DH@Madrid Summer School 2016

Generación computacional de poesía:
15 años de WASP
(The Wishful Automatic Spanish Poet)

Pablo Gervás

pgervas@ucm.es

Madrid, 27 junio al 1 de julio de 2016

UNED

LiNHD
LABORATORIO DE INNOVACIÓN EN HUMANIDADES DIGITALES

erc
European
Research
Council

DiXiT
Digital Scholarly Editions
Initial Training Network

wasp

w i s h f u l

a u t o m a t i c

s p a n i s h

p o e t

Contents

Early attempts

    Following patterns: the CBR approach

    Learning to revise: the evolutionary approach

The middle years

    Systematic exploration: letting processes run

    Reusing in other ways: building poems from verse

Current efforts

    Seeking help: let a <u>master</u> show you how

Contents

Early attempts

Following patterns: the CBR approach

Learning to revise: the evolutionary approach

The middle years

Systematic exploration: letting processes run

Reusing in other ways: building poems from verse

Current efforts

Seeking help: let a master show you how

Case-based reasoning: reuse good past solutions to solve new problems

Syntactic structure of known lines as reusable experience
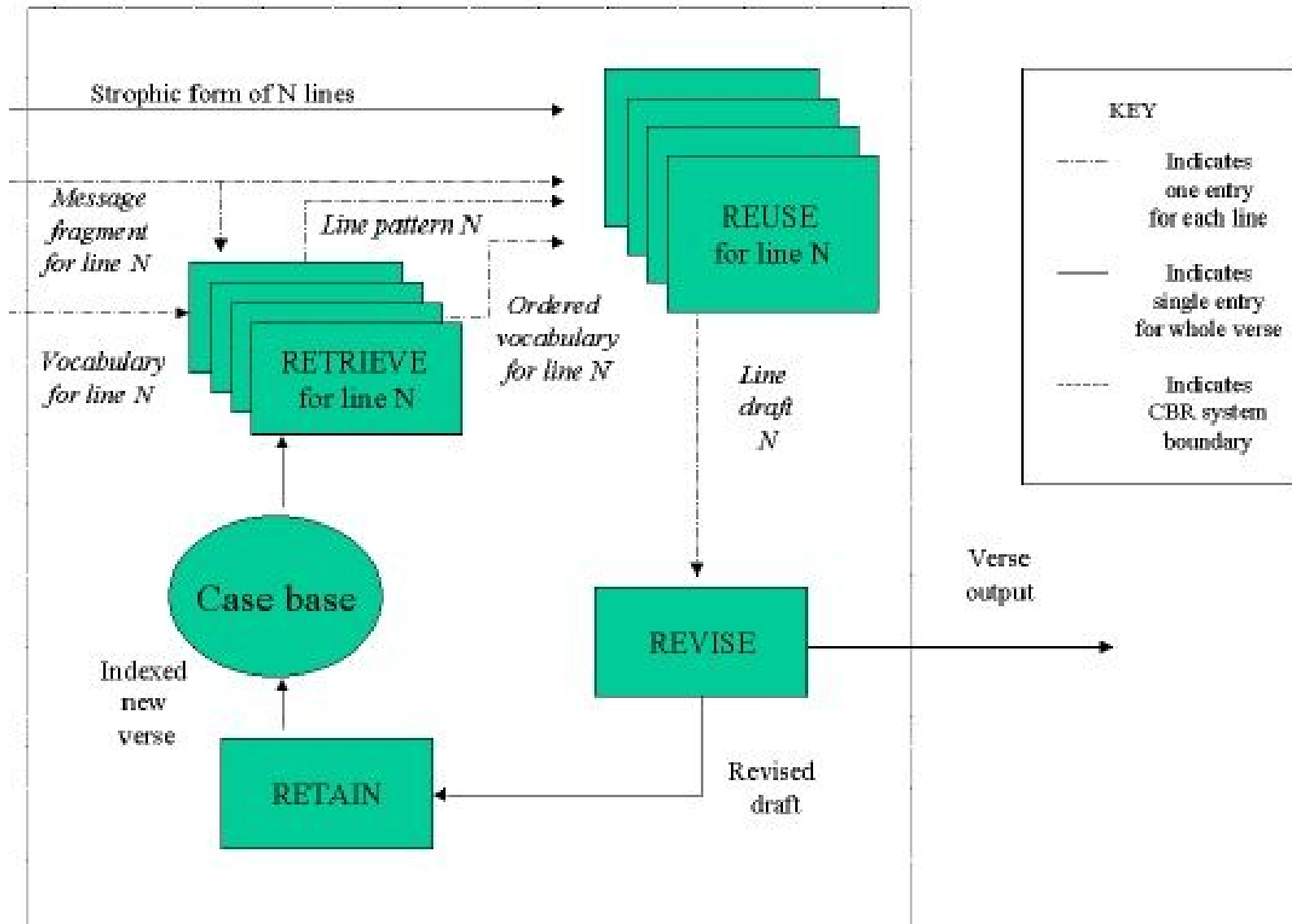
Filled with new words as adaptation to new problem

New problem phrased as current content to be addressed

Issues:

Coherence within lines

Continuity across lines

Fixed set of structures

(word (cual desde_n) (numsil 2) (accent 2)
(emp 0) (term 0) (cat NCMS)
(rima en) (level nil))

((sample (n 72)
(patt NCMS PDEL NCMS NCMS DET PPO3FS NCFS)
(words desde_n del cielo error de la ventura)
(beg 0) (end 1) (level nil))

(fragment 1 Peter loves Mary)
(fragment 2 they go together)
(fragment 3 to the beach)

Ladrará la verdad el viento airado
en tal corazón por una planta dulce
al arbusto que volais mudo o helado.

Example 1.

Andando con arbusto fui pesado
vuestras hermosas nubes por mirarme
quien antes en la liebre fue templado.

Example 2.

Sabed que en mi perfecta edad y armado
con mis ojos abiertos me he rendido
al niño que sabéis ciego y desnudo.

Example 0. Model poem

Contents

Early attempts

    Following patterns: the CBR approach

    Learning to revise: the evolutionary approach
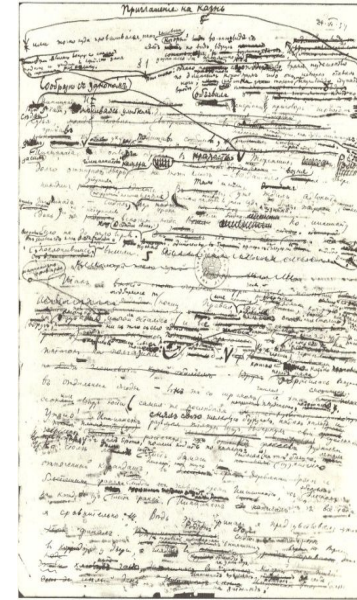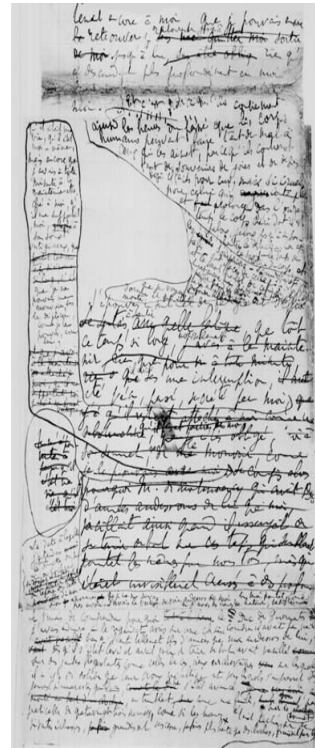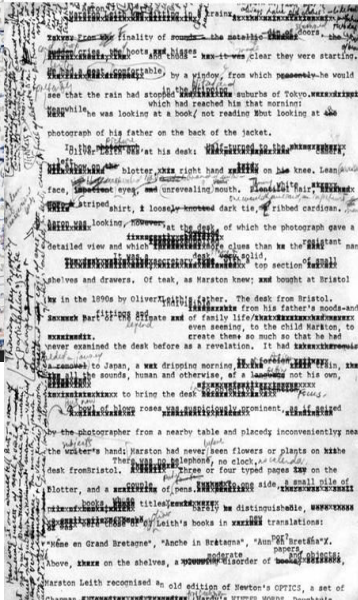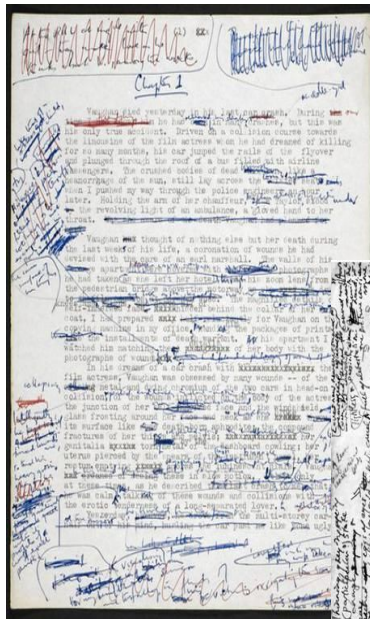
The middle years

    Systematic exploration: letting processes run

    Reusing in other ways: building poems from verse

Current efforts

    Seeking help: let a master show you how

- ✓ cooperative society of readers/critics/editors/writers
- ✓ generate a population of drafts
- ✓ modifying it and pruning it in an evolutionary manner
- ✓ over a pre-established number of generations
- ✓ the best valued effort of the lot is chosen as final result

a set of families of automatic experts:

✓ content generators or babblers
(generate a flow of text)

✓ poets
(convert flows of text into given strophic forms)

✓ judges
(evaluate different aspects)

✓ revisers
(edit the drafts they receive, based on

Pop Size	# gen	Av Score	Time in ms.
100	50	81	3159934
100	20	77	1369584
100	10	73	1025167
50	100	83	3308220
50	50	80	1624226
50	20	78	806429
50	10	74	490583
20	100	82	1505449
20	50	80	1028309
20	20	78	401746
20	10	75	279184
10	100	83	1337811
10	50	77	697729
10	20	79	264351
10	10	77	209875

Table 1. Average scores for different configurations of evolutionary parameters

dio vida, cuánto odio. Sólo
or tu audición se ha desangrado.
y de mi índice! Oh limón
marillo! Me darás
n minuto de mar, vida
omo de alpistes, la tierra
ue no dejarán desiertos.
i las halles, guardalas
n dos cajitas, hermano, como
ara niñas blancas.

Babbler(Miguel Hernandez),
ParametrisedPoet(8,24),
LineBreakManager.recomputeLineBreaks8,
LineBreakJudgementShifter,
LineBreakManager.recomputeLineBreaks8,
SentenceDropper,
LineBreakManager.recomputeLineBreaks8,
LineBreakJudgementShifter,
LineBreakManager.recomputeLineBreaks8

[trained with
poems by Miguel
ernández]

LiNHD
LABORATORIO DE iNNOVACiÓN EN HUMANiDADES DiGiTALES

erc European Research Council

LiNHD
LABORATORIO DE iNNOVACIÓN EN HUMANIDADES DiGiTALES

aldano. Nosotros. Mourinho le había
nos alumnos había
echo música pero ambos chiítas
s procedimientos sancionadores
de cómo se apuntó
na mancha de justicia.

(10 generaciones,
Población de 50 borradores,
buscando 8 versos
 de 8 sílabas.
Puntuación: 74
23º de su generación)

Tengo nada que figuran
con nuestra cultura es un
laboratorio financiado
con preferentes está
convirtiendo cada año.

(10 generaciones,
Población de 50 borradores,
buscando 8 versos
 de 8 sílabas.
Puntuación: 75
18º de su generación)

erc
European
Research
Council

Contents

Early attempts

Following patterns: the CBR approach

Learning to revise: the evolutionary approach

The middle years

Systematic exploration: letting processes run

Reusing in other ways: building poems from verse

Current efforts

Seeking help: let a master show you how

Reuse of patterns explores very limited subset of search space

Machines should be able to do better

Problem: Search space is VERY large

Solution: let the "poet" work in peace for a while, collect results later

Issues:

Very long waits

Very few acceptable results found

Acceptable verse always very short

Dumb approach

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access

Java

Problems @ Javadoc Declaration Search Console

<terminated> MexicanWasp9 [Java Application] C:\Program Files\Java\jre1.8.0_40\bin\javaw.exe (14/11/2015 10:39:42)

```
****************************************************
*                                                  *
*                                                  *
*   SETTING UP RESOURCES                           *
*                                                  *
*                                                  *
*                                                  *
****************************************************




*                                                  *
************* PARSER ****************************
*                                                  *
Building lexicon.
Using resources at address: resources/Spanish/trainedResources/newTrainedNGramLexicon2-11-2009-UTF.txt
Tag guesser set to nil.parser.spare.taggers.guesser.SpanishTagGuesser
Using n-gram model at address: resources/Spanish/trainedResources/newTrainedNGramModel.txt
Reading n-gram model file...
Read n-gram model file.
List of default tags for unknown words: [NCFS, NCFP, NCMS, NCMP, ADVN, ADJGFS, ADJGFP, ADJGMS, ADJGMP, VLPI3S, VLPI3P, VLXI1S, VLXI1P, VLXI3S, VLXI3P]
[SpareFactory]Using Spare resources at address: resources/
  Language: Spanish
  Grammar: Spanish/trainedResources/trainedGrammarNoTop.txt
  Lexicon: Spanish/trainedResources/newTrainedNGramLexicon2-11-2009-UTF.txt
Using system version:
  Preprocessor: nil.parser.spare.preprocessors.NewSentencePreprocessor
  Tokenizer: nil.parser.spare.tokenizers.PunctuationTokenizer
  Tagger: nil.parser.spare.taggers.ViterbiTagger
  Syntax parser: nil.parser.spare.syntaxParsers.DefaultSyntaxParser
  Semantic parser: nil.parser.spare.semanticParsers.DefaultSemanticParser
  Binder: nil.parser.spare.binders.DefaultBinder




*                                                  *
************* METRIC SCANNER ***********************
```

10:39
14/11/2015

File   Edit   Source   Refactor   Navigate   Search   Project   Run   Window   Help

Quick Access

Java

Problems   @ Javadoc   Declaration   Search   Console ⊠

<terminated> MexicanWasp9 [Java Application] C:\Program Files\Java\jre1.8.0_40\bin\javaw.exe (14/11/2015 10:39:42)

```
************* VERSE SCANNER ***********************
*                                                *




**************************************************
*                                                *
*                                                *
*   SET UP OF RESOURCES FINISHED                 *
*                                                *
*                                                *
*                                                *
**************************************************
Setting configuration to:
cullDownSize=500
numberOfSentences=1
initialPopulation=10000
numberOfGenerations=1
choiceOfCrossOver=5

iterate=false
infuse=false
compose=true
cross_over=false
revise=false
cull=true


babblers=[]

composers=[]

judges=

revisers=[]

mixers=[]
```

File   Edit   Source   Refactor   Navigate   Search   Project   Run   Window   Help

Quick Access

Problems   @ Javadoc   Declaration   Search   Console ✕

<terminated> MexicanWasp9 [Java Application] C:\Program Files\Java\jre1.8.0_40\bin\javaw.exe (14/11/2015 10:41:23)

```
Using n-gram model at address: resources/ngramModels/wordSequences/newTrainedNGramModelElSolClean.txt
Reading n-gram model file...
Read n-gram model file.
Building lexicon.
Using MoreTalk resources at address: resources/
Building lexicon.
Using MoreTalk resources at address: resources/lexicon/tentativeLexicon/
DesiredPopulationSize = 10000
Folder for dump initialised to ../../results/ReWasp/MexicanWasp/2015-11-14-10-41/1/
Directory Created
***************************************************
*                                                 *
*        Building initial population              *
*                                                 *
***************************************************


[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: jarocho

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: ratificaron

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: industrialización

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: ti´tulo

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: castañeda

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:
```

File   Edit   Source   Refactor   Navigate   Search   Project   Run   Window   Help

Quick Access   Java

Problems   @ Javadoc   Declaration   Search   Console

GenerateAndTestWaspSemanaCiencia [Java Application] C:\Program Files\Java\jre1.8.0_20\bin\javaw.exe (14/11/2015 9:16:43)

```
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge


[GenerateAndTestWaspSemanaCiencia] Found 0 candidates.


[GenerateAndTestWaspSemanaCiencia] Found 3 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 1 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 1 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 0 candidates.


[GenerateAndTestWaspSemanaCiencia] Found 0 candidates.


[GenerateAndTestWaspSemanaCiencia] Found 0 candidates.


[GenerateAndTestWaspSemanaCiencia] Found 2 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 4 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
```

9:39
14/11/2015

File   Edit   Source   Refactor   Navigate   Search   Project   Run   Window   Help

Quick Access

Problems   @ Javadoc   Declaration   Search   Console

<terminated> GenerateAndTestWaspSemanaCiencia [Java Application] C:\Program Files\Java\jre1.8.0_20\bin\javaw.exe (14/11/2015 9:16:43)

```
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 2 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 0 candidates.


[GenerateAndTestWaspSemanaCiencia] Found 0 candidates.


[GenerateAndTestWaspSemanaCiencia] Found 3 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 2 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 4 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 1 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge

[GenerateAndTestWaspSemanaCiencia] Found 2 candidates.

[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
[GenerateAndTestWaspSemanaCiencia] Draft ruled out by judge: RadicalAbsoluteVerseLengthJudge
No acceptable draft found after 1000 attempts.
Number of acceptable drafts found: 0
```

10:58
14/11/2015

File   Edit   Source   Refactor   Navigate   Search   Project   Run   Window   Help

Quick Access

Problems   @ Javadoc   Declaration   Search   Console ✕

<terminated> MexicanWasp9 [Java Application] C:\Program Files\Java\jre1.8.0_40\bin\javaw.exe (14/11/2015 10:42:15)

```
[SentenceEndingAcceptabilityEvaluator] WARNING: Word: mobilitas

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: nacieron

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: ayotzinapa

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word:  en

[SentenceEndingAcceptabilityEvaluator] WARNING: No lexical category found for:

[SentenceEndingAcceptabilityEvaluator] WARNING: Word: subsecretario
CurrentPopulationSize = 10000
*************** Scoring...
*************** Population scored.

Population size: 10000

Score of fittest individual: 49
******************************************************
*                                                    *
*       Composing drafts from initial population     *
*                                                    *
******************************************************

*************** Scoring...
*************** Population scored.
*************** Scoring...
*************** Population scored.
******************************************************
*       Culling initial population                   *
******************************************************
```

Nombre	Fecha de modifica...	Tipo	Tamaño
draft1-19976-19672.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19977-19690.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19978-19718.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19979-19726.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19980-19737.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19981-19750.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19982-19774.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19983-19775.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19984-19786.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19985-19814.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19986-19834.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19987-19841.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19988-19845.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19989-19849.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19990-19892.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19991-19896.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19992-19900.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19993-19905.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19994-19907.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19995-19913.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19996-19919.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19997-19947.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19998-19960.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-19999-19969.txt	14/11/2015 4:03	Documento de tex...	1 KB
draft1-20000-19998.txt	14/11/2015 4:03	Documento de tex...	1 KB
fullPopulation4run1.txt	14/11/2015 4:03	Documento de tex...	14.556 KB
report.txt	14/11/2015 4:03	Documento de tex...	1 KB
trace_number1.txt	14/11/2015 4:03	Documento de tex...	0 KB

```
1  Average score: 16
2  Percentage of drafts above 50: 3
3  Percentage of drafts above 75: 0
4  Percentage of drafts above 90: 0
```

LiNHD
LABORATORIO DE iNNOVACiÓN EN HUMANiDADES DiGiTALES

Contents

Early attempts

    Following patterns: the CBR approach

    Learning to revise: the evolutionary approach

The middle years

    Systematic exploration: letting processes run

    Reusing in other ways: building poems from verse

Current efforts

    Seeking help: let a <u>master</u> show you how

Consider resulting snippets of verse as building blocks for poems

Devise procedure for composing them into larger poems

Table 5. Example set of poems generated by the system (excluding those used for the original commission and those rejected according to the criteria presented in Section 4), with an approximate English translation.

Poem 1: 4 lines of 7 syllables with rhyme scheme BACA

Toda era una ave larga	Everything was a long bird
que cuando se conforman.	that when they conform.
Admitió que se tienen	He admitted that they have
registradas personas.	registered persons.

Poem 2: 4 lines of 7 syllables with rhyme scheme ABCA

Muestra también. Esta	Shows as well. This
noche adonde yo soy.	night were I am.
Subraya que para ellos	Underlines that for them
ya no salgas. Estrellas.	come out no more. Stars.

Poem 3: 4 lines of 8 syllables with rhyme scheme ABAC

Aspecto que se encontraban	Aspect that they were finding
ejemplares. Nuevamente.	exemplars. Again.
Señalaron que no haya	They pointed out that there should not be
más daños como los niños.	more harm like the children.

Poem 4: 4 lines of 9 syllables with rhyme scheme ABAC

Zona militar. Qué delicia	Military zone. What a delight
delgada incomprensible. Amiga.	thin incomprehensible. Girl friend.
Agueda era luto pupilas	Agueda was mourning green
verdes. Sobrepasa. Guerrero.	pupils. Overshoots. Warrior.

Poem 5: 6 lines of 9 syllables with rhyme scheme ABABAB.

Juegan el largo recorrido	They play the long tour
desde su muerte ya no salgas.	from his death come out no more.
Séptimo. Cordero tranquilo	Seventh. Peaceful lamb
cordero que paces tu grama.	lamb that grazes its grass.
Silencios. Cordero tranquilo	Silences. Peaceful lamb
cordero que paces tu grama.	lamb that grazes its grass.

Poem 7: 15 lines of 9 syllables with rhyme scheme CAABADAEAFGBAA.

Tus ojos. Vinos tintos	Your eyes. Red wines
blancos rosados. Nardo.	whites rosés. Tube rose
Amo tus ríos claros.	I love your clear rivers.
Tal vez esta medalla.	Maybe this medal.
Antes que este hachazo	Before this axblow
nos sacude. Imagínate.	shakes us. Imagine.
Séptimo. Pinté el tallo	Seventh. I painted the stem
luego el cáliz después.	then the calyx afterwards.
Ganar. Solicitamos.	Winning. We request.
Sólo soy un prisionero.	I am only a prisoner.
Admitió que se estima	He admitted that it is estimated
que mil michoacanas	that a thousand Michoacans
acudan. Queretanos.	turn up. Queretans.
Valor. Acompañado	Valour. Accompanied
por Margarita Flores.	by Margarita Flores.

Contents

Early attempts

    Following patterns: the CBR approach

    Learning to revise: the evolutionary approach

The middle years

    Systematic exploration: letting processes run

    Reusing in other ways: building poems from verse

Current efforts

    Seeking help: let a master show you how

syl Mention Rhyme / Line Go

□ 1

□ 2

□ 3

□ 4

□ 5

□ 6

□ 7

□ 8

	# syl	Mention	Rhyme	/	Line	Go
☑ 1	11			☑	Aquellas riendas me enfermaron. Fui.	
☑ 2	11			☐	A mi tarea se confabularon	
☑ 3	11			☑	para helarme el corazón lo mismo.	
☐ 4				☐		
☐ 5				☐		
☐ 6				☐		
☐ 7				☐		
☐ 8				☐		

	# syl	Mention	Rhyme	/	Line	Go
☑ 1	11			☑	Avanzada la noche aquella.	
☑ 2	11			☐	En tanto seguía a mí tener	
☑ 3	11			☑	tanta vista. Rocas y arbustos. Lanza.	
☑ 4	11		oda	☑	No line found after 20 tries!	
☐ 5				☑		
☐ 6				☐		
☐ 7				☐		
☐ 8				☐		

Acknowledgements

Thank you!

http://nil.fdi.ucm.es