND THE OX WEIGHT

- Francis Galton (16 February 1822 – 17 January 1911), cousin of Charles Darwin, was an English Victorian polymath, proto-geneticist, statistician...



- In 1906 Galton visited a livestock fair and stumbled upon an contest. An ox was on display, and the villagers were invited to guess the animal's weight after it was slaughtered and dressed.

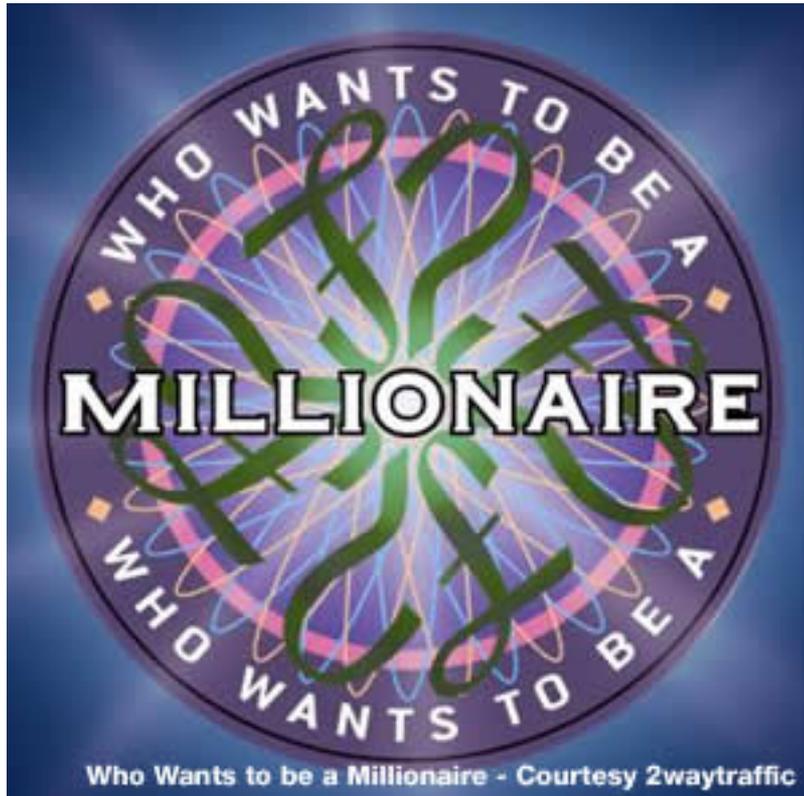


- Galton disliked the idea of democracy and wanted to use the competition to show the problems of allowing large groups of people to vote on a topic.

POWER OF AGGREGATING INFORMATION

- 787 people guessed the weight of the ox, some were experts, farmers and butchers, others knew little about livestock.
- Some guessed very high, others very low, many guessed fairly sensibly.
- Galton collected the guesses after the competition was over
- The **average** guess from the crowd was **1,197 pounds**
- The **correct weight** was **1,198 pounds!**
- What Dalton discovered was that in actuality crowds of people can make surprisingly good decisions IN THE AGGREGATE, even if they have imperfect information
- Many other examples can be found / mentioned ...

WHO WANTS TO BE A MILLIONAIRE?



- Compare the lifelines:
 - Phone a friend
 - Ask the Audience
- The correct answer is given:
 - Phone a friend → 65%
 - Ask the Audience → 91%

THE SPACESHUTTLE CHALLENGER



- On January 28, 1986, when the Space Shuttle Challenger broke apart 73 seconds into its flight, leading to the deaths of its seven crew members. The spacecraft disintegrated over the Atlantic Ocean, off the coast of central Florida
- The stock market did not pause to mourn. Within minutes, investors started dumping the stocks of the four major contractors who had participated in the Challenger launch:
 - Rockwell International, which built the shuttle and its main engines;
 - Lockheed, which managed ground support;
 - Martin Marietta, which manufactured the ship's external fuel tank; and
 - Morton Thiokol, which built the solid-fuel booster rocket.
- By the end of the day, Morton Thiokol's stock was down nearly 12 percent. By contrast, the stocks of the three other firms started to creep back up, and by the end of the day their value had fallen only around 3 percent.

THE SPACESHUTTLE CHALLENGER

- What this means is that the stock market had, almost immediately, labelled Morton Thiokol as the company that was responsible for the *Challenger* disaster.
- Months later it was discovered that it was in fact Morton Thiokol who caused the problem with the production of faulty O-rings.
- How did the stock investors know ?
- A good “explanation” is that, again, this is effect of the the wisdom of crowds.

GOOGLE PAGE RANKING

- How does Google work? (in a "simple" way)
- How does it classify pages so that typically the page you are looking for is in the first ten links it returns?
- It uses the PageRank algorithm, whose main idea is:
 - The more sites that link to a certain URL with a certain phrase, the higher the rating.
 - This works because each link is a *vote* for the connection between the phrase and the site.
- Again, this can be seen as a form of the wisdom of the crowds

EXPERTS VS. WISDOM OF CROWDS

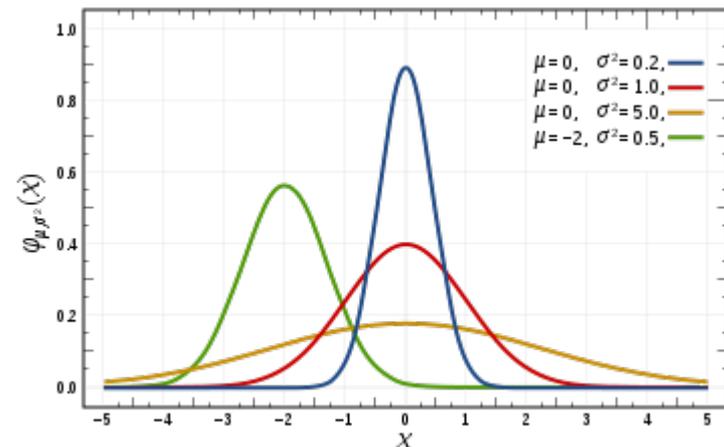
- It shows us that groups of people make excellent decisions and can select the correct alternative out of a number of options without any specific expertise (“maybe”)
- How could this be?
- One general observation is that *individual experts really aren't as smart as we think*, such that it might be difficult to find the “right” expert when the decision is fairly complex and involves multiple levels of knowledge and abilities
- An interesting experiments in this respect was done by Herbert Simon and W.G. Chase (1973), who explored the nature of expertise in the domain of chess.

EXPERTS ARE NOT KNOW-IT-ALLS

- They showed a chess-board in the middle of a game to an expert chess player and an amateur.
- They asked both to recreate the locations of all of the pieces on another boards, consistency the experts were easily able to reproduce the boards, whereas the amateur rarely could.
- So does this mean experts are smarter?
- No, because when they put the pieces on the board randomly, the expert and amateur both did equally as well.
- This shows how limited might be the scope of expertise.
- We normally assume people who are intelligent at one pursuit are good at all, but in actuality this is not at all the case.
- Chase said the intelligence and expertise is, in fact, “*spectacularly narrow*”

CROWD OF EXPERTS / NON-EXPERTS

- If a group of multiple experts for the domain is available, it is expected they collectively provide a better answer than they would do individually
- Value sampling from expert population (the *crowd*), each expert i outputs an estimate s_i , that can be seen as a random variable, and their sample mean c has the same expected value of the population
- If the population is of true experts, the estimates s_i will have (in the limit of large populations) a Gaussian distribution centred at the true value \bar{s} , and small variance
- The less expert the crowd is, the larger the variance
- If the crowd has no expertise at all, there's the risk that estimates will have a wrong bias (e.g., the green distribution)



DIVERSITY PREDICTION THEOREM

- How the crowd issues correct estimates / makes good decisions?
- It's a simple theorem (actually an identity)
- **Diversity Prediction Theorem:**

$$(c - \theta)^2 = \frac{1}{n} \sum_{i=1}^n (s_i - \theta)^2 - \frac{1}{n} \sum_{i=1}^n (s_i - c)^2$$

c is the crowd estimate, the sample mean of individual estimates s_i

θ is the ground truth

n is the number of individuals in the crowd

Crowd's (quadratic) error = Average (quadratic) error – Crowd diversity

- *Diversity*: spread of estimates / expertise in the crowd

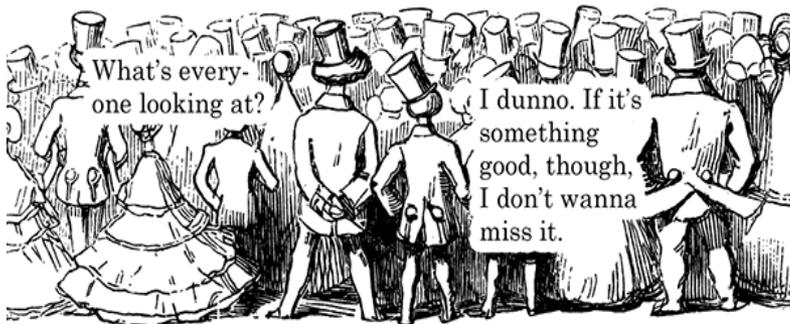


DIVERSITY PREDICTION THEOREM

- $[\text{Crowd's error}] = [\text{Average error}] - [\text{Diversity}]$
- How do we get a small Crowd's error?
 - A crowd of experts: $[\text{Average error}]$ is small, $[\text{Diversity}]$ will also be small, usually
 - A crowd of non experts: $[\text{Average error}]$ will be fairly large, but if we have a balanced large $[\text{Diversity}]$, we get a small error, we also need relatively large crowds to make the probabilities work

DIVERSITY PREDICTION THEOREM

- $[\text{Crowd's error}] = [\text{Average error}] - [\text{Diversity}]$
- When things can go wrong?
 - The non experts are **badly wrong** and have a (wrong) **bias** in their estimates, such that $[\text{Diversity}]$ can't counterbalance the $[\text{Average error}]$
 - When the estimates are **not independent**, such that, for instance, a *wrong bias* can be established because of social interactions, driving the crowd to the wrong answer



Jan Lorenz, Heiko Rauhut, Frank Schweitzer, Dirk Helbing, *How social influence can undermine the wisdom of crowd effect*. Proceedings of the National Academy of Sciences (PNAS), 108 (22) 9020-9025, May 2011

AGENT INTERACTIONS: INTERCONNECTION NETWORK

- *Social interactions* → Network, information sharing that propagates through a set of *interconnection channels*
- Interconnection Networks strongly affect how in a complex system information propagates, that in turn determines how individuals evolve over time
- How is an interconnection network represented **mathematically**?
- What **properties** do networks have? How are they measured?
- How do we **model** networks to understand their properties? How are real networks different from the ones produced by a simple model?
- What are useful networks for the task at hand?

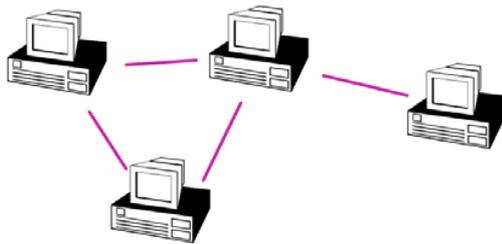
COMPLEX SYSTEMS AS NETWORKS

Many complex systems can be represented as networks

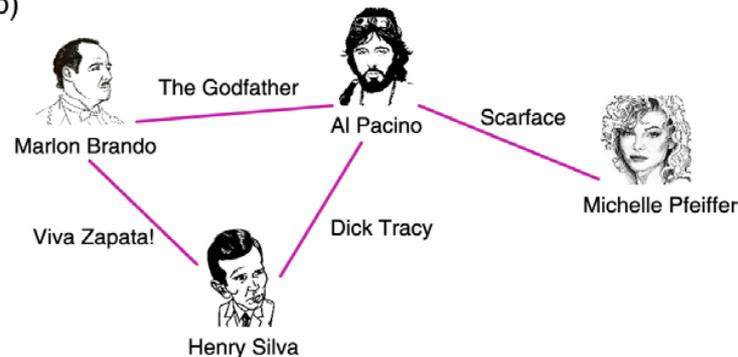
Any complex system has an associated network of communication / interaction among the components

- Nodes = components of the complex system
- Links = interactions between them

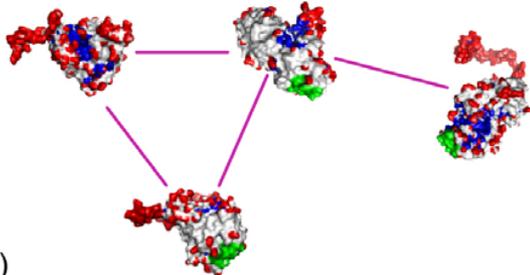
a)



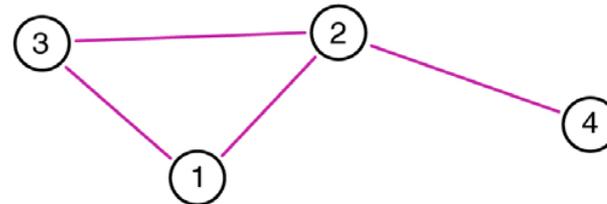
b)



c)



d)



RECOMMENDED READINGS

- Barabasi, “Network Science”
- Easley & Kleinberg, “Networks, Crowds, and Markets: Reasoning about a Highly Connected World”
- Newman, “Networks”

