Curiosity – A Base for Robots Learning and Development

Elena Vlahu-Gjorgievska[1], Nikola Rendevski[2], Blagoj Ristevski[3]

Abstract - **The curiosity is drive that push individuals for learning and self-development. This is why scientists see curiosity as field for develop mental robotics.**

In this article we will do retrospective about robot curiosity and what is achieved in recent years.

We will see curiosity drive as investigatory and manipulator on one hand and as exploratory on the other. There will be arguments about achievements of intrinsic motivation system as basic issue for robots learning (in which robot have to focus on situations which are neither too predictable or familiar, nor too unpredictable or situations where nothing can be learnt). Also there will be few words about the idea of reaching a learning goal by composing multiply simple machine learning methods which is basic idea of bootstrap learning.

As a conclusion, everything points that curiosity is a drive which pushes the robot towards situations in which it maximizes its learning progress.

Keywords **- curiosity, anticipation, intrinsic motivation, intrinsic development, self-motivating, bootstrap learning.**

I. Introduction

Curiosity is often referred as a drive whose satisfaction should generate positive emotions in the robot. Sometimes this type of learning is also referred as task-independent or task non-specific. During the years of research, curiosity is usually related to notions like: novelty, anticipation, surprise, exploratory behavior, interest, play.

But often curiosity is treated as kind of emotion. In the study of Arbib and Fellous (2004), it is written about the understanding of emotions in their functional context, distinguishing two aspects in emotions: external (emotional expression for communication and social coordination) and internal (emotion for organization of behavior – selection, attention and learning) aspect.

There is another aspect of seeing curiosity - like necessary drive to act and interact with the environment.

[1]Elena Vlahu-Gjorgievska is with the Faculty of administration and management information systems-"St. Kliment Ohridski" – Bitola, Partizanska, bb 7000 Bitola, Republic of Macedonia,
e-mail: evlahu@mt.net.mk

[2]Nikola Rendevski is with the Faculty of administration and management information systems-"St. Kliment Ohridski" – Bitola, Partizanska, bb 7000 Bitola, Republic of Macedonia,
e-mail: nikola@manu.edu.mk

[3]Blagoje Ristevski is with the Faculty of administration and management information systems-"St. Kliment Ohridski" – Bitola, Partizanska, bb 7000 Bitola, Republic of Macedonia,
e-mail: blagoj.ristevski@gmail.com

II. Latest Achievements in Artificial Curiosity

An interesting way of seeing robots' learning is bootstrap learning. Kuipers B., Besson P., Modayil J. and Provost J. are investigating how the foundations of spatial knowledge can be learned from unsupervised sensory-motor experience, relying on Piaget & Inhelders theory that common sense and hence most other knowledge is built on knowledge of few foundational domains such as space, time, action, objects, causality and so on. The basic idea is to reach a learning goal by composing multiple simple machine learning methods using weak but general learning methods to create the prerequisites for applying stronger but more specific learning methods.

The lowest level problem is putting a learning agent in an unknown environment with unknown sensors and effectors. The results are showing that learning even an apparently simple sensory-motor skill requires a large number of distinct learning algorithms, constructing a lattice of different representations of the sensory and motor capabilities of the robot. On the other hand Self-Organizing Distinctive-state Abstraction (SODA) is a new method for automatic discovery of high level perceptual features and large-scale actions for reinforcement learning in continuous environments. SODA combines perceptual abstractions of the agent sensory input into useful perceptual features and a temporal abstraction of the agent's motor output into extended high-level actions, thereby reducing both the dimensionality and the diameter of the task.

It is valuable for a robot to know its current position and orientation with respect to its map of the environment. This allows him to plan actions and predict their results using its map. A paradigm example of bootstrapping learning is place recognition. Here, weak learning method provides prerequisites for an abductive method – topological map-building. The experiment is built on the abstraction of the continuous environment to a discrete set of distinctive states and (assume is that) the agent has previously learned a set of features and control laws adequate to provide reliable transitions among a set of distinctive states in the environment. The place recognition problem is incorporating several different learning methods: unsupervised learning, supervised learning and includes learned topological maps.

For an agent to learn about an unknown world, it must learn to identify the objects in it, what their properties are, how they are classified, and how to recognize them. In this bootstrap learning scenario, the learning agent acquires working knowledge of objects from unsupervised sensory-motor experience. The representation of objects is constructed from dynamic sensor readings in four steps: individuation, tracking, image description and categorization. Dynamic readings are clustered and the clusters are tracked over time to identify objects, separating them both from the background of the static environment and from the noise of unexplainable sensor readings. This learning process leads the agent to experience the indoor environment with significant amounts of dynamic changes. So, the agent learned to individuate and track dynamic object in the scene and created categorization of shape models.

This article gave an important research direction – learning to use vision as a sensory modality. This kind of learning will straddle the developmental boundary.

The next case that we are going to discuss about is implementation of a habit system that automatically executes actions based on internal and external context ("A habit system for an interactive robot", Hsiao K., Roy D.). A robot system explained in this study is with both actions: habitual (actions based on concept) and intentional (actions performed in explicit service of a goal).

Actually, internal context of this system is a set of factors such as maintaining the mental model and external context includes factors coming from the environment, such as motor heat levels, proximity to surfaces and utterances from humans. So, the habits in this system include taking actions to reduce motor heat, avoid collisions with surfaces and interact coherently with humans.

The ability to interact with the environment is facilitated by robot's object-tracking mental-model, implemented as an internal three-dimensional simulation of it's environment. The result is a representation of objects in the robot's environment, along with the representation of the position of the human partner and the robot itself. Ripley (the robot) deals with verbal interactions by parsing the speech, finding word referents within its current mental model. If the robot determines that the situation is clear, there is a single course of action to take, it does it - take the action or respond. If no unique referent is provided the robot responds with question to resolve the ambiguity.

In this system curiosity refers to a drive that causes the robot to look around at various areas of the table and its surrounding environment. Having an up-to-date mental model is an anticipatory action that enables the robot to respond more quickly to requests from the human partner. If we compare this curiosity motivation with animal's drive to stay aware of its immediate environment, periods of exploratory learning would be analogous to an animal playfully trying previously untested actions, or to learn more about relatively unfamiliar objects. Another Ripley's behavior is spoken interaction. But, this interaction system is not very robust to interruption, so this allows all interaction-related actions to be completed before returning control.

What is accomplished is interactive robot capable of semi-autonomously assisting humans in various tasks. Conversational assistive robot is capable of learning about its environment and interspersing its own physical and mental needs with the desires of the interacting humans.

An intrinsic developmental algorithm designed to allow a mobile robot to incrementally progress through levels of increasingly sophisticated behavior is described in "Bringing up robot: Fundamental mechanisms for creating a self-motivated, self-organizing architecture" (Blank D., Kumar D., Meeden L., Marshall J.).

In the described developmental process robot starts with a basic, built-in innate behavior, exercises its sensors and motors, uses the mechanisms for abstraction and anticipation and discovers simple reflex behavior. The key components for the developmental algorithm are the processes of abstraction and anticipation in the context of a model of motivation.

In the first experiment the implementation of the intrinsic developmental algorithm is by using abstractions to govern neural network learning. The goal is to create an autonomous developmental learner based on neural networks. The robot will choose its own actions, initially based on its innate reflexes and eventually based on its internal motivations. Serious problem of neural network is that during repetition of similar sensor signals, a neural controller could get over trained on the appropriate behavior for some state and forgot how to respond appropriately other important states (catastrophic forgetting). In order to avoid this problem, in this case a network governor, an algorithmic device for automatically regulation of the flow of training patterns into the network, is used. This network governor is implemented as an RAVQ (resource allocating vector quantize), abstraction mechanism that can dynamically generate as many categories as needed for the given domain. A purpose of a governor is a training supervisor. Once training is complete, the network can stand alone to perform the tasks on which it was trained. The targets of the network are generated by the system so as to anticipate what movement follows from a set of sonar inputs; actually the hidden layer of the network must be making appropriate abstractions.

Second experiment is about using abstractions to create purposeful behaviors. The approach is for the network to operate on higher-level representational patterns derived from the sensory data, rather than on the raw sensory data itself. These higher-level representational patterns are created by an appropriate abstraction mechanism interposed between the environment and the controller network, serving to insulate the network from the robot's direct experience. In this experiment SOM (self-organizing maps) transforms the high-dimensional sensory data into a single compact, low-dimensional representation of the robot's perceptual state. These more abstract representations are then used by the controller network to determine the robot's next action. This self-organizing system is very different from that which a human might design, but it is exactly appropriate for the robot's sensors and view of the environment. The elements of intrinsic developmental algorithm are abstractions generated by the SOM's and used by the feed-forward neural network. Actually the network is trained on anticipated movements.

These intrinsic developmental algorithms are designed to allow a mobile robot to incrementally progress through levels of increasingly sophisticated behavior. The base concepts (essential mechanisms) of these algorithms are abstractions, anticipations and self-motivations and the ultimate goal of this developmental robotics program is to design a control architecture that could be installed within the robot so that when the robot is turned on for the first time, it initiates an outgoing, autonomous developmental process.

Method proposed in study “Active lexicon acquisition based on curiosity” (Ogino M., Kikuchi M., Asada M.) is based on the estimation of the co-occurrence probabilities between the words uttered by a caregiver and the visual features that a robot observes. This is a lexical acquisition model which makes use of curiosity to associate visual features of observed object with labels that is uttered by a caregiver. In this model the curiosity is based on the evaluated saliency and it affects to the selection of objects to be attended and changes the learning rate for lexical acquisition.

The system learns lexicons on shapes and colors of an observed object through communication with a caregiver. The robot selects one salient object; it acquires the visual features on shapes and colors through visual sensors and utters labels from its own knowledge. At the same time caregiver teaches a label that corresponds to the visual feature of the object that is unknown to the robot. So, curiosity that the robot feels has effects on the selection of the object to be attended. It consist two kinds of saliency: the habituation – low saliency for the visual features that is always observed and high for features that is observed for the first time; and the knowledge-driven, characterized by acquired knowledge – high saliency for the visual features that is not associated with any other label that is already learned.

What we can learn and conclude from this simulation experiment is that the learning model with curiosity acquires the given labels much faster then the simple Hebbian learning model and it shows better performance in the environment in which the number of exposed objects gradually increases. (Important is that the agent cannot associate the visual feature with the word uttered by the caregiver without understanding which feature the uttered word is intended and this is solved by associating the uttered label with the unlearned feature based on curiosity.) More important is that in this method the robot and the caregiver have joint attention. This experiment is only the beginning of investigations in this area. So, it gave us basics for developing this method to a real robot and combines other constraints such as grammar information.

Intelligent adaptive curiosity is a drive which pushes the robot towards situations in which it maximizes its learning process, is explanation of curiosity in paper “Intelligent adaptive curiosity: a source of self-development” (Oureyer P.-Y., Kaplan F.). They see curiosity as a mechanism of self-development while the complexity of its activity autonomously increases. In such environment the robot focus on situations which are nor too predictable, nor too unpredictable.

Important is that when a robot is put in a complex, continuous, dynamic environment it will be able to figure out by itself without prior knowledge which situations in this environment have a complexity which is suited for efficient learning at a given moment of its development.

The simple algorithm is: at each time step robot chooses the action for which the predicted learning progress is maximal. So viewing the learning progress as an internal reward, leads to a classical problem of reinforcement learning. The idea of the improved algorithm is that instead of comparing the mean error in prediction between situations which are successive in time, to compare the mean error in prediction between situations which are similar.

The machine for prediction of the robot is composed by a set of experts which are specialized in particular zones of the sensory-motor space, and each expert possesses a set of training examples, and each training example is possessed by only one expert. This set of examples is used to make predictions. Important is that at the beginning there is only one expert, and as new examples are added the expert should be split into two expert depending on some criterion (split when the number of examples is above threshold set to NS). But there is another criterion which decides how the set of examples is split in two parts which will be inherited by the new expert (finding a dimension to cut). The experiment of this algorithm shows a crucial result from the developmental robotics point of view. It allows a robot to autonomously scale its behavior so that it explores sensory-motor situations of increasing complexity and avoids being trapped exploring situations in which there is nothing to learn. The robot focuses first systematically on one kind of situations and then focuses systemically on another kind of situations.

The goal of this paper is to present a mechanism which enables a robot to autonomously develop in a process that is called self-development. IAC (Intelligent adaptive curiosity) algorithm allows a robot to autonomously scale the complexity of its learning situations by successively and actively focusing its activity on problems of progressively increasing difficulty. This is the first method which allows a developmental robot to go throw all steps autonomously and without prior knowledge.

In another paper of Oudeyer and Kaplan (Discovering communications) they use intelligent adaptive curiosity system as a cognitive architecture of the robot for development of communications skills. This system maximized only the expected reward, so problems related to delayed rewards are avoid, what makes possible using simple prediction system that later can be used in a straightforward action selection loop.

The explained experiment is called “Playground Experiment”. This involves a physical developmental robot capable of moving arms, neck, cheeks and producing sounds, which is installed into a play mat with various toys as well as with pre-programmed “adult” robot which can respond vocally to the developing robot in certain conditions.

What is shown is that more complex linguistic communication shares the same kind of special dynamics that distinguishes it from interaction with simple objects. Learning to predict the effects of the vocal outputs is different from predicting the effects of the motor commands directed towards non-communicating objects. Communication situations are characterized by such kinds of different learning dynamics. This doesn’t mean that they are more difficult to learn then how to interact with these objects.

In this system crucial difference is that the cognitive machinery as well as the motivation system is not specific to communication. Using complex reinforcement machinery brings biases which are specific to a particular method. While using such a method with intrinsic motivation system will be useful for the future research.

In the paper of Stojanov and Kulakov (On curiosity in intelligent Robotic Systems) they describe their understanding of curiosity based on thinking that it is better “a system to do something, rather than nothing”. So they think is good for the agent, if there is no specific goal, via the introduction of curiosity the agent to get rewards whenever it steps into the unknown, which would hopefully improve its world model and its performance on subsequent tasks. The model of the agent that they describe has a collection of inborn schemas that are self-motivated to get executed. So, the process is guided by the primitive internal value systems based on the satisfaction on agent’s drives, and one of them is curiosity. There are four critical mechanisms that guide agent’s development: *Abstraction mechanism* – which enables the agent to deal with more and more complex situations with the same or less cognitive effort; *Thinking and planning mechanism* – which hypothetically combines various previous experiences into new knowledge; *mechanism that provides emergence of more complex inner value and motivational systems* according to which new experiences are judged, foreseen and executed; *socialization mechanism* that enables the agent to interpret in a special way inputs coming from other intelligent agents. In this model agent’s interaction with the environment is represented by a graph with nodes and links (knowledge graph).

In their opinion curiosity is part of the motivational system and it can only partially influence the decision for taking actions, or may provoke internal interest for thinking about certain parts of agent’s environmental knowledge (represent by a knowledge graph). So, curiosity would only maximize the learning curve of an agent equipped with a mechanism for reinforcement learning.

In context of curiosity and understanding they are distinguishing two kinds of feeling of understanding in their agent architecture: feeling of understanding for the working memory and feeling of understanding for the whole agent’s environmental knowledge. In a particular situation, average confidence (of conceptual node within a distance of few links in knowledge graph) is used to judge the situation at hand. If the confidence is high enough, situation (represented in the working memory) seems to be well understood. If it is low the situation is perplexing. The feeling of understanding for the whole agent’s environmental knowledge is calculated as an average of confidence of all schemas in the graph. So, curiosity drive in this architecture is defined as proportional function of the both feelings of understanding. The purpose is to create a tendency to raise the confidence of the agents’ knowledge, because if the agent is generally perplexed it is hardly possible that it would learn something new. Only when the agent is confident enough in its current active knowledge, it is willing to continue to explore new situations and learn new things.

But the agent can understand something if it can find connection of new experience with something, or some experience, that it already understood. In the term of the architecture (the writers explain), the percept will be understood if a connection can be found between current percept and some percept that are part of some already understood schema. So, the transfer of knowledge occurs when a good analogy-mapping has been made. New schemas are constructed between nodes of the knowledge graph, and the interconnectivity is increased.

The most important characteristic of intelligence in explained architecture is expectations. Whenever the expectations are met, the agent does not have to bother what it will do next. The problem for the agent appears when it is surprised (by the detected mismatch between the expected and real percepts) and it has to figure out the solution for the situation. So, the curiosity drive has a function to increase the unexpectancy and uncertainty by adding new nodes in the knowledge graph (meaning something new to be learn) through imagination or thinking, or made by analogy-making.

The purpose of the paper is about curiosity not to be treated as simple driving forces which pushes agent to do something but as an elaborated mechanism which is inseparable from the internal knowledge representation and as guide for the process of thinking and imagination.

III. Conclusion

What we can say as conclusion of this paper is that after a half-century of continued research, the artificial intelligence is still far from developing any type of general purpose intelligent systems. But there are a lot of successful experiments that gave hope and that should be the basic issue for developing such system.

The main idea of integrating the curiosity in artificial intelligence is to recreate the world of a human infant, an entity with a sense of being, with a notion for exploring its environment, building a robot which develops.

References

[1] Stojanov, G. and Kulakov, A., 2006. On Curiosity in Intelligent Robotic Systems, (AAAI)

[2] Oudeyer, P.-Y. and Kaplan, F. 2006. Discovering communication. *Connection Science, Vol.18, No.2*

[3] Oudeyer, P.-Y. and Kaplan, F. 2004. Intelligent Adaptive Curiosity: a source of Self-Development

[4] Blank, D., Kumar, D., Meeden, L., and Marshall, J. 2005. Bringing up robot: Fundamental mechanisms for creating a self-motivated, self-organizing architecture

[5] Hsiao, K. and Roy, D., 2005. A Habit System for an Interactive Robot, *(AAAI)*

[6] Ogino, M., Kikuchi, M. and Asada, M., 2006. Active Lexicon acquisition based on curiosity

[7] Kuipers, B., Beeson, P., Modayil, J., and Provost, J. 2006. Bootstrap learning of foundational representations, *Connection Science, 18 (2), pp 145-158*

[8] Arbib, M. and Fellous, J-M., 2004. Emotions:from brain to robot, *TRENDS in Cognitive Sciences Vol.8 No.12*

[9] Barto, A.G., Singh, S., Chentanez, N., 2004. Intrinsically Motivated Learning of Hierarchical Collections of Skills