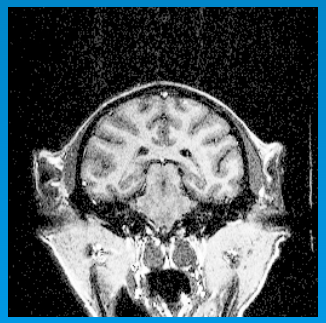
Natural Assessments in Animal Cognition

Rationality, representation, and comparative psychology



Two Related Classes of Questions in Cognitive Ethology

- Psychological explanation:
  When is it appropriate to use psychological (belief-desire) explanation?
- Rationality:
  Can animal behavior be evaluated according to some criterion of rationality?

And a challenge that unites both:
When is an animal judgment an inference and when is it “mere association”?

Mental Map of the Presentation

1. Review Bermudez' attempts to address these two questions
   - Inferentialism fails for non-linguistic creatures (and minimalism fails for at least some), but some belief-desire explanations informed by a careful study of **non-conceptual representation** are appropriate
2. Note that arguments B levels against Inferentialism for animals are also employed against Inferentialism for humans and point out similarities between human "intuitive" judgment and animal thought
3. Present a hypothesis about NCC that provides new research foundations for answering both questions
   - Don't expect a clear characterization of NCC, but I do
     - Offer some insights as to how non-conceptual content may be employed in reasoning
     - Show how one type of cognitively significant representation emerges from "mere association," and how it can be employed in at least one type of inference

Orienting Idea for My Project

- So there are these neurons, and when you put a lot of them together and give them lots of structured input, they can perform this really cool function of representation. Nobody understands exactly how this works, and there is great difficulty finding correlations between psychological items (beliefs) and neuronal activity.
- My idea: Let's not try just yet (as B and some minimalists do) to build up an atomistic, compositional psychological ontology from the ground up out of primitive perceptual belief elements. B is happy abstracting away from neurons too, but maybe too happy. In short, I don't like his compositional atomism and favor a more domain-oriented, connectionism-sensitive approach.
- So instead of focusing (primarily) on how to construct knowledge structures from primitive perceptual elements, perhaps we can focus on how these structures operate and work from there.
  - Give me some slack here; the projects should at least be compatible

Psychological Explanation and Critters

B feels that there are two basic options for a theory of animal rationality:

Inferentialism and Minimalism

- Inferentialism: The view that one can explain and predict the behavior of a creature by understanding the inferential relations between its propositional attitudes (beliefs, desires)

But Bermudez says:

B's Two Principal Objections

- Inferentialism (based on procedural rationality) fails for nonlinguistic creatures for two reasons:

1. Poverty of Empirical Evidence: Animals do not exhibit domain-general competence with these formal rules
   - Examples in literature of ability to use logical inference rules are dismissed by B as domain-specific, "mere conditioned responses" learned by awarding behavior
2. A Priori Impossibility: Our models of inference "understand" those rules in formal terms, as expressing formal (syntactic) relations between sentences, and non-linguistic creatures have no representations with suitable vehicles
   Remember B's Argument in Ch 8:
   - Either the premise-thoughts are available to deliberate introspection or not—if not, then they are unavailable for inference
   - No candidates introspectively-available for non-linguistic representations have suitable "vehicles" to fill similar functions

So we are after a form of representation that is non-propositional but has enough structure to explain animal behavior.

B's Characterization of the Search

Psychological Explanation and Representation

- Roughly, it's not sensible to talk about psychological explanation unless there's a significant degree of flexibility between behavior and environment; psychological explanation inappropriate when behavior:
  - Is always triggered by same [external] stimuli
  - Always takes the same form
  - Is present in all members of relevant species
  - Occurrence is independent of individual creature's history
  - Once launched, cannot be varied
- This kind of flexibility is only possible when we have some form of internal representation on which the creature can act; but if animal thought is non-propositional in nature, **we're after a form of non-conceptual representation with enough structure to explain animal behavior**

Where to Begin?
Minimalist Accounts of Representation as Perception

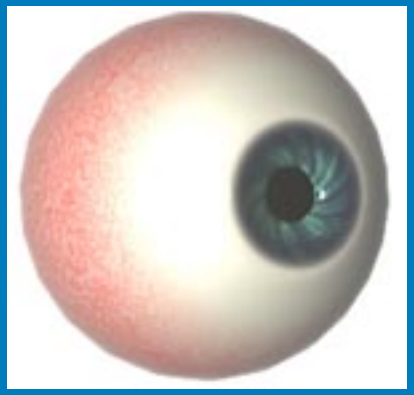
- Things get a little murky…
  - Dummett's "spatial images superimposed on spatial representations"
  - Cussin's cognitive trails: perception as registering ways of acting on environment, complex links called cognitive trails
  - Campbell's causal indexicality: perception as a set of causally indexical statements ("—tasty when I eat it," "—scratchable by me," "—I can mate with", etc.)
- B offers a nice, pseudominimalist story built out of perceptual affordances and instrumental beliefs

The Comparative Part:

In fact, inferentialism doesn't work so well for people either…

- *Human* judgments frequently violate formal inference rules
  - Conjunction Fallacy (Tversky + Kahneman)
  - Base-rate neglect
  - Law of Large (small) numbers
  - Affirming the Consequent, Denying the antecedent (Wason)
- In fact, only a small subset of human behavior is governed by formal inference; other decisions are governed by rules which have come to be called *inferential heuristics*
  - Competence/performance, fine, but let's stay on task
  - (I almost want to say that humans don't exhibit domain-general competency with formal inference rules either, the domain of language is just really big…)

For



Instance

(Anybody feel like claiming that text-parsing is serial or wholly syntactic?)

I cdnuolt blveiee taht I cluod aulacity uesdnatnrd waht I was rdgnieg . The phaonmneal pweor of the hmuan mnid, aoccdrnig to rsceearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoatnt tihng is taht the frist and lsat ltteer be in the rghit pclae.

The rset can be a taotl mses and you can sitll raed it wouthit a porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lterter by istlef, but the wrod as a wlohe. Amzanig eh? And I awlyas tohghut slpeling was ipmorantt..! And for toshe of you wtih mroe tmie tahn ohrets you wlli ntocie taht not olny are msot of the wrosd a mses but smoe of tehm are cpmlpoetley msis seplt awslel….!

What is an (Inferential) Heuristic?

- A rule or process of judgment which is constrained by cognitive economy. Heuristics are typically highly efficient, intuitive, and affected by learning—but which can produce judgments ruled faulty by formal inference rules. May be contrasted with **algorithmic** reasoning.

Examples of Tasks that Trigger Heuristic Judgment

- Is Linda more like to be a bank teller or a bank teller and active in the feminist movement? (Representativeness Heuristic)
- Do more words in *Animal Cognition* end with _ing or have 'n' as the second-to-last letter? (*Availability Heuristic*)
- Which river carries more gallons of water per second, the Amur or the Danube? (*Recognition Heuristic*)

So What's the Latest Line on How These Inf. Heuristics Work?

- T+K: Whenever we are called upon to formally evaluate some attribute, we may use an associated type of parallel assessment on a prototype instead.
  - Instead of attempting to formally assess likelihood of Linda's profession, we assess how similar Linda's description is to our salient feminist prototype
    - A high response on this assessment leads subjects to judgment, ignoring the fact that any time Linda is a bank teller and feminist, she is also just a bank teller (Conjunction Fallacy)
- Natural assessments performed on a prototype is one of the things we call "intuitive judgment"

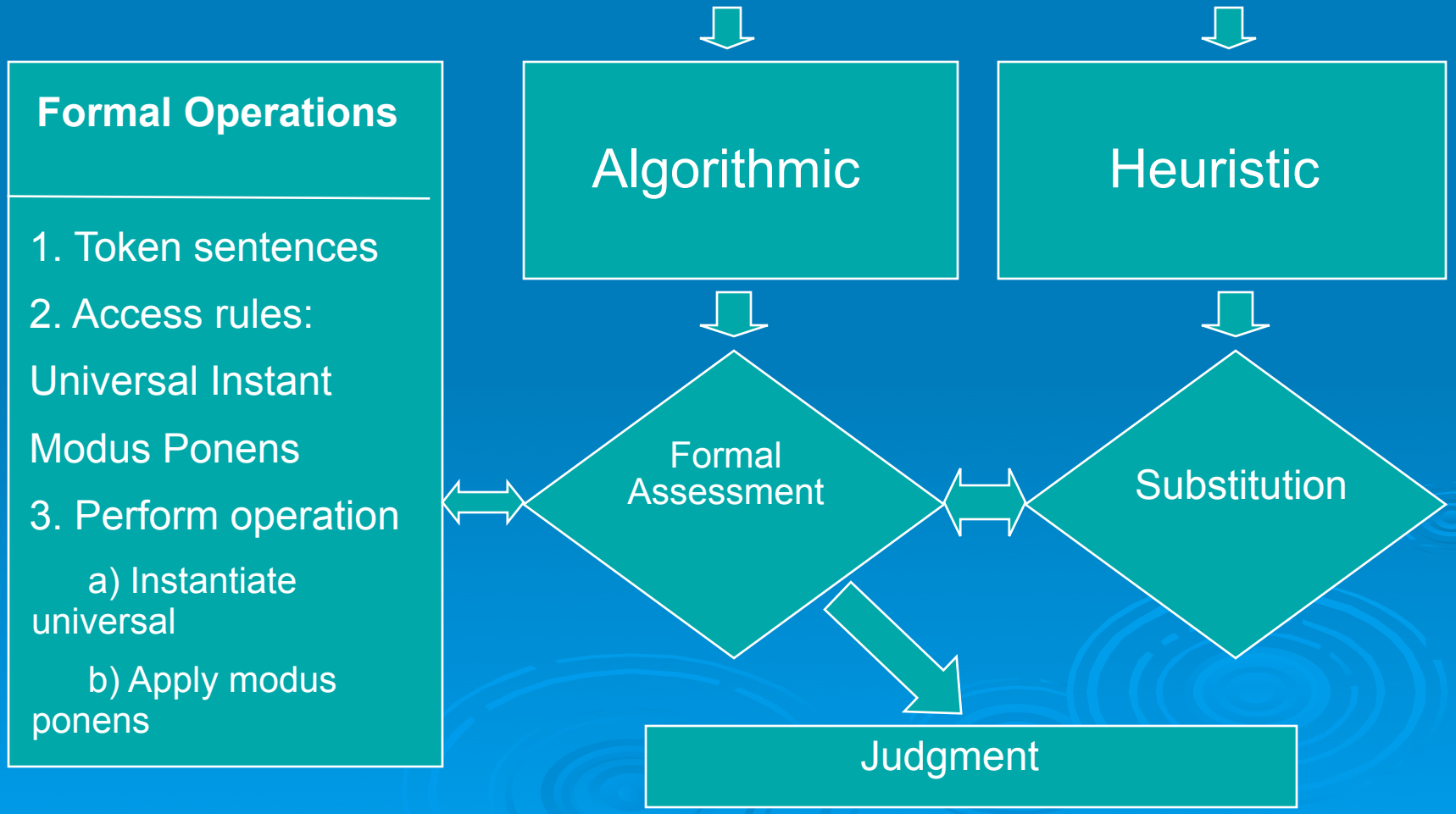
Why Are You Telling Me All This?

- Patience; this finally brings me to what I really want to talk about: **natural assessments**.
- "Some attributes are permanent candidates for this heuristic role because they are routinely evaluated as part of perception and comprehension, and therefore are always accessible. These natural assessments include *similarity, cognitive fluency in perception and memory, causal propensity, surprisingness, affective valence, and mood*." (T+K, 55)

So What's the Idea?

I ask an easy inferential (logic) question:

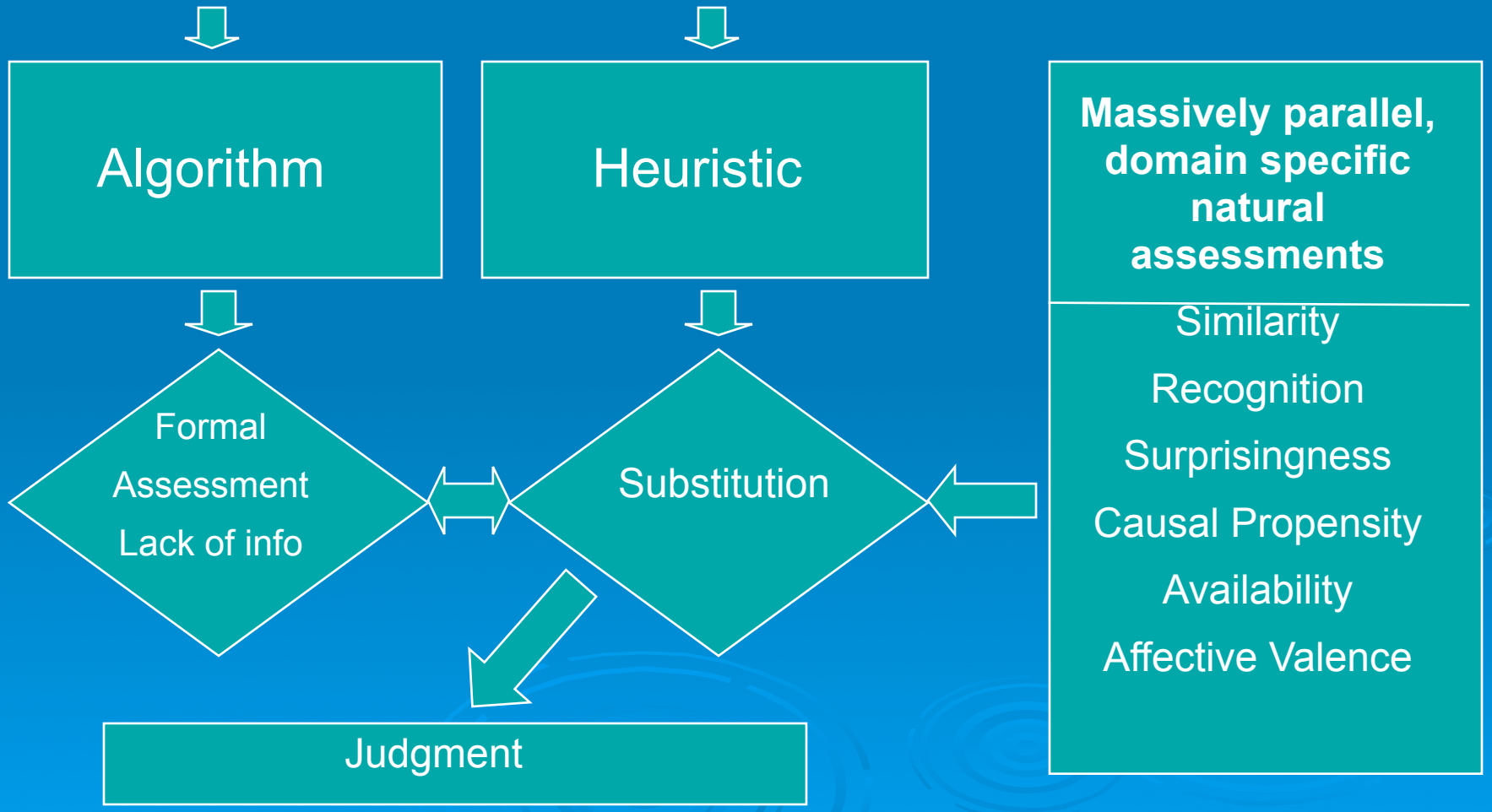
If Socrates is a man and all men are mortal, is Socrates mortal?



So What's the Idea?

I ask a hard question:

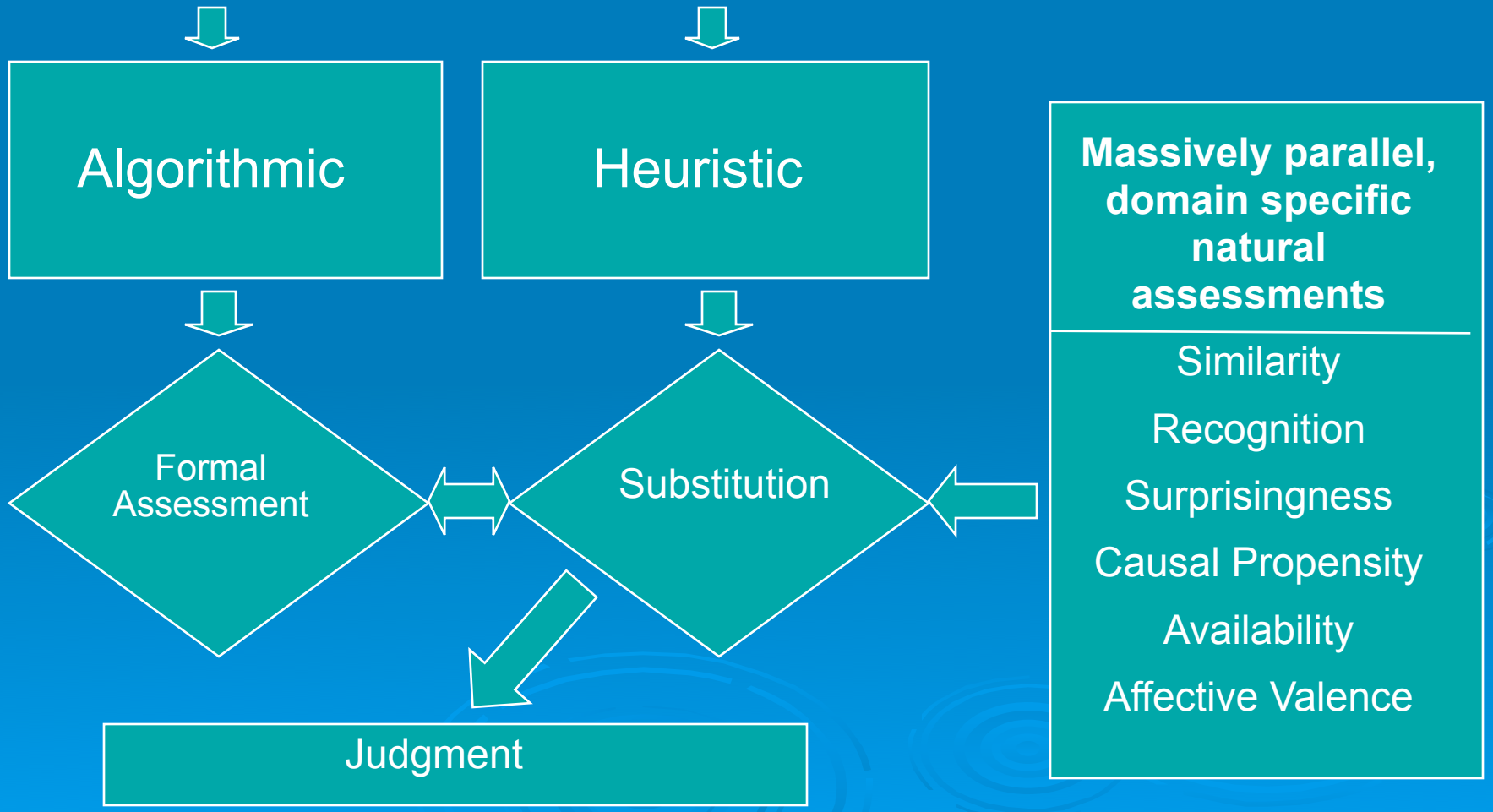
Which river carries more water per second, the Amur or the Danube?



So What's the Idea?

I ask an easy question the subject still screws up:

Is Linda more likely to be *a bank teller* or *a bank teller and active in the feminist movement*?



Why We Should Expect Critters to Employ Natural Assessments

- Argument from Connectionism: Natural assessments are essentially complex pattern-matching processes that can in principle be performed by neural nets
  - NA's are robust, fuzzy, parallel, affected by context, etc.
  - In other words, it's a lot "easier" to train neurons to assess similarity than use universal instantiation, and animals have much of the same neural stuff as we do
- Argument from Comp Psy: Pre-linguistic infants/ourangs show surprise when faced with causally impossible perceptual evidence, animals can categorize by similarity, etc.
- Argument from Evolutionary Continuity: There seems to be no principled reason why heuristic reasoning and natural assessment mechanisms would be unavailable to non-linguistic creatures (unlike, perhaps, procedural rationality), and lots of reasons to think they would be

Prototypes?

- The story we have so far isn't quite good enough
  - Where do we get these prototypes, and how are these natural assessments performed?
  - Why do we have different prototypes in different domains for the same category?
    - E.g. cop when speeding in the country, cop when walking down a dangerous alley
- Prototypes are domain-specific, representative instances of categories/concepts which are dynamically aggregated from interactions in that domain

The Origin of Prototypes

- If NA's are pattern-matching, then knowledge structures are the patterns
- Knowledge structures: Packets of organized information about the world that include information about the relations amongst features (prototypes)
- Tacit/Naïve theories: Knowledge structures that also have causal, explanatory features (higher-order causal relations)
- KSs are formed from experience with a class of thing ("purposeful," attentive)
  - Developed by forming associations from regularities; these include:
    - Mood/affect/drive associations with salient perceptual/cognitive cues
      --Especially involving past success/failure achieving goals
    - Propensity measures
    - Causal regularities (B reliably follows A)

Examples of Tacit Theories

- A tacit theory about physical objects posits the notions of object cohesion (objects move as wholes on continuous paths), boundedness (objects don't penetrate one another), rigidity (objects don't change shape as they move), and "no action at a distance" (objects move separately unless they come into contact).
- The "egocentric tendency" to assign a special value to ourselves causes us to find coincidences that happen to us a lot more surprising and special than equally unlikely coincidence that happen to others, to minimize the bad impact we've caused, believe in our power to influence random events, and that we are better than average at driving, sense of humor, parenting, reading others' minds, sexual prowess, etc.
- Tacit theories and non-linguistic creatures
  - Dishabituation attention research shows that infants have some form of tacit object theory
  - Hauser's Orang studies and fluid mechanics

Proto-Prototypes and Our Pets' Pet Theories

- The key move of this project is to extend the notion of knowledge structures/tacit theories to non-linguistic creatures
  - There's a sense in which everyone is struggling with the same thing, and a proto-KS is just a place-holder for a solution to the problem. The reason I favor this approach rather than the others mentioned is because is because I can give what I think is a more accurate account of how inferences are made from non-conceptual content that is useful for comparative psychology
  - The key move here is to make the notion of typicality broader than concepts/categories
  - Proto-prototype is a non-conceptual, fuzzy, aggregated measure of a class of experience which expresses what is "typical" for interaction with that class of activity—in other words, against which natural assessments can be made
  - So how will this work in the realm of animals? (Let's not try to be a bat here.)

Animal KSs

- Animal KSs are going to be broken down by domain (mode—e.g. predation, mating, foraging) and goal-desire (for types of stuff—e.g. type of food, mate, hiding location).
- All learning involves ability to extrapolate from experience to new situations

Animal Knowledge Structures

- I focus on KSs to leave open the question of causal links
- Question of the day: What will be the main difference between human KSs and animal KSs?
- Content: What kinds of regularities can be detected will be highly dependent upon:
  - Perceptual abilities
  - Memory/stable representation constraints (long-term storage, cognitive maps, etc.)
  - Innate attention/affect mechanisms (Garcia effect, affective valence)
  - "Purposeful" interaction with the domain—can only expect causal, goal-oriented contents if experience includes purposeful action in the domain (association of certain perceptual cues with satisfaction of certain drives, motor patterns)—strongly affected by success/failure in satisfying certain goal-desires

* And we should expect all these things to be highly attuned to a critter's natural environment

Animal Knowledge Structures

- I picture a continuum of elegance/power for knowledge structures varying along the following dimensions:
  - Specificity: How specific is the content of the KS? Can critter only experience a non-specific sense of surprise, or can it identify the reason for the surprise? Can critter only register dissimilarity or can it identify what is dissimilar?
  - Causal richness: Does the KS allow for causal or instrumental predictions/explanations?
  - Interconnectivity: How "networked" are the subcontents of the KS to each other? How interconnected is the KS to other KS?
  - Metacognitive access
- Bottom line: We should expect animal KS to be highly domain-specific, and optimal inference performance can only be expected in environments rich in cues to which animals are evolutionarily attuned
  - Predators should have more elegant KS in predation domains, social critters more elegant KS in social domains, *because* their perceptual abilities and innate mechanisms will make them more attuned/attentive to those regularities

Getting (a Little) More Precise

- At level of functional abstraction, we can characterize the relationship between judgment, natural assessment, and knowledge structures as follows:
  - NA(Input, KS) = Response level
    - Where NA is one of the parallel natural assessments, KS is the knowledge structure being used for the assessment, and Input is perceptual input or representation under consideration
  - From here we can see how judgments can be made; if we are considering two choices, response levels may be compared

Examples

- Say my cat is assessing whether to make a jump between two positions she never jumped between previously; say NA = similarity, KS = typicality measure for successful jumps in the past (w/ fuzzy sensitivity to perceptual cues, body motor patterns, affect cues, etc.)
  - Perceptual input will have variable affect on response
  - If response exceeds a fuzzy threshold, she will jump; if not, she will not
- “Choosing” between two options
  - If for some two inputs, NA(KS, Input1) > NA(KS, Input2), choose 1.
- Metacognition and dolphins
  - If response is open to another cognitive mechanism, then dolphin can register when it is uncertain; let NA be similarity, KS1 be typical perceptual conditions for successful performance sound assessment task, and KS2 be typical response conditions from KS1 on sound assessment task
  - NA(KS2, NA(KS1, SoundInput)) = response; if above a certain threshold, bail out

Representation Recap

- KS's and representation: just notice that they serve the role in B's picture for representation
  - Formed from experience
  - Allow flexibility of behavior via natural assessments

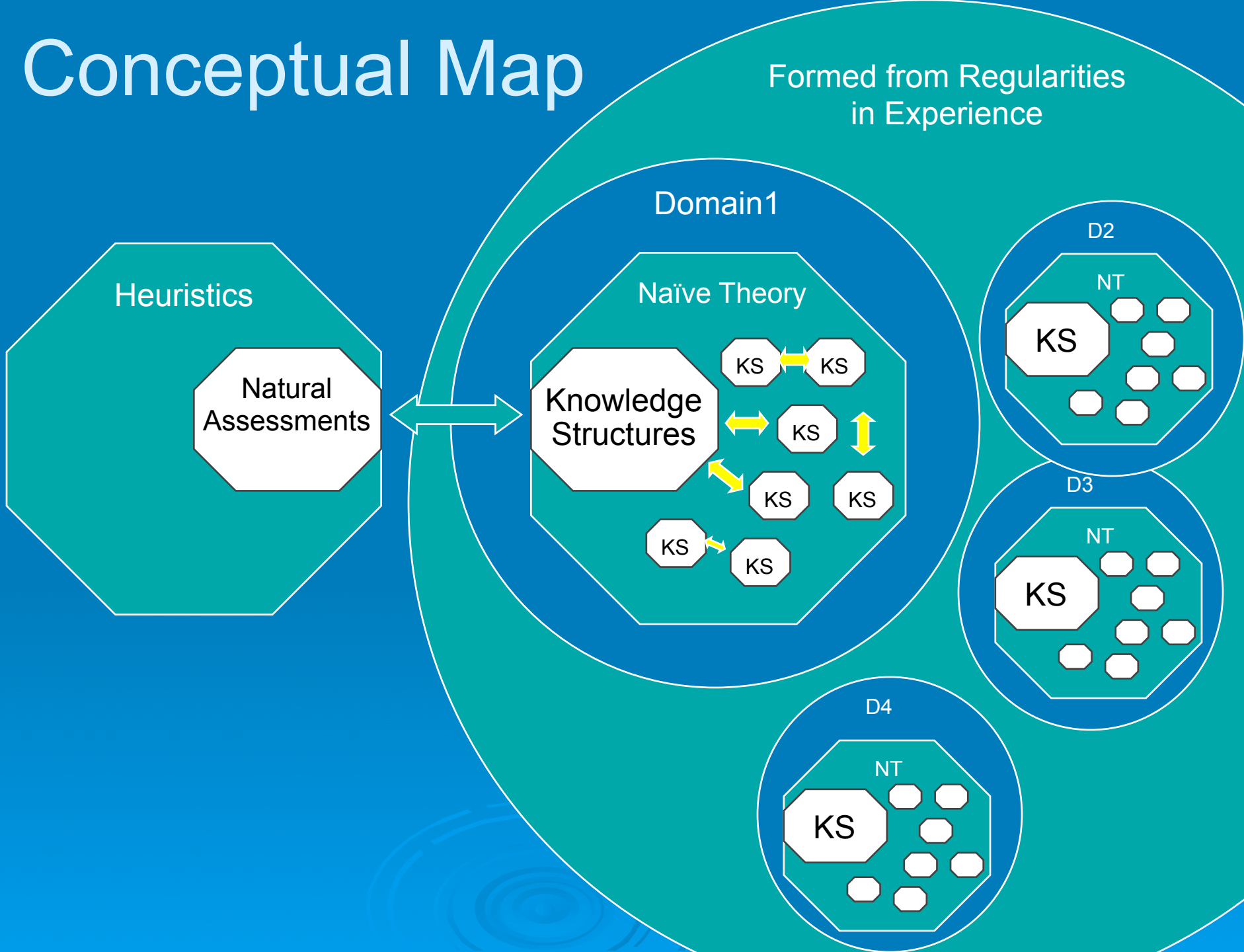
Theory Strengths

- Shift of focus to habituation paradigms to test which regularities animals are sensitive to via natural assessments
- Accommodates both KS sensitive only to perceptual gestalts up to higher-order property attributions
- Mid-level explanations that don't need to understand representation at level of individual neurons
- Shift to knowledge structure avoids project of trying to come up with determinate contents for proto-beliefs via success semantics
- Model can be used in addressing two questions stated in intro
- Compatible with other accounts
- Gains for comparative psych:
  - build on recent models for human intuitive judgment
  - Gracefully trim existing theories to accommodate our general reluctance to attribute higher-order cognitive operations to non-humans
- Good for evolutionary continuity—no good reason to suppose nature divides brains into human and non-human
- Accommodate terms like "mere association" without the bad aftertaste
- Fits other recent findings like Steve's metacognition for dolphins

Theory Weaknesses

- Need an account of reification
  - The specificity criterion is related to Quine's question of reification—how do we move from sensitivity to perceptual gestalts → individual properties → clusters of properties obeying higher-order properties → individuated elements
- Need a deeper account of how the different natural assessments are related to each other
  - How are they performed, what's being computed, are some built out of others, etc.
    - Optimistic that these answers can come out of research
- Domain talk is cheating
  - Domains will be interconnected, overlap, fuzzy boundaries, sub-domains
    - Nevertheless, I hear the animal folks talking about predating mode, mating mode, social mode, foraging mode with radically different behaviors; can get sub-domains out of searches for different goal-desires

Conceptual Map



Research Questions

- Which natural assessments are available to species when making decisions?
- Which regularities can each animal detect?
- How specific? Rich? Interconnected? Open to metacognitive access? (dishabituation paradigm, likelihood experiments)
- Which factors affect when animals use certain assessments and not others, and why? How do some assessments get associated with certain prototype and domains?
- And don't even get me started on Tinbergen's 4 Q's.

Research Difficulties

- Some of the domain variables will be internal and difficult to monitor. For example, when an animal is in **forage** mode, we should expect different results from natural assessments on identical inputs from when it is in **mate** mode; in other words, different domains may be difficult to distinguish, so regularities/irregularities in assessments may be hard to detect.
- The old bugbears introspection and intentionality: it's clear that there are complex relationships between the content of a KS and:
  - Introspection upon its reports
  - "Purposeful" interaction in the domain (allows for more complex conditional associations and goal-related associations)

So: Are these decisions inferences or "**mere associations**"?

Possible responses:

2. Shut up! I hate that question; it's boring, trite, and uninformative!
3. Do you want to know something normative? *Why are you asking me that?* Fine, for an evaluative model, B's levels of rationality ain't so bad, so look there—and have fun with success semantics.
4. By 'inference,' do you mean what B means? If so, see above.
5. What would we say about a domain-specific inferential mechanism built out of "mere association" whose output matches formal procedures?