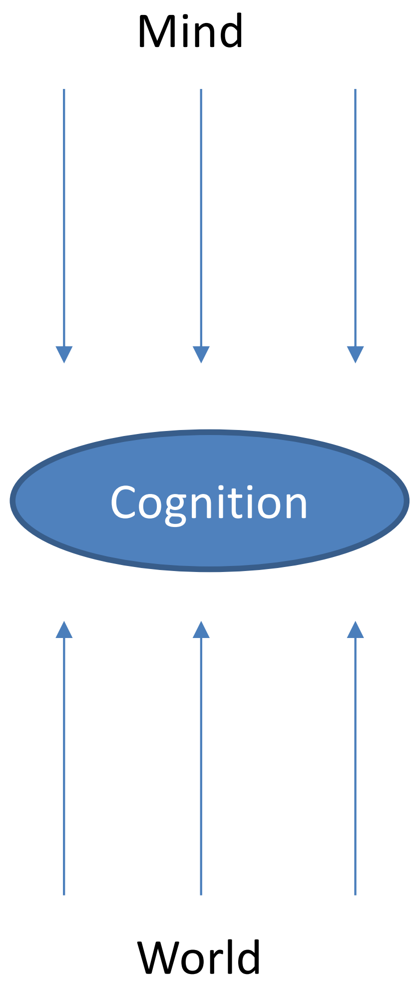
Association

Nisheeth

4th January 2018

Welcome to CS786 - a computational cognitive science course!



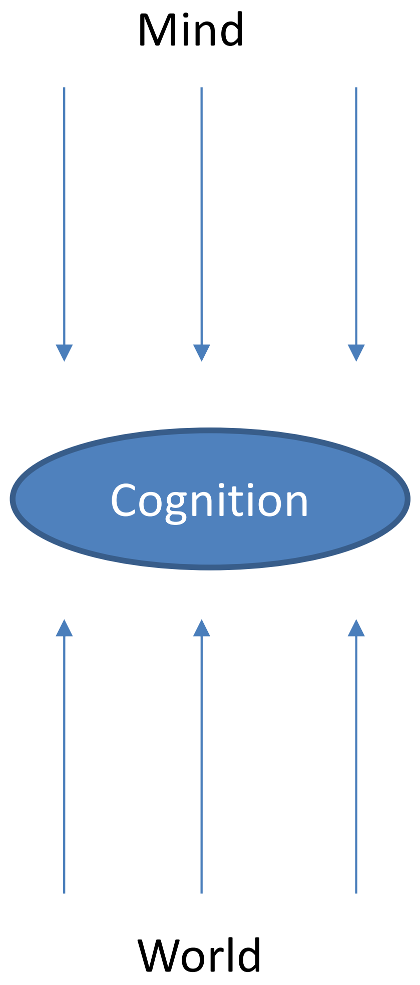
Cognition is the process by which the observer assembles what is observed into knowledge based on what the observer already knows

Does cognition have a cold start problem?

Developmental psychologists have found that even newborns come with a large bag of phenotypic and genetic experience

Cognition and computation

- Cognition is fundamentally path-dependent
- Purely analytic approaches fare poorly with path-dependence
- Computation maps on to cognitive path-dependence well



认知是基于路径的

About me

- I’m Nisheeth
- I sit in KD303
- Office hours for this course will be Friday 1500-1700?
  - Informal office hours right after the class everyday
- Email : nsrivast at cse.iitk.ac.in
- Phone: 7916

Course details

- CS786
  - TThF 0800-0850
  - KD101
- 7 quizzes (best 6 of 7 contribute 10% to course grade each)
  - May include take home programming components
  - Programming prep needed
    - ESC101
    - Courage
- No midsem or endsem
- 20% grade for class participation
- 20% grade for experiment participation

Course structure

- Sequence of seven modules
- Each module will last two weeks = 6 lectures
- Split into 5 lectures + 1 quiz-cum-discussion hour
  - Quiz will sometimes have a take-home programming component
  - Can ask TAs for help with programming problems
- Reading material will be assigned as web-links within each lecture
  - If you don’t read, you won’t be able to contribute in the discussion hour, which will draw upon these readings
  - Students are also encouraged to suggest their own readings; I will add them to the list if they seem relevant
  - Don’t have to read all the material assigned, but the more you do, the more fun you will have in the discussion hour

Course content

- Organizational principles
  - Foundations
    - Association
    - Reinforcement
    - Hierarchy
    - Perception-action-control loop
- Knowledge representation
  - Similarity and categorization
  - Perception
  - Memory
- Knowledge processing
  - Decision-making
- From knowledge to behavior
  - Motor control
  - Language

知识表征

Course policies

- Attendance is voluntary
  - But the class sessions will be the most important element of the course
  - You won't be able to keep up with the course just by following the slides
- Add-drop deadline is 12th Jan
  - Drops beyond that will require instructor and DUGC permission
  - My permission can be taken for granted
- Assuming good faith on your part (regular attendance and participation), the lowest possible grade you will get is C

Course philosophy

- This is a science course, not an engineering course
  - Emphasis is on following the chain of understanding where it leads, not developing technical competence
  - We will cover a lot of topics, many unrelated to each other
- Quizzes will be very easy
  - If you have come to class and read the reading material, you will have no trouble
- Collaboration in programming assignments is acceptable (with acknowledgement)
- There will be a lot of math
  - But only math to read, not math to do
  - Don’t let it scare you

This module - foundations

- Association (today)
- Reinforcement (tomorrow)
- Classical cognitive architectures (Tuesday)
- Modern cognitive architecture (Thursday)
- Quiz + discussion (Friday)

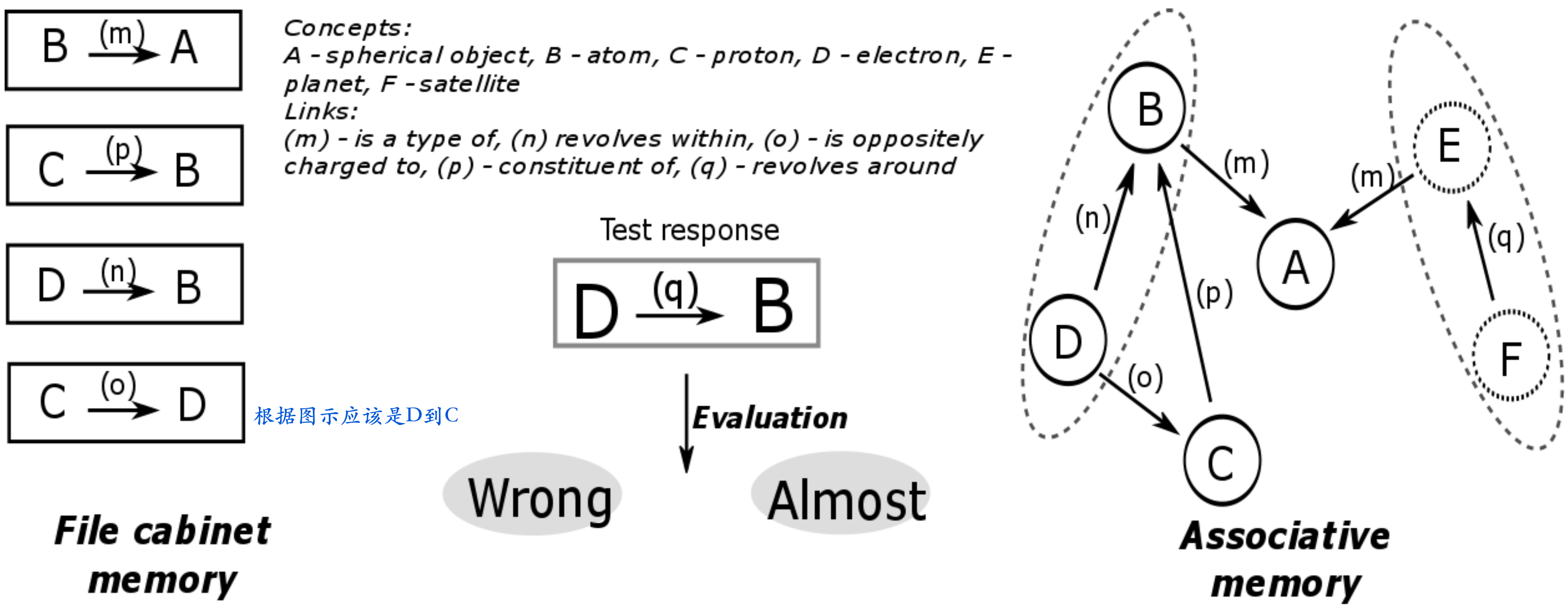
Association

- A computer stores information randomly. A mind contains concepts associatively

Knowing concepts associatively

同时发生的、并存的

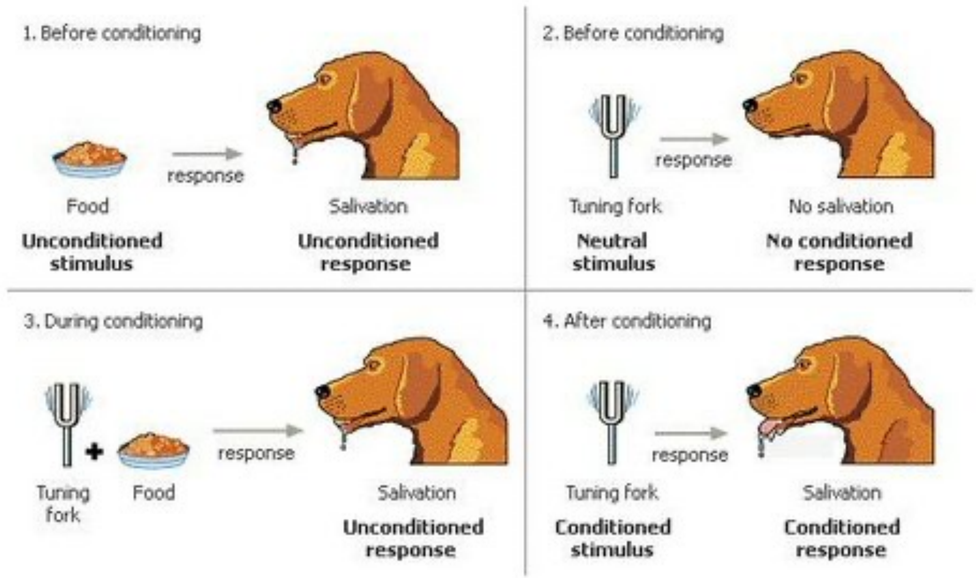
- *Related concepts* are activated concurrently



根据图示应该是D到C

What is relatedness?

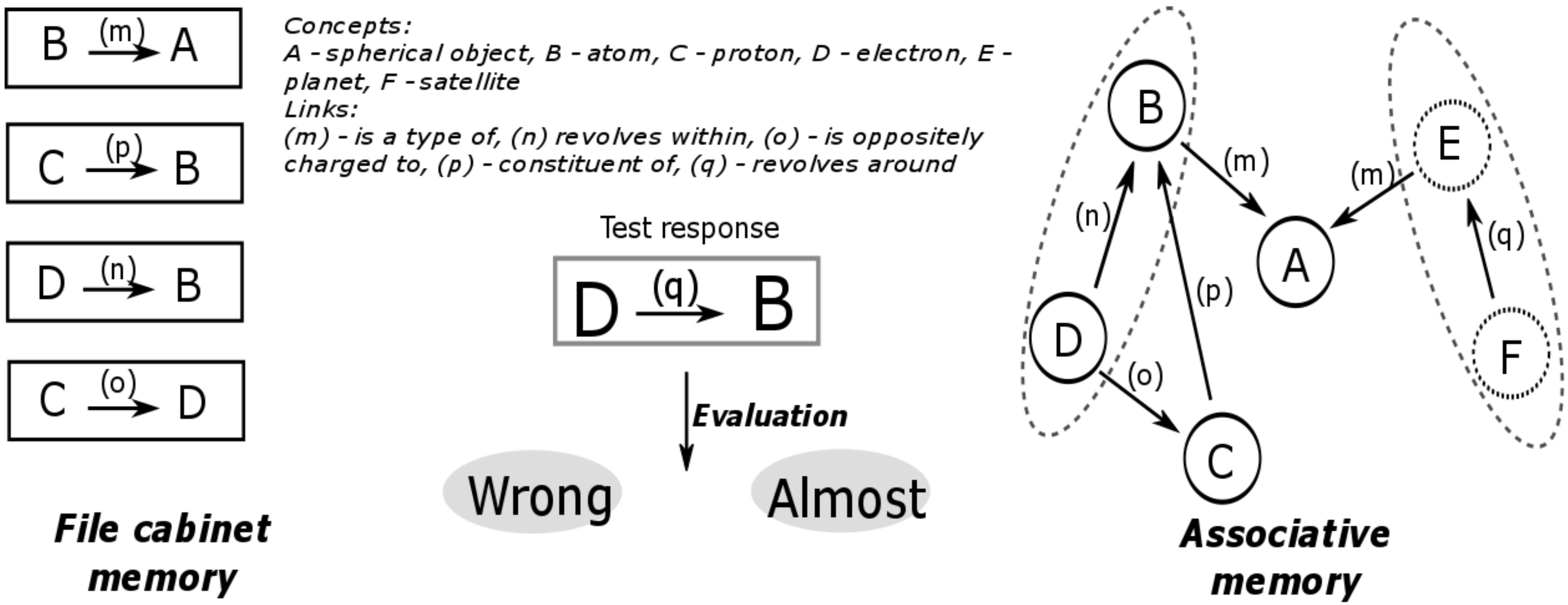
- Co-occurrence → Among the first behavior invariants discovered
- Functionally unrelated concepts become related when they are presented together
- Pavlov's dogs learned to *associate* sound with food.



Classical Conditioning (Source: schoolworkhelper.com)

Associativity is at the heart of human learning

- Spin a story that would make this student's associative error plausible



Studied systematically in conditioning experiments

- Not all as interesting as Watson's Little Albert experiment (pictured below)

Appearance of stimulus is closely followed by a particular behavior


CS: conditioned stimulus

US: unconditioned stimulus

Real-Life Examples of Classical Conditioning

Gustavson and Gustavson (1985) – Conditioned Taste Aversion

Coyotes killing sheep – problem to sheep farmers

Study conditioned coyotes not to eat the sheep

Sheep meat (CS) sprinkled with a chemical (UCS) that would produce a stomachache (UCR)

After coyotes ate the treated meat,
they avoided the live sheep (CR)

This humane application of conditioned taste aversion might be used to control other predators as well

Real-Life Examples of Classical Conditioning

Metalmikov & Chorine (1926, 1928) – Immune System

Injected Guinea Pigs with Foreign agents (non lethal) → antibodies → boost their immune system

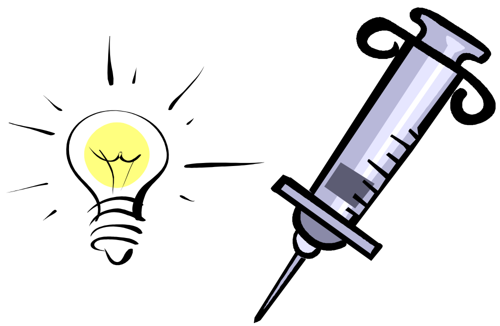
Then paired injections with Lights

Lights + Injections = better immunity

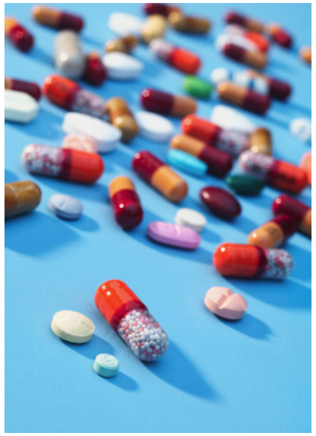
Lights alone = better immunity

霍乱

Later Injected Cholera: animals with prior conditioning better survival vs controls with no conditioning



In *A Clockwork Orange*, a mass murderer, is strapped to a chair and forced to watch violent movies while he is injected with a drug that nauseates him. So he sits and gags and retches as he watches the movies. After hundreds of repetitions of this, he associates violence with nausea, and it limits his ability to be violent.



Drug Overdose

- drug users become increasingly less responsive to the effects of the drug
- tolerance is specific to specific environments (e.g. bedroom)
- familiar environment becomes associated with a compensatory response (Physiology)
- taking drug in unfamiliar environment leads to lack of tolerance → drug overdose


ADVERTISING
MARKETING
PRODUCT
TELEVISION
COMMERCIAL
PUBLIC
ADVERTISEMENTS
RADIO
MEDIA
ADS
PRODUCTS
ADVERTISEMENT
BRAND
SERVICE
ADVERTISERS
EDIT
USED
ALSO
MANY
MESSAGE
HOWEVER
NETWORK
UNITED
MARKET
BILLBOARDS
MOBILE
AD
CONSUMER
OFTEN
STATES
ONLINE
INCLUDE
DIGITAL
TIME
CONTENT
COMMUNICATION
FIRST
MESSAGES
FORM
SOCIAL
INTERNET
INFORMATION
RESEARCH
EXAMPLE
BILLBOARD
MEDIUM
MAIN
CONSUMERS
CAMPAIGNS
AUDIENCE
ADVERTISE
GLOBAL
AGENCY
NEWSPAPER
BECOME
SERVICES
USE
PLACEMENT
VIDEO
CREATE
WEB
MASS
SHOWS
COUNTRIES
BROADCASTING
KNOWN
TV
MAY
NEW
SIGNS
DISPLAY
ARTICLE
SPONSOR
BUSINESS
COMPANIES
INCLUDING
LARGE
FOREIGN
SALES
PAID
VIRTUAL
ONE
LIKE
BOWL
SHOW
WORLD
STREET
COMMERCIALS
INSERTED
LOCAL
PROGRAMS
MAJOR
FILM
ANOTHER
CARS
LOGO
SPAM
E-MAIL
SEE
FOUR
MIX
WOMEN
TARGET
EVENTS
MODEL
SHOPPING
INCREASING
SUPER
VARIOUS
AGENCIES
IDENTIFIED
SPACE
MUSIC
EARLY
TYPICALLY
CHARACTER
CAMPAIGN
ORGANIZATIONS
REGULATION
BECAME
PROMOTE
INFOMERCIAL
MEANING
ABILITY
CUSTOMERS
INDUSTRY
NEWSPAPERS
COMPANY
PROCESS
SEMIOTICS
RISE
RECENT

Clinical therapies

- People keep trying to use conditioning-based methods, e.g. ‘flooding’ to treat phobias, fears and trauma-related disorders
- Doesn’t work very well – fear conditioning is much stronger than fear extinction
- For reasons that may become clearer as we go along

Can you think why? response more resistant to extinction

Modeling classical conditioning

- Most popular approach for years was the Rescorla-Wagner model

$$\Delta V_X^{n+1} = \alpha_X \beta(\lambda - V_{tot})$$

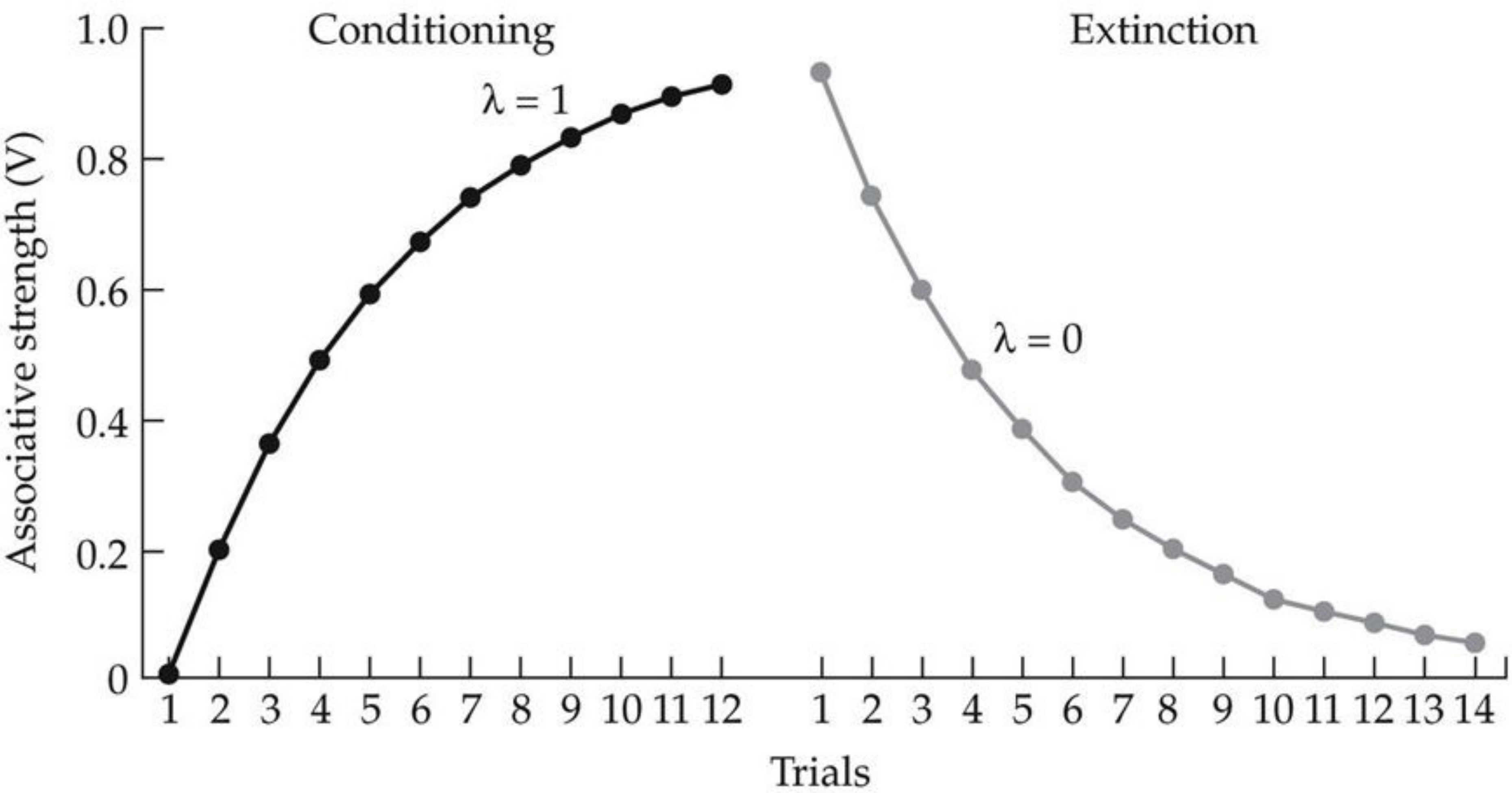
$$V_X^{n+1} = V_X^n + \Delta V_X^{n+1}$$

Some versions replace V_{tot} with V_x; what is the difference?

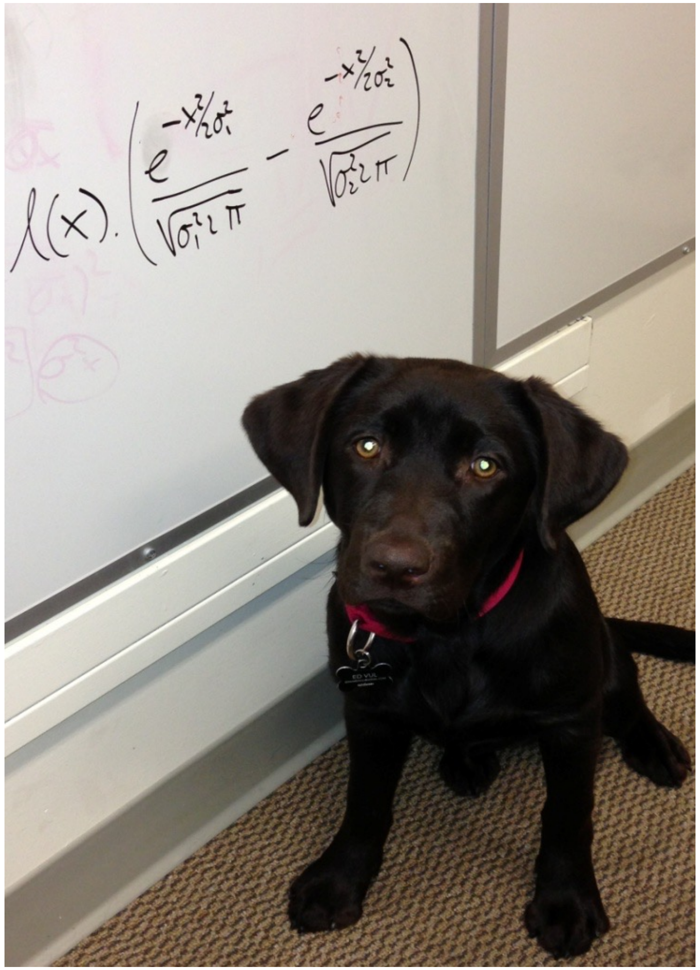
- Could reproduce a number of empirical observations in classical conditioning experiments

http://users.ipfw.edu/abbott/314/Rescorla2.htm

4.3 Conditioning and extinction in the Rescorla-Wagner model

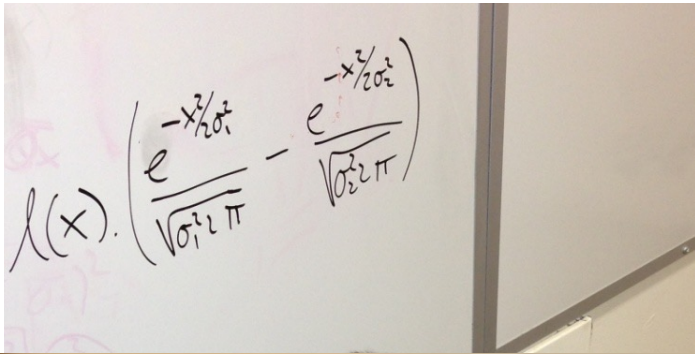


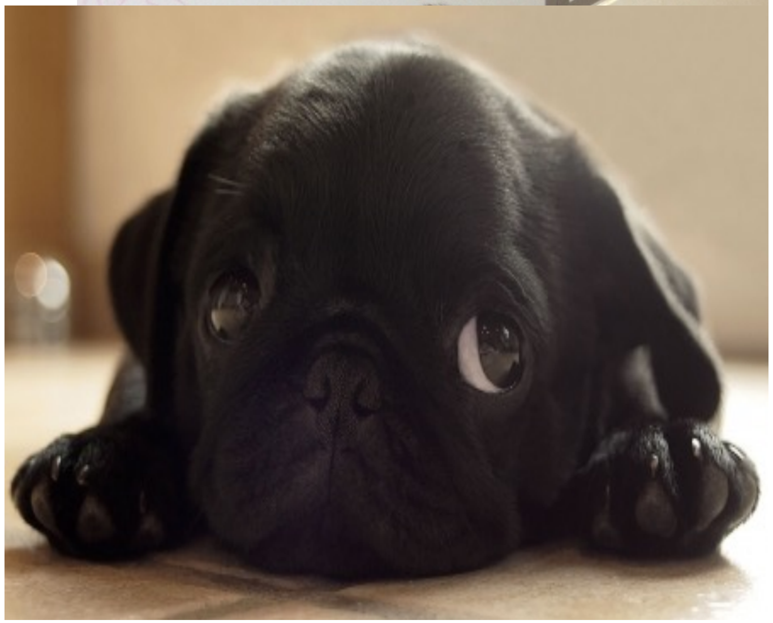
What RW could explain



What it couldn't

Pre-exposed

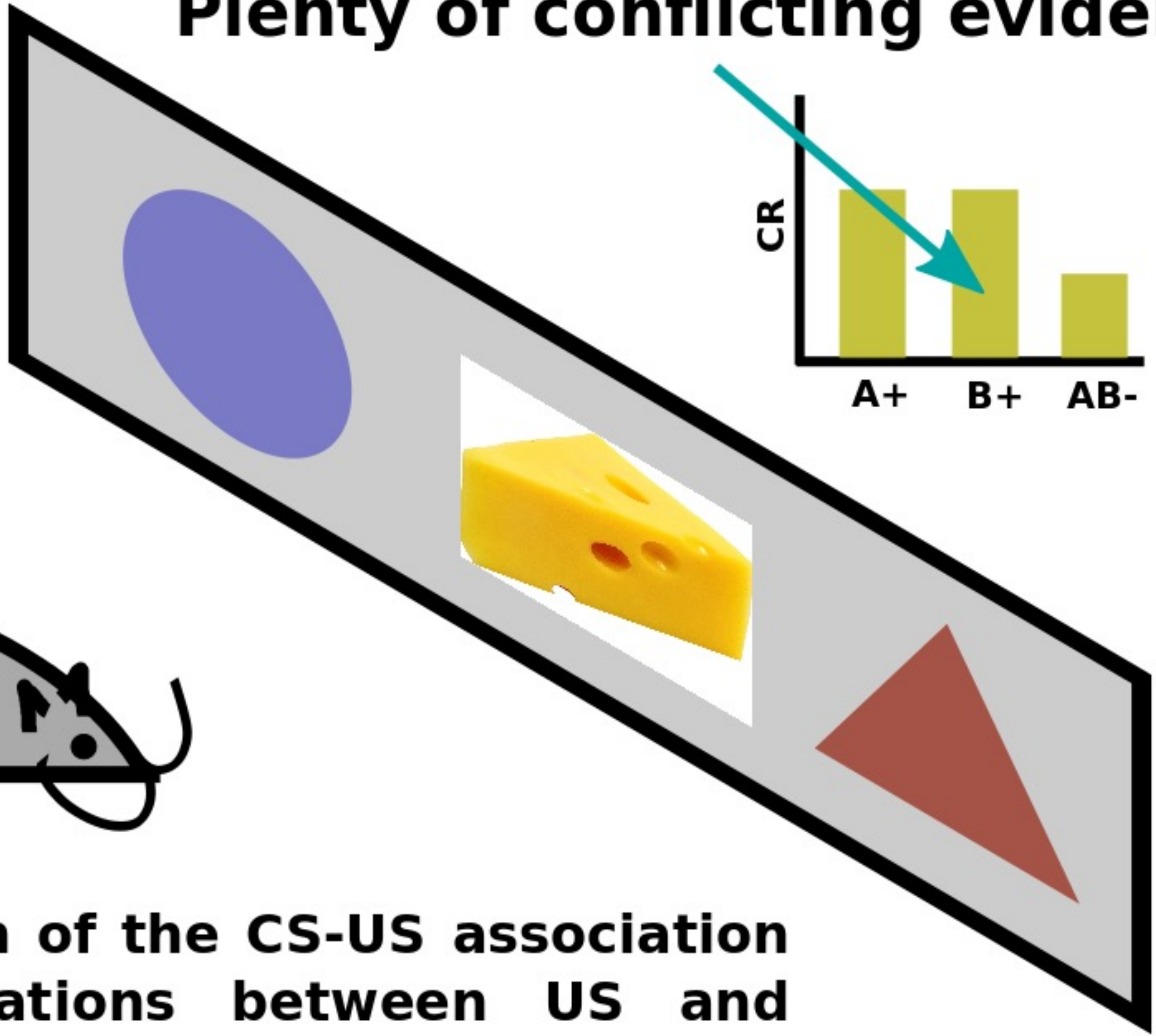


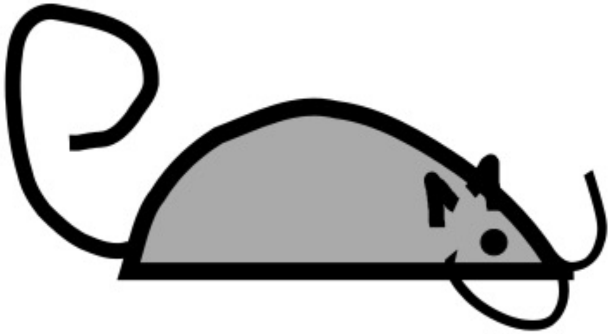


Latent inhibition

If value were just CS-US association, RW style models would be sufficient

Plenty of conflicting evidence





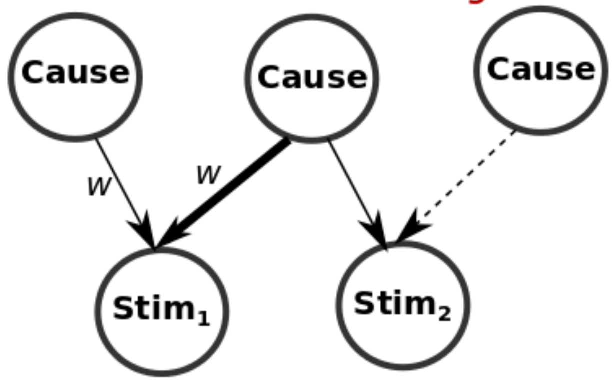
It looks like strength of the CS-US association depends on associations between US and other stimuli too. Why not model these relationships stochastically?

https://en.wikipedia.org/wiki/Rescorla%E2%80%93Wagner_model

Bayesian models of classical conditioning

(Courville, Daw & Touretzky, 2004)

LVs are causes that animals think lead to observations arising

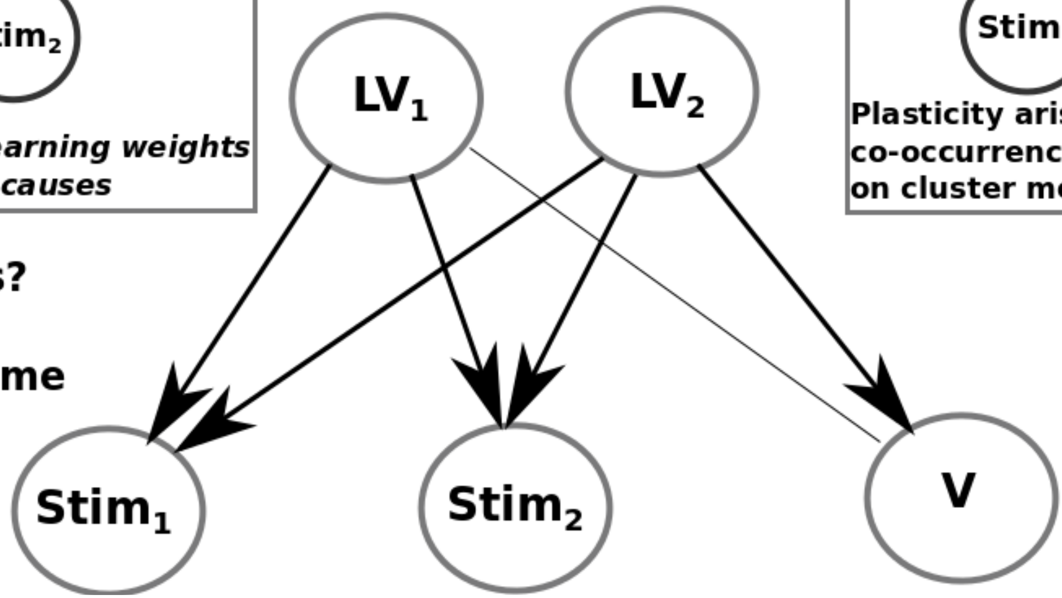


Plasticity arises from learning weights connecting stimuli and causes

How many causes?

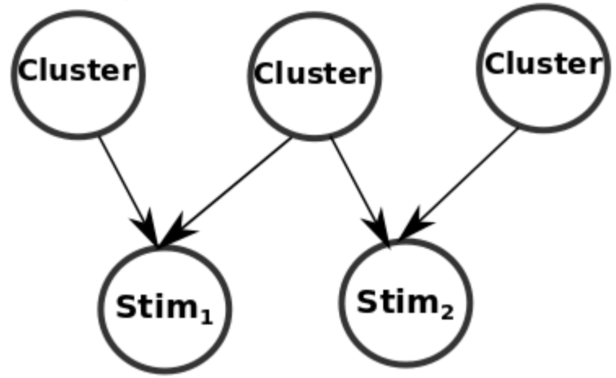
Where do they come from?

How to interpret latent variables?



(Gershman & Niv, 2012)

Animals group observed stimuli into clusters; these become causal LVs



Plasticity arises from learning stimuli co-occurrence frequencies conditioned on cluster membership (also learned)

Fails to predict summation-based conditioning effects.

What does clustering buy us?

Interpret latent variables as ***situations****.*
not ***causes***

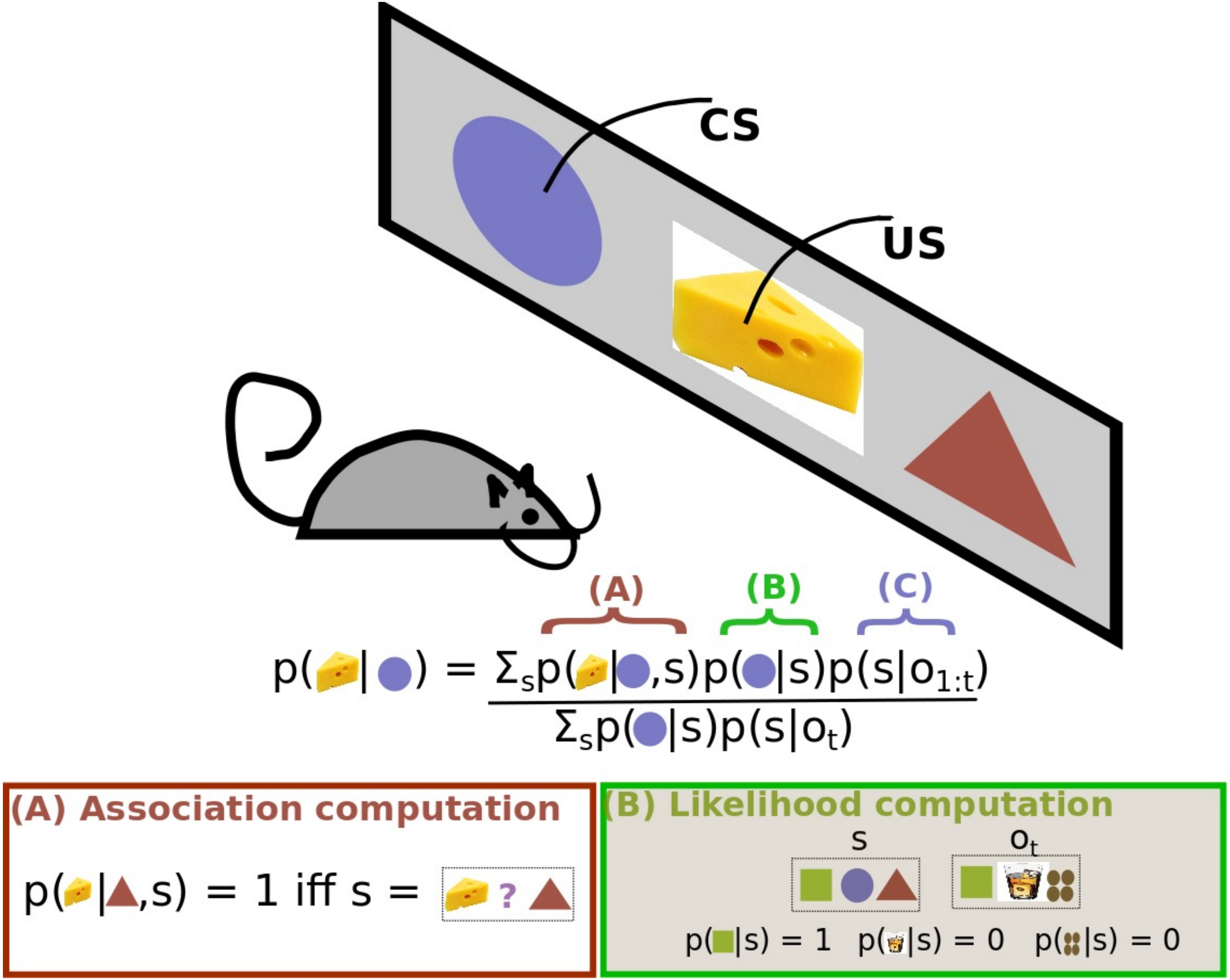
Index **situations** by stimuli co-occurrence patterns

s = **Yeah!**

s = *Smoooth …*

s = **Yikes!!!**

https://link.springer.com/article/10.3758/s13420-012-0080-8


CS
US
(A)
(B)
(C)
(A) Association computation
p(| ,s) = 1 iff s =
?
(B) Likelihood computation
s
o_t
p(|s) = 1
p(|s) = 0
p(|s) = 0

Problems for Bayesian conditioning models

- How to incorporate the role of time?
- How to incorporate the role of attention and salience?
- How to model the process by which animals learn the higher-order structure within latent causes?

Summary

- The mind learns by association
  - Associates novel with known, based on a number of ways of relation
- Association of novel to known causes generalization
- Association of known with known causes reinforcement
- We will talk about reinforcement next