Criticism of Hierarchical Temporal Memory

# Summary of Hierarchical temporal memory

Hierarchical temporal memory (HTM), developed by Jeff Hawkins, is a model attempting to simulate the neocortex. It is a hierarchical model with bottom-level neurons receiving detailed inputs and top-level neurons receiving abstract input. At each level, the representations are more abstract.

# Why HTM Fails

## The real world isn’t strictly hierarchical

HTM abstracts features by special locality, but Hofstadter pointed out that the human mind doesn’t not really do that. The Letter Spirit project has given examples of why you cannot recognize letters by abstracting features that are spatially close. HTM posits that you must recognize lines before you recognize shapes. But if you try this method on the letters below, you would fail. Shapes do not really come from the lines.

# http://www.cogsci.indiana.edu/farg/mcgrawg/letter-spirit.gif

We tell letters from the whole, not just from the sum of its parts. This phenomenon is called “emergence” in Gestalt psychology.

## Objects form a meaning when they relate to each other

Objects are really are fluid concepts rather than essential qualities. A “chair” for example does not always have four legs and does not always have a back. Even more, anything which is designed to be a chair could function like a chair. It does not necessarily have to look like a chair at all.



It’s therefore important to determine the connections between different objects. For example, to make sense, you have to determine if an object is supporting a person or if the person is just squatting over it. If an object is supporting a person, then the object is functionally a chair. If not (as in a person squatting above it) then the object is *not* functionally a chair. Therefore an object is recognized not just by shape but by its function relating to different objects.

A flaw in HTM is that it only takes account of the spatial qualities rather than its function and the connections to different objects. Because HTM only classifies an object by its visual qualities similar to its internal representation, it does not take account of other, non-visual factors which helps us to classify.

## Three-dimensionality cannot be learned from two-dimensional objects

According to Molyneux’s problem, it asks the question: "if a man born blind can feel the differences between shapes such as spheres and cubes, could he similarly distinguish those objects by sight if given the ability to see?"

In order to solve Molyneux’s problem, we could explore a related concept called the Müller-Lyer illusion. 

The diagram shows that the red lines are of the same length. However, an optical illusion exists in such that the topical viewer would perceive the first line being shorter than the second line.

This optical illusion tricks us in distance perception. We rely on cues to judge the distance. The first figure has arrows pointing away so we judge them to be closer. The second one has arrows pointing inward (like how the walls in a building merge) so we judge them to be farther. Civilized people should be more susceptible to this illusion because they live in rectangular environments, while rural people should be less susceptible to this illusion because they do not experience rectangular environments. Several researchers have confirmed that this is the case as they have found that some rural Africans are not susceptible to this optical illusion. This suggests that our ability to construct three-dimensional objects from two-dimensional objects is learned.

However, our ability to understand three-dimensional images is not learned. For example, blind people cannot see, but they do experience *spatial* perception. They can spatially navigate the world by touch and equilibrioception (sense of balance).

So our ability to understand and manipulate three-dimensional objects isn’t learned. It’s innate. But our process of constructing 3D objects from 2D images is probably learned. The problem with HTM is that it only addresses the latter, but not the former. So therefore it cannot understand 3D stimuli, like the “Cat-Dog Problem”.

In order to understand images, we must be sure that the actual 3D structure makes sense. Without the ability to understand and manipulate 3D objects, we cannot differentiate between Cats and Dogs and recognize chairs from non-chairs.

# Identifying the Meaning of Irrelevant Features

In order to identify the meaning of features, we have to make sense of the surrounding and distractions (irrelevant features).

In order to build an AI, we first have to start out recognizing noisy images or images with overlapping objects.

The AI has to filter out the noise and distractions and predict the real content.

There are two steps in filtering out noise.

1. Recognize the distractions as “symbols”.
2. Ignore the symbols to predict the underlying stimuli.

This process is known as “partial matching.”

# Partial Matching

We recognize objects by portions and hints (the glare of white in the Ziploc bag). We then reconstruct the complete object by our imagination. By the time we detect the object we have imagined it in our mind from small portions, we assume that the imagination is the "real thing" that we see.

We deduce the size of the objects also. If we see a *small* scrunched up pile of a flexible object, we could assume it is clothing.

I think our brain creates a map of "recognized" regions and "unrecognized" regions. It's like the familiarity bias in short-term memory. We seek novel stimuli and do not seek familiar stimuli.

# Unconscious Priming

Our brains might be multitasking our detection process. We simultaneously recognize it as a building and as a grate, but we consciously perceive it as a building.

The prefrontal cortex of the human brain has been suggested to play a part of “priming” stimuli.

# Motivation and Object Recognition

Motivation is essential for object recognition because we must detect connections between objects. Motivation is implemented by the machine processing objects and summarizing goal states and predictions in a compact form called “summary-images.”

# Storing Goal States

An "image" is an impression of the overall scenery, experience, or prediction.

It's a mix of emotions such as time, fear, pleasure and determines if the action was worth it.

When we recall the past, we merge all the pleasurable and painful stimuli to summarize the overall benefit of it.

# Evaluation of Goal States

“Summary-images” summarize all the feelings went through the prediction or past experience. If the summary-image tends towards pleasure instead of pain, then it will decide action which will lead up to the summary-image.

Summary-images can also manage processes. For example, how does one choose to read a sentence out loud or not? One imagines the summary-image of the result: the completion of reading-out-loud. It then focuses on how to attain the summary-image. It can activate a process which mimics the internal sound. That's how control and inhibition works.

Summary-images include the time necessary, the rewards, and the summary-image itself. It's thus makes long-term investments less rewarding than short-term investments.

# Combining Evaluations of Goal States

Some things can be desired as "boring" if it's a redundant task that isn't necessary to reach the desired summary-image.

Or if the combined summary-image of the task and the desired summary-image doesn't add up.

Writing is boring. But the task of teaching others is interesting. So the person comes up with an overall evaluation of the goal. It becomes interesting.

But let's say the person has another summary-image of becoming rich. The evaluation of that summary-image doesn't compare with being a teacher. So he chooses to reach the summary-image of becoming and being rich.

He then says "speaking is boring." What he really meant is that the summary-image of speaking is boring.

He says "teaching is interesting." What he really meant is the summary-image of teaching others is good, but the summary-image of becoming and being rich is better.

# How to Attain a Goal

The robot imagines the summary-image. It then tries to figure out how to achieve the summary-image to be certain.

The system can use "logic" via a series of deductions.

It has to be passionate about performing the tasks which lead to the certain summary-image. It has to be curious about other possibilities.

Boredom arises when it has explored many possibilities but failed to find a goal.

The agent also experiences temporal discounting.

The agent also has to develop passion in solving the problem itself.

# Action

Willpower can also be related to cognitive dissonance. We perform actions that are consistent with our beliefs. The more we emphasize our beliefs, the more actions that we will perform that are consistent with these beliefs.

Pain from painful thoughts is increased when we are not certain that we can overcome the painful thought. Attention is always focusing on the worst thoughts or the best thoughts.

The more we are certain that I will die from an action, the more I will refrain from doing it right now. The more we certain that we will experience short-term gratification, the more likely we will do it.

Action depends on certainty. Prediction and appraisal do not.

# References

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Hawkins, J., & Blakeslee, S. (2007). *On intelligence*. Macmillan.

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases.