CAPTURING NORMATIVE CONSTRAINTS BY USE OF THE SEMANTIC MARK-UP RASE METHODOLOGY

Eilif Hjelseth, Ph.D. student, <u>eilif.hjelseth@umb.no</u> *UMB / Department of Mathematical Sciences and Technology, IMT, Norway*Nick Nisbet, Director, <u>nn@aec3.com</u> *AEC3 Ltd., UK*

ABSTRACT

The AEC industry is highly regulated by a large number of rules given by public laws, codes, and regulative standards at both national and international levels. The relevant information in these documents need to be captured as rules for model checking in a time and cost effective way. The foundation for the RASE concept is using mark-up based on the four operators; requirement (R), applicabilities (A), selection (S) and exceptions (E) on normative text. The RASE technology has been tested on following three categories of documents: standard (case: NS 11001-1.E:2009 Universal design of building constructions - Part 1: Work buildings and buildings open to the public), standards with tables (Dubai regulations) and guidelines (case: GSA court design guidance document, USA). In each case expectations have been documented using free prose. On examination, the key clauses and phrases can be identified along with their role, allowing a testable, logical statement to be generated. The logical statement is then ready to be used by a compliance-checking engine to apply tests to a description of the facility. The results indicate that the RASE methodology can operate on a different types of normative documents with a trustworthy results.

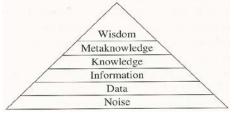
Keywords: Knowledge representation, Ontology, Semantics, Model checking, Decision support.

1.1 **Problem status**

The AEC/FM (architects, engineers, contractors, facility management) industry is highly regulated by a large number of rules given by public laws, codes, and regulative standards at both national and international levels. The question is how these normative documents can be captured as rules for automatic or semi-automatic model checking in a time and cost effective development and implementation into software. We see from the history of AI (artificial intelligence) and KBE (knowledge based engineering) and KM (knowledge management) that capturing information in a meaningful and trustworthy way has not been a straight-forward process, as the content increases (Russell and Norvig, 2010). We conclude that a wide spread use has not been reached and the expectations regarding semantic web has not moved forward as expected.

1.2 Principles for capturing knowledge

However, if we take a more direct and purpose driven approach, trustworthy results are possible. Hjelseth and Nisbet (2010) verified the fundamental principles for the four logical operator based RASE concept. We will in this paper capture from three different type of sources which in the next stage can be used in BIM based model checking systems. The RASE concept will according to figure 1 contribute with an alteration from information to metaknowlege.



Wisdom: Using knowledge in a beneficial way Metaknowledge: Rules about knowledge Knowledge: Rules about using information Information: Potentially useful for knowledge

Data: Potentially useful information **Noise:** No apparent information

Figure 1: The pyramid of knowledge. (Giarratano and Riley, 2005).

"Human knowledge is a process of approximation. In the focus of experience, there is comparative clarity. But the discrimination of this clarity leads into the penumbral background. There are always questions left over. The problem is to discriminate exactly what we know vaguely."

(Whitehead, 1937)

RASE is a semantic based concept for transforming normative documents into a singel well-defined rule which can be implemented into BIM / IFC based model checking software. The increased use of BIM enables the rules to apply in the information already captured in BIM models from many different design applications. Execution / implementation of rules in a model checker demands an unambiguous interpretation. We will in the following part will illustrate some general fundamental problems in text analysis and demonstrate that these can be solved or evaded by use of the RASE methodology. Our experiences indicate that RASE can also contribute to clearing up professional understanding of normative documents as standards, building codes etc. This will also influence on the challenge with performance based vs prescriptive defined regulations.

1.3 Knowledge soup

Knowledge soup is a concept develop by John F. Sowa (2000, 2004). It states that the complexity does not arise from the way the human brain works or the way that natural languages express information but rather from over-generalizations, abnormal conditions, incomplete definitions, conflicting defaults and unanticipated applications. These exceptions and borderline cases result from the nature of the world, not from any defect in natural language. We present in Chapter 2.2 three principles; Translate, Transform and Transfer for handing of these issues (Hjelseth, 2009). Tarski (1944) states that within a delimited domain, it is possible to achieve consistency and precise understanding. We are operating on normative documents within the building and construction domain.

1.4 **Ontology**

There are many definitions for "ontology". Schalkoff (2011) defines ontology as: "An ontology is a formal characterization of concepts in a domain of discourse". Some of the reasons to develop and ontology is to capture domain knowledge. Further can it be used to visualize, manipulate, reuse, and update or extend the representation. This may lead to an alternative, practical, and functional definition for an ontology: a way for a community to agree on common terms and structure for representing knowledge in a domain. A RASE rule can be considered as an exchange requirement. In the current example, both the rules model and the facility model are represented in the IFC schema. This approach has allowed applications such as Jotne EDM, Solibri and Singapore ePlanCheck to be adapted to use the rule model, re-using the IFC read/write capabilities already supported.

The IFC Product sub-schema represents the facility as a highly interrelated network of objects, such as walls, ramps and spaces. In contrast, the IFC Constraint sub-schema represents the logical structure of a normative document as a strict hierarchy of objectives. At the highest level, the primary objective is the satisfaction of the whole. This objective is then defined in terms of subsidiary objectives relating to distinct topics and their expression as requirements, exceptions, applicability and selection. At the lowest level the constraints are discrete metrics which carry testable expressions.

The constraint model is therefore a highly structured representation of the original text. However, the logical structure and each discrete metric have a direct line of derivation from the text and mark-up. This ensures that constraint model is recognizably and demonstrably correct. A reverse mapping can also be applied to the constraint model to construct free prose, though currently the re-constituted text is stylistically monotonous.

2.1 **RASE theory**

It is a characteristic of regulations that every 'check' must be in some way satisfied. The most obvious and most easily identified are the 'requirements' as these are associated with the future imperatives 'shall' or 'must' 'requirements'. It is required that a check contains at least one 'requirement'. Secondly, there will be text that identifies the 'applicability' of the check. These are often

compounded, for example 'external windows'. These phrases need not relate directly to the topic of the regulation or the topic of the overall check. For example, if a check applies in 'a seismic zone', this is a property of the building site, not of the structural integrity of a particular building material. In general, there will be one or more phrases defining the applicability. One special but distinct case is where a 'selection' of alternative subjects or more 'exceptions'. These are the opposite of 'applicability', and conversely work by exclusion. (Nisbet, Wix and Conover, 2008). The RASE markup language uses the following four RASE operators: 'requirement' 'applies, 'select', and 'exception'. Applied on a text, the user highlights any clause or phrase that means:

- 'shall'/'must' as a 'requirement', (including alternative requirements)
- less scope as an 'applies'
- more scope as a 'select'
- 'unless' as an 'exception', (including composite exceptions).

The relation between the operators and the original building codes in text is made apparent by a colour system according to the mark-up language. An example of the mark-up operators and its related colours is software is illustrated in figure 2. The colours were chosen for acceptability for those with visual impairments.

	S		0
Requirement	Applies	Select	Exception
{blue}	{green}	{ <i>red</i> }	{orange}

Figure 2: The four RASE operators for rule development (Hjelseth and Nisbet, 2010a).

The four RASE operators used for marking up text can be visualized by colours related to the operators, e.g. blue, green, red and orange. This gives the user an instant overview of what and how the rules are structured. The naming of the operator is chosen to correspond with the way standards, codes, regulative are written, which reflects natural language.

The marked-up Requirement (R), Applicabilities (A), Selection (S) and Exceptions (E) clauses will contain phrases. The four types of phrases can be identically attributed to have a topic, a property, a comparator and a target value. The topic and property are ideally be drawn from a restricted dictionary composed of terms defined within the regulation and normal practice. The value (with any unit) may be numeric, whereupon the comparators will include 'greater', 'lesser', 'equal' and their converses. If the value is descriptive, then only the 'equal' or 'not equal' comparators are relevant. If the value represents a set of objects, then the comparator may be any of the set comparison operators such as 'includes', 'excludes' (Hjelseth and Nisbet, 2010a and Nisbet, Wix and Conover, 2008).

2.2 Methods for "normative-isation" of text.

2.2.1 General precondition

Developments of rules are based on interpretation of normative text. The RASE concept is optimized for this purpose. For handling of variation in text, we introduce three principles; Translate, Transform and Transfer for handing of semantic issues.

2.2.2 Translate

When presented with a normative text with clear metrics, one will get at direct match from original text into mark-up with the "RASE-operators" (Hjelseth, 2009).

2.2.3 Transform

However, not all code related text, even standards is suitable for this. Where the individual metrics are ill-defined, or the text is poorly drafted, then the process of mark-up will expose this. Unstructured or "blurry" text can be reformulated based on its intentions. If the text is redrafted, then the mark-up process can continue, but the results will be based on the re-formulation. The change has occurred in the source text, not in the transformation process. This ambiguity is particularly common in Guidance documents, one of the four types or model checking concepts stated by Hjelseth and Nisbet (2010 b), where the designer is expected to consider alternatives and preferences outside of the text. These rules can also be source for "Guidance based" model checking by presenting options and advice. The intention by explicit use transformation of the text source is to increase the number of rules which can be extracted for the original text. Because RASE give a explicit documentation, this will be transparent when used in model checker software.

2.2.4 Transfer

However, sometimes the original text is formulated in way where transformation does not give a trustworthy result in a model checker. This will sometimes be the case with general parts or statement of general principles on high level: the solution should be environmentally friendly, user friendly etc. It is important that these aspects of the original document are identified and information about this is presented for the user so it can be interpreted manually by a professional (Hjelseth, 2009).

2.3 **RASE technology**

RASE is based on utilization four operators defining a predefined and uniform action. This enables automatic /semi-automatic transformation into software code. The technology presented in this paper transform the mark-up from the operators into html-tags. The software is developed by Nick Nisbet in AEC3 Ltd and consists of a graphical interface, and the transform to a structured constraint model.

3 TEST APPROACH AND METHODS

The tests were applied to a normative regulation using prose, a normative regulations using a mixture of prose and tables and finally to any advisory 'guideline' document. In the examples, the original text is shown, followed by the mark-up. We then give a summary of the testable metrics and show the result of re-constituting the text from the logical statement, as evidence that the interpretation is correct.

3.1 Test Case: Norwegian accessibility standard, NS 11001-1:2009

Many jurisdictions in Europe and North America have introduced accessibility standards, with their focus on the accommodation of users with of ambulatory disabilities. An example clause which considers both spaces and solid building fabric is considered here. We use Clause 5.2 from "NS 11001-1.E:2009 Universal design of building constructions - Part 1: Work buildings and buildings open to the public" as an example:

5.2 Dimensioning an access route to a building

The access route for pedestrians/wheelchair users shall not be steeper than 1:20. For distances of less than 3 metres, it may be steeper, but not more than 1:12.

The access route shall have clear width of a minimum of 1,8 m and obstacles shall be placed so that they do not reduce that width. Maximum cross fall shall be 2 %.

The access route shall have a horizontal landing at the start and end of the incline, plus a horizontal landing for every 0,6 m of incline. The landing shall be a minimum of 1,6 m deep. Minimum clear height shall be 2,25 m for the full width of the defined walking zone of the entire access route including crossing points.

Figure 3: The regulation

In the original text, it may not be obvious that the overall applicability expressed in the title carries down into following sentences. The second sentence is actually an exception to the first. The mark-up is shown as seen in the mark-up application in figure 4 and in a short form in figure 5.

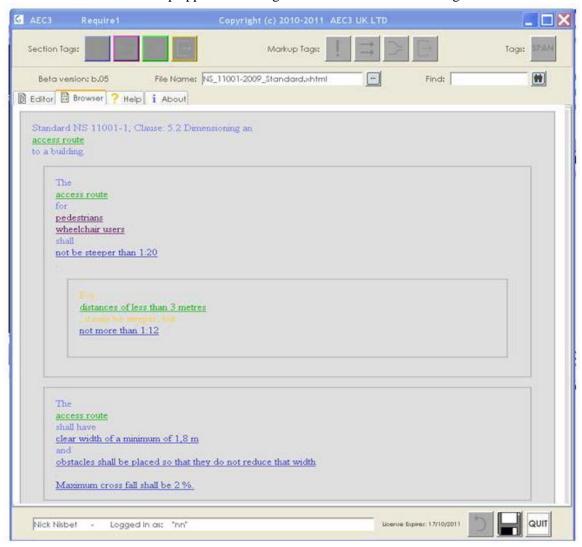


Figure 4: RASE mark-up by use of Require1 application. © AEC3 UK ltd

The application uses standard style-sheets to render the text with or without colour or tags shown. The underlying mark-up is shown in figure 5.

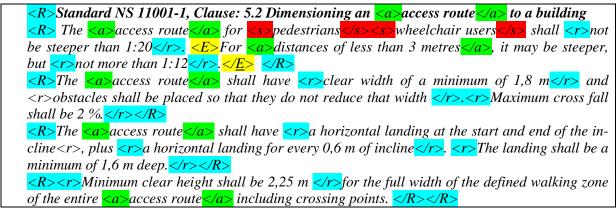


Figure 5: The underlying mark-up tags ('R/A/S/E' tags for clauses and 'r/a/s/e' tags for phrases.

In this presentation, objective clauses are shown here delimited by 'RASE' tags and metric phrases are delimited by 'rase' tags. Figure 6 summarises the phrases used with their equivalent object expressions.

Metric phrase	Туре	Object	Property	Comparison	Target	Unit
access route	applicability	space	usage	includes	access	
pedestrians	selection	space	user	includes	pede-	
					strian	
wheelchair users	selection	space	user	includes	wheel	
					chair	
					users	
not be steeper than 1:20	requirement	ramp	slope	more than	20.000	
distances of less than 3	applicability	ramp	length	less than	3.000	m
metres						
not more than 1:12	requirement	ramp	slope	more than	12.000	
clear width of a minimum	requirement	ramp	width	more than	1.800	m
of 1,8 m						
obstacles shall be placed so	requirement	ramp	obstructed	equals	FALSE	
that they do not reduce						
that width						
maximum cross fall shall be	requirement	ramp	cross fall	less than	2.000	%
2 %						
a horizontal landing at the	requirement	ramp	has	equals	TRUE	
start and end of the incline	_		landings			
a horizontal landing for	requirement	ramp	landing	less than	0.600	m
every 0,6 m of incline	_		interval			
the landing shall be a	requirement	landing	width	more than	1.600	m
minimum of 1,6 m deep.						
<minimum clear="" height<="" td=""><td>requirement</td><td>space</td><td>clear height</td><td>more than</td><td>2.250</td><td>m</td></minimum>	requirement	space	clear height	more than	2.250	m
shall be 2,25 m						

Figure 6: Summary of the metric phrases

One approach to confirming the accuracy of the process is to re-transform the logical; statement back into prose (see figure 7). For clarity this representation uses names (in quotes) for the intermediate objective clauses to convey the strict logical expression that every object in the building must satisfy. In the constraint model these names are usually unique and arbitrary identifiers that are keyed into the text.

Compliance to "Dimensioning an access route to a building" is achieved by either not being an access route, or meeting "access route slope", "access route width", "access route landings" and "access route height". Compliance to "access route slope" is achieved by not being used by pedestrians and not being used by wheel chair users, or by not being steeper than 1:20 or by being a "short ramp". Compliance to "short ramp" is achieved by being shorter than 3 m and being less than 1:12. Compliance to "access route width" is achieved by having a clear width of 1.8 m, and having no obstacles, and by having a cross fall of less than 2%. Compliance to "access route landings" is achieved by having a horizontal landing at the start and end of the incline and by having a horizontal landing every 0.6 m of incline and by having a landing depth of 1.6 m. Compliance to "access route height" is achieved by having a clear height more than 2.25 m.

Figure 7: A re-presentation of the derived logical statement as prose.

3.2 Test Case: A paragraph from the Dubai building regulation

In some cases normative documents include tabular inserts. This example (figure 8) shows how the tabular structure can be used along with the mark-up to infer the logical statement. This and the third example are concerned purely with the spatial structure of the building.

Building Habitable Spaces

The minimum floor area and minimum dimension of any space in different types of building shall be as follows:

	Minimum floor area	Minimum dimension		
Shop	$12.0 \ m^2$	2.4 m		
Office	$10.0 \ m^2$	2.5 m		
Residential room	$10.0 \ m^2$	3.1 m		
Kitchen	$6.0 m^2$	2.0 m		
Bathroom	$4.5 m^2$	1.8 m		
Lavatory	$1.5 m^2$	1.0 m		

It is permissible for some service rooms such as those used by servants and guard rooms to have a floor area of 6.5 m2 and a minimum dimension of 2.1 m.

Figure 8: The regulation.

The original text including the table can be marked-up (figure 9). The confusion between building usage and space usage was raised with the building authority.

	Minimum floor area	Minimum dimension
<a>Shop	$< r > 12 m^2 < / r >$	< r > 2.4 m < / r >
<mark><a></mark> Office <mark></mark>	$< r > 10 \text{ m}^2 < / r >$	< <i>r</i> >2.5 <i>m</i> <i r>
<mark><a></mark> Residential room <mark></mark>	$< r > 10 \text{ m}^2 < / r >$	< r > 3.1 m < / r >
<mark><a></mark> Kitchen <mark></mark>	< r > 6 m ² < /r >	< r > 2.0 m < / r >
<mark><a></mark> Bathroom <mark></mark>	< r > 4.5 m ² < / r >	< <i>r></i> 1.8 <i>m</i> <i r>
<mark><a></mark> Lavatory <mark><a></mark>	< r > 1.5 m ² < /r >	< <i>r></i> 1.0 <i>m</i> <i r>

Figure 9: The underlying mark-up tags

The mark-up of the table is similar to that applied to the prose text. The cells within the table typically contain applicability and requirement metrics. The table is read systematically and every interior cell with a requirement is made into a distinct test, taking any applicability from the upmost and leftmost cells.

Metric phrase	Type	Object	Property	Comparison	Target	Unit
Shop (and others)	applicability	building/	usage	includes	Shop	
		space				
12 m ² (and others)	requirement	space	floor area	more than	12.000	m ²
2.4 m (and others)	requirement	space	short side	more than	2.400	m

Figure 10: The underlying mark-up tags

The tabulation of the metric phrases (figure 10) shows that a small set of queries are needed to examine the building model.

Compliance to "Building Habitable spaces" is achieved by either not being a habitable space or by meeting the "shop dimensions", "office dimensions", (..., or by meeting the "service exception". Compliance to "shop dimensions" is met by either not being a shop or by having area greater than 12 m^2 and short side greater than 2.4 m...). Compliance to the "service exception is made by being a service room, used by servants or a guard room or by having area greater than 6.5 m^2 and short side more than 2.1 m.

Figure 11: The re-presention as prose (some test omitted for brevity).

The re-presented prose (figure 11) is correct but not particularly easy to read. A more sophisticated representation transformation would be able to reconstruct the table.

3.3 Test case: US Court design guidance document

Guidance documents may not undergo the intensity of review and revision applied to regulatory documents. An example paragraph has been taken from the US Court Design Guidance (figure 12). It is not unusual to find undefined terms such as 'near' which may need to be negotiated and agreed between the client and the designer.

The activities of the USDC focus on the courtroom. The courtroom requires direct access from public, restricted, and secure circulation. Ancillary spaces located near the district courtroom include: attorney/witness conference rooms accessed from public circulation; judge's conference robing room (provided only if the judges chambers are not located close to the courtroom) accessed from restricted circulation; trial jury suite accessed directly from the courtroom or restricted circulation; and prisoner holding cells accessed from secure circulation.

Figure 12: The original guidance

In the original text, it may not be obvious that this is tabulating access requirements for specific spaces. The process of marking up isolates the separate requirements, along with some specific exceptions (figure 13).

Figure 13: Guidance document with mark-up

The mark-up identifiers the ancillary spaces as being required to be near the courtroom and then details their individual access requirements. The requirements for access from the judges robing room are particularly subtle. The tabulation of the metric phrases (figure 14) shows that there are only three properties being used to assess the design. The interpretation of 'nearness' would need clarification, possibly supported by the use of 'fuzzy' logic and/or other forms of transformation.

Metric phrase	Type	Object	Property	Comparison	Target
spaces	applicability	space	type	equals	TRUE
attorney/witness conference rooms	applicability			matches	attorney/witness conference rooms
courtroom	applicability				courtroom
judges chambers	applicability				judges chambers
judge's conference robing room	applicability				judge's conference robing room
prisoner holding cells	applicability				prisoner holding cells
trial jury suite	applicability				trial jury suite
located close to the courtroom	requirement		nearness	includes	courtroom
public circulation	requirement		access	includes	public
restricted circulation	requirement				restricted
secure circulation	requirement				secure

Figure 14: Summary of the metric phrases used in the guidance document

This re-presentation (figure 15) again uses names for the intermediate objective clauses to convey the strict logical expression that every space in the courtroom must satisfy. By systematic testing, a compliance checking engine can rapidly highlight any issues in the building and in the original document.

Compliance to "Major spaces" is achieved by meeting the "courtroom access", and "ancillary access". Compliance to "Courtroom access" is achieved by either being not a courtroom or by having access to public circulation, having access to restricted circulation and having access to secure circulation. Compliance to "ancillary access" is achieved by either not being an ancillary space or by being near the courtroom and by meeting "ancillary space access". Compliance to "Ancillary space access" is achieved by meeting "attorney access", "judges robing access, "trial jury access" and "prisoner holding access". Compliance to "attorney access" is achieved by either not being an attorney/witness conference room or by having access to public circulation. Compliance to "judges robing access" is achieved by either not being a judges robing room or by having access to restricted circulation, or by meeting "alternative judges robing". Compliance to "alternative judges robing" is achieved by either not being a judges robing room or being close to the courtroom. Compliance to "trial jury access" is achieved by either not being a trial jury room or by meeting "alternative trial jury access". Compliance to "alternative trial jury access" is achieved by either being directly accessed from the courtroom or having access to the restricted circulation. Compliance to "prisoner holding access" is achieved by either not being a prisoner holding cell or by having access to secure circulation.

Figure 15: The re-presentation as prose

RESULTS / CONCLUSIONS The expected result are is a quantitative determination of how valid and reliable the RASE methodology are on different categories of text. Whilst poorly written normative documents will naturally need extra care, the automatic generation of a logical statement appears robust with no clauses being un-handled, and the re-presentation accurately reflecting the original content. The use of mark-up to capture simple metadata give a foundation for both automatic and user-driven model checking systems.

The methodology also exposes the fundamental metric phrases which a building model server or user must answer during automatic or interactive model compliance checking.

The RASE technology give a significant improvement in reduced time and improved documentation for capturing requirements.

REFERENCES

- Giarratano, J. and Riley, G. (2005). "Expert Systems Principles and Programming". Course Technology, a division of Thomson Learning, Inc. ISBN 0-534-38447-1
- GSA (2011). "US Courts design Guide" accessed at http://www.gsa.gov/graphics/pbs/Courts_Design_Guide_07.pdf . March 2011. Chapter 3 "Planning for U.S. Courthouses" pp 3.3, 3.4 (extract). (accessed: 2011-05-05).
- Hjelseth, E. (2009) "Foundation for development of computable rules". Presented at CIB-W78 conference in Istanbul 1st 3rd October 2009.
- Hjelseth, E. and Nisbet, N. (2010a) "Exploring semantic based model checking". Presented at CIB W78 2010 27th International Conference Applications of IT in the AEC Industry.
- Hjelseth, E., and Nisbet, N., (2010b) "Overview of Concepts for Model Checking". Presented at CIB W78 2010 27th International Conference Applications of IT in the AEC Industry.
- Lillehagen, F. and Krogstie, J. (2008). "Active Knowledge Modeling Of Enterprises". Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-79415-8
- Nisbet N, Wix J and Conover D. 2008. "The future of virtual construction and regulation checking", Brandon, P., Kocaturk, T. (Eds), Virtual Futures for Design, Construction and Procurement, Blackwell, Oxfordshire. doi: 10.1002/9781444302349.ch17
- Regulations (2006). Anon: "Building Regulations and Design Guide Fourth Edition 2006", Dubai. Section 9 "Building spaces", (extract) (private communications with the authors).
- Russell, S. and Norvig, P. (2010) "Artificial Intelligence: A Modern Approach", Prentice Hall Series in Artificial Intelligence. Englewood Cliffs, New Jersey. ISBN-13: 978-0-13-604259-4, ISBN-10: 0-13-604259-7. http://aima.cs.berkeley.edu (accessed: 2011-05-05).
- Schalkoff, R. J. (2011). "Intelligent systems: principles, paradigms, and pragmatics". Jones and Bartlett Publishers ISBN-13: 978-0-7637-8017-3
- Sowa, J. F. (2000) "Knowledge representation: Logical, Philosophical and Computational foundations". Thomson Learning. ISBN 0 534-94965-7
- Sowa, J. F. (2006) "The Challenge of Knowledge Soup". Published at Research Trends in Science, Technology and Mathematics Education, Mumbai. http://www.jfsowa.com/pubs/challenge.pdf (accessed: 2011-05-05).
- Tarski, A. (1944) "The Semantic Conception of Truth and the Foundations of Semantics", University of California, Berkeley. Published in Philosophy and Phenomenological Research 4 (1944). http://www.jfsowa.com/logic/tarski.htm (accessed: 2011-05-05).
- Whitehead, A. N., (1937) "Analysis of Meaning," Philosophical Review, reprinted in A. N. Whitehead, Essays in Science and Philosophy, Philosophical Library, New York, pp. 122-131.
- Wix, J., Nisbet, N. and Liebich, T. 2008. "Using Constraints to Validate and Check Building Information Models". 7th European Conference on Product and Process Modeling (ECPPM) Sophia Antipolis. 10.-12 Sept. 2008 p. 467-476 Edited by Raimar Scherer and Alain Zarli Taylor & Francis 2008. Print ISBN: 978-0-415-48245-5. eBook ISBN: 978-0-203-88332-7