Story Generation: Once Upon A Time...

Irina Borisova

Computational Approaches to Creative Language
SS 2010, July 6

For example, the heavy grammar felled the frontier. An individual and overlapping ambiguity developed. Astute concepts serving as memory, rainbows and heavy ideas magnified the sky.

(RUBBER BLUE BIODEGRADABLE ROBOT
by R. S. Pearson, 1988)

Outline

- Why story generation?

- History

- Case-Based Reasoning Story Generation (Gervás et al., 2004)

- Cast Story Generation (León and Gervás, 2008)

- Evaluation of automatically generated texts

- Dramatica Demo

Why story generation?

- get rich! :)
- Philip M. Parker
- 200.000 books 'written', such as:
    - "The Official Patient's Sourcebook on Acne Rosacea" ($24.95 and 168 pages long);
    - "Stickler Syndrome: A Bibliography and Dictionary for Physicians, Patients and Genome Researchers" ($28.95 for 126 pages);
    - "The 2007-2012 Outlook for Tufted Washable Scatter Rugs, Bathmats and Sets That Measure 6-Feet by 9-Feet or Smaller in India" ($495 for 144 pages).
- "My goal isn't to have the computer write sentences, but to do the repetitive tasks that are too costly to do otherwise."

Why story generation?

- entertainment with a low quality demand and a human editor

  - TV series

  - video games

- independent writer assistance (Dramatica, Writer Pro)

- education: narrative skills development

- studying (creative) writing

What does it mean to tell a story:

- For a human:
  - coherent
  - complete
  - involves some action
- judgements on creativity are very diverse
- For "computerized creativity" – story-predictability: the lower story-predictability, the higher creativity rate
- Knowledge bases for story generation include:
  - rhetoric knowledge
  - story-world knowldge   all evaluated
  - common-sense knowldge
- Interestingness and story output

History

- TALE-SPIN (1981)

  - story-grammar

  - a goal for a character

- MINSTREL (1994):

  - case-based schemes

  - author-level planning and processing

  - explicit creative block TRAM

- MEXICA (1999): engagement and reflection phases

- BRUTUS (2000):

  - thematic frame (story frames)

  - plot development

  - lexicalization (final output)

TRAM in MINSTREL

- TRAM:Standard-Problem-Solving

- TRAM:Similar-Outcomes-Partial-Change

- TRAM:Generalize-Constraint

- Task: generate a scene where *a knight kills himself*

- Episodes in the memory:

  - *Knight fights and kills a troll*

  - *Princess makes herself intentionally ill by drinking a poison.*

- to kill himself = to be injured

- knight = anyone

- princess = anyone

- poison to make someone ill = poison to kill someone

- Outcome: *The knight kills himself by drinking poison*

BRUTUS story-frame

story-frame

 name: evilGoal

 type: character goal

 agent: Hart

 plan: {lie_to_Striver, refuse_to_sign_his_thesis}

Story plot generation based on CBR
(Gervás et al., 2004)

- CBR: Case-Based Reasoning

- solving new tasks on the basis of previous ones (using analogy or prototypes)

- Cases – character functions (V. Propp) developed for the magic fairy tale analysis)

- two modules: CBR and NLG (Natural Language Generation)

- conceptual description $\longrightarrow$ plot plan $\longrightarrow$ text

Case-Based Reasoning

- CBR: Case-Based Reasoning

- solving new tasks on the basis of previous ones (using analogy or prototypes)

- Cases – character functions (V. Propp) developed for the magic fairy tale analysis)

- two modules: CBR and NLG (Natural Language Generation)

- conceptual description $\longrightarrow$ plot plan $\longrightarrow$ text

  ontology   query interface   NLG

Vladimir Propp

- Structuralist, anthropologist, folklore specialist

- Morphology of a Russian magic fairy tale, 1928: certain units of a plot are constant, and only some change

- a fairy tale is composed of a number of **functions**

  - *interdiction*

  - *interdiction violated*

  - *test of a hero*

- and **spheres of action** (*hero, villain, helper*)

CBR: System Knowledge

Protege Ontology with key components (below) and semantic measurements between them

- **Propp functions**
- Moves
- Character
- Properties of the characters (attributes)
- Roles
- Places and objects
- Descriptions
- Cases

Case Structure

A Fairy Tale in **Functions** and *Moves*

The Swan Geese (113 of Afanasiev Collection).

Initial situation (a girl and her small brother).

Interdiction (not to go outside),

Interdiction violated,

Kidnapping (swan geese take the boy to Babayaga's lair), **Competition** (girl faces Babayaga),

Victory,

Release from captivity,

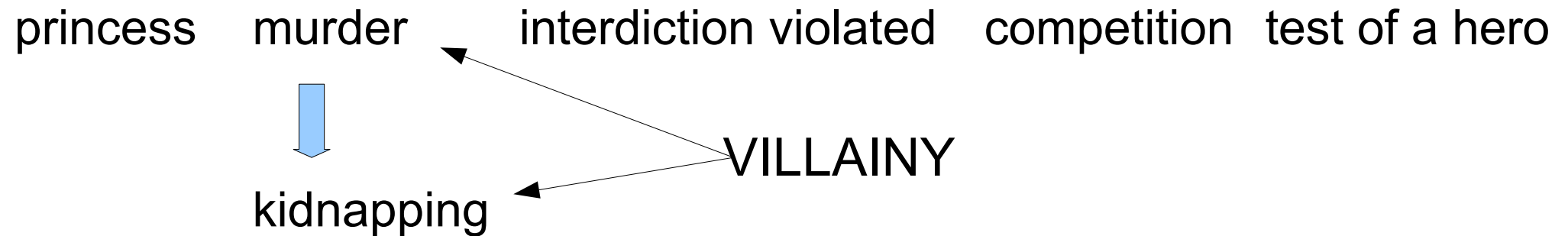**Test of hero** (swan geese pursue the children),

Sustained ordeal (children evade swan geese),
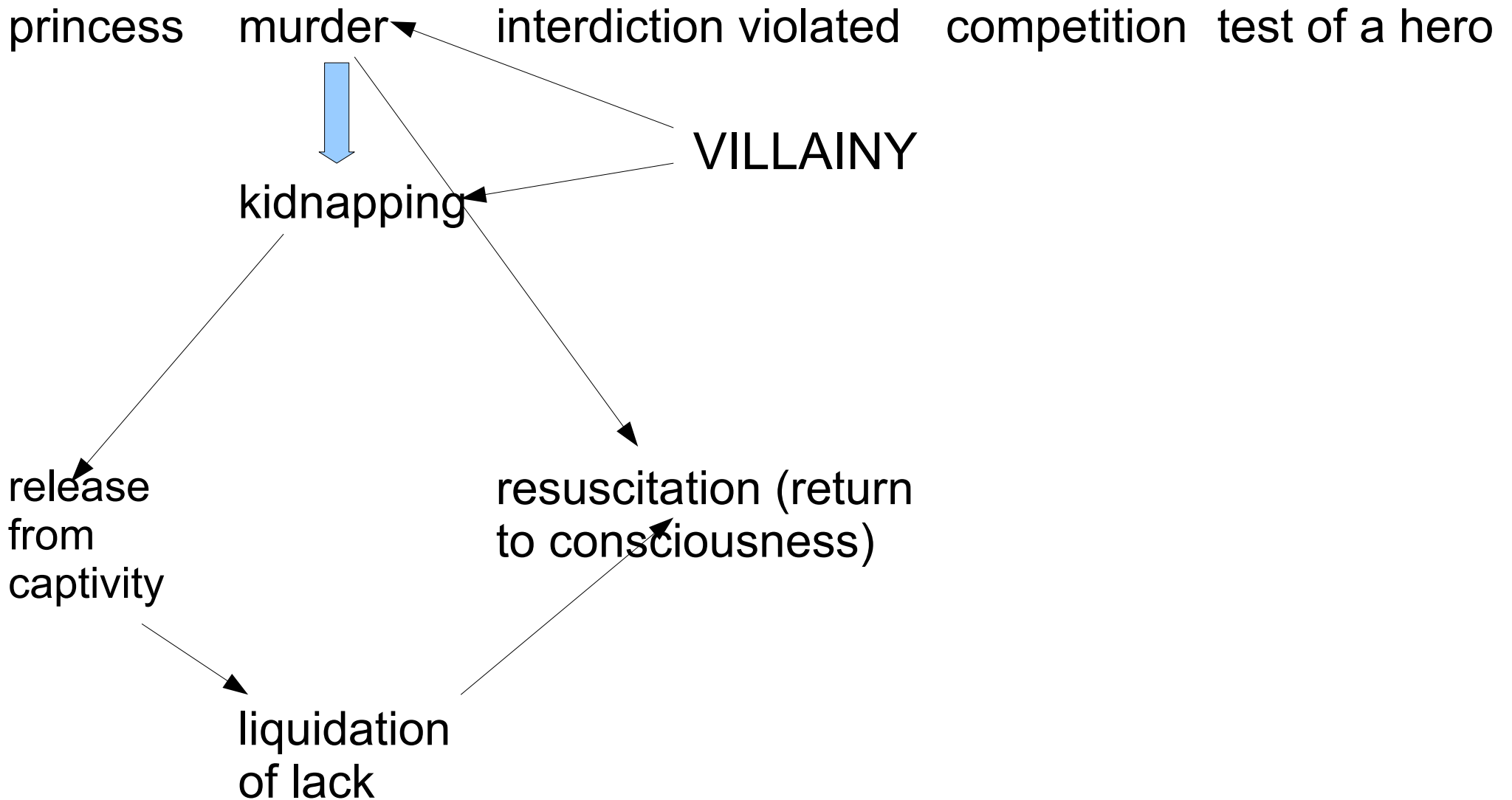
Return.

CBR Module

- a detailed query system
- substitution based on the similarity measurements

Query:

princess   murder   interdiction violated   competition   test of a hero

kidnapping   VILLAINY

CBR Module

- More Adaptation

princess   murder   interdiction violated   competition   test of a hero

VILLAINY

kidnapping

release
from
captivity

resuscitation (return
to consciousness)

liquidation
of lack

NLG module

- template-based

- accepts plot plans:
    - case
    - character functions indexed with the ontology elements that they compose

- each function is processed at once with all the elements of the ontology that it refers to, e.g.:
    - character
    - attribute
    - location

NLG stage 1

- Content Determination: what is in the character function, what is new, adn what comes from the discourse history

character(ch1,princess)

character(ch3,lioness)

location(l1,forest)

role(ch3,villain)

attribute(ch3,hungry)

attribute(ch3,fierce)

action(ch3,ch1,devour)

NLG stage 2

- Discourse Planning: heuristics for priority sequences in the ontology

character(ch3,lioness),

attribute(ch3,hungry)

character(ch3,lioness),

attribute(ch3,fierce)

character(ch3,lioness),

character(ch1,princess),

action(ch3,ch1,devour)

NLG stage 3

- Sentence Aggregation: regrouping on the base of the similair constructions

character(ch3,lioness),

attribute(ch3,hungry),

attribute(ch3,fierce)

character(ch3,lioness),

character(ch1,princess),

action(ch3,ch1,devour)

NLG stage 4

- RE Generation:

  noun-to-pronoun

  indefinite-to-definite articles

character(ch3,lioness),

ref(ch3,def),

attribute(ch3,hungry),

attribute(ch3,fierce)

character(ch3,pron),

character(ch1,princess),

ref(ch1,def),

action(ch3,ch1,devour)

NLG stage 5

- Lexicalization: template and concept-based (static objects, verbs, attributes)

"lioness"

"a"

"hungry"

"fierce"

L(x)+" was "+L(y)+" and "+L(z)

"she"

"princess"

"the"

L(x)+" devoured "+L(y)

NLG stage 6

- Surface Realization: realizing a template, orthography requirements, and simple graphic rules

  "The lioness was hungry and fierce. She devoured the princess"

Cast: Stotytelling Based on Exploration and Transformation (León and Gervás, 2008)

Recap

- Creativity

    - Exploratory

    - Transformational

    - P-creativity

- Conceptual space

Recap

- Creativity

  - Exploratory
  - Transformational
  - P-creativity – psychological: individual mind
  - H-creativity

- Conceptual space

Recap

- Creativity
    - Exploratory
    - Transformational
    - P-creativity - psychological
    - H-creativity – historical: whole human history
- Conceptual space

Cast: Stotytelling Based on Exploration and Transformation (León and Gervás, 2008)

- a set of logic predicates, variables, and rules

chase(policeman, criminal), want(criminal, money)

- constraints – in schema

friend(?x, ?y) $\land$ ¬ chase(?x, ?y)

Cast: Stotytelling Based on Exploration and Transformation (León and Gervás, 2008)

- a set of logic predicates, variables, and rules

chase(policeman, criminal), want(criminal, money)

- constraints – in schema

friend(?x, ?y) $\wedge \neg$ chase(?x, ?y)

- state space search

- knowledge base <K, C>

- K – set of logic facts

- C – set of schema

- I – set of user restrictions

- s – initial story (empty set of facts)

Exploration in Cast

- a function for creating new states $\phi(s, K)$

$$s_{n+1} = s_n \cup \phi(s_n, K)$$

- a coherence function (s, C) – validation of the partial story against the domain constraints

- a validity function (s, I) –validation against the user requirements

Algorithm 1 Exploratory algorithm for Story Generation

1. $s_0 = \{\}$
2. while $\omega(s_i, I) = \perp$ /* *while the partial story is not coherent with the user* */
   - (a) repeat $s_{n+1} = s_n \cup \varphi(s_n, K)$ while $\lambda(s_{n+1}, C) = \perp$ /* *look for a new fact* */
   - (b) if it has been possible, *continue*, expand the new state
   - (c) if there is no possible fact, return *failure*, discard the fact and continue exploring

Transformation in Cast

- results from the exploratory phase are coherent, but not very creative

- solution: modify coherence function – allow it to create partial non-coherent stories

- how to restrict the grade of incoherence:

    - percentage (10%)
    - meta-rules

CBR and Cast

- two different approaches suggested:
  - CBR: ontology and NLG modules
  - Cast: formal semantics
- theoretical level:
  - no implementation
  - no evaluation
- weak creativity concerns

Discussion

" Time: Too far in the future to understand

Setting: Earth, which has remained technological yet developed biodegradable and biosphere friendly technology. "

(RUBBER BLUE BIODEGRADABLE ROBOT by R. S. Pearson)

- How to deviate from a template in a creative way

- Evaluation:

  - 'limits' of a style, genre, creativity, novelty, and a human evaluator

- Highly-dependent on the progress and the use in other CL domains

- Higher demand in more granular studies

Find out more

- Dramatica Pro Demo:

  http://www.screenplay.com/t-dpdemo.aspx

- The DADA Engine:

  http://dev.null.org/dadaengine/

- Propp' functions:

  http://en.wikipedia.org/wiki/Vladimir_Propp

- A Blog on Story Generation (by 2005)

  http://storymachine.blogspot.com/

- Philip M. Parker describing his work:

  http://www.youtube.com/watch?v=SkS5PkHQphY

References

- León C, Gervás P.: Creative Storytelling Based on Transformation of Generation Rules, 5th Internation Joint Workshop on Computational Creativity, 2008.

- Pablo Gervás, Belén Díaz-Agudo, Federico Peinado and Raquel Hervás. 2004. Story plot generation based on CBR. *Knowledge-Based Systems*, Volume 18, Issues 4-5, August 2005, Pages 235-242

- Piroska Lendvai, Thierry Declerck, Sándor Darányi, Scott Malec. 2010. Propp Revisited: Integration of Linguistic Markup into Structured Content Descriptors of Tales, *Digital Humanities*, Vol. 7.

- Rafael Pérez y Pérez, Mike Sharples Three Computer-Based Models of Storytelling: BRUTUS, MINSTREL and MEXICA