Studies in Computational Intelligence 1130

Alexei V. Samsonovich
Tingting Liu *Editors*

Biologically Inspired Cognitive Architectures 2023

Proceedings of the 14th Annual Meeting of the BICA Society

MOREMEDIA

Springer

Studies in Computational Intelligence **1130**

Series Editor

Janusz Kacprzyk, *Polish Academy of Sciences, Warsaw, Poland*

Alexei V. Samsonovich · Tingting Liu
Editors

Biologically Inspired Cognitive Architectures 2023

Proceedings of the 14th Annual Meeting of the BICA Society

Springer

Editors
Alexei V. Samsonovich
Department of Cybernetics
Moscow Engineering Physics Institute
Moscow, Russia

Tingting Liu
College of Science and Technology
Ningbo University
Ningbo, China

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-031-50380-1 ISBN 978-3-031-50381-8 (eBook)
https://doi.org/10.1007/978-3-031-50381-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Crowdsourcing-Based Approbation of Communicative Behaviour Elements on the F-2 Robot: Perception Peculiarities According to Respondents

Liliya Volkova[2,3](✉), Artemy Kotov[1,3,4], and Andrey Ignatev[2]

[1] National Research Center «Kurchatov Institute», pl. Akademika Kurchatova 1, 123182 Moscow, Russia
kotov_aa@nrcki.ru

[2] Bauman Moscow State Technical University, 2-ya Baumanskaya ul. 5-1, 105005 Moscow, Russia
liliya@bmstu.ru

[3] Russian State University for the Humanities, Miusskaya pl. 6, 125993 Moscow, Russia

[4] Moscow State Linguistic University, ul. Ostozhenka 38 str.1, 119034 Moscow, Russia

Abstract. This article is dedicated to F-2 the companion robot and to interpretations of respondents' estimations of designed communicative multimodal behaviour. The affective robot is described: it represents a platform for implementing and verifying various individual behavioural traits for robots. F-2 interprets multimodal input: text, face orientation and tactile signals; it translates the input into facts, which are seeds for further affective behaviour. Facts trigger behavioural patterns for reacting—concurrent scenarios with their activation degrees varying over time. The most activated scenario is implemented via one reaction out of a pool of corresponding scenario reactions. Each reaction is multimodal and includes one or several components: speech, gestures, gazes. Robot behaviour is estimated by human assessors during conducted experiments on communication. Several notable effects were observed during perception of implemented communicative behaviour of F-2. These effects are discussed, they are presumed to be evidence to common human expectations transfer from human-human interaction to human-robot interaction.

Keywords: Human-robot interaction · Companion robots · Affective interfaces · Multimodal communication · Human-machine interaction

1 Introduction

In various languages "to respond mechanically" means to answer without involvement, without affection or taking the interlocutor's feelings into account. This is a marker of technologic insufficiency of most existing dialogue systems, as well as the fact that most people avoid chat-bots in online services, preferring communication and interaction with a living person. What prevents people from using most question answering

A. V. Samsonovich and T. Liu (Eds.): BICA 2023, SCI 1130, pp. 932–945, 2024.
https://doi.org/10.1007/978-3-031-50381-8_101

machines is not only their limitedness of scenarios and question types, but importantly a lack of personal touch and affection. People tend to treat robots as social actors, to subconsciously attribute personality traits to them [1] and to transfer psychological aspects of communication from human-human interaction onto human-robot interaction (HRI). In general, people are not only or not always interested in information, they crave for (affective) communication *per se*; this has been recently proved during worldwide lockdowns. It is in general use and of importance to show caring attitude in order to satisfy the addressee's needs in conversation. Applied psychology is highly concerned with enciphering and deciphering additional information from multimodal behavioural patterns, all in intention to increase the communication efficiency and to decrease the psychological discomfort brought up with inappropriate means of expressing one's attitude, of delivering the intended sense in cues. In HRI relative problems can be observed. Question-answering systems along with assistants in smart phones and houses are a global trend nowadays, and regarding the aforesaid, adding some affection into their mechanistic thinking is a problem to solve.

One particular field of dialogue systems is development of affective companions. Given a body, virtual or physical, such a companion would have an ability of nonverbal communication. On one hand, nonverbal communicative behaviour isn't the main channel for communication and is less controlled by the speaker as compared to speech. On the other hand, it is necessary for personal communication, as its elements can highlight the speaker's emotional state or one's attitude towards the topic, as well as it can give additional information about pragmatics of the message [2]. The nonverbal communicative behaviour can express up to 60% of the information [3] and gives a personal touch. The core of an affective reaction is associated with a communicative function (i.e. an intention to express some pragmatics, e.g. hesitation, negation, joy etc., referred to a particular stimulus by the addressee). Pragmatics of particular multimodal expressions can be investigated basing on a corpus of recorded multimodal emotional reactions, annotated with implementations of communicative functions. As to choice, corpora recorded with actors show pretty clear portraits of emotions; still, these are not fully applicable within a task of simulating communicative behaviour where communicative functions and their expressions can be blended. A more promising approach is investigating natural dialogue recordings, such as TV shows records in EmoTV [4], cinema in Multimedia Russian Corpus [5], oral exams and happy people recordings in Russian Emotional Corpus [2]. Affective reactions can be adapted onto companion robots and 3D avatars of dialogue systems.

Modern companion robots interact with people in different modes: via speech, gestures and gazes. Although the tactile channel is rarely the preferred mode of interaction for social robots, it becomes important within the two major areas. First, children tend to establish tactile contact with attractive robots, for they are used to tactile interaction with their toys. People contacting with robots at home or at a private space also tend to establish tactile communication with appropriate robots; that observation has helped to create a growing segment of tangible companion robots [6, 7]. Tactile feedback can be also of a great importance, if robots are used for therapy of people with health deficits or people passing medical treatment at a hospital [8–11]. Secondly, those who have just encountered a robot for the first time, try to evaluate its adequacy by checking some basic cognitive responses: people wave with their hand in order to attract the robot's attention and to check its gaze response, they say hello or ask a simple question (e.g. "What is your name?") to check its speech competence, and in many cases they slightly touch the robot to get its behavioural response to that. So, the social touch, although not frequent in modern social interaction space, becomes an important feature for companion robots. In this work we design a tactile input processing subsystem for the companion robot F-2, we extend its basic responsive features with those aimed at tactile stimuli, and organize the crowdsourcing procedure to check the perception of behavioural cues. Beyond this, we discuss a number of observations given by our respondents: the more the variety of opinions and of perception peculiarities is accounted, the more aspects of perception of robot's behaviour are discovered. F-2's behavioural diversity in social interaction is often source for hypotheses and sometimes for a more narrow experimental investigation of psychological aspects, transferred from human-human interaction onto human-robot interaction.

2 Overview

> *That's the way the emotional intellect affects the robot's behavior. The same event can affect us in different ways, depending on our mood or on circumstances.*
>
> *"Eva" by Sergi Belbel, Cristina Clemente, Martí Roca, Aintza Serra*

AI classics gave birth to a dream of a different kind of robots, exceeding the limitation of pure automation of heavy and sometimes dangerous labor. Alan Turing test has been finally passed (on the simulated ground of interlocutor's youth), the Loebner Prize has been awarded multiple times, which proves that dreams of humanity made their first steps in real life. Dialogue systems started resembling people in their mechanical minds and in the way they think (SIR showed simple logic for inference and solving straightforward logical problems [12]) and conduct conversation: ELIZA [13]—via rule-based reformulation of input, A.L.I.C.E. [14]—via heuristic matching of an input phrase to samples in its knowledge base, PARRY [15]—via dialogue strategies. SHRDLU [16] modeled commands understanding and performing, particularly it could state a lacuna and ask what would the missing term mean. CYC [17, 18] performed human-like reasoning and adapted to novel situations.

Affective robots are the next step on the stairway of robotic companions: we're looking forward to *a future sort of friendship, of a robot-human kind*, according to Cynthia Breazeal. Simulating surface emotional phenomena [19] is perceived in HRI as pleasant, it contributes to establishing closer contact and affection towards robots. The first robotic toy which shortly became object of attachment was Tamagotchi, and the corresponding effect of affection was named after it. As to toys, one particular problem is their limitedness of reactions: people get used to that and lose their interest. Hence, high variability of reactions scenarios is of big importance.

Several breakthroughs were made in the classic field of dialogue systems, inspired by AI pioneers. These are affective robots and agents based on text semantics processing and simulating surface emotional phenomena: IBug [20], SEMAINE [21], Greta [22], Kismet [19], Max [23, 24]. The latter is interesting for changing its emotional state over time:

it represents emotions via an activation axis and a valency axis. Each emotion is a point in this space, e.g. two 'happiness' points may differ: one having a high activation level, the other one having a low one—a passive sort of happiness. These projects model more natural reactions with behavioural elements which enable simulating surface emotional phenomena (referred to as affective behaviour), and often infer displaying emotional mimics or gestures adapted from human ones.

In communication, speech and gestures often cooccur and coexpress the speakers' message as a composite signal [25–27]. Hence, multimodal communication should combine elements of behaviour targeted at several layers of perception. The BML (behavior markup language, sort of a lingua franca for behaviourists in robotics [28]) format is handy: it is flexible and allows multiple layers in a behaviour frame, in particular speech, gestures, gazes, mimics.

Implementation of personality for social robots is a prospective HRI field. First, there is such an opportunity due to dialogue systems worldwide spread via smartphones. Second, while robots (and this general tendency covers AI as well, beyond fundamental science) are considered useful tools rather than social function carriers [29], the latter function is getting more and more actual for social robotics. It is stated that "personality is a key element for creating socially interactive robots", and that "studies on this dimension will facilitate enhanced human–robot interaction" [30]. This is due to providing users with better affordance, which makes it intuitive and natural for the users to understand the robot's behaviours [31]. Personality "represents those characteristics of the person that account for consistent patterns of feelings, thinking, and behaving" [32]. Particular traits of personality can distinguish one robot from others and impress people more, which leads again to personal touch [33].

3 F-2: The Platform for Implementing Affective Behaviour Elements and Personal Traits

3.1 A Communicative Agent Inside a Robot

—Max, what's your emotional level?

—Standard eighth.

—We're not accustomed to that much emotional robots. Turn it down to 5.

"Eva" by Sergi Belbel, Cristina Clemente, Martí Roca, Aintza Serra

F-2 first was developed as an agent comprehending text and expressing its artificial attitude to the input. Classic text analysis stages are executed. Morphological analysis operates with a morphological dictionary of 100,000 lemmas, resulting in a stack with annotated tokens. Next, syntax and semantics are extracted to form a dependency tree. The stack head is reduced when possible with any of the existing 850 syntactic rules: each rule can reduce a list of tokens in its right-hand side to the left-hand side head (e.g. to form a predicate group out of a predicate and its object). Several trees can exist simultaneously (in case different rules can be applied), unsuccessful trees are rejected. In the resulting tree, nodes are annotated with semantic roles (valencies) and with semantic features (semantic markers). Facts are to be extracted afterwards according to a set of templates, e.g., a subtree with following semantic valencies in nodes: {***agens***, ***predicate***, ***patiens***, ***instrumentative***} (subject, verb, object, instrument). It is similar to triads [34, 35] and frames [36], but we store more than one feature per valency in most cases, see example on Fig. 1.

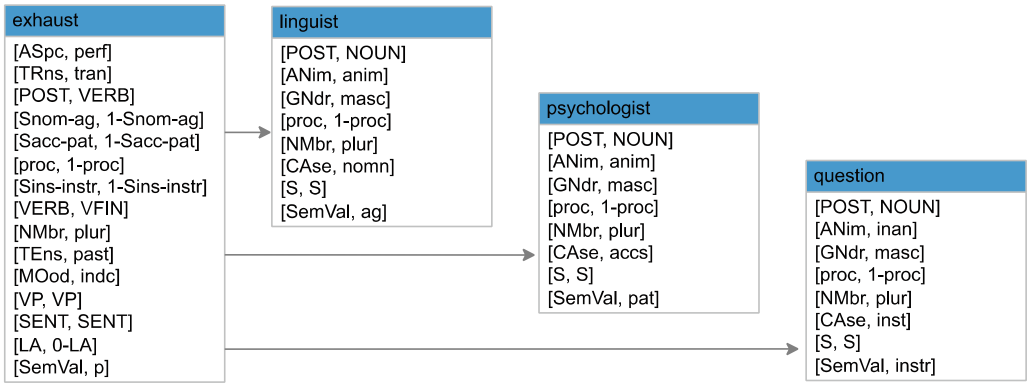

Fig. 1. A sample syntactic dependency subtree corresponding to a semantic fact template.

Facts are source for forming and expressing an attitude. We adopted the approach of concurrent states, the dominating one taking control in order to produce a reaction to the input: each input fact is matched to communicative goals (intentions to implement, standing for communicative strategies, e.g. to express happiness, to show negation, to draw attention, to hesitate etc.). The most activated goal implies selection of a multimodal response out of a pool of possible reactions corresponding to this goal. While the selected reaction is produced, all of the activated goals remain for further processing. Once the current reaction is complete, the corresponding goal is deactivated, and the next topmost goal is selected from those having non-null activation degree. There is a background goal implementing the standby mode via a small constant activation degree. In order to look awake when unoccupied, F-2 slightly moves all the time instead of freezing as if turned off: in between stimuli it looks at its hands or watches the ceiling, it moves its gaze as if thoughtful. All of the implemented elements of behaviour are selected and translated into behavioural elements for F-2 from our REC corpus (Russian Emotional Corpus) [2], and then approbated.

F-2 is taught to receive multimodal input as facts: from text (we use our own syntactic and semantic parser along with Yandex SpeechKit [37] for speech-to-text translation and for further speech synthesis) and from video (human face orientation is perceived [38]). The reaction is formed as a BML frame, its components are played by different modules of our robot control subsystem [38], performing speech synthesis, gesticulating, gazes. F-2 has 6 Dynamixel AX-12 motors (2 per each hand and 2 for its neck, enabling F-2 to nod and to turn the head left and right). An LCD monitor stands for the face. The robot is connected to a PC.

Research based on REC showed that dominant channels of nonverbal communication are hands (20% reactions), head (37%) and mimics (~ 30%), while body is used relatively rarely: the average is 6,6% cases per communicative function, except for 3 functions—absence/impossibility (35%), separation (97%), inspiration (34%) [2].

Hence, body movements are not obligatory for affective companions, as to Russian (different cultures have different gesticulation peculiarities and usage frequencies). This is the reason why F-2 gesticulates only with its hands and head.

3.2 Perception of Tactile Signals

People tend to develop excessive expectations of social robots due to generalization from behaviour and mental models of human kind [39]. The team and our respondents thought of F-2 only as of a talking head with mimics and hands, until one particular series of experiments. When making acquaintance with F-2, youngsters happened to involve the tactile channel of communication, trying to scratch or to tap the robot—and got upset for their touches were ignored. Thus an expectation gap [39] was discovered: our respondents expected F-2 to perceive touch as well, as pets do.

In order to fill this perception gap we developed a module perceiving tactile signals. Basing on an overview conducted in [40], we selected flexible resistive tactile sensors of appropriate size to place in palms, on cheeks and on top of head (2 on every cheek, 3 on top)—round sensors Ø 1.2 cm from the Interlink FSR 400 series [41]. We also placed 2 square sensors of the same series on F-2's belly. Since these sensors don't have an interface for connecting to a computer nor to a controller, they were connected using a third-party programmable controller, via an Arduino UNO board.

A software module was created for analysis of tactile signals: every sensor states the force of pressure. A training set was recorded; the overall data amount is 2.5 k recordings. The decision tree classifier showed the topmost precision, overwhelming logistic regression, Bayesian and support vector machine. The developed module classifies tactile input with precision from 75% to 100% for 5 touch types: tap (or simple touch), smoothing over, scratching, handshake, hit. Smoothing over and scratching are registered for head, cheeks and belly zones. Hit is registered for head and cheeks. Tap is allowed for all of the zones, including hands. Following four target communicative goals are implemented: to focus on the touched zone, to show discomfort, to express pleasure, to express sadness and pain. These communicative goals are part of F-2's set of goals; they cover five touch types and four touch zones.

4 Results

4.1 Evaluation of Communicative Strategies Implementation

Since we bridge the affective divide that parts human interlocutors from most AIs, human estimations of F-2's affective reactions are what matters in the first place. The most important and the main criteria for estimating designed behaviour in whole and in part are assessing and interpreting impressions of human interlocutors. We conduct experiments with assessors conversating with F-2 (or two F-2 robots which show slightly different behaviour). Assessors are asked to interact with F-2 in particular settings and to evaluate the robot's reactions in terms of "human-like", "realistic", "nice". Beyond the approbation, priceless remarks sometimes give us food for thought. We aggregate feedback on different stages of F-2's evolution. This is crowdsourcing: all of the respondents are volunteers invited to interact with the robot.

In one series of experiments we check one particular function of F-2's, reducing its functionality—of maintaining dialogue, recognizing people, reacting to gazes—in order to evaluate the function subject to research. Particular results on emotional expressions evaluation can be found in [42], F-2's gaze behaviour is analyzed in [38], its use of politeness strategies—in [43]: F-2 helps learning words in Latin, it gives a word in Russian and expects the answer in cycle. When one tries to guess the word and fails, this is an emotional situation, a threat to one's social face [44]. We essay at working with this embarrassment: in its multimodal response (a) F-2 points at the error ("Incorrect!"), (b) it ignores it, giving a hint ("it's similar to…") or (c) uses a politeness formula (a hedge: "no, *a bit* incorrect"), mitigating the face loss. In the reverse setting, F-2 loses its social face and uses multimodal hedges to compensate this: hedges contribute not only to expressing politeness, but to expressing emotional and cognitive states as well—the robot is perceived as nervous and hesitating [43].

4.2 Assessing Perception of Reactions on Tactile Input

Feedback from respondents is the most important factor for developing the robot. 24 respondents were invited to evaluate F-2's reactions to tactile input. The feedback was systematized and interpreted; results are given in Table 1. Overall expressed attitude is positive. Respondents express positive opinions to the new functionality of F-2.

Table 1. Feedback from respondents: attitude grouped by touch type.

Touch type	Positive	Negative	Neutral	Number of responses
Smoothing over	90%	1%	9%	68
Tap	53%	28%	19%	78
Hit	82%	11%	7%	44
Scratching	68%	6%	26%	31
Handshake	52.5%	27.5%	20%	40
Overall	71%	13%	16%	249

During experiments, our respondents introduced 2 new touch types: tickling the belly (associated with scratching) and holding or shaking both hands (handshake touch group). In Table 1 these are included in associated basic groups, in Table 2 both are stated separately and marked with asterisks, thus basic types being presented clearer.

In comparison of handshake in Table 1 and handshake subdivided in Table 2, basic handshake positive estimations changed from 52.5% cases to 60%. This difference is due to neutral or negative perception of F-2's reactions to an unforeseen touch type "hold hands", which was classified as one of predefined touch types. Produced reactions didn't meet respondents' expectations—an expectation gap again. On the other hand, separating "tickling the belly" from basic scratch makes it evident that this new type seems to fit existing patterns of reacting. Obviously, it is prospective to highlight new touch types and to develop corresponding reactions patterns.

Table 2. Feedback, attitude by touch type: basic and 2 new types introduced by respondents.

Touch type	Positive	Negative	Neutral	Number of responses
Smoothing over, tap, hit	72.6%	14.7%	12.6%	190
Scratching, basic	75%	8%	17%	24
Tickling the belly *	43%	0%	57%	7
Handshake, basic	60%	29%	11%	35
Hold both hands *	0%	20%	80%	5
Overall, basic types	73.5%	16.5%	10%	237
Overall, new types *	25%	8%	67%	12

Figure 2 shows an example: F-2 expresses pleasure (with lifting its head, closing its eyes and saying “cool” with slight raising hands) as reaction to smoothing over.

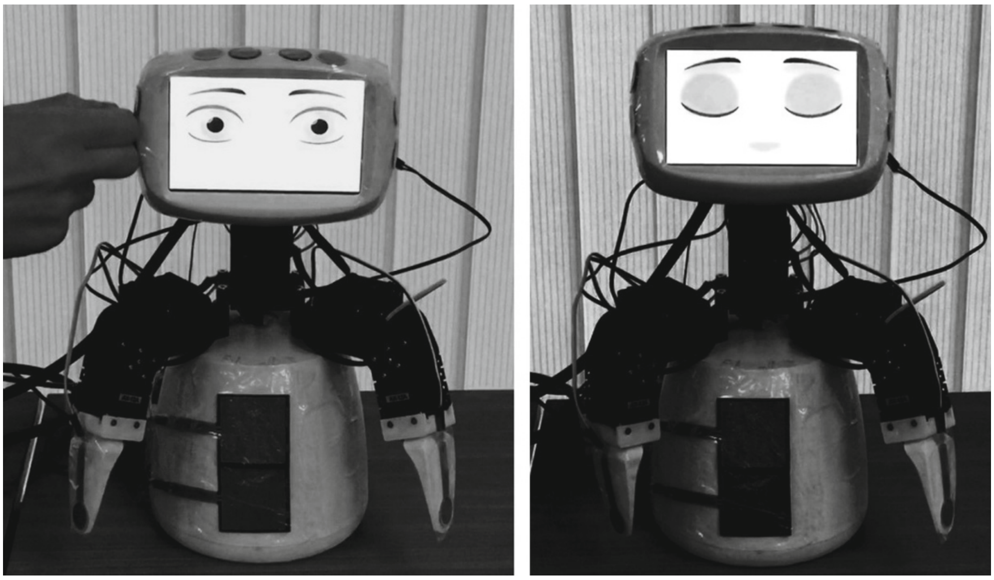

Fig. 2. F-2 modification equipped with touch sensors reacts to smoothing over its cheek.

5 Discussion

5.1 Perception of Positive Reactions in Case of Low Variability of Reactions

—Is there anything I can help you?

—Yes. Turn your emotional level back to 8.

“Eva” by Sergi Belbel, Cristina Clemente, Martí Roca, Aintza Serra.

During approbation of the tactile input detecting subsystem, our experiment wasn’t aimed at detailed conversation, thus respondents weren’t asked to give cues for further discussion with F-2. So the robot didn't have the opportunity to be talkative: it was mostly touched and sometimes hit (no robot was harmed). F-2's reactions spectrum was limited by the context to reactions associated with the predefined communicative goals enlisted in Sect. 3.2. In this setting, we collected comments from assessors who came in small groups and actually saw the robot communicating with other respondents for a while before or after taking one's turn. Their prolongated observations resulted in a few particular surprising comments, which can be summed as follows: the robot mostly shows positive reactions, it is "too happy", which was estimated by two people as "unnatural, for life isn't mostly positive".

Keeping in mind that the reactions observed were limited by target goals within the experiment under discussion, the F-2 team concluded the following. First, assessors tend to lose interest in case of prevailing homogeneous reactions in context of less informative communication (as compared to the context of applying core F-2 functions of conducting conversation and commenting on cues with a touch of personal touch [42, 45]). As this less talkative setting is possible during exhibitions when people first meet F-2 and gather first impressions (contrary to sharing thoughts during conversation), in such a case the robot's behaviour could be balanced with at least two more communicative strategies: to make acquaintance (a face recognition module along with a long-term faces memory is in progress) and to present oneself to unfamiliar people who showed interest via tactile actions. Second, a goal of expressing the need in distancing (which was previously considered as possible, but not primary for implementation) is prospective, perhaps with a shade of irony. E.g. the "tap" touch on the head or on the belly might result in a cue "I'm not THAT social" with shaking its head, or in a cue "Stay away from wild robots!", which is a not rude variation of "keep your hands off me" (the latter harsh one is absolutely not to be implemented).

Our respondents were asked to punch the robot in different zones, but most of them stated they didn't want to or weren't willing to. Those who tried punching can be subdivided into two groups: (1) parameterizing the punch force in order to find the threshold separating punch from simple touch, and (2) writing afterwards that they had feared to hurt the robot. The 2nd subgroup outnumbered cold-minded naturalists (and partly intersected it). A possible research direction arises: as we developed a communicative robot intended to be a pal, a friend, it is of interest to estimate general expectations of assessors in such sort of HRI (the approach described in [39]).

5.2 Righty or Lefty?

Assessors sometimes point out particular observed behaviour elements and draw their conclusions, which surprise us. During approbation of F-2's reactions to tactile stimuli, one assessor concluded the robot was right-handed: most reactions were shown with the right hand. Waving or pointing at something are performed by the right hand in the gestures database, which seems to correspond to the following consideration: people most frequently perform gestures with the leading hand. On the other hand, it is only one of possible points of view, depicting only part of the whole. Gestures of the second hand can be associated with the emotional intelligence, with enhancing metaphor explanation [46]. Particular investigations show that semantics partially determines hand choice for gesture production. According to [47], spatial aspects of a message determine the choice of the right or left hand for gesturing (e.g., use of left hand to gesturally depict an object moving in the relative left position). Speakers tend to use their dominant hand to represent messages with positive connotations in political debates [48]; an assumption was made that emotional valence (positive–negative) matched to the way right- and left-handers represent valence (e.g., the dominant side, either left or right, is positive) may determine hand choice for gesturing [49].

In future research, a robot can be specified as a lefty or as a righty, or balanced towards ambidexterity. Non-dominant hand usage for expressing rather emotional gestures (as contrary to strictly informative) is prospective as well, and it is subject to investigate in context of approbation with invited respondents and psychologists.

5.3 Gender Perception

According to [50], many of the gender-related perceptions and expectations formed in human-human interactions may be inadvertently and unreasonably transferred to interactions with social robots. Human perception is affected by gender-related expectations when judging both humans and robots with minimal gender markers, such as voice or even a name (which is the case of F-2: it recently got a higher voice as compared to the previous, both synthesized). We didn't conduct intentional experiments on gender perception yet, but our respondents already provided us with first results.

An experiment was carried out with two F-2 robots in order to compare perception of gazes-based communicative behaviour. When human face orientation changed, the video analysis software module detected one's gaze direction. If it was towards one of robots, the corresponding robot performed its reaction. In our simulation the left robot turned towards the interlocutor and looked straight at one's face, while the right one looked away after eye contact. The interesting fact is, the person with the highest emotional intellect estimation among present respondents (according to a test we conducted; emotional intellect is briefly defined as the ability to perceive, understand, and manage emotions [51–54]) assumed the left robot to be a boy, and the right one to be a girl. The motivation is as follows: the robot which doesn't look away demonstrates traits of "courageous and masculine behaviour", compared to the "shyness" of the right robot, associated by the interlocutor with rather feminine behaviour. Such perception peculiarities depend on a particular person's experience, for the cited traits may be associated not only with gender, but also with context. In any case, this difference might be one of important traits for developing different personalities for robots.

During the experiment dedicated to robot's reactions on tactile signals, our respondents either didn't state F-2's gender at all, or asked if it was a girl. This happened after we re-designed the robot's eyes: they were surrounded with gradient grey color for a more expressive look. That new feature turned out to look like cosmetics when viewed from the side, and was interpreted as a feminine feature by most respondents among that third part of those who paid attention to the gender aspect.

5.4 Effect of Enlightenment of Mutual Understanding

Last but not least, every human being starves for being accepted and understood, and people tend to transfer their communicative expectations from human-human to human-robot interactions [50]. F-2 was created as such a character showing its understanding and replying affectively (as taught to simulate it): it doesn't perform commands, but it can hold conversation via expressing its attitude to human interlocutor cues. We conduct a series of experiments on storytelling: F-2 comments on phrases, expressing its artificial attitude, and/or roves its gaze, which is often interpreted as thinking on the story [38], as a perceptual ability (though not necessarily staring at the interlocutor, as in [55]). During one experiment, the robot gave its cues as usual, this time including "Hmm" and "I understand" repeatedly. One respondent gave a feedback of awe, pointing at being understood. Thus, one can see the evidence to success of the adopted approach of imitation modelling [56], which is transferring human behaviour elements associated with emotions in human-human interaction onto the robot, and simulating affective reactions driven by communicative goals. Successfully selected and adopted communicative behavioural strategies meet human expectations in HRI, which is proved experimentally, according to feedback given by respondents.

6 Conclusion

F-2 is designed as a communicative robot, which comments on interlocutor's cues and multimodal behaviour and expresses its opinion with artificial affect. In this article we presented the introduction of a tactile input channel along with approbation results. Human-machine communication was set up with respondents who gave their opinions on what they experienced. We discuss selected observations grounded on human estimations of implemented behavioural strategies of F-2, as this crowdsourced feedback is the most important quality metric for this sort of HRI projects. Effects of transfer of expectations and perception from human-human interaction to human-machine interaction are highlighted. Presence of these effects enables us to conclude that F-2 is a successful project as a robot-companion: not only its behaviour is perceived as human-like and estimated mostly positively, but our respondents also treat it like a personality (while we fill their expectation gaps when discovered).

Due to the F-2 architecture, more behavioural strategies can be implemented along with personality traits. In our architecture, input facts trigger several communicative goals, which is quite universal, as this mechanism enables researchers to balance existing strategies via varying activation functions decrease principle: slowly fading activation for relatively important events, and no fading for background processes (e.g. the stand-by mode). Furthermore, tempers can be simulated on a robot basing on implementing different sets of activation functions, e.g. following an approach stated in [57]: 4 basic tempers could be modeled via activation and deceleration. With varying goals importance via initial activation values and fading rules (i.e. the activation function form), different tempers can be tuned. These are considered as prospective research directions along with conducting comparative analysis of perception of various culture-related, gender-related, social role-related behavioural patterns.

Acknowledgements. The reported study was supported by the grant of the Russian Science Foundation № 19–18-00547, https://rscf.ru/project/19-18-00547/.

The F-2 team wishes to express gratitude to all of our respondents, including students of the Power Engineering department of BMSTU. All of the crowdsourced feedback is precious for us as source of knowledge and as seeds of future research.

References

1. De Graaf, M. M.A., Ben Allouch, S.: Expectation setting and personality attribution in HRI. In: HRI, ACM/IEEE International Conference on Human-Robot Interaction, pp. 144–145. IEEE, Piscataway (2014)
2. Kotov, A.A., Zinina, A.A.: Functional analysis of non-verbal communicative behavior (in Russian). In: Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue", 14(21) vol. 1, pp. 308–320. RSUH, Moscow (2015)
3. Engleberg, I.N., Wynn, D.R.: Working in Groups: Communication Principles and Strategies, My Communication Kit Series, p. 133. Allyn & Bacon, Boston (2006)
4. Martin, J.-C., Devillers, L.: A multimodal corpus approach for the study of spontaneous emotions. In: Affective Information Processing, pp 267–291. Springer-Verlag, Heidelberg (2009)
5. Grishina, E.A.: Gestures and grammatical features of speech act. In: Multimodal Communication: Theoretic and Empiric Investigations (in Russian), pp. 25–47. Buki Vedi, Moscow (2014)
6. Gansohr, C., Emmerich, K., Masuch, M.: Hold me tight: a tangible interface for mediating closeness to overcome physical separation. In: Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST, vol. 178, pp. 74–85. Springer, Heidelberg (2017)
7. Paiva, A., Chaves, R., Piedade, M., Bullock, A., Andersson, G., Höök, K.: SenToy: a tangible interface to control the emotions of a synthetic character. In: Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 1088–1089. ACM, New York (2003)
8. Shibata, T.: Therapeutic seal robot as biofeedback medical device: qualitative and quantitative evaluations of robot therapy in dementia care. Proc. IEEE **100**, 2527–2538 (2012)
9. Shibata, T., Kawaguchi, Y., Wada, K.: Investigation on people living with seal robot at home. Int. J. Soc. Robot. **4**, 53–63 (2012)
10. Inoue, K., Wada, K., Shibata, T.: Exploring the applicability of the robotic seal PARO to support caring for older persons with dementia within the home context. Palliat. Care Soc. Pract. **15** (2021). https://doi.org/10.1177/26323524211030285
11. Takayanagi, K., Kirita, T., Shibata, T.: Comparison of verbal and emotional responses of elderly people with mild/moderate dementia and those with severe dementia in responses to seal robot, PARO. Front. Aging Neurosci. **6**(SEP), 257 (2014)
12. Raphael, B.: SIR, a computer program for semantic information retrieval. In: Minsky, M. (ed.) Semantic Information Processing, pp. 33–144. M.I.T. Press, Cambridge (1968)
13. Weizenbaum, J.: ELIZA. Commun. ACM **9**, 36–45 (1966)
14. Wallace, R.S.: The Anatomy of A.L.I.C.E. In: Epstein, R., Roberts, G., Beber, G. (eds.) Parsing the Turing test, pp. 181–210. Springer Science+Business Media, London (2009)
15. Colby, C.M.: Artificial Paranoia: A Computer Simulation of Paranoid Processes. Elsevier Science Inc., New York (1975)
16. Winograd, T.: Understanding Natural Language. Academic Press, New York (1972)

17. Lenat, D.B., Borning, A., McDonald, D., Taylor, C., Weyer, S.: Knoesphere: building expert systems with encyclopedic knowledge. In: Proceedings of the Eight International Journal of Conference on Artificial Intelligence, IJCAI'83, vol. 1, pp. 167–169. ACM, New York (1983)
18. Lenat, D., Prakash, M., Shepherd, M.: CYC: using common sense knowledge to overcome brittleness and knowledge acquisition [sic] bottlenecks. AI Mag. **6**(4), 65–85 (1986)
19. Breazeal, C.: Designing Sociable Robots. MIT Press, Cambridge (2002)
20. i bug. http://ibug.doc.ic.ac.uk/. Last Accessed 18 May 2023
21. Shröder, M.: The SEMAINE API: towards a standards-based framework for building emotion-oriented systems. Adv. Hum. Comput. Interact. **2010**, 319406 (2010)
22. Pelachaud, C.: Greta: a conversing socio-emotional agent. In: Proceedings of the 1st ACM Sigchi International Workshop on Investigating Social Interactions with Artificial Agents, pp. 9–10. ACM, New York (2017)
23. Max. http://cycling74.com/products/max/. Last Accessed 18 May 2023
24. Becker, C., Kopp, S., Wachsmuth, I.: Simulating the emotion dynamics of a multimodal conversational agent. In: ADS 2004, LNAI, vol. 3068, pp. 154–165. Springer, Heidelberg (2004)
25. Engle, R.A.: Not channels but composite signals: speech, gesture, diagrams and object demonstrations are integrated in multimodal explanations. In: Gernsbacher, M.A., Derry, S.J. (eds.) Proceedings of the Twentieth Annual Conference of the Cognitive Science Society, pp. 321–326. Psychology Press, London (1998)
26. Kelly, S.D., Ozyürek, A., Maris, E.: Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. Psychol. Sci. **21**, 260–267 (2010)
27. Kendon, A.: Gesture: Visible Action as Utterance. Cambridge University Press, Cambridge (2004)
28. Vilhjálmsson, H., et al.: The behavior markup language: recent developments and challenges. In: Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pelé, D. (eds.) IVA 2007, LNCS, vol. 4722, pp. 99–111. Springer, Heidelberg (2007)
29. Severinson-Eklundh, K., Green, A., Huttenrauch, H.: Social and collaborative aspects of interaction with a service robot. Robot. Auton. Syst. **42**, 223–234 (2003)
30. Miwa, H., Takanishi, A., Takanobu, H.: Experimental study on robot personality for humanoid head robot. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1183–1188. IEEE, Piscataway (2001)
31. Hara, F., Kobayashi, H.: Use of face robot for human-computer communication. In: IEEE International Conference on Systems, Man and Cybernetics, Oct. 1995, pp. 1515–1520. IEEE, Piscataway (1995)
32. Pervin, L.A., John, O.P.: Personality Theory and Research. Wiley, New York (1997)
33. John, O.P., Srivastava, S.: The Big-Five trait taxonomy: history, measurement, and theoretical perspectives. In: Pervin, L.A., John, O.P. (eds.) Handbook of Personality: Theory and Research (2nd edn.), pp. 102–138. Guilford Press, New York (1999)
34. Simmons, R.F.: Storage and retrieval of aspects of meaning in directed graph structures. Commun. ACM **9**, 211–214 (1966)
35. Simmons, R.F., Burger, J.F., Long, R.E.: An approach towards answering English questions from text. In: Proceedings on fall Joint Computer Conference, pp. 357–363. Spartan, New York (1966)
36. Minsky, M.: A framework for representing knowledge. In: Winston, P.H. (ed.) The Psychology of Computer Vision. McGraw-Hill, New York (1975)
37. Yandex SpeechKit. https://cloud.yandex.com/en/services/speechkit/. Last Accessed 18 May 2023
38. Velichkovsky, B.M., Kotov, A., Arinkin, N., Zaidelman, L., Zinina, A., Kivva, K.: From social gaze to indirect speech constructions: how to induce the impression that your companion robot is a conscious creature. Appl. Sci. **11**(21), 10255 (2021)

39. Kwon, M., Jung, M., Knepper, R.: Human expectations of social robots. In: 11th ACM/IEEE International Conference on Human-Robot Interaction, pp. 463–464. IEEE, Piscataway (2016)
40. Ignatev, A., Volkova, L.: On equipping F-2 the affective robot with tactile sensors. In: Uvaysov S.U., Ivanov I.A. (eds.) Information Innovative Technologies: Materials of the International Scientific–Practical Conference, pp. 123–127. Association of Graduates and Employees of AFEA Named After Prof. Zhukovsky, Moscow (2021)
41. InterLink Electronics. FSR 400 Series. https://www.interlinkelectronics.com/fsr-400-series. Last Accessed 18 May 2023
42. Zinina, A., Zaidelman, L., Kotov, A., Arinkin, N.: The perception of robot's emotional gestures and speech by children solving a spatial puzzle. In: Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue", vol. 19 (26), pp. 811–826. RSUH, Moscow (2020)
43. Malkina, M., Zinina, A., Arinkin, N., Kotov, A.: Multimodal hedges for companion robots: a politeness strategy or an emotional expression? In: Selegey, V.P., et al. (eds.) Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue", Issue 22, pp. 319–326. RSUH, Moscow (2023)
44. Brown, P., Levinson, S.C.: Politeness: Some Universals in Language Usage. Cambridge University Press, Cambridge (1987)
45. Kotov, A.: D-script model for synthesis and analysis of emotional speech. In: SPIIRAS (ed.) Proceedings on SPECOM'2004: 9th Conference "Speech and Computer", Saint-Petersburg, Russia, September 20–22, 2004, pp. 579–585. ISCA Archive (2004). http://www.isca-speech.org/archive/specom_04. Last Accessed 18 May 2023
46. Argyriou, P., Mohr, C., Kita, S.: Hand matters: left-hand gestures enhance metaphor explanation. J. Exp. Psychol. Learn. Mem. Cogn. **43**(6), 874–886 (2017)
47. Lausberg, H., Kita, S.: The content of the message influences the hand choice in co-speech gestures and in gesturing without speaking. Brain Lang. **86**, 57–69 (2003)
48. Casasanto, D., Jasmin, K.: Good and bad in the hands of politicians: spontaneous gestures during positive and negative speech. PLoS ONE **5**(7), e11805 (2003)
49. Park, W., Choi, W., Jo, H., Lee, G., Kim, J.: Analysis of control characteristics between dominant and non-dominant hands by transient responses of circular tracking movements in 3D virtual reality space. Sensors **20**(12), 3477 (2020)
50. Chita-Tegmark, M., Lohani, M., Scheutz, M.: Gender effects in perceptions of robots and humans with varying emotional intelligence. In: 14th ACM/IEEE International Conference on Human Robot Interaction, pp. 230–238. IEEE, Piscataway (2019)
51. Salovey, P., Mayer, J.D.: Emotional intelligence. Imagin. Cogn. Pers. **9**(3), 185–211 (1990)
52. Bar-On, R.: The emotional Intelligence Inventory (EQ-i): Technical Manual. Multi-Health Systems, Toronto (1997)
53. Bar-On, R.: The Bar-On Model of Emotional-Social Intelligence. Psicothema **18**(Suppl.), 13–25 (2006)
54. Ciarrochi, J., Forgas, J.P., Mayer, J.D.: Emotional Intelligence in Everyday Life: A Scientific Inquiry. Psychology Press, Philadelphia (2001)
55. Iwasaki, M., et al.: That robot stared back at me! Demonstrating perceptual ability is key to successful human-robot interactions. Front. Robot. AI **6**, 85 (2019)
56. Rudakov, I.V., Paschenkova, A.V.: A hierarchical method for verification of software algorithms via hierarchical Petri nets. Eng. J. Sci. Innov. **2**(14) (in Russian). BMSTU Press, Moscow (2013)
57. Karpov, V.E., Karpova, I.P., Kulinich, A.A.: Social Communities of Robots. URSS, Moscow (2019). (in Russian)