

Arildo Marconatto and Edson Sestari

Algorithm and Language

Abstract: *The aim of this paper is to highlight characteristic elements of the relationship between algorithmic data processing and human language. We also seek to analyze the interactions of human language with the evolution of algorithms and their uses. We conclude that human language is progressively influenced as the goals to be achieved by the algorithms progress and evolve.*

Keywords: *Algorithm, Philosophy of language.*

Resumo: *O objetivo deste artigo é destacar elementos característicos da relação entre o processamento de dados feitos por algoritmos e a linguagem humana. Buscamos ainda analisar as interações da linguagem humana com a evolução dos algoritmos e dos seus usos. Concluímos que a linguagem humana sofre influência progressiva conforme progridem e evoluem os objetivos a serem atingidos pelos algoritmos.*

Palabras-chave: *Algoritmo, Filosofía da linguagem.*

1. Introduction

More and more humans interact with their creations and in recent decades the most prominent human creation are algorithms, and among them the algorithms that use human language as a substrate to be worked on. The functioning of human language structures seems to be influenced and modified by the increasing use of algorithms, especially algorithms that use virtual databases to achieve their goals.

Initially we will define and conceptualize what is an algorithm, its main functions and some of its uses. Then we will list and develop the understanding of some of the relationship between algorithm and language, from this understanding we try to better understand the influences between language and algorithms.

Furthermore, we will try, beyond the multiple influences, to understand how much each one has from the other and if it is possible to separate the parts, whether there are parts at all. We assume that it is important to know what are the goals of the algorithms and whether they are compatible or contrary to the goals of humans.

For conceptual purposes we understand *language as an instrument and a convention for communication.*

2. Algorithm

2.1. What is an Algorithm?

Every computer program is an algorithm, but not every algorithm is a computer program. As what interests us here in this article are the virtual algorithms –*logical sequence of immaterial data processing that seek to achieve a goal*– we focus on them.

It is noteworthy that algorithms can also be found in physical nature, as well defined by Daniel Dennett “What Darwin discovered was not really an algorithm, but a large class of related algorithms” (Dennett, 1998, 53). The existing algorithms in the material world will not be objects of study of this article.



“The word algorithm (and the idea of studying it) comes from Al-Khowarazmi, a 9th century Persian mathematician” (Russell & Norvi, 2013, 31), but the first algorithm is attributed to Euclid who used it to calculate the greatest common divisor of a number. Algorithms are born from the studies of formal logic, which is one of the great fields of philosophy’s research, and developed as such from the 19th century with the works of mathematician George Boole (1815-1864) and which we now know as Boolean propositional logic.

But it has been in recent decades that studies of algorithms have developed increasingly, especially driven by computer data processing. And here appears one of the characteristics of algorithms, which is to compute something, and computing, or computation, is a concept that is difficult to define, because it can be understood either as counting something, as calculating something, or even as adjust, compare, parallel, include or exclude something from a count. We can compute numbers, symbols, images, sounds and a multitude of other elements. But the computation that interests us here is the computation of language, especially the written language in words.

2.2. Algorithms and Words

Algorithms have goals to be fulfilled and their existence is due to this: to achieve a goal. The substrate –in our case the words– and the process exist only to reach an end. In other words, the end is the purpose of an algorithm. For an algorithm, the computation or the way it handles words is directly related to the end it wants to achieve, which can be many. For example, a database search algorithm –such as Google– will treat words differently if its goal is to teach, communicate, or disclose something.

The ability of an algorithm to achieve its goals is increasing depending on the refinement of word processing and the size of the databases it has access to. The probability that an algorithm will achieve its purpose increases if the processes it uses to handle words are well designed and if it has access to a large number of words. The quality and complexity of an algorithm influence

the result as long as it has a big amount of information and processing power.

Algorithms are problem solvers and tend to imitate with improvement other algorithms that have the same or similar goals. “And it is through the ‘imitation’ of these aspects that computational algorithms learn and specialize in solving specific problems” (Florão, 2017, 52). Algorithms are classified primarily by the purpose they seek to achieve, and algorithms that use words as a substrate may have various functions to fulfill, such as relation, search, translation, selection, deletion, misrepresentation, exchange, word disclosure, and so on, and they use other words to achieve what they are meant to achieve.

An algorithm can also be a word patterner. “An algorithm is designed to extract and use statistical patterns in data” (Meira, 2017, 57). Thus, an algorithm can also standardize text by whether exchanging words for words of similar meaning or not, and it can do so depending on the purpose for which it was created. This, of course, can change the meaning of the texts and the original purpose for which they were created.

According to Squirra, algorithms can also use words to create new words or texts as well as to develop facts triggered by those words. “Algorithms are used in journalism to mine large volumes of data for the purpose of predicting certain facts” (Squirra, 2018, 139). Because of their large data processing capabilities, algorithms can scour large databases to perceive words in contexts that humans could not. Algorithms relating words that relate to other words in certain virtual environments can draw conclusions about facts that we could not draw. Algorithms developed for journalistic purposes are a good example for this algorithms’ capability. “Algorithms are currently being used by journalists both to filter content from the internet and produce news as well as to distribute it in a personalized way, allowing consumers to get more than they like and less than they don’t like” (Magalhães, 2017, 244).

A final feature of the relationship of algorithms to words is that they can also create new texts. According to Carreira “The algorithms that generate automated texts are the result of several technological advances, such as artificial intelligence and the generation of natural language.”

According to this author, the creation of texts is already relatively common among journalistic institutions.

One of the problems with this way of using algorithms is “algorithmic cannibalization”, that is, one algorithm uses the text of another algorithm to produce text that will be used by another algorithm and so on. The outcome of this process is still poorly known or studied, but may lead to poor language or standardization of language or texts with little originality, poetry, innovation, literature or dullness. At least until whenever an algorithm is created to improve texts with these characteristics.

3. Algorithm and language

3.1. Relationship

In the history of language development there was a time when language was the main relationship with itself, language created new languages, and in the relationship between different languages new ones developed, this era passed away. Today languages have between them (between it and itself) several relation elements that interpose in linguistic links, one of these elements is the algorithm.

For a long time the goal of language was to communicate something, to conceptualize, indicate, relate, mean and express it. Today, especially in the digital world, language is increasingly being used as a means to achieve the goals of algorithms. Word manipulation is done in order to achieve what the algorithm proposes to achieve. And what the algorithm proposes to achieve is something that is increasingly difficult to know because in addition to algorithm creators keeping their purpose secret, many algorithms create other algorithms or modify themselves to improve their performance and increase their chances of achieving their goals.

An algorithm modifies itself by improving its performance by eliminating paths that did not reach its objectives, searching for new paths, seeking relations of equality, discrepancy or similarity between the words and contexts in

which they are used. As Rodrigues highlighted (2017, 99) “computers can analyze large volumes of data and extract knowledge from them. They use resources like natural language and computer vision to identify patterns through machine learning algorithms.” Language has thus become more than a means for the algorithm to achieve its ends, language is a means for the algorithm to improve itself and to develop as such. Algorithms grow using and relating languages. Words have become the food for digital algorithms.

Algorithms already have a rudimentary form of learning. Many of them are created for more than a particular purpose: they are created with the ability to use human language to nourish and develop, it is almost a biological relationship, only virtual. In the words of Rodrigues (2017, 100) “artificial intelligence is used to perform tasks previously performed by humans; linguistics adds understanding of language and information extraction in texts”.

3.2. Utilization

Words have become for algorithms what they are for humans: an instrument and a convention for communication. Rodrigues (2017, 101) highlights another capability of algorithms, “variability: refers to data whose meaning is constantly changing, depending on the time and context a word may have other meanings”. Algorithms are making a rudimentary form of interpretation, and they use human language as an instrument to interpret contexts and better achieve their ends.

In addition to this form of interpretation, the algorithms are communicating with each other to better achieve their goals: “The Internet of Things (IoT) allows multiple devices to ‘talk’ to each other over the Internet or a private network” (Rodrigues, 2017, 101). But in this process of communication between algorithms the human language is not the main language used, in this communication the algorithms use “machine language”, which is a human creation.

The algorithms use “machine language” to work with human language. Algorithms designed to work with human language use non-human language.

3.3 Modification

In the introduction we defined language as an instrument and a convention for communication. We chose this definition because we believe that algorithms have changed the three concepts by which we define language: *Instrument*; *Convention*; *Communication*.

3.3.1 Instrument

An instrument is a resource we use to achieve some result; thus, language is one of the tools humans use to communicate, and algorithms have changed the resources we use to communicate. This change has influences not only in language, but in various spheres of the human being in society, as it is well asked by Pastor (2019, 285): “In the digital networks, Big Data, Internet of Things, computational mediations, digital data, algorithms, smart cities, devices, etc., how to think –in sociological, philosophical, anthropological, communication-terms– the social environment, social dynamics and the practice of the social sciences themselves?”

Algorithms have changed language’s diffusion logic, but they have also changed the speed of transmission, and in many cases, reshaping the logic and speed of language transform the language itself. For example: More and more digital communication is done by images and less by texts.

3.3.2. Convention

The term convention here refers to a set of rules that are created by the relationship and communication between individuals, language is therefore a set of rules that have been developed throughout the evolutionary process. We understand language as the result of the evolutionary process, but also as a driver of this same evolutionary process. We believe that any theory about language that does not take this into account tends to be poor in its explanation.

Language seen as a social convention took thousands of years to evolve, and the more it

evolved, the more it drove human evolution, as it was one of the primary tools for transmitting knowledge. For thousands of years oral language has been the primary means of conveying knowledge, and only four thousand years ago writing has slowly begun to take place in part of this transmission, and less than four decades ago algorithms have changed much of the communication rules.

In the digital world the rules of communication are different, the language is different and the language goals are different. This differentiation was only possible through the use of algorithms. For example: The linguistic rules of communication in social networks are distinct from the rules of oral or written language in a book.

3.3.3. Communication

Because we are especially concerned with written language in this article, we will define communication simply as the process of transmitting knowledge. But the process is not neutral, it interrelates, that is, influences and is influenced by the knowledge to be transmitted. The “sender-message-receiver” theory no longer has space in the digital world of algorithms.

The algorithms allowed a hybridism between language, process and transmitted knowledge. For example: A word placed on a digital social network tends to have a different and often dubious meaning, and when what is communicated by that word is perceived by tracking algorithms, algorithms tend to change the relationships and environment of existing words with future words. For example: an algorithm of an instant messaging application such as WhatsApp tends to suggest new words based on the written word database and to change new word suggestions depending on the written word. Another example: Google autocomplete is an algorithm based on all searches made by Google, and the choices we make in our searches will influence new searches.

Pastor defines this process as “complex of interrelated entities” and in it there is no longer “a radical separation between nature and culture, human and non-human” (Pastor, 2019, 271).

Another concept used by this thinker, which helps us to understand this process, is the author-network, where the individual is already a crowd and every action of this individual produces difference, either by imitation or modification of communicative language. “What matters is the microrelationships, the multiple diffuse and infinitesimal relationships that occur between or within individuals.” And yet “the small repetitions, oppositions and adaptations, or their sociological correlatives, the imitations, hesitations, and inventions that constitute subrepresentative matter and, as such, they refer not to individuals, but [...] to flows and waves of beliefs and desires” (Pastor, 2019, 273).

In large part, algorithms, using written language, will influence and modify the communication of this flow of beliefs and desires. For example, the algorithms of social networks tend to increasingly propagate the terms that are most used by users of this social network, influencing the beliefs and desires of these users.

In the digital world driven by algorithms “things are not simply “are” (static things), they “are” at this moment” (Pastor, 2019, 277). And “being at this moment” influences being and communicating not only in the digital world, but in the “real world,” whatever it is. Trying to exemplify: When someone is on a social network, he or she is there for his/her beliefs and desires that may have already been influenced by algorithms whose purpose was to get people there. But when that person is on the social network or messaging application the algorithms know he or she is there, and whatever they do in these virtual environments will influence the algorithms that can influence other algorithms in a network that includes other people as well that are in this virtual world. The algorithms’ relationship to words occurs in a similar way.

And where does the “real world” stand? For the virtual world probably in the Global Positioning System (GPS). Using GPS the algorithms know from where most of the machines are connected to them, and also know the language used in these places and will know which words can be used to achieve their intended purposes.

Algorithms influence the words of the virtual world and the “real world” of which the virtual world is increasingly part.

4. Concluding Notes

Realizing that human language is progressively influenced by algorithms that work with words is not difficult, the hard thing to know is how much one influences and is influenced by the other. Probably, this process feeds back constantly and finds a balance by eliminating the most glaring discrepancies.

Eliminating discrepancies can also be the major negative effect of using algorithms in language, as with “algorithmic cannibalization” language tends to be standardized so that it becomes increasingly accessible to more people. It tends to become average, mediocre. Poetry and literature tend to have less and less effect on the meaningful world of people, and widely divergent opinions tend to be planned.

Probably the algorithms are also controlled by finding a common point where the action of each one is the most advantageous for all. We think that another limiting and controlling factor is language itself, because words and their meanings and relationships are finite and finite are also the goals pursued by each algorithm.

We started this study by seeking to know how much of the goals of the algorithms are also the same goals of humans. We ended up by believing that technology and algorithms have become an ecosystem of which the human is also part. We can no longer separate actions from algorithms from human actions, we are one movement, to where the humans go the algorithms tend to follow their course, and the opposite also tends to be true. We have the algorithms in a dance without rhythm, but with many movements.

From our studies we have also concluded that such a text will probably be able to be written by algorithms in a short time, if they already cannot do it right now.

5. References

- Carreira, K. (n.d.). *Robôs executam as tarefas mais elementares do jornalista*. Disponível em: <https://www.academia.edu/29738887/JORNALISMO_AUTOMATIZADO_-_perspectivas_de_re_defini%C3%A7%C3%A3o_de_fun%C3%A7%C3%B5es_entre_algoritmos_e_jornalistas?fs=rwc> Acesso: 15 de jun. 2019.
- Carreira, K.; Squirra, S. (2018). As inovações tecnológicas levam o jornalismo aos agentes autônomos artificiais. *Revista Animus*. UFSM. Santa Maria. V.17 n.33.
- Dennett, D. (1998). *A perigosa ideia de Darwin*. Rio de Janeiro: Rocco.
- Florão, M. (2017). Computação cognitiva e a humanização das máquinas. *Revista Fonte - prodemge*. Ano 14 - N° 17.
- Magalhães, D. (2017). Precisão, rapidez e robôs: um panorama atual do Jornalismo algorítmico. *Revista Temática*. Ano XIII, n. 08. agosto/2017. NAMID/UFPB.
- Meira, W. (2017). *Computação cognitiva e a humanização das máquinas*. *Revista Fonte - prodemge*. Ano 14 - N° 17.
- Pastor, L. (2019). Das mônadas às redes: o resgate de um social associativo para a Sociologia Digital. *Plural, Revista do Programa de PósGraduação em Sociologia da USP*, São Paulo, v.26.1.
- Rodrigues, B. (2017). Computação cognitiva e a humanização das máquinas. *Revista Fonte - prodemge*. Ano 14 - N° 17.
- Russell, S.; Norvi, P. (2013). *Inteligência Artificial*. Rio de Janeiro: Campus.

Arildo Marconatto (arildomarconatto@hotmail.com). PhD Student in Philosophy - Unisinos - SEDUC - MT.

Edson Sestari (edson.sestari@facgmt.edu.br). PhD Student in Philosophy - Unisinos - FACCMT.

Received: 6 June 2019
Approved: 20 June 2019