**CAP 242-1: Exercícios de Escrita**

Versão 0.1, 08/03/2021

1) Para cada frase abaixo, marque em que parte de um documento científico ela deveria estar.

(I = Introdução, T = Trabalhos Relacionados, M = Metodologia, R = Resultados e Discussão, C = Conclusões, F = fora do documento – algumas frases podem ter mais de uma resposta)

( ) a. We calibrate our model with empirical evidence from 1985 to 2011.

( ) b. During the simulation we found a bug that forced us to fix the original algorithm.

( ) c. The methodology presented in this work might be applied to other regions with similar dynamics.

( ) d. Regarding the theoretical aspects, the dichotomy between of the pixel origin, strongly related to the measurement, and its application, aggregated in highly semantic concepts needs to be further studied and developed.

( ) e. Even though Morocovsky’s Quadrature is still an open problem, this article presented substantial advances towards a solution to such problem.

( ) f. According to the simulation outputs, the algorithm has sub-linear growth.

( ) g. The main limitation of this work is related to its computational efficiency.

( ) h. Creating the database took one whole year of the doctorade to be concluded.

( ) i. This work is written as follows.

( ) j. Deforestation models have been proposed since early 2000’s.

( ) k. Table 3 summarizes the simulation time of each experiment.

( ) l. Big data infrastructures is an active research field.

( ) m. We could not compile the system using Windows so we decided to use Linux.

( ) n. We expect our work to impact on the building of data cubes of analysis-ready data from satellite imagery.

( ) o. Therefore, investigate how to improve the computational efficiency of such algorithms is an important task.

( ) p. To assess cloud detection algorithms over Amazonia, we chose five areas representative of its climate heterogeneity.

( ) q. The world’s tropical forests are essential places for environmental sustainability and the future of our planet.

( ) r. We hope to improve the land use and cover mapping in the BDC project using SAR data cubes.

( ) s. Snow presents a comparison of three methodologies for label placement.

2) Reorganize as frases a seguir, separando-as em dois parágrafos.

(a) Primeiro parágrafo:

Segundo parágrafo:

1. Most spatial algorithms can be designed to be independent of spatial data structure, relying instead on basic properties that most of them provide.
2. Abstract data type definitions have an externally viewable set of operations and a set of axioms applicable to them [17].
3. The design of the Field data type is based on the ideas of generic programming.
4. To find the mean value of an attribute in a spatial data set, it is irrelevant whether the data structure is a TIN, a grid, or a set of polygons.
5. Generic programming uses abstract data types, which are formal tools that allow an objective evaluation of computer representations [3].
6. Generic programming is well-suited for building GIS [9].
7. The operations are generic, so they work for different data structures and different implementations.

(b) Primeiro parágrafo:

Segundo parágrafo:

1. We now consider some hypothesis that could account for such significant differences.
2. Fmask 4 had an overall accuracy of 90%, followed by Sen2Cor (79%), MAJA (69%) and s2cloudless (52%).
3. As noted by Baetens et al. [26], for satellites without thermal bands cloud detection methods use thresholds.
4. Different thresholds are set for the visible bands, the 1.38 µm band, and the Normalized Difference Snow Index.
5. For this reason, no single study can provide definitive guidance.
6. The results of this study show that the Fmask 4 algorithm consistently performs better than the alternatives for Sentinel-2 images of the Amazon rain forest.
7. Our results are different from those of Baetens et al. [26] who concluded that MAJA and Fmask 4 perform similarly with an overall accuracy around 90%, while Sen2Cor had an overall accuracy of 84%.
8. These approximations address important challenges for cloud detection methods: distinguishing clouds from snow, mountain tops, bright deserts, and large built-up objects.
9. Since each cloud detection method relies on different ad hoc hypotheses, its usefulness varies from scene to scene.
10. Studies that target specific regions, such as the current paper, provide valuable advice even though its results cannot be generalized to non-forest areas.

(c) Primeiro parágrafo:

Segundo parágrafo:

1. A special diagrammatic technique is introduced as a tool for database design.
2. Semantic ambiguities in these models are analyzed.
3. This model incorporates some of the important semantic information about the real world.
4. Some implications for data integrity, information retrieval, and data manipulation are discussed.
5. A data model, called the entity-relationship model, is proposed.
6. The entity-relationship model can be used as a basis for unification of different views of data: the network model, the relational model, and the entity set model.
7. Possible ways to derive their views of data from the entity-relationship model are presented.
8. An example of database design and description using the model and the diagrammatic technique is given.

(d) Primeiro parágrafo:

Segundo parágrafo:

1. In 2019, five petabytes of images were produced by the satellites Landsat-7/8, CBERS-4/4A, Terra/Aqua, and Sentinel-1/2/3 [3].
2. Remote sensing satellites revisit the same place over the Earth’s surface at different times and produce image sequences of the same location over time.
3. Image time series analysis and machine learning methods have been widely used for land use and cover classification and change detection with good results [6–11].
4. Time series derived from these sequences are very useful to investigate the dynamics of the environment over time.
5. Nowadays, researchers have free access to an unprecedentedly large number of remote sensing images collected by different satellites and sensors with distinct spatial, temporal, and spectral resolutions.
6. Big Earth observation data sets bring new challenges and opportunities, including the novel generation of technological solutions to store, process, disseminate, and analyze them [4].
7. One of these opportunities is the use of time series analysis to extract landscape change information from many remote sensing images [2,5].

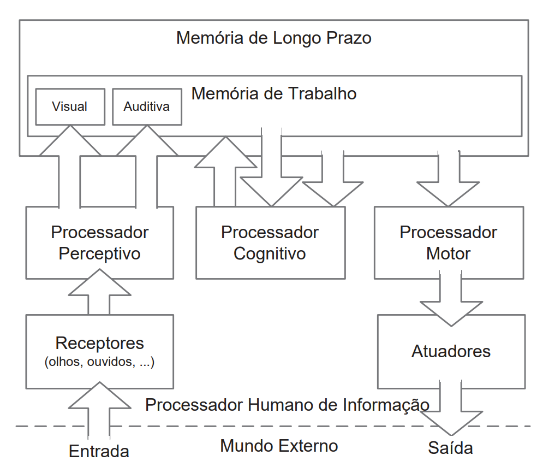
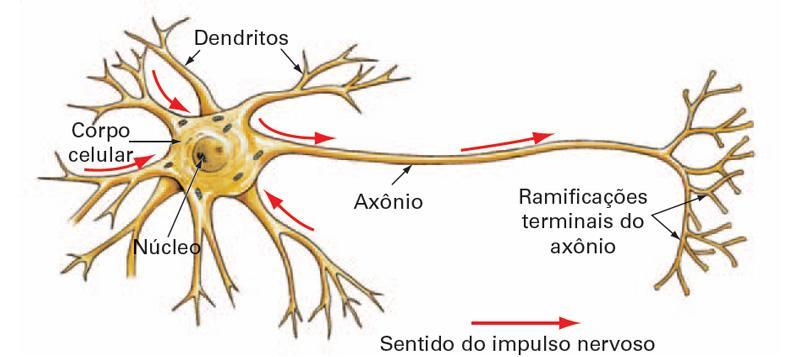
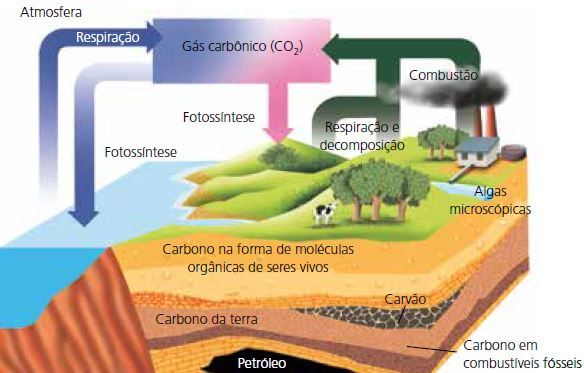
3) Reescreva os trechos abaixo com o objetivo de aumentar a legibilidade e remover erros de gramática.

1. A máxima entropia (MENT) descreve o maior valor encontrado em todas as simulações. Quanto maior o valor da máxima entropia (MENT), pior é a qualidade dos resultados.
2. Para identificar quais foram os usos e a cobertura terrestre do Estado, foi usado técnicas de aprendizagem de máquina para classificar séries temporais de imagens de satélite. Foi treinado uma base de dados de 1,892 amostras entre os anos de 2000 à 2015 em um modelo SVM.
3. O modelo é inicializado com os dados relativos ao ano de 2000 e simulou até 2020
4. Finalizado a classificação, é realizado uma série de pós-processamento para melhorar a qualidade da classificação com aplicação de um filtro Bayesian e máscaras de área urbana, cana-de-açúcar, água, floresta e cerrado. Essas duas máscaras são obtidas dos projetos PRODES Amazon e PRODES Cerrado, ambos produzidos pelo INPE.
5. O modelo A (alfa = 2), reproduz apenas 30% dos dados observados. O modelo B (alfa = 5), alcançou 0.8, muito mais.
6. Dessa forma, a memória do computador de quem realizar o processamento não é sobrecarregado.
7. Figueiredo et al. [2] apresenta um estudo da evolução anual do DAP para as principais espécies da Amazônia. O trabalho de Silva [3] apresenta um estudo mais detalhado da taxa evolução anual do diâmetro à altura do peito (DAP), apenas para uma espécie.
8. O modelo usa imagem de satélite de 2000 a 2015. As amostras de usos da terra foram fornecidas para os anos de 2000 a 2015. Elas possuem quatro classes: pasto, soja, milho, floresta, cerrado.
9. Após finalizado a criação do experimento, o processo é dividido em três processos e a simulação roda em paralelo.
10. Na Figura 2 é demonstrado o gráfico de evolução da população brasileira nos últimos 30 anos (até 2010).
11. A partir do modelo treinado, são classificados seis conjuntos de imagens de satélite do produto MOD13Q1 do sensor MODIS. Esse produto são caracterizados por possuírem uma resolução espacial de 250 metros e frequência temporal de 16 dias em uma projeção sinusoidal.

4) Melhore os títulos abaixo.

1. Investigando estratégias para a otimização dos modelos de dispersão de poluentes
2. Contagem de planetas anões utilizando os algoritmos de aprendizado de máquina com redes neurais
3. Levantamento da Literatura sobre Potencial de Localização Espacial de Aterros Sanitários em Todo o Estado de São Paulo
4. Um Estudo sobre a Granularidade Semântica nos Sistemas de Informação Geográfica: Aspectos Conceituais e Práticos
5. Comparação de Diferentes Técnicas de Identificação e Reconhecimento de Nuvens em Imagens Digitais Sentinel-2
6. Análise Estatística para a Estimativa de Produção Agrícola Total para o Estado de Mato Grosso
7. Novas Técnicas para Fusão de Imagens de Satélite usando o Aprendizado Profundo
8. Estimando tamanho de população e quantidade de consumo de energia usando imagens de satélites noturnos
9. Um modelo livre de falhas para a estimativa de ocorrências de focos de incêndios
10. Técnicas de Teste de Software Aplicados à Estratégia de Simplificação Expandida de Polígonos Vetoriais
11. Os resultados de um estudo sobre a composição química do asfalto em um conjunto de estradas federais

5) Escreva um ou dois parágrafos para descrever cada uma das figuras abaixo, supondo que elas não estão no texto.

1. 
2. 
3. 

6) Simplifique o abstract a seguir para que tenha apenas 12 linhas. Explique o motivo pelo qual os trechos foram removidos.

Land use change is a very important research topic. It is the result of interactions between complex processes operating at different spatial, temporal, and behavioral scales. Simulation models at regional to global scales are often incapable of including locally determined processes of land use change, usually dealing with very large datasets. This paper introduces a modeling approach that integrates demand-driven changes in land area with locally determined conversion processes. The model is illustrated with an application for European land use, using two layers of data, one with 10x10km cells and the other 1x1km cells. Interactions between changing demands for agricultural land and vegetation processes leading to the re-growth of (semi-) natural vegetation on abandoned farmland are explicitly addressed. Succession of natural vegetation is simulated based on the spatial variation in biophysical and management related conditions, while the dynamics of the agricultural area are determined by a global multi-sector model. The simulations run from 2000 to 2010, with scenarios exploring future possibilities until 2050. The model was implemented in C, using a MySQL database. The results allow an exploration of the future dynamics of European land use and landscapes. The model approach is similarly suitable for other regions and processes where large scale processes interact with local dynamics. The tool is available for different operational systems.