This is a repository copy of *Is Epicurus the father of Reinforcement Learning?*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/122198/

Version: Submitted Version

Proceedings Paper:
Vasilaki, E. orcid.org/0000-0003-3705-7070 Is Epicurus the father of Reinforcement Learning? In: Sheffield Machine Learning Retreat 2017. Sheffield Machine Learning Retreat 2017, 02 Jun 2017, Sheffield, UK. . (Unpublished)

Is Epicurus the father of Reinforcement Learning?

Eleni Vasilaki[1]

[1]*Department of Computer Science, University of Sheffield, Sheffield, South Yorkshire, United Kingdom*

September 30, 2017

The following text is largely based on my talk at the Sheffield Machine Learning Group Retreat, on the 2nd June 2017. Reading again Ancient Greek Philosophy for my own entertainment, I was astonished by the Epicurean view of pleasure, not quite what I was expecting, and its close correspondence to the notion of the Reinforcement Learning objective function. Following a Google search, I was even more astonished that nobody appeared to have made the link. Trapped in a cycle of urgent but non-important matters, I ended up writing up the talk as a text a couple of months later, during my summer vacation, which I further revised based on the kind feedback of Professors Lúcia Specia, Shimon Edelman and Neil Lawrence. I also thank Professor Lawrence for encouraging me to write this up in the first place, Professor Georg Struth for initial discussions on Epicurean Philosophy and Dr Tom Schaul for giving the thumps up. Special thanks to Professor Andy Barto for pointing out that Plato's ideas were also related to Reinforcement Learning, and for further discussions.

When we think of Greek Philosophy, for most this constitutes the names Plato and Aristotle. This is not a surprise given the depth and the amount of their work, which, for the largest part, remains intact till today. In the famous painting of Raphael, Plato and Aristotle are portrayed right in the middle, dominating the scene. It is hard to pay attention to the young man at the bottom left-hand corner, who is said to be the philosopher Epicurus.

Epicurus is perhaps the most misunderstood philosopher of the Ancient world. Those who have heard his name before associate it with pleasure. Epicurus is commonly perceived as a hedonist, as someone who only promoted the immediate pleasures of the flesh. This, however, is a superficial and largely misleading presentation of his philosophy.

Born in the island of Samos in 341BC [1, 2], also birthplace of the famous mathematician Pythagoras, Epicurus consequently lived in Athens, Lesbos, and Asia Minor. He was aware of the Aristotelian philosophy, which was a mainstream philosophical movement at the time, however he was dissatisfied with this view of the world and adopted instead the earlier view of Democritus:

«ἀρχὰς εἶναι τῶν ὅλων ἀτόμους καὶ κενόν, τὰ δ᾽ ἄλλα πάντα νενομίσθαι» [3]

"The beginning of everything are atoms and space, everything else is in your mind."

Much of his philosophy builds upon the philosophy of Democritus, complementing ideas that Aristotle had vigorously criticised. With his views not being very popular in Athens, where the Aristotelian philosophers dominated, Epicurus moved to Lesbos. However, his philosophy was again not accepted; even worse Epicurus was accused of serious crimes and had to escape leaving the island in stormy weather, to reach the Asia Minor coast.

At the time, there were several Greek cities-colonies in Asia Minor. These cities were considered more progressive than the mainland cities; a suitable environment for the Philosophy of Epicurus. For instance it was not uncommon for the women of the colonies to get an education, in fact Asia Minor was the birthplace of many famous hetairas who are remembered by history. Hetairas are thought to be highly educated, high-end courtesans, companions of philosophers. One such example is the eminent Aspasia of Miletus, companion of

Pericles who was the most renowned Athenian leader in the 5th BC century, the golden age of the Ancient Greek world. Aspasia is said to have written Pericles' most inspiring speeches.

If there was any place in the Ancient world that could have accepted the revolutionary Epicurean philosophy, this was Asia Minor. Indeed Epicurus established a famous school in the city of Lampsacus [1] which, similar to the school of Pythagoras, accepted women and slaves among the students, something unacceptable at the elitist schools of Athens at the time. To put the matter in some perspective, Girton College in Cambridge was established as late as 1869AC to accept female students [4]. Epicurus, similar to Pythagoras, accepted students based on merit, not gender or status. He was even accused of praising his female student Themista, wife of Leonteus more than worthy male philosophers. Themista, according to one source [1], had previously been a hetaira. In 307BC, Epicurus, having gathered several followers returned once again in Athens, where he died in his early 70s [1, 2].

The world of Epicurus, similar to Democritus, consists of a world of interacting atoms, which account for the perceived properties of bodies. According to Epicurus, perception can be explained based on the "interaction of atoms with the sense-organs" [5]. He added to the atomist theory the notion that "all perception is true" and that we can learn to pursue natural pleasures rather than misleading desires imposed by society [2]. This view is much compatible to the general framework of Reinforcement Learning, where an agent interacts with its environment and receives feedback from it. It is reflected even closer by a most elaborated version of the agent-environment interaction proposed by Andy Barto and colleagues [6], according which the external world provides "sensations", that are internally interpreted by the agent (or organism) as reward signals, and that together with the perceived internal states will lead to decisions and actions. In the Epicurean philosophy, there is a clear mapping of good as pleasure, or in the Reinforcement Learning terminology positive reward, and evil as pain, or negative reward, punishment. There is also a clearly defined objective function:

«...τὴν ἡδονὴν ἀρχὴν καὶ τέλος λέγομεν εἶναι τοῦ μακαρίως ζῆν» [7]

"We say that pleasure is the beginning and the end of a happy life",

Epicurus states in his letter to Menoeceus [8].

However, pleasure in the Epicurean philosophy does not merely have the notion of an immediate reward, as perhaps misinterpreted by the critics (or the followers!) of Epicurus. Instead, it has the notion of a function to be maximised given future rewards or punishments, as the consequence of present actions. In the simplest form, we can write pleasure as a function R, parametrised by

time t, which consist of a finite series of rewards of the future r at time t+1, t+2, t+3...T:

$$R_t = r_{t+1} + r_{t+2} + r_{t+3} + \ldots + r_T$$

As in Reinforcement Learning [9], the reward r, in this context, can be positive, zero or negative, capturing also potential neutral or negative effects of an action. According to Epicurus, choices of actions depend on future long-lasting pleasures:

"For we know pleasure as our first and inborn good, and we have it as the beginning of our every choice and avoidance, and we turn to it in using sensation as a standard for judging every good. And since pleasure is our first and natural good, for this reason we do not choose every pleasure, but there are times when we pass over many pleasures, when greater difficulty would result from them for us, and we consider many pains to be better than pleasures, whenever a greater and long-lasting pleasure will follow for us after we have suffered the pains. So every pleasure is a good through being naturally agreeable, nevertheless not every pleasure is to be chosen; just as it is also the case that every pain is bad, but not every pain is always by its nature to be avoided." (Translation taken from [8]).

In this context, the word "pleasure" has two different meanings, one referring to immediate pleasures (or pains), and one referring to the general concept of pleasure, the objective function that is captured by R_t. In the same letter, the Philosopher offers advices regarding the future consequences of the present actions. This is equivalent of thinking that the Philosopher has obtained, by his experience and observation, the state-action values Q(s,a) defined in Temporal-Difference Reinforcement Learning, which expresses the total reward that can be collected by an agent in state s when taking action a. This prior knowledge of the Q-values might save others the time and effort that would be required to discover if an action is beneficial or harmful for their "souls", or which types of pleasure should be pursued:

"For continual drinking and partying, or taking one's enjoyment of boys and women [...] do not produce a pleasant life, but sober reasoning which both examines the basis for every choice and avoidance and drives out the opinions which cause very great turmoil to take hold of our souls." (Translation taken from [8]).

This is precisely the turning point for all the presumed ideas of a simplistic hedonism in the Epicurean philosophy. It is not just the pleasure of now that one has to be concerned with, but its future consequences. Drinking, for instance, would have an immediate posi-

tive reward at time t+1 but a negative reward at some other time in the future (i.e., hangover!). In general, the values of the actions in the Epicurean Philosophy seem independent of the state of the agent, as they capture rather general concepts. For instance, Epicurus valued friendship beyond other pleasures and advised against pursuing romantic love too much as this often brings jealousy and pain in the long run. Considering friendship investment as an action, its value is independent of the state of the agent (human) and higher than the value of the action of pursuing romantic love. In fact, to the best of my knowledge, Epicurus never got married. He lived a long life, and a happy one to the end, despite the difficulties that he faced. Epicurus, in his letters, advised that death is nothing to be afraid of as when we are here, death is not present, and when death is present, we do not exist [8].

The evaluation of an action based on the addition of the consequent punishments and rewards is not unique in the Epicurean philosophy. Plato had earlier suggested much the same. In Protagoras, Socrates argues:

"Then your idea of evil is pain, and of good is pleasure. Even enjoying yourself you call evil whenever it leads to the loss of a pleasure greater than its own, or lays up pains that outweigh its pleasures." (Translation taken from [10])

Further in the text, however, Socrates claims that it is difficult to measure good and evil, it requires a special skill, an art or science, and Protagoras agrees with him. This claim goes against the idea of Reinforcement Learning, where a naive agent can learn a task with a trial and error approach (exploration/exploration).

In Reinforcement Learning, the reward function is formulated identical to the pleasure function we defined above for finite time. In the case of an infinite series of future rewards, it includes a discount factor as a mathematical necessity. The discount factor also captures the fact that an immediate reward is preferable to a future reward of the same size: we would rather have a million pounds now than in five years time. I have not found the notion of discounting future rewards in the Epicurean philosophy, and for this, the pleasure function is defined as a finite series. This simplification, however, does not affect the importance of the Epicurean philosophy, which can be viewed as a sophisticated reward maximisation practice.

Why is Epicurus, with his progressive thinking, only at the corner of Raphael's painting, and in the mind of most merely a hedonist? For one, most of his work is not available; what we know of him today is mainly due to the letters he wrote (Letter to Herodotus, Menoeceus, Pythocles) [1, 2]. The prolific writing of Plato and Aristotle also mesmerised academics and didn't let much space for appreciation of the Epicurean philosophy available via limited sources. There may be one more reason though; Plato with his theory of Ideas

(Forms) and Aristotle with his notion of the final cause [2] offered a framework very much compatible with the existence of a God. Epicurus with his Democritean view of the world did not impose Gods as an implicit yet strong necessity of his philosophy. On the contrary, he believed that the soul died with the body and that he could disprove the notion of afterlife [2].

Is Epicurus the father of Reinforcement Learning? Perhaps most, including the author, would save the term for Rich Sutton and Andy Barto who wrote the bible of Reinforcement Learning [9]. Indeed, in the Epicurean philosophy there is no algorithmic procedure for learning the action values similar to Temporal Difference Reinforcement Learning. Instead, the Philosopher suggests to use reason and experience [11] to choose the best action in terms of immediate and future pleasure, avoiding these that can cause great "turmoil" in the future. In this sense, the Epicurean model does not include a sufficiently detailed exploration/exploitation approach as in Reinforcement Learning, and yet it does state that sensation can be used to judge every good [8]. On the contrary, Plato considers the idea of knowing good and evil a difficult matter that requires specialised expertise [10], while Aristotle, similar to Epicurus [11], appears to have embraced learning by trial and error [12] with his following statement, loosely linked to Reinforcement Learning:

"For the things we have to learn before we can do them, we learn by doing them" (Translation taken from [13]).

Aristotle, however, believed in an intrinsic teleology, explaining nature and objects via the purpose that they serve. In the context of his *Metaphysics* [14], this idea might suggest that brains have been designed for a specific purpose. To Aristotle, God is "the ultimate cause and explanation of rational change and order in nature" [14], while Epicurus explained physical phenomena in a proto-Darwinian way, as "a process of natural selection" [5], a truly advanced idea for his time. The Epicurean understanding of the world implies that brains have been evolved rather than designed to purpose.

Epicurus' view of pleasure appears to be tightly related to the concept of Dynamic Programming, a key ingredient of Temporal Difference Reinforcement Learning. In particular, it reminds of the Bellman optimality equation, which determines that the unique, optimal policy of choosing actions is greedy (i.e., choose the action with the highest Q-value) yet assumes that Q-values are known or can be explicitly calculated [9]. As such, the notion of the Reward function in Reinforcement Learning, as a finite sum of present and future rewards is perfectly and, perhaps, surprisingly captured by the philosophy of Epicurus.

Bibliography

[1] *Encyclopaedia Britannica*. Encyclopædia Britannica, Inc., 2010.

[2] *The Stanford Encyclopedia of Philosophy*. The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 2012.

[3] Diogenes Laertius. *Lives of Eminent Philosophers, Book IX, Chapter VII Democritus*. `http://remacle.org/bloodwolf/philosophes/laerce/9democrite.htm`: Charpentier, Libraire-Editeur, 1847.

[4] *Girton College*. `https://www.girton.cam.ac.uk/girtons-past`.

[5] Tim O'Keefe. *Epicurus*. The Internet Encyclopedia of Philosophy, ISSN 2161-0002 `http://www.iep.utm.edu`. 2017.

[6] A. G. Barto. "Intrinsically motivated learning in natural and artificial systems". In: Springer Berlin Heidelberg, 2013. Chap. Intrinsic motivation and reinforcement learning.

[7] Epicurus. *Letter to Menoeceus*. Reproduced in Lives of Eminent Philosophers, Book X `http://www.perseus.tufts.edu/hopper/text?doc=urn:cts:greekLit:tlg0004.tlg001.perseus-grc1:10.1`.

[8] Epicurus. *Letter to Menoeceus*. Translated in English by The Philosophical Garden `http://www.philosophicalgarden.com`.

[9] Richard S. Sutton and Andrew G. Barto. *Introduction to reinforcement learning*. Cambridge: MIT Press, 1998.

[10] Plato. *Protagoras*. The Collected Dialogues of Plato. Pantheon Books, a Division of Random House Inc. NY, 1963.

[11] Carl J. Richard. *Twelve Greeks and Romans who Changed the World*. Rowman & Littlefield Publishers, 2003.

[12] W.K.C. Guthrie. *A History of Greek Philosophy: Volume 6, Aristotle: An Encounter*. Cambridge University Press, 1981.

[13] Aristotle. *Nicomachean Ethics*. Translated in English by W.D.Ross, `http://classics.mit.edu/Aristotle/nicomachaen.2.ii.html`.

[14] Vasilis Politis. *Aristotle and the Mataphysics*. Routledge, 2004.