Rule authoring is a creative process

Recently I attended a workshop on controlled languages on a very nice island named Marettimo in Italy. The majority of the attendees were academics, and they were struggling with their definition of a Controlled Natural Language (CNL). To create more clarity in the field and to help build a true CNL community, the idea of creating a *CNL Manifesto* was introduced during the workshop. The *Business Rules Manifesto* has served the same purpose and has been reasonably successful in establishing this goal (based on the dozen translations and large number of citations it has had so far). So that seems to be a good idea.

But this column is not about the *Manifesto*. As an observer of what was going on in the CNL community I was intrigued by the two approaches to CNL that were presented. One approach is called the 'naturalist' approach, and the other is called the 'formalist' approach.

The objectives of the *naturalist* approach to CNL were defined as:[1]

> *"the naturalist approach to CNL treats CL interpretation as a simpler form of the full natural language (NL) processing task in which ambiguity still resides, only to a lesser extent. One might say that the goal is to make English more tractable (understandable by computers) by simplifying the complexity of the English handled. In this approach, as with full NL, multiple interpretations of a sentence are possible, and the CL interpreter uses all the standard NL machinery to search for a 'best' parse and interpretation, but with the task being constrained to just the subset of English covered by the CL."*

The objectives of the *formalist* approach to CNL were defined as:[1]

> *"[the formalist approach to CNL] views a CL more as an English-like formal/programming language that is well-defined, predictable, and easier to use than a normal formal language. One might say that the goal is to make logic more tractable (easier to use by humans) by rendering it in human-readable terms that non-mathematicians can understand. Given that many existing formal languages are somewhat cryptic to untrained users, and good NL processing techniques already exist, it is a natural step to have users work*

> with CLs instead that are more readable and translate deterministically into the formal language. A 'formalist' approach would view this as an end in itself, and make no claims that the CL necessarily says anything about NL processing in general."

In writing business rules, the objectives of both approaches are appealing. Not surprisingly we see that the SBVR Structured English objectives have a strong overlap with the objectives of the formal approach to CNL. SBVR says about SBVR Structured English:[2]

> "… using English that maps mechanically to SBVR concepts. It is not meant to offer all of the variety of common English, but rather, it uses a small number of English structures and common words to provide a simple and straightforward mapping."

But the objectives of RuleSpeak have more overlap with the naturalist approach to CNL. The latest document on RuleSpeak says:[3]

> "The purpose of the Sentence Forms is to ensure that written Business Rules are more easily understood. They also help ensure that different practitioners working on a large set of Business Rules express the same ideas in the same way. Such consistency would not be possible if Business Rules were expressed in a completely 'free-form' manner. "

Being a spoiled only child, of course I want both. Is this realistic? According to the CNL community it is at least very ambitious.

What should we do if we want both?

To start with an answer to this question I suggest that we view the authoring of a business rule as a creative process.

The first version of a rule that a business expert writes down is often not the complete, truthful,

non-ambiguous, and perfect verbalization of the rule but instead *a braindump of the most important conditions that apply to a situation*. In this stage a naturalist approach to controlled language may help the user to structure the resulting sentence by offering support in the removal of ambiguous structures from the sentence and by suggesting vocabulary enhancements that ensure completeness.

Only when this first step of the rule authoring process is finished can we introduce the CNL of the formalist kind. Often this results in updating the sentence again to make it machine-processable.

After this step we can offer other verification checks, such as looking for conflicts and subsumptions. Validation can take place both after finishing step 1 and after finishing step 2. Also, it is very possible that several versions of a rule are created during steps 1 and 2, improving the rule with every version.

Does this sound inefficient to you? Then this is your wake-up call! Stop day-dreaming!

Even an experienced rule author with a background in formal logics (as I am) is not able to write a good rule without the first step. Omitting the first step blocks the creative process and leads to rules that erroneously simplify the situation at hand. So please do not ask your business users to write these machine-processable statements without supporting the creative phase of rule authoring! And do not feel ashamed if you are a business expert and find it difficult to write the right rule at once.

My conclusion is that we need both types of CNL for business rules, where the rule authoring process makes use of both approaches.

I hope that the CNL community and the rules community are able to connect the two approaches together in some way. Furthermore, I would recommend that the 'N' of CNL be taken seriously. Can we speak of a CNL if the resulting sentence is not natural and grammatically correct? I don't think so! And it would definitely not be useful to the business rules community. So there are some ambitious goals to reach in the near future. We are

working on it.

This article was originally published by BRCommunity (link).

[1] Peter Clark, P. H. (2009). "Naturalness vs. Predictability: A Key Debate in Controlled

Languages." In N. E. Fuchs (Ed.), *Workshop on Controlled Natural Language.* Marettimo

Island: Pre-Proceedings

[2] *Semantics of Business Vocabulary and Business Rules* (SBVR), v1.0. Object

Management Group (Jan. 2008). Retrieved from OMG: www.omg.org/spec/SBVR/1.0/PDF

[3] R. G. Ross. (2009). *RuleSpeak.* Retrieved from RuleSpeak: www.rulespeak.com