Jacob Devasier

4/16/2021

Review of Intrinsic Motivation Systems for Autonomous Mental Development

This paper confronts the question of whether or not a machine can be given the ability to be curious and to explore different activities based on intrinsic motivation alone. Human children switch between playing with different toys very often due to this intrinsic motivation and this paper discusses how such an idea can be embedded into a machine. The positive implications of such a feature in a machine are abundant; primarily, this machine would be able to learn new things by itself through trial and error much like children do, additionally, it would learn to understand a wide domain of information by itself because of the explorative nature of such a machine.

This paper introduces the Intelligent Adaptive Curiosity, an intrinsic motivation system which, much like human children, pushes the machine towards environments that it can learn a lot from. As the authors mentioned, the robot focuses on environments which are not too predictable, but still have some uncertainty to learn from. The paper mentions four aspects of human learning that I think are insightful: development should be progressive, incremental, autonomous, and active. Meaning that the way a machine should learn is to start from small tasks and slowly abstract and build onto those tasks, to do this the machine should take its own steps toward such a goal. But, while the goal of the paper is important, we still run into the problem of relying on reward mechanisms to get there. Ultimately, we should aim to move away from reward systems as they tend to cause the machine to work hard at getting a high reward rather than understanding what is happening.