

Managing Machine's Motivations

Janusz Starzyk^{1,2}, James Graham¹, and Leszek Puzio²

¹ School of Electrical Engineering and Computer Science,
Ohio University, Athens, OH, USA
{jg193404, starzykj}@ohio.edu

² University of Information Technology and Management, Rzeszow, Poland
lpuzio@wsiz.rzeszow.pl

Abstract. This paper presents concepts for the development and management of motivations in learning agents, which are critical for motivated learning. We suggest that an agent must be equipped with a mechanism referred to as a *nonspecific formative process* to trigger higher level motivations. Resource and action related motivations are discussed as examples of implementing such process in a virtual world learning scenario.

Keywords: Motivated learning, cognitive agents, reinforcement learning, goal creation.

1 Introduction

In this paper, we examine various ways to establish a motivation mechanism for the agent to develop. This is an extension of our earlier work on creation of goals by an autonomous learning agent [1].

Autonomous learning agents are needed to establish a path towards intelligent machines. Today, these agents find many applications in industry such as: robotics, video games, remote sensing, image recognition, quality control, warfare, assisting humans, entertainment, etc. and their importance is growing steadily. There are several concepts for organizing motivational systems. One, introduced by Pfeifer [2], shows motivation as a result of the developmental process. Another concept, based on external reward signals, is known as reinforcement learning (RL). It was initialized by the work of Sutton and Barto [3], followed by Brooks [4], Pfeifer [5], Schmidhuber [6] and many others. The intrinsic motivation system based on artificial curiosity was proposed by Oudeyer [7].

Merrick pointed out that RL robots do not have internal drives to maintain their resources within an acceptable range [8]. To address this problem a motivated learning (ML) system was proposed to allow the agent to develop its own motivations and goals [9]. Merrick introduced motivated reinforcement learning (MRL) and used motivated exploration in video games [8]. Motivated learning based on the need for resources was used to develop a coordinated learning strategy in a multi-stage stochastic game [10].

Motivated learning showed promise in supporting the development of intelligent systems. But a nagging question arises; where do the motivations come from? How should a system be motivated to develop? What are the conditions for motivations to reach higher levels of abstraction and sophistication in an agent's interaction with the environment? This paper tries to answer some of these questions and proposes a mechanism for creating higher level motivations fundamental to agent's mental development.

2 Basic Concepts in Motivated Learning

A ML agent has predefined *needs* (for instance need for shelter, food, or energy level). Agent *motivations* are to satisfy its needs. Thus, in order to introduce new motivations, an agent must develop new needs. A basic mechanism to create new needs for resources was described in [9]. This was extended in recent work to a mechanism that is used to create needs related to actions by other agents.

In order to clarify our discussion let us define some critical concepts used in ML.

Definitions

A *primitive pain* is associated with each predefined need and measures how far the agent is from satisfying its need. The pain is larger if the degree of satisfaction of a need is lower. For example the following function can measure resource related pains:

$$P_i = w_i * \frac{\varepsilon + R_d(s_i)}{\varepsilon + R_c(s_i)} \quad (1)$$

where R_d is a desired level of needed resource s_i , w_i is a weight that increases with the increased importance of resource i , R_c is the current level, and ε is a small positive number to prevent numerical overflow. *Pain reduction* in ML is equivalent to a *reward* in RL.

When an agent is introduced to a new environment, it does not know how to satisfy its needs and must experiment with various resources and available actions until one of its needs is reduced. A new *abstract need* is created for the resource used to reduce the primitive need. The agent will reduce the new abstract need in the same manner it reduced its primitive needs by trying various actions. An *Abstract pain* measures how far the agent is from satisfying its abstract need and is computed based on (1).

Once introduced, an abstract need can be satisfied by acting on another resource. This leads to another higher level abstract need (for this new resource) and related abstract pain. This simple mechanism allows the agent to build a potentially complex "network of needs" and such mechanism is a foundation of *resource based motivations*.

3 Developing Motivations

In this paper, we ask questions as to what other cognitive mechanisms should be considered for building motivations, and to allow a ML agent to develop even more abstract motivations, such as a motivations to gain love, friendship, recognition in society, self

esteem, or the human tendency to actualize itself as fully as possible. The last one, known as self-actualization, is considered by psychologists as the final level of psychological development that can be achieved when all basic and other mental needs are satisfied. According to Goldstein [11] self-actualization is not an ultimate objective, but rather a process driven by the tendency to actualize all self capacities, the entire potential available at any given moment and in given conditions.

The question is as follows: can all these levels of abstract motivation be derived from the *fundamental mechanisms* that create needs in ML? Are they necessary consequences of mental development? Are they characteristic of successful intelligent behavior, whether this behavior is conducted by a human, an animal, or a machine? Can these abstract motivations be achieved before the lower level needs are satisfied and if not, are they symptomatic to the friendliness, specific level of sophistication, and reciprocal support from the environment in which such growth is possible and useful?

Reasoning along this line we may ask: is the state of the environment related to the motivational levels and the level of mental development of its learning agents? We may also ask a more direct question: is the state of the environment a result of and measure of the state of development of the most successful individuals that inhabit it? These are not existential questions, although they may relate to such questions. The aim of asking these questions is to specify the necessary environmental conditions for developmental robots.

We have affirmative examples confirming mutual dependence of the environment and its inhabitants in humans. The more advanced the state of the environment in terms of technological support, tools efficiency, and ease of satisfying basic human needs; the more capable, more motivated and better developed individuals become, even when their brains do not change much.

There is a difference of opinion in psychology as to whether higher order needs and motivations are a driving force for human development or whether they are a prespecified ideal hierarchy of motivations and needs. This difference can be rephrased in a question: do we develop our needs as we grow or is there a given hierarchy that is fixed and specified independently of individual capacities to reach them that perhaps only few of us can reach, living ordinary lives. This question is important in view of developmental learning in machines where we do not put limits on the development or needs, but try to justify higher levels by what works and what makes the machine more successful in its interaction with the environment. A related and equally important question is about the basis of these motivations. Is there a mechanism that creates them and if so what is it?

Psychologists studying successful people like Albert Einstein and Charles Darwin found that these people were focused on finding solutions to societal problems rather than to their own personal problems, were open minded, had a strong sense of self, valued life and human dignity, and had a small group of close friends. This helped them to succeed where others could not. But this observation of personal traits in people who succeeded when extended to a general population may miss an important link in the developmental process, the one that justifies why such motivations are useful for growth of individuals and society. Instead some of the highest level

motivations are compared to the norms of morality in a given society. Are these norms invariant to the level of growth and sophistication of social interactions? Most likely not. We know that such norms change as society changes, as social consciousness and understanding of human needs and behavior grows. Thus, such norms or prespecified hierarchy of motivations cannot be considered as a constant part of the mechanism responsible for the developmental process.

Alderfer's *ERG theory* [12] stresses *existence*, *relatedness*, and *growth*, where *existence* focuses on material existence requirements, *relatedness* focuses on the need to relate and maintain social interactions, and *growth* focuses on personal development. But are they the mechanisms that resulted in current social organization or rather they reflect it? In this search for cause and effect we cannot accept existing norms as a foundation for the development of higher level motivations. Norms were not given but were derived from the development of humans and accepted by society. Subsequently, they became a part of the current environment that influences further growth. However, further growth must come from within individual's drives which are modified but not predetermined by the current social order.

Ryan and Deci [13] who promote self-determination theory focus on three elements: humans master their drives and emotions, have a tendency to grow and develop, and optimal development does not happen automatically. This theory is hard to accept as a drive for development since it does not explain what a driver for such motivations is or why they appeared in the first place. They do not provide a causal relationship in behavioral development that would yield these kinds of drives.

The Need for Achievement theory [14] of motivational growth stresses social motives like dominance. According to this theory people will take calculated risks, establish attainable goals, and fear failure. They also want to be praised for their accomplishments and receive feedback from others. Such and other theories debated by psychologists focus on explaining the human motivational system, the way it is and the way expresses itself, but they do not answer the important question from the developmental point of view. How did the human motivational system develop?

What is important for building intelligent machines is to describe the *nonspecific formative processes* (NFP) that the agent may use to develop motivations at a certain abstraction level. Nonspecific means that these processes are not prespecified to obtain certain motivations (like the need for food or shelter) but rather are used to develop motivations that help to solve a group of existential problems.

3.1 Nonspecific Formative Processes

An example of a NFP process is the way an agent creates goals and motivations in order to acquire resources needed for its survival as discussed in Section 2. Such a NFP process includes evaluation of changes in the environment and creates motivations to collect or to avoid certain resources in the environment. The resources are defined as objects in the environment that do not initiate actions by themselves. Rather they can be used by the agent to satisfy its needs (for instance food or water) or to be avoided since they may harm the agent (for instance poisons or toxic substances).

Another example is to equip the agent with an NFP process to evaluate actions by other agents (non agent characters or NACs) and learn how to encourage or discourage such actions. Such a NFP evaluation process is nonspecific because it makes no assumption as to what kind of agent performs the action, what action it performs, or whether or not this action has any effect on the ML agent. Learning how to discourage or encourage an action is also nonspecific as the ML agent has no preconceived knowledge whether or not the action performed by NAC is beneficial or harmful.

The need for action to discourage NAC action (that increases ML agent pain) will be a result of the NAC action pain signal created by the agent. The NAC action pain can be computed using:

$$B = \gamma * \frac{-\delta_a * \overline{P(y)}}{1 - \delta_a - 2 * L(y) + \epsilon} \quad (2)$$

where $L(y)$ is a likelihood that NAC agent will act on the ML agent, y identifies the NAC action, $\overline{P(y)}$ is the average pain to ML agent caused in the past by a NAC action, and δ_a is desirability of such action. δ_a is 1 when the action by NAC is desired and δ_a is -1 when it is not. The value ϵ is a small positive number to prevent numerical overflow, and $\gamma > 0$ regulates how quickly pain increases.

Another example will be an NFP process to evaluate actions by other intelligent agents that can learn to modify their actions according to the response of ML agent. Since no such mechanism has been developed yet, it is hard to speculate how general it can be and what it will involve.

It is this category of ML agent mechanisms that need to be investigated, designed and implemented in order to provide the agent with cognitive support to reach higher levels of motivational development as described by Maslow [15]. This paper addresses some of the issues and poses open questions to discuss scenarios, current developmental skills, levels of sophistication of the environment in which such development is useful or possible.

Developmental skills determine the internal state of the agent ready for further growth and development of its motivations and mental abilities. Thus, some motivations cannot be developed before others and they will naturally form a hierarchy of motivations and related skills. The developmental process is a function of itself and depends on its history as well as the agent's ability to accommodate new challenges in the environment. For instance, before the agent learns how to interact with another intelligent agent, it must possess skills necessary to respond to NAC actions. Similarly, in order to learn how to respond to NAC actions that are damaging an agent's resources, the ML agent must first learn the values of such resources, and have motivations to collect or protect them. Thus, it must have developed a resource NFP process before it can develop a NAC NFP process and related motivations.

To some degree, the structure of motivational drives develops gradually and can be compared to the genetic development of species. The difference is that it concerns only a single individual rather than generations of individuals. The similarity between these two processes is that both mental development as well as evolutionary development are incremental and depend on the current state of the developmental process. Similarity also lies in the randomness of the incremental changes that take place.

In genetic development it is survival of the fittest as a way to adapt to changes in the environment (by random mutations and crossovers), while in mental development it is the opportunity to learn an interesting trait to improve the way that an individual works (by random trial and error).

3.2 Social Agent

A learned response to NAC action pain (2) may not work well when the NAC is intelligent. An intelligent NAC agent may change its strategy, such that an initially successful action may fail. Another way in which an intelligent NAC may respond is to fight back, causing additional pain that was not there before. This may be a primitive pain inflicted on the ML agent, so there is no need to use bias estimation. To avoid such pain, the ML agent must learn that its action caused the NAC's response and modify its behavior. If the ML agent observes that total pain from all pain sources increased as a result of its action, the action should be avoided.

So, the question is what recourse does the ML agent have when an intelligent NAC destroys its resources or otherwise inflicts pain on the ML agent? No action against the NAC will be painful. An action that causes a response from NAC may be painful, but if the total pain increase is smaller than when no action is taken, such an action is acceptable. For instance, the ML agent may fight the NAC and even when it suffers some pain, the overall pain reduction may be greater than the pain inflicted by the fight, particularly if the ML agent is stronger than the NAC and makes it go away.

If the NAC is stronger, the pain inflicted by it may exceed the pain reduction and such action should be avoided. We then see a typical flight situation, where the agent will suffer the pain without fighting back.

But what if the NAC can fight back with equal resolve and inflict equal pain to the agent? Both agents will suffer a significant amount of pain without any benefit. For instance when they fight over the food supply, neither will get food, and they suffer the pain inflicted by their opponent. The obvious solution to this situation will be sharing the resources, rather than fighting for them. Such a decision, when neither agent fights for the resource will be "agreed" upon if the two agents know each other's ability to inflict significant pain and decide not to fight. Such an estimate can be obtained using (2) and evaluating the likelihood of aggression by the NAC directed against the ML agent.

The only modification we need in this case is to replace $L(y)$ by a likelihood estimate based on learning the NAC's behavior. This can be accomplished using RL which will learn the likelihood of NAC actions and predict the negative reward in response to the ML agent's action.

Similar analysis can be performed in the case where intelligent NAC action is desired. The likelihood of the NAC's action is learned using RL. If this likelihood is low, the ML agent will suffer the pain described by (2) proportional to the amount of reward (pain reduction) that the NAC action can bring.

3.3 Directing the Machine

A good way to accelerate learning is to use help from a teacher. The problem is how to introduce a teacher within the framework of ML? The simplest (but not the most desirable) way is to assume that any instruction given by the teacher motivates the robot to complete this instruction to a satisfaction of a teacher. We can use voice commands as sensory input that must be recognized and interpreted by the robot. We assume that a spoken command generates a primitive pain signal that is removed by the teacher once its command was correctly implemented by the robot. The level of the pain signal is between threshold and the maximum pain and depends on the articulation. A sharp and angry command will carry higher pain than a soft and gentle instruction. So the agent must learn not only to recognize the command but also to determine its emotional content. The teacher will reward the agent by pressing a prespecified key. A more developed agent may learn to use verbal praise for a reward.

4 Simulation Scenario

To illustrate the process of developing motivations, we designed a learning environment for the ML agent. In our simulation scenario, presented in Fig.1, we have five primitive pains: Sweet-tooth, Bee stings, Hunger, Thirst, Curiosity.

In Fig. 1 all resources are represented by ovals and actions by rectangles. Acting on a resource that inhibits a pain is indicated by inhibitory links (with solid black circle). An excitatory link (with arrow) triggers a NAC action pain or a NAC appearance. In Fig. 1, only simplified relations are shown to avoid clutter. Each resource symbol can be interpreted as an inhibitory interaction on the amount of resource and the lack of resource pain as shown in Fig. 2. Similarly, a related action on the resource causes a small inhibitory feedback link to the resource being utilized as shown on Fig.2 b from the action 'Plant Flowers' to the resource 'Flowers'.

As illustrated in Fig. 2, the top inhibitory link inhibits the abstract pain 'Lack of Flowers'. When the 'Lack of Flowers' is inhibited the Flowers resource is automatically activated. A resource is restored through proper action by the agent, in this case by buying flowers. As the resource is used up by frequent action (planting flowers) the inhibition from Flowers to the Lack of Flowers is weaker and the abstract pain 'Lack of Flowers' increases. In Fig. 2, the resource nodes and lack of resource nodes are automatically activated unless inhibited. The forward arrow links are excitatory, so the Flowers resource activates the non-agent character (bees).

Curiosity pain drives agent to explore the environment, *i.e.*, to learn useful actions on resources. Curiosity pain stays high until the agent learns all valid actions. Other primitive pains like Thirst, Hunger, Bee stings, Sweet-tooth, may be satisfied by various actions. For example, the Thirst pain can be satisfied by the 'Drink water from Cup' action. Hunger, Thirst, and Sweet-tooth pains increase with time, while the 'Bee stings' pain increases when Bees are around. We call those actions 'useful motor-sensor pairs', because the agent can satisfy one of its needs performing a motor action on a resource.

Resource related abstract pains are introduced as soon as a resource is determined useful for satisfying a primitive pain. For instance, as soon as the agent learns that it can satisfy its thirst pain by drinking water from the cup, it develops an abstract need to have water in the cup. Some resources may be used for multiple actions, like Money can be used to: Buy Food, Buy Flowers, or Buy Cigars. Similarly, a resource pain can be satisfied by more than one action.

To illustrate NAC action pains, we introduce non-agent characters like a Bug and Bees. Those characters interact with the agent or with the resources. Bees produce Honey from Flowers, the Bug eats Food, and Bees can also sting the agent. We wanted to create characters which do useful or harmful actions for the agent. In this example, we have two NACs. The Bug only engages in harmful action by stealing food, which cause the agent ‘Lack of Food’ pain. Bees do both useful and harmful actions. They produce honey and they sting the agent. In this example, we show the ability of our model to easily accommodate more characters, which can perform both useful and harmful actions. All useful motor-sensor pairs and their outcomes are presented in Table 1.

Table 1. List of Resources, useful Resource-Motor pairs and their outcome

Motor action	Resource name	Agent’s pains	Outcome		
			Increase	Decrease	Pain reduced
Eat food from	Bowl	Lack of Bowls		Bowls	Hunger
Drink water from	Cup	Lack of Cups		Cups	Thirst
Eat honey from	Honeycomb	Lack of Honeycombs		Honeycombs	Sweet tooth
Smoke	Cigar	Lack of Cigars		Cigars	Bee sting
Take food from	Fridge	Lack of Fridges	Bowls	Fridges	Lack of Bowls
Pour water from	Bucket	Lack of Buckets	Cups	Buckets	Lack of Cups
Plant	Flowers	Lack of Flowers	Honeycombs	Flowers	Lack of Honeycombs
Buy food with	Money	Lack of Money	Fridges	Money	Lack of Fridges
Pull water from	Well	-	Buckets	-	Lack of Buckets
Buy flowers with	Money	Lack of Money	Flowers	Money	Lack of Flowers
Buy cigars with	Money	Lack of Money	Cigars	Money	Lack of Cigars
Work for money with	Tools	Lack of Tools	Money	Tools	Lack of Money
Study for job with	Book	Lack of Books	Tools	Books	Lack of Tools
Play for joy with	Beach ball		Books		Lack of Books
Kick	Bug	-		Likelihood	Bug eating food
		Hunger -	primitive pain		
		Thirst -	primitive pain		
		Sweet tooth -	primitive pain		
		Bee sting -	primitive pain		
Any		Curiosity -	primitive pain		Curiosity

4.1 Simulation Implementation

Based on this scenario, we created a simulation in NeoAxis. The NeoAxis 3D Engine is an integrated development environment for 3D projects of any type and complexity. The environment is intended for use in such areas as the creation of video games, development of simulators, and development of virtual reality and visualization software.

It includes a full set of tools for fast and logical development of modern 3D projects. We created 3D models of resources and characters, and animations of characters in Autodesk 3ds Max. Fig. 3 presents sample resources and character's models. Some animations were captured by Kinect, which provides also full-body 3D motion capture, facial recognition and voice recognition capabilities. The agent logic was written in C++ with use of the Boost library. The simulation environment was implemented in C#.



Fig. 3. a) Resources: apple, ball, banana, bucket, bowl, cup, b) Characters: agent, bug, bee

Based on the presented simulation scenario, we created a simulation map with all resources and characters mentioned in Fig. 1. During creation of the simulation environment, we set the desired level for each resource. We implemented NACs and their actions (both types, i.e. desired and undesired). We also implemented agent actions on resources from Table 1. Fig. 4 presents such a map with resources and NACs.



Fig. 4. Simulation map

Multiple simulation tests proved that the ML agent was able to use the nonspecific formative processes to develop resource and NAC action based motivations, and was able to successfully learn how to manage its needs in a dynamic environment. To our knowledge no other cognitive agent can learn how to manage its resources and learn

how to respond to other characters' attacks. We challenge anyone in the machine learning community to try a simpler version of such scenario - Autonomous Learning Challenge - described on <http://ncn.wsiz.rzeszow.pl/?p=39>.

The simulation window presented in Fig. 5 shows: in the left top corner - a list of resources available to the agent and their quantities; in the top center - the current task performed by the agent; at top right - a list of abstract and primitive pains and their levels; bottom left - the state of agent's memory.



Fig. 5. Simulation window with a list of resources, current task, pains, and memory. a) the agent starts its learning process, b) the agent learned all actions.

When the simulation starts, the agent may interact only with three resources (Bowl, Mug, Honey) to satisfy its primitive pains (Hunger, Thirst, Sweet-tooth, Curiosity). This situation is presented in Fig. 5 a). The memory window in this screenshot is almost grey, indicating that the agent has learned very little. After some time in simulation, the agent's knowledge increases. It introduces new abstract pains, and discovers new applications for resources. As a result, the agent is able to keep its pain under control and keep resources at satisfied levels. Its memory indicates that it learned how to maintain resources and control the NAC related pains. After satisfying all pains (i.e., keeping their levels below threshold), the agent goes to rest on a mattress. Multiple tests proved that in less than half an hour of real time simulation, the agent was able to learn all useful actions and keep its pains under control.

5 Conclusions

This paper examines how motivations in a motivated agent are formed and how they can become increasingly abstract. We examined our earlier work and the recent changes we have made to introduce actions by non-agent characters (NACs), and determined how this work could be extended. The agent must learn increasingly abstract behaviour interacting with its environment and communicating with other agents. We posit that this can be done via the use of nonspecific formative processes by giving the agent the ability to reason about the cause and effect of its actions on

itself, the environment, and other agents. We also suggest that higher level motivations can only develop when the environment is sophisticated enough to stimulate and support further growth of the agent's mental powers. Brooks stated 20 years ago that intelligence cannot develop without environment. We append this by stating that the development is mutually dependent. Developing agents change their environment, and this change is necessary for further growth in the complexity of agents' motivations.

Acknowledgements. This research was supported by The National Science Centre, grant No. 2011/03/B/ST7/02518.

References

1. Starzyk, J.A., Graham, J.T., Raif, P., Tan, A.-H.: Motivated Learning for Autonomous Robots Development. *Cognitive Science Research* 14(1), 10(16), 10–25 (2012)
2. Pfeifer, R., Bongard, J.C.: *How the Body Shapes the Way We Think: A New View of Intelligence*. The MIT Press, Bradford Books (2007)
3. Sutton, R., Barto, A.G.: *Reinforcement Learning: An Introduction*. MIT Press (1998)
4. Brooks, R.A.: Intelligence without representation. *Artificial Intelligence* 47, 139–159 (1991)
5. Pfeifer, R., Scheier, C.: *Understanding Intelligence*. MIT Press, Cambridge (1999)
6. Schmidhuber, J.: Curious model-building control systems. In: *Proc. Int. Joint Conf. Neural Networks*, Singapore, pp. 1458–1463 (1991)
7. Oudeyer, P.-Y., Kaplan, F., Hafner, V.: Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Trans. on Evolutionary Computation* 11(2), 265–286 (2007)
8. Merrick, K.: A Comparative Study of Value Systems for Self-Motivated Exploration and Learning by Robots. *IEEE Transactions on Autonomous Mental Development* 2(2), 119–131 (2010)
9. Starzyk, J.A.: Motivation in Embodied Intelligence. In: *Frontiers in Robotics, Automation and Control*, pp. 83–110. I-Tech Education and Publishing (October 2008)
10. Teng, T.-H., Tan, A.-H., Teow, L.-N., Starzyk, J.A.: Motivated Learning for Coordinated Multi-Agent Reinforcement Learning in Multi-Stage Game. In: *13th Int. Conf. on Autonomous Agents and Multiagent Systems*, Paris, France, May 5–9 (2014)
11. Goldstein, K.: *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man*. Beacon Press (1963)
12. ALD72, C.: Alderfer, Existence, relatedness and growth. Free Press (1972)
13. Ryan, R., Deci, E.L.: Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Cont. Educational Psychology* 25(1), 54–67 (2000)
14. McClelland, D., Atkinson, J.W., Clark, R.A.: *The Achievement Motive*. Literary licensing (2012)
15. Maslow, A.: *Motivation and personality*. Pearson (1997)