CSCE 636: Slide 13

- Langlois and Garrouste (1997)

Cognition, Redundancy, and Learning in Organizations

Langlois and Garrouste (1997)

- The prevalence of the information-processing view of organizations (biological or social).
- Limitations:
  1. The **measure of information** does not account for the **economic value** [or **quality** (YC)] of information.
  2. Information structures are usually seen as **fixed**, and even when they are learned, they are only updated to probability distributions on **known and given** contingencies.
- The question: How organizations create categories of understanding in the first place, and how information builds a **knowledge structure**.

Knowledge and Structure

- Distinction between knowledge and information: a stationary stock vs. a flow?
- Not quite! Knowledge is about **structure**:
  - Knowledge is not just an accumulation of information.
  - Knowledge must be regarded as a **structure**, and it **changes**.
- For a message to **stick** to the structure, it must be **meaningful** to the receiving system.
- Choosing an information structure involves investment: Information structures develop or evolve **slowly** and **cannot be recreated** or reengineered quickly or costlessly.

What Makes Structure *Knowledge*?

- A structure constitutes knowledge if it is ordered in a way that **produces results**, either immediately or later on.
- New knowledge can be acquired **without new information** being received!

Discussion (YC):

- In our context, there needs to be a **potential for action** in knowledge, which is distinct from pure information.
- What role does **memory** play in making sensory input into **potential** action, not just **immediate** action?

Structure and Meaning, According to MacKay

- A system's structure defines **conditional states of readiness** on which signals operate.
- The **overall configuration** determines meaning of a message.
- Meaning only becomes relevant when we consider the **range of other states of readiness**.
- A change in meaning implies a different selection from the range of states of readiness.
- A meaningless message is one that makes **no selection** from the range.

Whence the Structure of Knowledge?

- **Example**: at a railroad switching yard, the **shape** of the key (i.e., the message) determines the **configuration** of the various tracks. The meaning consists in the **change** the message **effects** in the arrangement of the tracks.
- However, where does the **structure of knowledge** – the railroad switching yard – come from?
- **Hayek's answer**: knowledge is primarily a system of rules of **action** assisted and modified by rules indicating equivalences or differences of various combinations of stimuli.

Summary

- An organism must respond appropriately to the stimuli provided by the environment.
- The neural system of the brain creates, with experience, a **semipermanent structure** or "map" that guides action.
- Learning is a matter of **self-organization**, that is, of **creation of structure**.
- A framework for understanding **self-organization** is thus put forward.

The Theory of Information, Redundancy, and Learning

- Quantity of information, or entropy, is defined as:

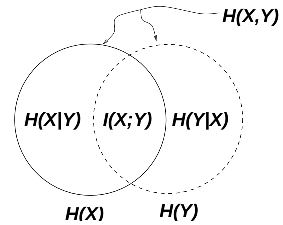
$$H(X) = -\sum_{x \in X} P(x) \log_2 P(x),$$

  where $P(x)$ is the probability of event that the symbol $x \in X$ will occur.
- This quantity is maximal when all $P(x)$ are equiprobable.

Example – Order in an Organization

- **Example:** Organization with members that are
  1. Equally skilled: random placement results in the same performance, i.e., the microstates are **equiprobable**.
  2. Highly specialized: random placement degrades performance, i.e., the microstates are **not equiprobable**.
- Thus, an organization with **division of labor** is a **lower-entropy**, or a more **ordered**, system.

Joint and Conditional Entropy



- Joint entropy:

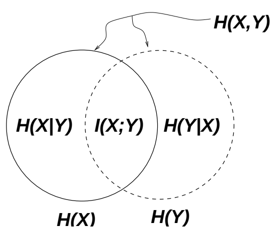
$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(x,y)$$

- Conditional entropy:

$$H(X|Y) = -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(y|x)$$

- Property: $H(X,Y) = H(X) + H(Y|X)$.

Joint and Conditional Entropy



- Conditional entropy:

$$\begin{aligned} H(X|Y) &= \sum_{x \in X} H(Y|X=x) P(x) \qquad (1) \\ &= \sum_{x \in X} P(x) \left(-\sum_{y \in Y} P(y|x) \log_2 P(y|x) \right) \\ &= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(y|x) \end{aligned}$$

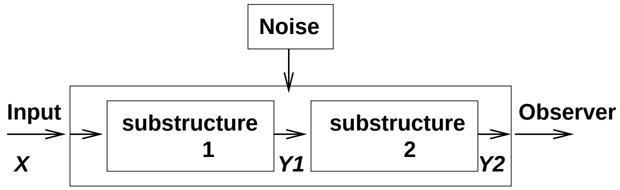
Redundancy

- Systems are vulnerable to perturbations or **noise**.
- To combat noise, we need to introduce **redundancy**:
  1. redundancy in the transmitted message, or
  2. redundancy in **system structure**.
- As more redundancy means less entropy, redundancy can be defined as:

$$R = 1 - \frac{H_{\mathrm{R}}}{H_{\max}},$$

where H_{R} is the entropy with redundancy and $H_{\max}$ without redundancy/noise.

Atlan's Idea of Ambiguity and Redundancy



- Totally constraining substructures (a perfect replica):

$$H = H(Y_1) = H(Y_2),$$

like a library full of identical books.

- Completely independent substructures (no shared information):

$$H = H(Y_1) + H(Y_2),$$

like a library full of books that do not refer to any other book at all (i.e., each book is totally independent of each other).

Atlan's Idea of Ambiguity and Redundancy

- If information is totally shared or totally partitioned among substructures, we get an uninteresting situation, i.e., no structure.
- To be interesting, we want the substructures to share information, which means nontrivial residual uncertainty (looking at Y gives you **some** information about X), which again means **ambiguity**.
- There are two forms of ambiguity in $H = H(Y_1) + H(Y_2|Y_1) - H(Y_1|X)$:
  1. **Destructive ambiguity**: signal loss $H(Y_1|X)$.
  2. **Autonomy ambiguity**: newly introduced information $H(Y_2|Y_1)$.
- Introduction of **autonomy ambiguity** in a system could lead to **self-organization**!

Atlan's Formulation of Ambiguities

- Since redundancy R is defined as

$$R = 1 - \frac{H_{\mathrm{R}}}{H_{\max}}$$

and $H_{\mathrm{R}} = H$ is the information in our system with redundancy, we get

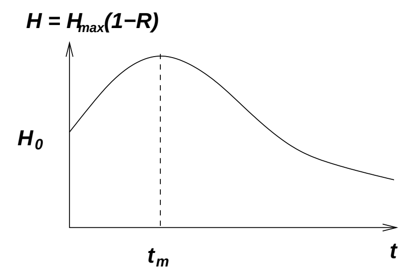
$$H = H_{\max}(1 - R).$$

- Differentiating this formula over time t, we get:

$$\frac{dH}{dT} = \underbrace{H_{\max}\left(\frac{-dR}{dt}\right)}_{(1)} + \underbrace{(1-R)\frac{dH_{\max}}{dt}}_{(2)}.$$

- Since R and $H_{\max}$ should **decrease** over time under noise, (1) is positive (autonomy ambiguity), and (2) negative (destructive ambiguity).

Atlan's Formulation of Ambiguities



- Integrating the previous equation can result in a plot like the above.
- Initially, gain in autonomy ambiguity supersedes the destructive force of noise, thus H increases.
- During this period, **self-organization** is supposed to occur (increase in complexity; decrease in R is converted into increasing H), until time reaches t_m.

von Foerster's Formulation

- Atlan's formulation can be a bit unintuitive, since when we talk about self-organization, we think of increase in order, or **redundancy**.
- This is exactly what von Foerster argued: Self-organization occurs when R increases.
- Thus, a system is self-organizing if $\frac{dR}{dt} > 0$.
- By differentiating

$$R = 1 - \frac{H}{H_{\max}}$$

we get

$$\frac{dR}{dt} = -\frac{H_{\max}\frac{dH}{dt} - H\frac{dH_{\max}}{dt}}{H_{\max}^2}.$$

Continued...

von Foerster's Formulation

- For $R > 0$ to hold, we need

$$H\frac{dH_{\max}}{dt} > H_{\max}\frac{dH}{dt}$$

- There are two cases we can consider:
  1. When $H_{\max}$ is a positive constant:

     $$\frac{dH}{dt} < 0,$$

     i.e., the entropy in the system must decrease (rearranging letters into order).
  2. When H is a constant:

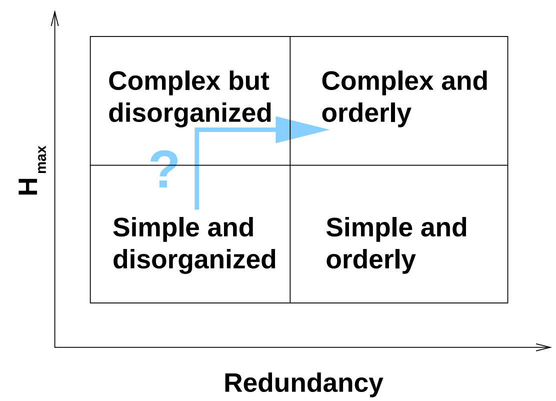
     $$\frac{dH_{\max}}{dt} > 0,$$

     i.e., the maximum entropy of the system must increase (introducing new letters).

Self-Organization: A Synthesis

- Either Atlan's or von Foerster's formulation alone does not suffice.
- Atlan equates self-organization with the increase in information content, which is unsatisfactory.
- The **redundancy** of von Foerster seems like a more appropriate measure of self-organization, but it is a relative measure, thus a system with a very small $H_{\max}$ and an even smaller H can still be seen has having maximum redundancy.
- Thus, the two aspects, **increase in information content** (Atlan) and **increase in redundancy** (von Foerster), must be considered together.

Self-Organization: A Synthesis



- There can be many routes from **Simple and disorganized** to **Complex and orderly**.

Discussion (YC)

- The formulation of L&G is a good candidate measure of the **degree of self-organization** in organisms.
- However, how do we design/build a learning agent to actually self-organize? Nothing is said about this.
- One idea: using this measure, we could probably derive a learning rule to **maximize** the degree of self-organization as described here (possible application to Kohonen SOM).
- We need to be cautious though since maximization algorithms can easily find loopholes in the objective function.
- Can **degree of self-organization** be used as a measure of **quality**?
- Is structure enough to hold **knowledge**? What role can **action** play in this formulation?

References

Langlois, R., and Garrouste, R. (1997). Cognition, redundancy, and learning in organizations. *Economics of Innovation and New Technology*, 4:287–299.