Algorithms and semantic infrastructure for mutation impact extraction and grounding

Submitted by rene [1] on Mon, 2010-12-27 09:57

- Open Mutation Miner [2]
- Bioinformatics [3]
- Text Mining [4]

Title	Algorithms and semantic infrastructure for mutation impact extraction and grounding
Publication Type	Journal Article
Year of Publication	2010
Authors	Laurila, J. [5], N. Naderi [6], R. Witte [7], A. Riazanov [8], A. Kouznetsov [9], and C. Baker [10]
Refereed Designation	Refereed
Journal	BMC Genomics
Volume	11
Issue	Suppl 4
Pagination	S24
Date Published	12/2010
ISSN	1471-2164
Abstract	Background

Background
Mutation impact extraction is a hitherto unaccomplished task in state of the art mutation extraction systems. Protein mutations and their impacts on protein properties are hidden in scientific literature, making them poorly accessible for protein engineers and inaccessible for phenotype-prediction systems that currently depend on manually curated genomic variation databases.

Results
We present the first rule-based approach for the extraction of mutation impacts on protein properties, categorizing their directionality as positive, negative or neutral. Furthermore protein and mutation mentions are grounded to their respective UniProtKB IDs and selected protein properties, namely protein functions to concepts found in the Gene Ontology. The extracted entities are populated to an OWL-DL Mutation Impact ontology facilitating complex querying for mutation impacts using SPARQL. We illustrate retrieval of proteins and mutant sequences for a given direction of impact on specific protein properties. Moreover we provide programmatic access to the data through semantic web services using the SADI (Semantic Automated Discovery and Integration) framework.

Conclusion
We address the problem of access to legacy mutation data in unstructured form through the creation of novel mutation impact extraction methods which are evaluated on a corpus of full-text articles on haloalkane dehalogenases, tagged by domain experts. Our approaches show state of the art levels of precision and recall for Mutation Grounding and respectable level of precision

but lower recall for the task of Mutant-Impact relation extraction. The system is deployed using text mining and semantic web technologies with the goal of publishing to a broad spectrum of consumers.

Notes	This article is part of the supplement: Ninth International Conference on Bioinformatics (InCoB2010): Computational Biology
URL	http://www.biomedcentral.com/1471-2164/11/S4/S24 [11]
DOI	10.1186/1471-2164-11-S4-S24 [12]
Copyright	© 2010 Laurila JB, Naderi N, Witte R, Riazanov A, Kouznetsov A, and Baker CJ.; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0 [13]), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Impact Factor	3.76
History	Published: 2 December 2010

Attachment	Size
1471-2164-11-S4-S24.pdf [14]	1.19 MB

Source URL (retrieved on *2026-01-09 16:03*):
https://www.semanticsoftware.info/biblio/algorithms-and-semantic-infrastructure-mutation-impact-extraction-and-grounding

Links:
[1] https://www.semanticsoftware.info/users/rene
[2] https://www.semanticsoftware.info/category/project/mutation-miner
[3] https://www.semanticsoftware.info/category/topic/bioinformatics
[4] https://www.semanticsoftware.info/category/topic/text-mining
[5] https://www.semanticsoftware.info/biblio/author/66
[6] https://www.semanticsoftware.info/biblio/author/22
[7] https://www.semanticsoftware.info/biblio/author/1
[8] https://www.semanticsoftware.info/biblio/author/67
[9] https://www.semanticsoftware.info/biblio/author/14
[10] https://www.semanticsoftware.info/biblio/author/68
[11] http://www.biomedcentral.com/1471-2164/11/S4/S24
[12] http://dx.doi.org/10.1186/1471-2164-11-S4-S24
[13] http://creativecommons.org/licenses/by/2.0
[14] https://www.semanticsoftware.info/system/files/1471-2164-11-S4-S24.pdf