Synthetic temperament of robots

Nikola Petrović, Andrea Bonarini and Aleksandar Rodić

Abstract—**The paper proposes a novel model to design synthetic temperament, based on classic child development research conducted by Thomas, Chess and Birch. The advantages of children's personality psychological models over models related to adult humans' personality are discussed from the perspective of usage in robotics. The mapping between children's temperamental traits and mobile robot behavioral patterns are proposed. Fuzzy set theory is used to map natural language expressions, used to describe children's temperamental traits in the psychological theory, to measured sensory inputs and robot's actuators outputs.**

Index Terms—**Synthetic temperament; Synthetic persona; Artificial personality.**

I. Introduction

There are many definitions of personality proposed by scientists based on their theoretical positions (1). For the purpose of this paper, personality is a set of psychological characteristics that uniquely influence a person's behavior patterns, cognition, and emotions. Most of the scientists agree that heredity and environment interact to determine one's personality. The personality in children is often referred to as temperament. The temperament is based on heredity and represents built-in features that are expanded by interaction with environmental influences in the formation of an adult human's personality.

Since Hippocrates, philosophers and scientists have tried to create personality models and to classify personalities into personality (or psychological) types. Most personality models are based on a set of personality traits, enduring personal characteristics that are revealed in a variety of situations. One of the frequently implemented models, nicknamed the "Big Five" or "FFM", defines openness to experience, conscientiousness, extroversion, agreeableness, and neuroticism (or emotionality) as major personality traits. The classifications of personalities are usually based on binary discretization of personality traits. For example, regarding extroversion, people are classified as introverts and extroverts. Building on the theoretical work of Jung (2), Isabel Briggs Myers and her mother, Katharine C. Briggs, constructed the Myers-Briggs Type Indicator (MBTI) assessment, a psychometric questionnaire designed to classify people in 16 psychological types.

Robots that have to interact with people, virtual personal assistants and other intelligent agents are more effective if they manifest a personality, perceive the personality of the intelligent agents they interact with and adapt their behavior to the personality of humans they serve. Synthetic personality can be exploited for robots in order to make them more believable. This would enable people who interact with robots to develop deeper relationships with them. Yet another use of personality models in robotics is to test psychological theories in experiments using controlled scenarios that include robots.

Just like with human personality, some aspects of synthetic personality may be innate while others may be developed by learning. This paper refers to the innate, built-in base of synthetic personality as synthetic temperament.

A milestone on the road toward the robots that can manifest and detect human-like personality may be the robots that manifest a child-like personality, the synthetic temperament, and detect temperament of the agents in their environment. This paper proposes a model of synthetic temperament in mobile robots, assuming that even motion patterns could convey temperament, and that people may attribute affect and personalities to movements of even non-bio-similar entities based on the way they move, only. (3)

II. State of the Art

There are several theoretical models and implementations of intelligent and emotional agents that use existing personality models as a component. Clarke Elliott's Affective Reasoner (4), Layered Model of Affect - ALMA (5), Artificial Emotion Engine (6) and SIMPLEX (7) are computational models of affect that represent emotions, moods, personality, and their particular relations. Rodić and Addi developed EI-controller (8), a model of emotion-driven behavior in robots based on Myers-Briggs theory on personality. The serious problems in implementation of those models are the mappings of detected behavior of persons in the environment into personality traits, and mappings of agent's own personality traits into behavior, or modification of behavior. The source of the problems is that personality traits, like neuroticism or extroversion, are high level concepts, and their detection and behavioral implementation are too complex for cognitive and interaction features of today's robots, out of narrow use cases and laboratories.

Julian and Bonarini (9) reported the results obtained from two experiments performed to study whether features different from face and bio-inspired bodies could convey emotions. The study has been done with a non-bio-inspired robot base, having an intentionally unusual (for a robot) shape. The results show that it is possible to convey emotions using features that can be implemented also in a non-bio-similar embodiment. It is not well known that the Disney's team at the beginning of the cartoon era was successfully exploring the possibility to express emotions with "the simplest of shapes": the half-filled flour sack (10).

Rudolf von Laban, one of the pioneers of modern dance in Europe, laid the foundations of methods and language for describing, visualizing, interpreting and documenting all varieties of human movement (11). "Laban Effort" is a

Nikola Petrović is with Morena inženjering, Bulevar Nikole Tesle 17/3 18106 Niš, Serbia (e-mail: nikola@morenaict.com). Andrea Bonarini is with the AI & Robotics Lab, Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy (e-mail: andrea.bonarini@polimi.it).

Aleksandar Rodić is with the Robotics Laboratory, Mihailo Pupin Institute, Volgina 15, 11060 Belgrade, Serbia (e-mail: aleksandar.rodic@pupin.rs).

system for understanding the way a movement is done with respect to inner intention. The paper (3) describes adaption of the Laban Effort System to authors' motions control of flying robots, and the results of a formal experiment that investigated how various Laban Effort System parameters influence people's perception of the resulting robotic motions. Laban notation is being considered in the human-robot interaction community a useful tool to describe motion, and, in particular, emotional movement, but Laban's contribution is mainly aimed at representing dance movement for people, giving for granted the qualitative terms used in the descriptions, which should reach the specificity of control actions to be implemented in robotics. The same applies to our reference model, described in the next section.

III. Theoretic Background

The New York Longitudinal Study (12), started in 1956 and continued over several decades thereafter, is regarded as a classic study into temperament traits. The study, conducted amongst young children by Thomas, Chess and Birch identified nine temperamental traits that could be reliably scored on a three-point scale (medium, high and low): activity level, distractibility, intensity, regularity, sensitivity, initial reaction, adaptability, persistence and default mood.

1. *Activity Level*, the level and extent of motor activity. This is the child's idle speed or how active the child is generally. Does the child use gross motor skills like running and jumping more frequently or rely more on fine motor skills, such as drawing and putting puzzles together? The available scores, as proposed by Thomas, Chess and Birch (13), are HIGH, MEDIUM, and LOW.

2. *Distractibility* refers to the ease with which external stimuli interfere with ongoing behavior. An easily distracted child is engaged by external events and has difficulty returning to the task at hand, whereas a rarely distracted child stays focused and completes the task at hand. The available scores are DISTRACTIBLE, MEDIUM and NOT DISTRACTIBLE.

3. *Intensity*, the energy level of a response whether positive or negative. The available scores are INTENSE, MEDIUM and MILD.

4. *Regularity, rhythmicity*, or degree of regularity of functions such as eating, elimination, and the cycle of sleeping and wakefulness. The trait refers to the predictability of biological functions like appetite and sleep. Does the child get hungry or tired at predictable times? The available scores are REGULAR, MEDIUM and IRREGULAR.

5. *Sensitivity, sensory threshold or threshold of responsiveness* is related to how sensitive the child is to physical stimuli. It is the amount of stimulation (sound, taste, touch, temperature change) needed to produce a response in the child. The available scores are LOW, MEDIUM and HIGH.

6. *Initial reaction, approach or withdrawal*, the child's characteristic response to a new object, person or environment, in terms of whether the child accepts the new experience or withdraws from it. Does the child eagerly approach new situations or people? Or does the child seem hesitant and resistant when faced with new situations, people or things? The available scores are POSITIVE, MEDIUM and NEGATIVE.

7. *Adaptability* of behavior to changes in the environment is related to how easily the child adapts to transitions and changes, like switching to a new activity. Does the child have difficulty with changes in routines, or with transitions from one activity to another? The available scores are ADAPTABLE, MEDIUM and NOT ADAPTABLE.

8. *Persistence*, the span of the child's attention and his persistence in an activity. This is the length of time a child continues in activities in the face of obstacles. Does the child continue to work on a puzzle when he has difficulty with it or does he just move on to another activity? Is the child able to wait to have his needs met? Does the child react strongly when interrupted in an activity? The available scores are LONG, MEDIUM and SHORT.

9. *Mood*, the child's general mood or "disposition", whether cheerful or given to crying, pleasant or cranky, friendly or unfriendly. This is the tendency to react to the world primarily in a positive or negative way. All children have a variety of emotions and reactions, such as cheerful and stormy, happy and unhappy. Yet each child biologically tends to have a generally positive or negative outlook. A baby who frequently smiles and coos could be considered a cheerful baby, whereas a baby who frequently cries or fusses might be considered a stormy baby. The available scores are POSITIVE, MEDIUM and NEGATIVE.

IV. Temperamental traits in mobile robots

The most of temperamental traits may be implemented as the mobile robot behavioral patterns visible to an observer. Some traits, e.g., persistence, may be perceived only while the robot executes a task, while others, e.g., sensitivity, are perceived better when there's no ongoing task. A simple demo task, line following, will be used in this paper to explain the general principles of temperamental traits manifestations during a task execution. We will use the term "idle state" to refer to the fact that no explicit short-term task is assigned to the robot. While the most of the mobile robots do nothing in the idle state, those having implemented some kind of artificial curiosity (14) (15) (16) use an idle state for self-improvement, e.g., learning by exploring their physical and virtual environments. Some traits, especially distractibility, sensitivity and initial reaction, are manifested in the presence of an external stimulus. To keep the traits easily observable and differentiated, we will use the sound patterns of various rhythm and base pitch as a sample external stimulus, having in mind that they don't interfere with the main task. The level of robot curiosity caused by such a stimulus is related to the novelty of the sound pattern, and the novelty can be calculated by comparison with known sound patterns.

We now go through the nine traits and discuss in details how they could be perceived.

1. *Activity level* of a robot is manifested during a task execution and during idle state. During the execution of the demo task, a robot with a high activity level observes the line position and orientation rarely, corrects its orientation using sharp turns and then goes straight in calculated direction at a high speed. A robot with low activity level follows the line at a low speed, observing the line position

and orientation often and adjusting its orientation with soft turns. PID controller tuning parameters may be numerical variables that fit this trait.

During the idle state, a curious robot with a high activity level explores the surroundings by changing its position and orientation fast. A robot with low activity level prefers the use small (e.g., servo) motors to change the positions of its sensors or changes its position and orientation at a low speed.

Activity level is strongly related with Laban's "flow" effort parameter and its adaptation in (3). Flow parameter may take two values: BOUND and FREE. BOUND refers to a robot that moves through the movements more carefully to execute the succession of the motion precisely and corresponds to LOW activity level. FREE describes a robot that moves without caring about the precision (uncontrolled movements) and corresponds to HIGH activity level.

2. *Distractibility* in robots, in its simplest form, is a binary feature related to the robot behavior when it receives the command to execute a new task during execution of some other task. If a distractible robot were in that situation, it would switch to the new task immediately, while a non-distractible one would complete the ongoing task first.

The others, more human-like manifestations of distractibility are related to curiosity and novelty of perceived stimuli. A tightly programmed robot would never interrupt the execution of an ongoing task just for curiosity sake, but one that is allowed to create tradeoffs between immediate tasks execution and learning may do so. The tendency to be curious and leave the duty for new experiences and stimuli in robots could be useful in exploring new ways to perform a task. On the other side, it could bring the robot away from its goal. This feature might be annoying for a user, but also interesting to involve her/him in a helpful relationship of mutual care, which may strengthen the rapport between the robot and the user, possibly an aged person or an autistic child. If a robot is allowed to be curious even during task execution, the measure of distractibility is inversely proportional to the novelty of the stimulus that causes delay of ongoing task execution.

During the idle state and exploration of a perceived stimulus, a curious, distractible robot would stop a current exploration and switch to a new one easily, while a non-distractible robot would stay focused on the on-going exploration even if another, more novel stimulus is perceived. The measure of this kind of distractibility is inversely proportional both to the novelty of the stimulus that causes attention shift and to the difference of novelty between the stimulus that causes attention shift and the currently explored stimulus.

Distractibility trait is related to the Laban's *space* parameter and its adaptation to robots in (3). While a robot that takes a “single-focused approach to the environment”, moves towards the next immediate goal with little or no distractions, the one with “multi-focused approach to the environment” meanders and wanders more being distracted by external stimuli.

3. *Intensity* is an attribute of the reaction, so this trait becomes observable as the energy level of a response when the robot is exposed to a stimulus. Besides being treated separately, this trait may be regarded as a property of other traits. Following some researchers (3) that put intensity in relation with *effort* in the Laban's modeling of human actions (11), we consider that the acceleration of a motion caused by a stimulus is a good way to express intensity. For instance, the base variable for approval/withdrawal is the direction of movement, but the acceleration of approval/withdrawal is related to the intensity trait. Variations of intensity are perceivable by the user only if the robot can exploit a sufficiently wide range of speed and acceleration.

4. *Rhythmicity* with children refers to the predictability of biological functions. This trait is hardly visible to a short-term observer. The best map to robots is predictability of recharging intervals. Assuming that robots have battery level sensors, and that the critical battery level is L, a robot with regular rhythmicity signals “battery low” and enters a “power safe” mode at regular time intervals, as long as the current battery level l is within a predefined range, e.g. $L<l<L+30\%$. A robot with irregular rhythmicity alerts battery low condition and enters the power safe mode strictly when a predefined level is reached, e.g., $l=L+20\%$, no matter when the last charging was occurred. The measure of rhythmicity is inversely proportional to the standard deviation of time spans between recharging requests.

5. *Sensitivity* (threshold of responsiveness) is related to the intensity of the stimulus to trigger a response, e.g. sound level or the speed of an approaching body. The threshold rules the sensitivity level. A robot might have different sensitivity on different sensors, as it happens in animals and people.

6. *Approval/Withdrawal* (initial reaction) is the response to a new stimulus/object/person. A robot manifests this trait by approaching to the source of stimulus to interact and explore, or by getting away of it. The base variable that describes this trait is the direction of movement. The robot with neutral attitude keeps the same distance to the stimulus source. According to the Laban’s effort factor named *space*, movement may be direct or indirect. The robot with strong positive/negative initial reaction approaches/withdraws straight along the line that connects the robot and the stimulus source, while the one with moderate positive/negative initial reaction approaches/withdraws using indirect path, weaving around that line, or even spiraling up/down to the target.

7. *Adaptability* concerns how much robot adapts to changes in the environment. An adaptable robot should be able to explore and find new solutions to achieve its goal despite non-matched expectations; for instance, if an unexpected obstacle is on its path, it should be able to re-plan to find its way to the goal. This might seem a desirable property, but it is often seen a source of potential problems by device (e.g., washing machines) producers that may see adaptability as a way to lose control of the product. So, although this trait might seem always desirable, it is interesting to explore its variability in robotic application that have to cope with real world.

Being not adaptable may have some advantages too. Non adaptable robots are faithful, stabile, predictable and respond only to its master’s commands. It is less likely that non adaptable robot will rush into dangerous situations. On the other hand, the adaptable robot has its drawbacks like unpredictability, volatility and inconstancy.

Adaptability refers to how easily or quickly the robot can adjust to changes in its environment after the "initial reaction" is manifested. We doubt that "adaptability" and "initial reaction" are not completely independent traits and that the most of the robots that are slow-to-adapt (adaptability trait value: NOT ADAPTABLE) are also more likely to withdraw when first confronted with new object/person (initial reaction: NEGATIVE, withdrawal).

8. *Persistence* is the amount of time a robot continues in activities in the face of obstacles. It can be represented by the amount of time dedicated to an activity, before spontaneously leaving.

9. *Quality of mood* is quite rich and articulated. There are different scales, related to the mood ("cheerful or given to crying, pleasant or cranky, friendly or unfriendly") for each of them a different base variable could be defined, e.g., the pitch and rhythm of sound emitted by the robot, the quality of movement (smooth or jerky, it can be obtained by modifying the PID parameters for the motors), distance and direction of movement for friendly or unfriendly. If we only consider the mood as POSITIVE, NEUTRAL or NEGATIVE, as in the reference model, this is related to the *valence* of an emotion, a parameter used in the often adopted Russell's circumplex model (17). If we look at that framework, on the negative side are emotions like anger, fear, sadness and disgust, on the positive side are happiness, calmness. The differences among them are in the other Russell's parameter (*arousal*), which is not considered in our reference model. As a first approximation we might take regular movements as expressing the POSITIVE mood and irregular movements (trembling, going back and forth and wandering) as expressing NEGATIVE mood.

From the above presented excursus on temperament traits, we have now a set of variables which can be implemented to characterize behaviors of the robot. For each of nine traits there are 1-2 variables. Each variable have 3 base values. This means that we have the possibility to represent a huge variety of composite temperaments.

V. From Linguistic Variables to Numeric Physical Properties

Each personality trait is represented by a linguistic variable (13). While variables in mathematics usually take numerical values, linguistics variables take natural language expressions. For example, trait/variable "Sensitivity" takes expressions "Low" and "High". A more granular term-set for each trait may be created by using linguistic modifiers (so called "hedges", e.g., "very", "somewhat", "quite", "more or less", "a bit"…) applied to the original natural language expressions. For example, the term-set for the trait *Sensitivity* may be {"Very low" , "Low", "Somewhat low", "Moderate", "Somewhat high", "High" , "Very high"}.

Besides being represented by a linguistic variable, if a trait is related to a measurable physical property, it may be expressed by a numerical, the so-called "base variable". For example, the sound sensitivity may be expressed as the sound level needed to produce a response in the robot, expressed in decibels. On the one hand, the robot's software has to deal with numerical sensory input data and to output numeric data to the actuators' controllers. On the other, it has to understand the personality settings expressed by the values of a linguistic variable and to express its findings about someone's personality traits in a natural way. So, a mapping between linguistic and numeric values that represent a trait (in other words: between the term-set and the range of possible base variable values) is required.

The most traits may be expressed both by the numerical variable (e.g. Sound Sensitivity=30dB) and by the value from the term-set (e.g. "Low"). A linguistic value of a measurable linguistic variable may be modeled as a fuzzy set that provides a linguistic interpretation for values of the corresponding base variable. The membership function μ(x) assigns to each value of the numeric base variable a degree of membership to the fuzzy set labeled by a linguistic value.

Therefore, we can build a model that can interpret data in terms of values for the traits, and relates them with possible values to define the outputs.

For instance, a robot distractible by sounds, where "distraction" is defined as a fuzzy set on the base variable "sound level" (see Figure 1), might decide to stop following the rule:

IF SoundIntensity is Distraction THEN Speed is Zero

Various distractibility traits may be implemented using slightly different membership function of *Distraction*, representing the robot's sensitivity for a given range of sound intensity. This kind of modeling gives the possibility to define a continuous input/output space to describe temperament and the respective expression, as it happens with people.

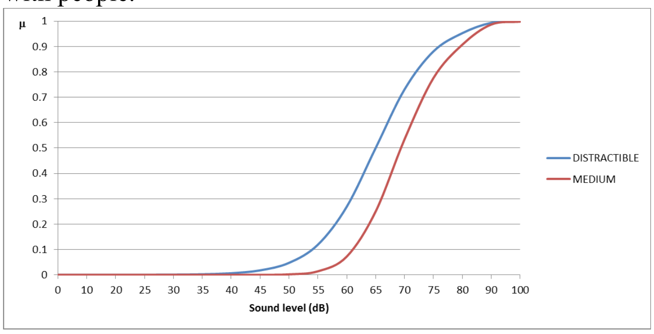

Fig. 1: Distraction membership function for Distractibility trait values "DISTRACTIBLE" and "MEDIUM"

Moreover, composition of trait effects can also take place. For instance, a robot with positive attitude to novelty (POSITIVE initial reaction), might consider the variation of sound intensity (sound is just the physical quantity we are considering in this example, but there is one of them for each available sensor) as modeled by a fuzzy set. If there was a quiet environment and the sound is strong, then SoundIntensityVariation would be HIGH, so introducing a high level of novelty. If the robot has a positive attitude can follow the rule:

IF SoundIntensityVariation is HIGH
THEN GoToSoundDirection AND Speed is HIGH

Here, *GoToSoundDirection* is an action that brings the robot to the direction of the sound. The rules related to distraction and initial reaction interfere, and the robot goes at a medium speed to the sound direction. The intensity trait modifies the membership functions used to interpret data.

VI. Conclusion and future work

We have presented a model to implement temperamental traits on robots by relating sensorial data and actuators activity to the traits of the Thomas, Chess, and Birch model, through fuzzy sets and rules. The model has to be tested in a real environment.

We plan to implement on a robot the proposed temperamental traits as behavior modifiers and to test them in a real environment (e.g., a tennis school) with different sets of traits. The first aim is to check whether people perceive implemented traits as expected. Second aim is to identify the best trait for the specific robot to be accepted in its environment.

Then we will create a robot that can perceive temperamental traits of the previous one by sensing its execution of a known task or by sensing its reaction in a specific situation like approaching.

The final goal will be a robot that can perceive temperamental traits of people by sensing their execution of a known task or by sensing their reaction in a situation like approaching.

VII. References

1. **Engler, Barbara.** *Personality theories (8th ed. ed.).* Boston : Houghton Mifflin, 2008. ISBN 9780547148342.

2. **Jung, C. G.** *Psychological Types (The Collected Works of C. G. Jung, Vol. 6).* s.l. : Princeton University Press , 1976 (1921).

3. *Communicating affect via flight path Exploring use of the Laban Effort System for designing affective locomotion paths.* **Sharma, M., et al., et al.** s.l. : IEEE, 2013. Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction. pp. 293-300.

4. **Elliott, Clark.** *The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System (PhD Thesis).* s.l. : The Institute for the Learning Sciences Technical Report #32, 1992.

5. *ALMA - A Layered Model of Affect.* **Gebhard, Patrick.** Utrecht : s.n., 2005. Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems.

6. *The Artificial Emotion Engine Driving Emotional Behavior.* **Wilson, I.** s.l. : AAAI Spring Symposium, 2000.

7. **Kessler, Henrik, et al., et al.** SIMPLEX – Simulation of Personal Emotion– A new Model of Artificial Emotions. *Affective Computing.* s.l. : InTech, 2008.

8. **Rodić, A. and Addi, K.** Mathematical Modeling of Human Affective Behavior Aimed to Design Robot EI-Controller. *New Trends in Medical and Service Robots.* s.l. : Springer International Publishing Switzerland, 2014.

9. *Studying people's emotional responses to robot's movements in a small scene.* **F., Julian M. Angel and Bonarini, Andrea.** Edinburgh : IEEE, 2014. 23rd IEEE International Symposium on Robot and Human Interactive Communication. ISBN: 978-1-4799-6763-6.

10. **Thomas, F. and O., Johnston.** *The illusion of life: Disney animation.* s.l. : Disney editions, 1984.

11. **Laban, Rudolf and Ullmann, Lisa.** *Laban, Rudolf and Ullmann, Lisa. Modern educational dance. Praeger New York.* New York : Praeger, 1968.

12. **Thomas, Alexander, Chess, Stella and Birch, Herbert G.** *Temperament and Behavior Disorders in Children.* New York : New York University Press, 1968.

13. *The Origin of Personality.* **Thomas, Alexander, Chess, Stella and Birch, Herbert G.** 1970, Scientific American, pp. 102-109.

14. *Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010).* **Schmidhuber, Jürgen.** NO. 3, s.l. : IEEE, 2010, IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, Vol. VOL. 2.

15. *Intrinsic Motivation Systems for Autonomous Mental Development.* **Pierre-Yves Oudeyer, Frédéric Kaplan, and Verena V. Hafner.** 2, s.l. : IEEE, 2007, IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, Vol. 11.

16. *A Comparative Study of Value Systems for Self-Motivated Exploration and Learning by Robots.* **Merrick, K.E.** 2, s.l. : IEEE, 2010, IEEE Transactions on Autonomous Mental Development, Vol. 2, pp. 119 - 131.

17. *A circumplex model of affect.* **Russell, J. A.** 39, 1980, Journal of Personality and Social Psychology, pp. 1161-1178.