THE INKA QUIPU ENIGMA

OLIVER KNILL

1. Introduction

1.1. The history, mathematics and database technology of the **quipu** [1], the “talking knots” of the Inka empire is a fascinating subject. Quipu are an original approach to number systems, database structures. Unlike marks on bones, tally sticks or clay tablets, ink on wood, papyrus or paper, it is a topological encoding, similarly in nature than genetic code is woven from protein knots. The first scientific study of quipus began by **L. Leland Locke**. His important article [11] documented in a very clear way how knots were used for recording numbers. In his introduction, Locke also pointed out that also in other parts of the world, like China, knot records have preceded the knowledge of writing.

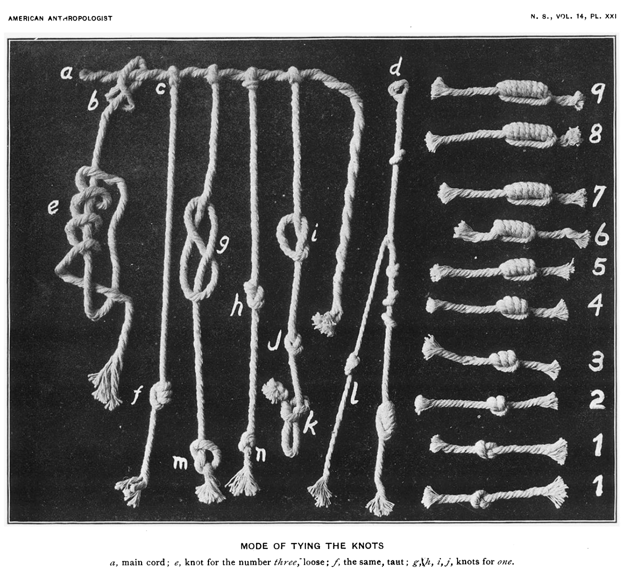

Figure 1. A page from [11].

Date: 10/22/2018, updated 1/31/2022 for Math E 320.

[1] Quipu and Khipu are equivalent spelling variations

1.2. Recent research pointed to **Rosetta stone break-through discoveries** leading to publications for a general audience like [6, 14, 5, 20]. In 2018, when this document started, there there was also quipu exhibit at the Boston museum of fine arts. Naturally, these popularizations or reports hide the work which is needed to investigate the topic. There is the field work of digging out, cleaning, reading and then cataloging the information, then to place the data into the context of the history, linguistic, and culture of the time and finally to translate interpret and cross referencing the data. In the quest to decode the quipu cypher, there has been spectacular progress for post-colonial quipus [7, 13] and progress in better understanding non-numerical pre-conquest quipus [4].

2. Knots, Links and Graphs

2.1. Strictly speaking, for a mathematician, a quipu is neither a **knot** (a closed loop in space) nor a **link**, a collection of non-intersecting knots in space. But they are links in a generalized sense in that they would be links if the ends of the individual ropes were connected. It is not so much the topology of quipu which is of interest for researchers but the information content which is encoded topologically. Because only three different type of knots appear in Inka style data (simple knots, figure eight knots and long knots) (whose topology is well understood), these entities could be replaced by symbols like $L4, S, E$ standing for a long knot with 4 turns, a single knot or a figure eight knot. A quipu can be described as a graph on which scalar and vector data are attached. The scalar data assign to a node the knot type or the attachment type, if the node is branching off there. The vector data which describe the connecting strings are determined by ply and spin direction, attachment type, color and the material of the knot. For a computer scientist a quipu is an example of a **graph database**.

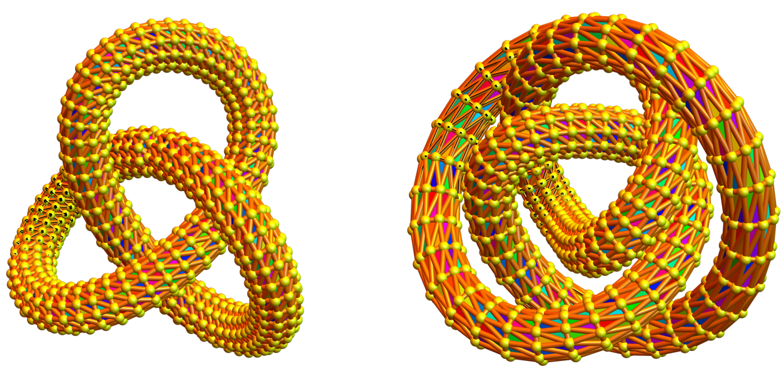

Figure 2. The trefoil knot and a figure eight knot.

3. Crypto riddles

3.1. The **Inca code** is a cypher which has not yet cracked. While the numerical data encoding are quite well understood, the problem lies in understanding non-numerical signs. Crypto riddles have always attracted the interest of the general public. Examples are the **Maya code**, the **Egyptian hieroglyphs**, the **German Enigma** during the second world war or the fictional alien pictorial language which is at the center of the movie "**Arrival**". An other example of an outstanding riddle is the **"Antikythera" instrument** which is believed to be an early **analogue computer** used for astronomical computations. More riddles have presented themselves when trying to decode texts from **Palimpsests**, texts which are hidden beneath other texts. An example is the Archimedes Palimpsest. In those cases, reading the text requires first to reveal the structure as the text had been erased and written over.

Figure 3. The Anticitera, the Maya and the Rosetta stone.

4. Rosetta stone moments

4.1. Also in the case of the German enigma code which was cracked at Bletchley Park, the problem was not entirely mathematical. One had to wait for Rosetta stone moments, clues like knowledge of the **weather code**, or rely on **planted information** which allowed the cryptographers to attack the code using **crib-based decryption** techniques. In the case of the quipu, the task is harder because there are no known cribs (at least from the pre-colonial time) and many of these documents were destroyed in the wake of the colonial conquest and because quipu from regions with high precipitation deteriorated rapidly if not preserved as they are made of organic material like wool.

4.2. There are less than 1000 quipu known today. The **Berlin collection** contains about a third. The decoding problem has linguistic, historical and anthropological context. Understanding the content of a coded text or new language needs "Rosetta stone moments" like in the case of the hieroglyphs, where Champoleon and Young have been able to crack the code of the hieroglyphs. The quipu form a cryptological riddle in which plain text information is missing. Since the information is believed to be non-phonetic, the problem is harder than in the case of hieroglyphs, the cuneiforms or the Maya code.

FIGURE 4. Some quipu researchers: Leland Locke, Marcia and Robert Ascher, Sabine Hyland.

5. Algebra with strings attached

5.1. For a mathematician, quipu can be fascinating in various ways. One knows already quite a bit about the numerical aspects [2]. But mathematics can be understood as a more general concept, not only as the science of numbers, or the quest to understand algebraic or geometric objects but more generally as a **science of structure**. For a mathematician, a **language** is a mathematical structure, usually a subset of a monoid of words, in which a grammar and axiom systems define what is meaningful in this language. [10].

5.2. In formal language theory, a language is a set of strings over some finite alphabet A. There is an operation on the set of string, which is concatenation. This is associative. Together with the zero element, the empty string, one has a **monoid**. A linear order on A defines a lexicographical ordering on the language. If we look at language encoded on a quipu however, then the monoid structure is gone. There are algebraic operations on graphs, like disjoint union or joins (which both can serve as additions) or product operations which complement them rendering the category of finite simple graphs into rings, but these structures have no meaning in language. The addition of strings to a quipu needs more information as strings can be attached in different ways.

5.3. Communicating with knots is a completely different approach to writing. The sentences are not elements in a monoid because there is a spacial **nonlinear approach**. One can encode a quipu as a weighted graph, where the nodes are the knots, which are labeled by the value of the knot, the spin or attachment direction.

The edges can be equipped with color, ply direction and hierarchy data too. Numbers are encoded using three different type of knots, but they can also be arranged in different ways leading to more information content than anticipated.

6. Nonlinear languages

6.1. But having content written down in a linear narrative way is not unique. This has also be developed by other cultures. We use pictures for example to represent mathematical statements, we use tables represent data, we use graphs to represent relations, mind maps are examples of graph information containers which are non-linear. In our time of electronic documents, we can add a parameter "detail" to a mathematical text. Varying the detail level allows then to zoom in and out in the knowledge landscape, similarly as we do when we look at a map of the earth. A map is a highly non-linear representation of data.

Figure 5. A **mind map** as an example of a non-linear language.

7. Popular culture

7.1. Intersections of linguistic and mathematics appear frequently in **pop culture**. The reason is that mathematics related to linguistics is more approachable than mathematics related to algebraic structures. We can mention the novels of Dan Brown, in which a Harvard symbiologist Robert Langdon is the hero. The field of

Symbology does not exist. We should also mention the movie “Contact”, in which an alien language is broadcast from an other planetary system to us. The decoding of the language needed spacial insight as it was a three-dimensional document. Also remarkable is the movie “Arrival” in which a linguist and physicist work together to get access to a strange smoke ring based language spoken by two aliens “Abbot and Castello”.

FIGURE 6. Linguistic in pop culture: Arrival and the Dan Brown story, Inferno and Contact.

8. Topological writing

8.1. First of all, the mathematical approach of the quipu is unique. It is a three dimensional writing, dealing with topological objects known as knots and links which are of interest to mathematicians, and physicists. For a computer scientist it is a **graph database**. Using spacial, material and color information, the Inkas have placed information onto the strings. An introduction about this fascinating topic is [20].

9. Seeing real samples

9.1. The museum of fine arts in Boston currently has an exhibit showing off some of the quipus from the Peabody museum at Harvard. While quipu are always mentioned in the context of the origins of number systems, there had been much progress recently in understanding more about these Inka code. This and some art installation must have prompted the exhibit at the museum.

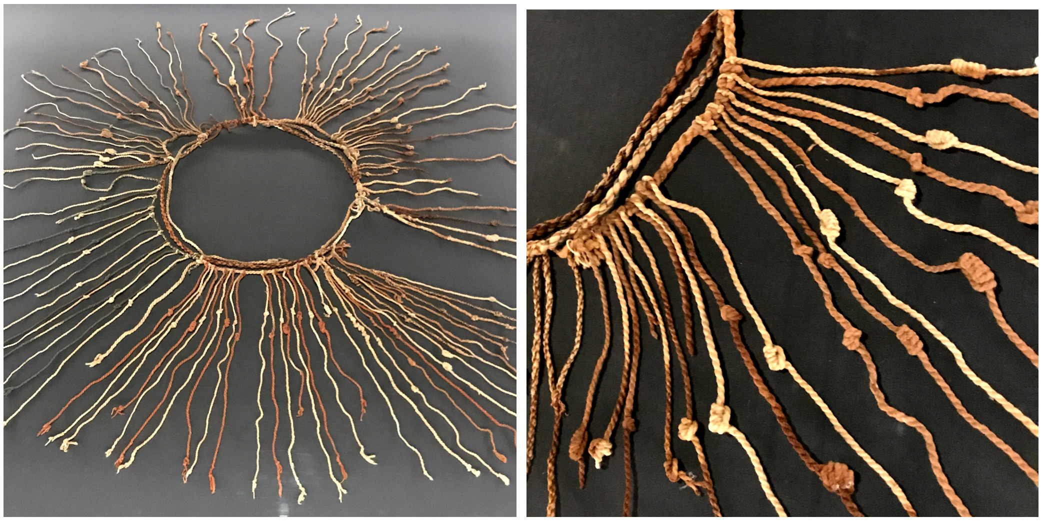

FIGURE 7. Photos of quipus from the exhibit at the Boston museum of fine arts.

10. Mathematical notation

10.1. The quipu research sheds light onto the origins of mathematical notation, the origins of number systems and even the philosophy of mathematics [17]. Talking knots are a highly original approach to language and naturally are of extreme interest in linguistics, semiotics, sociology and anthropology. There are also relations to **education** because the way human cultures develop mathematics is similar to how students acquire it. Recent battles about notation syntax (the PEMDAS wars) illustrate how "antropological" mathematical notation is: computers and humans empirically disagree with reading mathematical content like $6/2(1+2)$. Most humans get 1 while computers get 9. The syntax laws are ambiguous [3]. The PEMDAS wars are silly because it is a battle in a realm where no consensus has been built by authority. It is a heated battle because there are some, who religiously defend their own interpretation of mathematical syntax.

11. Management

11.1. An efficient **record keeping system** was necessary for building the Inca empire. Engineering projects involved calculations, recording of data, calculating ratios and proportions. As the largest empire in the pre-Columbian new world between Ecuador and Columbia, it stretched 5000 km along the Andes. The **Tawantinsuyu**, "the four parts divided together" or "land of the four quarters" as it was called covered a complex environment reaching deserts, the Andes or the Amazon.

11.2. The quipu story leads to insight in management and organization theory [20]. The Inka empire which lasted only for a short time (1400-1532) was able to develop and run effectively because of technology. (Of course also using military force but it appears that the power of organization can complement military conquests. This has been proven to be true even up to very recent times. Failed military adventures failed at providing management.) The quipu technology was essential for management and administration of such a complex structure. In some sense it must have enabled progress similarly as the **modern internet** now does: the Inka road network [19, 1] compares to the internet backbone and the quipu are the files.

11.3. To conclude, the story of the Inkas is an **allegory** for our time: investement in infrastructure, in language, in organization can be as powerful as military power. It is modern because in our time, power is also more and more established by entities which know how to gather process and understand information.

12. Genetic code

12.1. Interestingly, the Inkas stored information similarly than our **genetic code** is stored, on knots: our DNA consists of strands of twisted **DNA molecules** while quipus use twisted rods. It appears that modern bioinformatics is getting inspired by quipus [16]. Color encodings of electronic parts like resistors encode numbers in colors.

12.2. Mind-boggling is also that the Inkas used binary encoding [18] using spinning, plying and knot directionality and the markednes theory in linguistic. Binary steps were also done in Chinese hexagrams where the binary encoding had six bits ($2^6 = 64$).

12.3. Binary symmetries appear also in biology and modern physics, where chirality and parity are important. In physicist it is the weak force which shows an asymmetry, in biology it is the orientation of the DNA, which is dominant. There is also a left handed Z-DNA. In the Urton terminology of markedness, this would be the marked version.

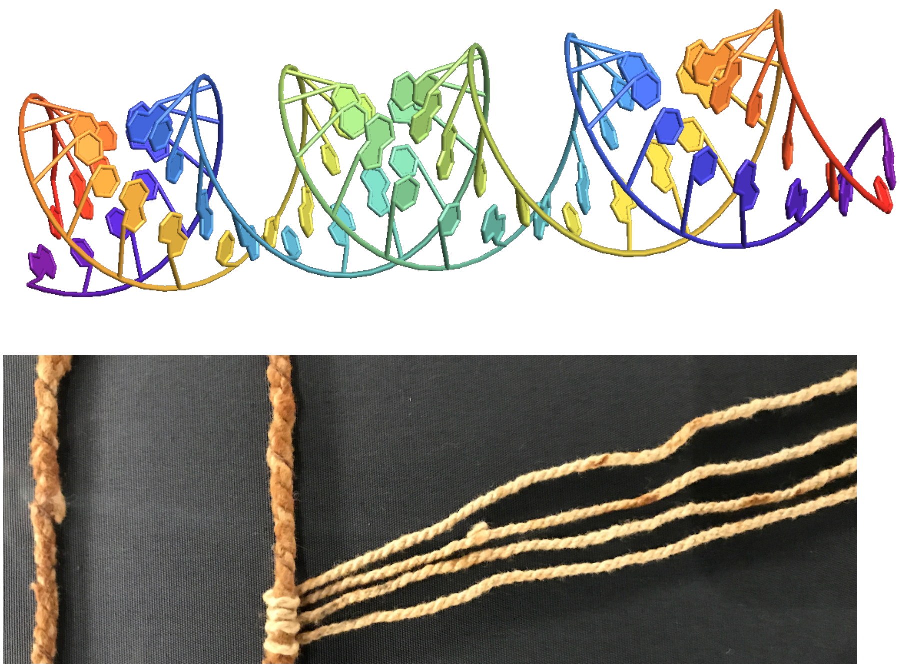

FIGURE 8. DNA and Quipu both have orientations.

13. Graph Database

13.1. Organization through inka decimal administration, required time accounting, census data to be organized on quipu by **quipukamayuq** (professional quipu writers), who had an **information technology** (IT) structure with a consolidated database in **Cuzco**, a prerunner of the "cloud". The cloud is just a modern word for a decentralized "main frame". If one would compare the quipu with files of modern computing, the analogue of the internet was realized by **Chaskis**, the quipu runners.

13.2. In our time of information technology, we deal a lot with three dimensional physical space: we have **augmented reality**, **computer vision**, **3D scanning** or **3D printing** technologies. Usual writing is two dimensional. The quipu system therefore appears very modern. Our own genetic code is encoded on knotted devices, the DNA, we use **graph based** databases, like Neo4J. Graph databases are an alternative to **relational databases**. They appear to be superior if the data structures are complicated.

14. The Unix philosophy

14.1. But one does not have to go far. One of the most popular databases used is the **Unix file system**, which organizes information in a **tree**. This technology allows comfortably to work with a half a dozen tera bytes of data at the finger tips and split different things into different tree branches (**directories**). This smaller quipu project (which occupied me over a few weeks) is an independent tree in my Unix database. Like every course, every website I maintain, my library with thousands of electronic documents, programming parts etc, they are all comfortably separated and organized like on a quipu.

14.2. One of the most important insights could be that **Like the Unix file system, the quipu database system is a paradigm**. It is a **gem in simplicity** and efficiency and very close to the **UNIX idea**. This principle was adopted also for our AI experiment of 2003 [9], where the AI bot was just a Unix file system and the intelligent agent just parses a sentence the travels the file system to do things as the nodes of the file system can be programs or little scripts which look things up. One advantage of this **quipu way** is that it is highly scalable. The industry uses it even with peta bytes of data while any conventional data base would get challenged. Extending a quipu data base is very easy, just add an other strang of nodes or produce more subsidary nodes. Similarly, a Unix file system can virtually have unbounded capacity.

15. The problem of backup

15.1. The only limitation in scale is the size of the harddrives. I personally currently have my files on 5TBytes external drives which are then stored in a frozen and of course encrypted form also in different locations (as the Inkas did). There are currently about 8million files there. They can be tiny text or program fragments, or larger documents like books, pictures or movies. But I would not store this in the cloud as one can also learn from history. The most obvious one is that services and companies die or change their focus, cutting off things which are no more profitable. Companies are no charity. The Inkas saved things in the "cloud" which was then their main capital "Cusco". And we all know what happened when the Spaniards invaded the place. Many major databases were destroyed and less than 1000 quipus survived.

15.2. My own data would even survive if Boston would be annihilated by a nuclear catastrophe (the analogue of a colonialization disaster) or all cloud services would have bit the dust. [One can easily imagine scenarios in which they could disappear in the near future. Examples are CPU leaking concerns, lawsuits due to copy rights or then that companies running the business will simply die or forced by some rogue government to make things accessible.] In the past, the surviving quipu were stored and backed up in hidden decentralized places. Unfortunately for us, we can not read most of the non-numerical data. The Inkas somehow used to encrypt things (even so this had not been the main intention it had the advantage of some privacy as the quipus contained what we would call today bank information or services owned). Also this is a lesson: never store information in a form which is not accessible by simple tools for which public domain or at least open source tools exist to read it.

15.3. My own small quipu project is a small branch in a bigger Unix tree of my work stations (which are synced regularly). My "quipu pendant string" contains only 500 MBytes of data currently but it includes scanned books, documents, the Harvard khiup database, articles and pictures as well as texts about quipu. If in future, more things should appear, I would add it as "subsidiaries" as the Inkas did when adding more information to a primary cord. I have absolutely no problem to find things like that as it is in of of the 3 major project branches in my Unix file hierarchy. It is nice to see that this simple but efficient storage paradigm is actually Inka technology.

16. Reversed Polish Notation

16.1. An important feature of the Quechua language is **agglutination**, which allows that operators often can be found at the end. Like "ni=I" appears at the end of **Runasimi-ta yacha-ku sa-ni**. (People language, learn, now, I) or **Oliver, Wasi-Ta ruwan** which translates as Oliver house builds. [8]. Despite that linguists call Quechua a SOV language (Subject, Object, Verb), the aggultinative part makes it possible to put a subject suffix and have the subject at the end.

16.2. This reminds of the reverse polish notation RPN (still used in stack oriented programming languages like **Poscript** or **Bibtex**). One sees also reverse order in numbers like Quepchua: 13, "ten, possessor of three", while we say "thirteen". In a **stack oriented language**, you say $23x$ **rather than** $2x3 =$. We don't need the equal sign. Operators come to the end, which is more efficient and does not need equal signs. So, it appears that at least for addition and multiplication, no computing device is needed. And unlike for pebbles (bad for transportation) and tally sticks (we can not subtract), the computation with knots can do that.

16.3. The advantage of RPN is also that no brackets are needed. We use the RPN often when doing quick computations. For example, to compute the sum of the squares of the square roots of the first 100 primes, one can use the RPN notation in Mathematica: **Range[100] // Prime //Sqrt //N // Total** which has the advantage that I see in each step what has been computed. The traditional (written way) is to write: **Total[N[Sqrt[Prime[Range[100]]]]]** which gives just the end result but requires to write a nested sequence of brackets.

FIGURE 9. Non-RPN and RPN calculators.

17. SEMIOTICS

17.1. The Swiss **Ferdinand de Saussure** (1857-1913) was a pioneer in linguistic and semiotics. Saussure was eclipsed vastly both in scope and originality by his contemporary **Charles Sanders Peirce** (1939-1914) who only later would be recognized as one of the greatest thinkers and philosophers of his time. Frank Salomon suggests the quipus reference system to be a general purpose semasiography [15]. Semasiographic signs were present in multiple Andean systems.

17.2. Highly successful and persistent non-phonetic scripts are not only used in math notation (figures or combinatorial diagrams like Dynkin or Ferrers diagrams or commutative diagrams) but also in physics (Feynman diagrams for example), they also are common in music notation, programming flow charts, chemical formulas, choreographic notation, and knitting and weaving codes.

17.3. Notation is important in mathematics and it is linked to mathematics itself: Barry Mazur was cited in "Enlightening Symbols" [12] that *A seemingly modest change of notation may suggest a radical shift in viewpoint. Any new notation may ask new questions.* This also applies to the quipu language. It is a completely new angle to the origin of mathematical language and illustrates the richness and diversity with which the art of expressing mathematical thought has begun.

Figure 10. Two pioneers in linguistics: Ferdinand de Saussure (1857-1913) and Charles Sanders Peirce (1839-1914).

References

[1] M. Anderson. 5 reasons the inka road is one of the greatest achievements in engineering. *In Anthropology, History and Culture*, 20, 2015.

[2] M. Ascher and R. Ascher. *Mathematics of the Incas: Code of the Quipu.* Dover Publications, 1981.

[3] F. Cajori. *A history of Mathematical Notations.* The Open Court Company, London, 1928.

[4] J. Clindaniel. Toward a grammar of the inka khipu: Investigating the production of non-numerical signs. Harvard dissertation, department of Anthropology, 2018.

[5] A. Shapiro (Host). Harvard student cracks incan code. https://www.npr.org/2017/12/28/574314933/harvard-student-cracks-incan-code.

[6] S. Hyland. Unraveling an ancient code written in strings. *Scientific American, Sapiens*, November 11 2017, 2017.

[7] S. Hyland. Writing with twisted cords: The inscriptive capacity of Andean khipus. *Current Anthropology*, 58(3):412–419, 2017.

[8] P. Jorgensen. Quechua - the living language of the incas. https://www.youtube.com/watch?v=KlXj28dXPAU, 2017.

[9] O. Knill, J. Carlsson, A. Chi, and M. Lezama. An artificial intelligence experiment in college math education. http://www.math.harvard.edu/knill/preprints/sofia.pdf, 2003.

[10] M. Kracht. *The Mathematics of Language*, volume 63 of *Studies in Generative Grammar.* Mouton De Gruyter, 2003.

[11] L.L. Locke. The ancient quipu, a peruvian knot record. *American Anthropologist*, 14:325–332, 1912.

[12] J. Mazur. *Enlightening Symbols, A short history of Mathematical notation and its hidden powers.* Princeton University Press, 2014.

[13] M. Medrano and G. Urton. Toward the decipherment of a set of mid-colonial khipus from the santa valley, coastal peru. *Ethnohistory*, 65:1–23, 2018.

[14] J. Radsken. Undergrad deciphers meaning of knots, giving native south american people a chance to speak. *Harvard Gazette*, August 25, 2017.

[15] F. Salomon. *The Cord Keepers, Khipus and Cultural Life in a Peruvian Village*. Duke University Press, 2004.
[16] A. Stasiak. Much like the khipu system, dna knots contain precious information. https://www.sib.swiss.
[17] G. Urton. *The Social Life of Numbers*. University of Texas Press, 1997.
[18] G. Urton. *Signs of the Inka Khipu*. University of Texas Press, Austin, 2003.
[19] G. Urton. Engineering a world with strings attached. Smithonian Institute Symposium, 2013.
[20] G. Urton. *Inka History in Knots*. University of Texas Press, 2017.