



Article

# Rilkean Memories and the Self of a Robot

Antonio Chella 

RoboticsLab, Department of Engineering, University of Palermo, Italy & ICAR-CNR, 90133 Palermo PA, Italy; antonio.chella@unipa.it

Received: 21 February 2019; Accepted: 22 April 2019; Published: 25 April 2019



**Abstract:** This paper discusses the concept of Rilkean memories, recently introduced by Mark Rowlands, to analyze the complex intermix of hardware and software related to the self of a robot. The Rilkean memory of an event is related to the trace of that episode left in the body of the individual. It transforms the act of remembering into behavioral and bodily dispositions, thus generating the peculiar behavioral style of the individual, which is at the basis of her autobiographical self. In the case of long-life operating robots, a similar process occurs: the software of the robot has to cope with the changes that happened in the body of the robot because of damaging events in its operational life. Thus, the robot, in compensating the damages of its body, acquires a particular behavioral style. The concept of Rilkean memory is essential in self-adapting robotics technologies where human intervention on a robot is not possible, and the robot must cope with its faults, and also in applications concerning green robotics.

**Keywords:** robotics; self; embodiment; autobiographical memories; Rilkean memories

## 1. Introduction

A question in social robotics [1] concerns the possibility for the robot to have a *self*. This argument recently received attention not only in philosophical research but also in robotics and multi-agent systems community, see Chella & Manzotti [2] and Pitt [3].

The concept of self is complex and multifaceted (see, e.g., [4] for recent discussions). It is true that it may be questionable to know whether a robot may have a genuine self. However, the employment of robots in the study of the self may be useful in two ways: on the one side, philosophers, neuroscientists, and cognitive scientists may employ robots as innovative tools to better understand the self. On the other hand, the cognitive, neuroscientific, and developmental studies on the self may help to build better robot architectures and systems.

Two approaches are typically followed. An approach correlates the self of the robot to the capabilities to reason about its knowledge by employing meta-reasoning paradigms based on higher-order logic. In this framework, Bringsjord et al. [5] describe a complex reasoning system implemented on a NAO robot which can pass the human test of self-consciousness proposed by Floridi [6].

According to the other approach, a robot with a self is capable of recognizing its body in static or motion state. Then, a robot with a self can identify itself in front of a mirror, even in the presence of distractors, as described in [7], and it should be able to describe the movements of its body [8].

The emerging view of the self of a robot [9] concerns the capabilities to reflect about itself, its perceptions and actions during its life. The self of a robot thus grows up from the content of the agent perceptions, recalls, activities, and reflections in a coherent life-long narrative.

The current AI studies on robot capabilities take into account only partially the role of the body: the typical approaches consider a robot as a computer with some sensors and actuators (see, e.g., [10]). Notably, Pfeifer and Bongard [11] analyze the strong relationships between the robot body and its cognitive functions, along the lines of the philosophical framework of embodied cognition (see, e.g., [12]).

The concept of Rilkean memories, recently introduced by Mark Rowlands [13,14], is employed to analyze the complex intermix of hardware and software related to the self of a robot. Briefly, the Rilkean

memory of an event is related to the trace of that event left in the body of the individual. A Rilkean memory is in place when the act of remembering an episode is transformed into behavioral and bodily dispositions, and then in a peculiar behavioral *style* of the individual, which is at the basis of her autobiographical self.

Now, forms of Rilkean memories may be considered even for robots. The software of a life-long operating robot has to adapt to the many changes that occurred in the body of the robot, to cope with the damaging events that happened during its life. Thus, the robot software generates a peculiar behavioral *style* of the robot, at the basis of the robot autobiographical self. Preliminary ideas at the base of the present discussion are reported in [15].

These concepts are essential in all the robotics applications where human intervention on a damaged robot is not possible, and the robot must cope with its faults and malfunctions. These concepts are relevant also in the emerging field of green robotics because the reuse of a robot would avoid waste of electronics and plastic materials.

In the following, Section 2 briefly outlines the concept of Rilkean memories introduced by Mark Rowlands and their relations to the autobiographic self; Section 3 presents the applications of the idea of Rilkean memories to robotics; Section 4 describes examples taken from robotics literature, and finally, Section 5 discusses the proposed concepts and outlines some conclusions.

## 2. Rilkean Memories

Rilkean memories are a kind of memories recently introduced by Mark Rowlands [13,14], who took inspiration from the novel “The Notebooks of Malte Laurids Brigge” by Rainer Maria Rilke. Briefly, the *embodied* Rilkean memory of an event is related to the trace of that episode left in the whole body of the individual, and not only in her brain. According to Rowlands:

The memory has become “blood.” It is now “glance and gesture,” “nameless and no longer to be distinguished” from the person whose memory it was. ([14], 54).

Moreover:

This form of memory is, typically, embodied and embedded. It is a form of involuntary, autobiographical memory that is neither implicit nor explicit, neither declarative nor procedural, neither episodic nor semantic, and not Freudian ([13], 141).

The Rilkean memories of a person are responsible for the *style* of that person. Consider the motion style of a person: for example, the habit of a person may be to walk on a particular side of a path because of a traumatic episode occurred in the past. The person may not remember the episode explicitly, but the event entered into her *blood*, and it became a part of her behavioral style.

A Rilkean memory is in place when the act of remembering an episode by episodic memory is transformed into behavioral and bodily dispositions, even when the original content of the event may be lost.

Rowlands reports the following example. Consider the case of a runner that experiences calf muscle tears. Then, the coach suggests her to lift the knee higher to employ larger muscles better. The runner changes the posture but, shortly after, she experiences several knee problems. The runner forgot that in the past she already had the same sorts of knee problems and that her low knee posture was a way to cope with the knee problems. Now, there are two possibilities.

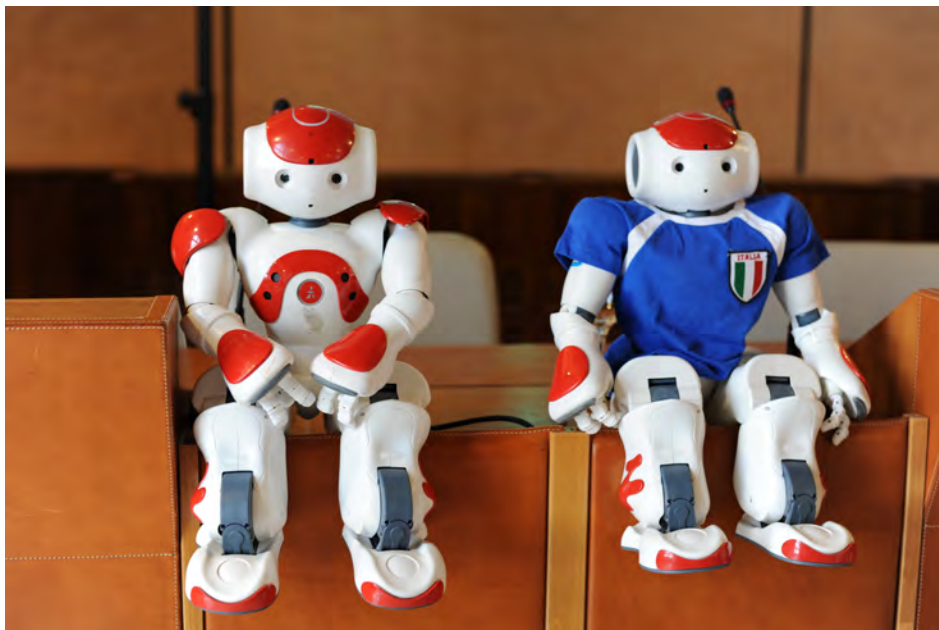
The first one is that the runner remembers her past knee problems through the episodic memory. The second case is that the runner does not consciously remember her issues. In this case, the low knee posture is the only remaining memory of her knee problems experienced in the past. According to Rowlands, both cases are examples of Rilkean memories. Notably, in the second case, the Rilkean memory is the single link to the past experiences of the runner concerning her knee, because the relative episodic memory has vanished.

Rilkean memories are relevant to the self of a person. Widespread problems concerning the self are related to the identification of the unity and persistence conditions for that person, i.e., the uniqueness of the person to other persons. Rowlands maintains that there is a strong relationship between the autobiographical self and the Rilkean memories: they play an essential role in holding the self together thanks to the unity and persistence of the person's behavioral style. Different people, with different histories and various episodes, acquire a distinct and recognizable personal style. This style survives even when the episodic memories of the events happening in the life of the person are lost, as in the case for example of Alzheimer's disease. Rowlands asserts that:

Rilkean memories are like the dark matter of memory. They hold everything together. And because they are the result of transformation from episodic memories, they link us to our past in concrete and, often, important ways, and so provide a basis for the unity and coherence of the autobiographical self ([13], 154).

### 3. Robot Rilkean Memories

The concept of Rilkean memories is relevant in the case of life-long operating robots. Consider for example the NAO, a small humanoid robot built by SoftBank Robotics and commonly adopted in many research laboratories. The RoboticsLab of the University of Palermo acquired two new NAOs since 2010 (Figure 1). In the beginning, they were two samples of the very same robot. They had the same body and the same software. Thus, both robots were able to perform the same tasks in the same operational style.



**Figure 1.** The two NAOs of the RoboticsLab.

The two robots were employed for different research projects with different purposes. Over the years, the two NAOs slightly changed their morphology in different ways. After executing and repeating different actions, their motors did not work correctly in peculiar ways. Their joint movements were not regular enough, and several bumps occurred in various parts of their body.

At some point, a software bug on one NAO caused the NAO to bump against the wall. Because of this hit, the left leg joint started not to operate correctly, and a slight malfunction of the joint occurred. The Rilkean memory of the episode is in place: the memory of the bumping event is transformed into a minor malfunction of the joint controlling the left leg. Now, the robot may be able or unable to access the episode depending whether its log file has been deleted or not.

In case the robot log file is not deleted, the software system of the NAO may correctly attribute the cause of the malfunction to the previous bumping. In this case, the log file acts as a simplified form of episodic memory. Otherwise, the software system can detect the malfunction of the leg but without the possibility of attributing any cause. The past bumping event has been transformed into a peculiar walking style of the robot. This situation strongly resembles the case previously described concerning the Rilkean memory of the knee posture of the runner. In particular, in the case of log file deleted, the peculiar walking style of the robot is the only memory of the previous bumping event.

A similar case happened when a student dropped one of the robots when turned off and wrecked a joint. The occurred hardware damage is the Rilkean memory of the dropping event.

At some point, due to a series of traumatic events, a joint of one NAO broke and was renewed. The diagnostic system of the NAO allows it to recognize its parts, its operating states, and the occurring faults. So, when the joint is replaced, the robot can identify it as a new one and different from the old, broken joint. Then, the identification of the renewed joint is the Rilkean memory that summarizes the series of events that led to the break of the previous joint.

It is to be noticed the difference between the renewing of an ordinary object by substituting its parts and the renewing of the body parts of the robot. An object does not have reflective capabilities, so it is not able to report its inner state and to reason on new renewed parts which are different from the old ones.

Since the two robots have been employed in different experiments and for various purposes, after many years, eventually they own slight different bodies and different behavioral styles. Consider the case of the NAO that showed a problem with the joint controlling the left leg, which is the Rilkean memory of a previous bumping event, as previously discussed. The software system of this NAO will take into account the motion constraints of the joint to generate a suitable motion plan that minimizes the use of the compromised leg. Instead, the other robot does not present problems with the left leg, but instead with the right arm, due to a different traumatic event. In the same way, the software system of this robot will generate a different motion plan to minimize the movements of the damaged arm.

Moreover, one of the two NAOs tends to overheat a motor due to previous episodes of excessive use of the corresponding joint. Thus, this NAO has to interrupt its actions quite often, and then the robot performs its tasks in a characteristic, rough and fragmented behavioral style.

Another example is related to the robot cameras: the NAO has two cameras mounted at the top and the bottom of the head. One of the two NAOs has a broken camera due to a previous head crash, and then it has to move its head in a particular way to use the only operating camera to compensate for the breakdown. This peculiar head motion is the Rilkean memory of the previous head crash.

Therefore, in all the mentioned cases, the two NAOs acquire different Rilkean memories. It should be noticed that the Rilkean memories of the two robots have an immediate effect on the different behavior styles of the two robots: after years of operations, two different behavioral styles emerge. As described, these different operational styles dependent on the intricate interweaving of the two NAOs' hardware and software, and their various biographies.

When the two robots are brought together and involved in the same task, the software of each robot has to take into account the capabilities and hardware constraints to generate two different operational plans for the two robots, for the same task. The two robots perform the same job in two different ways, based on their different behavioral style acquired during their operational life.

#### 4. Examples from the Robotics Literature

Examples of robots discussed in the literature and relevant for the discussion of Rilkean memories in robots are the *starfish* robot by Bongard et al. [16] and the *hexapod* robot by Mouret and collaborators [17].

The starfish robot is a robot able to build a 3D model of itself thanks to a sort of motor babbling, i.e., by generating casual movements of its body parts and analyzing the feedback from the sensors. The robot employs a genetic algorithm to find the suitable combination of 3D elements to build a correct

representation of its body which is coherent with the sensorimotor actions. Thus, the robot employs the 3D model to explore the combination of movements of body parts that allow it to move forward.

After breaking a leg, the starfish robot is no longer able to walk. However, thanks to the inner 3D model, it can reconfigure itself by searching a suitable set of body movements that allow it to walk again with a limp, with a peculiar behavioral *style*. According to our discussion, the acquired walking style of the robot is the Rilkean memory of the episode of the leg break.

A sophisticated example is the previously cited hexapod robot. This robot can build a 3D model of its motion as in the case of the starfish. It employs this 3D model to simulate possible different styles of movements. Thus, the robot builds a sort of inner behavior space, where each point is a particular motion style. When damage occurs on a leg of the robot, then the robot performs a trial and error search to rebuild and correct the behavior space to find a suitable way to employ the remaining legs. Also, in this case, the new walking style is the Rilkean memory of the episode of the occurred damage.

As pointed by Rowlands concerning Nader [18], the personal style pervades the autobiographic memory of a person. For example, when a person starts walking with a limp after an accident occurs, then she may remember the episodes of her past life where she walked with a limp even before the damage occurred.

As in the cases of the starfish and the hexapod, a robot can create an internal 3D representation of its body. Now, consider the possibility that the robot can store in its episodic memory all the motions episodes during its entire operational life, before and after the leg damage occurred.

Thus, the robot equipped with a 3D model of itself may be able to recall an episode of itself walking through a door, which happened before the leg damage. It may be natural for the robot to rebuild the memory of the event by substituting the robot body as it is after the leg damage. Therefore, the robot may recall the episode of itself walking through the door, which happened before the leg damage, as if the event occurred with the injured leg. In this sense, following Rowlands, the motion style of the robot with a limp is a unifying entity of the autobiographic self of the robot.

## 5. Discussion and Conclusions

This paper discusses the concept of Rilkean memories in life-long operating robots to put into evidence the complex intermix of hardware and software related to the robot self. The software of the robot has to adapt to the occurring changes in the body to cope with the events occurring in the operating life of the robot. In particular, the software has to compensate for the damages and malfunctions occurring to the body of the robot through its operational life, generating a peculiar behavioral style of the robot.

The proposed discussion is relevant in the scarcely investigated problem of aging robots, and it may open a new line of research that parallels the well-known research in developmental robotics [19], i.e., how the studies of aging processes in humans (see, e.g., [20]) may inspire the studies of aging in robots.

Moreover, the discussion highlights the needs of the robot to self-adapt to the occurring body changes and malfunctions. Self-adaptation is necessary for many applications, as in battlefields, hostile environments, nuclear reactors, planet explorations, and in general in all the applications where human intervention on a damaged robot is not possible, and the robot must cope with its faults and malfunctions.

A present-day challenge in robotics technologies and more in general in Information Technology concerns sustainable and green robotics [21,22]. The capability of reusing and employing old robots, even with malfunctions and defects, goes in the direction of avoiding dangerous waste of electronics elements and plastic materials.

An essential aspect of the proposed discussion concerns the relationship between Rilkean and the autobiographical self of the robot. Rilkean memories in robots, different from the standard memories implemented as data structures in the software system of the robot, are related to the tight correlation between the hardware of the robot. The Rilkean memories play an essential role in holding the self of

the robot together thanks to the unity and persistence of the behavioral style of the robot. Different robots, with different histories and different occurring episodes, acquire distinct and recognizable personal styles. This style survives even when the episodic memories of the events in the operational life of the robot are lost.

However, not only the body of the robot changes during its operational life but also the software is not a static entity. Robotics software evolves by employing machine learning strategies (see, e.g., [23]). Then, a robot during its operational life becomes different from a brand new robot because the software structures evolve and its body changes.

Thus, a long-life operational robot possesses some form of uniqueness. It is true that the state of a robot could be copied to another robot; it may be more comfortable for the software part than for the hardware part, where it is necessary to reproduce all the hardware faults and malfunctions, but it may be possible from a theoretical point of view. However, the copy of the robot will be “a reproduction” and not the original robot.

It is also true that it is possible from a theoretical point of view to restore the robot to its initial conditions by substituting all the damaged body parts and by uploading a brand new version of its software. As previously discussed, in many scenarios the human intervention may not be practically possible, and the robot has to cope with its faults and malfunctions. However, restoring the robot to the initial conditions has the price for losing all the evolved hardware parts and software structures during the operational life of the robot.

In conclusion, a robot, after a long operational life, should be considered as a sort of *artwork*. Then, copying its state to another robot is similar to copying an artwork as a painting to a new canvas. It is possible, but we take into consideration the importance and uniqueness of the original artwork, even if it may be copied and reproduced. The same discussion holds in the case of the restored robot: it would be a similar operation as restoring a canvas to the brand new state by removing the artwork on it.

**Funding:** This research received no external fundings.

**Acknowledgments:** The author wants to thank Mark Rowlands and Riccardo Manzotti for their comments on previous versions of this paper.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Dumouchel, P.; Damiano, L. *Living with Robots*; Harvard University Press: Cambridge, MA, USA, 2017.
2. Chella, A.; Manzotti, R. (Eds.) *Artif. Consciousness*; Imprint Academic: Exeter, UK, 2007.
3. Pitt, J. (Ed.) *The Computer after Me*; Imperial College Press: London, UK, 2015.
4. Gallagher, S.; Shear, J. (Eds.) *Models of the Self*; Imprint Academic: Exeter, UK, 1999.
5. Bringsjord, S.; Licato, J.; Govindarajulu, N.S.; Ghosh, R. Sen, A. Real Robots That Pass Human Tests of Self-consciousness. In *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, Kobe, Japan, 31 August–4 September 2015; IEEE Press: Piscataway, NJ, USA, 2015; pp. 498–504.
6. Floridi, L. Consciousness, Agents and the Knowledge Game. *Minds Mach.* **2005**, *15*, 415–444.
7. Gold, K.; Scassellati, B. Using Probabilistic Reasoning over Time to Self-Recognize. *Robot. Auton. Syst.* **2009**, *57*, 384–392.
8. Chella, A.; Gaglio, S.; Pirrone, R. Conceptual Representations of Actions for Autonomous robots. *Robot. Auton. Syst.* **2001**, *34*, 251–263.
9. Chella, A.; Frixione, M.; Gaglio, S. A cognitive architecture For Robot Self-consciousness. *Artif. Intell. Med.* **2008**, *44*, 147–154.
10. Levesque, H.; Lakemeyer, G. Cognitive Robotics. In *Handbook of Knowledge Representation*; Van Harmelen, F., Lifschitz, V., Porter, B., Eds.; Elsevier: Amsterdam, The Netherlands, 2008; pp. 869–886.
11. Pfeifer, R.; Bongard, J. *How the Body Shapes the Way We Think*; MIT Press: Cambridge, MA, USA, 2007.
12. Shapiro, L. *Embodied Cognition*; Routledge: New York, NY, USA, 2011.
13. Rowlands, M. Rilkean Memory. *South. J. Philos.* **2015**, *53*, 141–154.

14. Rowlands, M. *Memory and the Self: Phenomenology, Science and Autobiography*; Oxford University Press: Oxford, UK, 2016.
15. Chella, A. Rilkean Memories for a Robot. In Proceedings of the AISB Annual Convention, Symposium on Assessing Agency, Moral and Otherwise: Beyond the Machine Question, Liverpool, UK, 4–6 April 2018.
16. Bongard, J.; Zykov, V.; Lipson, H. Resilient Machines Through Continuous Self-Modeling. *Science* **2006**, *314*, 1118–1121.
17. Cully, A.; Clune, J.; Tarapore, D.; Mouret, J.B. Robots That Can Adapt Like Animals. *Nature* **2015**, *521*, 503–507.
18. Nader, K. Memory traces Unbound. *Trends Neurosci.* **2003**, *26*, 65–72.
19. Cangelosi, A.; Schlesinger, M. *Developmental Robotics—From Babies to Robots*; MIT Press: Cambridge, MA, USA, 2015.
20. Craik, F.; Salthouse, T. (Eds.) *The Handbook of Aging and Cognition*, 2nd ed.; Lawrence Erlbaum: Mahwah, NJ, USA, 2000.
21. Alves Filho, S.E.; Sá, S.T.L.; Burlamaqui, A.M.F.; Aroca, R.V.; Gonçalves, L.M.G. Green Robotics: Concepts, Challenges, and Strategies. *IEEE Lat. Am. Trans.* **2018**, *16*, 1042–1050.
22. Murugesan, S. Harnessing Green IT: Principles and Practices. *IT Prof.* **2008**, *10*, 24–33.
23. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2017.



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).