BEL Script V1.0 Specification

Table of Contents

Introduction

This document describes the format and syntax for BEL Script. BEL Script is a syntax and format for encoding BEL Statements and associated annotations as human-readable documents that can also be efficiently processed by the BEL Framework. BEL Script is an alternative to encoding BEL Statements using the XBEL XML format.

Both BEL Script and XBEL are document formats, that is, they have a specialized structure and have several sections that enable the BEL Framework to determine how to process and interpret BEL Statements. This document describes the overall document structure and provides examples of encoding BEL Statements using the BEL Script format.

Assumptions

This document assumes that the reader is familiar with the BEL language that is documented separately.

Version Changes

There are no version changes associated with this document.

Additional Resources

The following documents provide additional information.

- *Biological Expression Language V1.0 Overview*

- *BEL Framework V1.0 Annotation Types and Namespaces User Guide*

BEL Script Overview

BEL Script is a syntax and document format for storing multiple BEL Statements in a single text file and associating statement annotations, namespaces, and provenance with the statements.

BEL Script can be written using any text editor or word processor application that can save documents as simple text files. Each BEL Script is composed of records (rows) in the file which either:

- specify a BEL Statement using BEL,

- provide a control record, or

- define a comment.

Sections

BEL Scripts are composed of three (3) functional sections:

- Properties

- Definitions

- Statements
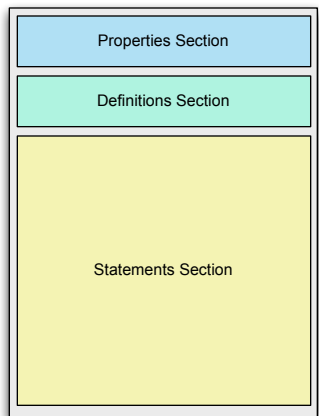
BEL Scripts are processed from top to bottom, thus the ordering of sections within the document and of records within each section is important.

Properties Section
The Properties section contains a set of control records that define the provenance for the document. This section includes control records that can identify when the document was created, the organization that created the document, copyright information if needed, and additional information that can be included to identify how the BEL Statements were processed.

Definitions Section
The Definitions section contains a set of control records that define the Namespaces and Annotation Types used within the document. Namespace control records are used to identify external vocabularies and ontologies that are used to identify entities within BEL Terms. Annotation Type control records define the set of annotations that can be used by BEL Statements within the document.

Statement Section
The Statement section is the last section in the document and contains a set BEL Statements and control records that associate Annotations and comments with the BEL Statements. Each BEL Statement records a biological fact.

Control Records

Control records are records that begin with a control action keyword. The BEL Script syntax supports the following control actions:

- SET

- DEFINE

- UNSET

SET
The SET action instantiates an object of a given type and assigns specific properties to that object reference. This action is used in the Properties section to associate provenance information with the statements that follow in the Statements section, and in the Statements section to associate annotations with specific BEL Statements. The SET action can also define a new Statement Group.

Usage:

Sets an annotation value, statement group name, or a document level property. Document-level properties must be set before all other annotations and BEL Statements.

Once set, document-level properties cannot be cleared. If the same document-level property is set multiple times, the last value takes precedence.

If a statement group is set, all following statements will be associated with the statement group. If no statement group is set, all statements will be associated with the default (unnamed) statement group.

The annotation will be associated with all BEL Statements after this command unless it is explicitly cleared with the UNSET action. If the annotation is previously set, the SET action changes its value. Use commas (,) to separate multiple values for the same annotation.

Syntax:

```
SET [ DOCUMENT ] name = value | {value, value,…}

SET STATEMENT_GROUP = value
```

Examples:

```
SET BodyPart = {"lung", "liver", "kidney"}

SET DOCUMENT Version = "1.0"

SET DOCUMENT License = "LGPL"

SET DOCUMENT Name = "Test Document"

SET STATEMENT_GROUP = "PubMed 123456"
```

DEFINE

The DEFINE action instructs the BEL Script interpreter to define a new object type and associate specific properties to that object type.

This action is used to define new Namespace references and new Annotation Types in the Definitions section of the BEL Script.

Usage:

Define a namespace or annotation in the document. Three value types are supported:

URL	A Uniform Resource Locator that can be used to retrieve a Namespace or Annotation Type definition.
LIST	A comma separated list of possible values, list surrounded by { and } characters. Can only be used with the ANNOTATION keyword.
PATTERN	A regular expression pattern that can be used to constrain the values associated with an Annotation Type. Can only be used with the ANNOTATION keyword

Syntax:

```
DEFINE [ ANNOTATION | [DEFAULT] NAMESPACE ] name AS [ URL | LIST
| PATTERN ] value | { value, values, …}
```

Examples:

The following example defines the **EGID** Namespace value. The Namespace identifier **EGID** is associated with the Namespace defined using the URL.

3

```
DEFINE DEFAULT NAMESPACE EGID AS URL \
    "http://resource.belframework.org/belframework/1.0/ns/entre
    z-gene-ids-hmr.belns"
```

This next example defines an Annotation Type named **ExposureTime** and restricts allowable values to the regular expression.

```
DEFINE ANNOTATION ExposureTime AS PATTERN "[0-9]+ hours"
```

This next example defines an Annotation Type named **Dosage** and restricts allowable values to the predefined set of values.

```
DEFINE ANNOTATION Dosage AS LIST {"low", "medium", "high"}
```

UNSET

The UNSET action clears one or more previously SET annotations such that any following BEL Statement does not get annotated with the annotations. If the UNSET action is applied to STATEMENT_GROUP any statements following the UNSET command will be assigned to the default statement group and all annotations associated will be restored to the state prior to the SET STATEMENT_GROUP command.

Usage:

The UNSET action can unset a single previously SET annotation, a list of previously SET annotation types or unset all currently set annotation types. Statements after the UNSET action will not be associated with the annotation unless they are explicitly set again with the SET action.

If the Annotation Type identified in the UNSET action is not defined, an Error will be generated during processing.

If the Annotation Type identified in the UNSET action is not SET, a Warning will be generated during processing.

Syntax:

```
UNSET [ name | { name, name,…} | ALL ]
```

```
UNSET STATEMENT_GROUP
```

Examples:

The following example will UNSET the ExposureTime Annotation Type.

```
UNSET ExposureTime
```

This next example will UNSET all currently set Annotations.

```
UNSET ALL
```

This example will UNSET the Annotations names rating, quality and species.

```
UNSET {ExposureTime, Dosage, Species}
```

This last example will UNSET the current Statemetn Group

```
UNSET STATEMENT_GROUP
```

Language Characteristics

This section provides information about other characteristics of the BEL Script format.

Reserved Keywords

The following keywords are reserved for use in BEL Script documents:

ALL	ANNOTATION	AS	DEFAULT
DEFINE	DOCUMENT	LIST	NAMESPACE
PATTERN	SET	STATEMENT_GROUP	UNSET
URL			

Case Sensitivity

BEL and BEL Script are case sensitive. All relationship types, function names, and reserved keywords are case sensitive.

Annotation Type names and annotation values are case specific.

Namespace identifiers are case specific. The case sensitivity of namespace values is defined within the Namespace itself but most Namespaces default to being case sensitive.

Blank Lines

BEL Script allows blank lines in any section. Blank lines are ignored by the BEL Script processor.

Quotes

All URLs used in Annotation Types or Namespace definitions in the Definitions section need to be quoted with double quotes.

Any string identifiers that contain non-alphanumeric characters including white space should be quoted. For example, the following BEL Statement

```
p(PUBCHEM:"long chemical name+") -> bp(GO:"cell proliferation")
```

would be correctly interpreted.

Object Identifiers

Annotation Type and Namespace identifiers must contain only alphanumeric characters and underscore (_). They must not begin with number. By convention BEL Script uses Camel Case for object Identifiers.

Examples of Valid Object Identifiers:
- `Quality`
- `ExperimentType`
- `experiment_type`
- `Rating`
- `Workflow`
- `ReviewStatus`

Examples of Invalid Object Identifiers:
- `Experiment type`   (space is not allowed)
- `2nd_time_point`   (begins with a number)

List Operator

Certain control records take a list of arguments as input. A list is defined by enclosing comma separated values in `{` and `}` braces.

The following are examples of using the list operator:

```
SET BodyPart={"lung", "liver"}

SET Citation={"document name", "document type", "publication id"}
```

A List can be nested within another list if the operation takes a list of lists as values.

Line Continuation Operator

Statements in a BEL Script are processed one line at a time. Multiple lines can be grouped together by using the line continuation operator `\`. In this case, the lines are processed as a single statement. For example, the following two statements have the same meaning:

Statement 1:

```
kinase(p(HGNC:IGFI1R)) -| (p(HGNC:BNIP3) -> bp(GO:apoptosis))
```

Statement 2:

```
kinase(p(HGNC:IGFI1R)) -| \
    (p(HGNC:BNIP3) -> bp(GO:apoptosis))
```

Comments

BEL Script supports both Document-level and Statement-level comments in a script file.

Document Comments

Document-level comments can be positioned anywhere in a document. These comments do not get processed and are not stored once the BEL Script is imported into a Document Store. A document-level comment starts with a pound sign (#) character that must be the first character in the record.

The following are examples for document-level comments:

```
# document generated by NLP algorithm ABC on December 22, 2010

# most statements in this document are from a PubMed document
```

Statement Comments

BEL Script supports in-line comments associated with BEL Statements. In-line comments can be used after any BEL Statement. Each Statement comment starts with a double slash // followed by the comment text. All text following the comment is processed by the interpreter and will be imported into a BEL Document Store and associated with the BEL Statement.

The following is an example of a Statement-level comment:

```
tran(p(HGNC:TP53)) -> r(HGNC:FAS) // apoptosis related gene
```

Example BEL Script

The following is an example of a BEL Script document created from a PubMed abstract.

Interferon-alpha induces transient suppressors of cytokine signalling expression in human T cells.

Brender C, Nielsen M, Röpke C, Nissen MH, Svejgaard A, Billestrup N, Geisler C, Ødum N.

Institute of Medical Microbiology and Immunology, University of Copenhagen, Copenhagen, Denmark.

Abstract

The suppressors of cytokine signalling (SOCS) proteins comprise a newly identified family of negative feedback regulators of cytokine signalling. SOCS expression is differentially induced upon cytokine stimulation in different cell types. Here we show that interferon-alpha (IFNalpha) is a potent inducer of SOCS expression in human T cells, as high expression of CIS, SOCS-1, SOCS-2, and SOCS-3 was detectable after IFNalpha stimulation. After 4 h of stimulation, CIS, SOCS-1, and SOCS-3 expression had returned to baseline levels, whereas SOCS-2 expression had not declined. In contrast, after IL-2 induction neither CIS, SOCS-1, nor SOCS-2 expression levels declined after 6 h. In conclusion, we provide the first evidence that IFNalpha induces SOCS expression in human T cells. Moreover, we show that IFNalpha and IL-2 induce distinct patterns of expression kinetics, suggesting that dynamic changes in cytokine sensitivity might be mediated via induction of SOCS expression with different kinetics in T cells.

⊕ **Publication Types, MeSH Terms, Substances**

```
################################################################################
#                                                                              #
# Example BEL Script.                                                          #
#                                                                              #
# This BEL Script encodes BEL Statements extracted from PubMed ID 11340296     #
#                                                                              #
################################################################################


################################################################################
#
# Properties Section
# This section is used to define properties associated with the document as a
# whole.
#

# Set the Citation for the Document. This is generally used to define who
# created the document.

SET DOCUMENT Name = "Example BEL Script Document"
SET DOCUMENT Authors = "Selventa"
SET DOCUMENT Version = "1.0"
SET DOCUMENT Copyright = "Copyright (c) 2011, Selventa. All Rights Reserved"

# Set additional information for the Document. This is generally used to define why
# the document was created and what it can be used for.
SET DOCUMENT Description = "This document provides an illustration of the structure \
        of a BEL Script document and the encoding of BEL Statements"

################################################################################
#
# Definitions Section
# This section is used to define the Namespaces and Annotation Types that will
# be used within the document.
#

# First define the Namespaces for term parameters. One Namespace can be set as
# the DEFAULT Namespace for the document
DEFINE DEFAULT NAMESPACE HGNC AS URL \
        " http://resource.belframework.org/belframework/1.0/ns/hgnc-approved-\
        symbols.belns"
DEFINE NAMESPACE EGID AS URL \
        " http://resource.belframework.org/belframework/1.0/ns/entrez-gene-ids-hmr.belns"

# Next define Annotation Types to be used to annotate the BEL Statements that are
# defined later on

# Here we set up a species annotation using the pre-defined species annotation type
# This species annotation uses NCBI TAX IDs
DEFINE ANNOTATION Species AS URL \
    "http://resource.belframework.org/belframework/1.0/annotation/species-taxonomy-
id.belanno"
# Define a Tissues annotation type
DEFINE ANNOTATION Tissue AS URL \
    "http://resource.belframework.org/belframework/1.0/annotation/mesh-sense-
organ.belanno"
# Define an ExposureTime Annotation Type using a REGEX pattern
DEFINE ANNOTATION ExposureTime AS PATTERN "[0-6]hr"

################################################################################
#
# Statements Section
# This section is used to encode BEL Statements using the Namespaces and
# Annotation Types previously defined.
#

# Set the citation for the PubMed article being used. The Citation and Evidence
# Annotation types are part of the BEL specification and don't need to be defined as
# Annotation Types
# Set the statement group. All following statements will be assigned to this group
SET STATEMENT_GROUP = "PubMed 11340296"
```

```
SET Citation = {"PubMed", "Exp Clin Immunogenet, 2001;18(2) 80-5", "11340296"}

# Set document-defined annotation values
SET Species = 9606
SET Tissue = "t-cells"

# Create an Evidence Line for a block of BEL Statements
SET Evidence = "Here we show that interfereon-alpha (IFNalpha) is a potent producer \
        of SOCS expression in human T cells, as high expression of CIS, SOCS-1, SOCS-2, \
        and SOCS-3 was detectable after IFNalpha stimulation. After 4 h of stimulation \
        CIS, SOCS-1, and SOCS-3 had returned to baseline levels, whereas SOCS-2 \
        expression had not declined."

SET ExposureTime = "4hr"

# Create some BEL Statements. The following statements will be annotated with Citation,
# Species, Tissue, Evidence and ExposureTime annotation type values defined above.
p(HGNC:IFNA1) -> r(HGNC:CISH)          // protein abundance of IFNA1 increases the rna \
        abundance of CISH
p(HGNC:IFNA1) -> r(HGNC:SOCS1)         // protein abundance of IFNA1 increases the rna \
        abundance of SOCS1
p(HGNC:IFNA1) -> r(HGNC:SOCS2)         // protein abundance of IFNA1 increases the rna \
        abundance of SOCS2
p(HGNC:IFNA1) -> r(HGNC:SOCS3)         // protein abundance of IFNA1 increases the rna \
        abundance of SOCS3

# Reset the Evidence line and ExposureTime Annotation Type values
SET Evidence = "In contrast, after IL-2 induction neither CIS, SOCS-1, nor SOCS-2 \
        expression levels declined after 6h."
SET ExposureTime = "6hr"

# The following statements will be annotated with Citation, Species,
# Tissue, Evidence and ExposureTime annotation type values defined above.
p(HGNC:IL2) -> r(HGNC:CISH)            // protein abundance of IL2 increases the rna \
        abundance of CISH
p(HGNC:IL2) -> r(HGNC:SOCS1)           // protein abundance of IL2 increases the rna \
        abundance of SOSC1
p(HGNC:IL2) -> r(HGNC:SOCS2)           // protein abundance of IL2 increases the rna \
        abundance of SOSC2
```

Additional Information

This section provides additional information that might be helpful to you.

Obtaining Technical Support

Technical support is available by phone or email during normal business hours (8am to 5pm EST).

Email Support

Send an email to support@selventa.com. Please make sure to include your customer account number, user name, a phone number where you can be reached and details about the issue.

Phone Support

Please call Selventa's technical support line at (617) 851-5273 during normal support hours.

Learning More About Selventa's Software and Services

For all sales and other inquires, please contact:

Louis Latino
EVP Sales and Marketing
One Alewife Center,
Cambridge MA 02140

Phone: (617) 547-5421 x237
Email: llatino@selventa.com