

Learning

Tom Rochette <tom.rochette@coreteks.org>

February 6, 2021 — 9686e64b

0.1 Context

0.2 Learned in this study

0.3 Things to explore

- Unsupervised classification learning (learning without others giving us feedback or without reward feedback loops)
- Importance of training from simple to complex problems?
 - Incapability of jumping from one state to another without going through the transitive states?
- By learning enough examples, the brain is able to reconfigure itself in order to be more efficient and by the same means generalizes certain concepts, thus increasing its understanding
- Hierarchy of abilities (sounds, words, sentences, ideas)
- Is there a maximum amount of information in the universe?
 - Given “Nothing is lost, nothing is created, everything is transformed.”, we can assume that there’s a certain amount of medium on which to store data, and so the answer is yes
 - Is it possible to be “done” with learning all the facts in the universe?
 - * If one “is” the universe, yes
- How can one discover the most appropriate order in which to learn unknown and new knowledge?
 - Learn from others
 - Randomly sample (trial and error)
- How do we learn to differentiate background noise from what we are actively trying to listen to?
- At what level of abstraction should we be thinking about learning?
- Is it more worthwhile to know a lot of skills at the 80% mark or master a few at the 90%+ mark, considering that to get from 80% to 90% requires a major time investment compared to 0 to 80%?
- Can I improve my current methods?
 - If so, where am I inefficient the most?
- Is one of my learning methods/techniques weak?
- How can I prevent myself from making errors?
 - Simulate alternatives, define a time limit for simulations (budget)
- Is learning about input and output?
 - Can learning occur without any transformation happening?

1 Overview

Learning is lossy encoding (similar to lossy compression).

2 What is learning?

- The process of reducing the distance between the behavior of an agent in reaction to a stimuli and what the optimal behavior is (supervised learning since we have examples of “optimal” (or at least better than random) behavior)

- The process of discovery where the agent is trying to accomplish some goal while minimizing/maximizing some set of metrics (for instance energy, speed, duration, power consumption, etc.)
- The process of understanding something and being able to recall this information when needed (for instance, discovering De Morgan's laws and then making use of them when solving logical problems)
- The acquisition of knowledge or skills through experience, study, or by being taught¹
- How agents ought to take actions in an environment so as to maximize some notion of cumulative reward²

3 What are the methods to learn?

4 Objectives

- Minimize the amount of time spent learning
- Increase the rate of learning
 - Minimize inefficient learning
 - * Learning what is already learned
 - Minimize mistakes
- Increase retention and recall rates
- Spend as little time as possible refreshing/preserving old memories

5 Parameters

- Curriculum

6 Learning rate

In the machine learning world, the more examples you have, the better your algorithm becomes at properly predicting the outcome based on its training data. We correlate this improvement in prediction skills with saying that the machine appears to be “more intelligent”. If we were to make a parallel with human learning, would it be appropriate to extrapolate that in order to become more intelligent, one has to learn more examples? Is intelligence the accumulation of examples with the purpose of building a general association-trigger system (you see something that looks like an example of something you’ve seen in the past and it triggers related memories)? According to Pólya’s *How to Solve It*³ method for solving mathematical problems, one of the important steps to problem resolution is the ability to relate a current problem to previously exposed problems that may help us to solve the problem at hand.

7 Learning from example

- Neural networks are often “blamed” for requiring a lot of data while human beings appear to learn (and extrapolate) from very few examples
- How long does it take for a child unfamiliar with some experience such as what a dog/cat is before they can recognize one?
 - Can a child that has never seen the real animal (only has seen them in children books) extrapolate and recognize one when they see one for the first time?

8 Learning according to reinforcement learning

- Greedy policy

¹<https://en.oxforddictionaries.com/definition/learning>

²https://en.wikipedia.org/wiki/Reinforcement_learning

³Pólya, George. *How to Solve It: A New Aspect of Mathematical Method*.

- Some form of exploration to prevent being blind to better alternatives
- An objective/goal must be defined, otherwise learning is done aimlessly

9 Learning from different mediums

9.1 Objectives

- Determine and extract important information

9.2 Books

- Read a chapter
- Write a short summary of what was discussed in the chapter
- Requires the ability to use language

9.3 Videos

- View the video
- Write a short summary of what was discussed in the video
- Can require the ability to use language

10 Types of things to learn

- Facts
- Relations/pathways
- Associations
- Rules

11 How do I learn?

- Read theory
- Observe examples
- Experiment
- Generate questions and answer them
- Organize my understanding
- Teach others to discover flaws in my own understanding

12 Ways to learn

- Segment problems into subproblems and tackle them individually (divide and conquer)
- Start from where you are and slowly work toward your objective (“forward kinetics”)
- Start from your objective and slowly work toward where you are (“reverse kinetics”)

13 Pólya’s method

- Understand the problem
- Obtain a plan
- Carry out the plan
- Look back at the solution

14 Applications

14.1 How would you optimize learning about 100+ topics on Wikipedia?

Relevant information could be extracted and used to create SRS (space repetition system) cards. These cards would be reviewed based on the SRS itself.

The list of topics/articles could be inserted in a software that would build a relation graph between all those topics/articles. Various metrics could be extracted from the articles (article first/last edit dates, size of article, number of incoming/outgoing links, number of editors, article quality, etc.) in an attempt to sort the articles by some “relevancy” order (think Google Page Rank).

Plan the order in which you are going to read the articles. By using the previously mentioned tool, the relation graph could give you an idea of the order you should read the articles. It is generally easier to read about topics in chronological order (from oldest topic to newest topic) as it will explain how understanding has developed over time.

The content of articles could be extracted. With enough advanced NLP tools it would be possible to order the claims/facts such that one claim can induce the next, making it easier to read a collection of articles.

Track what you read. This can be done either by writing down the articles you’ve read or using the watchlist feature of Wikipedia. If you are writing down which articles you’ve read, I suggest it is written in a computer document so that you can verify if you’ve already read the article. Furthermore, I’d write down the date at the time the article was read so that if you want to read the article again later on, you can look at only the changes that were applied since then. However, I strongly recommend simply using Wikipedia’s watchlist system which does this (it does not track when the article was added to the watchlist though).

While reading an article, write down related topics. New knowledge is better remembered when it is well integrated within our graph of knowledge.

15 Instructional theory

- Core ideas for the post-industrial paradigm of instruction
 - Learner centered vs teacher centered
 - Learning by doing vs teacher presenting
 - Attainment based vs time based progress
 - Customized vs standardized
 - Criterion referenced vs norm referenced
 - Collaborative vs individual
 - Enjoyable vs unpleasant

Source: https://en.Wikipedia.org/wiki/Instructional_theory

16 First Principles of Instruction

- Task/Problem-centered - Students learn more when the instruction is centered on relevant real-world tasks or problems, including a series of tasks or problems that progress from simple to complex
- Activation - Students learn more when they are directed to recall prior knowledge
- Demonstration - Students learn more when new knowledge is demonstrated to them in the context of real-world tasks or problems
- Application - Students learn more when they perform real-world tasks or solve real-world problems and receive feedback
- Integration - Students learn more when they are encouraged to integrate their knowledge into their life through reflection, discussion, debate, and/or presentation of new knowledge

Source: https://en.Wikipedia.org/wiki/First_Principles_of_Instruction

17 Andragogy

- Adults must want to learn
- Adults will learn only what they feel they need to learn
- Adults learn by doing
- Adults learning focuses on problem solving
- Experience affects adult learning
- Adults learn best in an informal situation
- Adults want guidance and consideration as equal partners in the process

Source: <https://en.wikipedia.org/wiki/Andragogy>

18 Unsorted

- Learning to talk is not necessarily about associating visual cues with auditory cues as apparently babies born blind can still learn to speak and listen
- Capacity of source location based on sound only? Similar to how the brain can determine the source of a sound due to the small delay of propagation that occurs when reaching each ear canal (external structure)
- Difference between recognition and recollection
- Learn from others
- Learn by comparing yourself with a copy/example you consider to be better (you already need the ability to evaluate what is better/worse than your current abilities)
- The ability to assign a score to a performance/agent function in order to pick the best
- Recognize areas of variability which are potential locations for optimization
- Determine the areas where you are the furthest from doing properly, these should be where you can make the most gains (this requires a distance metric)
- It's important to track the amount of information one is able to absorb, as well as how much is retained over different periods of time
- It is interesting to understand what one is able to understand rapidly and remember for extended periods of time without requiring constant memory maintenance

19 Scratchpad

- Learning is about being practical (making use of the knowledge that was acquired)
 - Either to predict or to suggest
- Based on prior experience, suggest what output to produce
- Construction of a mapping function between inputs and outputs, however complex it needs to be
- If I were to read about learning, what would my goal be?
 - Spend less time on learning (in other words, learn faster, more efficiently) so more time can be spent on producing value

20 See also

- [The Strategic Student Approach for Life-Long Exploration and Learning](#)

21 References

- <https://deepmind.com/blog/alphago-zero-learning-scratch/>

21.1 Self-help oriented

- <https://www.quora.com/What-are-the-best-books-on-rapid-learning-and-meta-learning-to-increase-ones-learning-acquisition-rate>
- <http://jacobmorch.com/metalearning/>