SBVR's Approach to Controlled Natural Language

Silvie Spreeuwenberg[1] and Keri Anderson Healy[2]

[1] LibRT, Amsterdam, Netherlands
[2] BRCommunity.com, USA
Silvie Spreeuwenberg, silvie@librt.com
Keri Anderson Healy, kandersonhealy@brcommunity.com

Abstract. The "Semantics of Business Vocabulary and Business Rules" (SBVR 1.0) is one of the initial specifications in the OMG's family of business-focused specifications. SBVR covers two aspects: Vocabulary (natural language ontology) and Rules (elements of guidance that govern actions). However, SBVR does not standardize any particular language for expressing vocabularies and rules. Instead, SBVR uses 'semantic formulation', which is a way of describing the semantic structure of statements and definitions. This approach of specifying structures of meaning, with its sound theoretical foundation of formal logic, provides a formal, language-independent means for capturing the semantics of a community's body of shared meanings. By taking this approach, SBVR can support multiple forms of representation.

Keywords: Business Rules, Business Vocabulary, Semantics, Semantic Formulation, SBVR.

1 Introduction

In December 2007, the Object Management Group (OMG) published "Semantics of Business Vocabulary and Business Rules" (SBVR 1.0), one of the initial specifications in the OMG's family of business-focused specifications. 'Business' in this context means 'human' — as contrasted with computer systems or technology. The term should be interpreted in its broadest sense, pertaining to any human/organizational activity, and not be interpreted narrowly (i.e., as only pertaining to commercial activity).

SBVR covers two aspects:

- *Vocabulary*: natural language ontology ('terminology' to ISO) — concepts and their representations (terms, names, definitions) as a cohesive set, rather than a simple list of terms and definitions.
- *Rules*: elements of guidance (policies, rules) that govern business actions of an organization.

However, SBVR does not standardize any particular language for expressing vocabularies and rules. Indeed, it is specifically *not* the intention of SBVR to mandate that any particular notation be used.

2 SBVR's Approach to meaning

At the heart of SBVR's approach is its being concept-centric, rather than word-centric. To this is added the notion of 'semantic formulation', which is SBVR's way of describing the semantic structure of statements and definitions. It is important to note that semantic formulations are not *expressions* of meaning — rather they are *structures* of meaning (the logical composition of meaning). With its sound theoretical foundation of formal logic, SBVR provides a formal, language-independent syntax for capturing the semantics of a community's body of shared meanings.

The structures of meaning are not used directly by people. Instead people will use a language that has a mapping to the structures of meaning in SBVR. This language can be in a graphical representation, but is most likely in a textual form. RuleSpeak® and SBVR Structured English are two example languages with a mapping to SBVR structures of meaning. They are controlled languages since they deal with a restricted subset of a language with respect to the mapping to the SBVR structures of meaning. However, the 'controlled' language can also be extended using the SBVR structures of meaning.

The SBVR standard itself is described as an SBVR vocabulary and gives, as such, a start for describing a controlled language. The structures of meaning can be used to extend this controlled vocabulary with domain-specific terminology. In particular, the domain-specific concepts must be defined using the structures of meaning provided by SBVR. Core notions in the structuring of meaning are (among others) the following:

- *Noun concept*, defined as: concept that is the meaning of a noun or noun phrase
- *Individual concept*, defined as: concept that corresponds to only one object [thing]
- *Verb concept*, defined as: concept that is the meaning of a verb phrase that involves one or more noun concepts and whose instances are all actualities

In a car rental business, typical noun concepts might be represented by the terms 'driver', 'vehicle', 'rental', etc. An example of individual concepts — usually only a small part of the total vocabulary — are 'Dollar' and 'Euro', each the name of a currency.

Verb concepts provide the ability to define connections between concepts that are of interest to the organization. These connections provide the business-level semantic structure required to find information about such relationships in text documents and relational databases, as well as providing the ability to specify business rules formally and unambiguously. For example, in a financial business, the connection between the concepts 'driver' and 'rental' might be defined by an associative fact type 'authorizes' ('rental authorizes driver'). Different kinds of verb concepts provide a powerful means to build ontologies that are semantically equivalent to Ontology Web Language (OWL). Three hierarchical relations are defined to describe: *assortments* (relationship between individual and general concept), *specializations* (hierarchical relationship between a concept and a category such that an instance of the concept is also an instance of the category), and *parts* (a given part being in the composition of a given whole).

By taking this approach, SBVR can support multiple forms of representation. For example, a fact type can be readily understood in both its forward (person *rents* vehicle) and reverse (vehicle *is rented by* person) readings as being for the same meaning. Also, both noun and verb fact type forms can be interpreted as one meaning as in the following two example fact type expressions: "person *has* phone number", "phone number *of* person".

2.2 SBVR's approach to representation

The approach of semantic formulation, with its logic grounding, supports two essential features of SBVR. First is the mapping of a semantic community's body of shared meanings to the vocabularies (and thereby the expressions and communication forms) used by its speech communities. For example, a rule (that prohibits crossing the railroad tracks) can be expressed in various national languages:

- Überschreiten der Gleise verboten [in the German-speaking community]
- Défense de traverser les voies [in the French-speaking community]
- Vietato attraversare i binari [in the Italian-speaking community]
- Crossing the railway lines is prohibited [in the English-speaking community]

Second is the mapping to XMI that enables interchange of concepts, facts, and business rules between languages (and supporting tools) that implement SBVR. For example, these are equivalent expressions of the same rule, according to the language conventions of (respectively) RuleSpeak and SBVR Structured English:

- RuleSpeak: The renter of a vehicle must have exactly three phone numbers.
- SBVR Structured English: It is obligatory that the renter of a vehicle have exactly three phone numbers.

Each word or word phrase in the sentence is mapped to the equivalent element of meaning in SBVR. For the second sentence the following table provides that mapping:

Representation in SBVR controlled English	***SBVR element***
It is obligatory that	Obligation formulation
the renter	Noun concept
of	Fact symbol in fact type "renter of vehicle" ("vehicle has renter")
a	Existential quantification
vehicle	Noun concept
have	Fact symbol in fact type "renter has phone number"
exactly three	Exactly-3 quantification
phone numbers	Noun concept

To perform this mapping automatically all words or word phrases have to be defined as SBVR elements of meaning. A parser that can deal with grammatical issues like

plurals and tense and has an understanding of the language conventions is needed to create the mapping with the meaning. These components are and have been built by several vendors of supporting software tools.

The difference between RuleSpeak and SBVR Structured English is related to readability and ease of use. The design decisions one can make in creating a mapping from SBVR to a controlled language are worth more investigation, as are evaluation methods to evaluate the resulting language.

The SBVR initiative is intended to capture business facts and business rules that may be expressed either informally or formally. Business rule expressions are classified as formal only if they are expressed purely in terms of concepts in the pre-declared schema for the business domain, as well as certain logical/mathematical operators, quantifiers, etc. Formal statements of rules may be transformed into logical formulations that are used for exchange with other rules-based software tools. Informal statements of rules may be exchanged as un-interpreted comments.

2.3 Logic Grounding

SBVR's logic foundation is first-order predicate logic with some restricted extensions into higher-order logics, with some limited extensions into modal logic — notably some deontic forms (for expressing obligation and prohibition) and alethic forms (for expressing necessities and possibilities). SBVR's use of modal logics yields provably-equivalent patterns of rule expression. For example, a given business rule can be stated in the form of prohibition, obligation, or restricted permission and be assured to represent the same underlying meaning. Again, these are three semantically-equivalent natural language expressions of one rule:

- It is prohibited that an open rental has an intoxicated driver.
- It is obligatory that no open rental has an intoxicated driver.
- It is permitted that a rental be open only if the rental does not have an intoxicated driver.

Assuming the characteristics 'person is intoxicated' and 'rental is open' are part of the vocabulary and that 'driver' specializes 'person', the semantic formulation underlying these statements can be expressed as:

It is obligatory that
. Not
. . Exists v1 : 'rental' where '<u>rental</u> is open'(v1)
. . . Exists v2 : 'driver' where '<u>person</u> is intoxicated'(v2)
. . . . '<u>rental</u> has <u>driver</u>'(v1, v2)

Or, equivalently:

It is obligatory that
. For all v1 : 'rental' where '<u>rental</u> is open'(v1)
. . For all v2 : 'driver' where '<u>person</u> is intoxicated'(v2)
. . . Not
. . . . '<u>rental</u> has <u>driver</u>'(v1, v2)

Acknowledgments

We would like to express our appreciation to all the members of the team who contributed to the SBVR specification. It is their willingness to share knowledge and their individual dedication to quality work that is making this body of work a success. In particular, we would like to thank those who have devoted their time and effort to reviewing, and improving, this paper.

References

1. Business Rule Solutions: BRS RuleSpeak® Practitioner's Kit. Business Rule Solutions, LLC (2001- 2004), http://www.rulespeak.com
2. Girle, R.A.: Modal Logics and Philosophy. McGill-Queen's University Press (2000)
3. Halpin, T.A., Girle, R.A.: Deductive Logic. 2nd ed. Logiqpress, Brisbane (1981)
4. Halpin, T.A.: Object-Role Modeling: An Overview. San Francisco: Springer, San Francisco (2000), http://www.orm.net/pdf/springer.pdf
5. International Organization for Standardization (ISO): Terminology work — Vocabulary - Part 1: Theory and Application. English/French ed., ISO (2000)
6. Larson, R, Segal, G.: Knowledge of Meaning: An Introduction to Semantic Theory. The MIT Press (1995)
7. Nijssen, S., Bijlsma, R.: A Conceptual Structure of Knowledge as a Basis for Instructional Designs. In: Kinshuk R., Koper P., Kommers P., Kirschner D., Sampson G., Didderen, W.E. (eds.) ICALT'06, IEEE: 6th Int. Conf. on Advanced Learning Technologies, 7-9. IEEE (2006)