Adjusting Autonomy by Introspection

Henry Hexmoor, Matthew Lafary, and Michael Trosen

University of North Dakota
Grand Forks, ND 58202

Abstract

In this paper we discuss autonomy and agents with abilities for changing their level of autonomy. We suggest enriching agents with introspective abilities so their choice of behavior can correspond with their current level of autonomy.

1. Autonomy

Let's consider a device as a complex machine that appears capable of tasks commonly performed by intelligent organisms. An automobile equipped with cruise-control, road-sensitive traction, and self-inflating tires is such a device. Such devices receive input from the environment and follow an algorithm provided by the device designer to produce an output. In general, devices cannot tell how well they perform. Furthermore, devices may only manipulate a fixed ontology of their surrounding. I.e., they represent things in the world in a fixed way. Many robotic applications would qualify them as devices. To the extent, if a system can be aware of its performance and can improve its pre-programmed ability to interact with its surrounding, or can alter its ontology it is autonomous. Autonomy is self-governance over output. A system such as smart chess-playing programs can be intelligent without being autonomous. This notion of autonomy is desirable in systems such as autonomous space applications needed in long-duration space missions.

Perhaps it can be argued that agents who make more decisions, can deeply introspect about their decisions, or in some ways possess higher cognitive abilities have a larger domain over which they are autonomous. However, in this paper we are not concerned about quality of autonomy but with an agent's ability to adjust its level of autonomy. In the remainder of this paper, we will not be interested in measuring an agent's autonomy for purposes of inter-agent comparison but focus on intra-agent autonomy.

Assume multiple agents share responsibilities in an environment. They may use a set of shared resources. Their tasks may overlap or even be unclear. The changes in the environment may make it difficult to predict events or action consequences. In such an environment, agents need to be ready to act independently, cooperatively, and with varied level of responsibility. In this scenario we need agents with adjustible level of autonomy.

2. Autonomy Change

Very often, an agent interacting with other agents changes its level of autonomy to correspond to its rank. When it runs into a higher ranking agent, an agent that is fully autonomous will lower its autonomy and take orders. Although authorities might be externally imposed, it is still largely the agent's introspective abilities that determine the agent's true level of autonomy. Methods of action selection and resource usage are changed for an agent by introspection. We assume following authority necessitates a proper level of autonomy. As a first approximation we consider the following ranks and corresponding autonomy levels: fully autonomous, boss, cooperative, underling, instructible, and remote control.

Apart from externally imposed authority among agents, an agent may change its autonomy due to personal choices rather than responding to the environment. An agent might lower its autonomy if it finds that it is no longer meeting its goals, or decides it will not be able to meet them and there is another agent available to take control. Take for instance, a high ranking agent that is controlling the pursuit of a criminal through a warehouse. The agent is using vision as its major source of input, and the criminal realizes this, so turns the lights out. The high ranking agent recognizes that it can no longer successfully track the criminal, so it requests that one of the underlings with a low-light camera take over the pursuit so as to not lose the criminal. An agent may also want to increase its autonomy. Take the same example of the high ranking agent that chose to become an underling in the dark, if it gets back into where it can see, it might want to regain its superior role. If a high ranking agent has underlings which are not accomplishing their goals in a timely manner, it may relieve them, and become fully autonomous and do all tasks itself (it may decide to follow the old adage that *if you want it done right, do it yourself*).

A more mechanical method of autonomy change is by treating the agent as a system that is service of a human user. This paradigm is more consistent with autonomous control systems being developed by NASA. In such systems, the agent will treat the human user having roles like the Boss and the "God". Perhaps issues with these systems are responsiveness and preserving goals while taking human guidance.

3. Levels of Autonomy Mirroring Authority Levels

The fully autonomous agent makes decisions and reacts to the environment without much input from an outside source. An example would be a robot that has determined that its battery level is low, so it must now proceed to its docking station to recharge. On the way to the recharge station, the robot may run into problems, such as obstacles, which it must overcome to be successful. The success of the robot reaching its internal goals can be measured and adjusted for efficiency.

A boss is an agent that has control over other agents as well as over itself. A boss would use the underling agents to complete it's tasks. The task can either be given to it by an outside source or enabled by a reaction to the environment. The reason that this agent is less autonomous than the fully autonomous agent is because it relies on its underlings to do their tasks and does not have complete control over the underlings. For example, the boss robot might tell an underling to go and close a door, it is relying on the underling to be able to figure out how to get to the door, and know the procedure for closing it. The boss does not have complete control over the actions that the underling takes, it gives the underling a goal rather than controlling it remotely, this results in the boss not having complete control over everything that it is doing.

The third level of agent autonomy is cooperative agents. This differs from the boss agent in that each agent communicates with the others and as a whole they come up with a coarse of action. Subsequently, a division of tasks determines task assignment.

An underling simply receives all task information from its boss. The underling will start to execute the command to the best of its ability, if it gets into a situation it doesn't know how to deal with, it will ask the boss for appropriate instructions on how to proceed. This continues until the underling has successfully completed its task or is instructed to complete a different task from its boss. This agent still has control over the details of its movements, but not the goals that it must accomplish. For example, it may be told to fetch a beer from a fridge. The underling will figure out how to go out of the room, and follow the corridor down to the kitchen and get a beer from the beer-o-matic and then return. It has control over all of its actions to accomplish its goal, but the goal is provided by a boss. Without the provided goal, it would just sit there waiting for a goal.

The next lower level of autonomy is the instructible agent. This agent is like the underling but it does not have complete control over what it is going to do. It has higher autonomy than a simple remote control agent because it may have a reactive layer which is still functioning. It also takes reasonably high level commands, like 'follow the wall on your left,' or 'go to that point'.

The remote control agent is the one with the least autonomy. This agent does not make any decisions on its own. For instance, in order to get it to go along a wall, another agent (or other outside source) would need to command it to move forward and then keep adjusting its right and left heading.

4. Adjustible Autonomy in the Warehouse

Consider the following scenario. Assume three agents that need to guard a warehouse. When an intruder enters, each agent is capable of doing these three things: closing the doors to shut the intruder in, shoot the intruder, and call the police.

Let's consider agents in our warehouse scenario to have preferences for their three actions. This will guide agent choice of action. Preference might be born out of agent's experiential tendencies, biases from skills or habits, or even an analysis of circumstances that allows them to generate priorities. We will not be concerned with the origin of preferences in this paper.

Furthermore, let's assume each agent has a different level of competence in its ability to carry out actions.

In the following table, P and C refer to priority and capability respectively. The numbers refer to the individual agents. The capabilities reflect how capable each agent is for doing the action at this particular moment in time. E.G., this does not mean how well it shoots, but how well it can shoot this intruder, not how well it closes a door, but if it can get to the door and close it before the intruder gets out. We assume that each agent can only do one thing.

Act	**P0**	**C0**	**P1**	**C1**	**P2**	**C2**
Door	**8**	**2**	**10**	**1**	**2**	**10**
Shoot	**5**	**5**	**8**	**6**	**2**	**0**
Police	**1**	**6**	**6**	**4**	**10**	**5**

Fully autonomous: Since the agent is fully autonomous it doesn't care what the other agents are doing, thus it needs to find the best combination of preference and capability. The easiest way would be taking the product of the P and Cs, the products would be 16, 25, 6 so it would decide to shoot the intruder.

Boss: Assume agent 0 is the boss, so it uses its preference to decide what will be done, then finds out how capable each of the other agents is of doing each action and assigns them based on who is most capable of doing each job. The easiest way of doing this is to take the product of the boss' priority and the capability of the best robot. The three highest products are of Agent 2 closing the door, Agent 1 shooting the intruder, and Agent 0 calling the police. So the boss would tell Agents 2 and 1 what to do, and then go to call the police.

Cooperative: At this level of autonomy the agents would want to come to a unified decision. One way of doing this would be to average the priorities, then use the average priority to find products and come up with the best combinations of priorities and capabilities.

Underling: In this mode the agent would just wait for the boss to tell it which action to do.

Instructible and remote control: In these levels of autonomy, agents would wait for low level commands.

5. Impact of Resources on Adjustible Autonomy

Preference and capability are only two of the factors that can be used in determining autonomy. Resource management is another key factor for consideration. Here we will limit our attention to reusable resources. We posit that an agent's level of autonomy affects the agent's method of resource selection and usage.

Consider the following resource parameters: overall availability of resource or resource scarcity (RA), suitability of resource for each action (RS), cost of using a resource with each instance of action (RC).

Let's assume a fully autonomous agent with a function that considers a fixed degree for resource parameters in its resource selection. For instance, the agent might value resource scarcity, suitability, and cost equally. However, when the agent's autonomy is changed to Boss, the agent will consider scarcity more heavily than before for delegation but will consider cost more heavily in its own choice of resources. If the agent's autonomy is changed to Underling, the agent will consider scarcity more heavily than the other factors. Also, the Boss agent may instruct an Underling to use a resource that it may not choose itself.

In the warehouse example, let's consider our agents had two weapons, one is an electrical stunner, and the other is a gun. Electrical charge might be in limited supply but highly desirable and using it once uses a lot of charge (low RA, high RS, high RC). Let's consider bullets weren't as scarce but less desirable and single bullet is highly effective (high RA, low RS, low RC).

Only upon Boss command, an agent may use its stunner but the agent might be more ready to use the gun.

6. Summary and Discussion

We suggested that changes in an agent's autonomy are often due to an agent's rank when interacting with other agents. We pointed out that changes in autonomy are sometimes due to an agents personal biases, choices, and goals.

We believe adjustment in autonomy can be implemented by including information that augments an agent's knowledge of its actions. In this paper we considered the augmented knowledge to include an agent's ability in performing tasks, priorities and tendencies for action choice, and resource parameters. A rich set of knowledge about actions is needed for introspection and behaving at different levels of autonomy.

Since it is easier to enrich reactive rules with parameters than to devise inference mechanisms in plans, we believe a reactive implementation is more amenable for adjustible autonomy.