Genome, Gene, Interval and Ontology

Yu Lin [1], Norihiro Sakamoto
Department of Sociomedical Informatics,
Graduate School of Medicine, Kobe University,
7-5-1, Kusunokicho, Chuo-ku, Kobe 650-0017 Hyogo, JAPAN
[1]linikujp@gmail.com

ABSTRACT. Biological sequence information is the most primitive data for bioinformatics data exchange and management, especially the genetic sequence information. However, a multitude of formats and models are being used now in different systems; Ontologies, which represent the structural relations of different types of entities, provide the basis of interoperation between systems. In this paper, we first compared the variable definitions existing in different biomedical ontologies. Then, by applying the Allen's Interval Calculus theory, a universal Genetic Interval and interval relations were defined both logically and ontologically. We showed how some genetic entities, such as genome, gene, DNA, chromosome, genetic marker, allele, polymorphism, and genetic susceptibility can be defined in OWL-DL using Protégé 3.3.1. By this framework, we could establish a data model for computing the co-localization of genomic elements with genes.

1 Introduction

Originally *ontology* was used as a philosophical term for the study of the nature of existence. In more recent years, the term is used for the structured representation of the types of entities and relations existing in a given domain in a way that is designed to support exchange and reuse of data and information.

In biomedical informatics, ontologies are used to support automatic retrieval and exchange of data across disciplinary boundaries. Ontologies provide the basis for interoperability between systems, and they are considered to be an important technology for the Semantic Web. It is only during the last decade that the importance of biomedical ontologies in addressing some of the grand challenges of genomic research has been recognized. [Collins et al. 2003] The use of biomedical ontologies has grown dramatically since the Gene Ontology (GO) Consortium was initiated in 1998 with three model organisms groups: FlyBase (Drosophila), the Saccharomyces Genome Database (SGD) and the Mouse Genome Database (MGD). [Ashburner et al. 2000] Further milestones were the establishment in 2001 of OBO (Open Biomedical Ontologies) to serve as "an umbrella body for the developers of life-science ontologies" and to provide an OBO ontology repository, which led in turn to the creation in 2005 of the OBO Foundry, an experiment directed towards the creation a suite of interoperable ontology modules designed to support life science research. [Smith et al. 2007] The goal of OBO is to provide computable definitions that can be used across all biological systems to describe and understand biological structures and processes.

However, the huge volume and complexity of biological data present new challenges. Primary data obtained in different locations or contexts and at different times may always be re-interpreted, and hence it is important never to throw it away. If scientists

or data curators are unable to communicate data and results reliably to one another, biological interpretation will be meaningless. Ontologies are the key here, but in moving from genotype to phenotype the problem becomes harder and less well defined.

A good example of relations between genotype and phenotype is genetic susceptibility to disease. Let's use the gene TCF7L2, which is a susceptibility gene to Type 2 Diabetes, as a case study here.

TCF7L2 (transcription factor 7-like 2) gene is a high mobility group (HMG) box-containing transcription factor implicated in blood glucose homeostasis. It is located in the cytogenetic band 10q25.3 according to the Entrez cytogenetic map. Many SNP's alleles of TCF7L2 are considered to be genetic susceptibility factors to Type 2 Diabetes.

Table 1. The susceptibility or resistance alleles of markers to T2D

Marker	**Allele**	**Susceptibility or Resistance to T2D**
DG10S478	0 allele	resistance
	non-0 allele	susceptibility
rs7903146	C allele	resistance
	T allele	susceptibility
rs12255372	G allele	resistance
	T allele	susceptibility
rs7895340	G allele	resistance
	A allele	susceptibility
rs11196205	G allele	resistance
	C allele	susceptibility
rs7901695	T allele	resistance
	C allele	susceptibility
rs12243326	C allele	susceptibility
rs4506565	T allele	susceptibility

According to the research on original papers, the first connection was established when Reynisdottir *et al.* (2003) reported a genome-wide significant linkage to chromosome 5q, 10q and 12q in the Icelandic population. Later, by genotyping 228 microsatellite markers spanning a 10.5-Mb interval on chromosome 10q in Icelandic individuals, Grant *et al.* (2006) revealed the susceptibility of a microsatellite DG10S478 to Type 2 Diabetes. Many researchers replicated the above results, and more SNPs located in TCF7L2 were reported susceptibility to Type 2 Diabetes, as we see in the above table.

DG10S478 is located inside the intron 3 of the TCF7L2 gene and within a well defined LD block ("exon 4 LD block of TCF7L2", that encapsulates part of intron 3, the whole of exon 4 and part of intron 4). The most important information here is the

location of the DG10S478, according to the location information we can confirm that DG10S478 is a susceptibility genetic factor.

As far as we know, the genetic variation which is located inside the intron, exon or regulatory elements of a gene could possibly affect the product of the gene; therefore it has an impact on the initiation of the development of disease. Many other genetic variations which located in the intergenic genomic region may have an impact on the transcript or they may through some currently unknown mechanism contribute to a predisposition to some given disease. This leads to one of the applications of our ontology to store the ongoing result of biomedical research in a structural form which will enable further knowledge to be gained in such circumstances.

2 Genetic entities in OWL-DL

Many genetic entities involved in the representation of genetic susceptibility to disease, such as genetic markers (SNP maker, microsatellite marker), gene, linked interval, allele, haplotypes, genotypes, and so on. For representing these terms, we need to reconsider the formalization of the basic genetic entities: *DNA, Chromosome, Genome* and *Gene*.

2.1 DNA , Chromosome and Genome

DNA is one of the biological structural entities; it is a biological macromolecule with a double helix structure. Genetic information is encoded by the linear sequence of bases in the DNA strands, which is also called the primary structure. Ontologically defined, *DNA_Molecule* is a *Nucleic_Acid* which *hasPrimaryStructure* Only *DNA_Sequence* and *hasPrimaryStructure* Some *DNA_Sequence*.

Human DNA is structured into chromosomes, which change their shape and content during the process of cell division. However here we ignore the differences of chromosomes during cell division process. The class *Chromosome* is a subclass of *Organelle*, which is a subclass of *GO:cellular_component*. By ontological definition, *Chromosome* is an *Organelle* which *hasStructurePart* Only (*DNA* or *Histone*).

A genome is the total set of different DNA molecules of an organelle, cell or organism. The human genome consists of 25 different DNA molecules, the mitochondrial DNA molecule plus the 24 different chromosomal DNA molecules. [Strachan and Read 2004] In our definition, Genome is a *Biological_ObjectAggregate*, which *hasComponentPart* Some *DNA_Molecule*.

The difference between DNA molecules and genes is that DNA molecules are naturally intact with mass and 3-D dimension. Whereas genes are not naturally intact and are segment of a DNA molecule located in chromosome. Such a segment of DNA molecule is designated a “gene” by scientist because it encodes functional protein or RNA.

Gene

The definition of gene is changing along with the rapid pace of observations obtained from the genomics field. Recently, Gerstein et al. proposed their definition of gene as: "The gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products" [Gerstein et al. 2007] based on the observations from ENCODE consortium project. Gingeras TR. proposed in his article [Gingeras et al. 2007] that instead of the currently usage of gene as a "genomic region producing a polyadenlayted mRNA that encodes a protein", transcripts need to be considered as "operational unit of a genome".

We investigated the existing definition provided by other ontologies.

We searched National Center for Biomedical Ontologies via BioPortal and OLS by keyword "gene", and found four ontologies holding the term "*Gene*" as a class. Their definitions and *hasSubclass* or *is_a* hierarchy are listed out in Table2. (SO can only searched by OLS, FMA only by BioPortal search)

Table 2 Differences of Gene term's definition and hierarchy in current biomedical ontologies

Ontology	**ID**	**Definition**	**Hierarchy (hasSubclass or is_a)**
Cell Cycle Ontology	CCO:U0000004	Each of the units of heredity which (except for polygenes) may be regarded as the controlling agents in the expression of single phenotypic characters and are usually segments of a chromosome at fixed positions relative to each other.	Entity - Continuant -- Gene
Gene Regulation Ontology	GRO: Gene	A unit of inheritance; a working subunit of DNA that contributes to phenotype/function and carries a particular set of instructions, usually coding for a particular protein.	DNA - DNA region -- gene
Foundational Model of Anatomy	Structural gene FMA :74402	Biological macromolecule which is a segment of a DNA molecule and which consists of sets of nucleotide pairs organized into regions of coding, noncoding and regulatory sequences.	Anatomical Entity - Physical anatomical entity -- Material anatomical entity --- Anatomical Structure ---- Biological Macromolecule ----- Structural Gene
Sequence Ontology	SO:0000704	A region (or regions) that includes all of the sequence elements necessary to encode a functional transcript. A gene may include regulatory regions, transcribed regions and/or other functional sequence regions.	Sequence Ontology - sequence feature -- region --- gene

In above ontologies, *Gene* is a subclass of *DNA region* (CCO) or *region* (SO); in FMA, *Gene* is a *Biological Macromolecule* which is a segment of DNA. In CCO and GRO, *Gene* is defined as the inheritance unit. All ontologies agree on two generic characteristic of a gene: 1) a gene is a region (or segment, or subunit) of DNA on chromosome; 2) a gene encodes a transcript.

We try to define a gene as generically possible. Thus, in our ontological definition:

Class: Gene SubClassOf: DNA_Segment

EquivalentTo: DNA_Segment that

hasPrimaryStructure ONLY DNA_Sequence AND

hasPrimaryStructure SOME DNA_Sequence AND

hasEndProduct (Protein OR RNA_Molecule) AND

isTranscribeTo SOME Transcript AND

hasGeneLocus SOME Gene_Locus

Genetic Interval

In mathematics, an interval is a certain subset of an ordered set. Allen JF has developed an Interval Temporal Logic [Allen et al. 1984, Allen et al. 1990] to present a concise, formal axiomatization of "interval-based" time. Hobbs, J. R. and Pan, F. developed a time ontology to describe the temporal entity and relations for semantic web. [Hobbs and Pan 2004] In biological investigations, many genetic entities, such as genes, alleles, haplotypes and genetic markers are based on their sequence information, which is a physical subset of an ordered DNA or RNA base set. Based on their similarities, and the needed of co-localization of some genetic entities with genes, we presented genetic intervals as physical material intervals, which start and end at a certain point or boundary.

Here, we use a universal named "*Biological Interval*" to describe the spatial continuous physical entity which contains ordered biological sets (DNA segments, Genetic Markers, Nucleic Acid Base Residues, RNA segments, Protein segments) between two boundaries: start boundary and end boundary on a chromosome, RNA or protein. *Biological Interval* is distinct by three intervals: *Genetic Interval*, *Interval Base Residue* and *Protein Interval*; *Genetic Interval* is distinct by *DNA Interval* and *RNA Interval*. *DNA Interval* has primary DNA sequence structure by its definition; whereas RNA Interval has primary RNA sequence structure.

We discuss mainly *Genetic Interval* in this paper. *Genetic Interval* is continuous, so that neither a genome (a collective of chormosomes), nor a genotype of a diploid (a collective alleles coming from different chromosomes), nor a gene family (a collective of genes of same homolog located to different chromosomes) are *Genetic Interval*. Whereas gene clusters which are juxtaposition genes on chromosomes can be a subclass of *Genetic Interval*. By length, the longest *DNA Interval* is the interval with the same start point and end point as the chromosome, and the smallest is when one start point and end point are equal (one residue of DNA or RNA).

By ontological definition, *Genetic Interval* is a *Biological Interval*, which hasEnd-Point *End_Boundary_of_Interval* and hasEndPoint Exactly 1, and hasStartPoint *Start_Boundary_of_Interval* and hasStartPoint Exactly 1, and hasIntervRelations with *Genetic Interval*, and is (*DNA Interval* or *RNA Interval*). The first fold subclasses of *DNA Interval* are: *DNA segment*, *Genetic Marker*, *Probe*, *Amplifier*, and *Flanking Sequences*. *DNA Segment* has such subclasses as: *Gene*, *Allele*, *Exon*, *Intron*, *Gene Regulatory Elements*, *Haplotype*, *Intergenic Segment*, *Linkage Interval*, *LD block* and so on. *DNA segment* is a *DNA Interval* that is located on a chromosome.

Interval Relations

Based on the interval relations we will introduce later, the locus of a susceptibility gene or region in a chromosome can be inferred from the location of other genetic markers.

The following axiom states one of the object properties of Genetic_Interval : hasIntervalRelation

<owl:ObjectProperty rdf:about="#hasIntervalRelation">
<rdfs:range rdf:resource="#Genetic_Interval"/>
<rdfs:domain rdf:resource="#Genetic_Interval"/>
</owl:ObjectProperty>

We use logic to define genetic interval and its relations:

geneticInterval(X) X is genetic interval
hasStartPoint(x1,X) X has start point x1
hasEndPoint(x2,X) X has end point x2

If the start and end of a genetic interval are identical, the genetic interval is a point interval.

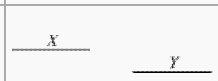
$pointInterval(X) \leftrightarrow geneticInterval(X) \wedge hasStartPoint(x,X) \wedge hasEndPoint(x,X)$

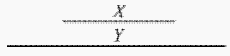
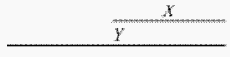
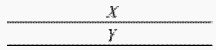
The start point and end point is point interval.

$pointInterval(x1) \leftrightarrow geneticInterval(X) \wedge hasStartPoint(x1,X)$

$pointInterval(x2) \leftrightarrow geneticInterval(X) \wedge hasEndPoint(x1,X)$

table 3 relations compare with Allen's Interval calculus

Relation in Allen's Interval	**Illustration**	**Relations of Biological Sequence Interval**
X < Y Y > X		isLocatedBefore isLocatedAfter
XmY YmiX		isAdjacentBefore isAdjacentAfter
XoY YoiX		isOverlapStartWith isOverlapEndWith

XsY YsiX		isStartsWith
XdY YdiX		isContainedIn
XfY YfiX		isEndsWith
X=Y		

isLocatedBefore
isLocatedBefore is both a Class-Class and an Instance-Instance relation.

isLocatedBefore(X,Y) genetic interval X is located before genetic interval Y

If genetic interval X is located before genetic interval Y, then the end of X is before the start of Y.

isLocatedBefore(X,Y) ↔ *hasEndPoint*(x,X) ∧ *hasStartPoint*(y,Y) ∧ *isLocatedBefore*(x,y)

isLocatedAfter
isLocatedAfter(X,Y) genetic interval X is located after genetic interval Y

If genetic interval X is located after genetic interval Y, then the start of X is after the end of Y; or the end of Y is before the start of X. If *isLocatedBefore*(X,Y), then *isLocatedAfter*(Y,X)

isLocatedAfter(X,Y) ↔ (*hasStartPoint*(x1,X) ∧ *hasEndPoint*(y1,Y) ∧ *isLocatedAfter*(x1,y1))

∨(*hasEndPoint*(x2,X) ∧ *hasStartPoint*(y2,Y) ∧ *isLocatedBefore*(y2,x2))

isLocatedAfter(X,Y) ≡ *isLocatedBefore*(Y,X)

isStartsWith
isStartsWith(X,Y) genetic interval X starts with genetic interval Y

If genetic interval X starts with genetic interval Y, then the start of X and the start of Y are identical and the end of X is before the end of Y.

isStartsWith(X,Y) ↔ ((*hasStartPoint*(x1,X) ∧ *hasStartPoint*(x1,Y))

∧ (hasEndPoint(x2,X) ∧ hasEndPoint(y,Y) ∧ isLocatedBefore(x2,y))

isEndsWith
isEndsWith(X,Y) genetic interval X ends with genetic interval Y

If genetic interval X ends with genetic interval Y, then the end of X and the end of Y are identical and the start of X is after the start of Y.

isEndsWith(X,Y) ↔ ((*hasEndPoint*(x1,X) ∧ *hasEndPoint*(x1,Y))
∧ *(hasStartPoint(x2,X)* ∧ *hasStartPoint(y,Y)* ∧ *isLocatedAfter(x2,y))*

isOverlapStartWith

isOverlapStartWith(X,Y) genetic interval X overlaps the start of genetic interval Y

If genetic interval X overlaps the start of genetic interval Y, then the start of X is before the start of Y and the end of X is before the end of Y and the end of X is distinct from the start of Y.

isOverlapStartWith(X,Y) ↔ (*hasStartPoint*(x1,X) ∧ *hasStartPoint*(y1,Y) ∧ *isLocatedBefore*(x1,y1))
∧ (*hasEndPoint*(x2,X) ∧ *hasEndPoint*(y2,Y) ∧ *isLocatedBefore*(x2,y2))
∧ ((*pointInterval*(x2,X) ≠ *pointInterval*(y1,X))

isOverlapEndWith

isOverlapEndWith(X,Y) genetic interval X overlaps the end of genetic interval Y

We can define isOverlapEndWith by isOverlapStartWith, if *isOverlapStartWith*(X,Y) , then *isOverlapEndWith*(Y,X)

isOverlapStartWith(X,Y) ↔ *isOverlapEndWith*(Y,X)

Another approach is straightforword. If genetic interval X overlaps the end of genetic interval Y, then the start of X is after the start of Y and the end of X is after end of Y and the start of X is distinct from the end of Y.

isOverlapEndWith(X,Y) ↔ (*hasStartPoint*(x1,X) ∧ *hasStartPoint*(y1,Y) ∧ *isLocatedAfter*(x1,y1)) ∧ (*hasEndPoint*(x2,X) ∧ *hasEndPoint*(y2,Y) ∧ *isLocatedAfter*(x2,y2))
∧ ((*pointInterval*(x1,X) ≠ *pointInterval*(y2,X))

isContainedIn

isContainedIn(X,Y) genetic interval X is contained inside genetic interval Y

If genetic interval X is contained in genetic interval Y, then start of X is after the start of Y and the end of X is before the end of Y and the start of X is unique with the start of Y, the end of X is unique with the end of Y.

isContainedIn(X,Y)
isContainedIn(X,Y) ↔ *(hasStartPoint(x1,X)* ∧ *hasStartPoint(y1,Y)* ∧ *isLocatedAfter(x1,y1))*
∧ (*hasEndPoint*(x2,X) ∧ *hasEndPoint*(y2,Y) ∧ *isLocatedBefore*(x2,y2))
∧ ((*pointInterval*(x1,X) ≠ *pointInterval*(y1,X))
∧ ((*pointInterval*(x2,X) ≠ *pointInterval*(y2,X))

isAdjacentBefore

isAdjacentBefore(X,Y) genetic interval X is adjacent to the start of genetic interval Y

If genetic interval X is adjacent to the start of genetic interval Y, then the start of X is before the start of Y and the end of X is before the end of Y, and the end of X is identical with the start of Y.

isAdjacentBefore(X,Y) ↔ (*hasStartPoint*(x1,X) ∧ *hasStartPoint*(y1,Y) ∧ *isLocatedBefore*(x1,y1))

∧ (*hasEndPoint*(x2,X) ∧ *hasEndPoint*(y2,Y) ∧ *isLocatedBefore*(x2,y2))

∧ ((*pointInterval*(x2,X) ≡ *pointInterval*(y1,X))

isAdjacentAfter

isAdjacentAfter(X,Y) genetic interval X is adjacent the end part of genetic interval Y

If genetic interval X is adjacent to the end part of genetic interval Y, then the start of X is after the start of Y and the end of X is after the end of Y, and the end of X is the same as the start of Y.

isAdjacentAfter(X,Y) ↔ (*hasStartPoint*(x1,X) ∧ *hasStartPoint*(y1,Y) ∧ *isLocatedAfter*(x1,y1))

∧ (*hasEndPoint*(x2,X) ∧ *hasEndPoint*(y2,Y) ∧ *isLocatedAfter*(x2,y2))

∧ ((*pointInterval*(x2,X) ≡ *pointInterval*(y1,X))

3 Conclusion

In biological research for genetic susceptibility to diseases, the observed subjectives are entities of *Genetic Interval*, such as linkage intervals, alleles, SNPs, genetic markers and so on. Nowadays, the study of genome-wide association, which examines genetic variation across the human genome, will discover hundreds SNPs associated with phenotype. However, only the SNPs, which can be co-localized within gene or gene regulate elements, will be meaningful for the further research.

In bioinformatics, the single-letter sequence code that describes a DNA or RNA is a simplified representation of a 3D chemical entity. The 3D structure of the DNA or RNA will become more significant after we explore more sequence information. We distinguish the 3D structure from other defined classes in the current existing ontologies, and introduced a universal term of *Genetic Interval* to represent the linear sequence data. By using the OWL-DL language, we defined some important basic terms in genetics: *DNA*, *Genome*, *Chromosome*, *Gene* and *Genetic Interval*.

Some groups have been developing a similar term, for example, BioPAX describes a class "*SequenceInterval*" as an interval of a sequence with only one *sequenceIntervalBegin* and one *sequenceIntervalEnd* [biopax URL]; no detailed interval relations were described. The GDB Human Genome Database uses ab object-oriented database to describe objects, classes and attributes. There is a "*Genomic Segment*" class in GDB, which is similar to *Genetic Interval*. The "*Genomic Segment*" comprises of any named region or set of regions of a genome. However, the subclasses of Genomic Segment are too broad. For example, some entities such as *Chormosome*, *Fragile Site*, *Cell Line*,

Contig, *Library* and *mapping panel* are subclasses of *Genomic Segment* class. Since it is an object-oriented database, its design and mission are different from our ontology.

Adopting the Allen's Interval Calculus to define the relations between *Genetic Interval*, the automatic computing and classification of meaningful genetic variants become feasible.

The following rules were established based on the ontology:

1) *isStructurePart(?z,?x)* $\wedge$ *isStartsWith(?z,?y)* $\rightarrow$ *isOverlapStartWith(x?, y?)*

If ?z is the structure part of ?x, and ?z is stars with ?y, then ?x overlaps start of ?y.

2) *isStructurePart(?z,?x)* $\wedge$ *isEndsWith(?z,?y))* $\rightarrow$ *isOverlapEndWith(x?, y?)*

If ?z is structure part of ?x, and ?z is end with ?y, then ?x overlaps start of ?y.

3) *isLocatedBefore(?x,?y)* $\rightarrow$ *isLocatedAfter(?y,?x)*

If ?x is located before ?y, then ?y is located after ?x

4) *hasStartLocationOnMap(?x,?y)* $\wedge$ *hasStartPoint(?y,?z)* $\wedge$ *isStarsWith(?z,?m)* $\rightarrow$ *hasStartLocationOnMap(?x,?m)*

5) *hasStartLocationOnMap(?x,?y)* $\wedge$ *hasStartPoint(?y,?z)* $\wedge$ *isStarsWith(?m,?z)* $\rightarrow$ *hasStartLocationOnMap(?x,?m)*

If ?x is the start location of ?y, and?y is the start point boundary of ?z, and whenever ?z is the start part of ?m or ?m is the start part of ?z, ?x is also the start location of ?m.

However, our ontology is silent about the order of the subsets inside a *genetic interval*, which we will tackle in future work; Also we need to extend the logical definition of *genetic interval aggregate*, as well as the ontological definition. Furthermore, the evaluation and validation of this piece of work needs to be done in the near future.

Reference

[Collins et al. 2003] F.S. Collins, E.D. Green, A.E. Guttmacher, M.S. Guyer, A vision for the future of genomics research, *Nature* **422**(2003), 835-847.

[Ashburner et al. 2000] M. Ashburner, C.A. Ball, J.A. Blake, et al, Gene ontology: tool for the unification of biology, The Gene Ontology Consortium. *Nat Genet* **25**(2000), 25–29.

[Smith et al. 2007] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L.J. Goldberg, K. Eilbeck, A. Ireland, C.J. Mungall, OBI Consortium, N. Leontis, P. Rocca-Serra, A. Ruttenberg, S.A. Sansone, R.H. Scheuermann, N. Shah, P.L. Whetzel, S. Lewis, The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nat Biotechnol* **25**(2007), 1251-1255.

[Strachan and Read 2004] T. Strachan, A.P. Read,Human Molecular Genetics 3, Garland Science, New York, 2004, 635.

[Gerstein et al. 2007] M.B. Gerstein, C. Bruce, J.S. Rozowsky, D. Zheng, J. Du, J.O. Korbel, O. Emanuelsson, Z.D. Zhang, S. Weissman, M. Snyder, What is a gene, post-ENCODE? History and updated definition, *Genome Res*. **17**(2007), 669-681.

[Gingeras et al. 2007] T.R. Gingeras, Origin of phenotypes: genes and transcripts, *Genome Res*. **17**(2007), 682-690.

[Allen et al. 1984] J.F. Allen, Towards a general theory of action and time, *Artificial Intelligence* **23**(1984), 123-154.

[Allen et al. 1990] J.F. Allen and P.J. Hayes, Moments and Points in an Interval-Based Temporal Logic, *Computational Intelligence* **5**(1990), 225–238.

[Hobbs and Pan 2004] J.R. Hobbs and F. Pan, An Ontology of Time for the Semantic Web, *ACM Transactions on Asian Language Processing (TALIP): Special issue on Temporal Information Processing*, **3**(2004), 66-85.

[biopax URL] URL: http://www.biopax.org/release/biopax-level3.owl