Concepts

Overview

- Concepts are a general form of mental representation of knowledge
    - Captures what we know about the things in our world and how they operate, as well as how we know them as they are
    - More general/broad than logical implications
- Kant believed that our minds create our conscious experience by applying concepts to our senses

- Two possibilities of how concepts are acquired
    - Innate
    - Learned from experience

Theories of Concepts

- What sort of cognitive entity is defined as a concept?
    - Definitions: Concepts are akin to definitions in a dictionary
    - Prototypes: Concepts are abstractions of the things they apply to
    - Exemplars: Concepts are good examples of the things they apply to
    - Causal: Concepts are explanatory categories

Definitions

- A definition gives the necessary and sufficient conditions required by a thing to be an instance of that term
    - Works well for technical concepts (Ex: Bachelor, triangle), but is limited
- Natural-kind concepts, concepts applied to things that aren't man-made are more difficult to determine conditions for
    - Ex: Tiger - if it had its stripes painted over and is now a herbivore, we'd still call it a tiger despite its requirements being "striped" and "carnivore."
        - This suggests that certain properties of a tiger may be dispensible
- Also cannot account for typicality effects
    - Every property in a condition should be important equally, since they must all be fulfilled for the definition to hold
        - But some properties might be considered more important by people compared to others

Prototypes

- A list of features that are typically found in members of a particular group
    - Members in the group don't have to require every feature in the list
- To determine typicality, we compute the weighted sum of the features that the object shares with the prototype minus the difference
    - Expressed in Tversky's contrast rule
        - $Sim(I,P) = a \cdot f(I\&P) - b \cdot f(P-I) - c \cdot f(I-P)$
- One of the strengths of the prototype theory is that it accounts for the typicality effects

- Has been applied to other, non natural-kind concepts, such as artifact, psychological and psychiatric concepts
- Difficulties
    - Technical concepts like a triangle have typicality effects, despite it not having an influence in determining their category
    - Features don't add weight independently
- Prototype theory suggests that people don't access info about the size / variability of a class or specific examples of a class, but they do use both kinds of info

Exemplars

- A good instance of a particular concept
- Similar to the prototype theory, people make concept judgements by looking at a similarity score
    - However, they don't have the idea of prototypical features in mind, but determining if something fits into a category is computed by
        - Looking at the similarity between the instance and one or more exemplars
        - Prototype is made out of exemplars, then similarity is computed with prototype theory
- Difference between exemplar and prototype is that exemplars don't remember any info about a category
    - Instead, comparisons are made
- Addresses shortcomings of prototype theory
    - Takes into account the fact that people know how large and variable a given category is
    - Preserves info about how features affect category judgements in a correlated/dependant manner
        - Different from prototype (independant)
- Problems
    - How are exemplars affected by learning something new about a concept
    - Why are there "general" concepts at all
        - Instead of a bird concept, why not have multiple bird instances

Causal

- A concept of X is part of a theory about what it takes to be a part of the concept
- Accounts for centrality of some features
    - Ex: Straight banana is more typical than a straight boomerang
- Problem
    - Contribution of similarity to diagnosis, rather than a theory (or may have less weight placed on it)

Evaluation of Concepts

- Summary of concepts:
    - Represent typical features that an object or event has, as well as causal information on how they tend to act
- Causal and exemplar theories seem to have the most plausibility

Representational Power

- While logic tells you a lot of what you can do in a given scenario, it is unable to tell you what not to do
- Minsky proposed to collect all information that is relevant in a given scenario into one representation, called a frame
    - A list of information that usually applies to a given situation
        - Consisting of slots and filler
            - Slots can be filled in with the proper information depending on the situation
            - There can be an example slot for exemplars
            - Relations slot to combine prototypes with exemplar and causal info
    - Frame is usually called a schema in cognitive science
        - A more specialized form is called a script
            - Sequence of events typically performed
- Note that a frame can be turned into a series of rules as well
    - But we lose the advantage of having all the info in a frame

- Concepts are organized hierarchically, with the most abstract at the root

Computational Power

- Inheritance is similar to applying a rule
    - Something **inherits** something else from the hierarchy
    - Could be simulated with rules through forward chaining
    - Similar to rules, inheritance supplies default information
        - There may be some exceptions where the default does not apply, the special case would override the default
    - Inheritance doesn't require the search that rules require
- Matching is similar to searching for rules, but applied to concepts
    - Process of finding out which concept fits the best with the given scenario or object
    - Visualized as a spreading activation so that the concept that is activated the most is the one that is matched
        - Things that satisfy the criterias are active through excitatory links in the relevant frames, those that do not are inactivated with inhibitory links
    - Exploits the fact that human memory is organized based on content
        - When confronted with a situation, we recall past information with the most similarities and least dissimilarities
            - Frames represent this with slots with categories like ("a-kind-of," or "a-part-of")

Psychological Plausibility

Planning

- As previously noted, when planning problems, logic is not efficient because if asked to produce a plan n times, it will make the exact same inferences n times
    - Not what humans do, we will recall past experiences
    - In a frame-based system, we can call on scripts rather than starting anew
        - May need to sometimes be modified, so cannot be used blindly

Explanation

- Schemata can provide pre-packaged explanations for previously experienced things
    - Faster than applying a rule backwards, or using abductive inference
- Limitations
    - Easy to over generalize
    - Not flexible
        - If given an unfamiliar scenario, a schemata based explanation cannot adapt, rule-based might be better

Learning

- Learning a concept is concept acquisition
- One method to learn a concept is by definition, i.e. by having the concept's features explained to us
- Another is by specialization, tweaking existing concepts to fit a new one
- Learning by copying is also prevalent, copy original concept with little tweaks
    - Examples unneeded
- Learning by generalization
    - Leave some slots blank in the frame
- Learning by combining two concepts together