

Generating Mathematical Word Problems

Sandra Williams

Computing Department, The Open University,
Milton Keynes, U.K.

A mathematical word problem

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Contents

- Motivations
- Anatomy of a Mathematical Word Problem
- The Semantic Web and NLG
- Generating a MWP from the Semantic Web – a worked example
- Varying cognitive difficulty
- Conclusion

Motivations

- Level of numeracy in the general population is low
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution

Motivations

- Level of numeracy in the general population is low
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution
- ITS question bottleneck
 - Natural Language Generation (NLG) could provide a solution

Motivations

- Level of numeracy in the general population is low
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution
- ITS question bottleneck
 - Natural Language Generation (NLG) could provide a solution
- NLG data source bottleneck
 - Semantic Web could provide a solution

Motivations

- **Level of numeracy in the general population is low**
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution
- **ITS question bottleneck**
 - Natural Language Generation (NLG) could provide a solution
- **NLG data source bottleneck**
 - Semantic Web could provide a solution
- **Cognitive complexity of Mathematical Word Problems**
 - Numeracy skills
 - Literacy skills
 - Problem-solving skills

Motivations

- **Level of numeracy in the general population is low**
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution
- **ITS question bottleneck**
 - Natural Language Generation (NLG) could provide a solution
- **NLG data source bottleneck**
 - Semantic Web could provide a solution
- **Cognitive complexity of Mathematical Word Problems**
 - Numeracy skills
 - Literacy skills
 - Problem-solving skills
- **Challenging research issues for NLG**

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Question story

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Question story

Numerical expressions

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Question story

Numerical expressions

Relevant objects(s)

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Question story

Numerical expressions

Relevant objects(s)

Relevant relationships(s)

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Question story

Numerical expressions

Relevant objects(s)

Relevant relationships(s)

Request for unknown value

What is a mathematical word problem?

A formica japonica worker ant weighs 0.004 gram. It can walk while holding in its mouth an object weighing 5 times as much as its own body. How many grams can a worker ant carry?

Question story

Numerical expressions

Relevant objects(s)

Relevant relationships(s)

Request for unknown value

Extraneous material



Motivations

- **Level of numeracy in the general population is low**
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution
- **ITS question bottleneck**
 - Natural Language Generation (NLG) could provide a solution
- **NLG data source bottleneck**
 - Semantic Web could provide a solution
- **Cognitive complexity of Mathematical Word Problems**
 - Numeracy skills
 - Literacy skills
 - Problem-solving skills
- **Challenging research issues for NLG**

Motivations

- Level of numeracy in the general population is low
 - >1:5 adults in the U.K. have poor numeracy (Govt. report, 2003)
 - Tutoring required
 - Intelligent Tutoring Systems (ITSs) could provide a solution
- ITS question bottleneck
 - **Natural Language Generation (NLG) could provide a solution**
- NLG data source bottleneck
 - **Semantic Web could provide a solution**
- Cognitive complexity of Mathematical Word Problems
 - Numeracy skills
 - Literacy skills
 - Problem-solving skills
- Challenging research issues for NLG

NLG and the Semantic Web

- Types of Natural Language Generators
 - Data-to-text
 - Text-to-text (text reformulation)
- OWL (Web Ontology Language)
 - Adopted as a standard for the Semantic Web
 - Based on description logic (DL)
- The Semantic Web as input data for NLG
 - Verbalisation: OWL to Controlled NL (e.g., ACE, SWAT)
 - One sentence per OWL statement

DL Statement	English Realisation
$A \equiv B$	An A is defined as a B
$A \sqsubseteq B$	Every A is a B
$A \sqsubseteq \exists P.B$	Every A Ps a B
$X \in A$	X is an A
$[X, Y] \in P$	X Ps an A
$X \in A \sqcap \exists P.B$	X is an A that Ps a B

SWAT Natural Language Tools

- **Semantic Web Authoring Tools**
 - EPSRC project (PIs Richard Power, Open University and Robert Stevens, University of Manchester)
- **Derives a lexicon from ontology identifier names**
 - #isleOfScilly → “The Isle of Scilly”, “The Isles of Scilly”
 - #has_height → “has a height of”, “have heights of”
 - #dog → “dog”, “dogs”
- **‘Verbalises’ using Prolog grammar**
- **Builds a structured text**
 - Glossary-style entries (Williams, Third and Power, 2011)
 - Aggregation of similar information (Williams and Power, 2010)
 - Adds NL descriptions as OWL-XML annotations (Third, Williams and Power, 2011)
 - User evaluations (Stevens et al., 2010, Williams et al., 2010)
- **Available as a web application**
swat.open.ac.uk/tools/

SWAT Tools Output (swat.open.ac.uk/tools/)

The screenshot shows a web browser window with the address bar containing 'swat.open.ac.uk/swat-cgi-bin/cgi-upload.py'. The page content is organized into three distinct sections, each with a title and several key-value pairs. The first section is for 'TOLLS ISLAND (individual)', the second for 'TOPOLOGY VALUE PARTITION (class)', and the third for 'TORCELLO (individual)'. Each section lists 'Typology' and 'Description' (or 'Examples' and 'Distinctions' for the class). The text in the descriptions is underlined, suggesting it is hyperlinked. The browser interface includes navigation buttons, a search bar with the Google logo, and a home button.

TOLLS ISLAND (individual)
Typology Tolls Island is an island.
Description Tolls Island is member of the Isles of Scilly.
Tolls Island is occupied by England.

TOPOLOGY VALUE PARTITION (class)
Typology A topology value partition is an intrinsic property partition.
Description If X has as topology Y then Y is a topology value partition.
Examples The following are topology value partitions: solid topologies, single membrane topologies, and hollow topologies.
Distinctions No selector is a topology value partition.

TORCELLO (individual)
Typology Torcello is an island.
Description Torcello is direct part of Europe.
Torcello is member of Venetian Lagoon Archipelago.
Torcello is occupied by Italy.
Torcello is surrounded by Venetian Lagoon.

A prototype for generating MWP questions

- Uses data from the Semantic Web (OWL ontology)
- Extends SWAT Natural Language Tools
 - Searches an ontology for patterns conforming to a logical structure defined for a MWP
 - Extends aggregation of selected ontology statements to include refactoring
 - Extends the Prolog grammar to handle more complex sentences and templates for requests
 - Generates one question for each pattern found in the input ontology
 - E.g., 31,520 questions of a single type are generated from the travel ontology (TONES repository)

Underlying logical structures of MWP

- Define a logical structure for a MWP (a minimal set of ontology statements conforming to a pattern)
- Possible to define many of these (dozens)
- An example
 - 2 DataPropertyAssertions with *numerical values N1 and N2*
 - 2 *comparable (members of class C) individuals X and Y* with the same numerical data property P

$[X, N1] \in P$	dataPropertyAssertion(P X N1)	<i>X Ps N1.</i>
$[Y, N1] \in P$	dataPropertyAssertion(P Y N2)	<i>Y Ps N2.</i>
$X \in C$	classAssertion(C X)	<i>X is a C.</i>
$Y \in C$	classAssertion(C Y)	<i>Y is a C.</i>

Add other relevant, but extraneous, ontology statements

$[X,Z] \in O$	objectPropertyAssertion(O X Z)	<i>X Os Z.</i>
$[Y,Z] \in O$	objectPropertyAssertion(O Y Z)	<i>Y Os Z.</i>

Aggregation of statements with identical structures

classAssertion(C X) + **classAssertion(C Y)** → **classAssertion(C [X Y])**
X is a C. *Y is a C.* *X and Y are Cs.*

Refactoring statements

$$\begin{array}{l} X \in C \\ \textit{X is a C.} \end{array} \quad + \quad \begin{array}{l} [X, N1] \in P \\ \textit{X Ps N1.} \end{array} \quad \rightarrow \quad \begin{array}{l} X \in C \cap \exists P.N1 \\ \textit{X is a C that Ps N1.} \end{array}$$

Complete Logical Structure for one type of MWP Question

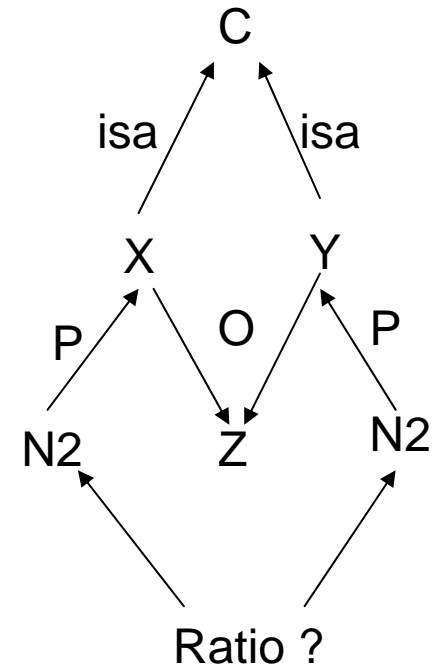
$X \in C \cap \exists P.N1$

$Y \in C \cap \exists P.N2$

$[X,Z] \in O$

$[Y,Z] \in O$

What is the ratio of the two <Ps>?



Realisation through an extended SWAT generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Realisation through a generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Question story

Realisation through a generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Question story

Numerical expressions

Realisation through a generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Question story

Numerical expressions

Relevant objects(s)

Realisation through a generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Question story

Numerical expressions

Relevant objects(s)

Relevant relationships(s)

Realisation through a generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Question story

Numerical expressions

Relevant objects(s)

Relevant relationships(s)

Request for unknown value

Realisation through a generation grammar

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Question story

Numerical expressions

Relevant objects(s)

Relevant relationships(s)

Request for unknown value

Extraneous material

Contents

- Motivations
- Anatomy of a Mathematical Word Problem
- The Semantic Web and NLG
- Generating a MWP from the Semantic Web – a worked example
- **Varying cognitive difficulty**
- Conclusion

Vary readability of the question text

- Low readability – complex sentences
- High readability – simple sentences
- Low-achieving students perform better on MWPs with simple sentences than MWPs with complex sentences
 - Wheeler and McNutt (1983)
- Sentence complexity is related to sentence length
 - Sentence length affects readability (Coleman, 1962)
 - Sentence length is a component of readability formulae (Flesch, 1949)
 - One means of minimising sentence length is to switch off refactoring and aggregation in the NLG system (but text is longer!)

Benbecula is an island. South Uist is an island. Benbecula is a member of the Uists and Barra Archipelago. South Uist is a member of the Uists and Barra Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Vary problem-solving difficulty

- Introduce/Remove distractor numerical values

- Add more DataPropertyAssertion statements from the ontology, e.g.:

Benbecula has an area of 124.

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. Benbecula has an area of 124. South Uist has a population of 1818. What is the ratio of the two populations?

- Introduce/Remove extraneous information

Benbecula and South Uist are islands that are members of the Uists and Barras Archipelago. Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Benbecula has a population of 1219. South Uist has a population of 1818. What is the ratio of the two populations?

Vary mathematical difficulty

- Vary the type of problem
 - In basic maths classes, concepts are taught in order of increasing difficulty. E.g., addition of single-digit numbers is taught before percentages
 - Vary the difficulty on a scale by consulting a maths curriculum
 - U.K. Qualification and Curriculum Authority (1999)
 - Corpus study of precision and mathematical form in discourse (Williams and Power, 2009)
- Vary the order of presentation of numerical values
 - There is a natural order of values for some calculations. E.g., 2:3 for a ratio but 3-2 for a subtraction (if +ve answer required)
 - Reversing the natural order increases difficulty
 - Students rearranged values to the natural order when asked to modify MWPs for their peers (Cohen and Stover, 1981)

Conclusions

- Automatic generation of MWPs is a potential solution for the ITS question bottleneck
- The Semantic Web provides a potential solution to the NLG data source bottleneck
- MWPs are an interesting research area
- Prototype system for generating MWPs implemented and walk-through presented
- System is still to be tested with an ITS system and with users.

Thank You !
Questions ?



Other factors that influence MWP difficulty

- Authenticity (Palm, 2009)

- Is the purpose of MWPs to teach students to apply abstract mathematical principles to real world problems?
- If so, some MWPs are unrealistic:

John's best time to run 100 meters is 17 seconds. How long will it take him to run 1 kilometer?

- 170 seconds is not a realistic answer (Verschaffel, Greer and Corte, 2000)
- Modelling the real world requires more complex maths
- How can we filter out unrealistic MWPs?
 - Not a problem with prototype question type