Summary: Steps Toward Artificial Intelligence by Marvin Minsky

12 September 2019 - Paul Meierling (6.844)

Overall summary of the paper

Marvin Minsky discusses the possibility of artificial (general) intelligence (AGI) and what concepts are needed to build a machine that is able to learn more than “only what it is told to do”. He believes by combining several methods we are able to build a machine that like humans can become greater than the sum of its parts. While our brains only consist of relatively simple Neurons we are still able to develop complex thoughts, a potential machine could be able to do the same.

Structure of the argument

Minsky argues that the combination of the following 5 components are critical to achieving AGI

I) Search: is for problems that are well defined (e.g. we can see if a solution is acceptable or not). Since problems are well defined we are able to brute-force an optimal solution by trial-and-error. We can improve on trial and error by using hill-climbing, which means moving in the steepest direction towards the solution. The direction can be calculated by partial derivatives. However, hill climbing may run into problems of local maxima as well as Mesa Phoneomenon (no changes for small steps).

II) Pattern recognition: describes the process of selecting the appropriate method for a given problem. One example are using prototype derived patterns (e.g. using original dies of printed letters on a page) to compare observed objects and classify them. We can use property lists in combination with normalization to better determine if an observed objects belongs to a given prototype. Furthermore we can combine properties (e.g. using bayes nets) to classes to better distinguish objects. Minsky argues that a recursive model for pattern recognition is needed (e.g. to determine multiple objects in a given scene) instead of purley statiscally one (like bayes nets).

III) Learning Systems: describe the process of using previous experience to solve new problems more efficiently. Minsky introduces reinforcement learning to show that a trainer and a reinforcement machine can reward certain behavior, which enables the machine to solve problems quicker in the future. Furthermore, this system can also be used to plan problems and generate predictions. Finally, he argues that to leverage such a system we can not use a complete decision tree (e.g. win, lose, draw of a chess match) but instead need deconstruct the problem recursively.

IV) Problem solving and Planning: is considering how we can use problem reduction and then select the right subproblems to solve. Two main factors are relevant when considering subproblem selection: centrality and difficulty. Minsky introduces the inheritance method, which enables us to more efficiently select subproblems without having to look at the entire structure everytime we update the problem.

V) Induction and Models: to have a machine become truly intelligent it needs to be able to use induction to predict new outcomes as well as generate more examples given existing rules. Furthermore the machine will also have to reason about itself, which requires an examination of its internal methods

Implications for State of the Art AI research

Overall the concepts discussed in the paper are still amazingly accurate even over 50 years in the future. Interestingly Minsky took a different direction than what a lot of current state of the technology is, namely he was focused on developing artificial general intelligence instead of neural networks that solve particular tasks (e.g. image recognition really well). Nevertheless his key concepts are very relevant for current state of the art research of neural nets: e.g. hill climbing is replaced by gradient descent, recursive problem solving for different objects is used by RNNs to detect multiple objects and many other concepts did stay the same.