A COMPUTATIONAL STUDY ON THE ROLE OF PLANNING AND AFFORDANCES IN LEARNING: PLAN TO LEARN VS. LEARN TO PLAN

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

ECE KAMER TAKMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COGNITIVE SCIENCE

SEPTEMBER 2015

Approval of the thesis:

A COMPUTATIONAL STUDY ON THE ROLE OF PLANNING AND AFFORDANCES IN LEARNING: PLAN TO LEARN VS. LEARN TO PLAN

submitted by **ECE KAMER TAKMAZ** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Science, Middle East Technical University** by,

Prof. Dr. Nazife Baykal
Director, **Graduate School of Informatics** ________________

Prof. Dr. Cem Bozşahin
Head of Department, **Cognitive Science, METU** ________________

Prof. Dr. Cem Bozşahin
Supervisor, **Cognitive Science, METU** ________________

Examining Committee Members:

Prof. Dr. Deniz Zeyrek
Cognitive Science Department, METU ________________

Prof. Dr. Cem Bozşahin
Cognitive Science Department, METU ________________

Assist. Prof. Dr. Cengiz Acartürk
Cognitive Science Department, METU ________________

Assist. Prof. Dr. Murat Perit Çakır
Cognitive Science Department, METU ________________

Assist. Prof. Dr. Murat Ulubay
Department of Management, Yıldırım Beyazıt Üniversitesi ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ECE KAMER TAKMAZ

Signature :

ABSTRACT

A COMPUTATIONAL STUDY ON THE ROLE OF PLANNING AND AFFORDANCES IN LEARNING: PLAN TO LEARN VS. LEARN TO PLAN

Takmaz, Ece Kamer
M.S., Department of Cognitive Science
Supervisor : Prof. Dr. Cem Bozşahin

September 2015, 53 pages

Our prior plans supervise our actions and help us form new plans, which turn out to be the modified versions of our related previous plans. This is the process that we call "plan to learn", a procedure involving self-supervision to learn more about the environment and extract meaning out of it. In cases where our prior plans fall short, it is not easy for us to produce efficient or complete plans. For instance, in a world very different from our own, with objects whose affordances appear to be false, encountering seemingly random situations prevent us from understanding our environment and how to act in it. If the differences are small enough, then we might be able to find patterns and adapt to that environment. However, if the distinctions are too large to make a meaning out of them, we can be stuck on such occasions, as we would not know how to make a reasonable plan in such a setting without supervision and without using our innate planning mechanism, thus we cannot "learn to plan". In this thesis, these two processes, "plan to learn" and "learn to plan" in structured problem domains are compared. To this end, we conducted two experiments using a video game involving object interaction and in the light of the outcomes, we developed a computer model that uses prior plans containing affordances to learn about the environment and to update its knowledge of the world.

Keywords: Planning, automated planning, learning, affordances, supervised learning

ÖZ

PLANLAMANIN VE SAĞLAYICILARIN ÖĞRENMEDE YERİ: ÖĞRENMEK İÇİN PLAN YAPMAK VE PLAN YAPMAYI ÖĞRENMEK

Takmaz, Ece Kamer
Yüksek Lisans, Bilişsel Bilimler Programı
Tez Yöneticisi : Prof. Dr. Cem Bozşahin

Eylül 2015 , 53 sayfa

Geçmiş planlarımız hareketlerimizi yönlendirirler ve yeni planlar oluşturmamıza yardımcı olurlar. Bu yeni planlar, eski planlarımızın yeni duruma uyarlanmış versiyonlarıdırlar. Bu süreç "planlama yaparak öğrenme" olarak adlandırılabir. Bu süreçte çevre hakkında daha çok bilgi edinilir ve çevredeki objeler ve hareketler anlamlandırılır. Geçmiş planlarımızın bizi yönlendirmek konusunda yetersiz kaldığı durumlarda, kullanışlı planlar oluşturmak kolay değildir. Bizim dünyamızdan oldukça farklı, yanlış sağlayıcılar (false affordance) taşıyan objelerle dolu bir dünyada, rastgele olayların gerçekleşmesi çevreyi anlamlandırmayı ve onun bu sistemde hareketin planlanmasın engeller. Eğer bu dünya ile o dünya arasındaki farklar azsa, bir takım tekrar eden kalıplar bularak o dünyaya uyum sağlayabiliriz. Ancak aradaki farkların çok fazla olduğu durumlarda bazı durumlarda takılı kalabiliriz. Bu gibi koşullarda sistemi nasıl anlamlandıracağımızı bilemediğimiz için doğuştan gelen planlama mekanizmamızı kullanarak bu durumda nasıl mantıklı planlar oluşturabileceğimizi de bilemeyebiliriz, yani bu yeni sisteme göre plan yapmayı öğrenemeyebiliriz. Bu tezde, "öğrenmek için planlama" ve "plan yapmayı öğrenme" süreçleri karşılaştırılmaktadır. Bu amaç için objeler ile etkileşimin olduğu bir oyun geliştirdik ve bu oyunun oynanmasını içeren iki deney düzenlendi. Bu deneylerin sonucunda elde edilen bilgilerin ışığında geçmiş planlarını ve sağlayıcı bilgilerini kullanarak çevresi hakkında bilgiler toplayan ve planlamada kullanacağı bilgileri güncelleyen bir bilgisayar modeli geliştirildi.

Anahtar Kelimeler: Planlama, otomatik planlama, öğrenme, sağlayıcılar, gözetimli öğrenme

To my family

ACKNOWLEDGMENTS

I express my sincere gratitude to my supervisor Prof.Dr. Cem Bozşahin, for letting me have the chance of being part of the Cognitive Science Department and also for his precious insights into planning and his help throughout the development of this thesis. I would like to thank Assist.Prof.Dr. Cengiz Acartürk, for his invaluable guidance from the very beginning of my master studies. For their comments and suggestions, I offer sincere thanks to the faculty members of the Cognitive Science department, Assist.Prof.Dr. Murat Perit Çakır, Assoc.Prof.Dr. Annette Hohenberger and Prof.Dr. Deniz Zeyrek. I would like to thank my fellow colleagues Tunç Güven Kaya and Gökhan Gönül, for being great officemates, with whom I could discuss basically everything. My friends Özge Nilay Yalçın, Yudum Paçin and Murathan Kurfalı, thank you for always being there when I needed some distraction from my studies. Students at the Cognitive Science Department and colleagues at Microsoft Language Development Center in Lisbon, Portugal, thank you for taking the time to participate in the experiments. Finally, I would like to thank my dad, my mom and my sister for the patience they demonstrated during my writing of this thesis and for their never-ending support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Planning plays a crucial role in human lives, when trying to attain our goals, to collect information, to speak, to search, to remember and even to make plans, we try to construct plans accordingly. Therefore, it should not come as a surprise that planning is made up of complex information processing stages. Planning problem in itself is one of the hard problems. The nature of planning, structure of plans and the formation of plans all require rigorous research on human behaviour and thoughts. Observing only the actions of humans would not be enough to grasp the nature of plans, as sometimes our plans may not be objectively observable and measureable from the outside.

Additionally, no account of the nature of planning would be complete without delving into learning theories. Before speculating about learning in humans, many studies were conducted on animals. In the psychology literature, there exist numerous different explanations of what the animals exactly learn when running around a maze and finding paths to the rewards. In addition to what they learn, how they learn is also a topic that attracted numerous scientists. One of the explanations for the learning behaviour is that it consists of conditioning (or association) and reinforcement in the presence of a reward. This kind of learning is response-learning in which the animals simply remember the sequences of movements they made and the responses they get from these movements around the maze. Response-learning is a kind of association or condition learning where the organisms learn to act in a certain way through associating the stimulus and response patterns. Other kinds of associations include stimulus-approach association, place-approach association or chains of these associations (Balkenius, 1994).

However, Tolman's place-learning experiments introduced a rather different idea that organisms create and fill in cognitive maps representing the spatial layout of the environment (Tolman, 1948). In this kind of learning, which is supported by latent learning, reinforcement is claimed to be not necessary. According to the place-learning account, when the rats run the maze, they pick up information about the spatial layout and store them in certain kinds of representations, even when there is no reward at the end of the experiment. However, the lack of this account is that it mostly focuses on the information gathering and storing part, and misses the point on how to use this information to create sequences of actions to perform in the environment to achieve a goal, that is to create and execute a plan (Miller et al., 1960). Yet, this account differs from its contemporaries in taking the cognitive structures in learning into consideration, as the constructivists did to a greater extent, and arguing that learning does not consist of just stimulus-response learning supported by behaviourists.

Planning in humans differ from the animal planning. Planning in animals is said to be reactive, that is, they can plan in the presence of a tool towards a goal. However, if the tool is absent they do not look for a certain kind of tool to achieve an end. In humans, deliberative planning takes place. In addition, available tools can be defined in many ways according to the needs of humans. A crucial aspect of the planning behaviour in humans is that it requires recursive, compositional and hierarchical aspects. In this way, it might be claimed that, as languages are also of recursive and compositional nature, language and planning can be related to each other.

One crucial idea that we take into consideration is that complex behaviour occurs with the help of prior planning and hierarchical organization, which indicate the compositionality of the learning and planning behaviour. As most of the problems we face are structured, we focus on finding solutions to structured planning problems with the help of our previous plans related to such structured problem domains. Therefore, in this study, we will try to investigate the theories on the nature of the learning and planning in a simple environment where object interaction and problem-solving takes place. In addition, we would like to explore the contribution of affordances to the learning and planning processes in structured worlds. To this end, we use automated planning software and refer to planning concepts. In this study, claims about the nature of learning and planning will be studied from a computational approach and an Artificial Intelligence perspective with the help of PKS, a knowledge-based planner, Planning with Knowledge and Sensing. There are Computer Science studies on planning and learning; nevertheless, most of them focus on rational agents, which on many occasions perform more efficiently than the human brain to provide correct answers to the problems they are given. However, what we are concerned here is not creating such agents. On the other hand, we would like to make use of theories of learning and planning to guide us in simulating the human brain on computers, which in turn may guide us in better understanding the inner workings of the human mind. It is beneficial to consider this process as part of the quest of exploring to what extent a computer can simulate learning and planning psychologically adequately and functionally correctly by following the concerns of Good-Old-Fashioned AI.

The organization of this thesis is outlined below:

Chapter 2, provides an account of learning theories, specifically focusing on reinforcement learning and cognitive maps. In addition, information processing systems and computational learning theories are also discussed.

Chapter 3, first subsection explains what planning is in terms of human behaviour and also what automated planning is, which is a branch of Artificial Intelligence. In this chapter, an account of a classical planning problem is also presented. In the second subsection, the concept of affordances is elaborated.

Chapter 4, is the part where the methodology is explained. First, details about the research questions and the related experiments that were conducted are presented, along with the results of the experiments. Then, the implementation details of the computational model according to the outcomes observed from the experiment results are explained.

Chapter 5, concluding remarks are made and possible future investigations and extensions are discussed.

CHAPTER 2

LEARNING

2.1 What is learning?

Learning is a central aspect of psychology and there exists numerous different ways of explaining what it exactly is. Learning could be demonstrated by adaptive changes in behaviour while taking an action, changes in thinking, reasoning or problem-solving activities (Lorenz, 1977). The reason learning is such a crucial part of psychology is that it helps an organism adapt to its environment that eventually leads to the survival of that organism and even the whole species. Without learning, it would be quite hard for an organism to adapt to its environment, since learning improves the organism's awareness and understanding of its habitat. In this way, with the help of experiential information they gathered through learning, organisms are able to build predictions about the environment. This would then advance the situation of that organism among its rivals in accordance with the laws of nature. Learning covers memorizing, information gathering, acquiring facts and skills, interpreting the world and so on. There are ongoing debates in psychology about exactly what happens when a rat presses a bar or runs a maze to gather food (Shettleworth, 2001). The reason for this debate is that animals sometimes seem to behave simply in a reactive to some stimulus , while sometimes also appearing to behave in ways that hint the inclusion of structural representations about the world.

Some claim that learning is a process where the animal gathers information, updates its views and modifies expectation etc, whereas others claim that it is a product, which is the end object (a different behaviour, for instance) of a process similar to what is defined before. The idea of learning as a product fits behaviourist ideas the most, since learning to them is manifested by observable changes in the behaviour. The idea of learning as a process matches information-processing theories of learning, as the whole encoding and decoding of information during learning indicates that learning is not just a product.

However, in most cases, learning is explained in terms of relatively permanent changes in the knowledge and behaviour of an organism as a result of experience. Organisms collect various sorts of data from their sensors and by observing the consequences of their actions. Behaviourists, for instance, assert that as we cannot objectively observe thoughts or mental states of an organism, we should disregard them in the scientific study of learning and we should only be concerned with what is objectively observable and measurable. Watson claims that learning is indicated by observable behaviours and not the conscious mind or mental because they cannot be measured in an objective manner (Watson, 1913). Although some behaviourists dismiss mental aspects altogether, some argue that they exist, however they should not be considered objective measures of learning. Therefore, on the whole, they observe learning in terms of changes in the visible and objectively assessable behaviour. At this stage, the question of whether performance is equal to learning comes forward. Some scholars state that as we can observe performance objectively, it should be considered the indicator of to what extent the organism learned new things. Yet, some maintain that even though performance might demonstrate certain aspects of changed behaviour, which in turn indicate learning, there might be hidden parts of what the organism learned, which could be manifested at a much later time.

One of the main contentions in the psychology of learning is concerned with the question of whether learning comprises of conditioning, association learning, stimulus-response (S-R) learning, or the processing of cognitive structures. In the following subsections, various views on this topic will be explained.

Learning theories, especially the ones that are related to the animal learning behaviour provide us with different explanations of the learning that takes place under certain conditions. Animal learning theorists propose several kinds of explanations such as response learning or place learning.

Several explanations of how and what the animals learn have been proposed such as conditioning, association, and reinforcement in the presence of a reward. Stimulus-response learning is a sort of conditioning in which the learner remembers the sequences of movements, the related responses, or any pattern it observes and utilizes this association-conditioning knowledge in its later runs. Therefore, the animal is said to "learn to act in a certain way through associating the stimulus and response patterns" (Tolman, 1948). However, this claim does not simply propose that there is only one kind of conditioning. Instead, there could be various kinds of association such as stimulus-approach association or place-approach association or chains of these associations (Balkenius, 1994).

Stimuli can originate from the environment which is an external stimulus or from internal resources. Sequential S-R learning was considered to be the basis of learning. Through observation and experience these S-R associations are thought to be formed and manipulated. At the beginning of the 20th century, behaviourism was very influential in psychology studies. Ivan Pavlov's classical conditioning, B.F. Skinner's operant conditioning (sometimes called instrumental conditioning) and Thorndike's Law of Effect were the seminal concepts in this context.

Ivan Pavlov was not a psychologist, rather he was a physiologist, who was studying the salivation in dogs. When he noticed that even at the sound of his footsteps the dogs start to drool, he started to theorize what is called "Classical conditioning" (Pavlov, 1927). Classical conditioning includes several kinds of stimuli and responses in each stage. Unconditioned stimulus (US) causes innate responses, which in the case of Pavlov's dogs, corresponds to the meat that makes the dogs salivate. The reason it is labelled unconditioned is that it naturally and automatically causes the dogs to salivate. Thus, salivation is the unconditioned response, UR. Neutral stimulus is the stimulus which is at the outset not related to the occurrence of the unconditioned response. In Pavlov's studies, neutral stimulus was the sound produced by a tuning fork. This tone originally does not have an impact on the salivation. However, when paired with the delivery of food, salivation, which initially occurs when the US is observed, also starts to happen at the presence of the sound. After conditioning has taken place, even when there is no food around, merely the sound of the tuning fork is enough to cause salivation, therefore, at this stage, the tone of the tuning fork is called conditioned stimulus (CS) and salivation which occurs in the absence of food and presence of tone is called conditioned response, CR.

Skinner explained behaviour in terms of reinforcers employed in the process (Skinner, 1953). He belonged to a part of the Behaviourism school that claimed that learning consists of stimulus-response associations and animals demonstrate that they learnt something via the changes in their behaviours or actions which are observable, rather than the changes in their mental representations which are not objectively observable. However, he also believed in the existence of mental representations, yet he dismissed them by arguing that they cannot be objectively observed.

There are different kinds of reinforcement learning. First one is "Positive reinforcement" which occurs when a positive reinforcer is introduced in order to increase the frequency of the previous behaviour. Thus, when the organism demonstrates an action and then detects that this it obtains a positive reinforcement, it would be predisposed to that behaviour more than before. Negative reinforcement occurs when after a certain display of behaviour; a negative reinforcer disappears from the environment. Punishment is the introduction of a negative element into the organism's environment after it performs an action. Hence, this negative element would result in a decrease in the frequency of the occurrence of the preceding action.

Operant conditioning is claimed to be voluntary and not reflexive as in the case of classical conditioning. In addition, the order of S-R is also different. What behaviour is learnt and repeated in operant conditioning depends on the consequences. Skinner acknowledged the role of thought in learning by claiming that organisms operate on the environment to achieve a goal and that operant conditioning is considered voluntary rather than reflexive or reactive. In addition, stimulus and response learning and reflexive actions are crucial for survival as they allow for immediate reaction, whereas operants are significant, because they ensue beneficial siutations (Skinner, 1953).

Reinforcers can be either of two types. Primary reinforcers are related to biological needs of the animals, such as food and water. Secondary reinforcers are not directly related to biological needs, however, they are later learnt to be associated with the biological needs, such as money or other kinds of rewards, even praises.

Vicarious learning and observational learning are also worth mentioning. Vicarious learning occurs when the animal achieves learning via trial and error processes. Observational learning is the imitation of observed behaviour that are demonstrated by other animals in the environment. Learning by perceptions involves the utilization of receptors. For instance, via visual learning, we learn about the visual properties of objects. Perceptual properties of the objects in the environment can help facilitate recall. It is also important in object recognition, where humans try to recognize objects by either utilizing structural descriptions, or geometric constraints, or feature spaces or shape-space approximation (Edelman, 1997).

Animals can also perform motor learning, spatial learning or observation learning. Most of the data for such learning comes from the visual receptors, and vast amounts of information we store are due to seeing (Laird, 1985). In Thorndike's study where the cats learn to push a button in a box in order to open the latch and reach the food inside, it is claimed that cats randomly happen to push that button and observe by trial and error that it leads to the opening of the latch (Thorndike, 1898). Thorndike's concept of rewards corresponds to Skinner's concept of reinforcers.

Thordnike's Law of Effect asserts that if a response produces satisfying outcomes, the organism would be more inclined to repeat it and if the outcomes of a response are not satisfying the frequency of the response's occurrence would decrease. A crucial aspect of operant conditioning concerns the schedule in which the reinforcers are introduced. Contingency of the reinforcement is a significant part of operant conditioning. If, after a while, the positive reinforcer is no longer introduced, then the frequency of the previous behaviour might decrease, which is explained by the Law of Extinction. Nevertheless, if that reinforcer is re-introduced after this extinction, frequency might increase again. Therefore, influence of the environment is considered to be an important part of learning and shaping the behaviour of subjects and this theory also embraces the existence of expectations of the organism, which might be considered cognitive elements. Another important aspect of operant conditioning, which is related to the classification of stimuli also seems to include certain cognitive processes such as stimulus discrimination and generalization. For instance, the differences in the tone of a signal that might be utilized in an animal learning experiment could lead to distinct responses, as the animals can differentiate between the tones.

We can learn things and we do not have to exhibit them, hence, performance does not always signify learning. Human behaviour appears to be comprised of more than just stimulus-response associations. Behaviour should also include thinking, remembering experiences or behaviour. Inferences about them can also be made with careful observation. Wilhelm Wundt, sometimes called "the father of psychology", asserted that mental processes of the mind are also crucial, hence, in order to study the thought processes, he utilized introspection. When we learn something, we modify our knowledge about the world. Certain stimuli draw our attention and by actively processing information about the environment, we learn and behave. Rather than passively learning, we organize information for storage and retrieval, we modify them and relating them to past events, we are able to build expectations.

Tolman's place-learning experiments introduce a rather distinct idea. Tolman asserts that there exists a learning type called "latent learning" in which the spatial orientation is quite crucial and the learner turns its attention to selective stimuli and forms hypotheses concerning the layout of the maze. Additionally, Tolman claims that learning does not consist of stimulus-response connections; instead, it is the construction of sets that function like cognitive maps formed as an organized collection of knowledge and expectations (1948). Spatial layout is represented by cognitive maps and the learner creates and fills in these cognitive maps. One of the most important aspects of this theory is that reinforcement is claimed to be not necessary for the learning to take place. However, some critics assert that reinforcement does not have to be a reward-like food. Reaching certain points in the maze or looking for an exit because the feeling of running around in the maze might cause anxiety-like behaviour in the rats; hence, these issues might also be taken as reward-bringing conditions.

McLellan asserts that the logical and computational model of a mind is on the rise due to the credit given to it by AI and cognitive science (McLellan, 1996). In terms of information processing theories of learning, organisms are said to code, store and retrieve information when learning and then later retrieve such information and manipulate it. This idea of learning puts the organism in the centre of the learning process, where the organisms are active participants that seek to gather information and not passive observers and reactors. Additionally, these theories involve analogies to the inner working of computers in terms of memory, storing information in databases, accessing them, manipulation of symbolic data. According to Marr, seeing involves computational processing of images "to know what is where by looking" (Marr, 1982) Hence, vision involves low-level processing of stimuli to extract information about the locations, positions and orientations, as well as the high-level processing of determining if there are familiar objects in the images. Low-level processing of visual stimuli corresponds to "seeing" and high-level processing corresponds to "seeing as", in terms of the distinctions made by Wittgenstein (1972) According to Balkenius, learning must be a computational process acting on a completely described system that was based on a descriptive framework (Balkenius, 1994).

According to the constructivist views on learning humans acquire various pieces of knowledge about the environment through problem-solving and discovery and they do not have to be spatial. Constructivist views are the opposite of behaviourist concepts that focuses on behavioural observations, whereas the former focuses on cognitive operations. The main proponents of behaviourism are Watson, Thorndike, and Skinner. As for the supporters of constructivism, we can name Dewey, Vygotsky Piaget, and Papert.

When we construct meaning, it is also essential to have certain structures that are capable of retaining knowledge. Bartlett suggested schemata, which are unconscious mental structures, as representations to hold world knowledge. Marvin Minsky's concept of frames also deals with the representation of knowledge in humans in order to construct similar structures in machines. Rumelhart, who is a cognitive psychologist, took Minsky's frames and formed a psychological theory pertaining to the mental representations of complex knowledge. Schank and Abelson's scripts and scenarios are constructs that focus on knowledge about sequences of actions in different domains.

Piaget maintains that natural curiosity to construct meanings out of the world causes the children to explore the world actively, instead of passive information collection or stimulus reactions. Vygotsky emphasises the role of the environment in children's development. Piaget and Vygotsky's work related to constructivism offered seminal ideas on the development of cognitive skills in children. In particular, Piaget suggested that learning is not cumulative, instead, it is more likely to be transformative. One of the most important claims by constructivists is that prior knowledge influences the learning process, and when solving novel problems, we try to connect them to our previous knowledge with the existence of similarities.

In the literature, different definitions of cognitive maps exist. However, the foremost argument is that they hold information about the environment and this information does not have to be acquired because of a reinforcer. In addition, a fundamental assertion of cognitive maps is that they are different from routes. If some parts of the path acquired after route learning is removed from the route, the learned path would no longer help the animal get to the point it desires. Nevertheless, cognitive maps allow for the removal of components and shortcutting of paths, as the map involves routes, paths, environmental relationships and other possible spatial representations. Bennett argues that that the idea of creating shortcuts indicates the existence of cognitive maps is not correct and he lists two reasons why this might not be the case: Animals might have been familiarized with the environment beforehand, or they noticed some familiar landmarks (Bennett, 1996).

A significant part of cognitive map-related discussions point out to a difference in the representations of these maps. They are not exactly like cartographic maps, but it is possible that the representations we create in our brains is similar to those maps. It might also be the case that cognitive maps represent the real environment in an analogue way, meaning that the representation more or less resembles what we see out in the environment. Analogue representation holds an image of the environment, objects, colours, shapes and sizes.

However, quite distinct from the other representations of cognitive maps, it is also plausible that they are actually represented with propositions such as north of this building is that park etc. Blind people are also able to create cognitive maps, therefore the whole experience of being in an environment and walking around in it could also help create the cognitive map. For that reason, it can be claimed that spatial memory does not always have to store visual information. In terms of both the visual and spatial data that is received, visuospatial sketchpad seems to be the structure retaining such information.

Intense inner motivations lead the animals to explore and learn about their environments (Strange, 1950). Exploration of novel stimuli through perceptual curiosity helps animals better act in familiar environments, as well as adapt to temporally and spatially changing environments (Berlyne & Slater, 1957). Natural curiosity and natural eagerness to learn are also crucial properties of humans (Laird, 1985). Through experiences and observations we form concepts and then to test our concepts and find solutions, we seek new experiences. This is a search for meaning, and constructivists claim that we construct this meaning mentally. The number of cues in the maze could affect the performance of the rats and their exploratory behaviour in the absence of a reinforcement during latent learning could facilitate a less random performance (Kimball et al., 1953).

Episodic memory, a term coined by Tulving, which stores events happened to us in the past, as well as when and where they happened, guides us in future planning (Tulving, 1972) An important aspect of episodic memory is that the person who remembers a past event that happened to him or her, remembers it with subjective awareness (Clayton et al., 2003). It is argued that some animals demonstrate future planning according to their desires or motivations, hence “mental time travel” is not unique to humans. When humans recall past events, we can obtain their verbal explanations of recollections. Nevertheless, this is not the case for animals, therefore, we must infer certain aspects from their behaviours. Particularly the case where the animals make predictions for their future desires and plan accordingly, compared to fulfilling their current desire, which is called the Bischof-Köhler hypothesis (Clayton et al., 2003). Whether animals have episodic memory is a debatable issue in psychology. The reason for this is that, when we recall something that happened to us in the past, we recall it happening to ourselves. Here, the concept of self and retrospective reminiscing capabilities are quite essential. Some animals are claimed to have episodic-like memories, such as the worm and nut caching scrub-jays who prefer the boxes they cached nuts after a long time (Clayton et al., 2003).

Spatial problem solving is a vital skill, as it provides the animals with possibilities to survive in their environment guiding the processes of foraging, escaping danger, chasing game. Hence, advanced spatial problem solving skills are advantageous in terms of evolutionary constraints. Moreover, sometimes when humans are under threat high level cognition is shut down, therefore, quick and effective use of information would be beneficial for survival. It is also crucial to have enough motivation to pay attention to relevant stimuli in such cases.

During the learning process, brain collects data from all sensory organs. A huge amount of data streams into the different locations in the brain causing the neurons to form new connections. Therefore, it would not be correct to confine the learning to just a few areas in the brain. However, particularly the role of hippocampus in place learning and the spatial memory is discussed frequently in the literature. It is now widely accepted that hippocampus plays a crucial role in memory-related processes (Eichenbaum et al., 1999). One seminal study conveyed the idea that a physical space representation, similar to cognitive maps, resides in hippocampus

(O'Keefe. & Nadel, 1978). Cues, topology of the environment and certain landmarks are encoded into hippocampal place cells; however, it should be noted that location-specific activity is the one that is being stored. There is some dispute about how the spatial information of running a maze is represented in the brain. To some authors, hippocampus stores only a component of spatial information and not a global map and not all processing in hippocampus is dedicated to spatial memory, thus alternatives to the cognitive map theory should be considered (Eichenbaum et al., 1999).

Most learning studies are based on animal research. There are also studies that compare the performances of human children in maze running to that of rats (Husband, 1929). Although similarities can be found between animal learning and human learning, it should be noted that due to species-related differences and innate properties, such as the differences in the physiology of the eye, brain nose etc. not all aspects of learning can be the same. Therefore, some results should be approached cautiously.

In the context of machine learning, supervised learning is the utilization of labelled training data or prior knowledge of concepts in order to classify novel instances (Mitchell, 1997). There can be several kinds of supervised learning. For instance, an agent can actively try to find out about the classifications of new examples, or it can be given the classified data directly, or the agent can ask relevant questions to an all-knowing oracle in order to improve the hypotheses. Unsupervised learning, on the other hand, involves determining and extracting the classifications from an unlabelled data set that consists of instances not previously encountered (Mitchell, 1997).

In the context of this thesis, the concept of "planning to learn" corresponds to supervised learning, and the concept of "learning to plan" corresponds to unsupervised learning. Similar to Jerry Fodor's claims in the Language of Thought hypothesis, about the mental representations and mentalese, the language of thought, which is claimed to have syntax like the natural languages and required for and supervises the process of learning natural languages (1975), we consider plans to be compositional and novel plans to be the combination or modification of simpler prior plans that supervise the formation of new plans and learning. Therefore, the use of prior plans as supervisors indicate self-supervision in planning and learning, which is orthogonal to reinforcement learning with the external concepts of reinforcers such as rewards.

In terms of learning in computer science, formal learning theory, or computation learning theory, is concerned with the acquisition of knowledge through observation and how an agent can reach correct outcomes from this process (Kearns & Vazirani, 1995). Valiant's theory of the learnable suggests the learning of generalizations that are Probably Approximately Correct (PAC) (Valiant, 2013). The most crucial claim is that humans can acquire new concepts without any explicit programming. Although the learning of some skills are genetically preprogrammed, or some explicitly memorized, there seems to be other skills that do not belong to these catergories. Valiant's choice of knowledge representation includes propositional calculus and formal grammars (Valiant, 1984).

CHAPTER 3

PLANNING AND AFFORDANCES

3.1 Planning in Humans vs. Planning in AI

Planning is the major focus of this thesis. Plans are sequences of operations or instructions that we form in order to achieve expected results. Hence, planning is the act of finding possible sequences of actions that would transform the current state or the start state into the goal state (Ghallab et al., 2004). Usually the goal is a certain one and the steps that should be taken to reach this goal are uncertain or unknown. Planning requires the use of complex cognitive processes and prefrontal cortex in the brain is claimed to be closely related to the ability of planning as lesions in the frontal lobe seem to be correlated with the occurrence of disorders of the planning function(Unterrainer & Owen, 2006). In addition, prefrontal cortex in humans demonstrates a more highly developed structure, which is also connected with a high number of complex sections in the cortex, than the one in other primates.

An instance of a planning problem might include recalling the past and predicting the future and when we are faced with intricate problems that call for the use of advanced skills or forming strategies, we are bound to plan our actions beforehand. Recalling the past is realized through the episodic memory, which holds procedural information about the conscious experience of one's recollections. This type of memory is different from the semantic memory that contains descriptive knowledge.

Humans can verbally describe contents retrieved from their episodic memories by explaining what happened, where it happened and when it happened. With that in mind, it can be said that plans that humans make can also have a verbal nature. When we plan to go to work, we can simply label the sequences of actions that are included in a plan of going to work with the hierarchically superior term of going to work. These actions could include leaving home, getting into the car, driving, reaching the office, which in themselves include more activities. For instance, leaving home involves opening the house door, going through it, closing it, unlocking it etc. All of these actions can be again dismantled into different components. One can even talk about the sequence of muscle movements required to open the door. Therefore, we can claim that the plans and the process of planning encompasses the utilization of hierarchical structures. General Problem Solvers developed by Simon and Newell also make use of the hierarchical structures, as well (Newell et al., 1959).

If we take planning in animals into consideration, this is not the case, as they cannot convey such information related to a subjective experience of an event, if exists, verbally. Nevertheless, it is also possible to make use of the visuo-spatial memory while planning (Unterrainer

& Owen, 2006). Tulving's definition of episodic recall indicates that what is retrieved from the episodic memory contains information about a unique event, what happened then, where it happened and when it happened (Tulving, 1972). Episodic memory is closely related to future planning. Hence, we cannot eliminate the possibility of animals having episodic memory. Some animals are claimed to have episodic-like memory. Birds caching nuts and worms prefer the caching locations that include nuts, if the time passed is long enough for the worms to decay, considering their previous experiences with caching worms. Animals might recall past events, but it is debatable whether they recall the action as it was done by themselves (or when they were around) or they simply recall some action happening under certain conditions. Yet, planning is about future events that have not so far happened, therefore an understanding of the future and expectation about it is necessary.

Planning in humans appears to include a sense of self or previous action done by self. Studies about the nature of planning in humans indicate that there is some heterogeneity in the courses of action that the humans take in certain kinds of tasks such as wayfinding in a maze. They use different kinds of remembering or recalling techniques or while walking around in a real or virtual maze they employ distinct ways of exploration.

It is possible to develop both online and offline plans. Offline plan means that the subject or the agent is given a problem definition and tries to form a plan before interacting with the environment. Online plans, however, differ in the way that they are formed during interaction. Tower of London studies demonstrate the differences between the number of operations performed and the number of errors in solving the problem after offline planning versus during online planning (Owen et al., 1995).

It is vital to note that not all plans in humans are rational or reasonable. We do not always choose the best possible course of actions. Indeed, sometimes we have so little information about the situation that we cannot even form proper plans. If the conditions are uncertain, or if we have incomplete information about the environment or if we are not able to predict what the results of our actions would be, we make assumptions and then move on to form plans or without plans, we start to perform some actions and after some exploration we begin to shape our plans. These deficiencies were about the actions that we can take in that environment and how their results would be. Ill-defined problems lack certain parts of a planning problem compared to well-defined problems. If the problem we encounter is an ill-defined one, we need to fill in empty slots by making assumptions using heuristics or statistical data. Nevertheless, it is hard to contemplate about a planning problem without the definition of a goal state. If we do not have a clear definition of the goal state, all of our actions could be meaningless or random. If we have a vague concept of the goal state, it is a better condition. Yet, it still has shortcomings that would prevent the formation of a proper plan. Therefore, goal-directedness is a crucial property of planning. Using a sequence of operators and actions, start state is transformed into the goal state. Planning is used when the steps to a goal state are unknown or when the steps are known but the order is uncertain. When we have a goal to achieve, we also usually have a plan in mind to modify and execute it in a certain context. This means that our actions are supervised by a sequence of operators we have previously brought together. Operators are the components of plans that are applied to a state to reach another state. Basically, they are action definitions that include preconditions and effects. Hence, this means that we have an idea of what might happen if we conduct an action on an object or with an object. However, this is also determined by our past experience with the objects or explorations using those objects.

Goal state is the solution state and it is also possible to define subgoals and create plans for reaching these subgoals to eventually reach the major goal. State transformations are done through the use of operators and they represent certain actions happening. Problem space includes all possible states that can be reached from the start state through the use of operators. Every operator cannot act on every state and there might be certain constraints or preconditions for an action to take place. Working memory and its capacity plays a crucial role in planning in humans. As far as the problem space is concerned, humans cannot possibly hold every state in their working memory. Instead of algorithms that can provide clear-cut solutions to a problem, humans might be making use of heuristics. The use of algorithms to reach a certain goal can take a very long time, although at the end the goal is reached (Miller et al., 1960). However, heuristics or certain strategies or tactics can help reach a solution faster. Nevertheless, this might not be the exact solution required at the beginning of the problem, but a similar state. Therefore, human planning is claimed not to be optimal. In addition, state space search is not objective, as people construct their own representations of the same problem and the same environment (Unterrainer & Owen, 2006).

In the scope of this thesis, parallel to the parse to learn – learn to parse arguments by Janet Fodor (J. D. Fodor, 1998a,b), we will refer to two different process: plan to learn and learn to plan. Plan to learn in this context indicates the usage of prior plans and performing them in order to learn more about the environment and to improve our domain knowledge of the world, which could be driven by curiosity about the environment and how things work. Learn to plan, however, indicates learning how to plan. This idea should not be confused with learning in order to plan, as learning in order to plan is in fact a part of the former process, planning to learn. In other words, after we consume prior plans in order to learn about the environment, we form new plans with our newly acquired knowledge. Learning how to plan, on the other hand, indicates a distinct process. Fodor claims that the parsing mechanism is innate in humans and we use this mechanism to acquire and grasp the inner workings of the language being used in our environment. Hence, the author asserts that we are not taught how to parse or we do not learn how to parse at the beginning, however we are capable of doing so in order to acquire the language. We are born knowing how to parse, it is just some parameters related to the language we encounter that gets modified during the acquisition. Parallel to Fodor's ideas, we investigate the nature of planning, which, by some authors, thought to be the inner structure that gave way to the development of natural languages. Plans can be mediated by verbal labels. In the same vein of Fodor's claims, plan to learn and learn to plan processes can be inquired. For instance, Steedman claims that the underlying structures of both the languages and plans could be the same and symbolic manipulation of verbal input might have developed on top of the planning mechanism (Steedman, 2002). When languages are taken into consideration, it is argued that not all languages are context-free. Context-free languages are equivalent to push-down automata. Therefore, push-down automata are not enough for simulating natural languages. For the recursive nature, an additional component to PDA is required for simulation. Embedded push down automata, which includes a stack of stacks, can simulate natural languages. Not only language, but also planning, reasoning and making inferences in a language require EPDA, due to their recursive nature (Steedman & Petrick, 2007). In particular, recursivity is inherent in collaborative multi-agent planning and communicative language. In this way, a tempting relation between sensory-motor action planning and language can be shown, if they can be both simulated by EPDA (Steedman, 2014).

Planning to learn seems quite crucial and beneficial in terms of survival. If we observe patterns in the actions we perform as causes and their effects to us or the environment, these patterns can help us in our further struggles with a planning problem. With the help of prior plans on domains similar to the domain of our actual problem or prior plans that had led us to similar goals, we can act deliberately and purposefully in order to achieve our goals. In terms of plans that include interaction with the objects in the environment, a great help is provided by the affordances, which will be explained in the next subsection. They help us be aware of the action possibilities in the environment hinted by their intended design functions. However it is crucial to note that there can be false affordances or hidden affordances, as well.

When there are fight-or-flight conditions, when someone is in danger or needs to take actions immediately due to a certain amount of time limitation, people may not demonstrate creativity in problem solving or planning. This is also related to the System 1 – System 2 proposed by Kahnemann (Kahneman, 2011). When the solution of a problem requires urgency, we make use of System 1, which helps us obtain quick answers to our problems, using certain heuristics. However, if we have time to ponder or the situation is not urgent, though it can be crucial, then the System 2, which is a slower but deeper process, starts to search for a solution.

This could also be the reason why in some cases, people are stuck at a state where they can think nothing but the intended function of an object, instead of thinking about new uses of that object, a situation called "functional fixedness" (German & Defeyter, 2000).Thinking about other uses of an object requires some free time. In addition to thinking about secondary uses of an object, it is also possible to create compound tools in such cases. On an occasion where someone is stuck because they cannot get past using a certain object in a certain way, it might be claimed that the attention on that object and the intention to use it in that certain way is so high that the person is temporarily incapable of thinking of other ways of using that object. System 1 is also related to our previous experience, therefore once we notice a different use of an object, it might be possible for us to utilize this function in our further problems.

We cannot think of a subject or an agent in isolation from the environment, as our plans, actions, movements all depend on the properties of the environment and the relation between the acting entity and the environment (Miller et al., 1960). Nonetheless, it should be noted that not all plans are for guiding the actions, there are also plans for gathering and altering information, as well as plans for making plans, i.e. metaplans. Planning appears to be so complex, involving complex structures and organizations, that, it is challenging for a simple stimulus and response arc to explain the inner workings of a process of forming and executing a plan.

We do not need to form elaborate plans at first. Cooking a meal involves many subgoals, preparing the ingredients, preparing the pots and utensils etc. We can refer to them with verbal labels. We do not need to go through all the details, just the rough sketches are sometimes enough.In addition, even though, we might have just enough knowledge about the environment, we may not act appropriately and reasonably in every situation. Our priorities, concerns or focus of attention might intervene with the formation and execution of a rational plan.

Plan can be descriptions that are part of our knowledge about the world. When they are transformed into control processes whose instructions are taken from the aforementioned descriptions, they can guide and control our actions, similar to how programs influence the processes of a computer. If there seems to be a gap between our knowledge database and the plan we try to execute, then we look for ways to compromise.

Plans are heavily influenced by our intentions, yet sometimes we have to perform actions that we do not intend to. In such cases, motivation and will must be taken into consideration. Attending to something, making an effort to understand its inner workings are evolutionarily advantageous actions that are driven by the high curiosity of animals. Although high curiosity makes the animals try out possibilities in the environment, if the animal gets the feeling that it is going into a loop and nothing changes, it might abandon that plan via stop rules, otherwise the loop can go on forever. The animal has a plan in mind and is willing to execute it, and this plan does not change much until the subject becomes bored with the repetition and looks for other ways of achieving the goal. As the effects of such actions cannot be observed, it would be hard to grasp the possibilities that object offers. Hence, the utilization of that object in a plan might be suspended until a meaning is attached to its behaviour. This means that we cannot alter our plans in accordance with the possible action-effect pair. However, we might remove the expected effect that occurred in similar previous trials, but did not take place in the current trials from our knowledge database temporarily, or sometimes permanently.

Plans are claimed to comprise hierarchical levels, similar to the Chomskyan levels present in the constituent analysis of verbal behaviour. Such an analysis might as well apply to actions. Planning mechanism and the utilization of some plans can be innate, therefore, the operation sequences observed in such cases can be similar. The same hierarchical structure, test and operate units can be found in similar plans.

Preparatory plans are the plans that we make before the realization of the actual plan. For instance, in order to open a door, we need make arm movements before even touching the door handle. These preparatory plans show our intentions and they make the subject and the environment ready for the actualization of the plan that would get the subject to the main goal.

There are several accounts regarding the formation of plans in humans. There can be several strategies used during the formation of a plan and also during the execution of it. We can make use of heuristics or we can exhaustively search a plan space. Heuristics are rules of thumb that we can employ to attain quick solutions. However, they do not always lead to desirable situations, hence they are risky, considered to algorithm-like search of the entire space of possible operation paths. Nevertheless, under certain circumstances, we need quick answers or short-cuts to our goals and in such situations, heuristics are of great help and importance. This is the System 1 that Kahnemann explains (Kahneman, 2011). Exhaustive searches can be used to get certain results. They are sure to help us reach a possible path from the initial state to the goal state. However, they can take quite a lot of time, even such applications might require indefinite amount of time to finish.

In forming our plans and particularly when executing them, we make use of two diverse but closely related process. One of them is the test phase, where we test the environment to see if we reached a certain goal state, or a subgoal state. Second one is the operational phase, where we act upon our environment to change the perceptual data that we get from it, meaning that making a change in our environment with the goal state in mind. However, it should be noted that operational part does not always include actions. Information gathering or transforming current information could also be part of the operational phase. If we consider the case where we perform an action in the operational phase, we expect to see some differences in our environment or in ourselves after the execution of this action. This is due to our previous knowledge about this action, its preconditions for execution and the effects that the operation of this action causes when the preconditions are present. Sometimes, the actions we perform do not cause the effects that we predicted beforehand. In such cases, we are puzzled and we try to understand why this could be the case. This is a situation where the intended action does not produce the expected outcomes. We might also try to perform the very same action several times just to be sure it is not producing the predicted results. This part might occur differently in different people because everyone might have different intentions, amount of motivation or patience. The number of operations performed in such a case is dependent on personal traits, as well. However, humans tend to stop trying after a certain amount of trials. Instead of being stuck in a loop, we might just give up and try different objects or different actions. Being stuck in a loop in this case means that the operational phase does not produce the relevant effects for the control to move on from the current test to the next. How do people know that they are stuck in a loop? For computers, understanding that a program is stuck in a loop is an unsolvable problem, named the Halting Problem. Therefore, humans probably make use of a certain rule to stop the inquiries once the number of operations performed, or the time passed etc. exceeds a reasonable value.

Our information about our environment, ourselves and the world, the universe all make up our Image (Miller et al., 1960). This designates the whole knowledge an agent has. When we are faced with a problem, first we need to grasp its requirements, its constraints. By collecting information about the problem, we make the problem a better-defined one, rather than a more ill-defined problem. This part actually corresponds to the modification of the knowledge the agent has. Formation of a plan might come later.

Regarding the formation of plans, it can be claimed that there can be several ways of forming a plan. We can make use of instinctual rules, heuristics or in a problem that requires motor skills, we can make use of imitation or the information we gathered from our previous explorations. Most of the plans we generate seem to be supervised by the Image an agent has. Sometimes we can make use of simple plans to construct more complex ones, as well. For simpler plans, we can even memorize a sequence of operations that we can recollect directly from the memory, by-passing a modification stage. In the modification stage, we can even talk about metaplans, which are plans to make plans. Here, it should be noted that we consider the plans to be hierarchically structured organizations. Parallel to the hierarchy that is observed in linguistics, the plans also seem to demonstrate hierarchy. There can be several subplans embedded into another greater plan. In such compound plans, it would not be easy to store the whole plan in the memory, rather, we might prefer a more economical way of keeping them in mind. For instance, when we think of something that we need to do on a certain day, we do not think of every single parameter that might be involved in the execution of our plan. This is partly because we can use our verbal abilities to name the plans by separating them into chunks. For instance, if we would like to perform a job that includes several subtasks, it is enough to name it and work the minute details during the execution of the plan. Otherwise, keeping a whole list of plans with their details would impose high cognitive load that prevent us from performing and this would mean that we are just observers and information-gatherers, rather than actors and experiencers in our environment. Cognitive map theory is focused on the former part of gathering and organizing information. However, it is not explained how the rats make use of such an organized map of information about the environments and themselves. Still, he advances upon conditional learning by adding a structure that plays a role in information processing. Gathering information or updating the information that we already have is the part where we modify our whole knowledge. However, planning does not simply consist of that component. It is, of course, a required component for the agent to be able to act sensibly in its environment. Nevertheless, it should be noted that the execution of plans just as the instructions in a computer program and the control moving from one operation to another is a whole different kind of process.

In cases where the plan is rather complex, we can make use of a previously used plan, and modify it to match our intentions. This modified plan would supervise our further actions. It must be noted here that our motivations and intentions lead our way towards the goals. However, these should not be confused with the reward system found in reinforcement learning. Following is very nice example by Lewin as recounted in (Miller et al., 1960): If we need to mail a letter, the priority of a mailbox would be very high and after we put our letter in a mailbox to be sent, this priority would be of less importance, it may even be quite unimportant. However, if consider the law of effect in the context of reinforcement learning, we would have expected its value to get higher, as it produced desirable results for us. Nevertheless, once we mail our letter, finding a mailbox is no longer a concern for us. This can be explained by taking motivations into consideration. Mailboxes are no longer a concern because we do not need them any more.

Instead of stochastic systems, we can think of plans as deterministic hierarchical systems. Human planning and problem solving are complex processes that show structure and organization. Such problems of organized complexity require the utilization of high cognitive processes and they allocate an important part in the working memory while they are being worked out. A basic planning unit comprises of two stages: test and operate (Miller et al., 1960) and these units can be nested into each other or they can be concatenated in order to account for more complex planning problems.

In the test stage, the focus is on whether a goal or a subgoal is achieved, if not, operate stage gets the attention and the agent generates and executes a new action, whose consequences are then evaluated in the test phase again. This is a circle that encompasses updating the world knowledge, confirming or infirming data. When an unexpected effect is encountered, subject becomes puzzled and tries to overcome this unexpected effect via constructing new plans, which would then control the instructions. The loop between the test and operate stages can take forever, hence, it is crucial to have a stop rule or a heuristic. If a number of operations or a certain amount of time is exceeded, agent is encouraged to try other options. In this part agents should exhibit low level creativity, removing constraints from familiar ones.

Planning is a vital part of human behaviour which makes use of experience and learning past events and future possibilities, automated planning is therefore, a significant part of Artificial intelligence. The reason why it is studied in AI is, of course, as it is a very crucial part of human behaviour. In addition, a second reason is that with the help of automated planning, many tedious works can be very practical (Nau, 2007).

In his book on automated planning Ghallab defines planning as the reasoning side of action, where one usually deliberately selects and orders actions by taking the consequences of these actions into consideration in order to attain a goal (Ghallab et al., 2004). Automated planning is a branch of AI that covers a broad range of applications from optimization and scheduling to games and dialog planning (Ghallab et al., 2004). Automated planning is concerned with the development of tools that are capable of making plans efficiently. Currently, such tools are being integrated into many complex systems such as spaces vehicles or robots. The focus in the automated planning research is to create safe, affordable and efficient planners that can be utilized in practical use, rather than in theory. However, such planners might not reflect what goes in our minds during planning. Our plans are not always efficient or safe. Indeed, they might leave possibilities for inducing risky situations. Hence, it can be claimed that not all planning activities that take in humans is rational. Human behaviour can be quite irrational from time to time, particularly when there are constraints that limit contemplating the situation we are in, such as certain dangers to our existence or time (Kahneman, 2011). The study of rational planning begins with "classical planning", which is a quite inflexible form of planning. Nevertheless, by relaxing certain variables, we are able to go beyond classical planning. The representation of a classical planning problem includes several indispensable components, which were explained before. Classical planning utilizes search algorithms to scan the state space to find a plausible path.

In addition to the classical planning applications, there are also planning with incomplete information and sensing and planning under uncertainty or with limited time and space resources. Since the uncertainty and incompleteness are utilized, probabilistic planning applications are also significant. The world and the state variables can be fully observable, partially observable or unobservable. Markov Decision Processes (MDP) can be utilized in fully observable cases. Since not all problems in certain domains are fully observable, Partially Observable Markov Decision Processes (POMDP) can be employed to solve the problems in partially observable situations. Nevertheless, it should be noted that the complexity of POMDPs would be very high and such problems are generally intractable (Even-Dar et al., 2012). Moreover, the length of a plan is exponential and the number of results and states could also be exponential. Planning applications can incorporate cost optimization or preferences. Whereas some planning applications do not make use of temporal properties while generating plans, some do employ time related components into the process. In addition to such aspects, plan generation requires good searching capabilities and planners should be able to produce hierarchical plans in order to alleviate the high complexity requirements of uncertain or unknown worlds.

In this thesis, we focus on structured problems, and with the help of our prior plans to such structured problems, we can solve new structured problems that are similar to the prior ones. Most of the problems we encounter bear structure. Hence, the agents would be learning about their environments in order to attain a profound understanding of how they can act and how they can generate plans. Moreover, the incorporation of structure would results in a decrease in the complexity.

Current state of the world in a domain is described by the state variables and operations transform one state into another. State space is the collection of states and it indicates possible, feasible or reachable paths from one state to another. State transform operations denote actions that gets the agent from one state to another. Planning in Artificial Intelligence is concerned with finding the sequences of actions via state space search that includes possible paths that would help the agent to reach a goal state from an initial state. This description of planning illustrates the main components of automated planning in AI. Classical planning is the kind of planning performed in the presence of complete information and certainty. It can be asserted that this kind of planning involves the formation of hierarchical plans.

Multi-agent planning is also another application area of the automated planning systems, where the agents participate, collaborate, and compete to consume the resources available to plan ahead (Weerdt et al., 2005).

Automated planning systems can be of three types: First plan then learn, where the agent generates plans to speed-up the learning stage, first learn then plan, where the agent learns about the values of states and then plans about them, finally systems that interchange planning and learning.

In addition to the classical planning applications, there are also planning with incomplete

information and sensing and planning under uncertainty. Since the uncertainty and incompleteness are utilized, probabilistic planning applications are also significant. Planning applications can incorporate cost optimization or preferences. Whereas some planning applications do not make use of temporal properties while generating plans, some do employ time related components into the process. In addition to such aspects, plan generation requires good searching capabilities and some planners utilize certain heuristics during search. Plan execution and the evaluation of the consequences each time an action was taken can also be added to a planner system. Hence, a planner can also perform the actions that it planned.

PKS, planning with knowledge and sensing, is an automated planning software and in order to define a planning problem, we need to determine several components pertaining to the agent's knowledge of the world and its capabilities similar to the situations, fluents and actions defined in situation calculus and the STRIPS notation for planning (Nilsson & Fikes, 1971). The definition of planning problem in STRIPS notation is denoted by the tuple consisting of the initial state, a set of possible actions and a set of goal states. STRIPS is the basic representation for planners, however, there are also STRIPS-like languages involving some other properties (Russell et al., 1995). Newell and Simon's definition of a problem space also comprises a similar tuple: a set of elements, a set of operators, an initial stage, a problem with a goal, and the knowledge available (Newell et al., 1959).

First, we must define a set of properties that represent the objects in the world and agent's knowledge. These include conjunctions of positive and negative facts, propositions, about the state the agent is in (Kf database in PKS), functions values, sensing effects (Kv), binary sensing effects (Kw) and exclusive-or facts (Kx) (Petrick, 2004). Another component is the set of state transforming actions that indicate what the agent is capable of given the conditions it is under, that is, preconditions, or the current state it is in. Two other crucial components are the previously mentioned "current state" and the "goal state" that the agent should arrive at after a series of state transitions. Current state consists of the initial world state and the agent's initial knowledge database. Goal state defines several conditions that should be satisfied at the end of the state transition process. After taking these planning components as inputs, automated planning applications generate a sequence of actions starting from the initial state and transforming it into the goal state at the end, PKS uses forward chaining and it is capable of performing depth-first search, breadth-first search of iterative deepening depth first search (Petrick, 2004).

Frame problem is one of the intriguing problems in AI introduced by McCarthy & Hayes (1969). Frame problem is intertwined with the concept of planning, as it deals with action effects. When we take an action, how do we know what effects would follow and which parameters would remain the same? PKS utilizes the STRIPS assumption in terms of implicitly overcoming the frame problem. This assumption indicates that if the database does not contain a certain literal, that literal is assumed to be false. Additionally, add and delete functions represent the effects of actions, therefore, it is assumed that every change that might occur is explicitly stated in the plan configuration. In PKS the assumption about the changes after the execution of an action comprise only of the effects that are defined in the related actions in the plan file takes place, when the preconditions of an action are met and the action is selected to be performed. If there are no explicit declarations related to a change, then it is assumed not to have happened. This is also called the "commonsense law of inertia". Because of the heavy cognitive demands of searching through all the variables in our vast database for world knowledge and the possible changes that might occur to them, humans seem to focus on a handful of variables after an action is performed, based on the location of attention and the intentions of the agent. In addition, humans are quite faster when noticing patterns, symmetries or obstacles in a problem-solving situation than computers. The attended variables are the ones that are relevant to the execution of an action, however, determining which of the variables are relevant and which are not is also a concern in planning in AI as no one would like an agent entering a deadlock, just as the robot Dennett described, while trying to find out about the relevant variables or identifying if this object is the same one as the one a second ago (Dennett, 1987a).

3.2 Affordances

Although concepts similar to affordances and action possibilities were first described by the Gestalt school (Uexkuell, 1920), the most seminal study on affordances were published by J.J Gibson. In fact, the term "affordance" was proposed by J.J. Gibson from the verb afford to indicate what is offered by the environment to the animal as possibilities for action (Gibson, 1977). In his view, actions that are available to an agent in a certain environment considering certain objects indicate the affordances that these objects provide. This idea reflects the ecological view on affordances, which argues that the perceived environmental situations are significant in planning and execution. Ecological psychology is concerned with the wayfinding, navigating behaviour and the subjects' actions by taking geographical attributes such as perceptual cues, landmarks or affordances of the environment into consideration.

According to Gibson, affordances can be directly perceived, meaning that when an agent looks around in the environment and scrutinize the objects available, then he or she is able to directly observe the actions that are possible to conduct on them. It is claimed that we can directly perceive the meaning of an object in the environment, that is, we can directly grasp what we can do with it or to it. This is called direct perception. Without noticing an object's attributes separately, the meaning is holistically obtained. The possibility of an action is directly comprehended. However, this might be problematic because the agent might at first explore the possibilities and then discover certain constraints on the actions it can take under those conditions. This is particularly the case for objects that bear abstractions, or that are related to certain concepts in the human world. It is hard to assert the affordances provided by such objects are directly perceptible. The idea of direct perception has been criticized by some authors who assert that affordances are not directly perceived and Gibson's theory lacks the involvement of cognitive processes and opposes the computational theories of perception (J. A. Fodor & Pylyshyn, 1981).

One of the arguments of Gibson, which was also criticized, is that affordances comprise all possible actions in the environment that are available to the agent with respect to the objects in the environment. If this was indeed the case, then choosing appropriate actions that would lead the agent to the optimum solutions towards the goal should not ever be a problem in planning. However, if the agent is not aware of even a tiny component in the environment, it might not be possible for him or her to consider an action related to that component in that scope. Hence, affordances should be better phrased as the actions available to the agent with respect to certain objects which are actively observed by the agent, meaning that the agent is aware of the presence of the object. Therefore, it would be possible for the agent to make use of that object by acting upon it. Here, one must notice the crucial part that attention and focus play in deciding which actions should be taken. If the agent pays his or her attention to certain objects, by taking working memory capacity into account, as well, it may not be possible for him or her to detect other objects and hence other possible actions. Short-term memory sometimes works on the basis of objects. In a case where the agent is actively trying to explore the environment, even though we may not talk about a fixed attention point, as the awareness is increased, it would be easier for the agent to detect possible actions.

Affordances are the attributes of neither the agent nor the object. They exist only in the presence of both and arise through the interaction between the environment and animal. Affordances help agents identify the actions that they can take in an environment. Affordances are claimed to be an intuitive notion in the physical world (St. Amant, 1999). We can also classify affordances as simple affordances, precondition execution, effect execution, precondition evaluation, and effect evaluation affordances(St. Amant, 1999). Affordances in user interface designs are also an important research topic in Human Computer Interaction. Norman studied common objects that we can interact with in our environments and the operations that are possible to be performed with or on them (Norman, 2002). In user interface studies using Norman's principles, UI components are designed in a way so that their functions would be ostensible from their appearances. It is sometimes asserted that "Form follows function", nevertheless some functional properties of objects might not be self-evident in terms of their physical appearances.

Functions of objects are determined by their physical characteristics in collaboration with the attributes of the agent and the environment. For instance, if the perceptual information we get from an object indicates the possibility of an action, while it is not actually possible to take that action in that environment, than this perceived affordance is said to be a false one (Gaver, 1991). If both the perceptual information and the real capabilities are consistent, then we associate a perceptible affordance with this object. It is also possible for an object to offer hidden affordance, where the possibility of the action is not visible from the perceptual information. Correct rejection of an affordance is the case where the object does not afford an action and the perceptual information does not cue the possibility of that action.

Therefore, for different agents, interaction with the same environment might offer distinct affordances. For instance, for a baby, a toy on a high shelf is not reachable, but it is reachable to her mother. Therefore, what an environment offers the subject is not constant, since it also depends on the subject's abilities of acting in that environment (Witt, 2011). Judgement of affordances, plan formation and execution can be affected by the agent's motor capabilities, perceptual capabilities, information-processing capabilities, prior knowledge, mood and motivation. For instance, a wall seem be climbable to someone of the height 180 m. and not climbable to someone of height 130 m, even though they can both climb if they actually try. Ratios of legs to the body and to the height of the wall are all factors in the difference between the affordances. Instead of some behavioural metrics, these parameters might offer better insights into the influence of motor abilities in perceptual judgements (Witt, 2011).

It is also necessary to consider the crucial relation between action and perception. Although the impacts of perception to action are more deeply researched, it is also possible to consider the effects of self-generated action on perceptual judgments (Schütz-Bosbach & Prinz, 2007). Theory of event coding asserts that perception and action are coded in a common representation, and thus, they are integrated to each other while also sharing resources, which is the main assertion of the ideo-motor principle (Hommel et al., 2001). Supporting evidence for this theory falls into two categories. First, induction paradigms, where one of the components induce the existence of the other. Second, inference paradigms that indicate the mutual interference between perception and control of action (Prinz, 1997). An important principle that is related to the common coding approach is the action effect principle, which maintains that during action planning and execution, the representations of expected effects of actions play an essential role (Prinz, 1997). This is particularly the case for action planning that involves data collection from sensorimotor units where the planning of goal-directed action benefits from perceptual information immensely (Hommel, 2005). If we take object properties and affordances into consideration, this idea seems to match the process of action planning by assessing the perceived affordances and expected effects of actions. Therefore, action selection, evaluation and execution are interwoven together with the previously encountered perceived and produced events (Hommel, 2005).

Different formalizations of affordances exist, for instance, Steedman's formalization of affordances indicates a relation between affordance-mediated planned action and natural language grammars (Steedman, 2002). Steedman bases his claim on two operations, which are shared both in language and plans, functional composition using combinators and type-raising. Hence, he asserts that the underlying mechanisms for both languages and planning seem to be intertwined, and languages might have developed on top of a primitive motor planning mechanism. Stoffregen defines affordances as emergent properties of the animal-environment system, and not the properties of the environment only (Stoffregen, 2003). Chemero defines affordances as the outcomes of the relations between the features of the environment and the abilities of animals (Chemero, 2003). These definitions are also highly related to the ecological approach.

Affordances play a crucial role in robotics particularly in determining what actions can a robot might take and which ones it should choose. As it can be clearly noticed, this component of action selection is related to planning. Affordances in AI represent a methodology to create computer models that are capable of perceiving affordances in their environments and making use of them so that it would act effectively and quickly (Nye & Silverman, 2012).

As it was explained in the previous chapter, humans are not merely passive learners. They learn through active interaction between each other or between themselves and the environment. Active learning helps to construct their own meanings of the environment and what takes place in the environment. Affordances are quite important in problem solving and decision making particularly in problems involving objects in the environment. It is possible to predict the effect of an action on a certain object. If we are given the object and the effects, we can also indicate what action might have caused this. In addition, if an action and its effect is given, we can recognize a certain object that this action-effect pair is closely related to (Montesano et al., 2008).

Subjects can learn affordances through social interaction by observing the actions of others or they can learn about them alone. The first case indicates imitation. However, if the animal observes the changes in the environment after an action has been performed by another, it might learn about the change itself, rather than the observed behaviour. Hence, it can be possible to learn about affordances by simply observing the changes after actions. Byrne proposes imitation to be a process of behaviour parsing, where the animal observes others' behaviours and tries to decompose it into statistically regular behaviour units (Byrne, 2003).

Objects that we use in daily life to achieve certain goals are designed and produced in order to achieve those goals. When we look at an unfamiliar object, we might try to guess what we could do with it or upon it. By making use of our prior information on familiar objects, actions that we can operate on them and the possible effects, we might infer the object-action-effect relations of novel objects, as well. Human-made objects have intended design functions. The concept of design stances by Dennett indicates that the fact that objects are designed to fulfil certain requirements of reaching a specific goal determines their behaviour, features, category that they belong to and existence (Dennett, 1987b), (Vermaas et al., 2013). We can also call them primed affordances. They signify the actions that the objects are primarily designed for. For instance, primed affordance of a knife is to cut. When we see a knife, we expect it to be able to cut certain objects or when we see a cup, we expect it to be able to hold liquids and also if it has a handle, we expect that we can hold the cup, as the primed affordance of the handle is the grab action. Therefore, intended design functions play an important role in our plans. If we see such objects under different conditions, we might also imagine other ways of using these objects. In such cases, our ideas of using that tool could depend on the requirements of a goal.

In cases where we try to use familiar objects, we can be primed about the actions to take under certain conditions. For instance, a door separates two rooms from each other or maybe the inside of the house from the outside. However, if we think more about the properties of a door, we can imagine several other uses of a door, as well. For instance, a door is a solid that mostly has a plain surface; hence, it could serve other purposes, as well. One can use it as a plank to sit on or make it a table top, or sail on it, of course when taken out of its hinges. Therefore, it is possible to imagine novel ways of using an object and an object has the potential of being used to accomplish distinct goals in distinct contexts. Humans are capable of inventing or discovering ways of making use of an object via considering its categorical features. This is specifically demonstrated when someone modifies an object or makes use of several objects to reach a goal. Here, what is important is that under certain conditions, some actions might be primed, meaning that the agents would notice certain affordances when compared to others, which might also be available to them, nevertheless misses their attention. When someone sees an object several actions may come to mind; however, if they are asked about other approaches or forced to take different actions, one might start to uncover other feasible actions, as well.

It is also possible for humans to look for objects while having a goal in mind. It is argued that this is what most of the animals lack in terms of tool use and planning. Even though some of the animals, especially primates, can utilize different tools to achieve a goal, they might be oblivious to the objects that are actually available to them but not currently in their perceptual views. In a case where they need to find a stick to reach food, even though they might have examined a stick-like tool in another room, they fail to go back to that room and fetch the stick (Koehler & Winter, 1926). Hence, they fail to make deliberative plans, and perform reactive planning. Another issue that the animals might lack in planning is the awareness of others and their mind states. This is related to the Theory of Mind, being aware of others' goals and intentions, and their knowledge, perception and beliefs (Call & Tomasello, 2008). An agent might be aware of other agents in the environment; nonetheless, this awareness is not enough, as it would be more beneficial if the agent can also guess what the others know about the environment and what their motivations would be. This requires an embedded Push-down automaton utilizing a stack of stacks.

We can relax the limits and start to see different usages of the tools. For instance, even though the intended design function of a knife is to cut, we can also throw a knife, hit some object with a knife, stab or punch into something, we can use its sharp tip and so on. We can use a cup for reasons other than holding liquids. However, this might not be the case under certain conditions. Sometimes, we are stuck on the idea of using the object to perform the intended design function. This is called "functional fixedness" when the subjects are fixed on the primary function so that they are not able to imagine other ways of using that tool immediately. The most seminal study conducted on functional fixedness was the candle problem (Duncker, 1945), where the subjects were either failed or very slow to find a solution to the problem of fixing a candle on the wall using only a box of tacks and matches. For instance, the box of tacks give them the impression of being a container, rather than a support; hence they fail to generate alternate constructions. When the tacks and the box are provided separately, however, subjects were more likely to find out the solution faster. Functional fixedness is claimed to occur less in children, as they are more goal-oriented and their concepts for objects' design stance are not yet developed and grounded as the adults (German & Defeyter, 2000). It was also shown that children categorize objects according to the intended design function of an object, just as adults did. However, children's judgment of the object's function was dependent on the goal at hand, whereas the adults preferred the design function (German & Johnson, 2002).

Functional fixedness could be a feature whose continuity was provided by the evolutionary processes. Particularly in cases where the disabled object functions are related to the goal somehow i.e. achieving it at an instant or achieving it ultimately, functional fixedness seems to occur for longer times. If we can overcome this constraint, then we are able to envisage using this tool in other ways or as a means of accessing another tool that might help us attain our goal. Hence, this is an important leap of ideas. We can use the object only in very few ways, if we keep thinking about its intended functions.

In a case where the intended function of an object is not available or somehow disabled, then we might be forced to think other ways of using this tool. If the scissors that we try to cut a piece of paper with do not cut the paper, then we start to realize that there is something wrong with the scissors and we start looking for other means of cutting that piece of paper. These ideas come to mind because object-object interactions and the relations of objects forming objects pairs also affect our decisions in making plans (Sun et al., 2014).

When the primed affordances are unavailable, this means that the everyday usages of objects is limited. In such cases, the object might be freed of concepts related to their design stance and intended design decisions. That is to say, they are now perceived in terms of their physical properties, rather than other concepts that might be attached to them. In such cases, we need to reformulate our action-effect pairs that are attached to this object. For instance, when we expect scissors to be able to cut, the cutting action that we can employ in our plans has the effect component in which the object being cut also shows some changes in its form. We can use this action-effect pair in our pre-plan. However, when we see that the scissors do not cut, we update our action-effect pair to show that this action on this object does not produce any effect.

In the scissors-and-paper case, at first we assume that we can cut with the scissors, as this is a perceptually salient affordance. Nevertheless, it turns out that we cannot apply the cut operator using that pair of scissors, we then turn into other uses of the scissors. On such an occasion where we cannot move further in our inquiries, we take other action possibilities into consideration. Here, we must also attach importance to the object categories a certain object belongs to. For instance, a door is used to get from one room to another, or from the inside of the house to the outside. Other than opening, closing or going through it, we can also consider its physical aspects to find out about the actions that a door affords. Doors are solid, large or small, wooden or metal planks that can be used as a table top or even as a very simple sail boat, if they are made of buoyant material. In such cases, an appropriate action that would lead us closer to the goal would be selected in terms of the physical categories that an object belongs to.

This part is where we use our plans to learn the environment and to achieve a goal. In a case where the objects behave randomly or in a way that defies the laws of physics, it might not be even possible to make use of prior plans that might lead us to the goal. In such a setting, one must first learn to plan. If the objects are familiar, then the subjects would think of the behaviours of similar objects. However, novel objects would be approached cautiously, inspecting their physical attributes more deeply. In addition, random or unlikely situations also seem to falter our planning mechanisms.

If we consider the affordances and their use in plans in the light of these claims, it can be said that they can be a part of the instinctual plans, plans that require experience in motor skills, or plans that can be generated using similar plans that were generated before. We can also think of them as patterns about the effects of performing an action, possibly on a certain object. These patterns are expected by the agent because these effects were observed in the past or they are expected to happen. If we go back to the affordances, when we act on an object that seems to afford an action to us, we expect that action to happen. In addition, our perception might even be altered due to this expectation of an action happening. It should be noted here that for different agents of distinct attributes, the same object can offer different affordances in the same environment. An agent might be able climb over the fence and another one might not. For that reason, what an agent can perform is defined by both the agent's, and the object's and environment's properties. For a tennis expert, the ball might seem bigger, or for someone who is experienced at climbing over various sizes of walls, a wall that appears too tall and hence unclimbable to a novice agent, can appear to be of reasonable height. These differences are explained in terms of action specific perception and perception of affordances based on expertise. Hence, it is essential that during the learning process, planning and affordances provide us with crucial information about the world and affect our judgements.

CHAPTER 4

METHODOLOGY

4.1 Research Questions

In this study, psychological claims about the nature of the learning and planning behaviours taking place in problem-solving stages is studied from a computational approach and an Artificial Intelligence perspective. Most importantly the interaction between learning and planning is taken into consideration.

In the Machine Learning and Automated Planning literatures, one can find out about the computational models involving both learning and planning. There are Computer Science studies about integrating planning and learning, but they use mostly the reinforcement learning algorithm such as Dyna-Q (Sutton, 1990). Nevertheless, the research conducted in these areas are focused on developing rational agents that are capable of producing results in under a handful of minutes. This is not to say that there is no study directed towards simulating the learning and planning processes in humans. Latter case is a whole different situation, where the researchers must also consider the human nature, according to some, the behavioural data, according to some others, introspection and so on. Building psychologically adequate and functionally correct models of learning capabilities that also involve planning processes requires a good knowledge of theory in psychology, since one must take them into account while developing an agent whose information processing systems and behaviour are relatable to the human information processing system and behaviour.

Our concern is about the nature of plans. In cases where everything happens congruently with our expectations or predictions, it is usually easier for us to act. This is because there is nothing much conflicting our knowledge about our environment or ourselves. However, under different circumstances, our expectations or predictions may not turn out to be true. In that case, we become puzzled as some things do not follow the pattern. Here, we use the term pattern to indicate action-effect pairs on objects, which were previously observed and utilized in plans and proved to be feasible when we performed those plans. These action-effect pairs on objects would be called affordances. When we expect to cut a piece of bread using a knife, and when it turns out that particular knife is not actually capable of cutting the bread, we need to update our knowledge about that knife, or the bread. Its sides could be blunt, or it could be plastic, or maybe the bread is too hard to be cut. Hence, the cutting action on these objects do not provide the desired effect. People are adept at updating the action-effect pairs. However, they might go back to their previous schemas about the knife or the bread soon after they see another knife cutting another bread.

When the constraints imposed on objects also allow the humans to update their knowledge they collected during their past experiences, humans can learn more about the environment and what sorts of outcomes they would get when they perform an action on a specific object. If an action-effect pair seems to be valid for an object most of the time, then this new pairing would become the rule.

Hence, if an object gives us the idea of affording a certain action, but when we try to perform that action on the object and we do not observe the intended consequential effects, it can be said that this object provided us with false affordance at the beginning. This is also the case for the knife that does not cut. As knives are quite familiar objects, which are designed to cut, we can see that something is not right about a knife that is not capable of cutting. At that stage, we are fixed on the function provided by the design intention. However, after a while we can get used to the defective condition of that knife and do not consider making use of it in our plans that might involve cutting. We might even consider using it for other purposes that seem novel to us. Nevertheless, we were drawn to the knife to cut at the beginning, probably because our previous plans that guided our actions to achieve similar goals made use of knives, scissors or other sharp objects that might help us cut. While taking an action, we were supervised, in a way, by our previous plans.

If we consider a situation where all the colours are inverted, or the gravity is in a different direction, it might seem strange to us. Nonetheless, although it might feel weird at the beginning, we may get used to them after a while. However, if the colours or the direction of gravity keeps changing seemingly randomly, it would be hard for us to adapt to that environment. This might be due to a lack of patterns for us to observe in the environment. When we consider a situation where objects seem to be moving randomly, or at some point they work, at some point they do not, it would be hard for us to consider utilizing such objects in our plans. In fact, in such cases we might not even be able to form plans considering the ostensibly faulty nature of those objects. This does not mean that there is no path from the initial state to the goal state, there can be one, but we are not able to notice it. Due to the existence of unexpected constraints, the complexity of finding out a sequence of operations that would lead from the start state to the goal state becomes larger.

In this thesis, we focus on making use of our plans as supervisors to learn about the objects in our environment and to update our knowledge to produce reasonable plans, develop novel ways of utilizing familiar objects. In contrast, we also consider the possibility of learning how to make plans versus an innate mechanism of making plans. In both of these concerns, we also take the learning theories into account and consider the shortcomings of reinforcement learning in terms of the lack of complex information processing stages and cognitive map learning in terms of the lack of explanation for the execution of plans rather than the storage of information. For this aim, we conducted an experiment and developed a computer model to compare the aforementioned differences in planning under different circumstances where our expectations and predictions do not turn out to be correct. In the experiment, we aim to compare the number of operations, the number of trials on affordance-bearing object pairs, the number of operations between certain milestone states, such as trying out a pair with related affordance and then turning the attention into subgoals. The details of the variables will be explained in the following section.

4.2 Methodology

This thesis includes two different parts in terms of methodology, they will be explained in this section separately. First part consists of the experiment stage, where we recorded the computer screen as a video while the subjects played a 3D First-Person-Controller Game. This game was a simple object interaction game where the players need to make use of the objects in their environment to achieve the goal they are given at the beginning of the game. The purpose of creating this game was to see which actions the subjects would take, how they would explore the game environment and how they would plan their ways to reach the goal. There were 2 different types of the game whose details will be explained shortly. Second part consists of a computer program, which is modelled after the components of the game. The model simulates the environment that are present in both of the games. Hence, there are also 2 parts of the model corresponding to the 2 different games. Computer model, just as the players, can act on the environment, make use of the objects to reach the goal that is instructed to attain. The implementation details of the model will also be explained in this section.

4.2.1 Experiments

For the experimental part, a 3D First-Person-Controller game environment was created using Unity. There were 2 different types of the game corresponding to Experiment 1, where we investigate the concept of planning to learn, and Experiment 2, the concept of learning to plan. The following explanations apply both to Experiment 1 and Experiment 2. The subjects are instructed to play this game and encouraged to think aloud, while the screen is captured as a video and the eye-movements of the subjects are collected. The environment consisted of 3 rooms and the player starts from the biggest one. The biggest room has a table inside and on the table, there is a piece of paper, a key, and scissors. This biggest room has two doors that connects it to the remaining 2 rooms. One of these doors is made of glass; therefore, the player can see the other room and its contents through the glass. The other door is wooden, preventing to see inside. The room with the glass door has a table and on the table, the subjects find pliers. Inside the room with the wooden door, there is a glass box, inside which rests a knife.

As the objects would seem familiar at first, their 3D images on computer screens would imply certain affordances to the subjects, nevertheless, some of these would be false affordances, as the predicted action possibilities would not be available for the subjects to utilize. Images of the objects do not afford the actions that would be performed upon them or with them in real life. Nonetheless, as the experiment stimuli consist of 3D objects and a 3D virtual environment in which the subject can interact using the keyboard or the mouse. Therefore, the objects' affordances would be phrased such as: "Dragging this 3D image of this object with the mouse and dropping it on a 3D image of another object would let the latter to get broken."

The user is given the goal of separating the paper into two. To avoid confounding the results, the goal was phrased in this way rather than giving the instruction as "the goal is to cut the paper". At first, the subjects are given the chance to familiarize with the environment and the controls. Moving the mouse around moves the in-game camera to change the point of view of the player. By using the left, right, up and down buttons, the player is able the move the First-Person-Controller around. It is also possible to make use of the A, D, W and S buttons,

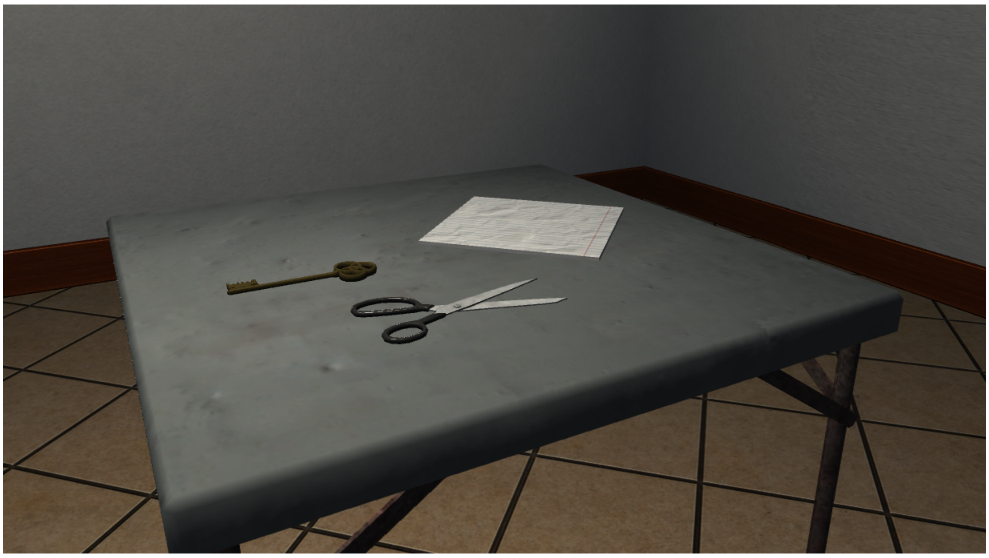

Figure 4.1: A screenshot from the game

respectively.

Except for the tables, doors and a box, all remaining objects (paper, scissors, key, pliers, and knife) can be grabbed and dragged around and dropped by the mouse controlled by the player. By moving the pointer on an object and left-clicking the mouse, the use can grab the object and by moving the mouse around he or she can make the object move on the screen. A grabbed object is dropped by releasing the left mouse button. It is enough for two objects to touch if something is expected to happen. This means that the player does not have to try different ways of bringing objects together. The subjects are informed about all of these before the experiment takes place. The object pairings that would produce effects in the first experiment are listed below:

- **Scissors + Glass Door** Glass is broken to allow entrance into another room, in which there exist pliers on a table
- **Pliers + Wooden Door** Wooden door is opened to allow entrance into another room, which contains a glass box. Glass box contains a knife.
- **Scissors + Glass Box** Glass box is broken to pieces to reveal the knife
- **Knife + Paper or Paper + Knife Paper** is separated into two pieces.

As it can be seen scissors do not cut the paper and the key does not open either of the doors.

The optimum operation sequence to achieve the goal is given below:

1. Break the glass door using the scissors.
   1.1. Grab the scissors
   1.2. Make the scissors and the glass door touch

    1.3. Drop the scissors

2. Get the pliers and open the wooden door.

    2.1. Grab the pliers
    2.2. Make the pliers and the wooden door touch
    2.3. Drop the pliers

3. Get scissors to break the glass box.

    3.1. Grab the scissors
    3.2. Make the scissors and the glass box touch
    3.3. Drop the scissors

4. Get the knife to separate the piece of paper into two.

    4.1. Grab the knife
    4.2. Make the knife and the paper touch
    4.3. Drop the knife

In the 4th stage knife and paper can switch places.

In Experiment 2 where we investigate "learning to plan", all these parameters are kept the same. However, this time some constraints in grabbing and dragging the objects are imposed. Scissors are made too heavy to easily grab and drag around. When the player grabs them, he or she is not able to see if the scissors is grabbed and moving around. Pliers do not open the wooden door at the first instance. It requires 3 trials. Grasping the key produces random movements in the first 6 trials. Then, the control returns back to the normal way of grabbing and dragging. The knife in this case also behaves differently. It can move in two dimensions but not along the x-axis. Therefore, it is not possible to take it out of the room it is located when it is found. Hence the paper should be grabbed and brought near where the knife is located. The optimum plan is the same.

It should be noted that although it seems very hard for the player to find a proper plan that would lead to the goal in Experiment 2 due to the constraints on the interaction between the player and the objects, there is actually a feasible solution.

A pre-experiment study conducted with 5 subjects by employing the earlier version of the game showed that there was a need to improve the 3D environment and the way the users interact with the objects in the environment. For instance, some subjects mistook the glass door with a mirror, hence, they avoided that part of the room. In addition, in most cases the subjects missed the point that it is enough for the objects to touch each other for an effect to occur. Therefore, the players in the real data collection phase are informed about this.

Different subjects participated in Experiment 1 and Experiment 2. Experiment 1 took about 5 minutes to 10 minutes, whereas experiment 2 took about 10 minutes up to 25 minutes.

4.2.1.1 Results

The video recordings of each subject's gameplay were watched annotated in terms of the actions taken by the subjects. All of the gameplays followed almost the same pattern. This pattern formed the basis of the algorithm utilized in the computer model. After the annotation, the number of operations required to reach certain stages that crucial in the game were calculated.

The first parameter that we compare is the number of operations performed in each session. The number of object pairs that indicate the intention of performing the primed function were separated into two parameters: The first one is the number of scissor-paper trials and the second one represents the number of trials with one of the doors and the key.

By taking the pattern observed in the gameplays, the number of operations that led the subjects towards the first trial of the goal object (paper) and primed object (scissors) pair that is directly relevant to the goal was another parameter. In addition to that parameter, the trials of other objects on the goal object was also investigated. After this phase, subjects move on to doors and create subplans to get rid of them. Hence, another parameter indicated after how many moves the subject directed his or her attention towards pair involving one of the doors and the key. When the combinations that a subject tries do not provide fruitful results, the attention is moved towards unlikely pairs. The number of actions taken after trying out the key-door pairs until the scissors is tried on the glass door was also a parameter. Some subjects after being able to grab the pliers took it directly to the paper and some to the wooden door. The number of actions between breaking the glass and unlocking the wooden door was also calculated. In order to check if the subjects would remember they broke the glass door with the scissors and would try to employ it to break the glass box, the number of actions taken between the unlocking of the wooden door and the breaking of the glass box is also recorded. Finally, how many operations the subjects performed after the knife was revealed is also another concern.

Seven subjects participated in the first experiment and seven different subjects participated in the second experiment. As the number of participants was not high enough, applying non-parametric test for the analysis of the data is considered. However, just to be sure normality tests and homogeneity of variance tests were also applied by taking each variable and comparing their values in the first experiment with the values in the second experiment. For instance, at first, the variable number of operations is selected and factored according the experiment number. The results of the normality tests indicate that although some data seems to have normal distribution, some others do not. The results were similar for the homogeneity of variance tests, as well. Variances were not homogeneous across data.

SPSS output table for the 2 independent samples non-parametric tests comparing the variables in Experiment 1 to ones in Experiment 2 is provided in Table 4.1. Mann-Whitney non-parametric tests indicated significance in terms of the following variables, when these two experiments are compared:

1. The number of operations
2. The number of operations until the first affordance trial on the goal object
3. The number of operations between the first trial of affordance and the trial of another object without affordance on the goal object

4. The number of operations between the first trial of affordance and the trial of subgoals using an object with related affordance

5. The number of operations between remembering the usage of a previous object with a related affordance and the usage of a novel object on the main goal object

Table 4.1: Non-parametric test results comparing Experiment 1 to Experiment 2

Test Statistics[b]

	ops	spops	kdops	firstaff	othergoal	subgoal	scgd	pp	pw	scgb	kp
Mann-Whitney U	,000	22,500	19,000	,000	9,000	3,500	20,500	13,500	14,500	15,000	,000
Wilcoxon W	28,000	50,500	47,000	28,000	37,000	31,500	48,500	41,500	42,500	43,000	28,000
Z	-3,134	-,260	-,789	-3,144	-2,193	-2,689	-,518	-1,505	-1,335	-1,221	-3,216
Asymp. Sig. (2-tailed)	,002	,795	,430	,002	,028	,007	,604	,132	,182	,222	,001
Exact Sig. [2*(1-tailed Sig.)]	,001[a]	,805[a]	,535[a]	,001[a]	,053[a]	,004[a]	,620[a]	,165[a]	,209[a]	,259[a]	,001[a]

a. Not corrected for ties.

b. Grouping Variable: exp

Abbreviation list for the variables in Table 4.1:

ops = Number of operations

spops = Number of scissors-paper trials

kdops = Number of key-door trials

firstaff = Number of operations until the first affordance trial on the goal object

othergoal = Number of operations between the first trial of affordance and the trial of another object without affordance on the goal object

subgoal = Number of operations between the first trial of affordance and the trial of subgoals using an object with related affordance

scgd = Number of operations between the first trial of subgoals using an object with related affordance and the first trial of subgoals using an object without related affordance

pp = Number of operations between the trial of subgoals using an object without affordance and returning back to the main goal object with an object without related affordance

pw = Number of operations between the usage of an object without related affordance on the main goal object and the trial of subgoals using an object without affordance

scgb = Number of operations between the trial of subgoals using an object without affordance and remembering the usage of a previous object with a related affordance

kp = Number of operations between remembering the usage of a previous object with a related affordance and the usage of a novel object on the main goal object

exp = Either 1 or 2, indicating which experiment this data belongs to

4.2.2 Computer Model

Computer model was implemented in Java using Eclipse and it also utilizes PKS software, a knowledge-based planning tool, Planning with Knowledge and Sensing, for automated planning. Computer model simulates the actions taken by the human players in the game and the knowledge updates that might be claimed to take place during the exploration stage.

In the Java code, a Game Object is defined. This indicates the scissors, paper, and doors etc., objects that are present in the environment. Game Class includes all the necessary methods to perform actions that the plan indicates and produce their effects according to the configuration of the objects and the environment.

First of all, an initial planning file, which is an XML file, is created by the agent. In this template file, model has access to action-effect pairs that represent agent's knowledge about its environment before the experiment. Pre-defined default actions are grab, drop and touch. The preconditions and effects of these actions do not change in the plan file. However, the definitions of cut, unlock and break actions get updated during the simulation. For instance, at first the agent has the knowledge that bringing scissors and a piece of paper together results in the paper being cut, as the intended design function of scissors is cutting. Additionally, keys are expected to unlock doors. The goal is set as the state of paper being cut. Preconditions indicate the requirements for that action to be performed. Parameters x and y denote the objects in question. Effects represent the outcomes that ensue after that action is performed. Below is the definition of the default grab action:

```
<action name="grab">
    <params>?x</params>
    <preconds>
        K(is_object(?x)) ^ K(is_free(?x)) ^ K(is_free(hand))
    </preconds>
    <effects>
        add(Kf, is_grabbed(?x));
        del(Kf, is_free(?x));
        del(Kf, is_free(hand));
    </effects>
</action>
```

PKS files are in the form of an XML file and they contain the definition of the initial state, goal conditions, agent's initial knowledge about actions and its capabilities in terms of state transforming actions. As the rooms are explored, the file is filled with information about the states, obstacles, revealed rooms or objects. Next, this file is fed into PKS, which generates a sequence of operations that gets the agent from the initial state to the goal state , a plan, using the information about the environment by making use of a search algorithm, depth-first, breadth-first or iterative-deepening depth-first search.

Using the template planning file, agent produces its first plan, which includes a sequence of actions listed in the output file as below:

- Grab(scissors)
- Touch(scissors, paper)
- Drop(scissors)

Then, the agent tries to perform these actions. When the firing of a method corresponding to an action does not produce the effects predicted in the plan, the execution is stopped and the plan configuration is updated in terms of the effects of the action taken. Hence in this case, the agent will be able to grab the scissors as the agent's hand is free and the scissors are graspable. Then, it is also possible for the agent to make the scissors and the paper touch each other. Touch action fires either one of the cut, unlock or break actions. Although in the plan file, agent expects that making scissors and paper touch each other would results in paper being cut, in the Java definition of the touch method, cut action is not fired. This is to simulate the scissors falsely affording the cut action.

After this part, as the goal is related to the paper, the agent performs exploratory actions on the paper. After running out of actions to perform on the goal object, the agent finds out about obstacles in the room, which are the doors. Then, the agent creates a subplan whose goal is to unlock one of these doors. The agent moves on to trying out the key on the doors, as in the plan file keys are still listed as affording the action of unlocking doors. After observing that this is also not the case for the current situation, agent starts to use the mobile objects other than the key on the doors. By keeping track of the object combinations it tried, the agent is finally able to break the glass door using the scissors. Finding the pliers in this room, the agent moves from the subplan of getting rid of the obstacles to the main goal of separating the paper and tries the pliers on the goal object, again to see no effects produced. Then, trying out the pliers on the wooden door opens the door. At this stage, the agent tries to make paper and the glass box touch, then pliers and the glass box. Using the scissors again, the box is broken and the knife is revealed. The agent cuts the paper using the knife.

In all of these stages, by using the base plan file, a plan to execute is generated. Agent keeps a list of objects that it sees and a list of object pairs that include pairs that it tried to bring in contact. The initial plan the agent generates include the actions that are guided by affordances of the objects visible to the agent. Hence, scissors cut the paper. After determining a sequence of operations, the agent tries to execute it. When the effects of these operations do not match the ones in the plan, the agent stops the execution and tries to alter the consequences of those particular actions in the plan definition file. In cases where the agent seems to be stuck, it looks for other ways of getting closer to the goal state, trying out unintuitive pairs, or checking if the obstacles such as doors can be eliminated.

The actions in the plans are then tried out by the agent to see if they give the desired outcomes. If not, the plan file is updated by either adding or removing action-effect pairs on objects and another plan is created. This goes on until the main goal is achieved. As the agent explores its environment, it learns about the objects, what it can and can't do with them indicated by action-effect pairs representing true affordances, whether there are obstacles or hidden objects in other rooms or boxes. All the information gathered during exploration are added to the agent's knowledge database as positive facts.

In the output file, the sequence of actions that would take the agent to the goal state is listed. Final plan produced by the model is given below:

1. Grab(scissors)
2. Touch(scissors, glassdoor)
3. Drop(scissors)
4. Grab(pliers)
5. Touch(pliers, woodendoor)
6. Drop(pliers)
7. Grab(scissors)
8. Touch(scissors, glassbox)
9. Drop(scissors)
10. Grab(knife)
11. Touch(knife, paper)
12. Drop(knife)

Below are the algorithms of the methods for assigning goals and subgoals and generating plans to explore the environment and reach the goals:

Main goal:

1. Check if there is an object whose primed affordance is related to the goal
    - 1.1 If there is, try to generate and execute the plan involving that object and the goal object
        - 1.1.1 If the desired effect is produced
            - 1.1.1.1 Terminate
    - 1.2 If there is no such object or the desired effect is not produced
        - 1.2.1 check the other objects in the environment
            - 1.2.1.1 if there is another object in the environment that is available and might be useful, try it on the goal object
                - 1.2.1.1.1 If the desired effect is produced
                    - 1.2.1.1.1.1 Terminate
                - 1.2.1.1.2 If the desired effect is not produced
                    - 1.2.1.1.2.1 Create a subplan for a subgoal

Subgoals**:**

1. Check if there is an object whose primed affordance is related to the subgoal
    - 1.1 If there is, try to generate and execute the plan involving that object and the subgoal object
        - 1.1.1 If the desired effect is produced
            - 1.1.1.1 Terminate and go to step 1 of the method for goals
    - 1.2 If there is no such object or the desired effect is not produced
        - 1.2.1 Check the other objects in the environment
            - 1.2.1.1 if there is another object in the environment that is available and might be useful, try it on the subgoal object
                - 1.2.1.1.1 If the desired effect is produced
                    - 1.2.1.1.1.1 Terminate and go to step 1 of the method for goals
                - 1.2.1.1.2 If the desired effect is not produced
                    - 1.2.1.1.2.1 Go to step 1.2.1 of the method for subgoals

CHAPTER 5

DISCUSSION AND CONCLUSION

The main goal in the experiment is to separate the piece of paper; however, it also includes sub-goals, as well. In the experiments, we noticed certain paths of operation sequences. For instance, subjects immediately turn their attention to the scissors and try to cut the paper using them, as this pair is the one most relevant to the goal. However, the nature of the experiment does not permit them to do so. Subjects need to look for other ways of using the scissors or separating the paper into two parts. For instance, after trying to cut the paper with the scissors, subjects try to cut it with different objects, at first ignoring the doors. However, the existence of a key cues that the doors might be unlocked, as well. Unlocking the doors to see if there are any objects that would be helpful in reaching the goal becomes a sub-goal itself. After seeing that the key does not unlock any of the doors, subjects start to try different objects to go through the doors and eliminate this obstruction. In a case where the key is unable to open the door, subjects do not get this idea immediately. At first, they think that they were not able to find the exact location where they should put the key to open the door.

The second part of the experiment includes a setting where the objects behave seemingly randomly and the actions done upon them create effects that seem to defy the laws of physics. In the first part of the experiment, we wanted to observe how the subjects become aware of the unavailability of the primed functions of objects and try to find out new ways of achieving their goals. As explained, in such cases functional fixedness occurs and the flexibility is reduced as the constraints seem to increase. At that point, subjects need to perform a creative leap to go past a limit to envisage novel ways, or maybe not novel but secondary ways of using objects. Therefore, the disabled functions of the objects and the goal must be related so that the subjects would think that the goal would be achieved by performing that function on that particular object. In a case where the disabled function is not relevant to the goal, subjects may not even notice that it is disabled. However, in a case where the situation seems to require the usage of the disabled function, the effect of functional fixedness would be observed to a greater extent where the subjects fail to immediately find other ways of employing that object or using other objects to achieve their goals. This is also the case for sub-goals, for instance, unlocking the doors.

As for the likely affordances part, the subjects also try to use the key on the paper, which is the goal object, then on the doors, which are obstacles that might be eliminated with the help of a key. After observing that the key is non-functional in terms of separating the paper into two or unlocking the doors, some subjects go into a state of breakdown, whereas some others start to look more actively at the environment. Then, at some point it occurs to them that they might use the scissors to pierce and break the glass. This stage requires some creative leap to overcome functional fixedness. After breaking the glass with the scissors, glass door is no longer and obstacle. A new object, pliers, is now available for use. Subjects tend to use the pliers on the paper, as it looks a bit similar to the scissors and might perform the required action. Some others tend to go directly to the wooden door to unlock it. Pliers unlock the wooden door, which is no longer an obstacle. Here the new object is a glass box, with a knife inside it. Subjects usually try the pliers on the glass box, probably because it is the one that they can currently see and that has the current attention. The latter options are the key and the scissors. To see if the subjects would remember that they broke the glass door with the scissors and whether they would try to do the same with the glass box, we made it so that the glass box is breakable by the scissors. Subjects seem to not immediately remember that they used the scissors to break the glass door with the scissors. After realizing that it could be done, they quickly use the scissors on the glass box to reveal the knife inside. Knife is a relief for subjects, because it seems to afford cutting. This part is interesting because although the setting consisted of objects whose primed affordances were not functional, subjects still think that a knife can cut. Indeed, this knife is capable of cutting the paper, and helping the subject achieve the goal.

Interesting trial combinations occurred during the second experiment, whereas in the first experiment subjects rarely directed their attention towards such ideas. For instance, one subject put forward the idea of repairing the key with the pliers, and one of them tried to bring pliers, scissors and the paper together. However, in the first experiment, subjects were able to move on faster towards the goal. This was probably due to the fact that as the subjects in the second experiment had to try harder to move objects around, they had more time and space to think about interesting combinations of objects that would help them get closer to the goal. The lack of additional actions was a constraint in the game and the computer model. The actions available to the players and the agent are grabbing, dropping, colliding two objects with each other to cut, unlock or break. However, in real life, humans can think of additional actions such as repairing or combining several objects. This indicates the limits of the experiment in terms of transforming real life interaction to the interaction present in the game. One of the limits of the computer model was again the lack of all possible actions, as we tried to keep the environment simple. The assumptions made about the agent's previous knowledge about its environment and action-effect patterns representing the affordances also impose constraints. In addition to affordances, it would also be beneficial to add information about the physical properties of the objects and plausible mechanisms to utilize this data during the process of planning.

First interaction occurs between the object whose state should be altered to attain the goal state, and the object whose primed affordance seems most relevant to the goal. Paper is the object to be separated into two pieces, and the scissors seem to afford the action of cutting it into two. Thus, the subject thinks that the cutting is required so he or she tries to perform that action with the scissors. However, the primed affordance of the scissor is disabled. After finding out about this peculiar property of the scissor in question, subjects try different combinations of objects to find out whether they would provide the desirable result. In addition, subjects also make use of the object whose primed affordance is disabled, as a means to eradicate obstacles, for a chance of obtaining another object behind or inside them that might also help us achieve the goal. In such cases, subjects are forced to use that object as an intermediary tool. It is also possible to make use of the same object, however this time having a different action in mind. Nevertheless, we did not utilize such interactions in the experiment to avoid confusion. An example to such a situation is trying to open a box with a screwdriver.

We might not be able to loosen the screws on the lid with that particular screwdriver, but we might consider using the screwdriver to form a fulcrum to elevate the lid of the box. Here, the primed affordance of the screwdriver seems to be required to open the box, but again it is disabled. The subject can try to hold the object in different ways or to make use of different ends of an object etc. to perform actions other than the primed design function. In a case where the action afforded by an object present in the environment is not relevant to the goal, it seems to boil down to physical properties.

All these conditions are summarized in the sentences below:

• Applying the primed affordance on the goal object

• Using the primed object in a different way than the intended design function on the goal object

• Using the primed object in a different way than the intended design function on different objects to obtain other objects or to eliminate obstacles.

• Using a non-primed object on the goal object

• Using a non-primed object on different object to obtain other objects or to eliminate obstacles

Via trial and error, subjects become aware of the disability of the primed affordances of certain objects. Then, they start using this object or other objects to perform other actions that would lead them the goal state. Sometimes, the subjects purposefully, deliberately search for objects or try actions. However, sometimes they find out about the environment with random actions or with the help of the actions that were not intended by them. For instance, they try to put the scissors on the table, put it falls on the paper, which was laying on the ground. They did not deliberately take the scissors towards the paper, nevertheless it happened.

This is specifically the case in the second experiment where the objects behave and move in odd ways. These random movements sometimes cause unintended actions to be performed. In the case where the objects move strangely, it is hard for the subjects to find out about the patterns of action-effects in the environment. These random action-effects might help them in their struggle.

It is crucial to note that the images of the objects in the game do not "literally" afford the actions that they afford in real life. For instance, a knife in real life affords cutting, nonetheless, the image of the knife in the game affords separating the image of the paper into two pieces. Therefore, one should think of the affordances in the game in terms of the actions that can be taken on the images of the objects and what they represent in comparison with the actions in real life. Hence, the image of the key not being able to eliminate or rotate the images of the doors indicates a false affordance. Additionally, the image of the pliers includes a hidden affordance of being able to rotate the image of the wooden door.

False affordances occur in the game because the perceptible images are the counterparts of the real life objects and they give the impression of being capable of performing the intended design function. However, by trying out this action possibility after being supervised by previous plans involving this object, subjects then reject this false affordance to search for possible hypothesis about the action-effect pairs of that particular object. This is where we claim "Plan to Learn" takes place. Presence of a goal or the absence of a goal would also make impacts on the operation sequences in the plan execution phase or the trial phases. Presence of a goal would lead the actions of the subject to be more goal-oriented and purposeful, whereas the absence of it could induce random exploration actions. The differences in these conditions might be exhibited in the eye-movements of the subject in terms of the number of fixations and saccades.

Subjects in the experiments act on the objects present in the environment at first. After noticing that these objects do not meet their expectations, they look for new objects in the environment. Hence, at first, a reactive kind of planning occurs. Observing that this plan does not perform well, subjects without changing the main goal, they update their knowledge database about the world and try to form new plans whose configurations are consistent with the current environment. After trying out the obvious choice of objects, scissors and paper, most subjects keeping the goal-relevant object paper in mind, look for other objects to cut it into two. Then, they notice the key on the table, which they later try to bring together with paper. Some subjects, however, instead of bringing together this unlikely couple, they go for the doors immediately with the key. Here, they are driven to the doors by the seemingly obvious affordance of the key. At this point, getting rid of the doors become the subgoal, as unlocking them might reveal new objects. This is a crucial point in planning, as the subjects move forward from the pure reactive planning to looking for objects that might yet be hidden, however, if found, can be useful in their quest for separating paper into two pieces. This part of the planning corresponds to the System 2 explained by Kahnemann, which includes higher cognitive processes taking place in the cortex (Kahneman, 2011).

If the planning mechanism is innate, the formation of plans should be similar in all cases. However, as in the case of parsing, certain parameters of the planning domain can affect the strategy used in solving the planning problem. Still, the brain looks for patterns to give meanings to. In the second experiment, these patterns are quite so observable, in fact at the beginning of the experiment objects seem to act in rather strange ways. The key randomly jumps from one place to another when the subject tries to grab it, or it is so hard to grab the scissors that when the subject is moving, he or she sometimes does not notice that the scissors fell to the ground. Under these conditions, there is some vagueness and uncertainty in the environment.

It seems that even though functional fixedness occurs, it is also observed that the subjects, after discovering that the scissors are not capable of cutting the paper into two, they look for possible objects that they might utilize while cutting. They start to think about objects that might cut, such as a knife or a piece of broken glass. After this phase, they become more aware of their environment and the objects around them. This part sometimes requires certain tips to be given to the subjects, as it takes quite a lot of time for some of the subjects to notice that they can employ objects in novel ways, as well as making use of their more familiar primary functionalities. It is also important to note that sometimes the context might change the interpretation of objects to a great extent.

When the operation sequence that comes to mind during the first trials fails, for instance, the operation to achieve to goal is cutting, using scissors applied on a piece of paper, new sequences are considered. In such cases, as we have mentioned above, the subjects start to look for objects that might be helpful in cutting. At this point, they do not look around randomly to construe possible ways of cutting with every object in the scene. Instead, they look for objects with specific physical properties, also manifested in their perceptual appearances. Hence, if the goal is cutting, they look for sharp, solid objects that can facilitate cutting. This indicates that we have certain object categories in mind which we associate with physical properties and the actions that we can perform with them. An object can be solid and it can also capable of smashing, whereas another solid object would fail to do so. In this case, one should consider another parameter that makes the former able to smash and the other unable to smash. This difference also constitutes a separate category. For instance, a metal knife and a broken piece of glass can be in the same category of objects that can cut, but only the glass belongs to the category of objects that allow seeing through.

This indicates a task-oriented search for potential objects that can help us achieve our goal. In the case where only the primary functions of objects are disabled, this is what happens in the experiment. With the goal in mind, subjects have a plan. The plan is to apply the cutting operation on the paper using a sharp, solid object that is capable of cutting. If there is such an object on the spot, then they tend to utilize it, without hesitating to consider other objects in the scene. It is as if they are not even available in the scene. Very quickly, the subjects fixate on the object with the primed affordance. This indicates how the action requirements given by the goal and the action possibilities implied by the perceptual stimuli cause a sort of selective perception. If they are not able to achieve this operation, seeing that it produces no desired effects, this is when they turn their attention to other objects. Here, it must occur to subjects that they cannot achieve the task directly meaning that the affordances implied by the stimuli were actually false ones, or it can just as well be that the knife had blunt edges, hence it could not cut, and nevertheless, they should use other means to achieve it.

Because of the constraints that are posed by the goal and the environment, the search space of the problem involving all possible states is narrowed down. This narrowing down is caused by the requirements of the goal state in terms of both the actions to be taken and the objects to be used. Hence, it can be said that with the goal in mind, the subjects already have a plan and during execution, they alter this space according to the recently discovered constraints that can only be observed after interacting with an object and seeing no desirable effects. This helps subjects to converge on a sequence of plans. In a case where the objects do not behave sensibly, for instance in a setting where the laws of physics are defied, it would be hard to get accustomed to the environment. In addition, if the objects are behaving randomly, for instance when the user clicks on the object and tries to drag it right with the mouse, the object sometimes goes to the right and sometimes to the left, maybe sometimes up or down, it is not possible for the subjects to grasp the meaning of the movements of that particular object.

Our brains look for patterns and sometimes work with the help of the patterns that we observe in the world. Patterns are quite useful in terms of the time and space that they help us to save. Quite importantly, they help us categorize objects, events or contexts. In a situation, where it is not possible to form patterns in mind, we cannot properly alter our plans that we have generated beforehand. Even though a feasible solution can be find, meaning that there is a possible sequence of operations that takes the subject from the initial state to the goal state, given the seemingly random effects of the actions taken, it would not be easy to guess the effects of our actions. It would be similar to learning everything again and again from the scratch when we see an instance of it. Nonetheless, patterns and categorization help us remember each time that we had observed that object or action before.

In the random effect case, one would try to use statistical learning. However, if the distribution of effects does not make sense, we cannot also create action-effect pairs on objects. Under such conditions, the goal would not guide us, as we struggle with the objects and the effects that actions upon them produce. Hence, no pre-generated plan would help. In this case, we expect meaningless moving around and saccades instead of fixations. Here, one must learn how to plan in these kinds of cases, which fails because many of the observations are quite counter-intuitive to the expectation coming from System 1. Compared to the first case, where one has a plan and the constraints do not seem to be that counter-intuitive, unlike the second case where the conditions make it hard to for the action-effect pairs to be fit into a pattern. Hence, the creative leap that is required to take place in order to get rid of the functional fixedness occurs later than the one in the first experiment.

Objects in our environment provide us with affordable actions, which can be related to our past experiences with those objects. Our motor plans are dependent on the objects in our and environment and the affordable actions that they provide us. We make plans according to these affordances that are constructed via our experience. Affordances in this context indicate action-effect pairs. Hence, when we are posed a problem in a familiar environment, we rely on our past experience and employ the related affordances in our pre-plans. Pre-plans are constructed in the mind before we take an action.

In the experiment, the first trials are always grabbing the scissors and bringing it closer to the paper. This is because the plan the subjects have in mind is a simple one that consists of a few actions. Subjects think that performing this plan would surely take them to the goal state. After becoming aware of their inability to perform the intended action, they move on the other ways of getting closer to the goal state. Plans we make seem to build on our previous plans. Instinctual plans are the most basic plans and the heuristic plans are the plans that operate on plans or plans of plans etc. Plans that require the usage of motor skills can also make use of muscle memories, which would need to be altered as the person grows physiologically.

When we construct a pre-plan, it can be either complete or incomplete. Incomplete parts may be filled in later. In simple cases where the environment includes 3 or 4 objects, it could be easier to make a complete plan. This is because the complexity is relatively low and the working memory capacity allows us to work on that kind of problems. Therefore, when we see a piece of paper and scissors, it is quite easy to make the plan as follows when the goal is to create two pieces of paper:

Grab(scissors), Touch(scissors, paper), Drop(scissors)

When we start to perform this pre-plan, we have certain expectations about the effects of these three actions grab, touch and drop. We expect that one of the results of touch(scissors, paper) would be a state of change in the piece of paper, making it cut. However, if what we perceive counters this expected effect, meaning that a different effect is observed, we need to update our pre-plan configuration and remove that effect from the action-effect pairs, as this is not observed in all cases. Here, one needs to start exploring his or her environment and update his or her plan configurations to match the real effects that are observed after performing a certain action. In a case where one can grasp a certain kind of pattern or even acquire some simple statistical expectation, humans can adapt to that kind of environment and make use of the objects even in an environment with a different gravity value or where the colours are all inverted. This is because even though they are different than the conditions we live in right now, one can adapt to them since they are permanent. If the gravity is in the x direction all the time, we can adapt to that as well. However, if it changes randomly from x to y to z or plus or minus, it would be harder to explore and update our plans, as we would not know what to expect in the next state. As one of the subjects indicated “I think of using the key to unlock the doors, but I tried to grab it, but it is not moving in a proper way and it did not work out. Now

I cannot think of anything else as I do not know what I can do with the key". This is called functional fixedness. In the first part of the experiment, subjects are better at recovering from this fixedness. Nevertheless, in the second part, they are stuck and they cannot even properly explore the environment and learn more about it in order to form plans. They are stuck and this might indicate that they are unable to see a pattern. It must be noted though, after a few tries, the objects return to normal and they could actually move properly. This is also shown in the model, where after a few search steps, it is actually possible to reach the goal from the initial state, and nonetheless, subjects fail to see that. There is a limit to the exploration drive, moving from curiosity to boredom. Curiosity is evolutionarily advantageous as it helps us become more aware of the environment and what we are capable of doing being situated in it, as well as our ability to detect relevant objects or properties in our environment.

With the experiments, we tried to demonstrate the differences between using plans to learn about the environment and learning how to make plans under certain conditions. Planning utilizes hierarchical structures, which makes the process a complex one. Most of the time when we encounter a problem, we can make use of our previous plans, modify them if required and execute the final plan. We can even alter our plans on the go while acting in the environment as online planning.

Thinking aloud helped us gain an insight into the minds of the players (Miller et al., 1960). Observing only the actions taken might cause us to speculate about the motives behind them and we might be mistaken. However, if someone talks about why he or she is taking that action that could provide us with valuable feedback. Sometimes, when a high cognitive load is imposed when the subject is unable to evaluate the situation, players might stop talking. It is also possible that thinking aloud might intervene with planning and make an impact on the outcomes.

Human planning makes use of verbal information and in some cases it might only rely on verbal structures. Instead of going into all the details and listing them one by one, a simple label would be enough to describe a whole range of activities. Due to the involvement of the phonological loop, thinking aloud might distort verbal planning activities. Languages are considered to have hierarchical organizations, which can also involve recursive structures. This could also be the case in planning, language and planning might be sharing the same form of underlying structure, a symbolic manipulation machine, universal Turing machine.

As in Janet Fodor's parse to learn vs. learn to parse argument, it is claimed that the underlying planning mechanism is innate (J. D. Fodor, 1998a,b; Steedman, 2002). As the child learns the language through parsing the linguistic information he or she gets from the environment, planning is performed to learn more about the environment we are situated in. No child learns how to parse, no one teaches them about the grammar or the parts of speech. In a similar vein it can be claimed that using pre-plans and supervision, agents are able to learn about their environment. We make and update plans and try to perform them to check the outcomes to see if the results confirm or infirm our knowledge. We do not learn how to accomplish these stages, we can only learn how to make better plans using strategies or heuristics or by adopting plans of someone else. This is when we learn about heuristics rules of thumb or better search algorithms. We can learn new plans, as they are simply a kind of information to add to our database of world knowledge. Nonetheless, we are not that good at seemingly random situations and making statistical judgements about them (Kahneman, 2011). There is a boundary to how much information we can keep in our working memory, time and space are all constraints to planning. Hence, we do not try to perform an action a thousand times to obtain the frequency of different effects happening. Maybe this is the case in our brains, what happens is that we repeat the idea of performing that action in order to achieve the effect we observed. Idea of performing an action is almost the same as performing the action (Miller et al., 1960; Prinz, 1997). Test phase of an operation does not last very long, humans make use of certain stop rules to avoid loops and the halting problem and trying out the same action over and over again. Observing once that an action does not occur under those conditions is usually enough. Instead of stochastic systems, Markov chains, deterministic hierarchical digital formal symbol manipulation systems seem more plausible for the case of planning as well as languages.

If the game could provide more feedback in terms of not occurring actions, the subjects would make more valid decisions. If the interaction would provide feedback, for instance the sound of wood or the sound of a locked door, then it would be easier for the subjects to find out that the interaction between these objects do not provide the effects previously desired by the subject. Hence, they would be able to navigate from the key-door interaction, and they would try to discover new ways of opening the door. This feedback could also include verbal information, such as: "The key does not seem to open the door, would you like to try other objects?"

Mann-Whitney non-parametric tests indicated significance in terms of some variables monitored in screen captures. First of all, the number of operations was significantly higher in the second experiment. This was one of the expected results, since the second environment contained objects moving randomly on top of disabled primed affordances. As the second case seems to impose more cognitive load and requires thinking more about the conditions, subjects struggled with the objects more in order to extract the meanings of their seemingly random movements.

In addition to the total number of operations, we expected the numbers of trials with affordance pairs to be greater. For instance, we expected that the number of scissors-paper trials would be significantly higher in the second case, since the conditions cause the subjects to remain more unsure. However, this was not the result we received from the tests.

One of the other significant parameters was the number of operations until the first affordance trial on the goal object. Again, the values of this parameter was greater in the second experiment. This disparateness could also be explained in terms of the struggle the subjects were going through. In both of the experiments, at the very beginning, subjects had a hard time grabbing and moving the objects around. However, they got accustomed to using the mouse and keyboard in synchronization in an advanced way about a minute later. Nevertheless, it was the first group who successfully tried the scissors-paper pair to check if it would ensue the desired effect.

The number of operations between the first trial of affordance and the trial of another object without affordance on the goal object was significantly greater in the second case, as well. This parameter refers to the usage of a seemingly irrelevant object on the goal object. In the game, this parameter corresponds to the first key and paper pair trial after the first trial with scissors and paper. Interestingly, this trial never occurred to some of the subjects in the first experiment who took the key directly to the doors. Hence, we can claim that in the second experiment, as the subjects are dealing with the strange movements of objects, they are not sure how those objects would react in familiar situations, therefore they try out all the combinations. Quite interesting ideas came up from the subjects' side. For instance, one of them thought about repairing the scissors with the pliers.

After trying out scissors-paper pair, some of the subjects directed their attention towards the key-paper pair, whereas the remaining ones went directly for the door with the key. The number of operations between the first trial of affordance and the trial of subgoals using an object with related affordance also turned out to be a significantly different in two experiments. It took more operations for the subjects in the second experiment to try out the key-door pairs, which form subgoals of getting rid of obstacles and which also include a primed affordance of the design intentions. Finally, the number of operations between remembering the usage of a previous object with a related affordance and the usage of a novel object on the main goal object was also found to be significant. This is the case where the subjects are required to break the glass box with the scissors. Subjects in the second experiment had more trouble overcoming the obstacle, which is the glass box containing the knife inside, while it was more easy for the subjects in the first experiment to recall that they broke the glass door with the scissors. Therefore, it can be said that as the cognitive load and the uncertainty was less in the first condition, subjects were able to recall more quickly the fact that they utilized the scissors to break the glass in the previous stages.

In the first case, they do not need to think about other combinations that much, as they can quickly grasp and move objects around, they also observe changes in the environment more quickly. Therefore, they do not seem to allocate greater amounts of resources to thinking about trying out uncommon object combinations. Indeed, this could be the favoured way in evolutionary terms. As it is quite important to act or decide quickly in fight or survival situations, using heuristics or prior plans rather than sitting and thinking about all the possibilities would be more advantageous. This indicates the importance of prior plans and the later modifications of them in the learning and further planning processes. In the second case, subjects do not begin to learn a different way of planning. Instead, they try to use their prior plans, patterns, affordances and other mental representations about their action and perception abilities. This is, of course, the same in the first case, as well. However, since the second experimental conditions impose more randomness and uncertainty, the underlying hierarchical mechanisms seems to make greater efforts to grasp meanings. It should also be noted that some subjects in the second experiment experienced breakdowns where they stopped trying out combinations or trying to move objects as they expressed that they do not know what to do because they are not able to predict the results of their actions.

Of course, one can try to keep the effects of every single trial in mind and to predict statistically what the outcome would be in the next trial. However, this is not the case we usually observe in humans. Humans tend not to rely on frequency data, particularly in the case of affordances. Since if a scissors do not cut in the first few trials, we do not usually try to perform the same action 1000 times to see if it would cut the piece of paper somehow. Yet, there might be exceptions.

First of all, the number of subjects in the experiment was not high enough to conduct parametric tests. Hence, this was a drawback of the experimental part. One of the drawbacks of the 3D environment was the difficulty of grabbing and moving objects around, as it required using both the mouse and the keyboard. Users experienced in playing video games, however, did not encounter difficulties in this case. Therefore, a pre-experiment familiarization part could be helpful. For instance, before starting the real experiment, there could be a part where the subjects try to explore a different simple environment, where they are free to grab and drag objects, navigate using the arrow keys and move the viewing angle using the mouse.

One of our concerns was about the number of objects in the environment and the number of operations that the subjects had to perform to achieve the goal. This did not turn out to be a problem, since most of the subjects were able to achieve the goal in under 10 minutes for the first part, under 20 minutes for the second part. In this duration, they were able to explore the environment, try out different object pairs to create the effects they desired and finally reach the goal state.

With utilizing a similar game environment, the presence or absence of rewards, the impact of cues and landmarks in learning and remembering could also be studied. Using eye-tracking methods would reveal the gaze patterns of subjects during the experiment. What we expect to see is that at first before any interaction, the subjects would perform fixations on certain objects that they think are relevant to their goal. Then, when they notice that the intended design functions, which could be related to the goal, are disabled, they would look for other ways of attaining the goal, where saccades would be greater in number. In the second group, we expect to see more saccades directly from the very beginning of the interaction. In terms of the step sizes represented by the number of actions taken to achieve the goal, we expect the number to be smaller for the first group.

In settings similar to the second experimental environment, sometimes our actions do not form meaningful sequences and we cannot grasp the cause and effect chains between our actions and the effects produced. Such cases are akin to memoryless state-by-state evaluators. When the adjoining states appear to have no reasonable relation with the preceding states, we need to decide what to do next by evaluating the current state at the time of decision and ignoring the past events. Similar-to-random experience in prior states do not present clear and reasonable ways to interact with the objects and the way of choosing available actions effectively to attain the goal in the next trials. In summary, we cannot learn how to plan under such conditions. During such unexpected situations, we still try to rely on our previous plans or previous experiences with familiar objects. Hence, we are reminded of possible actions via the help of perceived affordances and if the execution of those actions do not give rise to expected outcomes, we begin to try out the combination of objects or object-action pairs without heeding much attention to the familiar affordances they seem to convey to us. If the affordances turn out to be false ones, or our predictions never come true, we may feel lost and the situation would be a chaos. In the experiments, we included a handful of objects to interact with. However, if there were many more objects, as in real life, and if they all behaved in unexpected ways, there would be lots of possible combinations to try out and this would impose heavy cognitive loads on our memories. Particularly the short term memory would not be able to hold all of the object-object pairs or object-action-effect tuples to generate optimum plans. Unknown variables and uncertainty would hinder discovering patterns and structure in the environment and this means confusion on the side of the subjects. Additionally, when the affordances fall short, we may try to consider the physical properties of objects to see if they could be utilized to attain certain goals. This requires getting rid of the limits of human concepts and design intentions, which interestingly corresponds to a creative leap.

Bibliography

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349. Retrieved from `http://dx.doi.org/10.1016/j.cognition.2009.07.005` doi: 10.1016/j.cognition.2009.07.005

Balkenius, C. (1994). Biological learning and artificial intelligence. *Cognitive Science*, *30*, 1–19. Retrieved from `http://cogprints.org/3705/`

Bennett, A. T. D. (1996). Do animals have cognitive maps. *Journal of Experimental Biology*, 219–224.

Berlyne, D. E., & Slater, J. (1957). *Perceptual curiosity, exploratory behavior, and maze learning.* (Vol. 50) (No. 3). US: American Psychological Association. doi: 10.1037/h0046603

Byrne, R. W. (2003). Imitation as behaviour parsing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *358*(1431), 529–536. doi: 10.1098/rstb.2002.1219

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187–192. Retrieved from `http://scholar.google.de/scholar.bib?q=info:Dkt1DEhBHFEJ:scholar.google.com/&output=citation&hl=de&oi=citation`

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, 181–195.

Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003). Can animals recall the past and plan for the future? *Nature reviews. Neuroscience*, *4*(8), 685–691. doi: 10.1038/nrn1180

Dennett, D. C. (1987a). Cognitive Wheels: The Frame Problem of AI. In *The robots dilemma: The frame problem in artificial intelligence* (pp. 41–64).

Dennett, D. C. (1987b). *The intentional stance*. Cambridge, Mass.: MIT Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*, 270.

Edelman, S. (1997). Computational theories of object recognition. In *Trends in cognitive science* (pp. 296–304).

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., & Tanila, H. (1999). The hippocampus, memory, and place cells: Is it spatial memory or a memory space? *Neuron*, *23*(2), 209–226. doi: 10.1016/S0896-6273(00)80773-4

Even-Dar, E., Kakade, S. M., & Mansour, Y. (2012). Planning in POMDPs Using Multiplicity Automata. *CoRR*, *abs/1207.1388*. Retrieved from `http://arxiv.org/abs/1207.1388`

Fodor, J. A. (1975). *The language of thought*. Cambridge, Massachusetts: Harvard University Press.

Fodor, J. A., & Pylyshyn, Z. W. (1981). How direct is visual perception?: Some reflections on Gibson's "Ecological Approach". *Cognition*, *9*, 139-196.

Fodor, J. D. (1998a). Learning to parse? *Journal of Psycholinguistic Research*, *27*(2), 285–319. Retrieved from http://dx.doi.org/10.1023/A:1023258301588 doi: 10.1023/A:1023258301588

Fodor, J. D. (1998b). Parsing to learn. *Journal of Psycholinguistic Research*, *27*(3), 339–374. Retrieved from <GotoISI>://000073922600002

Gaver, W. W. (1991). Technology affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 79–84). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/108844.108856 doi: 10.1145/108844.108856

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review*, *7*(4), 707-712. Retrieved from http://dx.doi.org/10.3758/BF03213010 doi: 10.3758/BF03213010

German, T. P., & Johnson, S. C. (2002). Function and the origins of the design stance. *Journal of Cognition and Development*, *3*(3), 279-300. Retrieved from http://dx.doi.org/10.1207/S15327647JCD0303_2 doi: 10.1207/S15327647JCD0303_2

Ghallab, M., Nau, D., & Traverso, P. (2004). *Automated Planning: Theory and Practice.*

Gibson, J. J. (1977). The theory of affordances. *In R.Shaw and J. Bransford (Eds.) Perceiving, Acting, and Knowing. Toward an Ecological Psychology*, 67-82.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review*, 336–342.

Hayes, P. J. (1971). The frame problem and related problems in artificial intelligence. , 223–230. Retrieved from http://dl.acm.org/citation.cfm?id=891924

Holland, P. C. (2008). Cognitive versus stimulus-response theories of learning. *Learning & behavior : a Psychonomic Society publication*, *36*(3), 227–241. doi: 10.3758/LB.36.3.227

Hommel, B. (2005). Perception in action: Multiple roles of sensory information in action control. *Cognitive Processing*, *6*(1), 3–14. doi: 10.1007/s10339-004-0040-0

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849-878.

Husband, R. W. (1929). A comparison of human adults and white rats in maze learning. *Journal of Comparative Psychology*, *9*(6), 361–377. doi: 10.1037/h0074679

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kearns, M. J., & Vazirani, U. V. (1995). Computational Learning Theory. *ACM SIGACT News*, *26*(1), 43–45. doi: 10.1145/203610.606411

Kendler, H. H. (1947). An investigation of latent learning in a T-maze. *Journal of comparative and physiological psychology*, *40*(4), 265–270. doi: 10.1037/h0056024

Kimball, R. C., Kimball, L. T., & Weaver, H. E. (1953). Latent learning as a function of the number of differential cues. *Journal of comparative and physiological psychology*, *46*(4), 274–280. doi: 10.1037/h0060949

Koehler, W., & Winter, E. (1926). *The mentality of apes*. New York: Harcourt, Brace & Company, Inc.

Koller, A., & Petrick, R. P. a. (2011). Experiences with planning for natural language generation. *Computational Intelligence*, *27*(1), 23–40. doi: 10.1111/j.1467-8640.2010.00370.x

Laird, D. (1985). *Approaches to training and development*. Addison-Wesley.

Lorenz, K. (1977). *Behind the mirror*. New York: Harcourt Brace Jovanovich.

Marr. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. *Book*. Retrieved from `papers://e74d72ed-e60d-4d01-b249-70f43c2b74c1/Paper/p803`

McCarthy, J., & Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. *English*, *4*(463-502), 463–502. Retrieved from `http://works.bepress.com/cgi/viewcontent.cgi?filename=0&article=1002&context=jozsef_toth&type=additional` doi: 10.1.1.85.5082

McLellan, H. (1996). *Being Digital: Implications for Education*. Educational Technology.

Miller, G., Galanter, E., & Pribram, K. (1960). Plans and the structure of behavior: Holt, Rinehart and Winston. *Inc., New York*. doi: 10.1037/10039-000

Mitchell, T. M. (1997). *Machine learning* (1st ed.). New York, NY, USA: McGraw-Hill, Inc.

Miyata, H., Itakura, S., & Fujita, K. (2009). Planning in human children (Homo sapiens) assessed by maze problems on the touch screen. *Journal of comparative psychology*, *123*(1), 69–78. doi: 10.1037/a0012890

Montesano, L., Lopes, M., Bernardino, A., & Santos-Victor, J. (2008). Learning object affordances: From sensory–motor coordination to imitation. *IEEE TRANSACTIONS ON ROBOTICS*.

Nau, D. S. (2007). Current Trends in Automated planning. *AI Magazine*, *28*(4), 43. doi: 10.1016/0360-8352(94)00222-9

Newell, A., Shaw, J. C., & Simon, H. A. (1959). Report on a general problem-solving program. *IFIP Congress*, 256–264. doi: 10.1038/224923a0

Nilsson, N. J., & Fikes, R. E. (1971). STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. *Artifical Intelligence*, *2*(3-4), 189-208.

Norman, D. A. (2002). *The Design of Everyday Things* (Reprint Paperback ed.). New York: Basic Books.

Nye, B. D., & Silverman, B. G. (2012). Affordances in AI. *In N. M. Seel (Ed.), Encyclopedia of the Sciences of Learning*, 183-187.

O'Keefe., J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford University Press.

Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, *33*(1), 1–24. doi: 10.1016/0028-3932(94)00098-A

Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex.*

Petrick, R. P. (2004). PKS : Knowledge-Based Planning with Incomplete Information and Sensing.

Poggio, T. (1981). *Marr's approach to vision* (No. AIM.645).

Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. doi: 10.1080/713752551

Pylyshyn, Z. W. (1989). *Computing in Cognitive Science.*

Russell, S., Norvig, P., Canny, J., Malik, J., & Edwards, D. (1995). *Artificial intelligence: a modern approach.*

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: action-induced modulation of perception. *Trends in Cognitive Sciences*, *11*(8), 349–355. doi: 10.1016/j.tics.2007.06.005

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal Behaviour*, *61*, 277–286. doi: 10.1006/anbe.2000.1606

Skinner, B. F. (1953). *Science and human behavior.* New York: MacMillan.

St. Amant, R. (1999). Planning and user interface affordances. In *Proceedings of the 4th international conference on intelligent user interfaces* (pp. 135–142). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/291080.291103` doi: 10.1145/291080.291103

Steedman, M. (2002). Plans, Affordances, And Combinatory Grammar. *Linguistics and Philosophy*, *25*(5), 723–753. doi: 10.1023/A:1020820000972

Steedman, M. (2014). Evolutionary Basis for Human Language: Comment on "Toward a Computational Framework for Cognitive Biology: Unifying Approaches from Cognitive Neuroscience and Comparative Cognition" by Tecumseh Fitch. *Physics of Life Reviews*, *11*, 382–388.

Steedman, M., & Petrick, R. P. A. (2007). Planning dialog actions. In *Proceedings of the 8th sigdial workshop on discourse and dialogue* (pp. 265–272).

Stoffregen, T. (2003). Affordances as properties of the animal-environment system. *Ecological Psychology*, *15*(2), 115-134.

Strange, J. (1950). Latent learning under conditions of high motivation. *Journal of Comparative and Physiological Psychology*, *43*(3), 194-197.

Sun, Y., Ren, S., & Lin, Y. (2014). Object–object interaction affordance learning. *Robotics and Autonomous Systems*, *62*(4), 487 - 496. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0921889013002339` doi: http://dx.doi.org/10.1016/j.robot.2013.12.005

Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In *In proceedings of the seventh international conference on machine learning* (pp. 216–224). Morgan Kaufmann.

Thorndike, E. L. (1898). Animal Intelligence: an Experimental Study of the Associative Processes in Animals. *Psychological Monographs*.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189–208. doi: 10.1037/h0061626

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (p. 381–402). New York: Academic Press.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In: Frank AU, Campari I (eds) Spatial information theory: a theoretical basis for GIS. , 14-24.

Uexkuell, J. v. (1920). *Theoretische biologie*. Berlin: Gbr. Paetel.

Unterrainer, J. M., & Owen, A. M. (2006). Planning and problem solving: From neuropsychology to functional neuroimaging. *Journal of Physiology Paris*, *99*(4-6), 308–317. doi: 10.1016/j.jphysparis.2006.03.014

Valiant, L. G. (1984). A Theory of the Learnable. *Commun. ACM*, *27*(11), 1134–1142. Retrieved from `http://doi.acm.org/10.1145/1968.1972` doi: 10.1145/1968.1972

Valiant, L. G. (2013). Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World.

Vermaas, P. E., Carrara, M., Borgo, S., & Garbacz, P. (2013). The design stance and its artefacts. *Synthese*, *190*(6), 1131–1152. doi: 10.1007/s11229-011-9885-9

Viezzer, M., & Nieywenhuis, C. H. M. (2005). Learning affordance concepts: Some seminal ideas. *International Joint Conference on Artificial Intelligence, Edinburgh*.

Watson, J. B. (1913). Psychology as the Behaviorist Views it. *Psychological Review*, *20*, 158–177.

Weerdt, M. D., Mors, A. T., & Witteveen, C. (2005). *Multi-agent planning: An introduction to planning and coordination* (Tech. Rep.). In: Handouts of the European Agent Summer.

Witt, J. K. (2011). Action's Effect on Perception. *Current Directions in Psychological Science*, *20*(3), 201–206. doi: 10.1177/0963721411408770

Wittgenstein, L. (1972). *Philosophical Investigations* (G. Anscombe & R. Rhees, Eds. & G. Anscombe, Trans.). Oxford: Blackwell.