Cooperative Intelligence – A Humane Perspective

Bernhard Sendhoff
Honda Research Institute Europe
Offenbach am Main, Germany
bs@honda-ri.de

Heiko Wersing
Honda Research Institute Europe
Offenbach am Main, Germany
heiko.wersing@honda-ri.de

***Abstract*—In this contribution, we outline our concept of cooperation between humans and intelligent systems which we denote as cooperative intelligence. We argue from a human perspective and emphasize the advantages of keeping the human in the loop rather than targeting autonomous systems. Our focus is respecting human values such as retaining competences, sharing experiences, and self-esteem. We discuss process-oriented requirements for intuitive cooperation like joint goals and shared intentions and social dimensions like empathy, relations, and trust. Finally, we suggest that cooperative intelligence can be facilitated by integrating interaction episodes across multiple system embodiments and instances, achieving the best holistic service with regard to personal preferences and needs.**

***Index Terms*—cooperation, human-machine systems, trust**

I. Introduction

The field of artificial intelligence encompassing machine learning, intelligent systems, computational intelligence, neural computation and many other related research subjects has seen a remarkable growth in research activity, industrial engagement and public perception in the last 20 years. The widespread popular interest has led to some misconceptions or overly enthusiastic predictions of the performance of AI systems in the short-term future, however, an important discussion was triggered on how we as a society envision the role of AI systems. Ethical and societal implications and security and privacy are more relevant for already operational or soon available systems than the discussion on super-human intelligence and how it might endanger our species at large. Implications on the future organization of work [1] and privacy issues [2] are imminent challenges.

Almost ten years ago, deep neural architectures combined with huge data sets from the internet reached previously unseen recognition performance in computer vision as well as speech processing. This tremendous success of new machine learning methods applied to hard computational problems led to the aforementioned explosive growth of interest in AI systems well beyond the research community.

Autonomous driving is a good example for the sudden rise of AI technologies (even though first larger research programs were already carried out in the 80s) as well as today's insight that the arrival of fully autonomous vehicles was well overestimated by many. It is also a good example for the initial focus on autonomous systems that can replace humans (drivers, sales representatives, physicians).

978-1-7281-5871-6/20/$31.00 ©2020 IEEE

This focus on autonomy has changed in the last years. Autonomy has been augmented and in some cases even been replaced by humans and machines collaborating together to achieve optimal performance. Different terminologies have been used, sometimes synonymously, sometimes with a slightly different connotation. Research into man-machine collaboration has become one of the seven major AI research strategies outlined in the US AI R&D Strategic Plan [3], [4]. Leading universities have set up dedicated and strategic research programs and centers like the *Human Centered AI* Institute at Stanford University. In the domain of autonomous driving, the idea of cooperative driving[1] and cooperative mobility has been raised and promoted [5]–[8].

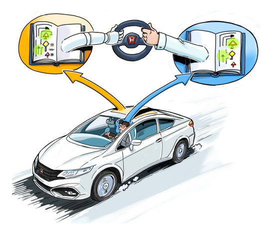

Fig. 1. Cooperation between driver and car requires both a shared understanding of the traffic situation and a correct prediction of the mutual intention of driver and car on how to negotiate it. Like in driver-driver cooperation correct prediction might require some form of communication (in [7] the authors extend the idea of visual gestures between drivers to a tactile communication channel between car and driver).

In this paper, we outline our human-centered concept of cooperation between humans and intelligent systems which we call *cooperative intelligence*. In Section 2, we first motivate why looking *beyond autonomy* is valuable for a future hybrid society composed of humans and AI systems. We then discuss in Section 3 requirements and conceptual approaches for achieving successful cooperation between humans and machines. Section 4 analyzes the role of maintaining a consistent user-centered approach to cooperative sustained interaction with multiple and variable system embodiments. We summarize our conclusions in Section 5.

II. Cooperative Intelligence - Concepts

A first definition of *cooperative intelligence* can be formulated as the capability of systems to work closely together with humans in a variety of ways and with different emphasis

[1] Here we primarily refer to the cooperation between car and driver and not to the cooperation between different vehicles.

towards solving a complex task in a demanding environment. However, this definition is almost entirely task driven. It fits well to the prevailing view of cooperation in the fields of human-machine cooperation or in collaborative robotics. We propose to widen the definition to include the evolutionary perspective on "benefiting from each other" as well as the societal view on "harmoniously living together". Therefore, *cooperative intelligence* is not just a capability with which something can be achieved but it is a property (a state) of a system that is a pre-requisite for defining a relation. Primarily, a relation between human and machine but also between machines, which becomes more important when machines acquire more freedom in decision making based on individual needs and wants that reflect individual user histories.

A. Definitions

The general definition of the term cooperation in the Oxford dictionary is "work or act together in order to bring about a result", emphasizing the importance of a goal in the cooperation process. Piaget [9] described the necessity of mutual adaptation for coordinated actions. Hoc [10] emphasized the necessity of managing interferences between the cooperation partners in real-time. Bratman [11] postulates the requirement of mutual support in his definition of *shared cooperative activity*, introducing a social dimension into the setting. Moll et al. [12] have emphasized the role of cooperation as the driving force of human intelligence development, rooting their work in the classic theoretic works of Vigotsky [13].

As early as 1988, Parker and Pin [14] used the term *cooperative intelligence* to refer to man-machine interaction. In Parker's *symbiont*, the mutual benefit of human-machine interaction has been a central element. However, the synergy that could be achieved by combining the advantages of humans and machines in a smart way has received less attention in the research community compared to cooperation between intelligent agents [15] or supervisory control in automation [16], which primarily aims at the guidance of the machine.

For cooperation in multi-agent systems Fryer and McKee [17] also used the term *cooperative intelligence*. Typically, the "intelligent" capabilities of each individual agent were limited and the expectation was that through the interaction collective intelligence could be an emergent property. In particular, in the field of heuristic optimization for global search this is often referred to as swarm intelligence [18], [19].

B. A Humane Perspective

Let us state in the following our main reasons why we consider cooperative systems as advantageous over fully autonomous systems from a human-centered point of view:

Inadequacy. The task to be solved is too difficult for current autonomous systems and the human is required for supervision (e.g. current driver assistant systems) or must contribute certain aspects of the required behavior that are difficult or impossible for current artificial systems [20]. Consequently, the organization of the interaction between systems and humans is usually rather basic and often limited to switching between the system and the human instead of carrying out a task together [21]. Here, cooperation can be a more flexible strategy to cope with the shortcomings of current systems. Krüger et al. [22] argue that increased autonomy should be accompanied by more adaptivity to the human, to ensure flexibility and alignment with human needs and requests. In Figure 2, we visualize the role of cooperation and competition between human and machine for achieving an optimal system performance according to the human needs.

Learning. The system (or the human, although practically it is mostly the system) shall advance its skills through the interaction with the human. Numerous ways have been suggested in the literature on how systems can learn from humans [23], [24]. Humans can give feedback on task performance through reinforcement learning and interactive optimization or provide ground truth in supervised labeling of environment and objects. In more elaborate learning environments the human can support the system by providing hints or strategies how a certain task can be achieved. One example is robot motion learning where the human can manipulate e.g. a compliant arm to show how a certain movement should be executed [25]. Systems with increased capability may monitor humans, e.g. in a manufacturing environment for the purpose of quality control, for optimizing operational procedures or for better ergonomics [26]. In such a situation, the human could learn from system suggestions [27].

Retaining. In general, humans have to exercise their cognitive and physical skills in order to maintain and keep them. If systems take over human tasks that have been previously mastered, humans tend to lose (or never acquire) those skills over time. Simple examples are navigation systems/map reading, hand-writing/typing[2] or simple arithmetic. Lost human skills are a problem if the replacing system is unavailable or they are a necessary component of other, more complex needed human skills. It is known, that physical skills have a positive effect on cognitive skills and that one type of cognitive skills (e.g. solving brainteasers) positively influences unrelated skills like taking care of daily routines. A "good memory" is a very general skill, that can be trained (or neglected) depending on tasks carried out (or "outsourced") to systems. In particular, for elderly people it is important to continue to execute daily tasks in their environment in order to stay "mentally fit".

Sharing. Humans are social beings and like to share experiences among each other. Therefore, even in cases where artificial systems might be able to accomplish a task without a human, for the human it might be more desirable to work as a team. Again, driving can serve as a good example. According to a study from Continental [28] most people enjoy driving most of the time. Therefore, even if completely autonomous driving systems might be available in the future, many drivers might not want to be driven fully autonomously all of the time. Sharing means *virtually* experiencing together, which we see

[2]In particular, in Asia people have increasing difficulty in mastering the correct stroke sequence for writing Chinese characters due to computer use.

as an important element for establishing a relationship. We will elaborate this aspect of cooperative intelligence in Section IV.

Self-Esteem. All humans want to feel competent and recognized by others for their achievement and capability. Several studies have shown that unemployment leads not just to financial shortcomings and to problems in participation but also to psychological and physiological disease patterns [29]. Humans need to advance their skills and their capabilities and contribute them to society. At the same time, as Dollard and Windefield [29] argue problems can equally arise through "overemployment". In general, the missing separation between work and free time (work-life balance) poses a problem for many people. Cooperative intelligence in artificial systems can support people at both ends of the work-life balance spectrum. Instead of replacing underskilled people by autonomous AI systems, cooperative intelligence can enable the less skilled workforce to provide a valuable contribution to the economy and to society. It is known that assistance systems for elderly people should not reduce their self-reliance but provide support that maintains and strengthens their cognitive responsibility. In this context, cooperation means ensuring safety but not necessarily reducing the task load or cognitive load.

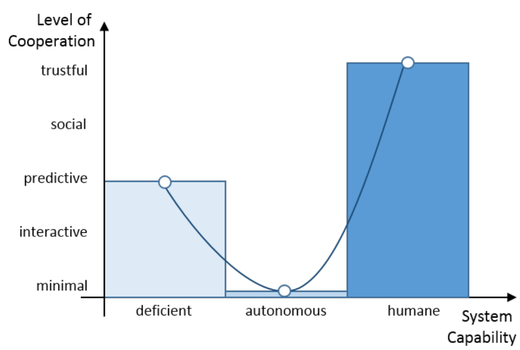

Fig. 2. Cooperation on the level of interaction and prediction is mostly sufficient to increase system performance with the support of the human. Autonomous systems only require minimal cooperation, e.g. by formulating objectives. Cooperation for a trustful and social relation is for the benefit of the human, i.e. the system is capable of behaving humanely, not in the sense of similarity but empathy.

Sharing and self-esteem are important psychological needs of people, as suggested in Maslow's hierarchy of needs [30]. Sharing is related to connecting to people and to social belonging. Relatedness, belonging or social affiliation have been identified across many different theories in social psychology as an important need for humans [31]. It is an interesting question whether the need of social belonging can be extended to the relation between humans and intelligent systems. The desire for competence and the need to advance our own skills (maybe even uniquely human) constitute the top two levels of Maslow's pyramid. Therefore, in order to address psychological needs of humans, future AI systems should not operate autonomously but in relation with humans.

In [32] Gambetta writes "basic forms of cooperation are inevitable if a society is to be at all viable". Therefore, if we envision a hybrid society where humans and intelligent systems "live" together, we need to expand the idea of cooperation from the relation between humans to those between humans and intelligent systems.

III. Cooperative Intelligence - Approaches

After discussing the main motivations for going beyond autonomous intelligent systems, we will outline relevant concepts and examples for the successful application of this approach.

A. *Joint Goals, Shared Intentions, and HMI*

A key challenge of human-machine cooperation is the establishment of a common ground of shared perceptions, attention, and intentions [33], thus allowing the negotation and installment of joint goals [11]. The process of cooperation must be continuously revised according to dynamic change of situation and context [10]. Gienger et al. [34] have realized a bimanual robot system for joint object handling (cooperative turning) that uses haptic feedback to negotiate in real-time a successful stable grasp and turn sequence between human and robot. Vinanzi et al. [35] have investigated estimating human intentions from body pose sequences in a joint block building game to support human robot cooperation. Bühler et al. [36] provide an approach of estimating the human belief about the task and environment state in a joint task sequence for human and robot. Their method allows an online prediction of next human actions to enhance cooperation. Mataric [37] has emphasized that the approach of socially assistive robotics can deliver a new quality of assistance by taking into account human social intentions and goals.

The human machine interface (HMI) plays a decisive role in facilitating communication between cooperative interaction partners. Norman [38] has emphasized the requirement for consistent, intuitive, and expressive interfaces that allow the human interaction partner to predict and understand intelligent machine or robot actions. In order to use the natural sensorimotor body representations of humans for this, Krüger et al. [39] have argued that the successful integration of car and driver for cooperative driving should provide interfaces that make the car an extension to the drivers body.

B. *Empathy, Relationships, and Trust*

Cooperation is strongly related to trust, a concept that is discussed differently in different disciplines and that is just at the beginning of being addressed in the context of intelligent systems. The necessity for building trust in AI and for building trustworthy AI has received significant attention with the increased capability and use of intelligent systems. Most publications focus on properties that AI systems should have in order to be trustworthy [40], [41]. Probably, the most well known is the Asilomar list of AI Principles [42] triggered by the 2017 Asilomar conference. Consequently, the criteria of transparency, explainability and interpretability have dominated the discussion. Interestingly, this somewhat contradicts the discussion of trust in sociology, where trust is related to limited predictability thus to a limitation of transparency. In general, the discussion of trust in AI is uni-directional whereas in sociology trust is always seen as a bidirectional concept of mutual interaction [32]. So, paradoxically a pre-requisite for cooperation can be the freedom to choose not to cooperate.

There is a profound difference between trust in the soundness of an automated technological device and the trust in the decisions and behavior of an intelligent system with certain degrees of freedom in its choices. Trust in technology can be attributed to the trust in the organizations that design and manufacture the technology. Consequently, trusting a pocket calculator or a computer (even though colloquially we might use such terms) does not make sense, because even though modern computers are highly complex machines, they do not have any behavioral freedom. In the discussion of trustworthy AI, the separation between trust in the developer and trust in the system itself is not always clear, however, it is important because only the latter is unique and will ultimately lead to a new kind of trust in a hybrid society.

In an evolutionary approach, it is more natural to regard trust as a result of cooperation rather than forming a pre-requisite. Indeed in evolutionary game theory, evolutionary stable strategies including cooperation emerge as a result of natural selection without the need of a concept of trust [43]. Either way instead of just focusing on properties it is important to consider what are constraints and conditions for trust and cooperation to emerge. Coercion or self-interest can lead to cooperation, however, not to trust. In particular, the recognition of a person that the cooperation partner only behaves in a trustful way because of self-interest is likely to lead to distrust and less cooperation. Coercion is not seen in sociology as a stable basis for replacing trust in cooperation, however, it is deeply rooted in the relation we anticipate with intelligent systems. Indeed, Asimov's famous three laws of robotics[3] from the 1950s are an example of coercion and of uni-directional trust. Up to now it is a philosophical discussion how much behavioral freedom is necessary for general artificial intelligence to emerge, however, trust through technological coercion would result in trust in technology instead of trust in AI in the above definition. In [45] the authors distinguish between "understanding trust" and "role-based trust" and they propose concrete guidelines on how trustworthy AI systems can be developed in a medical context in the even more complex triangular relationship between doctor-patient-system or in general between human-human-system. Falcone and Sapienza have developed a theoretical approach [46] for adjusting the level of autonomy to the appropriate trust level of an individual user of an IoT system.

One of the re-occurring pre-requisites for establishing trust (in the terminology of [45]) is a long-term relation of knowing each other and of goal and value alignment (item 10 in [42].) Indeed in sociology long-term arrangements are in general seen as a means for stabilizing cooperative behavior [32]. Gambetta writes: *"[..] here is a sense in which trust may be a by-product, typically of familiarity and friendship, both of which imply that those involved have some knowledge of each other and some respect for each other's welfare [..] "*, however, also emphasizing the differences between those notions: *" [..] trust, although a potential spin-off of familiarity, friendship, and moral values, must not be confused with them, for it has quite different properties [..]"*.

Familiarity, although conceptually different, can be one path towards trust in particular when combined with empathy. Familiarity leads to the sharing of goals and values. Here, sharing can be seen as an early stage of alignment. Complete alignment of goals and values does not seem necessary for developing familiarity, indeed, understanding and respecting each others differences seems to be sufficient. Familiarity can be a completely rational relation, however, empathy includes the understanding of each others emotions and non-rational motives. Where Gambetta [32] talks of familiarity and friendship, in the context of intelligent systems we would rephrase it as familiarity and empathy. Developing familiarity has two major aspects: closeness and length of interaction. Both are deeply related to the system's aspects of cooperative intelligence in particular in the context of the pervasiveness and ubiquitousness of cyber-physical systems, which will be discussed in the next section on *Distributed Intelligence*.

IV. Distributed Intelligence

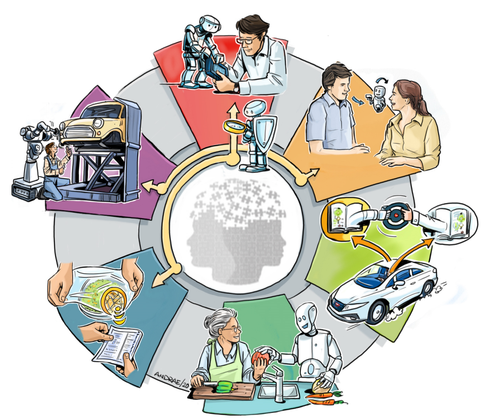

Fig. 3. An example of an intelligent cyber-physical system consisting of multiple embodiments and a shared knowledge representation that allows co-operative intelligence to emerge through sustained heterogeneous interaction.

A. Sustained Heterogeneous Interaction

We postulate that two ingredients are necessary to realize cooperative intelligence: sustained interaction (over a longer period of time) and heterogeneous interaction, i.e. the system experiences a variety of interactions and the human interacts with a variety of system instantiations sharing the same knowledge base. Therefore, intelligent cyber-physical systems will be an enabling technology for realizing cooperative intelligence for AI systems that are able to relate to humans

[3]**First Law:** A robot may not injure a human being or, through inaction, allow a human being to come to harm. **Second Law:** A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. **Third Law:** A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws [44].

with adequate levels of empathy and trust. We have argued in Section III about the relevance of prediction and shared intentions in cooperation. Predictive models require for their instantiation prior knowledge or data for learning that can be assembled from previous interactions [47]. Knowledge can be shared by either the cooperation partner directly through communication or by a different entity, e.g. another AI system or another instantiation of the same AI system.

Communication is a very rich means to share knowledge about intentions and general preferences, however, it also has some drawbacks. First, our preferences may seem unclear in relation to the capabilities of an AI system. Secondly, communication needs cognitive focus, which may divert attention from the task that the human/AI team wants to solve.

Interactions are heterogeneous because AI system instantiations are typically tuned towards tasks (driving, manufacturing, assistance). On the contrary, humans are tremendously versatile at the expense at not being well tuned to single tasks. Concentrating on single tasks limits interaction duration, narrows the scope of knowledge about the user, and thus renders sustained interaction and establishing a relation to the user difficult. For example, even a perfectly intelligent car will not be able to fully understand the physical capabilities of the human driver in a manipulation scenario and is limited to the actual driving time. Basically, every AI system instantiation can supply one puzzle piece of knowledge about the human. By combining several of those pieces at a sufficiently abstract level, cooperative intelligence will be an "emergent" systems property for building a sustainable relation to the human.

B. Multiple and Variable Embodiment

In the last section, we have argued why cyber-physical systems in a personal context (as opposed to an industrial context like in industry 4.0) will be an enabling technology for cooperative intelligence. However, we need to go a step further than the standard concept of cyber-physical systems, which are usually regarded as networks of embedded systems with some standardization for information exchange. In our context, we require a more centralist view, i.e., one system with multiple different embodiments or with a variable embodiment.

In Figure 3, different embodiment instances are sketched that are connected to a central knowledge representation. Some instances like a work space robot or an intelligent car are typical representatives of intelligent systems, others like intelligent food wrapping foil, smart paper or small humanoid assistants are less common. Each instance can contribute pieces of information: the work robot observes physical human state and preferred interaction modes, the intelligent foil samples nutritional parameters, the intelligent paper measures reading speed and intellectual interest. From the combination of the instances, a sustained interaction model can be obtained that allows a holistic understanding of the human for higher levels of cooperative intelligence. Information can also be exchanged, e.g. to provide nutritional suggestions for better physical robot interaction. However, note that a centralist view does not necessitate a centralist realization where a master instance receives and transmits information from and to different embodiment instances. Instead, in a distributed architecture all information from one system instances may "diffuse" to all other embodiments. Such a system will be more robust (and safer in a cyber security sense), however, the communication overhead can be severe.

Mühlig et al. [48] demonstrate a first example of a multiple and variable embodiment system, where two robots in different office locations and a smart office environment exchange information about the position of colleagues to allow coordinated search and guidance for requested colleagues. Whereas in our previous examples the interaction between different system instances happened on a larger time scale, the multiple embodiment system in [48]–[50] is a good example for one system that uses the different capabilities of its instances at once to solve a cooperation task.

C. Shared Representations

Cooperation requires the capability between human and system to communicate about processes, things, and abstract concepts like ideas or strategies that do not yet exist. Therefore, apart from the pursuit of trust between system and human, also practically a holistic understanding of the human (and of the system) is necessary to a certain degree to achieve cooperation. The aggregation of knowledge has to be sufficiently abstract to be useful for other instances. This requires going beyond classical constrained rule-based models and inference engines towards an open (holistic) approach including information from interactions between human, system instance and environment. The challenge is to abstract the information gained from one embodiment in a way such that the resulting knowledge can be made useful for the decision of another embodiment. In our previous example, the nutritional recommendation would be based on abstraction of the current work performance level and building relation to the physiological effect of certain foods.

This abstraction and interpretation process for each receiver and sender of information is a key component of the shared representation for cooperative intelligence.

V. Conclusion

The future of AI is in intense discussion, where autonomy is more and more augmented by the idea of cooperation between human and machine, to jointly exploit the strength of each interaction partner and alleviate their weaknesses. However, a sociological (and evolutionary) view on cooperation allows a complimentary and more comprehensive perspective on the relation that we envision between humans and intelligent systems. Here, cooperation goes beyond the idea of joint work and defines an empathic relationship, a state or frame of mind that is deeply related to trust and trustworthy behavior – another property of AI that is fervently discussed. Sustained interaction with the long-term sharing of experiences is one path towards establishing an empathic relation leading to cooperation. Cyber-physical systems in a personal context with a centralist knowledge representation can be the enabling technology for sustained heterogeneous interaction and for cooperative intelligence. We advocate a more sociological or humane perspective in the technologically dominated research in artificial intelligence to realize the vision of a hybrid society.

ACKNOWLEDGMENT

The concepts that we have outlined in this paper have been the result of many discussions among members of the Honda Research Institutes. In particularly, we would like to thank the following colleagues for their contributions: J. Eggert, K. Fujimura, C. Goerick, T. Hasegawa, M. Krüger, M. Olhofer, A. Richter, J. Schmüdderich, H. Tsujino, and C. Wiebel-Herboth. We thank E. Feigenbaum for discussions on the role of cooperation and empathy for future intelligent systems.

REFERENCES

[1] M. H. Jarrahi, "Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making," *Business Horizons*, vol. 61, no. 4, pp. 577–586, 2018.

[2] Y. Chen and A. S. Cheung, "The transparent self under big data profiling: privacy and Chinese legislation on the social credit system," *J Comp Law*, vol. 12, no. 2, pp. 356–378, 2017.

[3] L. E. Parker, "Creation of the national artificial intelligence research and development strategic plan." *AI Magazine*, vol. 39, no. 2, 2018.

[4] National Science and Technology Council (US), *The National Artificial Intelligence Research and Development Strategic Plan*. Exec Office President of the United States, 2016.

[5] B. Sendhoff, "Cooperative individual mobility," SAE 2015 World Congress, Session: Future of Mobility Innovation, 2015, invited talk.

[6] M. Walch, K. Mühl, J. Kraus, T. Stoll, M. Baumann, and M. Weber, "From car-driver-handovers to cooperative interfaces: Visions for driver–vehicle interaction in automated driving," in *Automotive User Interfaces*, 2017, pp. 273–294.

[7] C. Maag, N. Schneider, T. Lübbeke, T. H. Weisswange, and C. Goerick, "Car gestures–advisory warning using additional steering wheel angles," *Accid Anal Prev*, vol. 83, pp. 143–153, 2015.

[8] N. Schoemig, M. Heckmann, H. Wersing, C. Maag, and A. Neukum, ""Please watch right"–evaluation of a speech-based on-demand assistance system for urban intersections," *Transp Res F*, vol. 54, pp. 196–210, 2018.

[9] J. Piaget, *Études sociologiques*. Librairie Droz, 1965.

[10] J.-M. Hoc, "Towards a cognitive approach to human–machine cooperation in dynamic situations," *Int J human-computer studies*, vol. 54, no. 4, pp. 509–540, 2001.

[11] M. E. Bratman, "Shared cooperative activity," *Philos Rev*, vol. 101, no. 2, pp. 327–341, 1992.

[12] H. Moll and M. Tomasello, "Cooperation and human cognition: the Vygotskian intelligence hypothesis," *Philos Trans R Soc B*, vol. 362, no. 1480, pp. 639–648, 2007.

[13] L. S. Vygotsky, *Mind in society: The development of higher psychological processes*, 1980.

[14] L. E. Parker and F. G. Pin, "Man-robot symbiosis: a framework for cooperative intelligence and control," in *Space Station Automation IV*, vol. 1006, 1988, pp. 94–103.

[15] L. Panait and S. Luke, "Cooperative multi-agent learning: The state of the art," *Auton Agents Multi-Agent Syst*, vol. 11, no. 3, pp. 387–434, 2005.

[16] T. B. Sheridan, *Telerobotics, automation, and human supervisory control*, 1992.

[17] J. Fryer and G. McKee, "Towards distributed and cooperative intelligence in a networked robotic system," in *U of Salford*, 1995.

[18] M. Dorigo and L. M. Gambardella, "Ant colony system: a cooperative learning approach to the traveling salesman problem," *IEEE T Evolut Comput*, vol. 1, no. 1, pp. 53–66, 1997.

[19] R. Eberhart and J. Kennedy, "Particle swarm optimization," in *Proc IEEE IJCNN*, vol. 4, 1995, pp. 1942–1948.

[20] J. M. Bradshaw, R. R. Hoffman, D. D. Woods, and M. Johnson, "The seven deadly myths of "autonomous systems"," *IEEE Intell Syst*, vol. 28, no. 3, pp. 54–61, 2013.

[21] A. Schmidt and T. Herrmann, "Intervention user interfaces: a new interaction paradigm for automated systems." *Interactions*, vol. 24, no. 5, pp. 40–45, 2017.

[22] M. Krüger, C. B. Wiebel, and H. Wersing, "From tools towards cooperative assistants," in *Proc Int Conf Human Agent Interaction*, 2017, pp. 287–294.

[23] A. G. Billard, S. Calinon, and R. Dillmann, "Learning from humans," in *Springer Handbook of Robotics*, 2016, pp. 1995–2014.

[24] A.-L. Vollmer and L. Schillingmann, "On studying human teaching behavior with robots: a review," *Rev Philos Psychol*, vol. 9, no. 4, pp. 863–903, 2018.

[25] S. Manschitz, J. Kober, M. Gienger, and J. Peters, "Learning movement primitive attractor goals and sequential skills from kinesthetic demonstrations," *Robotics and Autonomous Syst*, vol. 74, pp. 97–107, 2015.

[26] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, "Postural optimization for an ergonomic human-robot interaction," in *IEEE/RSJ Int Conf Intell Robots & Systems*, 2017, pp. 2778–2785.

[27] P. Maurice, V. Padois, Y. Measson, and P. Bidaud, "Human-oriented design of collaborative robots," *Int J Ind Ergonom*, vol. 57, pp. 88–102, 2017.

[28] K. Sommer, "Continental mobility study 2011," *Continental AG*, 2013.

[29] M. F. Dollard and A. H. Winefield, "Mental health: overemployment, underemployment, unemployment and healthy jobs," *Adv Mental Health*, vol. 1, no. 3, pp. 170–195, 2002.

[30] A. H. Maslow, "A theory of human motivation." *Psychol Rev*, vol. 50, no. 4, p. 370, 1943.

[31] T. S. Pittman and K. R. Zeigler, "Basic human needs," *Social psychology: Handbook of basic principles*, vol. 2, pp. 473–489, 2007.

[32] D. Gambetta, "Can we trust trust?" in *Trust: Making and Breaking Cooperative Relations*, D. Gambetta, Ed., 1988, pp. 213–237.

[33] I. Wachsmuth, "Embodied cooperative systems: From tool to partnership," in *Collective Agency and Cooperation in Natural and Artificial Systems*, 2015, pp. 63–79.

[34] M. Gienger, D. Ruiken, T. Bates, M. Regaieg, M. MeiBner, J. Kober, P. Seiwald, and A.-C. Hildebrandt, "Human-robot cooperative object manipulation with contact changes," in *IEEE/RSJ Int Conf Intell Robots & Systems*, 2018, pp. 1354–1360.

[35] S. Vinanzi, C. Goerick, and A. Cangelosi, "Mindreading for robots: Predicting intentions via dynamical clustering of human postures," in *Proc IEEE Int Conf Dev Learning*, 2019, pp. 272–277.

[36] M. C. Buehler and T. H. Weisswange, "Online inference of human belief for cooperative robots," in *IEEE/RSJ Int Conf Intell Robots & Systems*, 2018, pp. 409–415.

[37] M. J. Matarić, "Socially assistive robotics: Human augmentation versus automation," *Science Robotics*, vol. 2, no. 4, 2017.

[38] D. Norman, *The design of future things*, 2009.

[39] M. Krüger, B. N. Walker, and L. Chuang, "The embodied vehicle," in *Proc AutoUI Adj Proc*, 2019, pp. 50–55.

[40] F. Rossi, "Building trust in artificial intelligence," *J Int Aff*, vol. 72, no. 1, pp. 127–134, 2019.

[41] K. Siau and W. Wang, "Building trust in artificial intelligence, machine learning, and robotics," *Cutter Business Tech J*, vol. 31, no. 2, pp. 47–53, 2018.

[42] "Asilomar AI principles," The Future of Life Institute. https://futureoflife.org/ai-principles/, 2018.

[43] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, p. 15, 1973.

[44] I. Asimov, *Runaround. I, Robot*, 1950.

[45] E. LaRosa and D. Danks, "Impacts on trust of healthcare AI," in *Proc AAAI/ACM Conf AI, Ethics, and Society*, 2018, pp. 210–215.

[46] R. Falcone and A. Sapienza, "On the users' acceptance of IoT systems: A theoretical approach," *Information*, vol. 9, no. 3, p. 53, 2018.

[47] S. Biundo and A. Wendemuth, "Companion-technology for cognitive technical systems," *KI-Künstliche Intell*, vol. 30, no. 1, pp. 71–75, 2016.

[48] M. Mühlig, S. Hasler, L. Fischer, and J. Deigmöller, "A knowledge-based multi-entity and cooperative system service architecture," in *IEEE Int Conf Hum-Mach Syst*, 2020.

[49] S. Fuchs, N. Einecke, and F. Eisele, "SmartLobby: Using a 24/7 remote head-eye-tracking for content personalization," in *Adj Proc ACM Int J Conf Pervasive & Ubiquitous Comp and Proc ACM Int Symp Wearable Computers*, 2019, p. 53–56.

[50] A. Sadik, C. Goerick, and M. Mühlig, "Modeling and simulation of a multi-robot system architecture," in *Int Conf Mechatronics, Robotics & Syst Eng (MoRSE)*, 2019, pp. 8–14.