Towards a Cognitive Humanoid Dancer

Agnese Augello,[1] Antonio Chella,[2] Salvatore Gaglio,[1,2] Ignazio Infantino,[1]
Giovanni Pilato,[1] Riccardo Rizzo,[1] and Filippo Vella[1]

***Abstract*— We outline a cognitive architecture for a humanoid robot implementing creative dancing behaviors. The architecture, starting from a repertoire of patterns of movements, is able to generate improvised dance actions and to evaluate them. The creative process is supported by the cognitive architecture which allows to robot to take into account a motivations mechanism in order to create robot choreographies.**

I. Introduction

A dancer, during a free, improvised performance should be able to perceive and to analyze music and at the same time to generate suitable body movements according with rhythm in order to let the audience to feel emotions. To accomplish these tasks, he/she should be able to interact with the mate dancer and to negotiate free space with other dancers; he/she should be able to figure out himself from the point of view of the audience, and so on. Dancing is therefore a complex cognitive ability [1].

The aim of this paper is to outline a cognitive architecture for a humanoid robot able of autonomously creating and executing different dancing actions, i.e., sequences of movements involving different parts of the robot body while perceiving music.

Some autonomous dancing robots have been proposed in the literature, see [2], [3], [4], [5]. The architecture outlined in this paper takes as a source of inspiration many cognitive studies about dance.

The creation of free, improvised choreographies by a robot is an important challenge for the research field of computational creativity [6]: in fact, an artificial agent able to improvise dances needs an *embodied* and *situated* creativity.

In order to evaluate such a robot system from the point of view of computational creativity, it is necessary to analyze the main aspects characterizing the creativity of the robot, both in general and with specific reference to the domain of dance. For this reason, the system will be assessed according to a well-defined methodology, following the guidelines of the SPECS procedure discussed by Jordanous [17], which has been proposed as a standard procedure for evaluating the creativity of a computational system.

II. Humanoid Dancer Architecture

Gärdenfors [8] discusses a program for musical analysis inspired by the framework proposed by Marr [9]. According to Gärdenfors, perception of time in music is processed by a hierarchy of levels related with temporal intervals, beats and rhythmic patterns. At the fourth and last level, pitch and time merge together.

Schack [10] proposes a similar hierarchy of action levels related with dance: the first level is the level of *sensorimotor control* and it is based on reactive behaviors; the second level is related to *sensorimotor representation*, responsible for information related with specific, automatic, movements. The third level is the *mental representation* level; this level stores mental representation of movements and it is able to transform desired outcomes in suitable sequences of movements. The fourth level is the *mental control* level and it is related with *basic action concepts* (BAC). At this level, different movements generated at the previous level are suitably chunked in order to generate complex actions and figures, as a *pirouette*.

Following the hierarchy of actions, Schack outlines a cognitive architecture integrating the level of action organization with appraisal of emotions and motivations. According to Schack, motivations and emotions directly influence the decision of the action plan to be performed. The architecture proposed by Schack also takes into account the case of problems occurring during the movements of the dancer: a failure in performance may negatively affect the motivations of the dancer himself.

The cognitive architecture discussed in this paper is inspired by the Psi model originally proposed by Bach [11] and then refined by Dörner and colleagues [12]. The Psi model in fact takes into account the tight link between motivations, perception, action planning and actions execution outlined by Schack. A previous application of the Psi model to computational creativity is discussed in [13].

A simplified diagram of the implemented architecture is shown in Fig. 1; the architecture is embedded in the Aldebaran NAO® humanoid robot. NAO is a robot presenting many useful characteristics. Notably, the robot operating system manages many complex low level tasks. Moreover, NAO is adopted by a large community of open source developers; it is in fact a standard humanoid platform.

III. Operation of the Cognitive Architecture

Briefly, the creative process is accomplished by means of a genetic algorithm (see [14] for a review) that allows for the generation of robot *sensorimotor patterns* each one encoding a rhythm pattern and the associated BAC.

The patterns are the repertoire of the robot, they are stored into the Long Term Memory (LTM) and they are analyzed

[1]Agnese Augello, Salvatore Gaglio, Ignazio Infantino, Giovanni Pilato, Riccardo Rizzo and Filippo Vella are with the Italian National Research Council, Viale delle Scienze, Building 11, Palermo, Italy augello@pa.icar.cnr.it

[2]Antonio Chella, Salvatore Gaglio are with University of Palermo, Viale delle Scienze, Building 6, Palermo, Italy

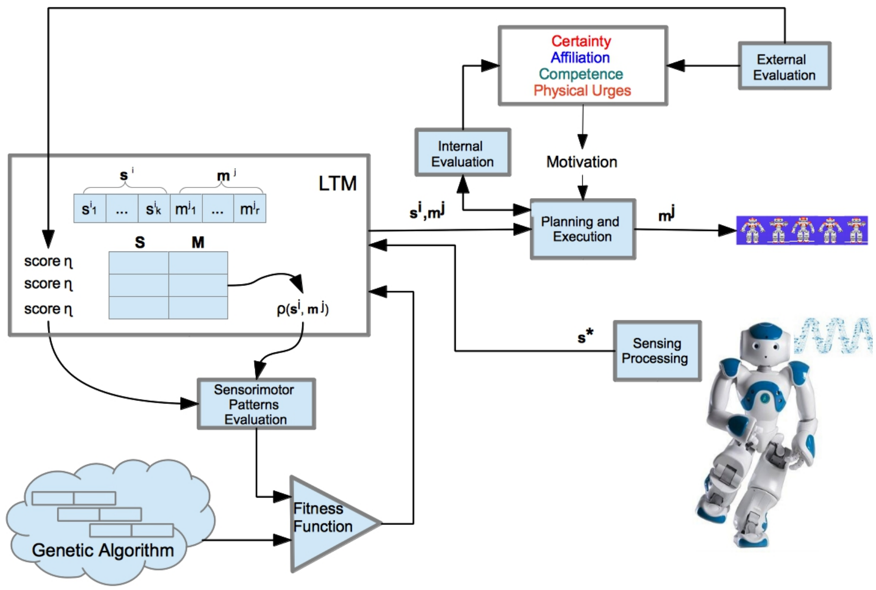

Fig. 1. The Humanoid Dancer cognitive architecture

during the execution phase by a *planning and execution* module. This module processes the perceived rhythm and searches the LTM for the most adequate BAC to be associated with that rhythm.

A motivation parameter influences the planning and execution module. Motivation is in turn indirectly influenced by external and internal evaluations through *urges*, i.e. components of the cognitive architecture able to reflect specific demands. Examples are physiological drives as *Physical Urges*, social drives as *Affiliation*, and cognitive drives, as *Certainty* and *Competence* [11] [15].

More in details, the LTM stores the robot repertoire as a set of binary vectors $\mathbf{d}^{(i,j)} = [s^i_1, s^i_2, \cdots, s^i_k, m^j_1, m^j_2, \cdots, m^j_r]$ whose dimensionality is $k + r$.

The first k bits, identified by the vector $\mathbf{s}^i$, code the rhythm pattern perceived by the robot, while the last r bits, identified by the vector $\mathbf{m}^j$, encode a BAC. Here, i and j are used in order to identify different configurations of bits, corresponding to generic rhythm patterns perceived by the robot and different BACs performed by the robot.

At the beginning, the repertoire is a random population of vectors $\mathbf{d}^{(i,j)}$. Then, the fitness function:

$$\rho(\mathbf{s}^i, \mathbf{m}^j) : \{0, 1\}^{k\times r} \rightarrow \mathbb{R}$$

is computed between the vectors $\mathbf{s}^i$ and $\mathbf{m}^j$ of each $\mathbf{d}^{(i,j)}$. The function $\rho(\mathbf{s}^i, \mathbf{m}^j)$ characterizes the specific *attitude* of the robot dancer. According to the standard genetic algorithm paradigm, the vectors $\mathbf{d}^{(i,j)}$ related with a fitness measure above a fixed threshold will survive and they will be recombined, constituting at the end the personal dancing *style* of the robot.

The fitness score encodes a sort of internal *resonance* between what the robot perceives (the rhythm stimulus) and how it reacts to the external stimuli, i.e. the corresponding BAC: if the BAC resonates with the perceived rhythm, as it happens for example when the BAC allows the robot to move its arm in sync with the *tactus* of the perceived rhythm pattern, then the fitness function ρ is high.

The rhythm patterns perceived by the NAO are processed by means of standard digital audio processing functions that code suitable rhythm features as a binary vector $\mathbf{s}^*$ of dimensionality k. The association mechanism allows for triggering of the BAC performance by encoding a dynamic binding that is at the same time internally and externally evaluated.

At regular intervals, the robot looks for a pattern $\mathbf{s}^*$ in its set of vectors $\mathbf{s}^i$ in order to find a match. If a match is found between $\mathbf{s}^i$ and $\mathbf{s}^*$, the corresponding BAC $\mathbf{m}^j$ of the vector $\mathbf{d}^{(i,j)}$ is executed until the next rhythm arrives. If more than one matching vectors $\mathbf{s}^i$ is found, then the sub vector $\mathbf{m}^j$ of $\mathbf{d}^{(i,j)}$ that has the highest fitness value is chosen. When two vectors $\mathbf{d}^{(i,j)}$ have the same fitness value, a random choice is then done among their corresponding sub vectors $\mathbf{m}^j$. If no match is found at all, then the vector $\mathbf{s}^i$ whose cosine with $\mathbf{s}^*$ is highest is chosen instead.

After the execution of a set of BACs, an external feedback from the audience is given. The feedback concerns the whole performance of the robot, and it is expressed as a score $\eta \in R$ ranging from -1 to $+1$ and it is determined by trials.

The NAO therefore associates the FF score:

$$FF^{(i,j)} = \alpha\rho(\mathbf{s}^i, \mathbf{m}^j) + (1 - \alpha)\eta$$

to the vector $\mathbf{d}^{(i,j)}$, where $\alpha \in \Re$ is $0 \leq \alpha \leq 1$.

After a set of evaluations, when the robot is "tired" or non motivated, then an introspective, self-criticism, process starts. During this process, the NAO analyzes the set of employed patterns and it re-evaluates them taking into account their associated $FF^{i,j}$ values.

The dancer then chooses a threshold so to erase the patterns having an $FF^{i,j}$ value below this threshold. The surviving patterns, together with the non used patterns, will constitute a new repertoire which will evolve as described above.

IV. Motivation

As previously stated, motivation influences the creative behavior of the dancer agent. The motivation mechanism is implemented by means of a real values parameter. This parameter reaches high values when all urges are satisfied, and it is lowered when some urgency is not satisfied yet.

When the motivation parameter is below a given threshold, the robot will react by performing some action. If this reaction has no effect and the motivation is still below the threshold, the robot stops the execution of the current task, and it executes instead some actions aimed at satisfying pressing needs (for example, to go to the charging station).

Here, we focus on cognitive urges as defined in [15]: the *Competence* which is *the effectiveness of the agent at fulfilling its needs*, e.g., the effectiveness at correctly performing a set of BACs, and the *Certainty* which is *the confidence of the agent's knowledge*, e.g., the amount of dancing style judged as being *pleasant*.

Internal and external evaluations will influence the cognitive urges, as shown in Fig. 1. As fully described in [7], the *Competence* will be influenced by the internal evaluation process, while the external one will have an impact on the *Certainty* urge.

The values of *Competence* and *Certainty* determine a degree of the dancer evolution. As an example, a *beginner* is a dancer having both low *Competence* and *Certainty*, while a dancer that has both of them high, is an acclaimed artist. Different combinations of these values could be referred to other interesting states of the performing dancer.

V. Evaluation of the proposed system

The robot dancer, during the execution of a BAC, evaluates its ability in order to accomplish the employed technique according to some external and internal expectations.

Internal expectations come from the comparison between an internal model and the product of execution of the BAC, considering objective parameters and aesthetic judgements. The adopted internal model is similar to the one described in [16], and it is part of the planning and execution module.

This self-evaluation then concerns the similarity between what is produced by the agent and the generated expectations. If the expectation of the agent is disregarded, and the outcome of a BAC does not reach the goal, something in the technique, style, movement should be changed. Instead, a correct accomplishment of a technique allows the agent to achieve a positive evaluation, with the consequence that the value of his competence increases, and consequently its motivation. The internal evaluation affects the value of the *Competence* urge.

The external evaluation regards a post-production evaluation of the dance, obtained by asking the audience to express their judgements. This opinion has an impact on the style used to produce that result. The external evaluation affects the *Certainty* of the robot.

As previously stated, future works will concern a structured evaluation of the robot dancer from the point of view of creativity according to the guidelines proposed in [17].

References

[1] B. Bläsing, M. Puttke, and T. Schack, Eds., *The Neurocognition of Dance - Mind, Movement and Motor Skills*. Hove, UK: Psychology Press, 2010.

[2] G. Xia, J. Tay, R. Dannenberg, and M. Veloso, "Autonomous robot dancing driven by beats and emotions of music," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, AAMAS'12. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2012, pp. 205 – 212.

[3] T. Takeda, Y. Hirata, and K. Kosuge, "Dance step estimation method based on HMM for dance partner robot," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 2, pp. 699 – 706, 2007.

[4] E. Avrunin, J. Hart, A. Douglas, and B. Scassellati, "Effects related to synchrony and repertoire in perceptions of robot dance," in *Proceedings of the 6th International Conference on Human-robot Interaction*, HRI '11. New York, NY: ACM, 2011.

[5] M. Michalowski, R. Simmons, and H. Kozima, "Rhythmic attention in child-robot dance play," in *18th IEEE International Symposium on Robot and Human Interactive Communication*, RO-MAN 2009. IEEE Press, 2009, pp. 816 – 821.

[6] J. McCormack and M. d'Inverno, Eds., *Computers and Creativity*. Heidelberg: Springer, 2012.

[7] A. Augello, I. Infantino, G. Pilato, R. Rizzo, and F. Vella, "Creativity evaluation in a cognitive architecture," in *Proceedings of the 2014 Annual International Conference on Biologically Inspired Cognitive Architectures (BICA 2014)*, Boston, MA, 2014.

[8] P. Gärdenfors, "Semantics, conceptual spaces and the dimensions of music," in *Essays on the Philosophy of Music*, V. Rantala, L. Rowell, and E. Tarasti, Eds. Helsinki: Philosophical Society of Finland, 1988, pp. 9–27.

[9] D. Marr, *Vision*. New York: W.H. Freeman and Co., 1982.

[10] T. Schack, "Building blocks and architecture of dance," in *The Neurocognition of Dance - Mind, Movement and Motor Skills*, B. Bläsing, M. Puttke, and T. Schack, Eds. Hove, UK: Psychology Press, 2010, pp. 11 – 39.

[11] J. Bach, *Principles of synthetic intelligence - Psi: an architecture of motivated cognition.* Oxford, UK: Oxford University Press, 2009.

[12] C. Bartl and D. D., "PSI: A theory of the integration of cognition, emotion and motivation." in *Proceedings of the 2nd European Conference on Cognitive Modelling*, 1998, pp. 66 – 73.

[13] A. Augello, I. Infantino, G. Pilato, R. Rizzo, and F. Vella, "Introducing a creative process on a cognitive architecture," *Biologically Inspired Cognitive Architectures*, vol. 6, pp. 131 – 139, 2013.

[14] P. Husbands, P. Copley, A. Eldridge, and J. Mandelis, "An introduction to evolutionary computing for musicians," in *Evolutionary Computer Music*, E. Miranda and J. Biles, Eds. Heidelberg: Springer, 2007, pp. 1 – 27.

[15] Z. Cai, B. Goertzel, and N. Geisweiller, "OpenPsi: Realizing Dörner's "Psi" cognitive model in the OpenCog integrative AGI architecture," in *Artificial General Intelligence*. Heidelberg: Springer, 2011.

[16] A. Chella, S. Gaglio, and R. Pirrone, "Conceptual representations of actions for autonomous robots," *Robotics and Autonomous Systems*, vol. 34, pp. 251–263, 2001.

[17] A. Jordanous, "A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative," *Cognitive Computation*, vol. 4, no. 3, pp. 246 – 279, 2012.