Automatic Semantic Annotation for the Easification of Action Rule Legislative Sentences for Specialist Readers

Sherry Maynard
The University of the West Indies, Cave Hill Campus, Cave Hill, St. Michael, Barbados

Abstract
This research has applied automatic semantic annotation to a text easification solution that aids non-legal experts in reading legislation as part of their work. It annotates the modality, actor, action, case and condition concepts within action rule legislative sentences. The research first analyzes the lexical and syntactic compositions of a corpus of legislation commonly read by a group of compliance professionals and then extracts data sets of action rule legislative sentences for annotation. The annotation is rule-based, fully automated and utilizes Tregex patterns and Tsurgeon operations. The resultant easified legislative sentences were confirmed by legal experts as having preserved the semantic integrity of the original sentences. In addition, the professionals who participated in the research, reported lower intrinsic and extraneous cognitive loads when they read the easified version of the legislative sentence, when compared to the loads experienced when they read the original version of the same sentence.

Keywords
Easification, semantic annotation, specialist readers, cognitive load, intrinsic load, extraneous load,

1. Introduction

This research fully automates the semantic annotation of five concepts found in action-rule legislative sentences. These concepts include modality, actor, action, case and condition. The semantic annotation is part of a larger goal of easifying the legislative sentences to aid the comprehension of specialist readers, i.e. non-legal experts reading legislation as part of their work. Specialist readers may include professionals in areas such as compliance, audit, finance, risk, information security, human resources and health and safety.

It has long been acknowledged that legal language is complex both in its construction of and the expression of its ideas. Syntactic contributors to this complexity include the density of prepositional phrases, the high degree of subordination, syntactic discontinuity and lengthy sentences [1-5]. In addition, the language is characterized by technical vocabulary, wordiness, repetition, nominalization and the excessive use of binomial and multinomial expressions [2, 4, 6, 7].

Even legal experts resort to reading the explanatory notes that accompany a bill rather than the legislative text itself [8, 9]. Similarly, some legislators and government officials have confessed that they do not understand much of the bills they vote on [10]. Nonetheless, organisations aiming to reduce cost and looking for skills beyond legal expertise, are seeking persons with investigative, audit and critical thinking skills to have primary responsibility for the legal compliance function within their organizations [11-13]. Hence, persons with training in organizational behavior, finance, accounting and information systems are being regarded as ideal candidates for this critical responsibility [14]. The legal compliance function is an important part of modern businesses

Proceedings of the Fifth Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL 2021), June 25, 2021, São Paulo, Brazil.
EMAIL: *sherry.maynard@cavehill.uwi.edu*

as they navigate aggressive regulatory environments, unconstrained by geographical boundaries [15], and while the cost of legal compliance is high, the cost of non-compliance is approximately three times higher [16].

2. The Corpus Analysis

The Barbados legislation that formed the corpus analyzed in this research are those commonly read by forty-five members of a compliance professional association in Barbados. Seventy four percent of these participants have no legal training and eighty-four percent experience challenges reading legislation. The challenges reported mirrored those associated with the syntactic and lexical features of legal language as outlined in the introduction. The Flesch reading ease scores of these commonly read Barbados legislations range from 28.1 – 36.6, i.e. they are difficult to very difficult to read [17]. The upcoming sections detail the syntactic and lexical features of the corpus.

2.1. Syntactic & Lexical Features

The corpus analyzed is composed of the following Barbados legislation:

- Exempt Insurance Act, 1983
- Companies Act, 1985
- Proceeds of Crime, 1990
- International Business Companies, 1992
- Financial Institutions Act, 1997
- International Financial Services Act, 2002
- Anti-Terrorism Act, 2002
- Money Laundering and Financing of Terrorism (Prevention and Control) Act, 2010
- Financial Services Commission, 2010

Overall, the corpus contains 192155 tokens and 3306 sentences. This size is sufficiently large because the conservative nature of legal discourse does not necessitate a large corpus to determine its linguistic features. Bhatia (1983) identified linguistic patterns in legislative text based on a single British Parliament act; these findings were later confirmed when similar experiments were repeated on larger corpuses of European, Hong Kong and Chinese legislative texts [18, 19].

The average sentence length of the legislation in the corpus range from 39 – 66 words, with the overall average of the corpus being 53 words. This average sentence length significantly exceeds Curtotti et al. (2015) recommendation of keeping legislative sentence lengths below 30 words [20]. Furthermore, it is more than double the average sentence length for English academic articles (26 words) [21] and the recommended length for general text of 15–20 words [22]. Sentence length in legislative writing, could be considered a secondary matter when compared to the benefit gained from having as much related ideas together in a single sentence to mitigate against taking the law out of context [23-25].

The corpus has on average three coordinating conjunctions per sentence. In calculating the usage of the coordinating conjunctions, detection rules were created to identify when ‘and’ / ‘or’ were used in binomial or multinomial expressions; these usages were deducted from the total conjunctions prior to calculating the ratio of coordinating conjunction per sentence. Therefore, the average represents phrasal or clausal conjoining. In the corpus, ‘or’, ‘and’ and ‘for’ are the primary conjunctions used, 46.58%, 27.16% and 20.62% respectively. On the contrary, the coordinating conjunction ‘but’ that marks contrast had only 2.16% presence in the corpus. Similarly, ‘nor’ and ‘so’ had only 3.10% and 0.38% usage respectively; ‘yet’ had no occurrences within the corpus.

In addition, the corpus had on average two subordinating conjunctions per sentence. Relative clauses are heavily used in the corpus, with relative pronouns making up 53.69% of the total subordinating conjunctions identified. As with coordinating conjunctions, contrast-type subordinating conjunctions (*e.g. while, whereas*) are seldom used within the corpus; they make up 0.06% of the total subordinating conjunctions. In addition, there is one occurrence of the similarity type conjunctions i.e. the term ‘likewise’.

Curtotti et al (2015) suggested, for improved readability of legislative text, to avoid using more than two conjunctions per sentence [20]. The multiple conjunctions create complex sentence structures and syntactic discontinuities that can make sentences difficult to read and understand. However, for every negative impact a given linguistic feature has on the readability of the legislative text there are corresponding benefits for the legal domain. For example, while the intensive use of conjunctions can result in cognitive overload for some readers, they usage serves the legal goals of precision and all-inclusiveness [18, 26, 27]. Achieving these goals could mean compacting all relevant information into a single, long, complex sentence that aids in minimizing the possibility of loopholes and evasions in the law [18, 28, 29]

A sample of 208 sentences (*45 – 115 words*) was extracted from the corpus and their dependency distance metric calculated. This metric can be used as an indicator of comprehension difficulty and has implication for the utilization of readers' working memory capacities. A recommended threshold is less than 3 words [30]. The average dependency distance metric of the sample sentences is 4 words; the lowest being 2 words and the highest 9 words. Therefore, on average four words separate two elements that share a syntactic relationship, which would typically reside alongside each other in the sentence structure.

Finally, the use of Latin and Old English terms in the corpus was assessed. The most commonly used archaic terms are "thereof", "forthwith", "thereby" and "thereafter"; i.e. 98, 61, 26 and 22 occurrences respectively. The most commonly used Latin term was "mutatis mutandis", which is used 12 times. However, overall the use of Old English and Latin words in the corpus is miniscule: 243 Old English words and 30 Latin words. In a corpus of 192155 words, these usages average less than zero for a term-to-sentence ratio. This lexical occurrences support the findings of a study by Dell'Orletta (2012) which showed no significant differences in the lexicon of a set of EU legislation and the stories from the Wall Street Journal. On the contrary, there was a noticeable difference in the underlying syntactic structure of the writings in the two domains [31].

3. The Semantic Annotation of Legal Concepts

The concepts annotated for the easification of action rule legislative sentences are defined in table 1 below. The concepts were adopted from Coode (1845) specification of the essential and optional elements of action rule legislative sentences [32].

Table 1: Concept Definitions

CONCEPT	DEFINITIONS
Modality	The auxiliary representing the action's modality
Actor	The person or class of persons performing or prohibited from performing a legal action
Action	The rights, privileges, powers, obligations or liabilities
Case	The circumstances / occasions in which the legal action applies
Condition	The prerequisites that must occur before the legal action becomes operable

The semantic annotations are rule based and utilize Tregex patterns and Tsurgeon operations [33]. They are fully automated and require no human intervention in pre-processing the sentences. The Stanford CoreNLP [34] pipeline was used to perform the typical NLP pre-processing tasks of tokenization, sentence segmentation, part of speech tagging and constituency parsing. The output of the parsed tree is the primary basis for the annotation rules. Nine Tregex pattern – Tsurgeon operation pairs were created to detect the five semantic concepts defined in table 1 above. The upcoming sections provide an overview of the Tregex rules specified in table 2 below.

Table 2: Rule Specification

<table>
<tr><th>CONCEPT</th><th>TREGEX PATTERN</th></tr>
<tr><td>Modality</td><td>MD = modal >> (VP > S) | >> (VP > (S > ROOT)) | >> (VP > (S >> SINV)) | >> (VP > SBAR)</td></tr>
<tr><td>Actor</td><td>NP = actor $++ (VP << MODAL) & ((> (S > ROOT) | > (S > (S > ROOT)) | > (S > (SINV > ROOT)) | > (S > (S $- (CC <, /and/ | <,/ but/))) | $+ (ADVP $+ (VP > S))))</td></tr>
<tr><td>Action</td><td>VP = action $-- (ACTOR.. MODAL) & (> (S > ROOT) | > (S > (S > ROOT)) | > (S > (SINV > ROOT)))</td></tr>
<tr><td rowspan="2">Case</td><td>SBAR = case < (SBAR < (WHADVP < (WRB < /Whe*/))) | < (WHADVP < (WRB < /Whe*/))</td></tr>
<tr><td>/,/ =comma $- CASE = case</td></tr>
<tr><td rowspan="4">Condition</td><td>ADVP = condition > S & < (RB < /^[A-Z][a-z]+/)</td></tr>
<tr><td>PP = condition (> S | > SINV) & < (IN < /^[A-Z][a-z]+/)</td></tr>
<tr><td>SBAR = condition < (IN < /^[A-Z][a-z]+/)</td></tr>
<tr><td>/,/ =comma $- CONDITION = condition</td></tr>
</table>

3.1. The Modality Concept

The first rule searches for modal auxiliaries within the sentence, primarily those at higher levels within the tree structure. The rule however is deliberately wide reaching to ensure that it captures the correct modal auxiliary needed for the identification of the 'Actor' and 'Action' concepts in subsequent rules. Generally, the targeted modal auxiliary is sandwiched between the 'Actor' and 'Action' sub-trees. The annotation rule identifies a modal auxiliary which is dominated by a verb phrase (*VP*). The verb phrase (VP) is in turn immediately dominated by either a declarative clause or a subordinate clause that is immediately dominated by the root of the parsed tree.

3.2. The Actor Concept

The actor rule detects the noun phrase that acts as the subject in the English language sentence structure. Therefore, it is a node that must be immediately dominated by nodes that are at high levels within the parse tree, i.e. clauses immediately dominated by the root node. The actor noun phrase (NP) is the left sister of the verb phrase (VP) that dominates the modal auxiliary detected in the modality rule. In addition, the rule accommodates instances where the connection between the NP and the VP is interrupted by an adverbial phrase and makes provisions for complex sentences joined by coordinating conjunctions, in which case the conjunction node acts as the head of the embedded sentence.

3.3. The Action Concept

The legal action within the legislative sentence is a verb phrase (VP) who is the right sister of the sub-tree that represents the 'Actor' concept" and which precedes the 'Modality' concept. The 'Action' verb phrase represents the predicate of the sentence and is therefore immediately dominated by high-level nodes in the sentence tree that have direct connections to the root node. The annotation rules covered to this point are the core or mandatory concepts in the action-rule legislative sentences.

3.4. The Case & Condition Concepts

The case rule captures the Wh-clauses in the initial sentence position, which typically represent the case concept. These clauses are subordinate clauses that immediately dominates a 'Wh-adverbial phrase, which in turn dominates a 'Wh-adverb' that begins with a upper case 'W' followed by a lower case 'h' and 'e' and then by any other characters. This regular expression detects clauses beginning with terms such as 'Where', 'When', 'Whence' and extensions such as 'Whenever'.

The condition rule identifies adverbial and prepositional phrases that are immediately dominated by a declarative clause and immediately dominates an adverb or a preposition respectively. In most instances, the case and condition clauses end with a comma. An additional rule searches for this comma and relocates it inside the case and condition sub-trees. The goal is to ensure that during the easification process an orphan comma is not left behind.

4. Related Works

Boella et al. (2013) implemented a legal concept detection mechanism using a Support Vector Machine binary classifier. They utilized syntactic dependencies to build triplets to train three classifiers to categorize the concepts of active roles, passive roles and objects [35]. They used the Italian TULE parser to create the dependency information for the legislative text [36]. The results of their approach showed high precision and recall for the detection of the active role (*precision 97.2% and recall 92.6%*), moderate performance for the passive role (*precision 100% and recall 26.8%*), and low performance for the object role (*precision 59.3% and recall 31.9%*). These results were negatively affected by the accuracy of the POS tagger and the syntactic parser. For instance, when the POS tagger did not recognized a noun, it missed an eligible word for a semantic label and the dependency parser could incorrectly label the semantic relations associated with that term [35]. One of the reasons given for the use of the machine learning classifier was to overcome the need for the sequential execution typically associated with pattern-matching rules.

Sleimi et al. (2018) utilized the traditional ordered set of pattern matching rules to detect a collection of legal concepts and attained high performance across the varying concepts [37]. The purpose for the annotation in this work is to support legal requirements engineering. Sleimi et al. (2018) used Tregex patterns to extract ten main phrase level concepts from constituency and dependency parsed trees. They established a set of markers for each concept type based on dictionaries and ontologies. These markers formed part of the pattern matching rules. For example, one of the patterns for the "Actor" rule (*subject dependency and NP < actor marker*) was represented as a noun phrase in the subject dependency position and one that immediately dominates a term from the list of actor markers. The accuracy of Sleimi et al. (2018) rule detections had overall precision and recall measures of 87.4% and 85.5% respectively using 200 statements from Luxembourg traffic laws [37].

The level of accuracy attained in the work of Sleimi et al. (2018) may result in part to the use of predefined terms within the relevant concept repositories. While this approach simplifies the rule construction, it requires human pre-processing to identify the terms that represent the markers for each concept. This technique was utilized in other tools such as, the Gaius T, [38] and the NomosT, [39]. It however has some drawbacks, for instance, where the repositories are inadequately defined, the performance of the detection rules will be negatively affected. In addition, new markers will need to be added to extend the detection capabilities of the annotation rules beyond the initial legislative domain. It is important to note that the work of Sleimi et al. (2018) also suffered challenges associated with the performance of the parser as with the work of Boella et al. (2013). Much of Sleimi et al. (2018) detection errors occurred from the constituency parser's inaccurate attachments of subordination, coordination and prepositional phrases and hence causing the dependency parser to infer incorrect dependency relationships amongst the nodes [37].

5. Research Experiment

The semantic annotations were done at a sentence level using three data sets containing action rule sentences that met the following criteria:

- contiguous and complete;
- a single legal action
- simple, complex & compound structures;
- a single or compound subject;
- at least one modal auxiliary in the upper level of the sentence tree;
- 40 or more words;
- dependency distance metric of 3 or more;

Contiguous and complete sentences are those with a non-bulleted format that end with a full stop and not a semicolon. The selective nature of the sentences in the experiment were driven primarily by the easification methodology utilized in the next stage of the experiment and the limitations of using a constituency parser not trained on legislative text.

A hundred development sentences (Dev-Set) were extracted from a set of Barbados intellectual property legislation and annotated by the author. These were used to iteratively test the annotation rules during construction. These legislation included:

- Trademark Act, 1985
- Patent Act, 2001
- Industrial Designs, 1981
- Copyright Act, 1998
- Telecommunications Act, 2001

An assessment of the syntactic composition of the intellectual property legislations was done and compared against those read by the research participants to ensure a degree of compatibility. The use of development sentences from a comparable but different legislative domain from those read by the participants was to ensure that the algorithm only processes sentences from the participants' domain after the rule development was frozen. Two test sets were extracted for the purpose of testing the performance of the annotation rules.

The first test set (Test Set A) contained one hundred and twenty-one sentences extracted from the legislation read by the participants. These legislation were primarily from the financial services sector. The average sentence length for Test-set A was 63 words and the average dependency distance metric was four. The author annotated Test-set A to provide a gold standard to assessment the performance of the annotation rules.

The second test set (Test-Set B) consisted of sixty-three sentences extracted from the Barbados Road Traffic Act 1981. The average sentence length for Test-set B was 60 words and the average dependency distance metric was four. Two legal experts independently annotated these sentences. The author was guided by the annotation procedures recommended by Hovy and Lvid (2010) [40], these included:

- The provision of guidelines that define the concepts and the method of highlighting each concept within the data set;
- Giving the annotators practice sentences to ensure the annotation process is understood and the instructions are clear;
- Using annotators with reasonably similar levels of education;
- A minimum use of two annotators and have them act independently;
- In the absence of a third adjudicator annotator, any sentences where the annotations differ should be discarded;

The annotators were two lawyers with equivalent educational training. They used the text highlight feature in Microsoft Office Word to highlight each concept using a specified color scheme. As a way of improving the speed and reliability of the annotations, the legal experts were instructed to annotate one concept at a time across all the sentences; for example, the first round of annotations highlights the actor concepts only, the second round the actions etc. [40]. Since two annotators were used in the experiment, the thirteen sentences where their annotations differed were deleted from the test set. Hence 50 sentences remained in Test-Set B, which represents a 79% agreement between the annotators. In addition, to maximize the limited time of the legal experts, a trade off was made where the experts annotated all of the mandatory concepts and the case concept; the optional condition concept was not annotated. The legal experts did not engaged the author during the annotation process.

5.1. Results of the Annotations

The precision, recall and F measures were computed for the development and the two test sets. Both lenient and strict computations were performed; the lenient computation assigned 0.5 points to partial annotations, while the strict computations assigned no points to partial detections, hence treating them as missed annotations. The measures were done using GATE Developer 8.0 [41]. Based on the application of the semantic annotation to the easification of sentences within the business context, the partial detections are unacceptable therefore only the strict computations were used. Table 3 below shows the results of the annotation rules using the Dev-set.

Table 3: Annotation Results for Dev-Set

CONCEPT	Truth	Extracted	Perfect Match	Precision %	Recall %	F Measure %
Modality	118	141	118	83.7	100	91.1
Actor	116	103	94	97.9	82.5	89.5
Action	117	98	92	98.9	79.3	88.0
Case	34	33	27	100	79.4	88.5
Condition	20	17	17	100	85.0	91.9
Overall	405	392	348	93	86.6	89.7

The rules detected 392 annotations from the development set. Of these 348 or 86% were perfect matches and 57 were missed or partially detected annotation (14%). Annotations were missed either because of the wrong text or no text being detected for a given concept.

Once the rule construction was frozen, the performance of the semantic annotation rules was tested using Test-Set A and Test-Set B. The algorithm had not seen any of the sentences in these test sets prior to the computation of the results shown in table 4 and 5 below.

Table 4: Annotation Results for Test-Set A

CONCEPT	Truth	Extracted	Perfect Match	Precision %	Recall %	F Measure %
Modality	142	159	142	89.3	100	94.4
Actor	141	134	129	97.0	91.5	94.2
Action	142	131	124	100	87.3	93.2
Case	47	47	41	100	87.2	93.2
Condition	34	30	28	100	82.4	90.2
Overall	506	501	464	95.7	91.7	93.6

Table 4 shows the detection results for Test-set A; of the 501 annotations detected, 464 were perfect match, i.e. 92%; 42 were missed or partially detected (8%). As expected, based on the strategy discussed earlier, the results for the modality concept showed a 100% recall. The recall for the condition concept was the lowest at 82.4%. Alternately, there were 100% precision results for the action, case and condition concepts. The F measures for all the concepts were above ninety, with the overall precision, recall and F measures being 95.7, 91.7 and 93.6 percentage respectively. These overall percentages are not averages of the individual concept measures, but rather computations based on the detection totals across the concepts.

The results presented so far, have been compared against truths annotated by the author. The results for Test Set B are compared against truths annotated by the two legal experts participating in the research; these are shown in table 5 below.

Table 5: Annotation Results for Test-Set B

CONCEPT	Truth	Extracted	Perfect Match	Precision %	Recall %	F Measure %
Modality	50	60	50	83.3	100	90.9
Actor	51	44	44	100	86.3	92.6
Action	51	41	41	100	80.4	89.1
Case	21	19	19	100	90.5	95.0
Overall	173	164	154	93.9	89.0	91.4

Of the 173 annotations detected, 154 were perfect match, i.e. 89%; 19 were missed or partially detected (11%). The performance results on Test-set B are comparable with those on the Test-set A. The overall precision was 93.9%; a 100% recall measure for the modality concept and the 'case' concept had a recall of 90.5%. The overall F-measure was 91.4%.

6. Discussion

Generally, the detection results of the semantic annotations were good, with values of 83 – 100 % for precision, 80 – 100% for recall and 89 – 94% for the F measure. To ensure the annotations were fully automatic and hence eliminating the human pre-processing, the implementation deviated from the use of concept markers utilized in tools such as, the Gaius T, [38], NomosT, [39] and the tool by Sleimi et al (2018) [37]. This made the detection rules more complicated but allows for scalability and applicability across multiple legislations in varying domains. As illustrated in the data sets, the annotation rules detection capabilities spanned the intellectual property, financial services and road traffic legislations.

The detection rules for the three mandatory components of the action rule legislative sentences have a high degree of dependence. Hence the risk of an initial failure in detecting the modality concept can be transferred into failed actor and action detections. To mitigate this drawback, the modal detection rule was designed to be all-inclusive in nature and in all the test sets had a 100% recall results.

The automated detection rules used in this research suffered from similar parser related difficulties experienced in other works [35, 37, 42, 43]. In the case of the Stanford constituency parser, while the support website recommended the most up-to-date version of the parser for the best performance, that recommendation did not hold true for the legislative text used in this study. The researcher found that the older probabilistic context free grammar parser generated less parsing errors than the newer shift-reduce parser.

The increase in the parsing errors was directly linked to the increase in the complexity in the sentence structures. Repeated errors occurred when the subject of the sentence had one or more embedded qualifiers, when prepositional phrases broke the continuity between the modal auxiliary and the main verb, and where compound sentences contained 'or' conjunctions. In addition, some sentences were tagged as fragments if the typical English sentence structure (*subject-verb-object*) was not detected. Another interesting parsing error occurred when the term 'issue' used in the context "shall issue to the applicant" was tagged as a noun instead of a verb. This miss tagging of the word 'issue' reflected the part-of-speech tagger's interpretation of 'issue' as a topic or problem, instead of the act of distributing something. This error is likely rooted in the differences in the genre of the material used in the training the part of speech tagger when compared to legislative text.

While the current work showed the applicability of the annotation rules across legislation in different domains, an expanded scope of the action rule sentences would further test the generalizability of the annotation rules. Therefore, future work includes the utilizing larger, more diverse datasets to test the annotation rules. However this will also necessitate the employment of techniques to overcome the limitations of the part of speech and constituency parsers.

7. The Semantic Annotation Applied to Easification

The semantic annotation of the legal concepts was a necessary step in the easification process. The diagram in figure 1 below shows how the semantic annotation fitted into the overall algorithm design. It added computer readable intelligence to the legislative sentence to facilitate the automation of the clarifying cognitive structuring easification device.

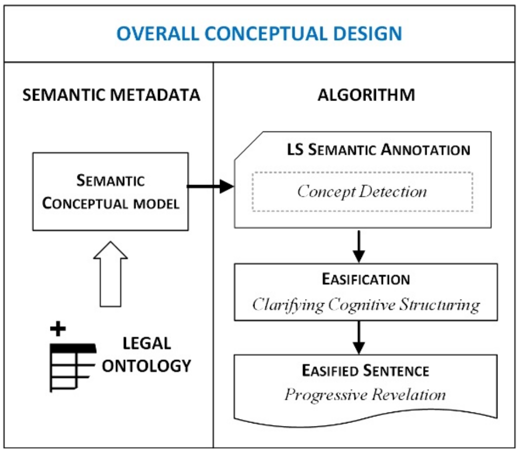

Figure 1: Semantic Annotation applied to Easification

The easification of legislative sentences is a viable alternative to text simplification and is suitable for specialist readers. Unlike text simplification, it focuses less on modifying the text and more on aiding the mental processes of the readers to facilitate the intake of the idea. Consequently, easification evades a major risk of text simplification, that of inadvertently altering the meaning of the legislative text. This shift in emphasis from the text to the reader increased the likelihood of easification preserving the semantic integrity of the legislative text.

The easification device, clarifying cognitive structuring makes the components, the structure and relationships of the action rule legislative sentences more apparent to specialist readers. It draws on cognitive load theory (CLT) [44], which offers insights into the consumption of working memory resources during task performance and learning. CLT is built on the following basic ideas about the human cognitive architecture (HCA) [45, 46]:

- HCA has a very limited working memory storage mechanism and a very large long-term memory storage facility;
- The demands on working memory occurs from conscious cognitive activities;
- Schematic structures are utilized to store information in long-term memory;

Cognitive load is the demand placed on the storage and processing resources of working memory. When the mental demands of the activities in working memory, at a given instance, exceed an individual's cognitive capacity, the individual experiences cognitive overload [45, 47]. Miller (1956) estimated that working memory stores approximately, 7 (+/- 2) amount of active information chunks, which decay within 15 – 30 seconds if not actively rehearsed [48]. Other researchers suggested a more precise capacity might be 3 - 5 chunks during information processing [49].

These working memory constraints have implications for sentence processing and comprehension. The capacity theory asserts that sentence parsing and memory processes compete for the same pool of resources. Therefore, if sentence processing demands a substantial amount of resources, the resources dedicated to storage would be reassigned to meet the processing demand; the resultant reduction in storage capacity can lead to forgetting part of the sentence; i.e. forgetting by displacement [50]. The longer and more syntactically complex the sentence, the more likely readers will lose track of the structural development of the idea [18]. This can occur when some of the components succumb to working memory decay before integration into the structure being built [51]. Typically, readers are unaware of the intricate resource allocations in working memory until they reach near full capacity and the resultant trade-offs in working memory distribution starts to occur [52].

For the purpose of this research two types of cognitive loads were measured, intrinsic load and extraneous load. The intrinsic load (IL) is the innate complexity of the information or task. This complexity is determined by element interactivity, which is the degree of interconnectivity amongst elements that necessitates them being processed simultaneous. Intrinsic load is essential for comprehension [47, 53-57]. The extraneous load (EL) is induced by the way information is presented and organized. It is considered the 'bad' load because it results in cognitive processing that is unrelated to learning and could impede learning. EL occurs when there is high element interactivity and suboptimal communication. The aim is to minimized extraneous load [58, 59].

7.1. Results of the Application to Easification

The easification algorithm performs the following functions utilizing the semantic annotations along with additional annotations. It searches and extracts the semantic annotated elements; annotates additional lower stratum elements, extracts the main legislative idea, inserts logic indicators and generates output formats for the readers.

Take for example section 48 (2) of the Barbados Securities Act 2002 as shown below:

> "Where a broker is charged with an offence involving fraud or dishonesty or where it is alleged that he has defaulted in the payment of moneys due to a self-regulatory organisation or to any other market actor, the Commission may, if it considers that it is in the public interest to do so, suspend the registration of the broker pending the final determination of the charge or allegation." [60]

This legislative sentence has 68 words and a dependence distance metric of 4.75. The easification algorithm generates the two outputs in figure 2 and 3 from the input sentence above.

The Main legal Idea

...the **Commission *may*** suspend the registration of the broker pending the final determination of the charge or allegation.

Figure 2: The Main Idea of Securities Act 2002 318A, s48 (2)

The main legislative idea shown in figure 2, consist of 18 words; approximately 74% less than the amount of words in the full sentence (68 words). In addition, the complexity of the sentence has been reduced in this transient phase of the sentence processing. The aim is to give the reader the opportunity to create a mental frame of the legislative idea prior to processing the details. The output in figure 3 below, adds the details with informative component labels and the If-Then construct that makes the cause and effect relationship more obvious.

The Full Legislative Idea

Circumstances where the legal Action applies	IF a broker is charged with an offence involving fraud or dishonesty or where it is alleged that he has defaulted in the payment of moneys due to a self-regulatory organisation or to any other market actor,
Legal Actor & nature of the legal Action	THEN the **Commission *may*,**
The legal Action	suspend the registration of the broker pending the final determination of the charge or allegation.

Figure 3: The Easified version of the Securities Act 2002 318A, s48 (2)

The output illustrated in figure 3 utilizes the following If-Then format proposed by Langton (2005) as an extension to the initial easification device [61]:

IF case(s)
IF condition(s), sub-condition(s)
THEN legal actor(s) ***modal***
legal action(s)

Four lawyers were asked to evaluate the similarity in the semantics of four pairs of action rule legislative sentences; the original-unmodified version and the corresponding easified version. There was an overarching agreement amongst the lawyers that the meanings of the original legislative sentences were retained in the easified versions.

An additional experiment was also conducted to identify the impact of the easified legislative sentence on the cognitive load of sixty-three professionals that participated in this part of the experiment. A modified version of Leppink, Pass et al (2013) cognitive load measurement instrument was used to capture the perceived intrinsic and extraneous load of the participants [62]. Confirmatory Factor Analysis was performed on the modified measurement instrument and it was found to be valid, reliable and the data collected showed good model fit. In the experiment, the control group was given the original version of the legislative sentence and the experimental group was given the easified version of the same legislative sentence. An independent sample t-test showed that the lower means for the intrinsic and extraneous loads of the experimental group, when compared to the control group were statistically significant.

Presenting the research participants with the main idea first, temporarily reduced the element interactivity of the legislative sentence. In addition, the use of progressive revelation allowed the participants to add the details incrementally, at their own pace; this further assisted them in managing their intrinsic load. The mean of the intrinsic load, of the experimental group was 3.33 and the control group is 4.57, with a statistically significant p value of .01038 and a 95% confidence interval. Similarly, the mean extraneous load of the experimental group was 4.16 and the control group was 5.43 and was statistically significant with a p value of .021 at a confidence interval of 95%.

8. Conclusion

This research assessed the lexical and syntactic composition of a corpus of Barbados legislation read by compliance professionals working in Barbados. This research bridged a gap, and developed a solution for specialist readers working in the business context where preserving the semantic integrity of the legislative text is critical to legal compliance. An algorithm was successfully developed to easify action rule legislative sentences. This included creating several semantic annotation rules to detect key legal concepts without requiring any human pre-processing of the text. The algorithm outputted an easified legislative sentence with multiple perspectives of the legislative idea. The easification of the action rule legislative sentence proved effective in lowering the intrinsic and extraneous loads of the specialist readers in the research sample, without compromising the semantic integrity of the legislative sentence. Future work will seek to expand the sample size of the participants and to explore the impact of informed ratings in the cognitive load tests.

9. References

[1] E. Mattiello, Nominalization in English and Italian Normative Legal Texts. SESP Across Cultures, 2010. 7 129 - 146.

[2] P.M. Tiersma, Language of Legal Texts, in Encyclopedia of Language & Linguistics, B. Keith, Editor. 2006, Elsevier: Oxford. pp. 549-556.

[3] R.P. Charrow, V.R. Charrow, Making Legal Language Understandable: A Psycholinguistic Study of Jury Instructions. Columbia Law Review, 1979. 79(7) 1306-1374.

[4] P.R. Macleod, Latin in Legal Writing: An Inquiry into the Use of Latin in the Modern Legal World. 39 B.C.L. Rev. 235, 1998.

[5] J. Crandall, V.R. Charrow, Linguistic Aspects of Legal Language. 1990.

[6] R. Hyland, A Defense of Legal Writing. University of Pennsylvania Law Review, 1986. 134(3) 599-626.

[7] D. Mellinkoff, The Language of the Law. 1963, Eugene, OR: Resource Publications. 526.

[8] J. Sheridan, Legislation.gov.uk and Good Law. Civil Service Quarterly, 2014.

[9] R. Heaton, When Laws Become Too Complex - A review into the causes of complex legislation. 2013.

[10] B.C. Jones, Don't Be Silly: Lawmakers 'Rarely' Read Legislation and Oftentimes Don't Understand It . . . But That's Okay. Penn State Law Review, Penn Statim, 2013. 118(7) 7 - 21.

[11] Deloitte, The changing role of compliance officers. 2014.

[12] Ernst & Young, Compliance seeks a path to regulatory readiness, in Insurace CCO survey. 2014, Ernst & Young Global: London.

[13] J.A. Tabuena, The Chief Compliance Officer vs the General Counsel: Friend or foe?, in Compliance & Ethics Magazine. 2006, Society of Corporate Compliance and Ethics: Minneapolis, MN. pp. 4-7 & 10-15.

[14] A. Gross-Schaefer, C.A. Cueto, Ethics & Compliance: The Game-Changer in the Business World. International Journal of Business and Social Science, 2017. 8(2) 57 - 65.

[15] Sovos, The State of Regulatory Compliance. 2017.

[16] Ponemon Institute The True Cost of Compliance with Data Protection Regulations. 2017.

[17] R. Flesch and A.J. Gould, The art of readable writing. 1949.

[18] V.K. Bhatia, Simplification v. Easification—The Case of Legal Texts. Applied Linguistics, 1983. 4(1) 42-54.

[19] V.K. Bhatia, N.M. Langton, J. Lung, Legal Discourse: Opportunities and treats for corpus linguistics, in Discourses in the Professions: Perspectives from Corpus Linguistics, U. Connor and T.A. Upton, Editors. 2004, John Benjamins Netherlands. pp. 203-232.

[20] M. Curtotti, E. McCreath, T. Bruce, S. Frug, W. Waibel, C. Nicolas, Machine learning for readability of legislative sentences, in Proceedings of the 15th International Conference on Artificial Intelligence and Law. 2015, Association for Computing Machinery: San Diego, California. pp. 53–62.
[21] G. Zhang, H. Liu, A Quantitative Analysis of English Variants Based on Dependency Treebanks. Glottometrics, 2019. 44 16-33.
[22] M. Cutts, Oxford Guide to Plain English. 2013, URL: http://site.ebrary.com/id/10775452.
[23] C. Williams, Legal English and Plain Language: an introduction. ESP Across Cultures, 2004. 1 111-124.
[24] E. Tanner, Sanctity of the Single Legal Rule/Single Sentence Structure. Monash U. L. Rev., 2000. 26.
[25] C. Renton, Renton Commitee Report on legislation. 1975.
[26] V.K. Bhatia, An Investigation into formal and functional characteristics of qualifications in legislative writing and its application to English for Academic legal purposes. 1982, Univ. of Aston in Birmingham.
[27] V. Bhatia, J. Engberg, M. Gotti, D. Heller, Introduction, in Vagueness in Normative Texts - Linguistic Insights Studies in Language and Communication, V. Bhatia, et al., Editors. 2005, Peter Lang International Academic Publishers pp. 9-21.
[28] H.E.S. Mattila, Comparative Legal Linguistics 2006: Ashgate Pub Co. 347.
[29] C. Frade, Legal Multinomials: Recovering Possible Meanings from Vague Tags, in Vagueness in Normative Texts Linguistic Insights Studies in Language and Communication, V. Bhatia, et al., Editors. 2005, Peter Lang International Academic Publishers pp. 133-153.
[30] H. Liu, Dependency distance as a metric of language comprehension difficulty. Journal of Cognitive Science, 2008. 9(2) 159 - 191.
[31] F. Dell'Orletta, S. Marchi, S. Montemagni, B. Plank, G. Venturi, The SPLeT–2012 Shared Task on Dependency Parsing of Legal Texts, in 4th Workshop on "Semantic Processing of Legal Texts. 2012: Istanbul, Turkey, May. pp. 42 -51.
[32] G. Coode, On legalislative Expression or the Language of the Written law. 1845, William Benning and Co.: London.
[33] R. Levy, G. Andrew. Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. Proceedings of the The Fifth International Conference on Language Resources and Evaluation (LREC'06). Genoa, Italy: European Language Resources Association (ELRA). 2006.
[34] CoreNLP, S. Using the Stanford CoreNLP API.2014, URL: http://stanfordnlp.github.io/CoreNLP/api.html.
[35] G. Boella, L. Di Caro, L. Robaldo. Semantic Relation Extraction from Legislative Text Using Generalized Syntactic Dependencies and Support Vector Machines. Proceedings of. Berlin, Heidelberg: Springer Berlin Heidelberg. 2013, pp. 218-225.
[36] L. Lesmo, The Turin University Parser at Evalita 2009. Proceedings of the 11th Conference of the Italian Association for Artificial Intelligence. Itally, 2009.
[37] A. Sleimi, N. Sannier, M. Sabetzadeh, L. Briand, J. Dann, Automated Extraction of Semantic Legal Metadata using Natural Language Processing. Proceedings of the IEEE International Requirements Engineering Alberta, Canada 2018, pp. 124-135.
[38] N. Kiyavitskaya, N. Zeni, T.D. Breaux, A.I. Antón, J.R. Cordy, L. Mich, J. Mylopoulos, Automating the Extraction of Rights and Obligations for Regulatory Compliance. Proceedings of. Berlin, Heidelberg: Springer Berlin Heidelberg. 2008, pp. 154-168.
[39] N. Zeni, E.A. Seid, P. Engiel, J. Mylopoulos, Building Large Models of Law with NómosT. Proceedings of. Cham: Springer International Publishing. 2016, pp. 233-247.
[40] E. Hovy, J. Lavid, Towards a 'science' of corpus annotation: A new methodological challenge for corpus linguistics. Interrnational Journal of Translation, 2010. 22(1) 13-36.
[41] Sheffield, T.U.o., GATE Developer. 1995, The Univ. of Sheffield: Sheffield, UK.
[42] F. Dell'Orletta, S. Marchi, S. Montemagni, G. Venturi, T. Agnoloni, E. Francesconi, Domain adaptation for dependency parsing at evalita 2011. International Workshop on Evaluation of Natural Language and Speech Tool for Italian, pp. 58-69. Springer, Berlin, Heidelberg, 2012.
[43] G. Boella, L. Humphreys, M. Martin, P. Rossi, L. van der Torre, Eunomos, a Legal Document and Knowledge Management

System to Build Legal Services. Proceedings of. Berlin, Heidelberg: Springer Berlin Heidelberg. 2012, pp. 131-146.
[44] J. Sweller, P. Ayres, S. Kalyuga, Cognitive Load Theory. Explorations in the Learning Sciences, Instructional Systems and Performance Technologies, ed. M. Spector and S. Lajoie. 2011, New York: Springer.
[45] J. Sweller, Human Cognitive Architecture. Instructional Science, 2004. 32(1) 9-31.
[46] J. Sweller, S. Sweller, Natural Information Processing Systems. Evolutionary Psychology, 2006. 4(1) 147470490600400135.
[47] J. Sweller, Cognitive Load Theory, Learning Difficulty, and Instructional Design. Learning and Instruction, 1994. 4 295-312.
[48] G.A. Miller, The Magical Number Seven, Plus or Minus Two Some Limits on Our Capacity for Processing Information. Psychological review, 1956. 63 81 - 97.
[49] N. Cowan, The magical number 4 in short-term memory: A reconsideration of mental storage capacit. Behavioural and Brain Science, 2000(24) 87-185.
[50] M.A. Just, P.A. Carpenter, A capacity theory of comprehension: individual differences in working memory. Psychological review, 1992. 99(1) 122-149.
[51] E. Gibson, The dependency locality theory: A distance-based theory of linguistic complexity, in Image, Language, Brain: Papers from the First Mind Articulation Project Symposium, Y.M. Marantz and W. O'Neil, Editors. 2000, The MIT Press: Cambridge, MA. pp. 95-126.
[52] P.A. Carpenter, M.A. Just, The Role of Working Memory in Language Comprehension, in Complex Information Processing: The Impact of Herber A. Simon, D. Klahr and K. Kotovsky, Editors. 1989, Lawrence Erlbaum Associates: Hillsdale, NY. pp. 31-68.
[53] K.E. DeLeeuw, R.E. Mayer, A Comparison of Three Measures of Cognitive Load: Evidence for Separable Measures of Intrinsic, Extraneous, and Germane Load. Journal of Educational Psychology 2008. 100(1) 223 - 234.
[54] W. Schnotz, C. Kürschner, A Reconsideration of Cognitive Load Theory. Educational Psychology Review, 2007. 19(4) 469-508.
[55] J. Sweller, Cognitive load during problem solving: Effects on learning. Cognitive Science, 1988. 12(2) 257-285.
[56] J. Sweller, P. Chandler, Why Some Material Is Difficult to Learn. Cognition and Instruction, 1994. 12(3) 185-233.
[57] J. Sweller, J.J.G.v. Merrienboer, F.G.W.C. Paas, Cognitive Architecture and Instructional Design. Educational Psychology Review, 1998. 10(3) 251-296.
[58] S. Kalyuga, Informing: A Cognitive Load Perspective. Informing Science: the International Journal of an Emerging Transdiscipline, 2011. 14 33-45.
[59] J. Sweller, Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. Educational Psychology Review, 2010. 22(2) 123-138.
[60] The Securities Act 2002, 318A, s 48(2). 2002, The Government of Barbados.
[61] N.M. Langton, Cleaning up the act:using plain English in legislation. Clarity - Journal of the International Association Promoting Plain Legal Language, 2005(54) 28-33.
[62] J. Leppink, F. Paas, C.P. Van der Vleuten, T. Van Gog, J.J. Van Merriënboer, Development of an instrument for measuring different types of cognitive load. Behavior Research Methods, 2013. 45(4) 1058-1072.