

*International Technology Alliance
in
Network & Information Sciences*

Constructing Controlled English for Both Human Usage and Machine Processing

*7th RuleML
Human Language Technology Special Track*

July 12, 2013

Stephen Poteet, Ping Xue, Anne Kao
Boeing Research & Technology

David Mott, Dave Braines
IBM UK



MINISTRY OF DEFENCE



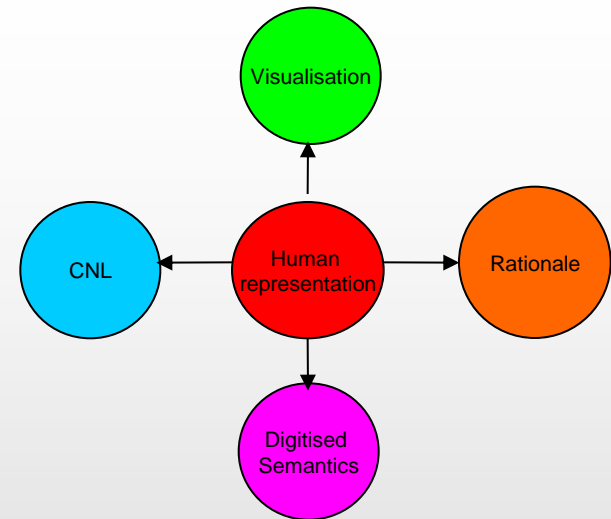
Fundamental Research Issues

- How do we assist people to create and use applications that **reason**?
 - **Modelling concepts**, relationships and rules of inference
 - **Grasping the basic logic** of the model and rules
 - **Understanding** the reasoning performed by **others**
 - **Sharing understanding** across the human team
 - **Sharing reasoning and results** across different systems



5 Steps Towards a Solution

- **Human Representation**
 - “conceptual models”
- **Semantic Representation**
 - formal definition of concepts and logical relations
 - machine processing
- **Controlled Natural Language**
 - communication and shared understanding
- **Rationale**
 - explanation and challenging of reasoning
- **Visualisation**
 - enhancing human understanding



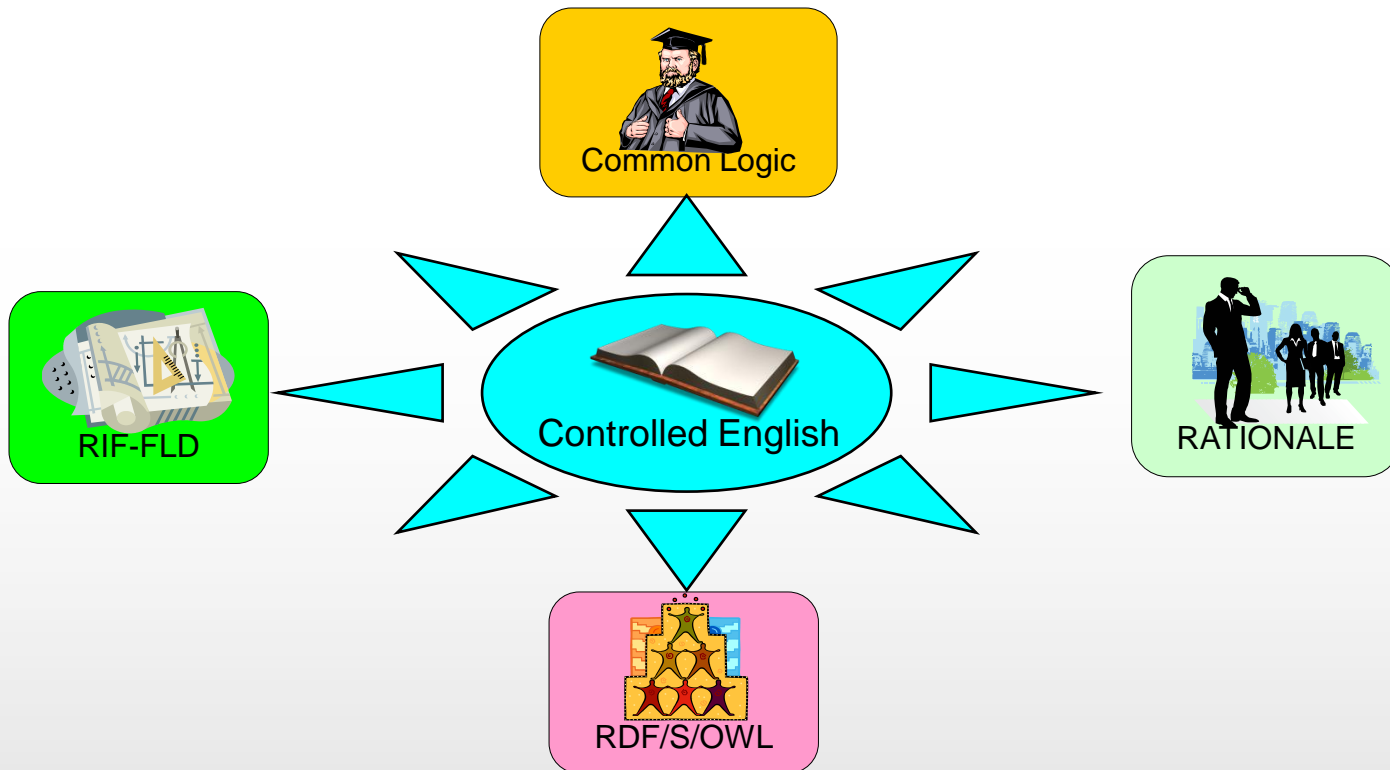


ITA Controlled English (CE)

- CE is a type of **controlled natural language**
- A controlled natural language is a subset of a natural language using a **restricted set of grammar rules** and a **restricted vocabulary**
- Traditionally, focus was either on improvement for ***human readability*** or for ***machine readability***
- We are concerned with both easy readability for human consumption and unambiguous representation for computer processing
- **Challenge:** how to balance user-friendliness and computational predictability



CE is Machine Processable



- ITA CE is semantically consistent with and partially mappable to:
 - First Order Predicate Logic
 - CE is based on Common Logic Controlled English (Sowa 2007), with extensions
 - Existing ontology modeling languages such as OWL



Statements and Queries in CE

- CE permits a set of plain English sentences for stating propositions referring to entity existence, properties and relations:
 - there is a person named Fred.
 - the person Fred has French as language.
 - the person Fred is married to the person Jane.
- CE also permits meta statements that specify information about propositions such as their truth status or whether they are assumptions:
 - it is true that** there is a person named John.
 - it is assumed that** the person Fred is married to the person Jane.
 - it is true to degree CV** that Fred is a father.
- CE also allows queries:
 - for which** P1 and P2 is it true that the person P1 is the parent of the person P2.



Building a Model in CE

- The user writes a conceptual model to fit their domain
- The model and the language are built up simultaneously
- The model is built from the more abstract levels to the more specific
- This is all done with the CE “**conceptualise**” statement
 - Adding an entity into the ISA hierarchy
conceptualise a ~ person ~ P that is an agent.
 - Adding an entity with certain attributes
conceptualise the person P
that has the value H as ~ height ~ and
has the value W as ~ weight ~
has the color C as ~ hair color ~.
 - Adding a relation
conceptualise the person P
~ is married to ~ the person P2.



Inference and Rationale in CE

- CE also allows statements of logical rules to allow inferencing:
if PREMISES then CONCLUSION

if (the person X has the person Y as brother) and
 (the person Z has the person X as father)
then
 (the person Z has the person Y as uncle)

- ... and statements that can be used to reconstruct the rationale for an assertion or conclusion:

CONCLUSION because PREMISES

the task T1 has the agent A1 as executor
because

the plan P1 has the agent A1 as executor and
the plan P1 contains the task T1.



CE Based Capabilities and Applications

- CE is designed to be easily extensible
- CE encourages a richer interaction and integration between human and machine reasoning capabilities
- CE is most useful in situations that have the following characteristics:
 - A high degree of human interaction, usually involving specialist users with complex needs in non-trivial environments
 - A likelihood of rapidly evolving or uncertain tasks, queries or other knowledge-based activities
 - The need for collaboration, between different people or teams and/or across different disciplines



CE Store: A CE-Based Tool Suite

- There is a “CE Virtual Machine”
 - Inputs/outputs/reasons/explains CE
 - Applications can be written to use this CEVM
- There is a Java CESTore and a Prolog CESTore
 - Java CE Store: **<http://ibm.co/RDla53>**
- The “CE Store” allows one to:
 - Perform basic CE sentence parsing
 - Define and extend any domain concept model
 - Assert any CE sentence conforming to the appropriate conceptual model(s)
 - Define and execute a CE query against a domain model using a combination of a visual query language and written CE
 - Define logical rules
 - Execute the logical rules to infer new CE information and assert it to the knowledge base
 - Display rationale for conclusions
 - Determine whether assumptions were involved in the inference

Constructing a Query in CE Store

CE Store browser - Alpha v1.2 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

localhost:8080/CeStoreWeb/

Most Visited Latest Headlines Getting Started

CE Store browser - Alpha v1.2

Last transaction (executeCeQuery) took 0.016 seconds. Store contains 272 instances and 742 sentences. Code version=1.2.4587 [refresh the page](#) [login](#)

Concepts

Filters: [\[pn\]](#) [\[s\]](#) [\[nz\]](#) [\[nr\]](#)
[refresh](#)

Showing 10 primary concepts:

- [cancer](#) (1, 1)
- [cancer scan](#) (1, 1)
- [diagnostic test](#) (1) [\[s\]](#)
- [disease](#) (1) [\[s\]](#)
- [lump](#) (1, 1)
- [man](#) (3, 3)
- [patient](#) (1, 1)
- [person](#) (9, 9) [\[s\]](#)
- [symptom](#) (1) [\[s\]](#)
- [woman](#) (1, 1)

General information **CE Query Builder (CEQB)** **Lexical information** **Extracted info**

Entity

Concept details for [person](#) below:

| | |
|---|---|
| Name | person [r] |
| parent(s) [all] or [direct] | <ul style="list-style-type: none">primary thingthing |
| children [all] or [direct] | <ul style="list-style-type: none">manwomanpatient |
| creation date | 1369063016762 |
| instances | 9 (simple list) (list details) (exact list) |

Diagram:

the person P1 [9] is the parent of [5] the person P2 [9]

for which P1 and P2 is it true that
(the person P1 is the parent of the person P2)

Query/rule name:

- [Clear query canvas](#)
- [Validate query](#)
- [Execute as](#)

Entity

Concept details for [person](#) below:

| | |
|---|---|
| Name | person [r] |
| parent(s) [all] or [direct] | <ul style="list-style-type: none">primary thingthing |
| children [all] or [direct] | <ul style="list-style-type: none">manwomanpatient |
| creation date | 1369063016762 |
| instances | 9 (simple list) (list details) (exact list) |

Sentences

Sources

Saved queries & rules

Errors (0) Warnings (0) Debugs (0) Alerts (0)

No errors were returned in the last request.

Query Results

CE Store browser - Alpha v1.2 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

localhost:8080/CeStoreWeb/ Boeing Search

Most Visited Latest Headlines Getting Started

CE Store browser - Alpha v1.2

Last transaction (executeCeQuery) took 0.016 seconds. Store contains 272 instances and 742 sentences. Code version=1.2.4587 [refresh the page](#) [login](#)

Concepts

Filters: [\[pn\]](#) [\[s\]](#) [\[nz\]](#) [\[nr\]](#) [refresh](#)

Showing 10 primary concepts:

- [cancer](#) (1, 1)
- [cancer scan](#) (1, 1)
- [diagnostic test](#) (1) [\[s\]](#)
- [disease](#) (1) [\[s\]](#)
- [lump](#) (1, 1)
- [man](#) (3, 3)
- [patient](#) (1, 1)
- [person](#) (9, 9) [\[s\]](#)
- [symptom](#) (1) [\[s\]](#)
- [woman](#) (1, 1)

Actions

Add CE

Search

General information **CE Query Builder (CEQB)** **Lexical info**

Query:

for which P1 and P2 is it true that
(the person P1 is the parent of the person P2)

Results (5 rows):

| # | P1 | P2 | CE |
|---|------------------------|------------------------|--|
| 1 | Bill | John | the person 'Bill' is the parent of the person 'John'. |
| 2 | David | Max | the person 'David' is the parent of the person 'Max'. |
| 3 | David | Nellie | the person 'David' is the parent of the person 'Nellie'. |
| 4 | John | Jean | the person 'John' is the parent of the person 'Jean'. |
| 5 | Maggie | John | the person 'Maggie' is the parent of the person 'John'. |

Entity

Concept details for [person](#) below:

| Name | person [r] |
|---|---|
| parent(s) [all] or [direct] | <ul style="list-style-type: none">primary thingthing |
| children [all] or [direct] | <ul style="list-style-type: none">manwomanpatient |
| creation date | 1369063016762 |
| instances | 9 (simple list) (list details) (exact list) |
| | <ul style="list-style-type: none">age (constant) [c] [r]belongs to (set of things) [c] [r] |

Sentences

Sources

Saved queries & rules

Errors (0) Warnings (0) Debugs (0) Alerts (0)

No errors were returned in the last request.

Constructing a Rule in CE Store

CE Store browser - Alpha v1.2 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

localhost:8080/CeStoreWeb/ Boeing Search

Most Visited Latest Headlines Getting Started

CE Store browser - Alpha v1.2

Last transaction (executeCeRule) took 0.0 seconds. Store contains 274 instances and 748 sentences. Code version=1.2.4587 [refresh the page](#) [login](#)

Concepts

Actions

Add CE

Add CE sentence:

conceptualise the person P1 ~ is the grandparent of ~ the person P2.

[Validate CE sentence\(s\)](#)

[Submit CE sentence\(s\)](#)

General information CE Query Builder (CEQB) Lexical information Extract

Entity

Sentences

Sources

Saved queries & rules

[Refresh saved query/rule list](#)

Queries:

- != test ([load](#), [show](#), [execute](#))
- = test ([load](#), [show](#), [execute](#))
- brother of ([load](#), [show](#), [execute](#))
- brotherless 1 ([load](#), [show](#), [execute](#))
- brotherless 2 ([load](#), [show](#), [execute](#))
- close relation to patient ([load](#), [show](#), [execute](#))
- grandparent ([load](#), [show](#), [execute](#))
- older brothers 30 ([load](#), [show](#), [execute](#))
- older brothers 5 ([load](#), [show](#), [execute](#))
- older brothers 65 ([load](#), [show](#), [execute](#))
- parent ([load](#), [show](#), [execute](#))
- patient lump close relation ([load](#), [show](#), [execute](#))
- patient presents with lump ([load](#), [show](#), [execute](#))
- person suffers from cancer ([load](#), [show](#), [execute](#))
- query1 ([load](#), [show](#), [execute](#))
- temporary ([load](#), [show](#), [execute](#))
- younger brothers ([load](#), [show](#), [execute](#))

the person P1 [9] is the parent of [5] the person P2 [9]

is the grandparent of of [conclusion] is the parent of [5] the person P3 [9]

[grandparent]

if

(the person P1 is the parent of the person P2) and

(the person P2 is the parent of the person P3)

then

(the person P1 is the grandparent of of the person P3)

Clear query canvas

Validate query

Execute as query or rule

Save query

Errors (0) Warnings (0) Debugs (0) Alerts (0)

No errors were returned in the last request.

Rule Execution in CE Store: Rationale

The screenshot displays the CE Store browser interface in a Mozilla Firefox window. The browser's address bar shows the URL `localhost:8080/CeStoreWeb/`. The page header indicates the last transaction (executeCeRule) took 0.0 seconds, the store contains 274 instances and 748 sentences, and the code version is 1.2.4587. There are links for [refresh the page](#) and [login](#).

The interface is divided into several sections:

- Left Panel:** Contains a sidebar with 'Concepts', 'Actions', and 'Add CE'. The 'Add CE sentence' section shows a conceptualization rule: `conceptualise the person P1 ~ is the grandparent of ~ the person P2.` Below this are links for [Validate CE sentence\(s\)](#) and [Submit CE sentence\(s\)](#).
- Top Panel:** Includes tabs for 'General information', 'CE Query Builder (CEQB)', 'Lexical information', and 'Extract'. The 'CE Query Builder (CEQB)' tab is active, showing a query:

```
[grandparent]
if
  ( the person P1 is the parent of the person P2 ) and
  ( the person P2 is the parent of the person P3 )
then
  ( the person P1 is the grandparent of of the person P3 )
```
- Right Panel:** Contains a list of entities (Entity, Sentences, Sources) and a section for 'Saved queries & rules'. It includes a link for [Refresh saved query/rule list](#) and a list of queries with their respective actions (load, show, execute).
- Results Section:** Displays the results of the query execution, showing 2 rows. The first row (1) shows that 'the person 'Bill'' is the grandparent of 'the person 'Jean'' because 'the person 'Bill' is the parent of the person 'John' and the person 'John' is the parent of the person 'Jean'' [grandparent]. The second row (2) shows that 'the person 'Maggie'' is the grandparent of 'the person 'Jean'' because 'the person 'Maggie' is the parent of the person 'John' and the person 'John' is the parent of the person 'Jean'' [grandparent].
- Bottom Panel:** Includes a search bar and a status bar showing 'Errors (0)', 'Warnings (0)', 'Debugs (0)', and 'Alerts (0)'. A message states: 'No errors were returned in the last request.'



Example CE Applications

- Fusion of hard data (from sensors) and soft data (human reports) for situation awareness on the battlefield
 - Expression of information needs
 - Description of asset capabilities
 - Conversion of information products into a machine-processable form
- Real-time integration of maps, photos, and messages about events
 - Assignment of objects (e.g. buildings, vehicles) to locations on a map
 - Association of photos of objects with icons on the map
 - Identification and location of objects from short messages on the map
- Collaborative planning
 - Specialized form of general problem solving
 - Support from higher level models for spatial and temporal reasoning
- Fact extraction from free text
 - See later discussion



Using CE to Quickly Instantiate a Map

Add CE

Add CE sentence:

```
there is a spatial thing  
named 'x' that  
  has '51.61721910' as  
latitude and  
  has '-2.74801318' as  
longitude.
```

[Validate CE sentence\(s\)](#)

[Submit CE sentence\(s\)](#)





Benefits of CE and CE-Based Tools

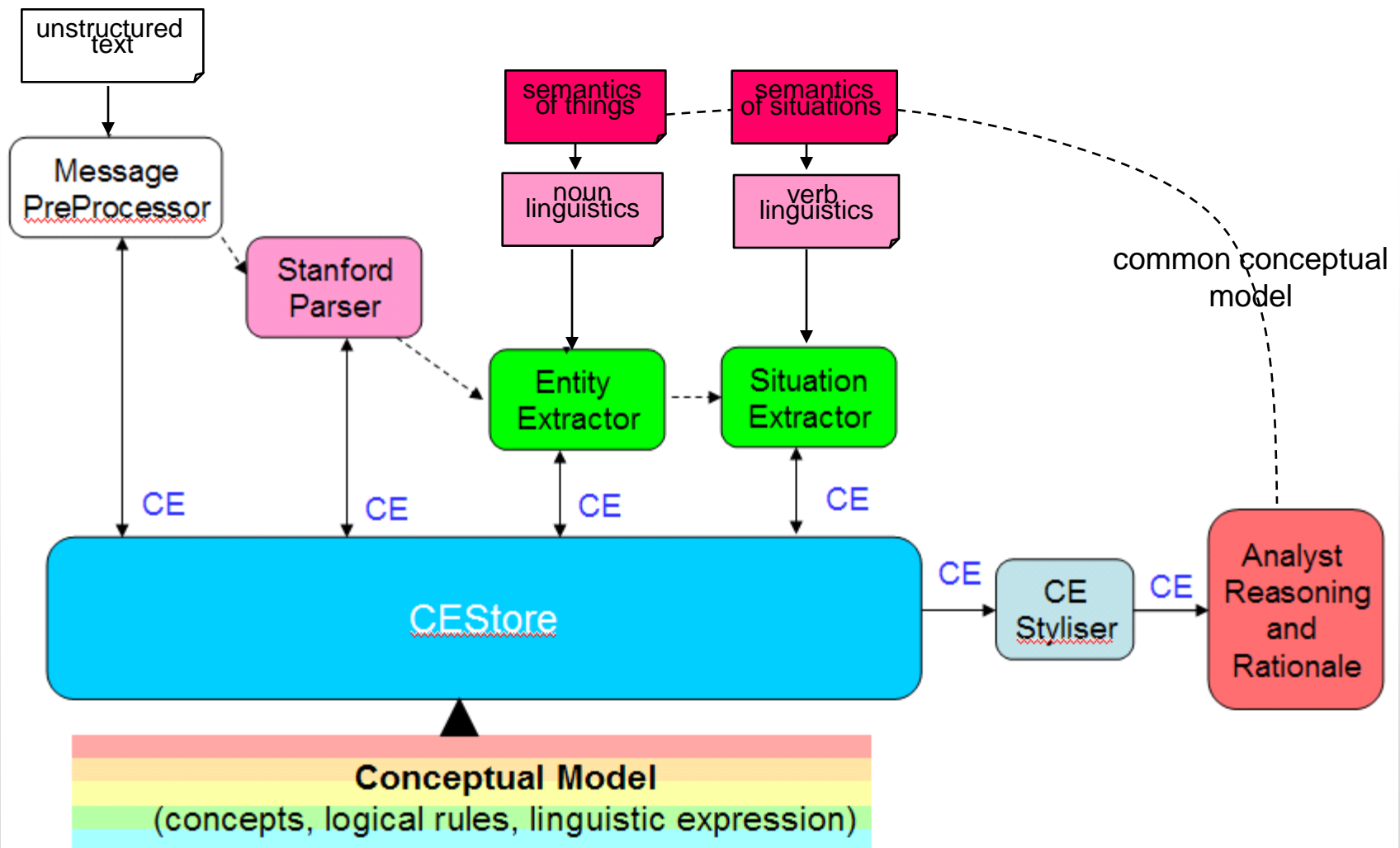
- Provide a simplified and common form of expression in English
- Provide automated tools to enable access to standardized terminology and reinforce consistent language use
- Provide a method to allow extension of standard terminology
- Provide communication aid tools that encourage human-machine interaction, best leveraging human knowledge and computer processing capabilities
- Provide capabilities to construct and extend knowledge models
- Provide an intuitive, CE-based capability for end users to query information from the available information sources



CE Plays Dual Role in Fact Extraction

- CE as the target of the NL processing
 - Facts in documents can then be used for further reasoning
- CE as a means of describing the NL processing
 - To share understanding of the linguistic processing
 - E.g. as part of rationale for a conclusion
 - To help configure NL tooling at various levels
 - Linguists and knowledge engineers can add rules and infrastructure
 - Domain experts can add domain rules
 - End users can add new words and concepts

Architecture for Fact Extraction and Reasoning





Making Our “Intuitions about Language” Accessible (I)

“Nouns stand for things”

if (there is a noun phrase named NP)
then
 (the noun phrase NP stands for the thing T).

“Nouns tells us what type of thing”

if (the noun phrase NP has the noun N as head and stands for the thing T)
and
 (the noun N expresses the concept C)
then
 (the thing T is a C).

“the call was monitored ...”



there is a communication named #26.



Making our “intuitions about language” accessible (II)

Verbs refer to “situations”

- A situation is “something happening in the world”:
 - an event, action, state (*from verb phrases*)
 - things (*from noun phrases*)
 - roles that these things play in the situation (*from phrase structure*)
 - location, time (*from prepositional phrases*).
- For example:

there is a communications monitoring situation named #39 that has
the call #15 as patient role and
has the thing #17 as source role and
has the thing #27 as destination role.



Domain Semantics

■ Modelling “Communications”

- Reports speak about monitoring **communications** between people together with the things that were said

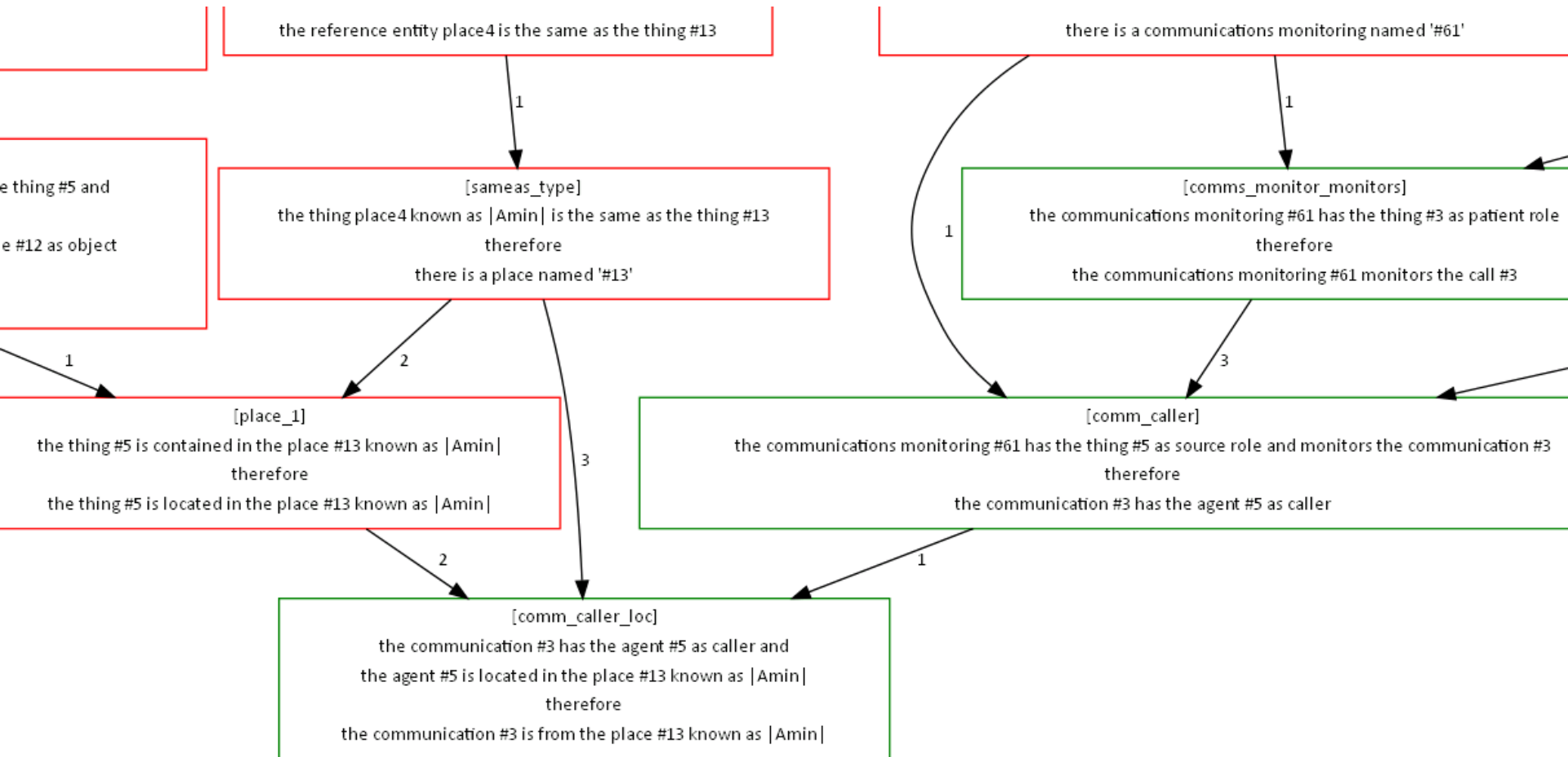
conceptualise

a “**communication**” C that
has the agent A as “**caller**” and
has the agent B as “**recipient**” and
has the value D as “**date**” and
has the value T as “**time**” and
has the value V1 as “**caller utterance**” and
has the value V2 as “**recipient utterance**” and
“**is from**” the place FROM and
“**is to**” the place TO.



Rationale Shows Steps Leading to a Fact

Specific
Semantics





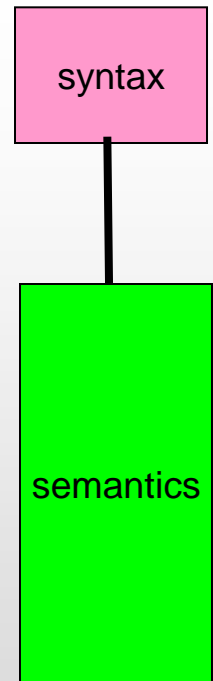
Extending CE: Linguistic Frames

there is a linguistic frame named np3 that
has 'a person' as example and
defines the noun phrase NP and

has the sequence
(the determiner DET and the noun COMMON)
as syntactic pattern and

has the statement that
(the noun COMMON expresses the entity concept EC)
as preconditions and

has the statement that
(the thing X realises the entity concept EC) and
(the noun phrase NP stands for the thing X)
as semantic statement.





On-going and Future Work

- Extend the CE lexicon by **leveraging publicly available lexical-semantic resources** such as WordNet and VerbNet
- **Leverage ERG** for CE fact extraction
- Enhance the representation of **uncertainty**, especially as expressed in natural language, and incorporate mechanisms for propagating uncertainty and supporting **argumentation**
- To extend the research in the **semantics of linguistic processing** to handle more complex language phenomena and to allow **guidance of language parsing via domain models**
- Extend CE syntax and semantics to enhance **semantic expression** and **stylistic felicity**, based on our work in natural language processing for fact extraction



The International Technology Alliance

ACKNOWLEDGMENT

This research was sponsored by the US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defence, or the UK Government. The US and UK Governments are authorised to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.



BACKUP SLIDES



CE, First Order Logic, and OWL

- CE and First Order Logic
 - Basic CE sentences are all given a FOL semantics in the definition of CE reference
 - There are some parts of CE that have not yet been given a formal semantics in CE
 - E.g. assumption-based logic
 - Not all of FOL can be represented in CE
 - E.g. certain combinations of existential quantifiers embedded in the scope of a universal quantifier
- CE and OWL
 - there are some parts of CE that cannot be represented in OWL
 - E.g. rules, assumptions, although potentially these could be represented in extensions to OWL such as RIF
 - There are some parts of OWL that cannot (easily) be represented in CE
 - E.g. lists of explicit values for properties
 - There are one or two fundamental differences in philosophy
 - E.g. we prefer to make (nearly) all of the rules of inference to be explicit, whereas in OWL there are many implicit rules of inference
- For more on the relationship between CE, FOL, and web modeling languages, see Mott, D.: The representation of logic within semantic web languages, ITACS, [url:https://www.usukita.org/papers/5242/details.html](https://www.usukita.org/papers/5242/details.html) (2009)