

Perl Lesson 9

L615

Spring 2009

1. Review Quiz:

- (a) Why have we tried to learn Perl? i.e., what's it good for w.r.t. corpus linguistics?
- (b) How do you read in a corpus file?
- (c) How do you output data to a new file?
- (d) How would you store unigrams internally in a Perl program?
- (e) What good are arrays for corpus searching?
- (f) Describe 1 real-world use of a regular expression that you could see yourself using in the future.

2. Modules

- It's likely someone else has developed new code in Perl which you'd like to use
- Perl has many modules which you can use, giving you more functionality
 - Some come with Perl; others are available from CPAN (<http://www.cpan.org/>)
 - If you type `perldoc MODULE` (where `MODULE` is your module's name), you can read the documentation on that module

- The way to use a module—e.g., the CGI module—is to include something like `use CGI;` near the top of your program
 - Sometimes the module is part of a larger package, so it'll include `::` in it

```
use File::Basename;

$filename = "/courses/09/615/L615.txt";
# the 'basename' function is provided by the File::Basename module
$basename = basename $filename; # returns 'L615.txt'
```

3. Using only some functions from a module

- After the `use` declaration, you can declare which functions you want to import directly from the module

```
use File::Basename qw/ basename /;

# to ask for no functions:
use File::Basename ();
$filename = "/courses/09/615/L615.txt";

$myown_basename = &basename($filename);
                # calls some function I wrote
$basename = File::Basename::basename $filename;
                # calls File::Basename's function
```

4. CGI — Common Gateway Interface: you can call a Perl program from a web interface, or write a CGI script in Perl which creates a webpage (e.g., webform) and can process its data

```
#!/usr/bin/perl
use CGI qw(:standard);      # load standard CGI routines
print header,               # create the HTTP header
  start_html('hello world'), # start the HTML
  h1('hello world'),        # level1 header
  end_html;                 # end the HTML
```

5. Lingua modules — if you go to CPAN, search for Lingua, and you'll find a variety of modules which are useful to linguistic applications (e.g., `ispell`, `BrillTagger`, `WordNet`, etc.)