Information Foraging

Brett Miller
Robert Travis
CS376

What is Information Foraging Theory?

- Hypothesis: in searching for and consuming information, humans adapt their strategies and environment to maximize information gain.
- Analogous to optimal foraging theory, which hypothesizes organisms adapt food-foraging strategies to maximize energy intake.

The basics of the theory

- Patches of information (e.g. an article, a web page)
- Cost associated with moving from one patch to another
- Information return function for within-patch foraging
- Information scent: cues that foragers use to estimate how much information they are likely to get from a patch and when to move on (think heuristics)

Examples from reading

- Analyst doing research to write a business intelligence newsletter
  - information patches: magazine articles
  - enrichment activities: filtering articles and organizing workspace
- MBA students doing research to write a strategic analysis report
  - information patches: various business articles
  - scent used to pick articles based on size (smaller is better) rather than perceived information
- Other examples?

Value of a quantitative model

- “The model is great, but how important is it if it confirms things we already knew?” --Neil Patel

- Is there something we should learn from the curves and specific inflection points drawn in the paper ? If so, I must have missed it. To me the equations are not validated by field studies and are therefore not useful -- “Sudheendra Hangal”

Limitations of experiment

- "The experiment used to test the hypothesis was done using the Scatter-Gather software. This raises some questions about generalizability" --Nick Briggs

- How realistic is the experimental evidence for validity of theory?

Limitations of theory in general

- “I tend to believe that the [search] space is multidimensional - scents can also lead you certain directions in your information search that is neither "good" or "bad" in a strict sense, but different. Such scents can be useful for exploratory tasks rather than goal-oriented tasks, such as in early stages of design” --Loren Yu

- “Also I think this aspect of foraging [enrichment] is less important today when much of the searching happens from a search engine, not with piles of books on a desk” --Neil Patel

Limitations of the Analogy

- “It's not as if information is a one-use item - once an animal eats an apple, the apple's basically gone” --Marcia Lee

- “ ... other animals' consumption of a resource doesn't make a resource more available - as it does in the case of google search. By observing what information people are consuming google or digg can make finding interesting information more easy. The analog in the natural world in this case isn't as clear.” --Michael Smith

Faceted Metadata for Image Search and Browsing

- A different approach to image search
- Given description of a picture, extract descriptive words and categorize them into different hierarchical facets
- Approach image search as a process of applying filters based on faceted metadata of images
- Flamenco relies on existing metadata and an algorithm that categorizes words into facets

Exploring Flamenco

- http://flamenco.berkeley.edu/demos.html

- Try doing a search for two different objectives:
  - Find a watercolor painting of a harp.
  - Contrast and compare depictions of musical instruments in Asia and Europe.

Discussion

- In what ways is using Google image search information foraging?
- In what ways is using Flamenco information foraging?
- How is it different and what are the optimizations? How does it improve information gain per unit time?
- Flamenco makes use of existing metadata; how would this scale to existing images without descriptions?