

\*new RFC# 87 targeted at 1.1.0\*

**Major revisions v1.0.0 (BBFRFC 84) to v1.1.0 (BBFRFC 87) are:**

- Clarified that SA.strand relates the sequence between parent and subComponent DC orientation.
- Adopted requirement to use SO for type as passed by vote on 10/12/2011
- All other minor revisions added up over time

**ALL REVISIONS**

UML diagram (Figure 5)

=====

**v1.0.0 to v1.0.3**

Removed "s" from "subComponents", the property is now "subComponent"

Changed datatype of "bioStart" from "int()" to "positiveInteger"

Changed datatype of "bioEnd" from "int()" to "positiveInteger"

**v1.0.3 to v1.0.8**

sync UML diagram versions with document

Specification document

=====

**v1.0.11 to v 1.0.12**

- Revert "... sa1.bioEnd is strictly less than sa2.bioStart" to "...sa1.bioEnd is strictly greater than sa2.bioStart"; "... sa1.bioEnd is strictly less than sa2.bioStart" was correct in v1.0.2

- Added: "Logical consistency of Location Data

Given sa1:SequenceAnnotation, sa1.bioStart and sa1.bioEnd are logically consistent if: sa1.bioStart is greater or equal to 1 (positive integer); and

sa1.bioEnd is greater or equal to sa1.bioStart; and

the DNA sequence length of the DnaComponent specified by sa1.subcomponent is equal to sa1.bioEnd - sa1.bioStart + 1." this requirement was passed by vote VaTech 2011, and had been missing in the specification text, thanks to Swapnil Bhatia for noticing this omission.

- Throughout the text the word "direction" was used when discussing strand of the subComponent(DC). Changed this to "strand orientation".

- Section 8.5.3 changes

- SA class definition now reads:

"Individual instances of the SequenceAnnotation class specify a relationship of the subComponent (DnaComponent) to the DnaComponent being annotated, its 'parent' DnaComponent. This relationship is the position and strand orientation of the subComponent relative to the parent DnaComponent orientation. The SequenceAnnotation location CAN be specified by the bioStart and bioEnd positions of the subComponent, along with the DNA sequence. Alternatively, the partial order of SequenceAnnotations along a DnaComponent can be specified by indicating the precedes relationship to other SequenceAnnotations. As a convention, numerical coordinates in this class use position 1 (not 0) to indicate the initial base pair of a DNA sequence. This convention is followed by the broader Molecular Biology community, especially in the relevant literature. The strand orientation, or direction, of the subComponent's sequence relative to the parent DnaComponent is specified by the strand [+/-]. For strand: '+' the sequence of the subComponent is the exact sub-sequence, and for '-' it is the reverse-complement of the parent DnaComponent's sequence in that region.

A SequenceAnnotation instance MUST specify a subComponent of type DnaComponent.  
SequenceAnnotations MUST belong to exactly 1 DnaComponent."

- SA.strand now says:

- "Zero or one value of type xsd:string. Strand orientation '+' or '-' is a 'subComponent relative to the parent DnaComponent orientation' flag. For a full explanation see Logical consistency of the subComponent's DnaSequence value.

- SA.subComponent now says:

- One and only one value of type DnaComponent. This property specifies the DnaComponent which is being annotated on the parent DnaComponent's sequence. Analogous to 'a feature' in other systems, the DnaComponent value serves to indicate information about the sequence at the position specified by the SequenceAnnotation's location data properties or the relative position object property. The sequence of the DnaComponent specified by the subComponent property MUST be logically consistent with the strand value.

- Logical consistency of the subComponent's DnaSequence value now says:

When present, the subComponent's DnaSequence MUST relate to the exact sequence found in the interval specified by the Location Data of the SequenceAnnotation.

The subComponent's sequence is logically consistent if the value of strand is:

'+' and the subComponent's sequence specifies the sequence of the parent DnaComponent in that region.

'-' and the subComponent's sequence specifies the reverse-complement of the parent DnaComponent's sequence in that region.

If the strand value is un-specified and the DnaSequence of the subComponent is present, the subComponent's sequence specifies the exact sub-sequence of the parent DnaComponent.

- Removed the reference to precedes='null' in "Well-formed constraint: Relative Position"

Section 9

Added Section 9.4 Example "DnaSequence of subComponent on the minus strand" Figure 10

#### **v1.0.10 to v1.0.11**

- formatting/page numbers have changed

- from Mandy's comments

- p. 5 "As software tools are adapt to progress ..." changed to "As software tools are adapted to progress ..."

- p. 6 "For example, DNA components can be hypothesized to have a biological function, deemed necessary for DNA assembly processes, or serve the synthetic biologist as landmarks in analysis." Verbs should be consistent...

now reads:

"For example, DNA components can be hypothesized to have a biological function, be deemed necessary for DNA assembly processes, or have served as landmarks in analysis."

- p 9 (Top) "SBOL version 1.0, focuses on the representation of DNA components as the SBOL:Core:model." (Shouldn't this be SBOL version 1.0.0 or 1.1.0?) corrected to 1.1.0

- Corrected wording in SBOL Vocabulary and Appendix to remove statements about an SBOL Editor maintained vocabulary, a proposal that rejected by vote. The revisions are based on the following comments from Mandy:

- Mandy: "SBOL:Vocabulary defines the key terms used in the core model." Didn't we get rid of the SBOL Vocabulary? I guess you talk later about some of the terminology (ie, DnaComponent) being part of the SBOL:Vocabulary, but I wonder if that will be confusing since there is also going to be an SBOL:Vocabulary extension (which you also refer to in this document.) You also have a section in this document called SBOL Vocabulary (8.3) -- I'm just wondering if we are overusing this.

You identify a couple of provisional terms that we are going to ask the Sequence Ontology folks to add (Gene Expression and DNA Construction, and you refer the reader to a list of provisional terms in the Appendix. The list in the appendix would appear to be examples of actual Sequence Ontology terms, not provisional ones (the two you name on page 9 are not in there, for example.)

Mike: removed the references to extensions, but I kept "SBOL:Vocabulary" which are terms specific to SBOL, with no overlap with SO.

Section 8.3 is the vocabulary used to define Core classes.

Section 8.1 now reads:

SBOL Vocabulary

In version 1.1.0, SBOL:Vocabulary defines the core concepts used by SBOL in the structured description of synthetic DNA designs. SBOL:Core terms are DnaComponent, DnaSequence, and SequenceAnnotation. To provide user-defined groupings of DnaComponents, SBOL:Core also defines a Collection. Additional terms, such as those used as used to classify DNA Components by type (see section 8.5.1) are defined by the Sequence Ontology (Eilbeck 2005, Mungall 2010). New terminology should be added in collaboration with the Sequence Ontology project.

"

Section 8.3 now reads:

"The SBOL:Vocabulary defines terms used in SBOL. Below we define terms for the Core. Non-SBOL term definitions such as those needed to classify DnaComponents by type can be obtained from the Sequence Ontology (Eilbeck 2005, Mungall 2010). For example, a promoter region, coding sequence, and transcriptional terminator are all defined by the Sequence Ontology (see the Appendix for a list of examples). Terminology outside of the scope of the Sequence Ontology should be submitted as new term requests to its curators (<http://www.sequenceontology.org/>)."

p 9 "start with lower cases letters and follow lower camelCase." Probably it is sufficient to say "lower camelCase", because by definition a lower camelCase term starts with a lower-case letter, but if you want to include the additional clarification, it should be "start with lower-case letters". - done

p.20 SequenceAnnotation.strand elements are missing colons, ie "strand +" instead of "strand: +" This is inconsistent w/p.19 -done

p.23 Cesar's email address is no longer valid -- should use cesarr@me.com, unless there is another email address he would prefer to have used. -done

## **v1.0.9 to v1.0.10**

pg 10 Clarified Version policy

"Major versions (X) ... the submission of the specification document.. (BBF RFC)" added "is REQUIRED".

"Minor versions (Y)" added "The submission of the specification document to ... (BBF RFC) is OPTIONAL."

"Patch versions (Z)" added "The submission of the specification document to the BioBrick Foundation as a Request For Comment (BBF RFC) is NOT RECOMMENDED."

Changed any mentions of v1.0.0 which needed to be updated to v1.1.0

pg 1 changed RFC# from 84 to 87; 87 replaces 84

pg 15 added table of nucleotide ambiguity codes to definition of DnaSequence.

p. 5 "As software tools are adapt to progress ..." should be "As software tools are adapted to progress ..." -done

p. 6 "For example, DNA components can be hypothesized to have a biological function, deemed necessary for DNA assembly processes, or serve the synthetic biologist as landmarks in analysis." Verbs should be consistent.

changed to ", or have served as landmarks in analysis." -done, check w/ Mandy

p 9 (Top) "SBOL version 1.0, focuses on the representation of DNA components as the SBOL:Core:model." (Shouldn't this be SBOL version 1.0.0 or 1.1.0?) -done

#### **v1.0.8 to v1.0.9**

Fixed appendix: SO URI's to use "purl" namespace

Fixed appendix: removed non-SO examples

Fixed appendix: removed sub-sections - Appendix is now just "examples" of SO

Fixed page breaks

Fixed other formatting

#### **v1.0.2 to 1.0.8**

pg 3 Table of Contents updated

pg 9 remove "rp" from SAnull and replace relative position (terminal SA) with position unspecified in Abbreviation key

pg 9 Add section # 8.2.1 to SBOL versions and releases

pg 11 Figure 5 updated to v1.0.2 "subComponents" changed to "subComponent".

pg 11 Figure 5 updated to v1.0.8 change bioStart and bioEnd datatype "int()" to "positiveInteger()"

pg 11 Figure 5 add current version # (v1.1.0) to diagram, keep the spec synced w/ UML diagram.

pg 13 remove "for" in DC.name last sent. "and meaningful, for as may be done" to "and meaningful, as may be done"

pg 13 DC.type replace "or provisional SBOL:Vocabulary extensions" with "(see Appendix for commonly used terms)"

pg 13 DC.type replace ", please contact the SBOL Editors (see section 11 for contact information)" with "(use 'Request a Term' <http://sequenceontology.org>)"

pg 15 change SA.bioStart xsd:integer to xsd:positiveInteger

pg 15 change SA.bioEnd xsd:integer to xsd:positiveInteger

pg 16 remove wording in SA.precedes "Finally, a null value for a precedes property indicates the terminal SequenceAnnotation within a given DnaComponent. Absence of the precedes relation indicates this knowledge is not known."

pg 16 change "... sa1.bioEnd is strictly less than sa2.bioStart" to "...sa1.bioEnd is strictly greater than sa2.bioStart"

pg 19 Figure 7.b add "strand: +" to SequenceAnnotations

pg 20 change "subcomposition" to "composition"  
pg 20 remove "two" from "in the two top DnaComponents"  
pg 22 change "SBOL version 1.0.0" to "SBOL version 1.1.0"

**v1.0.0 to 1.0.3**

pg 1 added authors Josh Natarajan, Carlos Olguin  
pg 2 Nicholas Roehner's affiliation changed to Bioengineering  
pg 15 move SA.uri to required Data Property section  
pg 19 Figure 7b "subComponents" changed to "subComponent".