

Computational Infra-structure and Machine Set-up

Gepoliano Chaves, Ph. D.

March 7th, 2024

Overview

After completing today's session, students should:

- ▶ Be able to define computational algorithms via the Arab influence in Europe
- ▶ Define bioinformatics
- ▶ Define pseudo-code
- ▶ Practice global Needleman Wunsch algorithm in their notebook
- ▶ BLAST a sequence
- ▶ Download R
- ▶ Think about downloading and aligning sequences from the State of Bahia

Software Development: Creating Algorithms and Pipelines

Algarismos, algorithms and the abacus

- ▶ According to the Wikipedia, the word *algarismo*, in Portuguese, originated “from the name Al-Khwārizmī (c. 780–850)”.
- ▶ Although Portuguese and Spanish are very similar, apparently this is a word that only exists in Portuguese and can be translated into English as digit and into Spanish as *dígito*.



Figure 1: Al-Khwārizmī: a mathematician whose name originated the word algorithm in English and *algarismo* in Portuguese.

Algarismos, algorithms and the abacus

- ▶ *Algarismo*, in Portuguese
- ▶ *Algoritmo*, in Portuguese; Algorithm, in English
- ▶ *Dígito* in Portuguese and Spanish; Digit in English



Figure 2: Al-Khwārizmī: a mathematician whose name originated the word algorithm in English and *algarismo* in Portuguese.

Algarismos, algorithms and the abacus

- ▶ *Algarismo* (Portuguese) and algorithm (English) are words that originated from the Arab language and was brought to Iberia by Muslims during their ruling in the European territories of Portugal and Spain.
- ▶ Therefore, historically by defending their faith, Spaniards and Portuguese contributed to the definition of the European identity (Wikipedia (n.d.)).



Figure 3: Islamic expansion in Europe between the 7th and 10th centuries and the construction of an European cultural identity.

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

The muslim **invasion** of the iberian peninsula



Figure 4: Christian kingdoms in the iberian peninsula describe the conflict with the Muslims as “invasion”. In 711, Muslims coming from Northern Africa cross Gibraltar and defeat disunited Christians in the battle of Guadelete.

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ Formation of the European states of Portugal and Spain
 - ▶ Portugal is the very first European state to come into existence
- ▶ Formation of the Kingdom (State) of Portugal (according to Bezerra (2024))

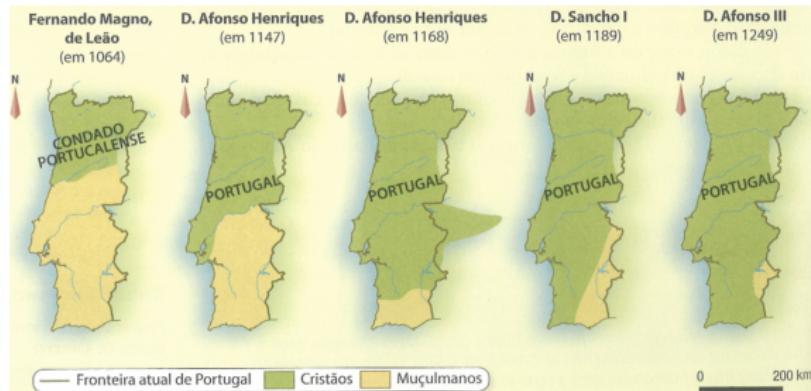


Figure 5: The formation of the Portuguese State in the context of the conflict of the Christian kingdoms in Portugal and Spain agianst the rule of Muslims.

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ Brazil and other Latin American countries inherited cultural and genetic assets from the Iberian colonizers



Figure 6: Condado Portucalense (*Portucalense County*): a territory in the Iberian Peninsula inherited by Dom Afonso Henriques in the 12th century. Dom Afonso Henriques would later become the first King of Portugal.

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ The formation of Portugal has, for obvious reasons, important consequences for Brazil
- ▶ However, after Brazilian independence, we were taught to pay less attention to our ties with Portugal



Figure 7: Afonso I of Portugal: The First king, National Portuguese hero and defendor of the Portuguese identity inherited by Brazil. Statue in Guimarães, Portugal.

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ In Brazil, we inherited the Iberian practice of *Limpeza de sangue* (Portuguese) and *limpieza de sangre* (Spanish) Umaeoutra (2013)
- ▶ So much so, majority of people are nominal Catholics in Brazil

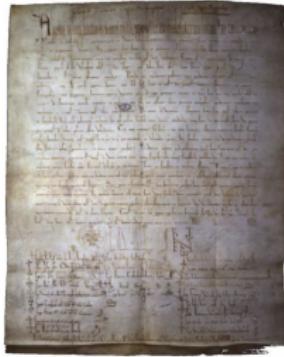


Figure 8: Bula papal Manifestis Probatum gives the title of king to Dom Afonso Henriques in 1179. Throughout history, Portugal and Brazil maintained allegiance to the Pope (Strauss (2014)).

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ At the origin of *limpeza de sangue* in Portugal and Spain, this practice aimed to eliminate the Muslim and Jewish influences in Iberia
 - ▶ Marriages between *Cristão-novo* and *cristão-velho* couple: enforced conversion to Christianity of former Jewish and Muslim individuals
 - ▶ Similar patterns of marriages in the colonies in the Americas
 - ▶ This practice was of ideologic and religious nature, not of biologic or racial genetic purity (eugenics) (Umaeoutra (2013))
- ▶ However, later in Brazil, this practice targeted indigenous people and descendants of African slaves

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ However, because of our past as a colony, we Brazilians disregard our shared history with Portugal

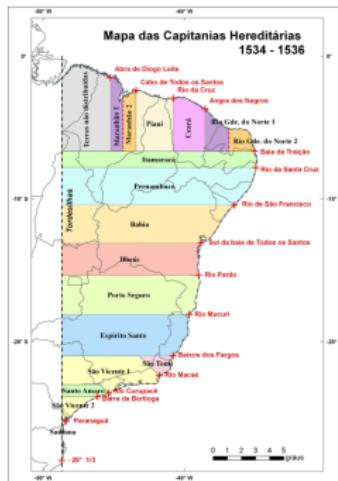


Figure 9: After conquering their territories back from the Muslins in the Iberian Peninsula, Portuguese and Spanish arrive in the New World. Organization proposed to the recently “discovered” Brazilian territory by Portugal in the 16th century. Figure from Cintra (2013)

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

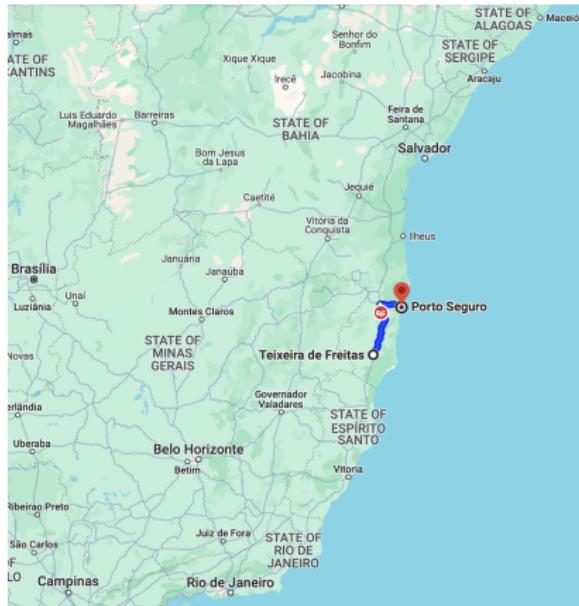


Figure 10: The common history of Portugal, Brazil, the capitaincy of Porto Seguro and Teixeira de Freitas in Southern Bahia can be seen in a map.

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ Dom Afonso Henriques, the Portuguese guardian of the Christian (Catholic) faith
- ▶ After a battle in Ourique in 1139, defeats Muslim leaders and is proclaimed Afonso I, King of Portugal
 - ▶ Less than 400 years later, Portuguese colonization of Brazil would start
- ▶ Brazilian social and genetic concepts are inherited from Portugal and Spain due to colonization

Limpeza de sangue (Portuguese) and *limpieza de sangre* (Spanish): Christianity and the Formation of Iberian States

- ▶ Why do we Brazilians call the arrival of Muslims in the Iberian Peninsula, **invasion** whereas we call the arrival of the Portuguese in Porto Seguro, **discovery**?



Algarismos, algorithms and the abacus

2700-2300 B.C.E.: humans invent a machine to calculate

- ▶ Al-Khwārizmī was a Persian mathematician, astronomer, geographer and scholar in Baghdad, whose name means “the native of Khwarezm”, located in modern-day Uzbekistan”.
- ▶ For historical, ideological and religious reasons, Iberians, seeking allegiance to the Catholic church, aimed to eliminate Muslims from the administration of the Iberian States, using the practice of *limpeza de sangue*
 - ▶ Mixed marriage of a *cristão-novo* and a *cristão-velho*
- ▶ Because of *limpeza de sangue*, do *NOT* see a deep impact of the Muslim influence in Portugal, Spain or Brazil

Algarismos, algorithms and the abacus

2700-2300 B.C.E.: humans invent a machine to calculate

- ▶ In ancient days like those of Al-Khwārizmī, an abacus was a tool used to support operations calculated by mathematicians in the same or similar ways that computers are used to support mathematical operations in the present.
- ▶ Even with *limpeza de sangue*, Portugal and Brazil learned how to count with the Muslims that “invaded” Iberia



Figure 12: An abacus, also known as a counting frame, is a hand-operated tool used since ancient times in mathematical operations.

Algarismos, algorithms and the abacus

2700-2300 B.C.E.: humans invent a machine to calculate

- ▶ To the best of my knowledge, *algarismo* is a word that only exists in Portuguese
- ▶ Due to *limpeza de sangue*, the Muslim religion was excluded from Portugal, but we inherited the word and the mathematics



Figure 13: An abacus, also known as a counting frame, is a hand-operated tool used since ancient times in mathematical operations.

Algarismos, algorithms and the abacus

- ▶ Although along the centuries the tools men use to calculate complex operations have evolved, it remains constant that learning to use these tools, as complex as the tools seem to be, has benefits.
- ▶ For example, mathematical operations using *algarismos* or digits evolved to become complex mathematical algorithms that continue to be consisted of sets of mathematical rules to be followed to obtain a result from the operation of factors.
- ▶ In the sense of the works of mathematicians such as Al-Khwārizmī, algorithms are sets of mathematical rules to be followed to obtain a result.

Algarismos, algorithms and the abacus

- ▶ An abacus is a basic machine used as a counting frame, constructed by humans to take an input, do calculations, and provide the user with an output.

Algorithms are a set of rules to be followed in calculations to obtain the expected outcome from a given input. The word algorithm has the same etymological origin as the word algarismo.

- ▶ For the purpose of this course, an algorithm can be understood as a machine just as a calculator or an abacus

Algarismos, algorithms and the abacus

- ▶ In the context of the present course, we will apply those mathematical or computational rules to analyse biological data using computational biology algorithms.
- ▶ We will also start developing our own computational biology algorithms.

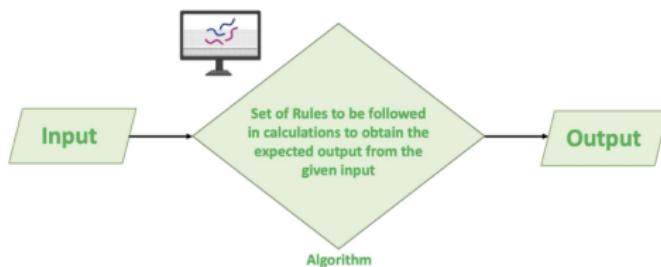


Figure 14: Definition of computational algorithms.

Algarismos, algorithms and the abacus

Conclusions

- ▶ Algorithms are **machines** just like an abacus or a calculator
- ▶ Algorithms can be much more complicated than that
- ▶ Portugal and Brazil learned mathematics from but excluded Muslims by banning their influence in the Iberian States and their colonies

Algarismos, algorithms and the abacus

Three types of algorithms for the programmer

<https://www.youtube.com/watch?v=Uym4-KhP3Lc>

Algarismos, algorithms and the abacus

INTENTIONALLY LEFT BLANK

Combining *algarismos*, algorithms and the abacus: bioinformatics

- ▶ The word algorithm can also be used as synonyme of the words script and pipeline.
 - ▶ By using human language, pseudocode and code, we can develop our own algorithms.
 - ▶ Bioinformatics is a discipline that uses computational algorithms to solve problems in biology.

Figure 15: Bioinformatics uses computer technology to solve problems in biology.

What is Bioinformatics

- ▶ According to the National Human Genome Research Institute
 - ▶ Scientific subdiscipline that involves using computer technology to collect, store, analyze and disseminate biological data and information, such as DNA and amino acid sequences or annotations about those sequences.
 - ▶ Scientists use databases that organize and index such biological information to increase our understanding of health and disease sometimes as part of medical care.



Figure 16: Definition of computational algorithms.

Bioinformatics courses in Brazil

- ▶ Universidade de São Paulo (IME)
 - ▶ Programa Interunidades de Pós-Graduação em Bioinformática
- ▶ LNCC (RJ)
- ▶ Bioinformática/UFMG
- ▶ Pós-Graduação em Bioinformática/UFRN

The language of bioinformatics: Human Language, Pseudocode and Coding Steps

- ▶ Describe in human language what the question is
 - ▶ How does a gene set affect the EFS and OS survival rates?
 - ▶ How do COVID-19 new variants escape the host immune-system become more lethal?
- ▶ Start searching for commands in a programming language, such as R, Bash or Python, that will accomplish a task with the data at hand
- ▶ Write command or script to complete the task

The language of bioinformatics: Human Language, Pseudocode and Coding Steps

- ▶ The human language, pseudocode and final script or pipeline steps ensure repeatable and reproducible experimentation.



Figure 17: Repeatability and reproducibility. Figure from Mackenzie (2021)

References

- Bezerra, Juliana. 2024. *A História e Formação de Portugal*.
<https://www.todamateria.com.br/formacao-de-portugal/>.
- Cintra, Jorge Pimentel. 2013. *Reconstruindo o Mapa Das Capitanias Hereditárias*. <https://www.scielo.br/j/anaismp/a/BmZzYkT6KTRDPBsmTkCzvJr/abstract/?lang=pt>.
- Mackenzie, Ruairi. 2021. *Repeatability Vs. Reproducibility*.
<https://www.technologynetworks.com/informatics/articles/repeatability-vs-reproducibility-317157>.
- Strauss, Arnold. 2014. *Da Reconquista Cristã Ao Reino de Portugal*. <https://www.slideserve.com/arnold/da-reconquista-crist-ao-reino-de-portugal>.
- Umaeoutra. 2013. *ESTATUTO DE LIMPEZA DE SANGUE*.
<https://umaeoutra.wordpress.com/2013/03/16/estatuto-de-limpeza-de-sangue/>.
- Wikipedia. n.d. *Scree Plot*. Accessed March 26, 2024.
[https://en.wikipedia.org/wiki/Scree_plot#:~:text=In%20multivariate%20statistics%2C%20a%20scree,principal%20component%20analysis%20\(PCA\)](https://en.wikipedia.org/wiki/Scree_plot#:~:text=In%20multivariate%20statistics%2C%20a%20scree,principal%20component%20analysis%20(PCA).).