Patent Summarization and Paraphrasing

David Cinciruk

February 25, 2015
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Patent Summarization

- Patent Summarization is the technique of summarizing and/or paraphrasing patent claims into a more readable format.
- Patent Summarization is a harder technique than regular text summarization especially when one considers dependent claims and how to relate them with the summary.
Overview of their System

- **Claim Dependency Structuring:** Relating dependent claims to the claims they depend on
- **Simplification:** Segments the original sentence and reconstructs them to obtain grammatical sentences
- **Regeneration:** Generates a deep syntactical structure from the above and summarizes them and/or combines them
Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Patent Claims

- Only legal portion of a patent
- Used to determine the extent of protection a patent provides
- Legally, the rest of the patent is only used to support the terms used in the claims
Sample Patent Claim

1. A personal digital assistant (PDA), comprising: a calendar function; an address book function; a processor operably coupled to a memory and a display; wherein the PDA dynamically generates a route path using the processor and the memory from a moveable location associated with the PDA to a destination of the PDA by repetitively dynamically expanding one or more adjacent locations and inserting the adjacent locations into a first data structure wherein one or more first locations of the first data structure are associated with a then existing least cost location in the route path, the route path is dynamically generated from the moveable location, the first locations of the first data structure, and the destination; and wherein at least a portion of the route path is dynamically communicated to the display.

2. The PDA of claim 1, further comprising an interface device operable to audibly communicate the route path.

3. The PDA of claim 1, wherein the first locations are removed from the first data structure as the first locations become part of the route path and the removed first locations combine to form a second data structure.
Grammatical Problems

- Claims are very dense structures that barely resemble standard English at times.
- The language is filled with words that don’t match the typical uses of words or which have very explicit meaning compared to how they are used in real life.
Relating Dependent Claims

- Dependent claims take items from the claim they are dependent on and narrow down the scope to a more implementable form.
- They are not part of the main scope of the invention but instead used to describe what the patentee wants the patent to be used for.
- Dependent claims may conflict with one another showcasing different ideas that the inventor had with his invention.
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
   Three Different Simplification Experiments

Deep Paraphrasing and Summarization
   Sidebar - Meaning-Text Theory
   Preprocessing
   Deep Syntactic Summarization
   Generation Proper

Evaluation of their System
Claim Structures

- Program determines the portion of the parent claim an independent claim refers to and maps them together.
- Able to set how much of this information is to be kept.
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Steps of Simplification

- **POS Tagging:** Standard POS Tagging with Tree-Tagger
- **Segmentation of Claim Sentences:** Uses Rule Based and Machine Learning based segmentation
- **Coreference Resolution:** Determine if two NPs that refer to the same object
- **Building a Clause-Discourse Tree:** Identifies clause structures
- **Reconstruction of Sentences:** Creates grammatically correct independent sentences
Clause-Discourse Structuring
Clause-Discourse Structuring

- **Clause Structuring:** Output is a binary tree whose terminal nodes are sentence segments and whose intermediate nodes specify whether the subtree is a subordination, coordination, or juxtaposition.

- **Clause Tree Flattening:** Binary Tree is flattened to account for n-ary constraints such as coordinates.

- **Projection of the Clause Structure onto the Disclosure Structure:** Intermediate nodes are enriched with discourse information based on Rhetorical Structure Theory labeling the discourse units (spans) as nucleus/satellites and determining the relationship between them according to a set of rules that use the type of constructs and lexical information.
Clause Structuring

- Search for the best clause structure in a space restricted by a set of weighted rules and constraints on the rules
- Rules encode fundamental features for the identification of coordination, subodination, and juxtaposition relations between spans
- Constraints ensure syntactical correctness and global coherence
- Rules, constraints, and weights were developed and adjusted in a series of iterative trials based on evaluation of a small development corpus consisting of 8 manually structured claims (featuring 104 segments)
Rule R is a quadruple $< F_1, F_2, W_0, W_c >$
- $F_1$ and $F_2$ are feature tuples describing the two spans $S_1$ and $S_2$ - consists of $< punct, coord, subord, syntagm, colon >$
  - punct - clause delimiter punctuation preceding the segment
  - coord - coordination marker
  - subord - subordination marker
  - syntagm - span's syntactical group (mainly S, NP, or VP)
  - colon - whether the clause contains a colon
- $W_0$ is the rule's initial weights
- $W_c$ is the set of weighted constraints (weights determine if it's a hard or soft constraint)
# Constraint Table

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Plain coordination, Coordination of subordinations:</em></td>
<td></td>
</tr>
<tr>
<td>C1) Spans have the same syntagm.</td>
<td>Yes=1, No=0.5</td>
</tr>
<tr>
<td>C2) If spans are not separated by a comma, then their core clauses</td>
<td>Yes=1, No=0.5</td>
</tr>
<tr>
<td>are contiguous (and the reverse is true).</td>
<td>(or reverse)</td>
</tr>
<tr>
<td>C3) Non-terminal S2 are constructed with a coordination.</td>
<td>Yes=1, No=0</td>
</tr>
<tr>
<td><em>Coordination of subordinations:</em></td>
<td></td>
</tr>
<tr>
<td>C4) Spans must be introduced by the same subordinated marker.</td>
<td>Yes=1, No=0.1</td>
</tr>
<tr>
<td><em>Semi-colon coordination:</em></td>
<td></td>
</tr>
<tr>
<td>C5) Non-terminal S2 is constructed with a semi-colon coordination.</td>
<td>Yes=1, No=0</td>
</tr>
<tr>
<td><em>Subordination:</em></td>
<td></td>
</tr>
<tr>
<td>C6) Favour right branching.</td>
<td>Yes=1, No=0.1</td>
</tr>
<tr>
<td>C7) Non-terminal S1 is coordination.</td>
<td>Yes=0, No=1</td>
</tr>
<tr>
<td><em>Subordination, Juxtaposition</em></td>
<td></td>
</tr>
<tr>
<td>C8) Favour spans with higher ranking coreferring NPs, taking into account</td>
<td>$1^{st}=1, 2^{nd}=0.9, 3^{rd}=0.8, \text{none}=0.7$</td>
</tr>
<tr>
<td>a focus rank of noun phrases: $1^{st}&gt;2^{nd}&gt;3^{rd} &gt; \text{none}$</td>
<td></td>
</tr>
</tbody>
</table>
Search Problems

> As number of segments grow the number of rules to link them grows exponentially

> A variation of local beam search is used to search among the various trees with the metrics calculated for each application of a rule serving as an objective function

> Goal of algorithm is a rooted tree so the algorithm can backtrack and explore alternatives
Sentence Reconstruction

- Set of rules to convert segments into sentences
- Adds missing nouns or verbs
- Conjugates verbs in non-finite clauses
- Removes initial markers
- Change the order of constituents
OPTIONAL - Discourse Structure Based Summarization

- **Depth:** all spans below a given depth are preserved

- **Discourse Relation:** a list of ordered disclosure relations can be provided, pruning occurs from lowest to highest
  - Purpose > Means > Elaboration-Object-Attribute > Elaboration-Location

- **Discourse Markers:** an ordered list of discourse markers can be provided for pruning
  - When > By > For
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
   Three Different Simplification Experiments

Deep Paraphrasing and Summarization
   Sidebar - Meaning-Text Theory
   Preprocessing
   Deep Syntactic Summarization
   Generation Proper

Evaluation of their System
Segmentation Experiments

- Manually constructed database of about 1500 patents
- Machine Learning - Weka’s J48 Decision Tree Learner
- Rule Based - Best result utilized semi-colons, commas, and about 20 lexical markers and expressions of the patent domain
- Evaluation counts 1:1 alignment between machine segmentation and hand segmentation as truth

<table>
<thead>
<tr>
<th></th>
<th># manual</th>
<th># automatic</th>
<th>#1:1</th>
<th>Prec</th>
<th>Rec</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development corpus, optical recordings:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule-based</td>
<td>3942</td>
<td>3522</td>
<td>2509</td>
<td>71%</td>
<td>63%</td>
<td>66%</td>
</tr>
<tr>
<td>J48</td>
<td>3942</td>
<td>3411</td>
<td>2217</td>
<td>64%</td>
<td>56%</td>
<td>59%</td>
</tr>
<tr>
<td>Test corpus, machine tools:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule-based</td>
<td>3101</td>
<td>2807</td>
<td>1867</td>
<td>66%</td>
<td>60%</td>
<td>62%</td>
</tr>
<tr>
<td>J48</td>
<td>3101</td>
<td>2606</td>
<td>1639</td>
<td>62%</td>
<td>52%</td>
<td>56%</td>
</tr>
<tr>
<td>Test corpus, optical recordings:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule-based</td>
<td>2781</td>
<td>2535</td>
<td>1664</td>
<td>65%</td>
<td>59%</td>
<td>61%</td>
</tr>
<tr>
<td>J48</td>
<td>2781</td>
<td>2352</td>
<td>1521</td>
<td>64%</td>
<td>54%</td>
<td>58%</td>
</tr>
</tbody>
</table>

Table 2: Evaluation results for segmentation (# = number of segments)
Coreference Resolution Experiments

- Database of 30 claims
- Precision of 83% and Recall of 79%
- Types of Errors
  - Chunking Errors - failed to mark a chunk as an NP
  - NP Modifier - identical NPs used as modifiers inside larger and more complex NPs that aren't the same objects
  - Abstract NPs - abstract nouns are often repeated but do not corefer to each other
  - Partial Match - some NPs may contain another NP inside it but they do not match completely
Clause Structuring Experiments

- Comparing automatic segmentation and coreferencing to manual segmentation and coreferencing and running experiments on both
- Baseline performs clause structuring based on right branching given the number of segments in the gold standard
- Evaluation involves counting the number of identical spans between automatic and manual
- Automatically map each segment of the raw input to its corresponding gold standard segment
- Development: F-score of 64%, Test: F-score of 42%
- This experiment was also useful for determining the most important rules and constraints
Table of Contents

Overview of Patent Summarization
Patent Claim Language
Claim Dependency Structuring
Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Stages for this Step

- **Preprocessing:** Dependency parsing and mapping to Deep Syntactical Structure
- *(Optional)* **DSyntS Summarization**
- **Structure Transfer:** Converts from one language model to another
- **Generation Proper**
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Meaning-Text Theory

- Theoretical linguistic framework, put forward in Moscow by Aleksandr Žolkovskij and Igor Mel’čuk, for the construction of models of natural language.
- Mapping of a semantic meaning to phoentics via intermediate stages.
Semantic Representation

- Web like structure that consists of a network of predication nodes with arrows running from predicate nodes to argument nodes.
- Nodes correspond to lexical and grammatical meanings as these are expressed directly by items in the lexicon or by inflectional means.

![Semantic Representation Diagram](image)
[The media] [harshly criticized the government for its decision to increase income taxes]
Syntactic Representation

- Uses Dependency Trees
- Consists of two levels - Deep Syntactic Structure (DSyntS) and Surface Syntactic Structure (SSyntS)
  - Deep Syntactic - trees represent dependency relations between lexemes and lexical functions. More about relating words directly with verbs. Two sentences may have the same deep syntactic relation but different other structures
  - Surface Syntactic - represents the language-specific syntactic structure of an utterance and includes nodes for all the lexical items
Deep vs Surface Syntactic Representation
Morphological representation

- Strings of morphemes arranged in fixed linear order reflecting the ordering of elements in the actual utterance
- Deep Morphological - consists of strings and morphemes
- Surface Morphological - converts morphemes into the appropriate morphs and performing morphological operations
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Minipar Dependencies vs. Stanford Dependencies
Minipar to Surface Syntactic Structures

- Uses Minipar to generate dependency models
- Develop a set of 137 rules to map Minipar dependencies to SSyntS

(4) Relative clause mapping
How to map Surface Syntactic to Deep Syntactic

- Verbal Tense Auxiliary Forms are mapped onto attribute-value pairs
- Determiners are removed and definiteness/indefiniteness is placed into an attribute
- Governed Prepositions are removed from the structure
- Some lexical units are reduced to abstract lexical labels
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification
  Three Different Simplification Experiments

Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
Deep Syntactic Summarization

- Small set of summarization criteria is based on specific patterns within the DSyntS input
- Takes into account discourse and dependency structure information
- Follows specific rules
Example Rules

- A noun has a postponed attribute:
  - The optical component is a shading member [arranged near the optical axis around the aperture plane of the optical system]

- A definite noun is modified by a full statement:
  - An automatic focusing apparatus comprises the actuator [which controls the focusing means depending upon the output of the phase detector]

- A noun in a dependent claim is modified by a has-part relation (in an independent claim, it can bear important information):
  - A unitary ridge is formed on the top face [having side surfaces constituting the first and second side chip deflector surfaces].

- A noun in a dependent claim is modified by a PURPOSE relation (for + Gerund):
  - The apparatus comprises a lens [for converting the light from the signal plane].
Table of Contents

Overview of Patent Summarization

Patent Claim Language

Claim Dependency Structuring

Simplification

   Three Different Simplification Experiments

Deep Paraphrasing and Summarization

   Sidebar - Meaning-Text Theory
   Preprocessing
   Deep Syntactic Summarization

   Generation Proper

Evaluation of their System
Stages of Generation

- Aggregation of repetitive and isolated fragments of the DSyntS
- Introduction of discourse markers to the SSyntS
- Generation of referring expressions in the SSyntS
- Syntactic Generation
Aggregation

- Fusion of several smaller sentence or phrase structures that share common parts into one structure
- Uses the coreferences determined in the previous part
- Example:
  - 1 - [An optical disk drive] [comprises] [a laser light source].
  - 2 - [An optical disk drive] [comprises] [an optical system].
  - (1+2) - An optical disk drive comprises a laser light source and an optical system.
Discourse Marker Insertion

- Some rules add discourse markers to the top verb of the SSyntS.
- Depending on the discourse relation to which is introduced, the marker can be retrieved from a discourse-marker dictionary.
- Examples of Discourse Markers: For, By, When.
Referring Expression Generation

- Bunch of rules that the program uses to try to make a more understandable summary/paraphrase
- Example: Introducing a relative pronoun as subject of a sentence if it is coreferring with the object of the previous sentence.
  - The phrase to become the relative clause must not be syntactically “too heavy” (i.e., contain “too many” nodes); if it is, the introduction of a deictic is preferred
  - The rule does not apply either if the object of the first sentence has undergone aggregation so as not to rebuild sentences that would be very long and potentially ambiguous
Syntactic Generation

- Stage responsible for completing the generation
- Realizes agreement, word order, and other surface morphological structures to convert the deep syntactical structures back to regular sentences.
Table of Contents

Overview of Patent Summarization
Patent Claim Language
Claim Dependency Structuring
Simplification
  Three Different Simplification Experiments
Deep Paraphrasing and Summarization
  Sidebar - Meaning-Text Theory
  Preprocessing
  Deep Syntactic Summarization
  Generation Proper

Evaluation of their System
System Setup

- Uses a software package ROUGE to automatically evaluate quality of summarizations
- Chose 30 patents at random to be summarized using their method as well as with Microsoft Word’s automatic summarizer
- Also ran experiments with Multilingual patents to determine how well translation work - native speakers rate translations given by their software versus Google Translate
Results and Explanations

- For summarization, their system achieved an F-Score of 61% while MS Word achieved 43%.

- Low score can be partially explained by the object/method dichotomy in some patent claims, which cannot reliably be identified in an automatic way. If a patent claim section contains claims referring to both the invented object and the method of applying this object, both kinds of claims tend to contain largely the same information.

- Multilingual Results - no meaningful information is lost during simplification.

<table>
<thead>
<tr>
<th></th>
<th>baseline</th>
<th>regeneration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligibility</td>
<td>49</td>
<td>58</td>
</tr>
<tr>
<td>Simplicity</td>
<td>49</td>
<td>74</td>
</tr>
<tr>
<td>Accuracy</td>
<td>47</td>
<td>51</td>
</tr>
</tbody>
</table>
Example Summarization

An optical disk drive comprising: a laser light source for emitting a laser beam; an optical system for conversing the laser beam from the laser light source on a signal plane of optical disk on which signal marks are formed and for transmitting the light reflected from the signal plane; one or more optical components arranged in the optical path between the laser light source and the optical disk for making the distribution of the laser beam converged by the conversing means located on a ring belt just after the passage of an aperture plane of the optical system; a detection means for detecting the light reflected from the optical disk; and a signal processing circuit for generating a secondary differential signal by differentiating the signals detected by the detection means and for detecting the edge positions of the signal marks by comparing the secondary differential signal with a detection level.
An optical disk drive comprises a laser light source, an optical system, a detection means, and a signal processing circuit. The laser light source emits a laser beam. The optical system converses the laser beam from the laser light source on a signal plane of optical disk and transmits the light. The optical disk drive also comprises one or more optical components. It is arranged in the optical path between the laser light source and the optical disk. The detection means detects the light. The signal processing circuit generates a secondary differential signal and detects the edge positions of the signal mark.