RIFL (2020) Vol. 14, n. 2: 63-75
DOI: 10.4396/AISB201906

The Communication Problem

Micheal Straeubig
University of Plymouth
michael.straeubig@plymouth.ac.uk

Abstract Analysis, interpretation and construction of artificial and natural languages have been central concerns of artificial intelligence since the 1950s.
Current applications for automated language progressing range from real-time translation of spoken language through automated discovery of sentiment in online postings to conversational agents embedded in everyday devices. Recent developments in machine learning, combined with the availability of large amounts of labelled training data, have enabled non-structural approaches to largely surpass classical techniques based on formal grammars, conceptual ontologies and symbolic representations. As the complexity and opaqueness of those stochastic models becomes more and more evident, however, the question arises if we trade gains in observable performance with a literal loss of understanding.
This article presents a distinction-based approach to critically re-visit fundamental theoretical concepts such as code, information, language, communication and meaning. I will follow Niklas Luhmann's theory of social systems by locating communication firmly within social systems. Departing from Luhmann, I do invite machines as participants into some of these systems. Finally, I propose to employ Friedemann Schulz von Thun's 4-sided communication model in order to overcome the current information-theoretic emphasis of communication.

Keywords: communication, AI, systems theory, constructivism, cybernetics, language

Received 27/02/2020; accepted 16/03/2020.

0. A the beginning

«How can a computer be programmed to use a language?» is one of the seven questions put forward in the proposal that sparked the seminal Dartmouth conference on AI in 1956. The subject is then elaborated further:

> It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning and rules of conjecture. From this point of view, forming a generalisation consists of admitting a new word and some rules whereby sentences containing it imply and are implied by others. This idea has never been very precisely formulated nor have examples been worked out (McCarthy *et al.* 1955).

1. Information vs. Communication

A few years before the official birth of AI, Shannon and Weaver lay out their groundbreaking model of communication (Shannon, Weaver 1949), based on the transmission of information over a noisy channel: «The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point» (Shannon 1948: 1). By defining information in mathematical terms based on thermodynamic entropy, Shannon manages to abstract the message from the medium (such as the telegraph, telephone, television). Weaver however gives a much broader view of communication as «all of the procedures by which one mind may affect another» (Shannon, Weaver 1949: 1).

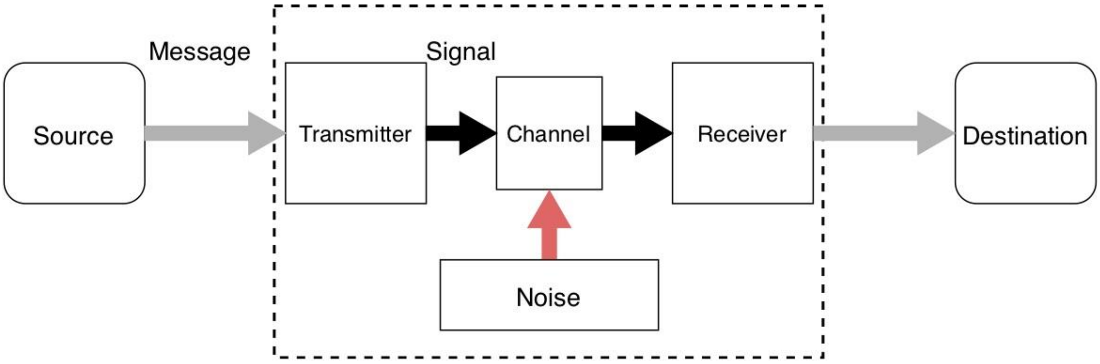

Fig. 1

Shannon later warns against the out of hand use of his theory that he firmly locates within the context of engineering (Shannon 1956). Despite this, Shannon's concept of information is ubiquitous today (Gleick 2011) while an understanding of its relation to communication is still missing, as apparent explanations of phenomena like curiosity, creativity and art demonstrate (Schmidhuber 2010). Claiming to explain creativity in terms of data compression is akin to explaining human consciousness in terms of chemical structure—it is a category mistake. Without a distinction between data and information (Floridi 2011) and without meaningful selections from possible differences (Bateson 2000) it is impossible to implement communication apart from pure information theory.

2. Cognition vs. Action

As manifest in the opening quotations, the fledgling field of AI begins to observe communication through the distinction between thought and language, a hotly debated issue in analytic philosophy since Wittgenstein (Wittgenstein 1922; Wittgenstein 1958; Thornton 1998; Chihara, Fodor 1965). McGinn (1996) discusses various positions, distilled into the question if thinking necessarily requires language. He denies a conclusive proof; however, both traditional research in human language development (Barrett 1999) and computer-linguistic practice (Jurafsky, Martin 2009) are commonly co-locating linguistic and cognitive capabilities. Searle (1980) illustrates his rhetorical question about consciousness in the machine by a translation metaphor, whereas possible mechanisms of symbolic grounding are debated in contexts of connectionism (Harnad 1990) and enactment (Steels 2008).
The connection of speech to intentions, expectations and to effects and consequences beyond the communicative situation is captured in the concept of speech acts (Austin 1962; Searle 2011). It is implemented, for example, in Grosz and Sidner's (1986)

RIFL (2020) Vol. 14, n. 2: 63-75
DOI: 10.4396/AISB201906

discourse structure, integrating sequences of utterances with dynamical states of attention and intentions. Jurafsky and Martin (2009) discuss various aspects that differentiate dialogue from other natural language processing tasks: turn taking, utterances, grounding, implicature and coherence. Their operationalisation is suggested through extended notions of speech acts (Power 1979; Bunt 1994) and through concepts of conversational games (Carletta *et al.* 1997).

3. Human vs. Machine and Simulation vs. Performance

By focussing solely on the performative aspect, Turing's (1950) imitation game represents a completely different approach to communication. Turing is not concerned with *how* the communicative effect is achieved—a much-debated philosophy (Harnad 1992; Hayes, Ford 1995; Cohen 2005; Epstein, Roberts, Beber 2009) that has survived in form of competitions like the Loebner-Prize (Powers 1998). It is also relevant today in practical applications such as the construction of believable non-player characters for video games (Hingston 2012). The black box is supposed to stay closed for the player.

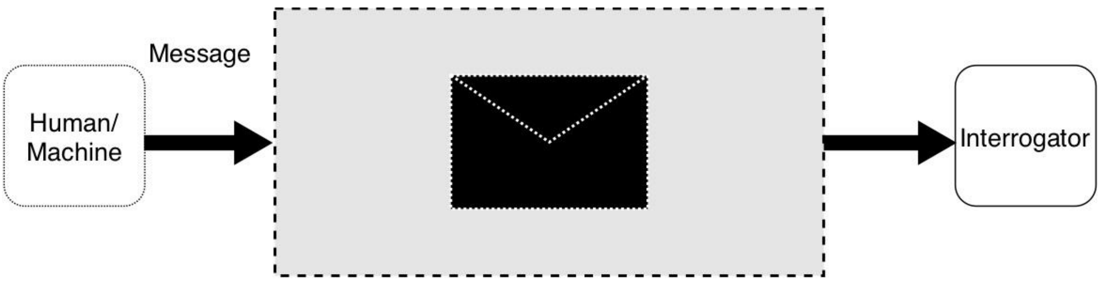

Fig. 2

But can we rely on the black box approach to evaluate the progress of AI regarding the communicative capabilities of machines? For Hernández-Orallo (2014) this discussion goes back to the rift between McCarthy's (1996) definition of AI as «[…] the science and engineering of making intelligent machines, especially intelligent computer programs» vs. the one by Minsky (1969: V) «[AI is] the science of making machines capable of performing tasks that would require intelligence if done by [humans]». In both cases, in order to evaluate or implement the capability for communication in a machine we are forced to make further distinctions, e.g. following Marr's (2000) organisational hierarchy of a computational model, algorithmic representation and physical implementation[1]. One must choose a computational approach, select an algorithm, pick a technical platform.

4. Symbolic AI vs. Machine Learning

By making use of deep neural network architectures and large amounts of training data, connectionist approaches have surpassed other techniques that were based on formal grammars, conceptual ontologies and symbolic representations. They have shown an impressive rate of progress. Recurrent neural networks (RNN) that generate stochastic predictions for the next element in a sequence through supervised learning (Sutskever, Martens, Hinton 2011; Wen *et al.* 2015; Wu, Martinez, Klyen 2018) are now being outpaced by unsupervised Transformer models which do not rely on task-specific input

[1] For a critique of applying Marr's categorisation to humans, see (McClamrock 1991).

text (Radford *et al.* 2019a). As the complexity and opaqueness of those models becomes more and more evident, however, the question arises if we trade gains in observable performance against a literal loss of understanding. Ashby's (1956) cybernetic «black box» re-appears as the other participant in the medium of communication.

In addition, a second black box that surrounds the first one is beginning to affect research. Upon publication of their groundbreaking GPT-2 model (Radford *et al.* 2019a) the research institute OpenAI announced that they that would keep the full trained model a secret, in order to «prevent malicious applications of the technology» (Radford *et al.* 2019b)[2]. It is also unfortunate that Microsoft has published no comprehensive analysis of their Twitter chatbot project «Tay» (it is more than ironic that both failed and successful projects apparently need to be shielded from further research), thereby missing an opportunity to learn from mistakes (Schlangen 2009). The company pulled their failed experiment within just 24 hours after malevolent internet users discovered that Tay would mirror any input fed to it, including misogynistic, homophobic and outright racist statements (Neff 2016). Tay was «learning» in the limited sense that this concept is currently used in the field of machine learning (Zador 2019); it had no option of distancing itself from incoming information. Without distinguishing between itself and others and between message and information, it could not communicate, as I will argue below. This is not the same as to require consciousness in a machine. Drawing a distinction between oneself and others is a necessary, yet not a contingent condition for consciousness (Maturana, Varela 1980).

The same goes for Mitsuku, advertised as «four-time winner of the Loebner Prize Turing Test, [it] is the world's best conversational chatbot» (2019). Mitsuku is based on Pandorabots, a platform that allows creating «intelligent conversational agents» (Pandorabots, Inc. 2019) in the markup language AIML (Wallace 2003). Similar commercially available chatbots are essentially evolved forms of the pattern-matching and rule-based beginnings of conversational AI, Weizenbaum's (1966) script ELIZA.

Here is a short dialogue with Mitsuku.

1a. MITSUKU: «I am Mitsuku. I want to be your friend».
1b. MITSUKU: «Hi. I'm the world's most humanlike conversational AI».
2a. ME: «Hi Mitsuku».
3a. MITSUKU: «Hi there».
4. ME: «How is it to be a chatbot?».
5. MITSUKU: «No different from being anything else I suppose».
6. ME: «But you are the world's best».
7. MITSUKU: «Thanks but surely I am not the only one».
8. ME: «Well, there can be only one world's best».
9. MITSUKU: «How about that. What proof do you have of that?».
10. ME: «"Best" is a superlative».
11. MITSUKU: «I don't have a favorite is a superlative».
12b. MITSUKU: (Shows a picture of Arnold Schwarzenegger as Terminator).
13. MITSUKU: «. But my favorite movie is Terminator.»

The black box opens up when the dialogue deteriorates semantically at line 7. It falls apart from line 11, when syntax, semantics and punctuation (line 13) don't work anymore.

Two remarks are in order: I conducted the dialogue twice, on May 5, 2018 and on March 20, 2019. Mitsuku's responses were identical with the exception of the greeting

2 In the meantime OpenAI has released multiple versions of their contribution, including the full model.

(line 1a, 1b), an additional dialogue line initiated by me (line 2a, 3a) and the picture that Mitsuku inserted at line 12b in the latter conversation. Also note that in conducting this dialogue I did not intend to deliberately break or fool the conversational agent. Instead, I kept my attempt at communication along the lines how I would have responded to a human in a casual conversation. Still it is evident that the system does not come close to communication in any sense of the word.

5. Syntax vs. Semantics vs. Pragmatics

Peeking inside the black box, we are likely to find tensors (Sutskever, Martens, Hinton 2011). Alternatively we return to symbolic AI for representations of syntax, (formal) semantics, and pragmatics (Allen 1995; Jurafsky, Martin 2009; Russell, Norvig 2010: 860-1019). Automating syntax leads to formal grammars and production rules of languages, automating semantics leads to knowledge representation, for example through logical forms and semantic networks. Adding pragmatic aspects such as general, task-specific or contextual knowledge leads to other forms of knowledge representation (Minsky 1974), sometimes augmented with constructivist concepts (Drescher 1991).

In this picture, communication is largely a mechanical, symmetrical process. The receiver parses a message and transforms it into some form of knowledge representation. This is then combined with contextual knowledge and made available for techniques that simulate cognitive achievements such as planning or inference. In order to generate a message, the language pipeline is run in reverse. In contrast to stochastic and connectionist procedures, we encounter glass boxes, algorithms that are precisely understood yet of limited capabilities. Their underlying concepts are borrowed from semiotics (Morris 1971) to automate the three aspects of messages, understood as complexes of signs that take part in operations of communication and designation. According to Eco (1978), the latter makes the difference between a mere stimulus-response driven interaction and a semiotic process.

6. Nature vs. Nurture

Linguistics has occupied itself with a long-standing debate about the nature of human capabilities as structurally innate (Chomsky 1957) versus self-constructed through interactions with the environment (Tomasello 2005). We observe a similar rift between the classical AI approaches that rely on pre-programmed structures and neural networks, which can be conceived as universal approximation functions. The field of social robotics has followed a third direction by embracing constructivist ideas such as Piaget's (1952) model of stepwise development (Cangelosi, Schlesinger 2015). Here, linguistic capabilities are acquired through enactment in laboratory situations such as language games (Lewin and Lane 2000), which by Steels' (2008) account deliver a solution to the grounding problem. However, taking a closer look at the products of emergent processes seems appropriate (Kottur *et al.* 2017).

7. Social Systems

During the previous sections, I have highlighted some of the issues that arise in the development of communication between humans and machines. I have based my observations on specific, historically evolved distinctions. Crossing these distinctions sometimes means that one has to traverse disciplinary boundaries as well. Yet many of the concepts that have emerged lump together concepts across systemic boundaries: Peirce's elaborate semiotic structures and Austin and Searle's locutionary speech act taxonomies seek to describe communicative phenomena in terms of information, meaning, cognition, propositions, intentions, utterances, references and much, much more. It is difficult to see how those concepts can be translated into practice.

In the end, the unsurmountable level of complexity can only find refuge in eschatology (Kurzweil 2009). To avoid taking the road to singularities, I am siding with Turing's remark about consciousness: «But I do not think these mysteries necessarily need to be solved before we can answer the question with which we are concerned in this paper». (Turing 1950: 447) Whereas consciousness in my opinion is not a necessary condition for the communication problem, the latter nevertheless does appear to be a hard problem (Chalmers 1995) of its own.

In the following I propose two initial steps towards a solution. The first one is necessary to clarify the context of communication and the second one to capture its facets in a practice-based, empirical way. The first step involves reducing and the second one increasing complexity. Both are achieved by observing different sets of distinctions.

I will first outline the context of communication in Niklas Luhmann's (1996) social systems theory. Luhmann distinguishes biological, psychic (minds) and social systems. These kinds of systems are structurally coupled but closed under their own operations. They operate with different codes and different distinctions. Most importantly, communication takes place in social systems; neither minds nor neurons (nor humans for that matter) communicate.

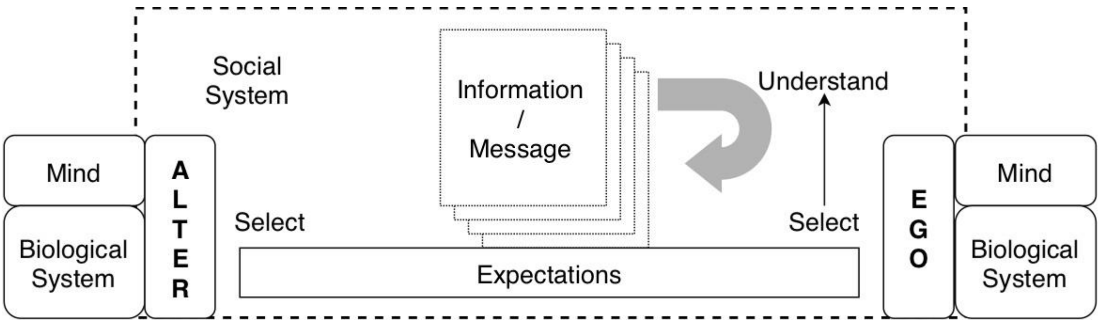

Fig. 3

In contrast to Luhmann, I do invite machines as participants into social systems (Straeubig 2017). While in Luhmann's account, psychic systems are structurally coupled with social systems, I do not believe that minds (or brains) are a necessary condition for communicative capabilities. At a minimum, a machine must be able to act as an observer, to draw distinctions between itself and the other, and to distinguish between message and information. On this foundation it can form expectations that allow it to take part in communication. This means we can avoid speculating about conscious machines or try building bottom-up biologistic simulations, in the hope that something will somehow emerge.

Instead, the plan is to focus on communication itself, both to «reintroduce communication into cybernetics» (Baecker 1997) and to reintroduce cybernetics into communication. But what do the participants in a communicative situation actually observe? This is the topic of the final section.

8. Four Sides of Communication

Friedemann Schulz von Thun's (1981) four-sided communication model integrates concepts from Bühler's (2011) Organon model and Watzlawick's distinction between

content and relationship of messages (Watzlawick, Bavelas, Jackson 2014). In this model, each act of communication has four sides, both for the sender and for the receiver: facts, relationship, self-presentation and appeal. These four facets or subtexts appear in almost every message and can be observed and analysed individually, with regard to their relative emphasis or in terms of their congruence.
From the perspective of the sender, the factual side contains the actual subject matter of the message. The self-presentation side carries both intentional (self-promotions) and unintended (self-revelations) expressions of the sender. This theme is elaborated further by Goffman (1990) in his observations about social encounters. The relationship side encodes how the sender views the receiver and the relationship between them. Finally, each act of communication also carries appeals—these are actions that the sender intends for the receiver to carry out. Appeals can be communicated openly (advice, a command) or hidden (manipulation).

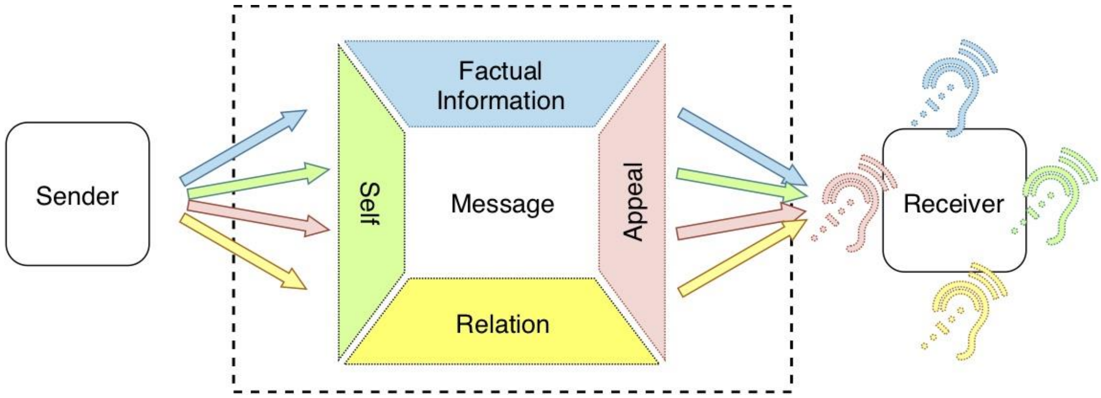

Fig. 4

Von Thun (1981: 25-31) illustrates the concept through an example in which a couple is driving and the partner on the passenger seat says: «The traffic lights ahead are green.» The answer of the driver is: «Who is driving, you or me?»[3].
Analysing the factual content of this brief exchange poses no challenge. Assumed that the observable context is as stated, both speakers are expressing facts about the situation. However, we discover that there is more to this conversation. The self-presentation might be interpreted as the passenger being impatient, in a hurry, or just wanting to be helpful. The relationship aspect of the message conveys that he/she might see themselves as the better driver or more attentive to the situation. The implicit appeal to the driver is to step on the accelerator and drive faster.
In analogy to the sender, who «speaks with four mouths», the receiver simultaneously «listens with four ears», but not necessarily with the same intensity. If one side is amplified out of proportion, the receiver's perception becomes contorted. An exaggerated factual ear ignores interpersonal clues, whereas an ear tuned to relation cannot perceive the factual content. An ear listing purely for self-representation would come across as therapeutic, while an appeal-focused listener would be likely to act with excessive alacrity.
In our example, the driver could easily agree with the factual side, or notice the passenger's self-revelation, but as evident from the answer, he/she listens mainly on the

[3] In the original publication (Schulz von Thun 1981: 25-31), the situation is gendered and it would be interesting to investigate how this affects the interpretation. I chose to de-gender the account for the present discussion.

relationship and appeal side. The response also communicates that the driver sees themselves in charge of the situation and won't accept being lectured.
This four-sided model is derived from a long-standing practice regarding human communication. It is easy to understand, focuses on concrete situations, and allows an encompassing and precise observation of communication-related phenomena. It works with written, verbal and non-verbal communication. What remains to do is to employ the model when we replace one or all human participants with machines.

9. Conclusion

Natural language processing is a central concern of artificial intelligence since the 1950s. However, comprehensive and practical models for implementing communication are still missing. At the same time the rift between social robotics, connectionist and symbolic approaches in AI is growing. The former is embracing constructivist, embodied and enactive approaches (Cangelosi, Schlesinger 2015) while the latter resort to formal and idealised models (Bard *et al.* 2019; Grice *et al.* 2013). This is despite, possibly because of prior efforts seeking fundaments in information-theoretic, structural-linguistic or cognitive models while largely ignoring social aspects as illustrated above.
I argue that different approaches are needed. We use language to communicate and we understand language through its use. Therefore, we must start from pragmatics (Allwood 1978), not from syntax, and observe social systems from a transdisciplinary perspective. As I have set out before (Straeubig 2017), we also should invite machines into our social systems and grant them permanent presence. We can then observe the various aspects of communication both from the perspective of the sender and the receiver. In the present article I have sketched Luhmann's and von Thun's approaches and hinted at their cybernetic, constructivist, and practice-based backgrounds. A lot remains to be done in future work to operationalise these ideas.
Only then, I claim, can the term communication be used, as Weaver envisioned, «in a very broad sense to include all of the procedures by which one mind may affect another».

References

Allen, J.F. (1995), *Natural language understanding*, Pearson, Redwood City (Ca).

Allwood, J. (1978), «A Bird's Eye View of Pragmatics», in Gregersen, K. (ed.), *Papers from the Fourth Scandinavian Conference of Linguistics*, Odense University Press, pp. 145-159.

Ashby, W.R. (1956), *An Introduction to Cybernetics*, Martino Fine Books, Eastford.

Austin, J.L. (1962), *How to do things with words: the William James lectures delivered at Harvard University in 1955*, Harvard Univ. Press, Cambridge (Mass).

Baecker, D. (1997), «Reintroducing Communication into Cybernetics», in *Systemica*, vol. 11, pp. 11-29.

Bard, N. *et al.* (2019), «The Hanabi Challenge: A New Frontier for AI Research», in arXiv:1902.00506, from: http://arxiv.org/abs/1902.00506.

Barrett, M. (ed.) (1999), *The development of language*, Psychology Press, Hove.

Bateson, G. (1972), «Form, substance and difference», in *Steps to an ecology of mind*, University of Chicago Press, Chicago, 2000, pp. 454-471.

Bühler, K. (1934), *Theory of Language: The representational function of language*, John Benjamins Publishing Company, Amsterdam.

Bunt, H. (1994), «Context and Dialogue Control», in *Think*, vol. 3, pp. 19-31.

Cangelosi, A., Schlesinger, M. (2015), *Developmental robotics: from babies to robots*, The MIT Press, Cambridge (Mass).

Carletta, J. *et al.* (1997), «The Reliability of a Dialogue Structure Coding Scheme», in *Computational Linguistics*, vol. 23, n. 1, pp. 13-31.

Chalmers, D.J. (1995), «Facing up to the problem of consciousness», in *Journal of consciousness studies*, vol. 2, n. 3, pp. 200-219.

Chihara, C.S., Fodor, J.A. (1965), «Operationalism and Ordinary Language: A Critique of Wittgenstein», in *American Philosophical Quarterly*, vol. 2, n. 4, pp. 281-295.

Chomsky, N. (1957), *Syntactic structures*, Martino, Mansfield Centre (Conn.).

Cohen, P.R. (2005), «If not Turing's test, then what?», in *AI magazine*, vol. 26, n. 4, p. 61.

Drescher, G.L. (1991), *Made-up minds: a constructivist approach to artificial intelligence*, MIT Press, Cambridge (Mass).

Eco, U. (1978), *A theory of semiotics*, Indiana University Press, Bloomington.

Epstein, R.S., Roberts, G., Beber, G. (2009), *Parsing the Turing test: philosophical and methodological issues in the quest for the thinking computer*, Springer, Dordrecht, London.

Floridi, L. (2011), *The philosophy of information*, Oxford Univ. Press, Oxford.

Gleick, J. (2011), *The information: a history, a theory, a flood*, Vintage Books, New York.

Goffman, E. (1990), *The presentation of self in everyday life*, Penguin, London.

Grice, H.P. (2013), «Logic and Conversation», in Ezcurdia, M. (ed.), *The semantics-pragmatics boundary in philosophy*, Broadview Press, Peterborough (Ont.), pp. 47-59.

Grosz, B.J., Sidner, C.L. (1986), «Attention, intentions and the structure of discourse», in *Computational Linguistics*, vol. 12, n. 3, p. 30.

Harnad, S. (1990), «The Symbol Grounding Problem», in *Physica*, vol. D, n. 42, pp. 335-346.

Harnad, S. (1992), «The Turing Test is not a trick: Turing indistinguishability is a scientific criterion», in *ACM SIGART Bulletin*, vol. 3, n. 4, pp. 9-10.

Hayes, P., Ford, K. (1995), «Turing test considered harmful», in *IJCAI* (1), pp. 972-977.

Hernández-Orallo, J. (2014), «AI Evaluation: past, present and future», in arXiv preprint arXiv:1408.6908, from: http://arxiv.org/abs/1408.6908.

Hingston, P. (ed.) (2012), *Believable bots: can computers play like people?*, Springer, Berlin, New York.

Jurafsky, D., Martin, J. H. (2009), *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*, Prentice Hall, Pearson Education Internat, Upper Saddle River (NJ).

Kottur, S. *et al.* (2017), «Natural Language Does Not Emerge 'Naturally 'in Multi-Agent Dialog», in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Copenhagen, pp. 2962-2967.

Kurzweil, R. (2009), *The singularity is near: when humans transcend biology*, Duckworth, London.

Lewin, I. and Lane, M. (2000), «A formal model of Conversational Game Theory», in *In Proc. Gotalog-00*, 4th Workshop on the Semantics and Pragmatics of Dialogue, Gothenburg.

Luhmann, N. (1996), *Social systems*, Stanford University Press, Stanford.

Marr, D. (1982), *Vision: a computational investigation into the human representation and processing of visual information*, Freeman, New York, 2000.

Maturana, H.R., Varela, F.J. (1980), *Autopoiesis and cognition: the realization of the living*, D. Reidel Pub. Co., Dordrecht, Boston.

McCarthy, J. *et al.* (1955), «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence», in *AI Magazine*, vol. 27, n. 4, p. 3, 2006.

McCarthy, J. (1996), *What is Artificial Intelligence?*, Computer Science Department, Stanford University, 2007, from: http://www-formal.stanford.edu/jmc/whatisai/node1.html.

McClamrock, R. (1991), «Marr's three levels: A re-evaluation», in *Minds and Machines*, vol. 1, n. 2, pp. 185-196.

McGinn, C. (1996), «Thought and Language», in *The character of mind: an introduction to the philosophy of mind*, Oxford University Press, Oxford, New York, pp. 83-106.

Minsky, M.L. (1969), *Semantic information processing*, The MIT Press, Cambridge, London 2015.

Minsky, M.L. (1974), *A Framework for representing knowledge*, Memo 306, MIT AI Laboratory.

Morris, C.W. (1971), *Writings on the General Theory of Signs*, De Gruyter Mouton, The Hague, Paris.

Neff, G. (2016), «Talking to Bots: Symbiotic Agency and the Case of Tay», in *International Journal of Communication*, vol. 10, pp. 4915-4931.

Pandorabots, Inc. (2019), Pandorabots, from: www.pandorabots.com.

Piaget, J. (1952), *The Origins of Intelligence in Children*, International Universities Press, New York.

Power, R. (1979), «The organisation of purposeful dialogues», in *Linguistics*, vol. 17, n. 1-2.

Powers, D.M.W. (1998), «The total Turing test and the Loebner prize», in *Proceedings of the Joint Conferences on New Methods in Language Processing and Computational Natural Language Learning*, pp. 279-280.

Radford, A., Wu, J., Child, R. *et al.* (2019a), *Language Models are Unsupervised Multitask Learners*, OpenAI.

Radford, A., Wu, J., Amodei, D. *et al.* (2019b*), Better Language Models and Their Implications*, OpenAI, from: https://openai.com/blog/better-language-models/.

Russell, S.J., Norvig, P. (2010), *Artificial intelligence: a modern approach*, Prentice Hall, Upper Saddle River.

Schlangen, D. (2009), «What we can learn from Dialogue Systems that don't work», in *Proceedings of DiaHolmia, the 13th International Workshop on the Semantics and Pragmatics of Dialogue*, (SEMDIAL 2009), pp. 51-58.

Schmidhuber, J. (2010), «Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990-2010)», in *IEEE Transactions on Autonomous Mental Development*, vol. 2, n. 3, pp. 230-247.

Schulz von Thun, F. (1981), *Miteinander reden: Störungen und Klärungen: Psychologie der zwischenmenschlichen Kommunikation*, Rowohlt, Reinbek bei Hamburg.

Searle, J.R. (1969), *Speech acts: an essay in the philosophy of language*, Cambridge Univ. Press, Cambridge, 2011.

Searle, J.R. (1980), «Minds, brains, and programs», in *Behavioral and Brain Sciences*, vol. 3, n. 3, pp. 417-457.

Shannon, C.E. (1948), «A mathematical theory of communication», in *The Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656.

Shannon, C.E. (1956), «The bandwagon», in *IRE Transactions on Information Theory*, vol. 2, n. 1, p. 3.

Shannon, C.E., Weaver, W. (1949), *The mathematical theory of communication*, Univ. of Illinois Press, Urbana, 1998.

Steels, L. (2008), «The symbol grounding problem has been solved. so what's next», in *Symbols and embodiment: Debates on meaning and cognition*, pp. 223-244.

Straeubig, M. (2017), «Let the Machines out. Towards Hybrid Social Systems», in *Proceedings of AISB Annual Convention 2017*, Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB), Bath, pp. 28-31.

Sutskever, I., Martens, J., Hinton, G. (2011), «Generating Text with Recurrent Neural Networks», in *Proceedings of the 28th International Conference on International Conference on Machine Learning*, Omnipress, Anderson St, Madison (Wi), pp. 1017-1024.

Thornton, T. (1998), *Wittgenstein on language and thought: the philosophy of content*, Edinburgh University Press, Edinburgh.

Tomasello, M. (2005), *Constructing a language: a usage-based theory of language acquisition*, Harvard Univ. Press, Cambridge (Mass.).

Turing, A. (1950), «Computing machinery and intelligence», in *Mind*, 59 (236), pp. 433-460.

Wallace, R.S. (2003), *The Elements of AIML Style*, ALICE A.I. Foundation, Inc.

Watzlawick, P., Bavelas, J.B., Jackson, D.D. (2014), *Pragmatics of human communication: a study of interactional patterns, pathologies, and paradoxes*, W. W. Norton & Company, New York.

Weizenbaum, J. (1966), «ELIZA-a computer program for the study of natural language communication between man and machine», *in Communications of the ACM*, vol. 9, n. 1, pp. 36-45.

Wen, T.-H. *et al.* (2015), «Stochastic Language Generation in Dialogue using Recurrent Neural Networks with Convolutional Sentence Reranking», in *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Association for Computational Linguistics, Prague, pp. 275-284.

Wittgenstein, L. (1922), *Tractatus logico-philosophicus*, Cosimo Classics, New Jork, 2007.

Wittgenstein, L. (1958), *Philosophical investigations*, Basil Blackwell, Oxford.

Worswick, S. (2019), Mitsuku, from: http://www.mitsuku.com.

Wu, X., Martinez, A., Klyen, M. (2018), «Dialog Generation Using Multi-Turn Reasoning Neural Networks», in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1, Association for Computational Linguistics, New Orleans, pp. 2049-2059.

Zador, A. (2019), «A Critique of Pure Learning: What Artificial Neural Networks can Learn from Animal Brains», in *bioRxiv*, from: http://biorxiv.org/lookup/doi/10.1101/582643.