Ontology Engineering Group, UPM

Nov. 12, 2015, Madrid, Spain

Gravitation and Fitness in Ontology Dynamics

Vadim Ermolayev

Dept of IT, Zaporozhye National Univ., Ukraine

http://ermolayev.com/

SemData

FP7 Marie Curie IRSES

The Plan …

- **Motivation**
  - Is ontology dynamics on the R&D agenda? Are there open issues? Where to seek insights?
- **Ontologies and their Fitnesses**
  - ~~What is an ontology? – conceptually and formally …~~
  - How to make an ontology a "shared specification" of domain knowledge?
  - Ontology **Fitness** – a key metric of being "shared"
- **The Law of Gravitation – for Domains and Ontologies**
  - What if the forces between domains and ontologies existed? How would these look like – in a model?
- **Supporting Techniques and Tools**
  - Now we have a model, so what? Is it adequate? Do we have instruments to measure things and compute forces?
- **An Outline of Potential Use Cases**
  - Even so, how could these be used? Would these make our lives better?

Motivation

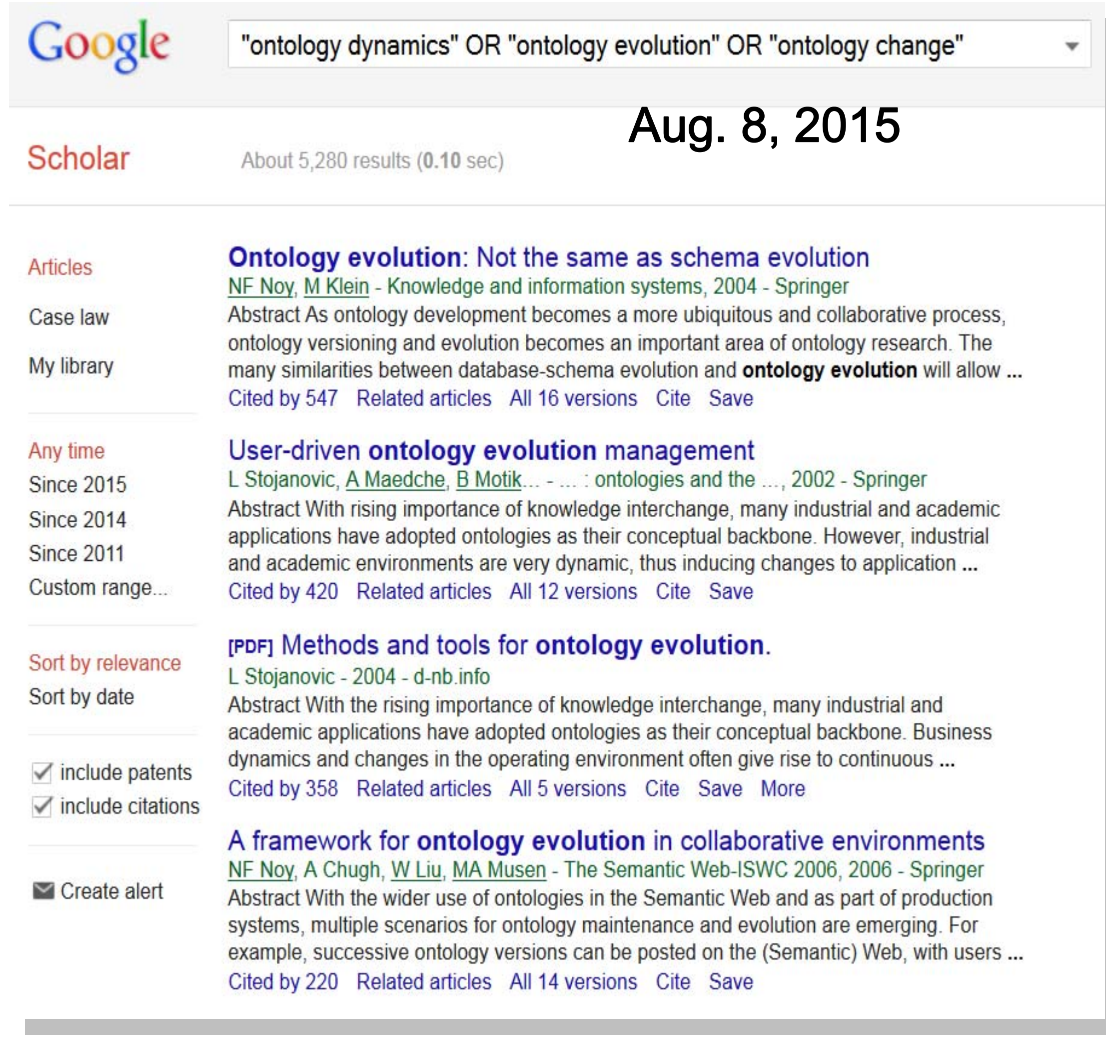
- Is ontology dynamics on the R&D agenda?
- Are there open issues?
- Where to seek insights?

So, is it on the Agenda?

… on the Agenda?

As of Nov. 11, 2015

- Google:
  - ~ 42 900
- Google Scholar:
  - Any time: ~5 420
  - Since 2015: 290



The State-of-the-Play

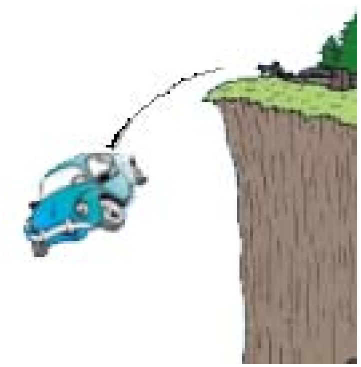
- Ontology **change** * – *changing an ontology <u>in response</u> to a certain need*
- Several facets -> research fields:
  - Ontology **evolution** (reactive response to a change in the domain or its conceptualization)
  - Ontology **refinement** (goal-directed, proactive change)
  - Ontology **versioning** (enable transparent access to different versions of an ontology)
  - Ontology **mapping** (identify related vocabulary elements)
  - Ontology **morphing** (map between vocabularies and axioms)
  - Ontology **matching** (map and measure semantic distance between vocabularies and axioms)
  - Ontology **alignment** (result of matching process)
  - Ontology **translation** (to a different representation language)
  - Ontology **integration**/ **merging** (fuse knowledge from ontologies covering similar/ identical domains)
  - Ontology **debugging** – **diagnosis** and **repair** (render an ontology consistent/coherent)

Are these about DYNAMICS?

* Giorgos Flouris et al.: Ontology Change: Classification and Survey. *The Knowledge Engineering Review*, 23(2), 117-152, 2008. doi:10.1017/S0269888908001367

What is Changed, not Why …

- Does this one fit to go faster than 100 mph?



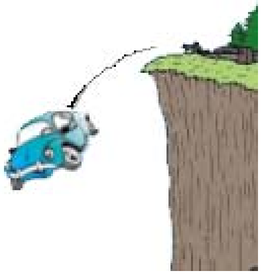
Ontology Change:
- Let’s measure the distance passed in 1 hour
- If <= 100 mi, then NOT

Would be happy to have:
- Perhaps NOT, judging by its current shape
- Could be able however if thrown from there – up the cliff

Open Issues?

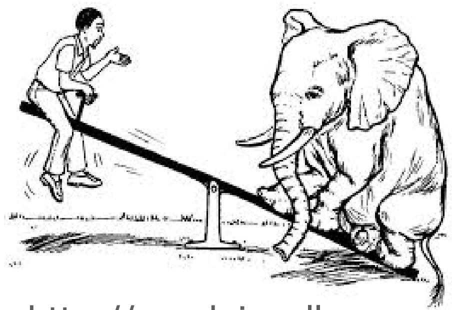
- Ontology Change – studies the ways to change an ontology in response to a need
  - C.F.: **Kinematics** – studies the **motion** of objects without a reference to its causes
- Ontology Change:
  - In fact does Kinematics but misses the study of the causes that trigger the need
  - C.F.: **Dynamics** - is concerned with the study of **forces** and **torques** and their **effect on motion**

- **Ontology Change** misses **Ontology Dynamics**
  - What would be the force **causing** this going faster then 100 mph?



Focus and Assumptions

- In our work:
  - Insights (so far) are taken from … Newtonian **Mechanics**
    - In particular: the models are based on the Newton's Law of **Universal Gravitation**



http://goodpixgallery.com/

- Differences (simplifications / complications):
  - Knowledge representations are discrete in their nature:
    - Continuous models -> **discrete models**
  - Objects are immaterial:
    - Masses -> **Fitnesses**

Ontologies and their Fitnesses

- How to make sure that an ontology is a "**shared** specification" of domain knowledge?

A Shared Spec …

> *"I find it critical to remember that every ontology is a treaty – a social agreement – among people with some common motive in sharing."*

Tom Gruber in the Interview for the Official Quarterly Bulletin of AIS Special Interest Group on Semantic Web and Information Systems, Vol. 1, Issue 3, 2004.

- Who are those **People**?
- What are their **Motives to share**? Why **Common**?

- Answers are sought for Domain Ontologies:
  - Those describing a well circumscribed Domain of Discourse
    - Like Gene Research, Financial News, Sport Cars, ICT Research on BIG DATA, etc.
  - A community of **people**, with **motives** to share domain knowledge, could be outlined
    - Domain Knowledge Stakeholders (or Subject Experts)
    - Common Motive – the description of Domain Knowledge to be based on their views
      - Thought of as **requirements**

Implies …

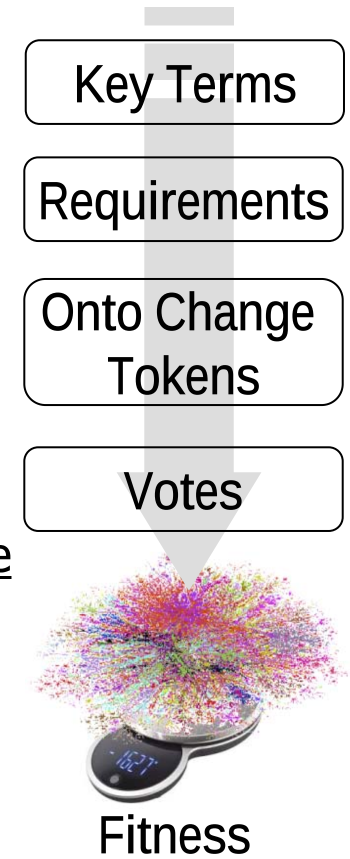
- An OE Methodology to be able to result in a Shared Spec.:
  - Elicit the (change) requirements from the Domain Knowledge Stakeholders as fully as possible
    - Measure completeness
  - Transform the elicited requirements – often informal
    - To the (changes in the) ontology – formal
  - Measure how well the result **FITS** the opinions of the Domain Knowledge Stakeholders
    - Completeness; Accuracy
- Otherwise:
  - The **Commitment** of the Domain Knowledge Stakeholders to the output ontology will be low
- So:
  - Higher the **Fitness** -> higher the **Commitment**
  - Higher the **Commitment** -> more **Shared** the Spec. is
- Which brings us about acknowledging the role of Ontology **Fitness** …

The Scales for Fitness

- Φ – the degree of the conformance of an ontology to the requirements of the Domain Knowledge Stakeholders
- Difficult to measure
- Having the following is all not easy:
  - The requirements
  - The ontology
  - These two compared and difference / conformance measured
- Done in our **OntoElect** OE methodology

Fitness in OntoElect

- OntoElect *: Φ is understood as proportional to the ratio of positive and negative **votes** of the Stakeholders regarding the assessed ontology
  - Votes collected indirectly – using a statistically representative professional Document Corpus:
    - Extract a **saturated** set of multi-word key terms
    - Select the most **influential** key terms – Requirements
    - Transform the natural language definitions of the terms to **formalized structural contexts** – Ontology Change Tokens
    - **Map** the structural contexts to the ontology – Votes
    - Compute change in **Fitness** – more or less Votes



* Tatarintseva, O. et al. (2013) Quantifying Ontology Fitness in OntoElect Using Saturation- and Vote-Based Metrics. In: Ermolayev et al. (eds.) *ICT in Education, Research, and Industrial Applications. Revised Selected Papers of ICTERI 2013*, CCIS 412, pp. 136–162

The Law of Gravitation

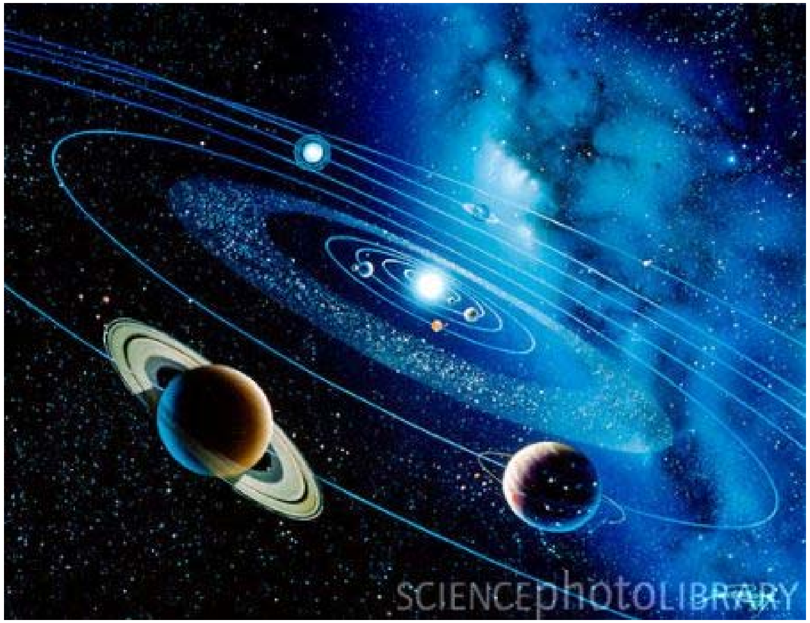
- What if the forces between domains and ontologies existed?
- How would these looked like – in a model?

A “Closed” Mechanical System

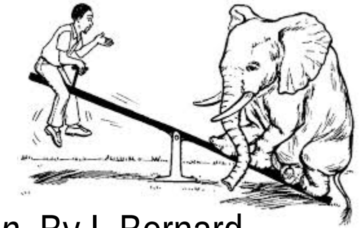
- Coordinate grid:
  - E.g. Helio-centric, 3D, Decartes
- Distance:
  - (K)m, from point (0,0,0)
- Mass:
  - Kg
- Force:
  - Newton’s Law of Universal Gravitation *

$$F = \gamma \frac{m_1 \times m_2}{r^2}$$

  - Which are the most influential objects in the system?
  - How do they influence the other objects?



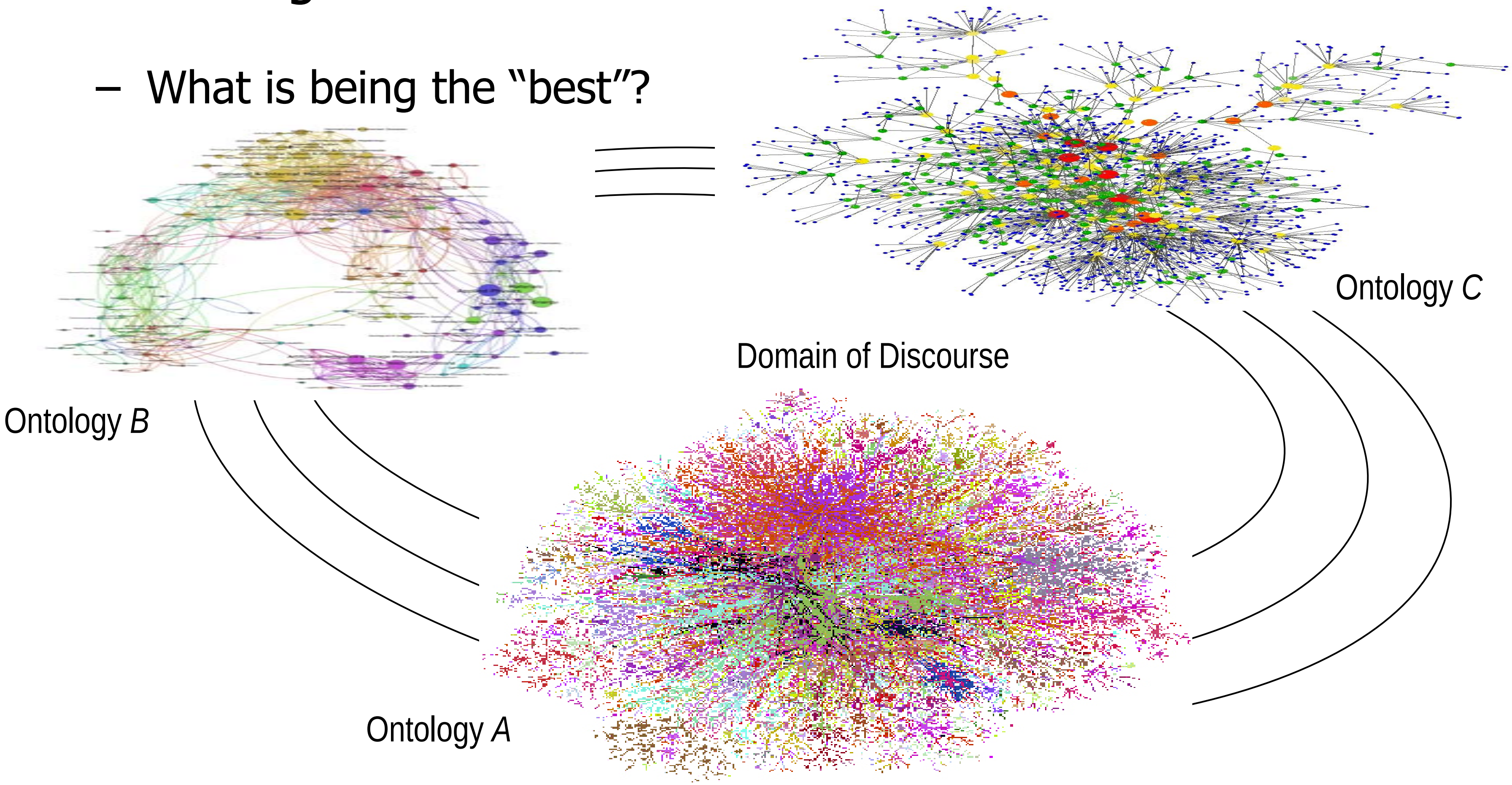
https://harifromtheworldday.wordpress.com/2012/06/06/the-solar-system-has-nine-planets/



* Newton, I.: *The Principia, Mathematical Principles of Natural Philosophy, a new Translation*. By I. Bernard Cohen and A. Whitman, preceded by "A Guide to Newton's Principia" by I Bernard Cohen, University of California Press, 1999

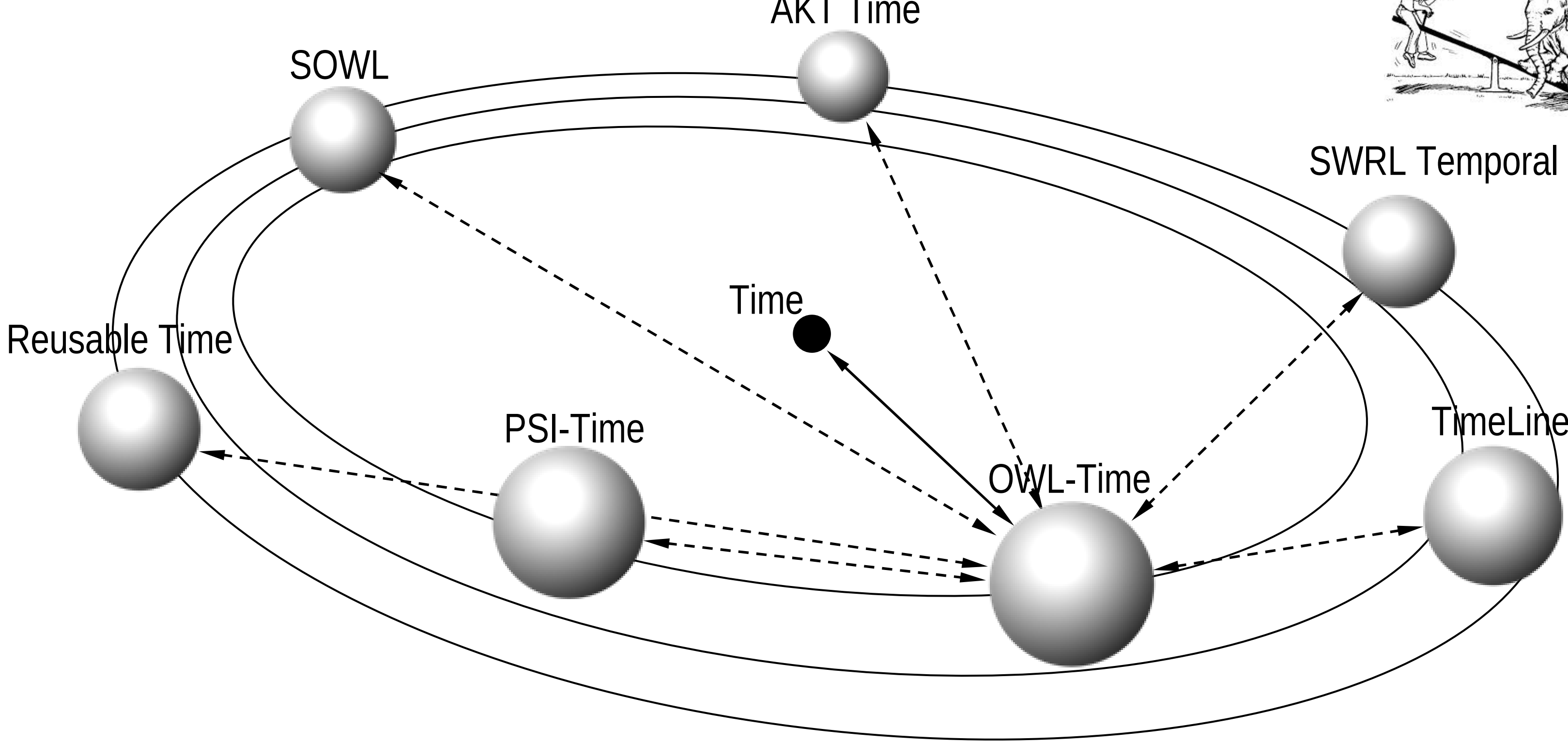
A Domain-Ontology System

- Which of the ABC is the "best" for (most influential in) describing the Domain?
  - What is being the "best"?



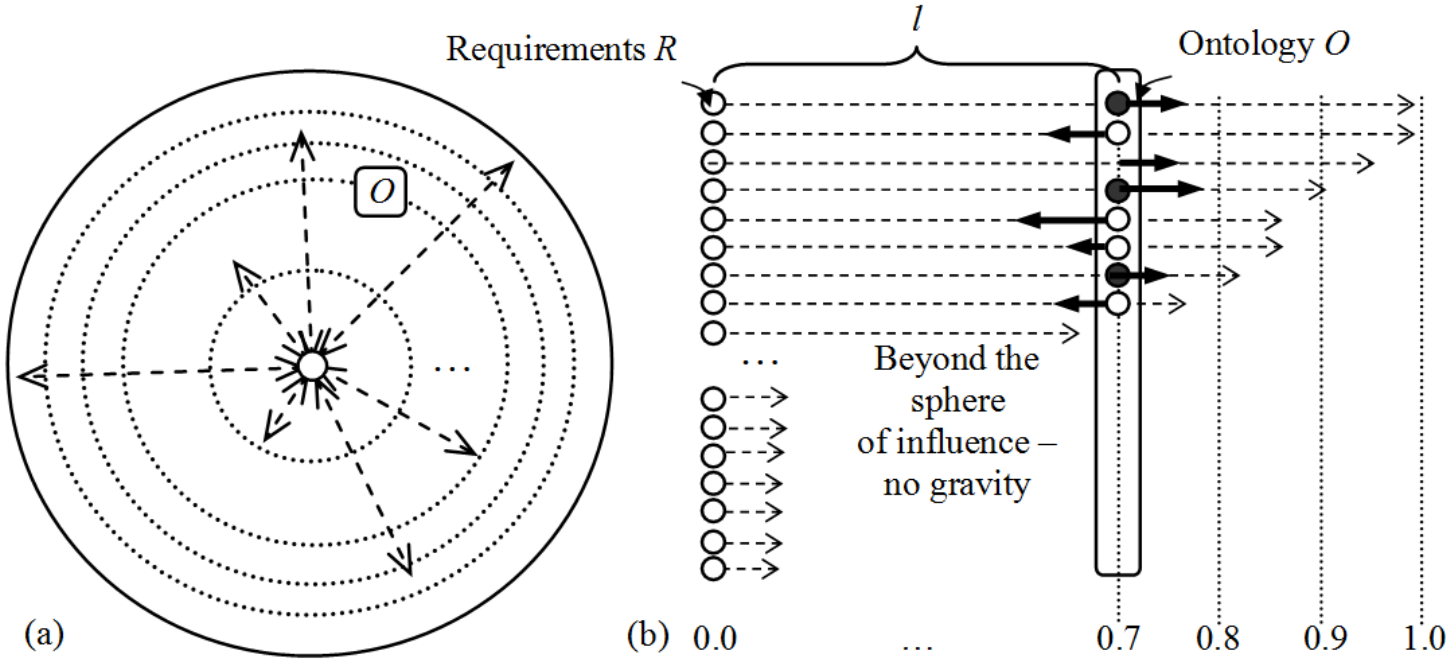
E.g. in a Time Domain* …

- The best – the most **fit** (Φ) to the requirements of the Domain Knowledge Stakeholders
  - How to <u>measure</u> and pragmatically <u>use</u> ontology **fitness**?



* Ermolayev, V., et al. (2014) Ontologies of Time: Review and Trends. *Int. J. of Computer Science & Applications*, 11(3), 57–115

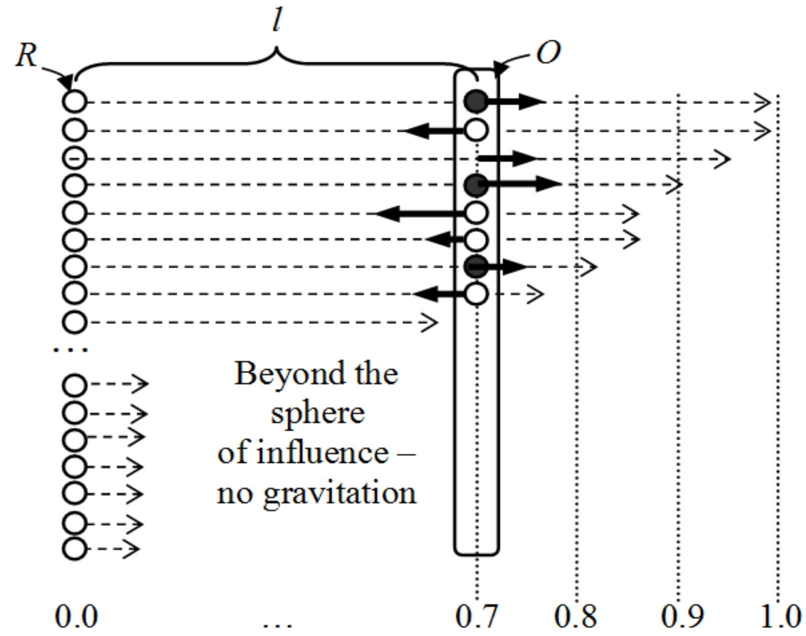
Req-s and Spheres of Influence



- All the requirements R are placed in the centre of the D
- They are not equal in their importance
  - Have different spheres of influence around the centre of gravitation, quantified using normalized scores ns ($[0,1]$) – *TF*-like metric
- O is positioned in D at a distance l from its centre ($[0,1]$)
  - 0 – perfectly fits; 1 – does not fit at all

A Gravitation Force

- O is checked against the $r \in R$
  - which spheres of influence reach the position of O: $ns_r \geq l$

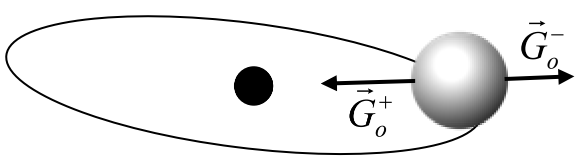


- Outcomes:
  - $o \in O$ meets r – increase in fitness

$$G_o^+ = \frac{1 \times \Delta\Phi_o^+}{(ns_r)^2}$$

  - o contradicts r or there is no o for the r – decrease in fitness

$$G_o^- = \frac{1 \times \Delta\Phi_o^-}{(ns_r)^2} \qquad G_O^- = \frac{1 \times \Delta\Phi_O^-}{(ns_r)^2}$$



- Overall gravitation force:

$$\left.\overrightarrow{G_O}\right|_D = \sum_{r \in R : ns_r \geq l} \left(\overrightarrow{G_o^+} + \overrightarrow{G_O^-} + \overrightarrow{G_o^-} \right)$$

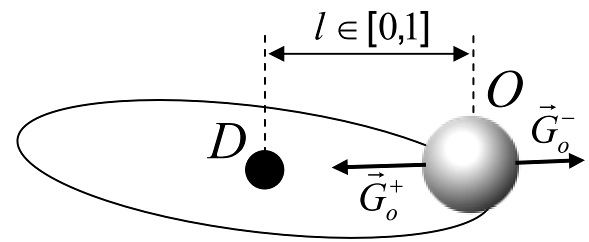
An Equilibrium State

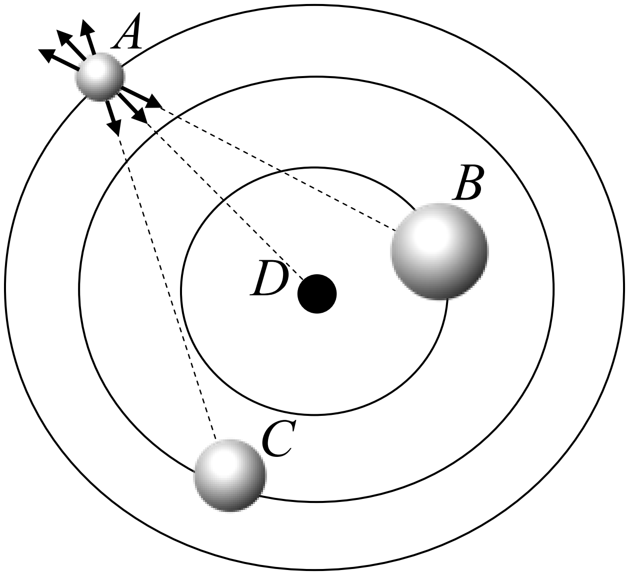
- What is the proper position of O in D ?
  - **Equilibrium state**
  - One ontology:
    - Zero sum of gravitation forces in the gravitation field of D

$$\left.\overrightarrow{G_O}\right|_D = \vec{0}$$

  - Several ontologies:
    - Also have gravitation between them

$$\left.\overrightarrow{G_A}\right|_D + \left.\overrightarrow{G_A}\right|_B + \left.\overrightarrow{G_A}\right|_C = \vec{0}$$

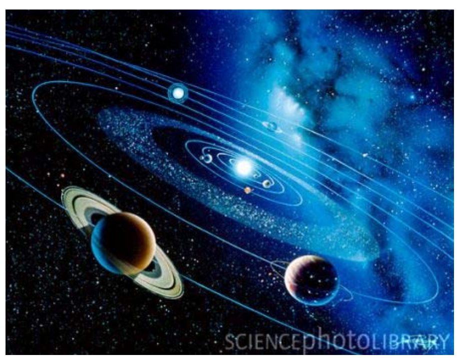




Compared to Mechanics

- Coordinate grid
  - E.g. Helio-centric, 3D, Decartes
- Distance
  - M, from point (0,0,0)
- Mass
  - Kg
- Force
  - Newton's Law of Universal Gravitation

$$F = \gamma \frac{m_1 \times m_2}{r^2}$$

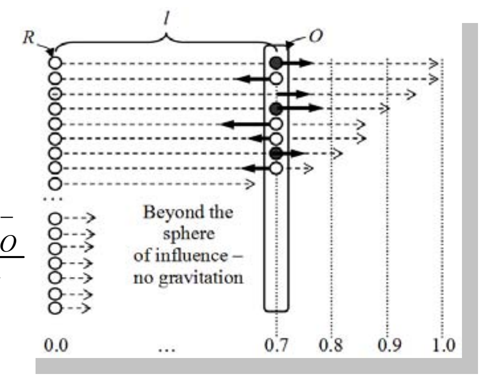


- To apply:
  - Measure r
  - Measure Masses

- Coordinate grid
  - Domain-centric, 2D, normalized
- Distance
  - Normalized, semantic: $l \in [0,1]$
- Mass
  - Fitness of O re R describing D
- Force
  - Fitness / semantic difference-based gravitation

$$G_o^+ = \frac{1 \times \Delta\Phi_o^+}{(ns_r)^2}$$

$$G_o^- = \frac{1 \times \Delta\Phi_o^-}{(ns_r)^2} \quad G_O^- = \frac{1 \times \Delta\Phi_O^-}{(ns_r)^2}$$



- To apply:
  - Measure ns
  - Compute Fitnesses

Supporting Instruments

- Now we have a model, so what?
- Do we have instruments to measure things and compute forces?
- Is the model adequate?

Where’s the Real Staff?

… Contributors to be mentioned …

The Real Staff: Workflow

- Elicit Domain Requirements
  - Multi-word key terms – with *ns*
  - Important key terms
  - Ensure that the set is statistically representative (completeness)
- Compute the Change in ontology Fitness
  - For elicited important requirements
    - Retrieve natural language definitions
    - Create structural contexts (ontology change tokens) – formal, explicit
    - Map structural contexts to the ontology
  - Compute votes
  - Compute the change in Fitness
- **Apply the Law of Gravitation**
  - To visualize the change in Fitness

E.g. in Time Domain

- Process is quite laborious, though
  - Could be automated in routine parts
  - Results could be re-used to a large extent

Elicit Requirements: Time*

- Domain: Temporal Representation and Reasoning
- Community: TIME Symposia series authors
- Document corpus: TIME Proceedings series, 1994-2013, ~440 papers, chronologically ordered
  - Incremental slices of the document corpus:

Slice ID	1994	1995	1996	...	2012	2013
D1						
D2						
...						
D19						
D20						

* Ermolayev, V., et al. (2014) Ontologies of Time: Review and Trends. *Int. J. of Computer Science & Applications*, 11(3), 57–115

For Each Incremental Slice

- Bag of terms extracted * and sorted by normalized term scores (ns)
- Termhood created by retaining
  - Valid terms – manual filter
  - Important terms – $ns > \varepsilon$ (such that the sum of ns above is a little higher than 50%)

- Termhood difference values computed using the THD algorithm **:
  - Absolute: $thd(T_{i-1}, T_i)$
  - Relative: $thdr = thd(T_{i-1}, T_i) / \sum_{T_i} ns_j^i$

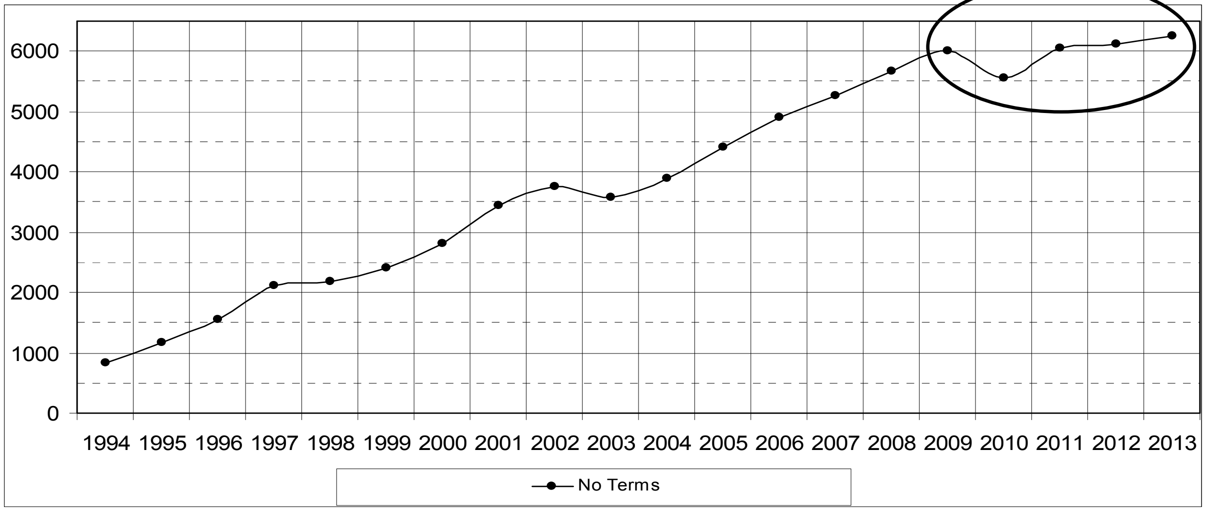
Software developed by Olga Tatarintseva as part of her PhD project

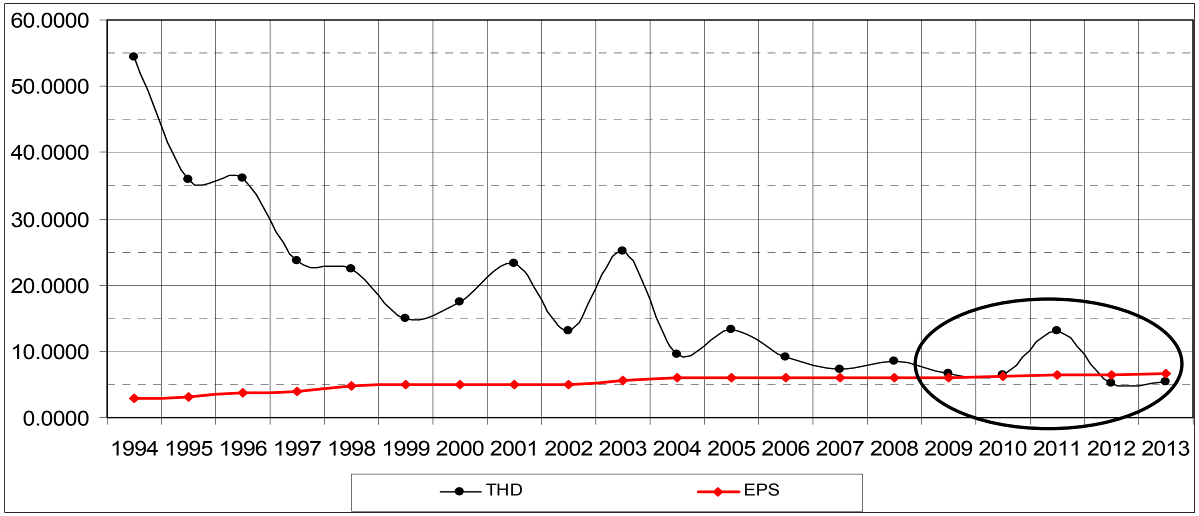
* Using TerMine service by the UK National Centre for Text Mining (NaCTeM, http://www.nactem.ac.uk/).

** Tatarintseva, O. et al. (2013) Quantifying Ontology Fitness in OntoElect Using Saturation- and Vote-Based Metrics. In: Ermolayev et al. (eds.) *ICT in Education, Research, and Industrial Applications. Revised Selected Papers of ICTERI 2013*, CCIS 412, pp. 136–162

Completeness Check

- Observed:
  - Saturation:
    - ~6,000 terms in the last 4 termhoods
    - *thd* below ε
  - Terminological drift
    - *thd* above 0
  - Terminology contribution peaks:
    - 2001, 2003, 2005, 2008, and 2011
  - The (representative) **majority vote**
  - Still too many terms retained



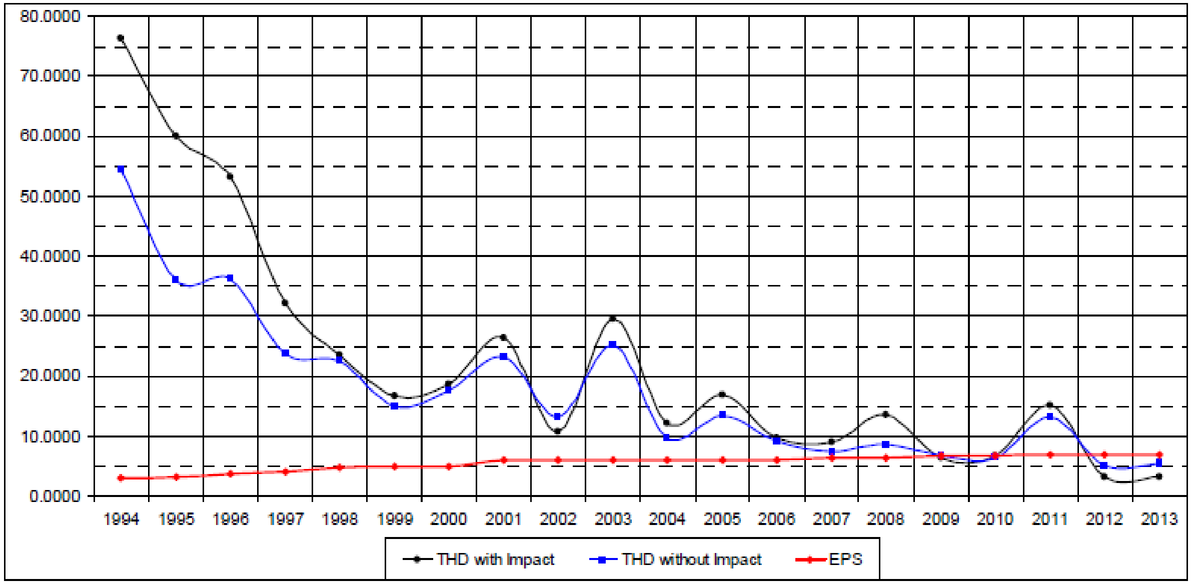


A Decisive Minority Vote

- Terminology contribution peaks: 2001, 2003, 2005, 2008, and 2011
- Account for **impact**:
  - Citation info collected (Google Scholar)
  - Paper impact computed based on citation frequency (cfr)

$$imp = \begin{cases} [0.2 \times cfr] + 1, cfr > 0 \\ 0, cfr = 0 \end{cases}$$

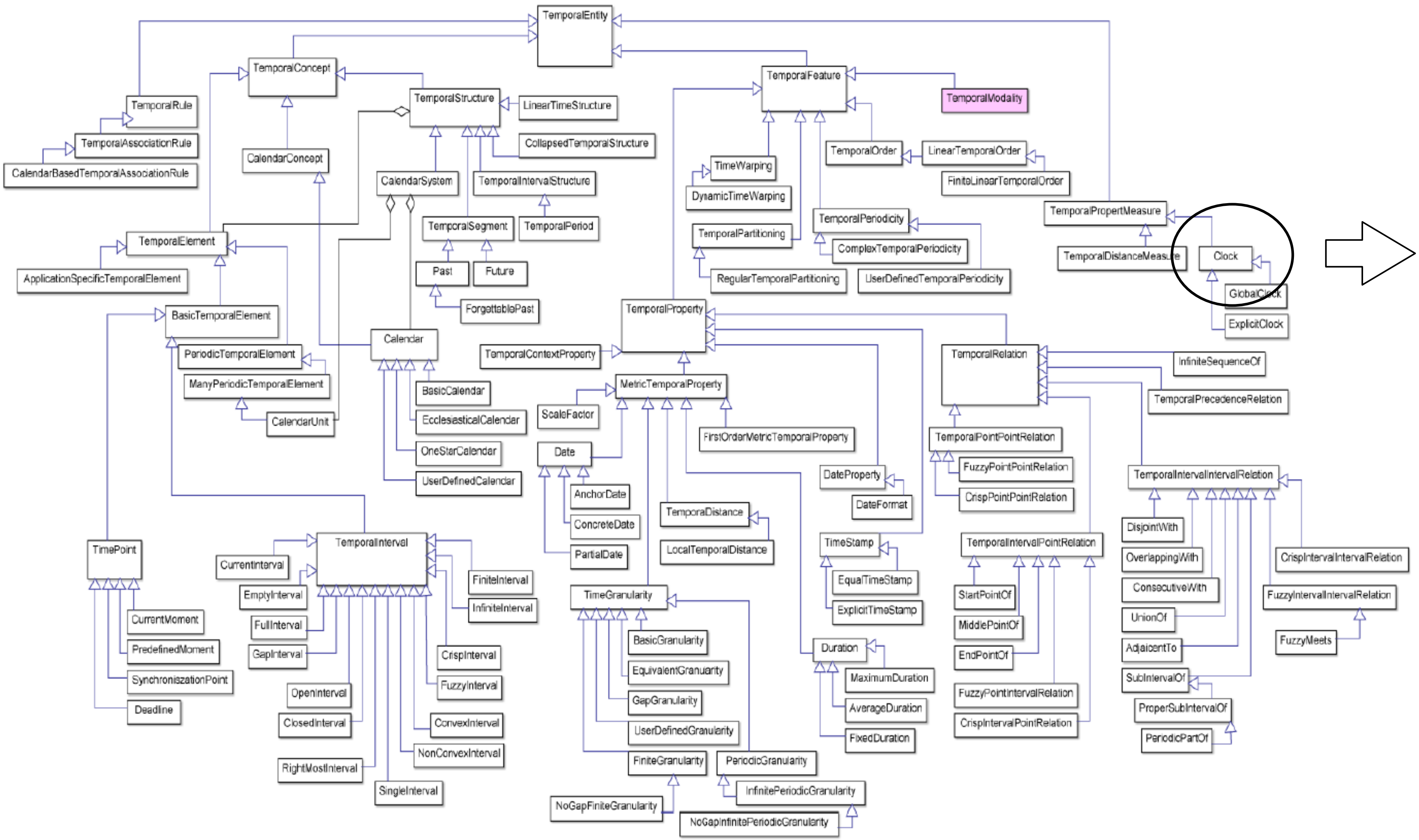
  - Papers with $imp = n$ replicated n times – changing the incremental slices
  - thd / $thdr$ / eps re-computed
- **Strong correlation**
- Termhood based on high-impact (**24**) papers only
- **686** Terms vs 6,109
  - The "**influence**" that triggers change



Classified Terms

Score	Term	Logic	Problem	Formula	Formalism	Operator	Method	Model	Reasoner	Domain	Language	Feature	Constraint	Instance	Pattern	Application	Project	Author
	Total No of terms: **686**	44	27	6	36	8	22	24	1	4	8	175	28	1	13	110	1	178
147.11	temporal logic	✓																
100.11	calendar pattern														✓			
86.54	temporal constraint												✓					
68.73	temporal operator					✓												
59.58	fuzzy match											✓						
52.25	temporal structure											✓						
49.83	calendar schema											✓						
46.25	temporal representation				✓													
41.00	temporal reasoning						✓											
40.00	freeze quantifier				✓													
37.73	fuzzy interval											✓						
36.36	xml document															✓		
36.00	crisp interval											✓						
34.00	satisfiability problem		✓															

Feature Taxonomy



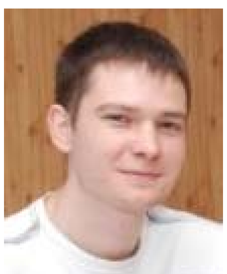
Natural Language Definitions

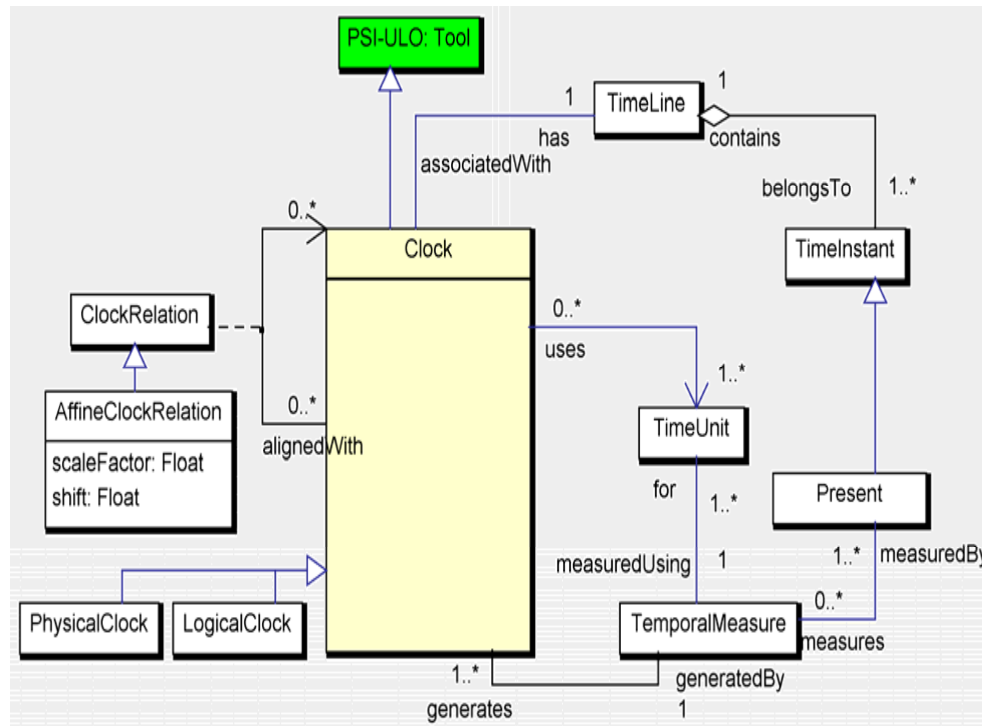
A **Clock** is an **Instrument** to generate the instances of a **TemporalMeasure** of a **TimeInstant** when this **TimeInstant** instance is also the instance of a **Present**. A **Clock** is always associated with a particular single **TimeLine** (there could be **TimeLines** with no **Clock** but also **TimeLines** having several different **Clocks** associated with them). Different **Clocks**, associated with the same **TimeLine** or different **TimeLines** may “run” differently, e.g. quicker or slower and also with offsets compared to each other. Some **Clocks** may be related to each other for enabling a proper comparison of the values they return. This is done by specifying a **ClockRelation**. A specific and most widely used kind of a **ClockRelation** is **AffineClockRelation** which allows aligning different time velocities (using the scaleFactor property) and also time offsets, like delays (using the shift property). A **Clock**, as a measurement instrument, may return a single value (a **TimeStamp** corresponding to a single **TimeUnit**) or several values (the parts of a **TimeStamp** corresponding to different **TimeUnits**). A **PhysicalClock** and a **LogicalClock** are the two disjoint specializations of a **Clock**.

Structural Contexts

Ontology (Change) Tokens (r from R)

A **Clock** is an **Instrument** to generate the instances of a **TemporalMeasure** of a **TimeInstant** when this **TimeInstant** instance is also the instance of a **Present**. A **Clock** is always associated with a particular single **TimeLine** (there could be **TimeLines** with no **Clock** but also **TimeLines** having several different **Clocks** associated with them). Different **Clocks**, associated with the same **TimeLine** or different **TimeLines** may "run" differently, e.g. quicker or slower and also with offsets compared to each other. Some **Clocks** may be related to each other for enabling a proper comparison of the values they return. This is done by specifying a **ClockRelation**. A specific and most widely used kind of a **ClockRelation** is **AffineClockRelation** which allows aligning different time velocities (using the scaleFactor property) and also time offsets, like delays (using the shift property). A **Clock**, as a measurement instrument, may return a single value (a **TimeStamp** corresponding to a single **TimeUnit**) or several values (the parts of a **TimeStamp** corresponding to different **TimeUnits**). A **PhysicalClock** and a **LogicalClock** are the two disjoint specializations of a **Clock**.



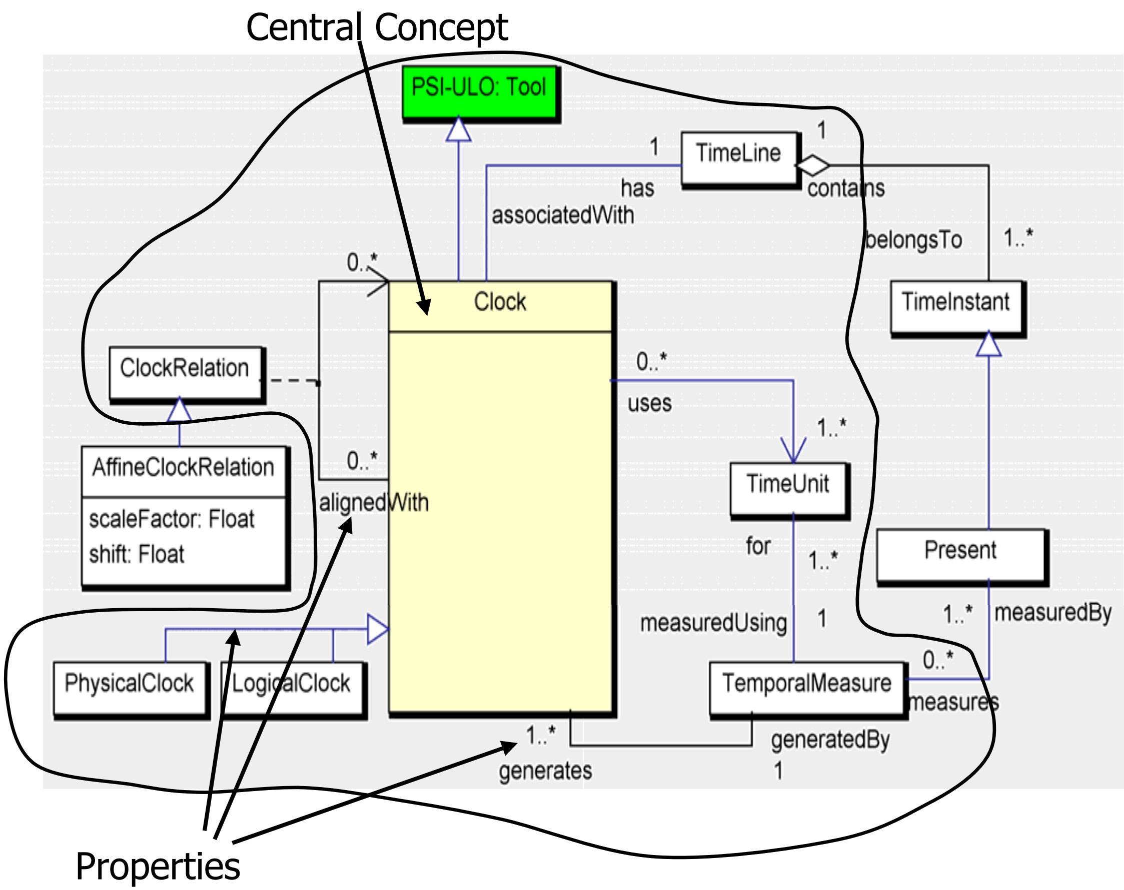


ongoing work by Eugene Alferov as part of his PhD project

+ OWL

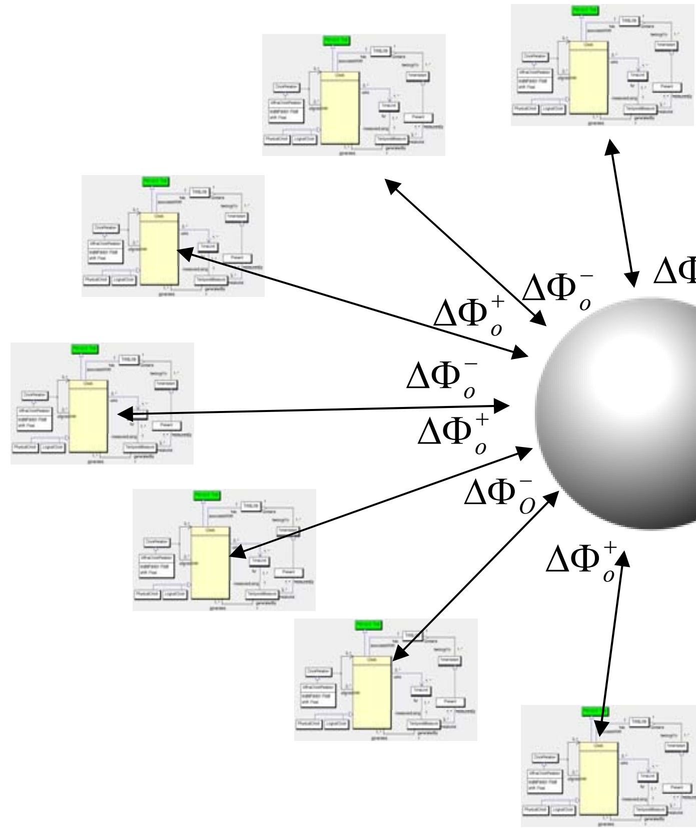
Ontology (Change) Tokens

- To be verified
  - Manual – knowledge engineer
  - Perhaps, refined
- To be mapped to the O
  - Auto – SDDE
  - Verified – ODV



Mapping *r*-s to Ontology

- Structural Difference Discovery Engine (SDDE)*
  - Agent-based software
- Meaning negotiation **
  - Type theory
  - Propositional substitutions
  - Generates mappings
    - Ontology alignment format
    - ABox transformation rules (instance migration)



Developed by Maxim Davidovsky

as part of his PhD project

* Davidovsky, M., Ermolayev, V., Tolok, V.: Agent-Based Implementation for the Discovery of Structural Difference in OWL-DL Ontologies. In: H.C. Mayr et al. (Eds.): UNISCON 2012, LNBIP 137, pp. 87–95 (2013)

** Ermolayev,V., Keberle, N., Matzke, W.-E., Vladimirov, V.: A Strategy for Automated Meaning Negotiation in Distributed Information Retrieval. In: Y. Gil et al. (Eds.): ISWC 2005, LNCS 3729, pp. 201 – 215, 2005

A Way to Verify
Ontology Difference Visualizer

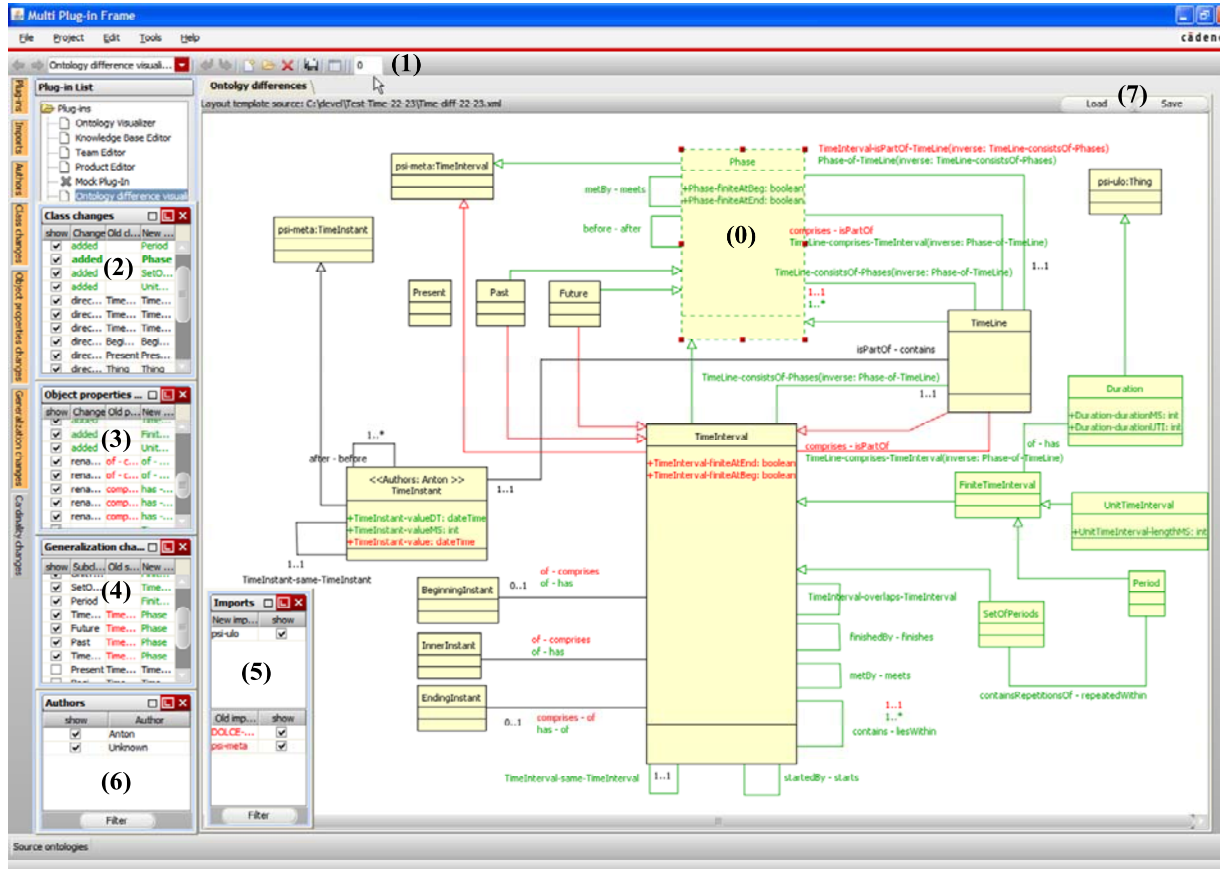
Structural Context:

(1) Neighborhood radius

Fine-tuning by inclusion or exclusion of:
(2) Concepts
(3) Object properties
(4) Subsumptions
(5) Module imports

Personification:
(6) Author context

(7) Saved as Layouts



Developed for Cadence Design Systems Inc. by myself and Anton Copylov

* Ermolayev, V., et al. (2010) Using Contexts in Ontology Structural Change Analysis.. In: Ermolayev, V., Gomez-Perez, J.-M., Haase, P., Warren, P, (eds.) CIAO 2010, CEUR-WS, vol. 626 (2010)

Compute Change in Fitness

- Token mappings: $\mu = (t, r, o, cf)$
  - *t* – ontology token (central concept, properties)
  - *r* – relationship {*equivalence*, *membership*, *subsumption*, *meronymy*, *association*}
  - *o* – ontology element
  - *cf* – confidence factor
- Positive votes: $v_o = ns \times w(r) \times cf$
  - *ns* – normalized score of the corresponding term (central concept)
  - *w*(*r*) – mapping relationship type weight
  - *cf* – mapping *cf*
- Propagated votes: $v_o^p = att \times v_{o^{sub}}$
  - Reflect the contribution of *o* to the semantics of the ontology element o^{sub} subsumed by o
  - *att* – attenuation coeff, chosen empirically
- Negative votes: $v_{t_i}^- = -ns_i$
  - No mapping – missing in *O* or contradicts to some *o*
- Change in **Fitness**

$$\Delta\Phi_O = \sum_{o \in O} v_o + \sum_{o \in O} v_o^p + \sum_{t \in T^{miss}} v_t^-$$

Part of Olga Tatarintseva's PhD project

Is the Approach Adequate?

- Not fully evaluated so far – ongoing work
  - Indications that the components work adequately:
    - OntoElect
    - Structural Difference Discovery Engine
    - Ontology Difference Visualizer

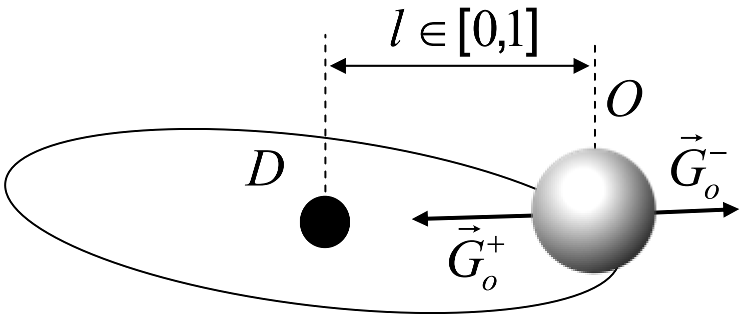
- A handful of instruments is in place
  - Not yet complete
  - Some have to be put together
  - Some have to be finalized – Def–to–UML+OWL
  - **OntoGrav** – visualizer to be implemented as a proof of concept
- Use cases have to be elaborated for validation

Potential Use Cases

- Even so, how could these be used?
- Would these make our lives better?

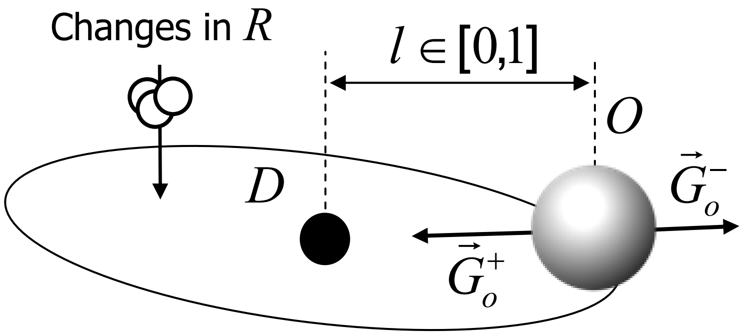
Ontology Refinement

- Implementation of the required changes in an ontology for making it better fit the changed stakeholder requirements
  - The equilibrium state before the changes in the requirements came



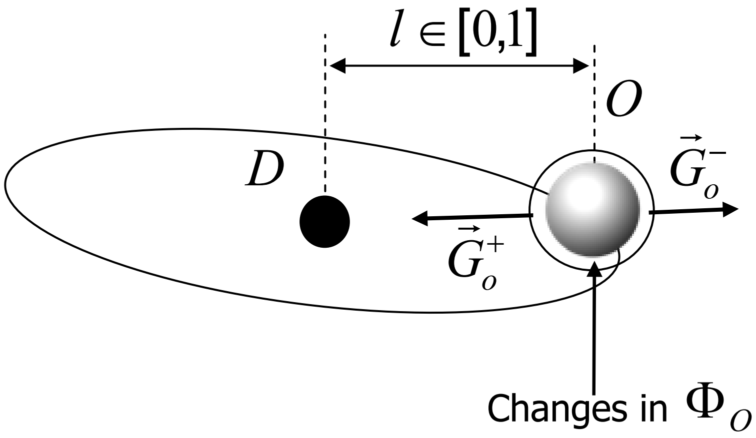
Ontology Refinement

- Implementation of the required changes in an ontology for making it better fit the changed stakeholder requirements
  - Changes in the requirements R cause the changes to the gravitation field generated by D



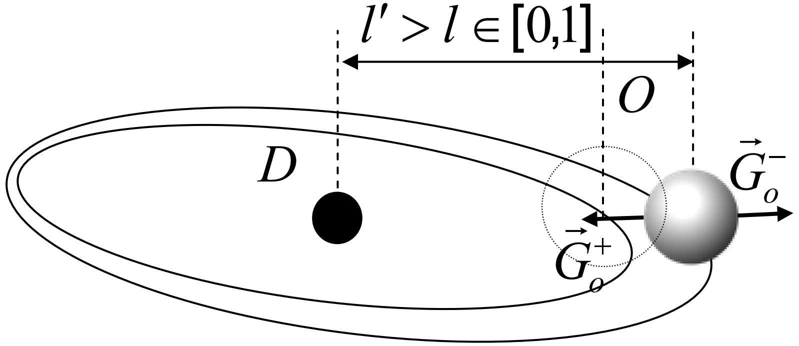
Ontology Refinement

- Implementation of the required changes in an ontology for making it better fit the changed stakeholder requirements
  - … and also the changes in the fitness of O
    - As presented before



Ontology Refinement

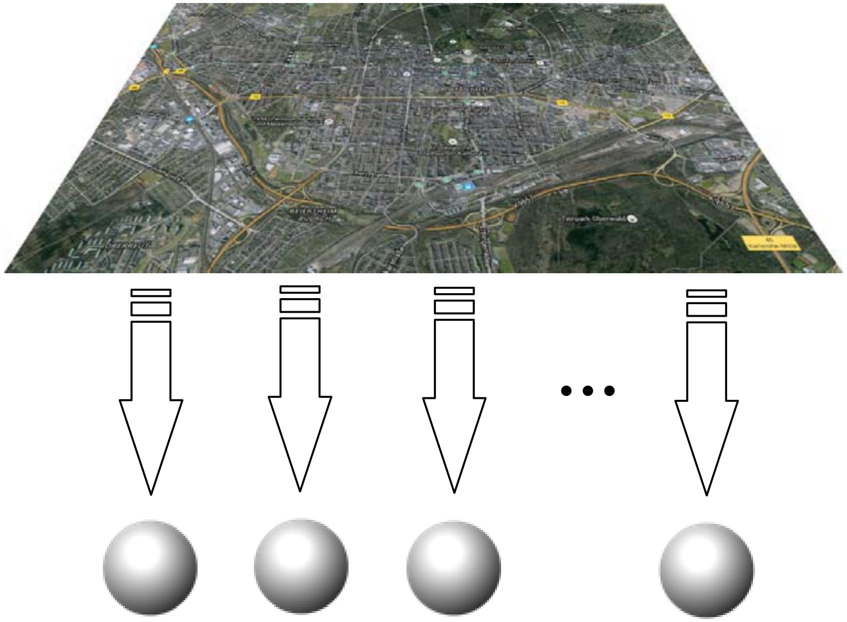
- Implementation of the <u>required changes</u> in an ontology for making it better fit the changed stakeholder requirements
  - O changes its position in D – finding a new equilibrium state
    - Could be many O –s
  -  A visualization tool (**OntoGrav**) development planned



New **equilibrium** state

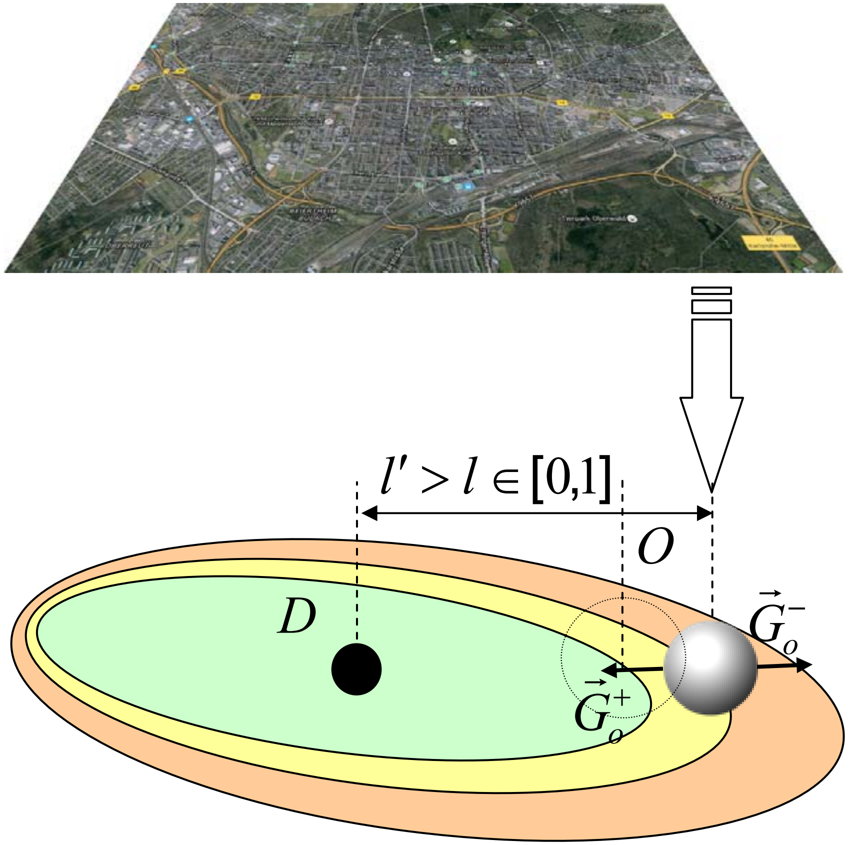
Anomaly Detection

- Reveal the parts of data that change beyond normal values
  - Hint about a potential or developing problem
- E.g., if a system is a civil community and its environment (**D**)
  - Requirements: comfort and safety of life
    - R_i for sub-domain D_i
      - E.g. certain minimal level of air pollution
  - Different streams of observation data (sensors), e.g.:
    - Outdoor temperatures
    - Water levels
    - Industrial/traffic emissions
    - Share prices
    - Cell phone activity
  - Populate respective ontologies O_i
    - Knowledge extraction
    - Causing ontology change



Anomaly Detection

- Static requirements
  - Normal
- Changing ontologies
  - Less fitness
  - More anomaly in the change



Concluding Remarks

Some Concluding Remarks

- Ontology Change is a hot topic
  - Clames Dynamics, but
  - Lacks Dynamics –
    - Responds to **needs** (Kinematics)
    - Does not analyze **causes** – an open issue
- Analogy in Mechanics
  - Newton's law of Universal Gravitation
  - <u>Fitness-based law of gravitation</u> in a Domain – Ontology system
    - Shows how new requirements **cause** fitness change
- Instruments that may help measure the change in fitness
  - Extract (change) requirements
  - Transform to knowledge tokens
  - Map to ontology and verify mappings
  - Compute fitnesses
- Use cases of potential interest in:
  - Ontology Refinement
  - Anomaly Detection

Will be happy to answer your questions ...

Will be also happy to continue discussions

vadim@ermolayev.com

Some Relevant Questions

- Are(n't) ontologies really(rarely) useful? Aren't these just data structures by another name?

- What if requirements come in different languages? Would the Law of Multilingual Gravitation be different?

- Is the saturation indicator of OntoElect valid? Would any arbitrary collection of documents "saturate" if big enough?

- Can we trust automatically generated mappings?

Just in Case …

Analogies and Insights …

- **Mechanics:**
  - **Dynamics** is concerned with the study of **forces** and **torques** and their **effect on motion**
  - **Motion**: **change** of position

- **System Dynamics** studies the behaviour of (complex) systems over time
  - The behaviour of the entire system is affected by internal **feedback (causal) loops** and time delays

Analogies and Insights …

- **Population Dynamics** studies:
  - Short- and long-term changes in the **size and age composition** of populations
    - E.g. ageing or population decline
  - Biological and environmental processes influencing those changes
  - **Deals with** the way populations are **affected by**:
    - Birth and death rates
    - Immigration and emigration

What is an Ontology?

- **Ontology** (in AI and Information Science):
  - A "formal, explicit specification of a shared conceptualization" *
- A descriptive theory for:
  - A Domain of Discourse
  - Common Sense
  - …
- To explain:
  - Conceptualization?
  - Specification of a Conceptualization?
  - Explicit Specification?
  - Formal Specification?
  - Shared Conceptualization?

* Rudi Studer, V. Richard Benjamins, and Dieter Fensel. Knowledge Engineering: Principles and Methods. *Data & Knowledge Engineering*, 25(1-2):161–197, 1998

Conceptualization:

- "an <u>abstract, simplified view</u> of the world that we wish to represent for some purpose " *
  - A **model** ...
- A **model** comprising "the <u>objects</u>, <u>concepts</u>, and other <u>entities</u> that are presumed to exist in some area of interest and the <u>relations</u> that hold [between] them" **
  - **Relation**s – also properties
- **Relation**s are understood in an intentional (conceptual) sense – i.e. focusing on their meaning, independent of a particular state of affairs ***

* Gruber, T. R. (1993) A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition* 5 (2): 199–220, doi:10.1006/knac.1993.1008

** Genesereth, M. R., & Nilsson, N. J. (1987) *Logical Foundations of Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann Publishers

*** Guarino, N. (1998) Formal Ontology in Information Systems. In: Proc FOIS'98, IOS Press, pp. 3-15

Is this a Conceptualization?

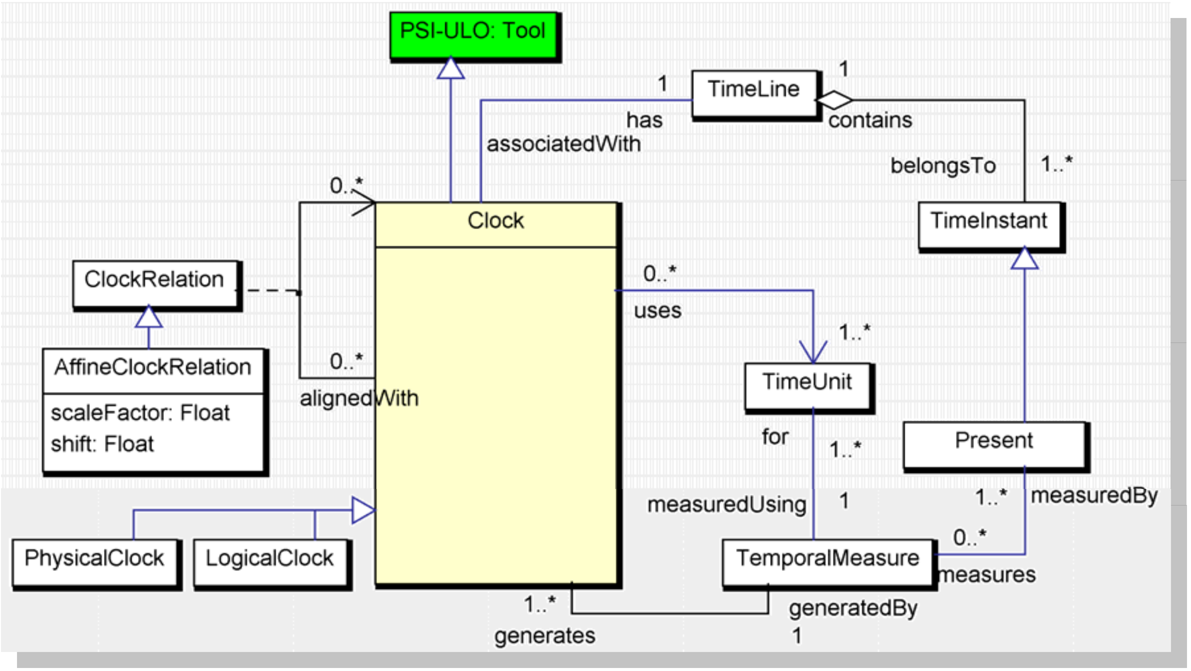
A **Clock** is an **Instrument** <u>to generate</u> the instances of a **TemporalMeasure** of a **TimeInstant** when this **TimeInstant** instance is also the instance of a **Present**. A **Clock** is always <u>associated with</u> a particular single **TimeLine** (there could be **TimeLines** with no **Clock** but also **TimeLines** having several different **Clocks** associated with them). Different **Clocks**, associated with the same **TimeLine** or different **TimeLines** may "<u>run</u>" differently, e.g. <u>quicker</u> or <u>slower</u> and also <u>with offsets</u> compared to each other
...

- **Concepts**? – in CaMel notation, bold
- **Objects**? – the instances of Concepts
- **Intentional Relations**? – underlined, properties in fact

• Is this a **Specification** of a Conceptualization?

Specification of …

- A Specification is "a detailed precise presentation of something or of a plan or proposal for something" *

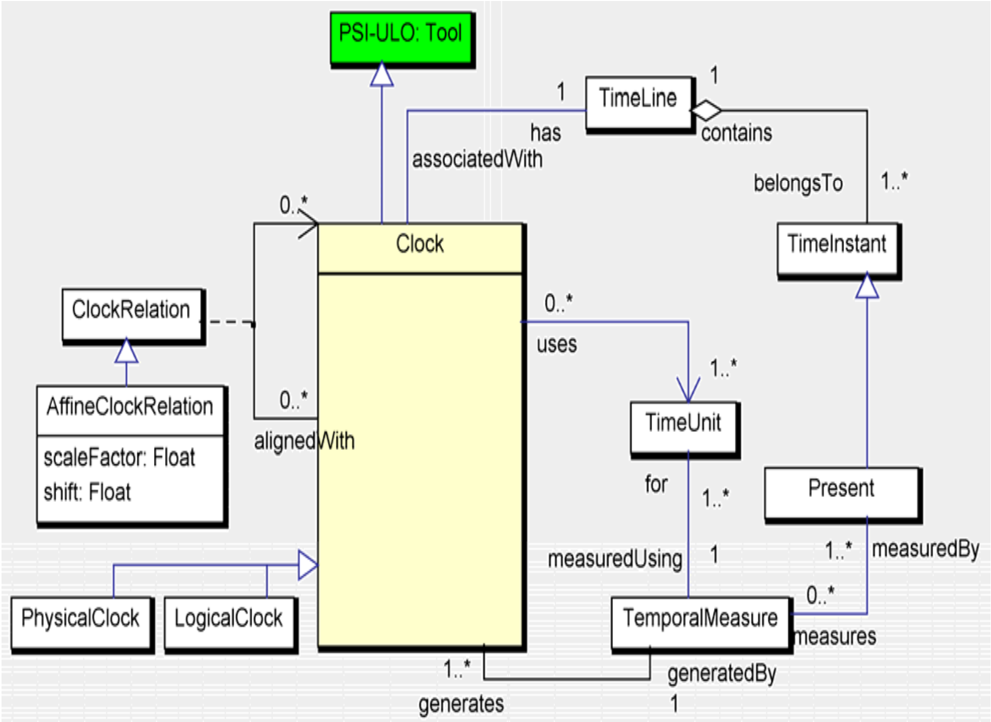


- Is this a Specification of the (above) Conceptualization?
  - Detail? Precision? --> … Explicit …

* Merriam – Webster Dictionary: http://www.merriam-webster.com/dictionary/specification

Explicit Specification of …

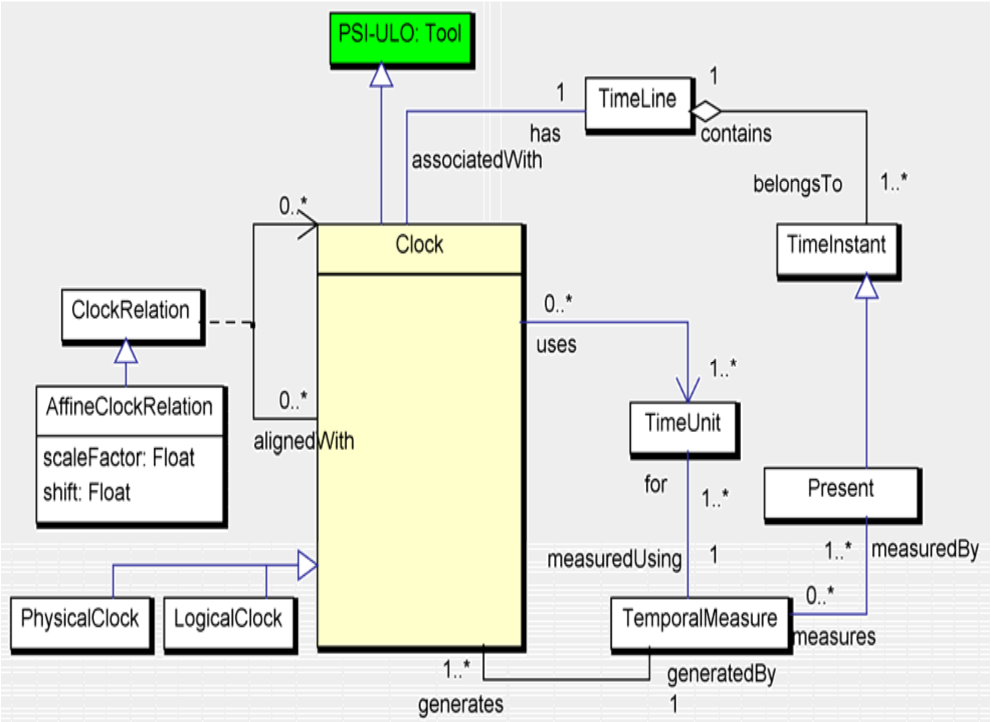
- Explicit: "very clear and complete, leaving no doubt about the meaning" *
  - Clear:
    - **Granular** enough to reveal the relevant bits of meaning in a way not allowing misinterpretation
    - Free of contradictions
  - Complete:
    - Providing **all important** bits of meaning



* Merriam – Webster Dictionary: http://www.merriam-webster.com/dictionary/explicit

Formal Specification of …

- Using a formally defined specification language
- For ontologies:
  - Such that logical inference is enabled
- Req of being Formal ->
  - Ontology is a **Logical Theory**



- Is that UML spec. Formal?
  - Ontology Specification Languages (such as OWL) have been developed for making ontologies formal

Ontology is a Logical Theory

- **Ontology *** – a tuple:

$$O = (C, P, I, T, V, \leq, \perp, \in, =)$$

- where the sets C, P, I, T, V are pair-wise disjoint and:
  - C – set of **concepts** (or classes)
  - P – set of **properties** (object and datatype properties)
  - I – set of **individuals** (or instances)
  - T – set of **datatypes**
  - V – set of **values**
  - $\leq$ – reflexive, anti-symmetric, and transitive relation on $(C \times C) \cup (P \times P) \cup (T \times T)$ called **specialization**, (**subsumption**) that form partial orders on:
    - C – concept hierarchy; and
    - P – property hierarchy
  - $\perp$ – irreflexive and symmetric relation on $(C \times C) \cup (P \times P) \cup (T \times T)$ called **exclusion**
  - $\in$ – relation over $(I \times C) \cup (I \times V)$ called **instantiation**
  - $=$ – relation over $I \cup P \cup (I \times V)$ called **assignment**

***** Euzenat J. and Shvaiko P. 2007. *Ontology Matching*, Berlin Heidelberg (DE), Springer-Verlag

Ontology: TBox and ABox

- An ontology comprises its schema and the assertional part:

$$O=(S, A)$$

- Ontology Schema S (terminological component – TBox) contains the statements describing:
  - The concepts of O
  - The properties of those concepts, and
  - The axioms over the schema constituents
- The set of individuals A (assertional component – ABox) contains the ground statements about:
  - The individuals, and
  - Their attribution to the schema

Problems with Requirements

- Acquiring a complete and accurate collection of Domain knowledge (requirements) is difficult:
  - **K**-s are subjective
  - **K**-s are tacit
  - **K**-s are partial
  - **K**-s are hard to get / not available
  - **K**-s specs are rarely explicit and formal
    - Contradictory interpretations
  - *unf* – challenging – expressive power