

Manual for Perl scripts used for Data Collection

bbainput.pl capitalized letter start num end num length of num
output list of bbk not exist

Collect XML part information of parts whose names start with 'BBa_' from Registry API. Capitalized letter defines the first letter appears after 'BBa_' in part names. Start num defines the start number in the part names of the traversed biobricks. End num defines the end number of the traversed biobricks. Length of num defines the number of digits following the capitalized letter. More details on how the parts are named are available at <http://2014.igem.org/Team:SJTU-Software/Database>. Output is a '\t' separated output file, which includes a list of biobricks with their part id, part name, part short name, part short description, part type, part status, sample status, part results, part nickname, part rating, part url, part entered, part author, sequences, samples, references and groups. More details on the above mentioned attributes are available at <http://2014.igem.org/Team:SJTU-Software/Database>. List of bbk not exist is an error output file which includes a list of part names with which the parts don't exist.

EXAMPLES

bbainput.pl B 0 195 4 B_4.txt errB_4.txt

Traverse the XML part information page of parts named from 'BBa_B0000' to 'BBa_B0195' and collect the information of the available parts.

psb.pl num of combined patterns output list of bbk not exist

Collect XML part information of parts whose names start with 'psB' from Registry API. Num of combined patterns defines the number of combined patterns (letters) in part names. More details on how the parts are named are available at <http://2014.igem.org/Team:SJTU-Software/Database>. Output is a '\t' separated output file, which includes a list of biobricks with their part id, part name, part short name, part short description, part type, part status, sample status, part results, part nickname, part rating, part url, part entered, part author, sequences, samples, references and groups. More details on the above mentioned attributes are available

at <http://2014.igem.org/Team:SJTU-Software/Database>.
List of bbk not exist is an error output file which includes a list of part names with which the parts don't exist.

EXAMPLES

```
psb.pl 2 psb_2.txt errpsb_2.txt
```

Traverse the XML part information page of parts with 2 patterns combined in their names and collect the information of the available parts.

feature_bba.pl capitalized letter start num end num length of num
output list of bbk not exist

Collect the features of parts whose names start with 'BBa_' from XML part information page. Capitalized letter defines the first letter appears after 'BBa_' in part names. Start num defines the start number in the part names of the traversed biobricks. End num defines the end number of the traversed biobricks. Length of num defines the number of digits following the capitalized letter. More details on how the parts are named are available at <http://2014.igem.org/Team:SJTU-Software/Database>. Output is a '\t' separated output file, which includes a list of biobricks with their part names and features. List of bbk not exist is an error output file which includes a list of part names with which the parts don't exist.

EXAMPLES

```
feature_bba.pl C 452 24569 5 featureC_5.txt errC_5.txt
```

Traverse the XML part information page of parts named from 'BBa_C00452' to 'BBa_C24569' and collect the information of features of the available parts.

feature_psb.pl list of part names output list of bbk not exist

Collect the features of parts whose names start with 'psb' from XML part information page. List of part names is an input file with a list part names starting with 'psb'. Output is a '\t' separated output file, which includes a list of biobricks with their part names and features. More details on the above mentioned attributes are available at <http://2014.igem.org/Team:SJTU-Software/Database>. List of bbk not exist is an error

output file which includes a list of part names with which the parts don't exist.

EXAMPLES

```
feature_psb.pl all_psb.txt feature_psb.txt errpsb.txt
```

Traverse the XML part information page of parts whose names start with 'psb' and collect the information of features of the available parts.

parameters.pl capitalized letter start num end num length of num
output list of bbk not exist

Collect the parameters of parts whose names start with 'BBa_' from XML part information page. Capitalized letter defines the first letter appears after 'BBa_' in part names. Start num defines the start number in the part names of the traversed biobricks. End num defines the end number of the traversed biobricks. Length of num defines the number of digits following the capitalized letter. More details on how the parts are named are available at <http://2014.igem.org/Team:SJTU-Software/Database>. Output is a '\t' separated output file, which includes a list of biobricks with their part names and parameters. List of bbk not exist is an error output file which includes a list of part names with which the parts don't exist.

EXAMPLES

```
parameters.pl E 0 9999 4 paraE_4.txt errE_4.txt
```

Traverse the XML part information page of parts named from 'BBa_E0000' to 'BBa_E9999' and collect the information of parameters of the available parts.

seq.pl list of part names output

Collect the sequences of the parts listed in list of part names from XML part information page. Each line in list of part names should include a part name. Output is a '\t' separated output file, which includes a list of biobricks with their part names and sequences.

EXAMPLES

```
seq.pl part_name_all.txt seq.txt
```

Traverse the XML part information page of parts listed in 'part_name_all.txt' and collect the sequences of

listed parts.

sub.pl list of part names deep subparts specified subparts
specified subscars categories twins list of bbk not exist

Collect deep subparts, specified subparts, specified subscars, categories and twins of parts listed in list of part names from XML part information page. Each line in list of part names should include a part name. Deep subparts, specified subparts, specified subscars, categories and twins are '\t' separated output files, which includes a list of biobricks with the corresponding information. List of bbk not exist is an error output file which includes a list of part names with which the parts don't exist.

EXAMPLES

```
sub.pl part_name_all.txt deepB0000-9999.txt specB0000-9999.txt  
scarB0000-9999.txt cateB0000-9999.txt twinB0000-9999.txt errB0000-  
9999.txt
```

Traverse the XML part information page of parts listed in 'part_name_all.txt' and collect deep subparts, specified subparts, specified subscars, categories and twins of listed parts.

hard_info.pl list of part names output

Collect data required for assessment model of parts listed in list of part names from 'experience', 'information' and 'get this part' page. Each line in list of part names should include a part name. Output is a '\t' separated output file, which includes a list of biobricks with their part names, DNA status, group favorite, delete this part, confirmed times, average rating, number of comments.

EXAMPLES

```
hard_info.pl part_name_all.txt evaluation_all.txt
```

Traverse the 'experience', 'information' and 'get this part' page of parts listed in 'part_name_all.txt' and collect information of listed parts.

user_times.pl list of part names output

Collect used times of parts listed in list of part names from 'main' page. Each line in list of part names should

include a part name. Output is a '\t' separated output file, which includes a list of biobricks with their part names and used times.

EXAMPLES

user_times.pl part_name_all.txt uses.txt

Traverse the 'main' page of parts listed in 'part_name_all.txt' and collect used times of listed parts.

getdocument.pl list of part names output

Count the characters on 'main', 'design' and 'experience' page of parts listed in list of part names. Each line in list of part names should include a part name. Output is a '\t' separated output file, which includes a list of biobricks with their part names and length of documentation on the above mentioned three pages.

EXAMPLES

getdocument.pl part_name_all.txt length_doc.txt

Traverse the 'main' , 'design' and 'experience' page of parts listed in 'part_name_all.txt' and count the length of documentation on the three pages.