A Repository of Frame Instances for Generation
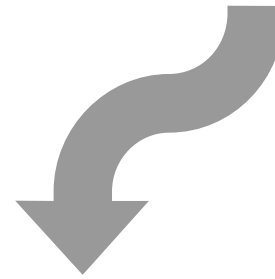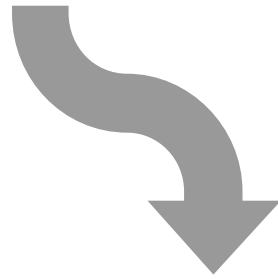
Valerio Basile

WebNLG 2016, Edinburgh
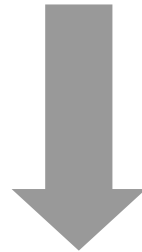
Something Old, Something New

Information Extraction   NLG

Large-scale resource

Web NLG

Background Story #1

Knowledge Extraction with Semantics

Basile, Cabrio and Schon (2016)
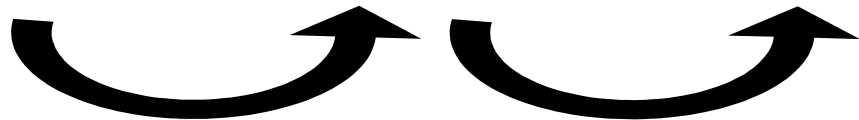
Default Knowledge Base
- about objects
- for domestic robots

AUTONOMOUS
LEARNING
OF THE MEANING
OF OBJECTS

ALOOF

X
Default Knowledge Base

✓
Default Knowledge Base

KnEWS
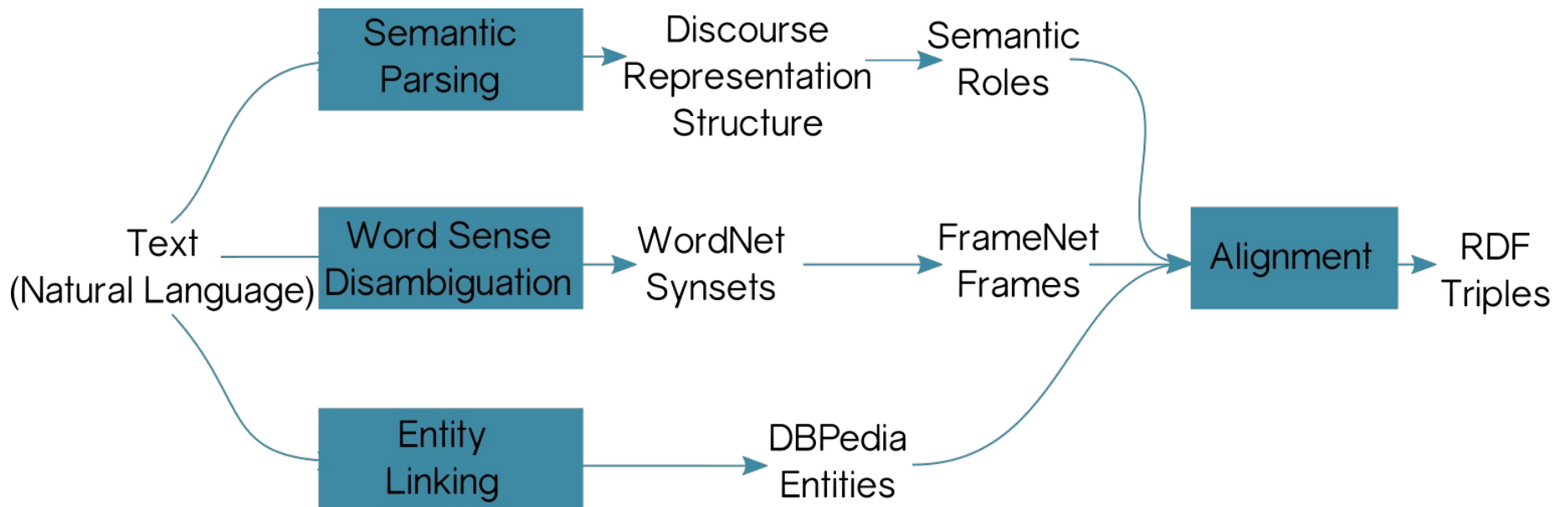
Integrates:
C&C tools, Boxer, UKB, Babelfy, Spotlight

KnEWS

DRIVE(e)
ROBOT(x)
AGENT(x, e)

The robot drives a car
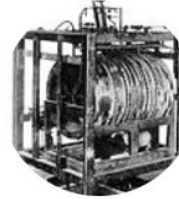
robot
A mechanism that can move automatically

drives
Operate or control a vehicle

car
A wheeled vehicle adapted to the rails of railroad

DRIVE
http://wordnet-rdf.princeton.edu/wn31/201934845-v
ROBOT
http://dbpedia.org/page/Robot

Frame Semantics

Frame Operate_vehicle
DRS event

Frame element Driver
DRS thematic role "agent"

Frame element Vehicle
DRS thematic role "theme"

Non-core Frame Elements
Route, Purpose, Speed, ...

KnEWS

```
fb: <http://framebase.org/ns/>
rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns>
wn: <http://wordnet-rdf.princeton.edu/wn31/>

fb:fi-Operate_vehicle_e7f56d25 rdf:type fb:frame-Operate_vehicle-drive.v
fb:fi-Operate_vehicle_e7f56d25 fb:fe-Driver wn:02764397-n
fb:fi-Operate_vehicle_e7f56d25 fb:fe-Vehicle wn:02961779-n
```

Demo at
   http://gingerbeard.alwaysdata.net/knews/

Frame Semantics

Schema by Framebase
www.framebase.org
(Rouces et al., 2015)

Alignments by
Framebase (Wordnet-FrameNet)
Semlinks (Verbnet-FrameNet)
Babelfy (Wordnet-DBpedia)

Background Story #2

Natural Language Generation
from Logical Forms

Basile (2015, PhD thesis)

NLG as machine translation

Word-meaning alignment

k_1: ¬ k_2:

	x_1 e_1
	CUSTOMER(x_1)
	PAY(e_1)
	AGENT(e_1,x_1)
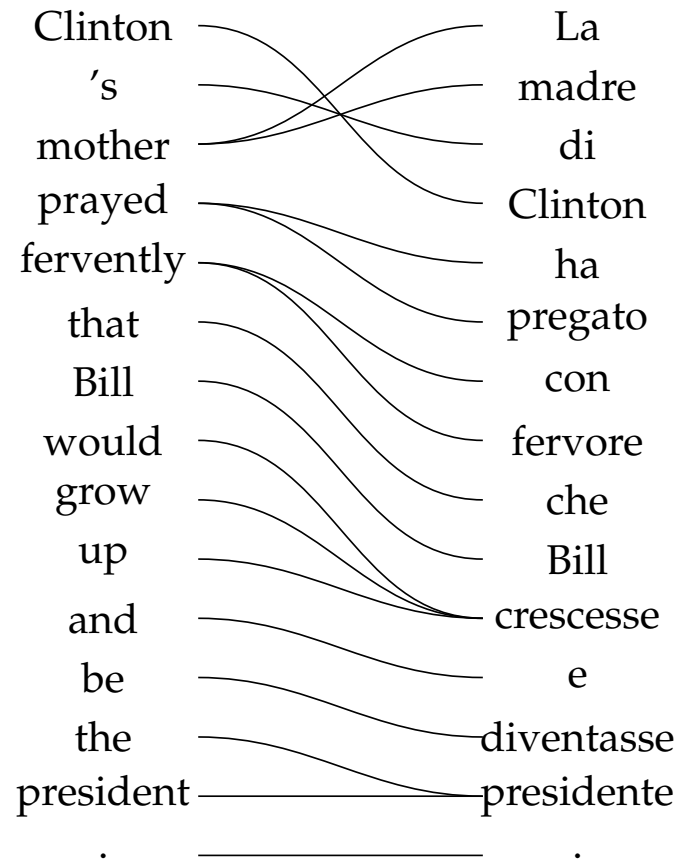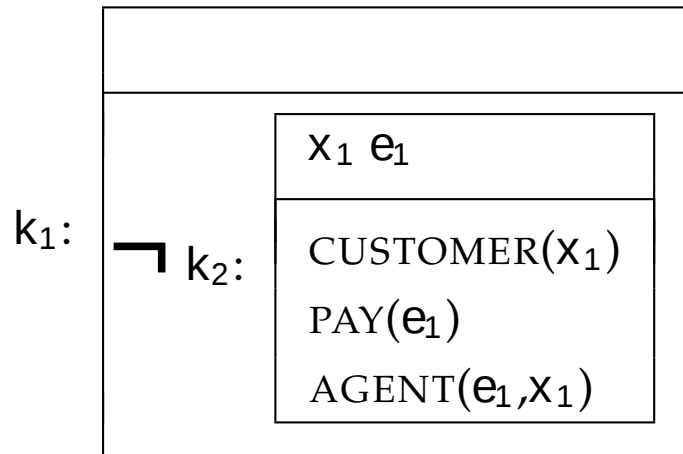
k_1	unary			
	scope	k_2		
k_2	referent	e_1		
k_2	referent	x_1	1	A
k_2	event	PAY		
k_2	concept	CUSTOMER		
k_2	role	agent		
CUSTOMER	instance	x_1	2	customer
PAY	instance	e_1	4	pay
agent	internal	e_1	1	
agent	external	x_1		
k_2	**surface**	e_1	2	**did**
k_2	**surface**	e_1	3	**not**
k_2	**surface**	e_1	5	**.**

Word-meaning alignment

$$
\frac{\begin{array}{c} \dfrac{x_1 : A\ customer \qquad e_1 : x_1\ did\ not\ pay}{e_1 : A\ customer\ did\ not\ pay\ .} \end{array}}{k_2 : A\ customer\ did\ not\ pay\ .}
$$

$k_2 : e_1$

Complete and incomplete surface forms

Combining the ideas

Word-aligned Frames and Frame Elements

Word-aligned frames and FEs

$$\cfrac{k_2 : e_1 \qquad \cfrac{x_1 : \text{A customer} \qquad e_1 : x_1 \text{ did not pay}}{e_1 : \text{A customer did not pay .}}}{k_2 : \text{A customer did not pay .}}$$

`wn:Customer` : "A customer"
`fn:Paying` : "*x* did not pay"

(In a particular instance)

```xml
<frameinstance id="Operate_vehicle_ce746f21-2d8d-4fe8-8981-
df95c9b0eb07" type="Operate_vehicle-drive.v" internalvariable="e1">
  <framelexicalization>k3:x1 is driving k3:x2</framelexicalization>
  <instancelexicalization>
    The robot is driving the car .
  </instancelexicalization>
  <frameelements>
    <frameelement role="Driver" internalvariable="x1">
      <concept>
        http://wordnet-rdf.princeton.edu/wn31/02764397-n
      </concept>
      <rolelexicalization>
        The robot is driving x2
      </rolelexicalization>
      <conceptlexicalization>The robot</conceptlexicalization>
    </frameelement>
    <frameelement role="Vehicle" internalvariable="x2">
      <concept>
        http://wordnet-rdf.princeton.edu/wn31/02961779-n
      </concept>
      <rolelexicalization>
        x1 is driving the car .
      </rolelexicalization>
      <conceptlexicalization>the car .</conceptlexicalization>
    </frameelement>
  </frameelements>
</frameinstance>
```

```xml
<frameinstance id="Operate_vehicle_ce746f21-2d8d-4fe8-8981-
df95c9b0eb07" type="Operate_vehicle-drive.v" internalvariable="e1">
    <framelexicalization>k3:x1 is driving k3:x2</framelexicalization>
    <instancelexicalization>
      The robot is driving the car .
    </instancelexicalization>
    <frameelements>
      <frameelement role="Driver" internalvariable="x1">
        <concept>
          http://wordnet-rdf.princeton.edu/wn31/02764397-n
        </concept>
        <rolelexicalization>
          The robot is driving x2
        </rolelexicalization>
        <conceptlexicalization>The robot</conceptlexicalization>
      </frameelement>
      <frameelement role="Vehicle" internalvariable="x2">
        <concept>
          http://wordnet-rdf.princeton.edu/wn31/02961779-n
        </concept>
        <rolelexicalization>
          x1 is driving the car .
        </rolelexicalization>
        <conceptlexicalization>the car .</conceptlexicalization>
      </frameelement>
    </frameelements>
  </frameinstance>
```

```xml
<frameinstance id="Operate_vehicle_ce746f21-2d8d-4fe8-8981-
df95c9b0eb07" type="Operate_vehicle-drive.v" internalvariable="e1">
   <framelexicalization>k3:x1 is driving k3:x2</framelexicalization>
   <instancelexicalization>
     The robot is driving the car .
   </instancelexicalization>
   <frameelements>
     <frameelement role="Driver" internalvariable="x1">
       <concept>
          http://wordnet-rdf.princeton.edu/wn31/02764397-n
       </concept>
       <rolelexicalization>
          The robot is driving x2
       </rolelexicalization>
       <conceptlexicalization>The robot</conceptlexicalization>
     </frameelement>
     <frameelement role="Vehicle" internalvariable="x2">
       <concept>
          http://wordnet-rdf.princeton.edu/wn31/02961779-n
       </concept>
       <rolelexicalization>
          x1 is driving the car .
       </rolelexicalization>
       <conceptlexicalization>the car .</conceptlexicalization>
     </frameelement>
   </frameelements>
 </frameinstance>
```

```xml
<frameinstance id="Operate_vehicle_ce746f21-2d8d-4fe8-8981-
df95c9b0eb07" type="Operate_vehicle-drive.v" internalvariable="e1">
    <framelexicalization>k3:x1 is driving k3:x2</framelexicalization>
    <instancelexicalization>
      The robot is driving the car .
    </instancelexicalization>
    <frameelements>
      <frameelement role="Driver" internalvariable="x1">
        <concept>
          http://wordnet-rdf.princeton.edu/wn31/02764397-n
        </concept>
        <rolelexicalization>
            The robot is driving x2
        </rolelexicalization>
        <conceptlexicalization>The robot</conceptlexicalization>
      </frameelement>
      <frameelement role="Vehicle" internalvariable="x2">
        <concept>
          http://wordnet-rdf.princeton.edu/wn31/02961779-n
        </concept>
        <rolelexicalization>
          x1 is driving the car .
        </rolelexicalization>
        <conceptlexicalization>the car .</conceptlexicalization>
      </frameelement>
    </frameelements>
  </frameinstance>
```

A Repository of Frame Instances for Generation

Data

ESL Yes: http://www.eslyes.com/

725 English short stories for English learners.
14,140 Sentences.

30,217 frame instances (420 unique frames).
1,455 concepts (1,201 WordNet synsets and 254 DBpedia entities) filling in 41,945 roles (161 unique roles).

29,409 role instances could not be mapped to FrameNet.

Evaluation

Experiment #1: re-generation

 7,366 instances correctly regenerated
11,996 incorrect instance lexicalizations (usually containing variables)
 5,211 subordination
 1,865 phrasal verbs or adverbs
 3,779 other of reasons (failure of the entity linking module, wrong syntactic analysis, ...)

Evaluation

Experiment #2: generation by composition

1. For each frame instance, produce four new frame instances by replacing one or both frame elements, either with similar concepts or with randomly chosen concepts.

2. Generate the lexicalization of the new frames by composing the frame lexicalization structure with the new concept lexicalizations.

Evaluation

Experiment #2: generation by composition

Replaced frame elements	Judgment nonsensical/informative/fluent
1, most similar	23/33/44
2, most similar	24/53/23
1, random	23/35/42
2, random	54/23/23

Issues

Alignment somewhat arbitrary

Long pipeline, many bumps on the road

Lexical choice vs. morphology generation

Future

Enlarging the resource horizontally
(with what text?)

Enlarging the resource vertically
(with what layers?)

Statistical generation

Towards a gold standard

http://dbpedia.org/resource/Gratitude

↕

Thank you!