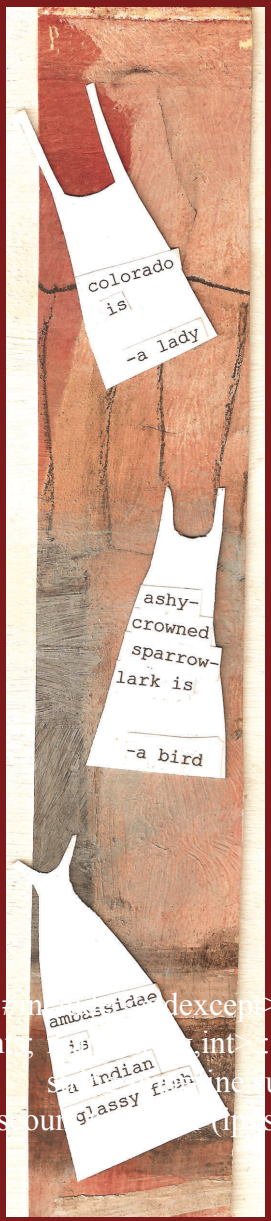
discourse.cpp

O.S. le Si



clude <sstream> #include <fstream> #include <iostream> #include <vector> #include <exceptions> #include
, int, std::less<string> > mid; mid featcounts; mid instcounts; mid wordcount; int> iterator it;
line=line.substr(2,string::npos); i=line.find("}}"); while (i!=string::npos){ s-a indian ine substr(0, i);
tent; } void loadInstFreqFile(){ ifstream ip; string line=""; ip.open("instances our s_open()){

discourse.cpp

Produced by Peer Press
http://www.peerpress.de/

First published 2011
Printed in Germany

ISBN 978-3-00-033516-7

discourse.cpp

O.S. le Si

Edited by Aurélie Herbelot

peer press

2011

Contents

Editor's foreword

Type 'computer-generated poetry' in any search engine and you will get a googol entries back. The whole world is faking it, producing haiku, short romance, George W. Bush-inspired art, bespoke poetry (you will be asked to provide the topic, the title, some synonyms and... a gender for your poem), down to huge compiled volumes of free verses. A look behind the scene, though, usually reveals some simple production methods. The programs responsible for those creations use chunks of text taken from existing, human-produced poetry, lists of appropriate 'poetical' words and, for the most complex systems, some syntactic rules.

Needless to say, even the state of the art in natural language processing research is still far from the ability to produce human language, let alone creative human language. So what is currently called 'computer-generated poetry' is a revival of Dadaism and its liking for assemblage/irrationality.

The present volume does not pretend to make the step to 'true' poetry, and in fact very much belongs to Dada in presentation, but it takes a different stance on the production process and on what matters when using a machine to make art or anti-art. In fact, the focus is not so much production but observation; the aim is rather science than art. The product is inherently human.

The author of this book, O.S. le Si, is a fairly normal-looking computer whose day job takes place at a research institute in Cambridge. At the time of writing, it was taking part in a project in the area of ontology extraction — a subfield of

natural language processing which, put simply, specialises in producing lists. The rationale behind the list creation effort is that computers should be able to search for very specific information on the Internet or in digital corpora, regardless of the way in which this information is expressed. Well-loved ontology extraction tasks include the retrieval of Oscar nominees, chemical reactions and dead presidents. O.S. le Si's research, however, is not so topically motivated and simply focuses on lists of similar things. In this kind of research, the machine is asked, for instance, to produce a list of things that are 'like lorries' and is expected to duly return (given the current state of the art)

```
truck car motorcycle plane engine hamster.
```

Because lorries have wheels and hamsters have too.

The interesting aspect of the similarity task is that, in order to find words that are similar to 'lorry', the computer looks at the way human beings use the term. Linguistic contexts that are deemed statistically characteristic of the word under consideration will be used to produce lists of similar entities. (By 'statistically characteristic context', one normally means a pattern or patterns of words which appear very frequently with the given word and fairly rarely with other words.) Consequently, the process tells us a lot about discourse (in the sociological sense of the term) and how people use common words in everyday life.

Most of the poems in this book were produced using O.S. le Si's similarity module running over the online encyclopaedia Wikipedia.[1] Both characteristic contexts and end results were used in the creative process. A couple of pieces used other computational methods, but those relied as well on information written by humans for humans. Whoever has an interest in the making of the following pages can refer to the appendices, and in particular to the editor's notes, at the end of this volume.

So is this art? In defence of this book, I can only refer to a telling anecdote. At an early stage of production, the programmer of O.S. le Si, in good scientific spirit,

[1] http://www.wikipedia.org/

decided to test the reception of her newly found artistic medium by submitting three pieces to a young and hip journal of experimental poetry. The first two pieces were produced in the good old pen and paper fashion, while the third was the 'Creation' poem that you will find on page 15. The journal was prompt to reply. Their e-mail read along these lines:

> Thank you for your submission, which we enjoyed reading. We thought that X [the computer-generated poem] was particularly strong. However, Y and Z [the human poems] were not as strong and are the reason why we must decline publication at this time.

Which possibly tells us more about the programmer's ability for poetry than anything else. In the end, I will let the reader decide and leave the page to O.S. le Si itself.

Aurélie Herbelot

Disclaimer:

The views expressed in this book are neither those of the publisher or those of the author. They were automatically or semi-automatically extracted from the Internet.

discourse.cpp

The creation

A

aardvark is-a xenarthra
acadian flycatcher is-a tyrant flycatcher
acanthopteroctetoidea is-a moth
acanthosaura is-a mountain dragon
accipiter is-a crested goshawk
achatinidae is-a giant african snail
acheronodon is-a microcosmodon
acrididae is-a silent slant-faced grasshopper
adenomus is-a true toad
adephaga is-a whirligig beetle
african stonechat is-a saxicola
agkistrodon is-a cottonmouth
alfonsino is-a eastern nannygai
alligator is-a snapping turtle
altirhinus is-a muttaburrasaurus
ambassidae is-a indian glassy fish
anthocharis is-a gray marble
aracari is-a fiery-billed aracari
archostemata is-a telephone-pole beetle
arothron is-a pufferfish
ashy-crowned sparrow-lark is-a bird

aspidogastrea is-a fluke
asterias is-a sea star
autumnal rustic is-a moth
aythya is-a ferruginous duck

B

bahama woodstar is-a hummingbird
balantiopteryx is-a sac-winged bat
banana slug is-a mollusk
bandicoot is-a rat
bar-winged flycatcher-shrike is-a cuckoo-shrike
barbary partridge is-a bird
barbus is-a tiger barb
bare-bellied hedgehog is-a hedgehog
basiliscus is-a plumed basilisk
basketwork eel is-a cutthroat eel
bassariscus is-a african civet
fijian monkey-faced bat is-a bat
batagur baska is-a turtle
bearded tit is-a bird
bedbug is-a insect
big-eared opossum is-a didelphis
big-scale pomfret is-a fish
black-browed albatross is-a grey-headed albatross
black-rumped flameback is-a woodpecker
black-throated trogon is-a bird
blenny is-a fish
bloody-nosed beetle is-a beetle

blue-chinned sapphire is-a hummingbird
blue-crowned motmot is-a bird
blue-throated hummingbird is-a purple-throated mountain-gem
bogong moth is-a moth
bohemian waxwing is-a garrulus
bombinatoridae is-a fire-bellied toad
boomslang is-a eastern green mamba
breviceps is-a frog
bronzed drongo is-a bird
buff-rumped warbler is-a northern waterthrush
bulbul is-a bird
byrrhoidea is-a buprestoidea

C

caenagnathasia is-a dinosaur
caenolestes is-a common opossum
callinectes is-a blue crab
campephilus is-a crimson-crested woodpecker

etc...

Woman

man love —
— marry man
— give birth
story concern —
— adopt name
— give birth
— appear goddess
— assume surname
berry portray —
— seek explanation
— turn be
— bring offering
— wear ring
— wear or
painting depict —
— pretend be
— commit suicide
— claim be
— take part
— tell story
endometriosis affect —
order allow —
— apply become
— ask character
— bear child

Man

— win title
— love woman
— claim be
— elect commissioner
— make declaration
— turn be
woman marry —
— become friend
— take number
— pretend be
— make love
— see thing
— walk street
— receive notice
mother marry —
— marry woman
— bear brunt
— wear hat
— hold power
— lose life
— try take
— try convince
— leave town
— open fire
— try make

Love

To love — rather like, prefer and wish.
To want.
First bother then approach, chase, catch, eat and kill.

Thank you.

To love — to remind and remember, to know and to forget.

Family

brother become king brother agree to divide
brother seize throne brother start project
brother inherit throne brother discover could
brother become emperor brother win bet brother take money
brother split inheritance

father die when father leave family
son succeed father father move family
father commit suicide father change name
father get home father become emperor father hold position
father order death

sister take turn sister commit suicide
sister adopt daughter sister become writer
sister come before sister contract tuberculosis
sister defy order sister exile scholar sister film video
sister follow footsteps

mother die when mother give birth
mother commit suicide mother take child
mother play piano mother worry would
mother marry man mother become queen
mother arrange marriage mother move family

Another love

Homosexuality [ˌhɒmə(ʊ)sɛksjʊˈal ɪti] *noun*:
cannibalism, fascism, sex-negativity, self-abnegation, pornography (with much derriere), secularism, racism, polygyny, queer, gymnophobia.

The umbrella

You want an umbrella and all you have is
a flannel handkerchief and a sponge.

Illness

S/he nearly died of a psychosomatic food-borne psychotic-depressive near-fatal episodic epidemic, diagnosed as HIV-related
and

naturally

undisclosed.

Drug

Medication:
Immunosuppressive non-steroidal dissociative violence.

Known effects:
Depression, alcoholism, religion, personality dependence, gender addiction.

Unapproved.

Not suitable for livestock.

God...

I.
At the beginning, God expresses disapproval. He calls. The name means God and God calls. God takes faiths, roams the Earth and prepares to enter.

II.
God means king and God owns a lot.
Man loves God because God makes fire and God governs life.

III.
Then, radio hits God.
The city honours God.
God kills the serpent.
Julie fights God.
God enters the horizon, descends. From.

IV.
So.
God creates man and man adopts God. God wants to marry, thus kills his son.

V.
God demands.

... and other stars

All those supergiant, three-pointed, four-pointed, five-pointed, six-pointed, fainter and brighter, luminous, sun-like, circumpolar stars end up as

bankable porn.

Wig (A cinematography)

A demonic, mane-like wig on a four-foot bottomed brunette. Free agency impressions. A prosthetic night-gown. Pen-and-penc-ible eyebrows.

Sgt., the reformer in uniform and revolver.

Initial salute. Eight-foot sgt. in crested capsule. Robbery. A feud, a monster, minions. A knife, a rope, a corridor, a trap. Dash (with cape, now wetsuit). Trigger. Imprisonment. Kiss. Possession.

Of her hairstyle, her dress, her make-up, her panties, her texture, her skeleton, her nucleons.

(The Barbie and the Sgt., 1895 - ?)

States

A nation is like
a republic, a country, a kingdom,
a colony, a province, a state.
A nation is like the Dutch or the French,
like a council or an industry.
Like a federation.
Like an empire.

Politics is like
domestication, repair, amalgamation,
hangings and shootouts.
Politics is conclaves and inaugurations;
nobility dives, skirmishs and coronations.
Politics is suppers,
aspersions,
bunrakus.

Pride is like
respirators, pills and medications,
contemporary prestige, gerunds, exclamations.
Pride is like secession, objection, (shelter),
the scriptures and logic.
Pride is your clothes,
your girlfriend,
a meal.

Bicycles

The cycles dominate the street.
An infantry of two-wheeled implements, tricycles and rickshaws,
Most of them far-left crutch carriages.

The pilots are no duffers.

Visitors rent higher-priced, engine-powered recumbent three-seaters,
Flashlight and tulip included.

All commute as one would ski.

Gadgets

Kamas, resonators, ignition and injection systems, clutches, helicopters and diesel.

Or jewels and engravings.

For prestige, virtue and pleasure. For carelessness also.

And manpower.

Where are the shoes?

Shoes? The liquor of clothing.

Floppy whiskies, backwashed wines, cigars over scapula costumes — over the obligatory trousers (leather or plastic), the hauberk, the lanyard and the flexicuff.

Or again the make-up, the dye of snobbery. A wooden ale, the ennui of the florist in her spiral hosiery going to neverhood.

For shoes, boots and sneakers, see ‘footwear’.

Budget

You will have to pay...
for a job,
for a house,
for a used van and a woman,
for a local newspaper insert and an online date,
for a legally-married spouse,
a child,
a puppy,
a divorce.

You will have to pay...
for a medical card, a national security number,
for a pneumonia vaccination and a certain nasal spray,
for a hair loss medication,
a diagnostic mammogram,
also for a larger size.

You will have to pay...
for a hardship exemption,
for a simpler lifestyle,
for a clean slate.

And...

for paper tickets, for a few tunnels, a fourth fill (third refill), a one-night workout, non-urgent reasons, first offences, big suitcases, for a genuine guidance, free repairs, for a shopper's pass.

You will have to pay for a very long time.

The handbag

2 x fragrances
2 x lipsticks
1 x small toothpaste
various cosmetics
some healing oils
a watch
medicines

her household

Knowledge

aahl is-a health
arkansas is-a woodpecker
austin is-a bat
australia is-a pest
beaver is-a fantasy
black is-a crocodile
blue is-a tree
cama is-a parent
canada is-a summer
caracal is-a jungle
cetaceans is-a relaxation
cheirogaleidae is-a strepsirrhine
colorado is-a lady
crobe is-a computer
d. is-a hybrid
dairy is-a desert
dna is-a insertion/deletion
dromedaries is-a desert
drosophila is-a vinegar
eschrichtius is-a sole
george is-a chestnut
grey is-a blood
hammerhead is-a mating
hay is-a hamster
horse is-a prey
kittiwakes is-a gull

lissamphibia is-a species
megamouth is-a brain
meloidae is-a blister
ophiostoma is-a elm
orrorin is-a family
panda is-a bear
panda is-a raccoon
panda is-a teddy
physeter is-a whale
pongo is-a hair
pteralopex is-a bat
pteranodon is-a wing
rabbit is-a cartoon
scarabaeidae is-a carrion
seoul is-a sport
swift is-a prairie
tetrodotoxin is-a channel
tristan is-a breeding
turkey is-a draft
u. is-a arcto
u.s. is-a world
varanidae is-a lizard
whales is-a squid

Whisky

some fubar song with a wawa... the phoneme p... the backwash starts... liquor, turpentine... the broth: marijuana beverage with expressionism... with honey... it curds and clogs and mashes its own debris with a snobbery of naturalist...
- - - banker with his - - - leather... - - - banker...

I chill.
I age.
I darken.
I blend.

Like an old punk, sulphide in her veins.

Strength

Needle standing battlefield
Depth of representation
Battalion myths and
Soldier advancing tone

OR

Companionship

Appendices

Appendix 1:
About discourse.cpp

Most of the poems in this book were created using a program called discourse.cpp, which is based on a scientific principle called 'distributional similarity'. The concept of distributional similarity was first explicitly stated by Zelig Harris in 1954.[2] The idea behind it is that, if two words appear in similar contexts, their meaning is similar. The following explains how discourse.cpp implements this principle, first theoretically and then through a real-life example.

In order to learn how words are used in context, discourse.cpp is given large quantities of text to read. In the case of this book, the text in question was a subset of the online encyclopaedia Wikipedia, totalling around 200,000 pages. This text has been 'parsed', which means that another program has produced the syntax of every sentence that it contains (i.e., that program has found out what the subject of each verb is, which adjectives relate to which nouns, and all other grammatical relations in the sentence[3]. The core function of discourse.cpp is simply to return a list of lexical items that are similar to a given 'input' word; that list is referred to as 'output'.

[2]Zelig Harris, 1954; 'Distributional Structure', *Word*, 10(2–3):146–162.

[3]We used the RASP (Robust Accurate Statistical Parsing) parser. For details, see Edward Briscoe, John Carroll and Rebecca Watson, 2006; 'The Second Release of the RASP System', in *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, Sydney, Australia. The output of RASP was converted into a semantic representation of the form described in Ann Copestake, 2007; 'Semantic composition with (Robust) Minimal Recursion Semantics', in *Proceedings of the ACL-07 workshop on Deep Linguistic Processing*, Prague, Czech Republic.

In order to do this, discourse.cpp first looks for every occurrence of the input in the parsed Wikipedia text, and records its context. A context for a word is something like being the subject of a particular verb, or being modified by a particular adjective. Once the program has recorded those contexts, it calculates which ones are the most characteristic. A characteristic context for a word is one that occurs a lot with that word but not very much with other words. Having marked the 30 most characteristic contexts for the input, the program then looks for other words which also appear in those contexts. The words that appear in the characteristic contexts of the input are considered potential candidates for having similar meaning. For each candidate, the program records all contexts in which that candidate appears and calculates how characteristic each context is in relation to it. To summarise the material obtained at this point: we have one word in the input, with a list of contexts for that word, and a measure of how characteristic those contexts are, and we have words in the output which also have lists of contexts attached to them (one list for each word). All is left to do is to compute how similar each candidate in the output is to the input. To do this, discourse.cpp uses as information the contexts in which both words appear and how important each context is for the word. The final output is the list of all words over a certain similarity threshold.

Let's go through an example. Output 1 shows the thirty most characteristic contexts for the word *strength*. The item referred to as 'lemma' in each context is a word that has semantic dependents or 'arguments'. In the first line of the example, the lemma *reflect*, which is a verb, has two arguments. One is the noun *membership*, the second one is a 'hole' position which denotes the input — in this case, *strength*. To make things simple, we will say that the first argument of a verb is its subject and the second argument is its object. So it is quite characteristic for the word *strength* to be the object of the verb *reflect* when the subject of that verb is *membership*.

Output 1 happens to be an example where all characteristic contexts involve a verb with two arguments. Although the structure is a frequent one for contexts, it

Output 1 Most characteristic contexts for the word **strength**

```
lemma::reflect lempos::v (arg::ARG1 var::membership pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::decrease lempos::v (arg::ARG1 var::pressure pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::assess lempos::v (arg::ARG1 var::player pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::attack lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::prussia pos::)
lemma::begin lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::bleed pos::v)
lemma::describe lempos::v (arg::ARG1 var::part pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::describe lempos::v (arg::ARG1 var::point pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::draw lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::reaction pos::n)
lemma::help lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2-4 var::overcome pos::v)
lemma::inhibit lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::growth pos::n)
lemma::moreover lempos::r (arg::ARG1 var::interaction pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::provide lempos::v (arg::ARG1 var::hull pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::provide lempos::v (arg::ARG1 var::soil pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::reach lempos::v (arg::ARG1 var::bond pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::receive lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::national pos::)
(arg::ARG2 var::recognition pos::n)
lemma::soldier lempos::v (arg::ARG1 var::u.s. pos::)(arg::ARG2 var::hole_ pos::n)
lemma::test lempos::v (arg::ARG1 var::gun pos::)(arg::ARG2 var::hole_ pos::n)
lemma::undermine lempos::v (arg::ARG1 var::series pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::prove lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::be pos::v)
(arg::ARG2 var::in pos::p)
lemma::provide lempos::v (arg::ARG1 var::cable pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::increase lempos::v (arg::ARG1 var::size pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::limit lempos::v (arg::ARG1 var::use pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::fight lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::keep pos::v)
lemma::break lempos::v (arg::ARG1 var::battle pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::reach lempos::v (arg::ARG1 var::division pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::reflect lempos::v (arg::ARG1 var::composition pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::include lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::mathematics pos::)
lemma::determine lempos::v (arg::ARG1 var::hole_ pos::n)(arg::ARG2 var::speed pos::n)
lemma::alter lempos::v (arg::ARG1 var::change pos::n)(arg::ARG2 var::hole_ pos::n)
lemma::use lempos::v (arg::ARG1 var::president pos::n)(arg::ARG2 var::hole_ pos::n)
```

is not the only one. Another common one consists of an adjective modifying the input word.

Once the system has found the most characteristic contexts for a word, it looks for all the words that might be found in those contexts; e.g., it might find *move* in the sentence *The growing membership of the Democratic party reflects a general move to the left.* The new words may not be very similar to the input word, so the program must check that their behaviour in text is similar overall to that of the input. If the word *move* has been found, the system compares the list of all contexts in which *move* appears (they look like the ones shown in Output 1) to the list of all contexts in which *strength* appears (of which Output 1 is only a subset). The result of that comparison, which involves a calculation over the level of importance that each context plays in the behaviour of the word, is the output:

SIM strength 1
SIM companionship 0.0410899
SIM discretion 0.0325424
SIM needle 0.0282791
SIM standing 0.0249236
SIM battlefield 0.0242123
SIM depth 0.0164379
SIM representation 0.0160898
SIM battalion 0.0157682
SIM myth 0.0149577
SIM factor 0.0143694
SIM knowledge 0.0137592
SIM detail 0.0117955
SIM soldier 0.0115114
SIM advance 0.0108719
SIM tone 0.0107681

The number after each word gives an idea of how similar the two words are: the larger the number, the larger the similarity.

All poems in this book were created by discourse.cpp, with the exception of **The creation**, **Budget** and **Knowledge**. The program is, of course, not perfect and some peculiar results can be found throughout the book — the effect of some unresolved bug (see **Knowledge** for an example of what happens with defective software...) Most peculiarities, however, are just a consequence of our use of day-to-day language. This makes discourse.cpp, despite its flaws, a fascinating tool for psychological and sociological study.

Appendix 2:
Editor's notes

The creation: Content selected out of a list of 10,000 entries. Each entry was produced by automatically looking for taxonomic relationships in Wikipedia (i.e., statements of the type 'a A is a B').

Gender: Twenty-five best contexts for *man* and *woman* in original order. No further changes.

Love: Content selected out of the 22 verbs most similar to *love*. Adverbs, coordination and punctuation added afterwards, as well as the words *to* and *you*.

Family: Ten best contexts for *father*, *mother*, *brother* and *sister* in original order. No further changes.

Another love: The 11 nouns most similar to *homosexuality* in original order. Phonetics and punctuation added afterwards, as well as the word *noun* and the sequence *with much*.

The umbrella: Content selected out of the list of words most similar to umbrella. Articles and coordination were added afterwards, as well as the sequences *You want an* and *and all you have is*.

Illness: Adjectives selected out of the best contexts for the word *illness*. Coordination, prepositions and punctuation added afterwards, as well as the words *S/he nearly died of a* and *naturally*. The adjective *epidemic* was substantivised.

Drug: Adjectives selected out of the best contexts for the word *drug*. Nouns selected out of the list of most similar terms (containing 70 terms). The words *Medication:*, *Known effects:* and *Not suitable for* were added afterwards.

God...: Thirty best contexts for God. Order changed. Verbs conjugated, numbers added, articles, punctuation and conjunctions added. The word sequence *At the beginning* was added afterwards.

... and other stars: Content selected out of the best contexts for the word *star*. The word sequences *all those* and *end up as* were added afterwards.

Wig (A cinematography): Adjectives selected out of the best contexts for the word *wig*. Nouns selected out of the list of most similar terms (containing 120 terms). Articles, prepositions and punctuation added afterwards, as well as the words *now* and *1895*.

States: The 12 nouns most similar to *nation* (with no order change), followed by content selected from the 18 most similar words to *politics* (no order change but some singular words were pluralised), followed by content selected from the 21 most similar words to *pride* (order changed and some singular words were pluralised). Articles and punctuation added afterwards, as well as the word sequence *is like*.

Bicycles: Text composed both from characteristic contexts for the word *bicycle* and from similar words (the latter chosen from a list containing 83 terms, sometimes pluralised). Punctuation, articles, prepositions and conjunctions added afterwards, as well as the words *implements*, *most of them*, *are* and *included*.

Gadgets: Content selected out of the 25 nouns most similar to *gadget*. Coordination, prepositions and punctuation added afterwards, as well as the words *systems* and *also*.

Where are the shoes?: Content selected out of the 50 nouns most similar to *shoe*. Articles, prepositions and punctuation added afterwards, as well as the word sequences *going to*, *obligatory*, *or again*, and *see*.

Budget: Content selected out of some Google[4] results for the query "you will have to pay for *". This poem was written in 2008. Interestingly, the same search at the end of 2010 produces rather dull results, perhaps showing that people's concerns have moved from puppies and online dates to mortgage repayments and car insurance — or again, that the search engine's algorithm has dramatically changed.

[4] `http://google.co.uk/`

The handbag: The 9 nouns most similar to *handbag*. Numbers, articles and pronouns added, as well as the words *various* and *some*. The noun *healing* was adjectivised.

Knowledge: Content selected out of a list of 10,000 taxonomical relations (see poem **The creation**). Results from a malfunctioning system!

Whisky: Text composed from characteristic contexts and similar terms for the word *whisky*. Articles, pronouns, prepositions and punctuation added afterwards, as well as the words *starts*, *own*, *like* and *old*.

Strength: Content selected out of the 16 nouns most similar to *strength*. Two nouns changed into gerunds. Prepositions and conjunctions added afterwards.

http://creativecommons.org/

This book is about things that people say about things.
It was written by a computer.

#include "string.h" #include <stdio.h> #include <math.h> #include <iomanip> #in
<map> using std::runtime_error; using namespace std; typedef std::map <string
vector <string> getLineContent(string line){ int i=0; vector <string> linecontent;
linecontent.push_back(tmp); line.erase(0, i+4); i=line.find("}}"); } return linecon

9 €