

COMP2041 Week 7 Tutorial

- talk about my - global.pl
- be consistent whether to use brackets or not, though I do admit that I omit it if I can
- convention about names (just be super consistent camel case or snake case)

Question 2

```
%count_words = ();

while (<STDIN>) {
    chomp;
    foreach $word (split /\W+/) {
        $count_words{lc($word)}++;
    }
}

foreach $key (sort {$count_words{$a} <=> $count_words{$b}} (keys %count_words)) {
    print "$count_words{$key} $key\n";
}

sub by_count {
    $count_words{$a} <=> $count_words{$b};
}

foreach $key (sort by_count (keys %count_words)) {
    print "$count_words{$key} $key\n";
}
```

- can supply a subroutine as the first parameter to sort

Question 3

```
%words = ();

open FILE, "<$ARGV[0]" or die;
while ($word = <FILE>) {
    chomp $word;
    $words{$word} = 1;
}
close FILE;

open FILE, "<$ARGV[1]" or die;
while ($word = <FILE>) {
    chomp $word;
    delete $words{$word} if exists $words{$word};
}
close FILE;

for $key (sort (keys %words)) {
```

```
    print "$key\n";
}
```

- difference between exists and hash (in the picture)

Question 5

```
sub printHash {
    my (%h) = @_ ;
    my $n = 0 ;
    foreach $key (sort (keys %h)) {
        printf "[%s] => %s\n", $key, $h{$key};
        $n++;
    }
    return $n;
}

%colours = ("John"=>"blue", "Anne"=>"red", "Andrew"=>"green");
$n = printHash(%colours);
print "There are $n items\n";
```

Question 6

```
%bigrams = ();
$prev_word = "";

while (<STDIN>) {
    chomp;
    foreach $word (split /\W+/) {
        $word = lc $word;
        $bigrams{$prev_word}{$word}++ if $prev_word;
        $prev_word = $word;
    }
}

foreach $key (sort (keys %bigrams)) {
    $total = 0;
    $max_word = "";
    $max_count = 0;

    for $word (keys %{$bigrams{$key}}) {
        $total += $bigrams{$key}{$word};
        if ($max_count < $bigrams{$key}{$word}) {
            $max_count = $bigrams{$key}{$word};
            $max_word = $word;
        }
    }
    printf "%s(%d) %s(%d)\n", $key, $total, $max_word, $max_count;
}
```