
Architecture of a Web-based Predictive Editor for Controlled Natural Language Processing

Stephen Guy
Rolf Schwitter

Stephen.Guy@mq.edu.au
Rolf.Schwitter@mq.edu.au

Overview

- 1. Introduction
- 2. Predictive Editor Requirements
- 3. Predictive Editor Architecture
- 4. Analysis of Predictive Editor
 - Review of Features
 - Future Research

1. Introduction

- Controlled Natural Language
- What is a Predictive Editor?
- Aims of this Predictive Editor & CNL System
- Client-Server Architecture Overview

1. Introduction - CNL

- Formal:
 - A controlled natural language (CNL) is a subset of a natural language whose grammar and vocabulary have been restricted in order to eliminate ambiguity and reduce complexity for automated reasoning.
- Informal:
 - CNL: Intermediary language between expressive (English) and unambiguous logic expressions
- Use for Specification System that can run logic

1. Introduction - Predictive Editor

- Formal:
 - Typically, the writing process of a CNL is supported by an authoring tool called a Predictive Editor.
 - The authoring tool provides feedback for the user about the coverage of the CNL.
- Informal:
 - A tool to support the learning and writing process

1. Introduction - Predictive Editor

- Prompt user with allowed options for next input
- Restrict user from entering non-allowed input
- Summarise entered information to user
- Prompt user when critical information is missing
- + ?

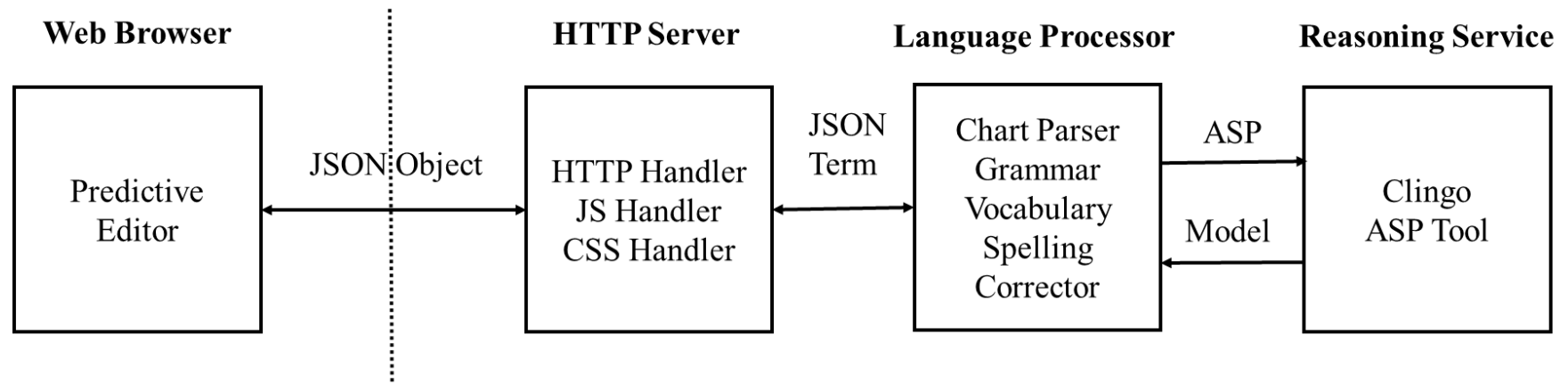
1. Introduction

Characteristics of this Predictive Editor

- Portable web-based system connected to remote server
- Allow text input and interactive pull-down menus
- Display multiple possible sentence completions
 - e.g user enters characters “The”
 - Completion: “There is...”
 - Completion: “The parent...”
- Cater for anaphoric references
- Ability to add new words to user-lexicon
- ASP allows for non-monotonic reasoning

1. Introduction

Predictive Editor & PENG(ASP) Architecture



- Portable web based Client-Server
- JSON & Asynchronous AJAX
- ASP & Weak Negation

1. Introduction

Predictive Editor User Interface

File ▾ Options ▾ Help ▾

PREDICTIVE EDITOR FOR CONTROLLED NATURAL LANGUAGE

CNL INPUT

Parents of a child normally

Lookahead Information ▾ Anaphoric Expressions ▾

"Parents of a child normally" + copula

DATA SUMMARY

ENTERED TEXT:

Sam is a child.
Every father of a child is a parent of the child.

GENERATED PARAPHRASES:

Sam is a child .
Every father of a child is a parent of <ana> the child </ana> .

ANSWER SET PROGRAM:

child(sam).
parent(A,B) :- father(A,B), child(B).

RESULT:

clingo version 4.2.1
Reading from asp.lp
Solving...
Answer: 1
child(sam)

2. Predictive Editor Requirements

- Edit information already typed
- Display next Lookahead (LAH) Information if possible
- Display current LAH if still possible
- Display multiple sentence completions
- Display anaphoric references
- Add new words to lexicon

2. Predictive Editor Requirements: Display Next LAH If Valid

CNL INPUT

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence: constraint

DATA SUMMARY

ENTERED TEXT:
Sam is a child.

GENERATED PARAPHRASES:
Sam is a child .

ANSWER SET PROCESSING:
child(sam).

RESULT:
clingo version 4.2.1
Reading from asp.lp
Solving...
Answer: 1
child(sam)
SATISFIABLE
Models : 1

copula	
determiner	A
existential there	An
name	Every
noun	For every
preposition	The
variable	
wh adverb	
wh determiner	
wh pronoun	

CNL INPUT

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence: constraint

DATA SUMMARY

ENTERED TEXT:
Sam is a child.

GENERATED PARAPHRASES:
Sam is a child .

ANSWER SET PROCESSING:
child(sam).

RESULT:
clingo version 4.2.1
Reading from asp.lp
Solving...
Answer: 1
child(sam)
SATISFIABLE
Models : 1

copula	
determiner	
existential there	There is
name	There are
noun	
preposition	
variable	
wh adverb	
wh determiner	
wh pronoun	

2. Predictive Editor Requirements: Display Next LAH If Valid

CNL INPUT

The

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence:

"The" + adjective

ENTERED TEXT: noun actor
Sam is a child.

GENERATED PARAI ordinal boxer
Sam is a child .

ANSWER SET PROGRAM: chef
child(sam). child

RESULT: department
clingo version 4.2.1
Reading from asp.lp
Solving... father
Answer: 1 guard
child(sam) husband
SATISFIABLE job

2. Predictive Editor Requirements: Display Current LAH If Still Possible

CNL INPUT

The

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence:

"The" + adjective

ENTERED TEXT:
Sam is a child.

GENERATED PARAI noun actor
ordinal boxer

Sam is a child . chef

ANSWER SET PROGRAM:
child(sam). child

RESULT:
clingo version 4.2.1
Reading from asp.lp
Solving...
Answer: 1
child(sam)
SATISFIABLE

department
father
guard
husband
job

CNL INPUT

The

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence: constraint

"The" + copula

ENTERED TEXT:
Sam is a child.

GENERATED PARAI determiner
existential there There is
name There are

Sam is a child .

ANSWER SET PROG
child(sam). noun

RESULT:
clingo version 4.2.1
Reading from asp.lp
Solving...
Answer: 1
child(sam)
SATISFIABLE

preposition
variable
wh adverb
wh determiner
wh pronoun

2. Predictive Editor Requirements: Display Current LAH If Still Possible

CNL INPUT

The|

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence: ▸

"The" + ▸ adjective ▸

Program Status: noun: sg ▸ actor

ENTERED TEXT: ordinal ▸ boxer

> chef

child

department

father

guard

CNL INPUT

The|

Lookahead Information ▾ Anaphoric Expressions ▾

start sentence: ▸ constraint ▸

"The" + ▸ copula: sg ▸

Program Status: determiner ▸

ENTERED TEXT: existential there ▸ There is

> name: sg ▸ There are

noun: pl ▸

preposition ▸

variable ▸

2. Predictive Editor Requirements: Add Word

CNL INPUT

Every father of a child is|

Lookahead Information ▾

Anaphoric Expressions ▾

"Every father of a child is" +

adjective

absent

DATA SUMMARY

ENTERED TEXT:

Sam is a child.

GENERATED PARAPHRAS

Sam is a child .

ANSWER SET PROGRAM:

child(sam).

RESULT:

clingo version 4.2.1

Reading from asp.lp

Solving...

Answer: 1

child(sam)

SATISFIABLE

cardinal

educated

determiner

female

negation

male

preposition

truthful

relational adjective

variable

CNL INPUT

Every father of a child is beautiful

Lookahead Information ▾

Anaphoric Expressions ▾

"Every father of a child is" +

adjective

add: beautiful

DATA SUMMARY

ENTERED TEXT:

Sam is a child.

GENERATED PARAPHRAS

Sam is a child .

ANSWER SET PROGRAM:

child(sam).

RESULT:

clingo version 4.2.1

Reading from asp.lp

Solving...

Answer: 1

child(sam)

SATISFIABLE

cardinal

absent

determiner

educated

negation

female

preposition

male

relational adjective

truthful

variable

2. Predictive Editor Requirements: Anaphoric References

CNL INPUT

Lookahead Information ▾ **Anaphoric Expressions** ▾

Sam
The child

DATA SUMMARY

ENTERED TEXT:

Sam is a child.

GENERATED PARAPHRASES:

Sam is a child .

ANSWER SET PROGRAM:

child(sam).

RESULT:

clingo version 4.2.1

Reading from asp.lp

Solving...

Answer: 1

child(sam)

SATISFIABLE

CNL INPUT

Lookahead Information ▾ **Anaphoric Expressions** ▾

Alice
Sam
the child

DATA SUMMARY

ENTERED TEXT:

Sam is a child.

GENERATED PARAPHRASES:

Sam is a child .

ANSWER SET PROGRAM:

child(sam).

RESULT:

clingo version 4.2.1

Reading from asp.lp

Solving...

Answer: 1

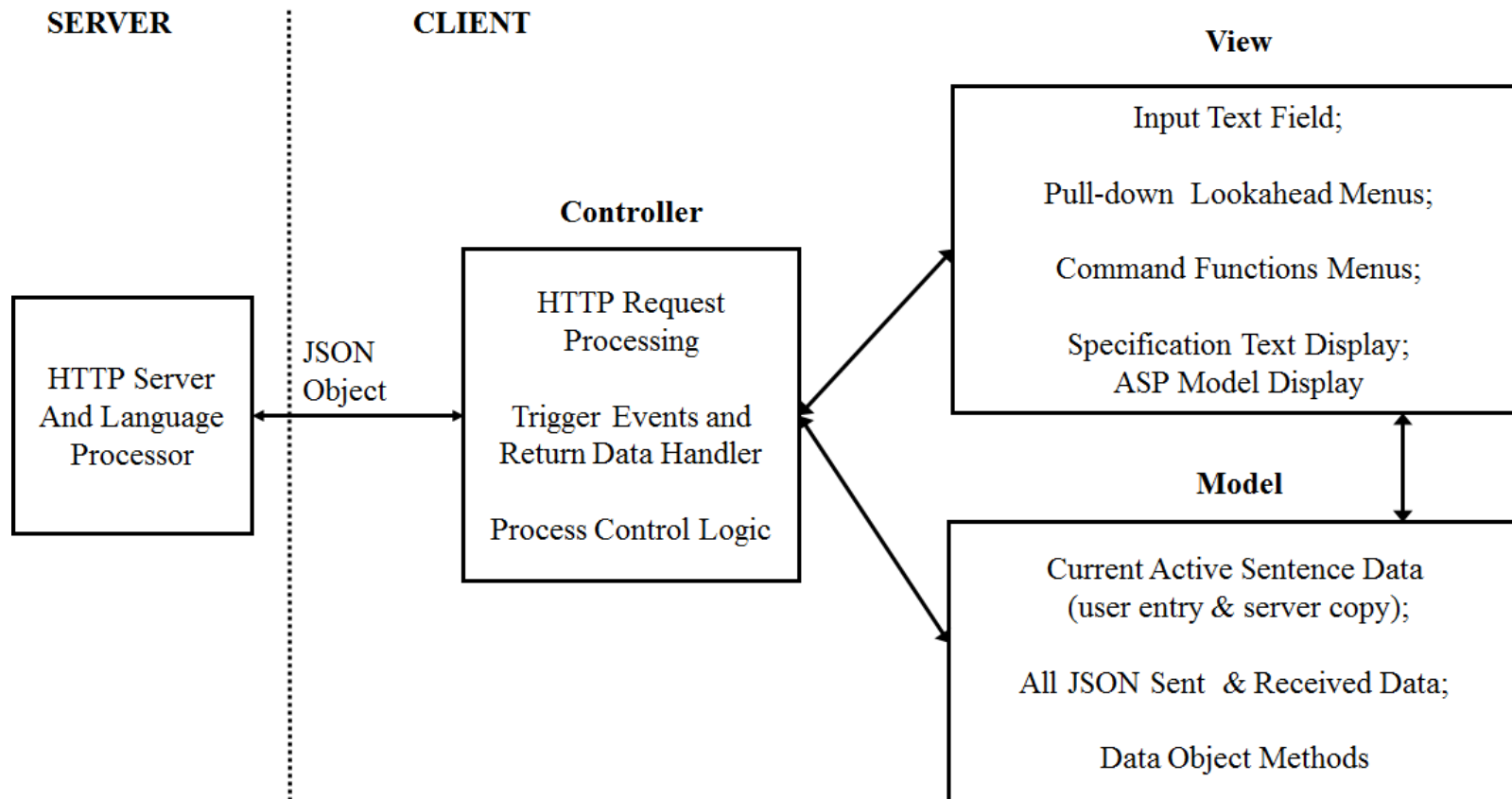
child(sam)

SATISFIABLE

3. Predictive Editor Architecture

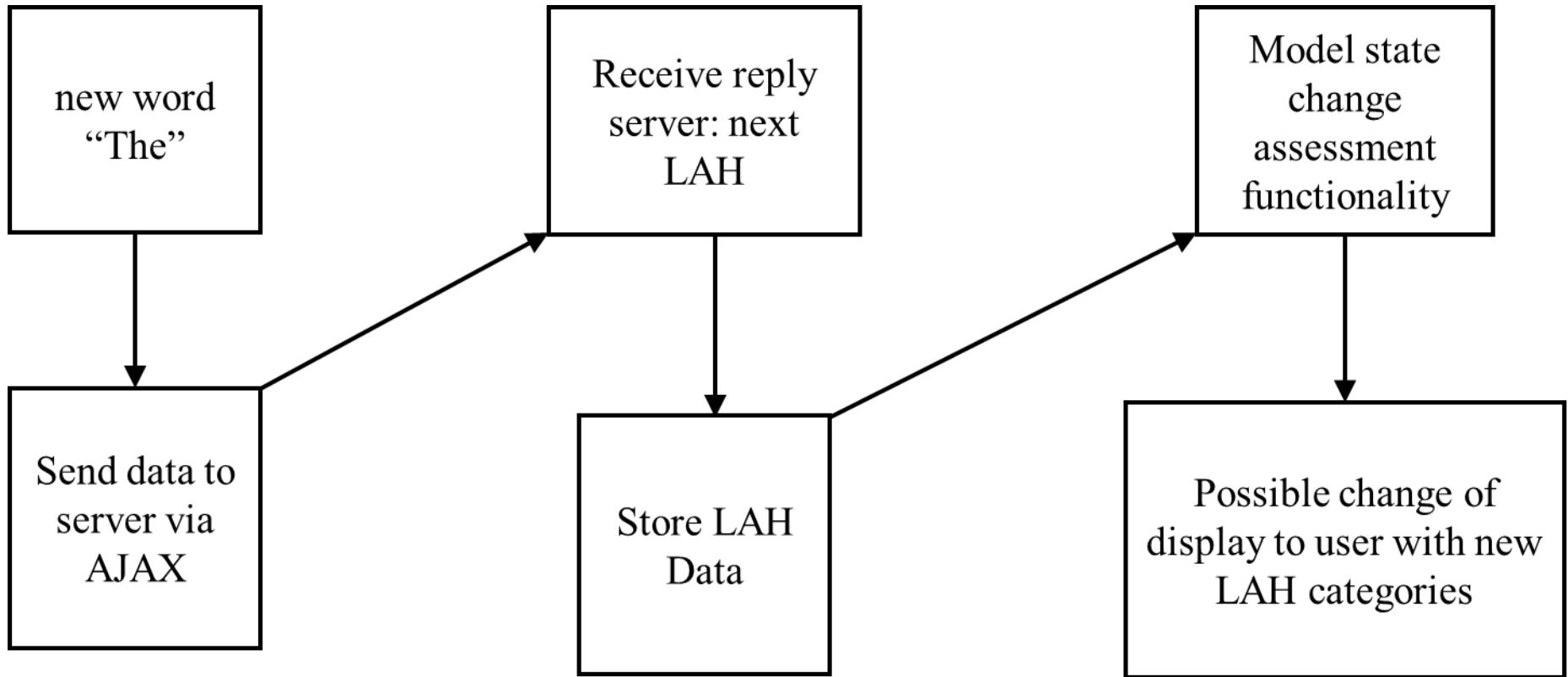
- Event Driven
 - Delays in Ajax communication with server
 - Delays in display structures
 - Only change view of model when required
- Based loosely on MVC
- JS Objects with associated methods, e.g .match_LAH()
- Recursive Pipeline Processing of Objects: Data, Processing Function & Next Object
- Client-Server communication not directly synchronized to user display

3. Predictive Editor Architecture - Client



3. Predictive Editor Architecture

Pipeline Processing



3. Predictive Editor Architecture

Input Stack System

- Represent 'sentence' text data as Stack of JSON tokens
 - Stack for tokens at input
 - Stack for tokens sent to server
- Compare Stacks of Input Data with Model sent to server.
 - Create Stack of difference input vs model ('Delta')
- If token in Delta stack matches LAH then send to server
- Display LAH data when required
- Input Stack System: Really a system based on an input stack comparison followed by a subsequent state decision system

3. Predictive Editor Architecture

Input Stack System – Overview: Word Level

Text Input	Model: Sent	Model: Input	Model: Delta	LAHs	Match LAH	Send Delta
""	Generate initialiser->	{"token": "", "pos":0}	{"token": "", "pos":0}		N/A	"" pos 0
"The"	""	"The"	"The"	The, There, A	Y	"The" pos 0
"The child"	"The"	"The child"	"child"	child, parent	Y	"child" pos 1
"The child works"	"The child"	"The child works"	"works"	works, lies	Y	"works" pos 2
"The child works."	"The child works"	"The child works."	."	and, or, "."	Y	." pos 3

3. Predictive Editor Architecture

Input Stack System – Character Level ->The

Text Input	Model: Sent	Model: Input	Model: Delta	LAHs	Match LAH	Send Delta
""	Generate initialiser->	{"token": " ", "pos": 0}	{"token": " ", "pos": 0}		N/A	Y
"T"	{ "token": " ", "pos": 0 }	{ "token": "T", "pos": 0 }, { "token": " ", "pos": 0 }	{ "token": "T", "pos": 0 }, { "token": " ", "pos": 0 }	add: The, There, A	N	N
"Th"	{ "token": " ", "pos": 0 }	{ "token": "Th", "pos": 0 }, { "token": " ", "pos": 0 }	{ "token": "Th", "pos": 0 }, { "token": " ", "pos": 0 }	The, There, A	N	N
"The"	{ "token": " ", "pos": 0 }	{ "token": "The", "pos": 0 }, { "token": " ", "pos": 0 }	{ "token": "The", "pos": 0 }, { "token": " ", "pos": 0 }	The, There, A	Y	Y
	{ "token": "The", "pos": 0 }			add: child, parent		

3. Predictive Editor Architecture

Input Stack System – Character Level ->The

Text Input	Model: Sent	Model: Input	Model: Delta	LAHs	Match LAH	Send Delta
""	Generate initialiser->	{" ", 0}	{" ", 0}		N/A	Y
"T"		{"T", 0} {" ", 0}	{"T", 0}	add: The, There, A	N	N
"Th"		{"Th", 0} {" ", 0}	{"Th", 0}	The, There, A	N	N
"The"		{"The", 0} {" ", 0}	{"The", 0}	The , There, A	Y	Y
		{"The", 0}		add: child, parent		

3. Predictive Editor Architecture

Input Stack System – The ->There

Text Input	Model: Sent	Model: Input	Model: Delta	LAHs	Match LAH	Send Delta
"The"		{"The", 0} {" ", 0}	{"The", 0}	The, There, A	Y	Y
"Ther"	{"The", 0} {" ", 0}	{"Ther", 0} {" ", 0}	{"Ther", 0}	add: child, parent; The, There, A	N N	N
"There"	{"The", 0} {" ", 0}	{"There", 0} {" ", 0}	{"There", 0}	child, parent; The, There, A	N Y	Y
	{"There", 0} {" ", 0}			add: is, are		

3. Predictive Editor Architecture

Input Stack System – The +

Text Input	Model: Sent	Model: Input	Model: Delta	LAHs	Match LAH	Send Delta
"The"	{ " ", 0 }	{ "The", 0 } { " ", 0 }	{ "The", 0 }	The, There, A	Y	Y
"The c"	{ "The", 0 } { " ", 0 }	{ "c", 1 } { "The", 0 } { " ", 0 }	{ "c", 1 }	add: child, parent; The, There, A	N N	N
"The ch"	{ "The", 0 } { " ", 0 }	{ "ch", 1 } { "The", 0 } { " ", 0 }	{ "ch", 1 }	child, parent; The, There, A	N N	N
"The child"	{ "The", 0 } { " ", 0 }	{ "child", 1 } { "The", 0 } { " ", 0 }	{ "child", 1 }	child, parent; The, There, A	Y N	Y
	{ "child", 1 } { "The", 0 } { " ", 0 }			add: works, lies		

4. Analysis of Predictive Editor -PENG ASP

PREDICTIVE EDITOR FOR CONTROLLED NATURAL LANGUAGE

CNL INPUT

Lookahead Information ▾

Anaphoric Expressions ▾

DATA SUMMARY

ENTERED TEXT:

Sam is a child.
John is the father of Sam and Alice is the mother of Sam.
Every father of a child is a parent of the child.
Every mother of a child is a parent of the child.
Parents of a child normally care about the child.
If a parent of a child is provably absent then the parent abnormally cares about the child.
John does not care about Sam.
Alice is absent.

GENERATED PARAPHRASES:

Sam is a child .
John is the father of <ana> Sam </ana> and Alice is the mother of <ana> Sam </ana> .
Every father of a child is a parent of <ana> the child </ana> .
Every mother of a child is a parent of <ana> the child </ana> .
Parents of a child normally care about <ana> the child </ana> .
If a parent of a child is <ins> not </ins> provably <ins> not </ins> absent then <ana> the parent </ana> abnormally cares about <ana> the child </ana> .
<ana> John </ana> does not care about <ana> Sam </ana> .
<ana> Alice </ana> is absent .

ANSWER SET PROGRAM:

4. Analysis of Predictive Editor -PENG ASP

GENERATED PARAPHRASES:

Sam is a child .

John is the father of <ana> Sam </ana> and Alice is the mother of <ana> Sam </ana> .

Every father of a child is a parent of <ana> the child </ana> .

Every mother of a child is a parent of <ana> the child </ana> .

Parents of a child normally care about <ana> the child </ana> .

If a parent of a child is <ins> not </ins> provably <ins> not </ins> absent then <ana> the parent </ana> abnormally cares about <ana> the child </ana> .

<ana> John </ana> does not care about <ana> Sam </ana> .

<ana> Alice </ana> is absent .

ANSWER SET PROGRAM:

child(sam).

father(john,sam).

mother(alice,sam).

parent(A,B) :- father(A,B), child(B).

parent(C,D) :- mother(C,D), child(D).

care(E,F) :- parent(E,F), child(F), not ab(d_care(E,F)), not -care(E,F).

ab(d_care(G,H)) :- parent(G,H), child(H), not -absent(G).

-care(john,sam).

absent(alice).

RESULT:

clingo version 4.2.1

Reading from asp.lp

Solving...

Answer: 1

child(sam) father(john,sam) mother(alice,sam) -care(john,sam) absent(alice) parent(john,sam) parent(alice,sam) ab(d_care(john,sam)) ab(d_care(alice,sam))

SATISFIABLE

Models : 1

Calls : 1

Time : 0.000s (Solving: 0.00s 1st Model: 0.00s Unsat: 0.00s)

CPU Time : 0.000s

4. Analysis of Predictive Editor Future Developments

- User Login:
 - User lexicon for added words
 - Administrator can add words to main lexicon for all users
 - Tailor complexity of grammatical display in menus
 - advanced user: `noun`
 - basic user: `noun:pl or name:sg`

 - advanced user: `adjective`, `adverb`, `noun`
 - basic user: `function words`
- Reasoning, summary and meta-data from ASP Data

4. Analysis of Predictive Editor Future Developments

- Upgrade help
 - Hover over section
 - Right click pull down menu item
 - What is a 'copula'? OR Transitive vs Intransitive
- Upgrade file loading, saving and editing
 - Save and load specifications entered
 - Edit previously entered sentences
- Usability & Appearance upgrades (& evaluation)
 - Hierarchical arrow menus on Summary Data Sections
 - Scroll menu items

4. Analysis of Predictive Editor Summary

- Web based, portable
- Allows for text input and interactive pull-down menus
- Displays multiple possible lexical categories
- Add additional words to lexicon
- Display includes anaphoric references
- ASP allows Weak Negation
- File loading and saving for specifications
- Use as tool for teaching logic