The Karyotype Ontology: A computational representation for human cytogenetic patterns

Jennifer D. Warrender (jennifer.warrender@newcastle.ac.uk)
Phillip Lord (phillip.lord@newcastle.ac.uk)
Newcastle University, Newcastle-upon-Tyne, UK

Introduction

One of the most commonly used terms in cytogenetics is the term karyotype - a description of the chromosomes present in a cell, such as the number of chromosomes, as well as the presence of abnormalities (if any). It is important due to its diagnostic application, as it can identify chromosomal abnormalities which cause a variety of genetic disorders. Human karyotypes are normally represented using a string, as defined by the International System for human Cytogenetic Nomenclature (ISCN). However they can be hard to parse, validate and query as they are not computationally amenable.

For this work, we are using an ontology - a model that captures knowledge as concepts and relationships between these concepts, to provide a strong computational and formal interpretation of karyotypes. The development of the ontology is potentially valuable for cytogenetics by reducing collections of karyotypes to a form that is easy to query, check and maintain.

Approach

For our methodology, we made use of a semi-agile approach, driven from the ISCN specification and other use cases. However, even with a semi-agile approach, there are many similar concepts. Therefore, we defined the ontology as a series of parameterisable patterns which expand to the full ontology [Warrender and Lord 2013], using a novel tool called Tawny-OWL [Lord 2013]. We are evaluating the ontology using examples defined by the ISCN2013.

Class Type | Implemented
--- | ---
Chromosome | 27
Chromosome Component | 1
Centromere | 25
Telomere | 25
Band | 1321
Event | 10
Feature | 18
Karyotype | 1
Base Karyotype | 16
ISCN Example Karyotype | 168
Named Karyotype | 18
Resolutions | 6
Total Number | 1636

Future Work

- Extend the scalability of the model to include $10^5$ karyotypes.
- Tools for visualisation of karyotypes.
- An attractive end user syntax for the computational encoding of karyotypes.

References

WARRENDER, J. D., AND LORD, P. 2013. A pattern driven approach to biomedical ontology engineering. SWAT4LS 2013 (Dec.).