Requirements Content Goodness and Complexity Measurement Based On NP Chunks

Chao Y. Din
George Mason University
Fairfax, Virginia, 22030, USA

ABSTRACT

In a typical software development project, a requirements document summarizes the results of the requirements analysis and becomes the basis for subsequent software development. In many cases, the quality of the requirements documents dictates the success of the software development. The need for determining the quality of requirements documents is particularly acute when the target applications are large, complicated, and mission critical. The purpose of this research is to develop quality indicators to indicate the quality of requirements statements in a requirements document. To achieve the goal, the goodness properties of the requirements statements are adopted to represent the quality of requirements statements. A suite of complexity metrics of requirements statements is proposed as the quality indicators and is developed based upon research of noun phrase (NP) chunks. A two phased empirical case study is performed to evaluate the usage of the proposed metrics. By focusing upon the complexity metrics based on NP chunks, the research aided in development of complexity indicators of low quality requirements documents.

Keywords: Complexity Measurement, Complexity Metrics, Cohesion, Coupling, NP Chunk, Requirements, and Software Quality.

1. INTRODUCTION

This paper asserts that a requirements document, or formally Software Requirements Specification, is the single artifact produced through the requirements engineering process. Its quality inevitably becomes the main focus of requirements management. Despite abundant suggestions and guidelines on how to write high quality requirements statements, good requirements documents are difficult to find.

The purpose of this research [4] is to develop a set of metrics to indicate the quality of requirements statements in a requirements document. The quality factors are presented by a set of goodness properties. The indicators will be able to identify requirements statements with low goodness property values.

This research [4] uses statistical and partial parsing approaches to obtain a subset of noun phrases, named Noun Phrase (NP) chunks. Abney indicated that chunks are the basic language parsing unit and they correspond to "the basic concepts" for human brains to comprehend a text document [1]. NP chunks are hence adopted as the basic processing units in this research.

This research [4] developed three complexity metrics: count of NP chunks (NPC-Count), cohesion of requirements sections (NPC-Cohesion), and coupling of requirements sections (NPC-Coupling).

A two-phased empirical case study was performed [4] to evaluate the proposed complexity metrics. Phase I of the case study compared the NPC-Cohesion and NPC-Coupling metrics with the cohesion and coupling metrics proposed by Ricker [18]. Ricker's research demonstrated the correlation between the complexity metrics and understandability, or comprehension, of the requirements. By demonstrating the consistency between the two sets of metrics, this research [4] proved to be correlated to understandability, one of the goodness properties, of the requirements statements. Furthermore, the case study showed that the NP chunk based complexity metrics possess the following two additional capabilities: (1) they differentiate nouns from other syntactic categories (or word classes) – an important capability to differentiate object methods and properties from object classes, and (2) they adopt the spatial distance of NP chunks as the measuring units – an important capability in a cognition complexity model [2].

The phase II case study then demonstrated how the three proposed metrics can be used to identify low quality requirements statements.

Based upon the two phased case study, it is assured that the proposed complexity metrics indicate the content goodness properties of requirements. The contribution of the research [4] is the construction of a suite of NP chunk based complexity metrics and the evaluation of the proposed suite of metrics.

2. RESEARCH PROBLEM AND ITS IMPORTANCE

How to identify low quality requirements statements in a requirements document is an intricate research question. This research [4] answers the question in a constrained environment where the current best practices of identifying requirements and eliminating requirements defects are adopted. The constraints are as follows.

1) A systematic requirements method such as Viewpoint Oriented Requirements Definition (VORD) has been followed to produce the requirements document [20].
2) Requirements are written in the IEEE standard requirements document format (IEEE/ANSI 830-1993).
3) The requirements document is grammatically correct and spelling errors have been checked.

4) Traditional requirements guidelines to avoid ambiguous terms (large, many, user friendliness) and week phrases (as applicable, as required, as a minimum) have been followed.
5) A domain thesaurus and/or company term definitions have been supplied.
6) A requirements inspection method has been adopted to eliminate requirements defects.

Note that requirements inspection can be effective only if the sections of requirements document is limited to 8-15 pages so that a requirements quality inspector can perform inspection of that sections within two hours time frame [11].

Another difficulty in identifying requirements defects is due to the spatial complexity of distant requirements, which refers to related requirements that are scattered far apart in a large requirements document. The current requirements inspection practice does not consider this kind of requirements defects.

The importance of the Research

The proposed suite of complexity metrics is supported by a tool researched and developed to identify high complexity and hence low quality requirements. Once low quality requirements are identified, analysis of the low quality requirements can be conducted so that they can be classified into categories of potential risks. Appropriate management actions can then be considered. [21] Identified several categories of system risks due to low quality requirements.

Low quality requirements are not only the source of system product risks but also the source of system development resource risks, which includes cost overrun and schedule delay. Using the proposed suite of complexity metrics as quality indicators, an impact assessment and threats classification of the identified low quality requirements can be performed. Such an early warning is vital to rescue a possibly failing project.

The development of high quality systems depends on management's awareness of such low quality requirements, their ability to expediently assess the impacts of those low quality requirements, and the capability to develop a plan to rectify the problem. This identification of low quality requirements and the subsequence risk analysis of those requirements provide the foundation for the development of high quality systems.

Requirements inspection is designed to identify requirements defects, which is also the goal of the proposed complexity metrics. However, the requirements inspection process usually requires significant time and efforts. Furthermore, defects due to related requirements that span more than 8 to 15 pages apart cannot be identified by requirement inspection [11].

On the other hand, the proposed complexity metrics are developed using computer programs that provide a way to quickly identify potential requirements defects. More importantly, identifying defects due to scattered but related requirements in large requirements documents is not an issue for the proposed complexity metrics.

3. BACKGROUND

Quality and Content Goodness Properties

Schneider proposed 11 goodness properties as a better coverage of quality factors [19]: Understandable, Unambiguous, Organized, Testable, Correct, Traceable, Complete, Consistent, Design independence, Feasible, and Relative necessity.

In this research [4], the main concerns are the four goodness properties: Understandable, Unambiguous, Organized, and Testable. These four goodness properties are named as *Content Goodness Properties* and are the only goodness properties on which the remainder of the research will focus.

Complexity, Complexity Metrics and Measurement

Complexity is a major system characteristic that controls or influences quality. It has been widely accepted as an indirect indicator of quality and hence the content goodness properties [6,8,10,12]. The remainder of the research focuses the discussion on complexity.

[17] Provides a survey of complexity metrics and identified five out of 375 metrics that are related to requirements. Unfortunately, none of those five metrics are used to measure the natural language descriptions of the requirements. [18], not listed on the survey, developed a set of requirements metrics: cohesion, context, and coupling. One of the contributions [18] made is the demonstration of a positive correlation between cohesion, context, and coupling metrics and understandability of requirements statements. In [18], the context metric is assessed by the relationships between the sentences of a section and their section title. This research [4] does not consider the context metric.

Readability Index

When measuring the quality of documents written in natural languages, the readability indexes or metrics must be considered. In general the written communication skills are measured in terms of readability and hence the use of readability indexes. Readability is a measure of the ease that a piece of writing can be read. Readability indexes are designed to access the suitability of a piece of writing for readers at particular grade levels or ages.

Factors considered in the readability indexes are number of words, number of syllables in words, number of words in sentences, …, etc. Scores of the readability indexes are compared with scales based on judged linguistic difficulty or reading grade level.

One of the arguments against readability indexes is that difficult text often contains difficult words because it discusses abstract ideas while easy text uses common words because it discusses concrete experiences.

Another argument against readability indexes is that Readability formulas are based on the simple surface characteristics of the text. Measuring text elements that are primarily based on surface characteristics does not adequately capture comprehension [15].

Graesser, and his colleagues conducted a research project named Coh-Metrix to develop new readability indexes that are based on cohesion relations, interaction between a reader's skill level, world knowledge, and language and discourse characteristics. Its modules use lexicons, part-of-speech classifiers, syntactic parsers, templates, corpora, latent semantic analysis, and other components that are widely used in computational linguistics [7].

Unfortunately, readability indexes, including Coh-Metrix, are not comparable with this research [4] for the following reasons:

1) The readability metrics are designed for the whole documents, instead of sections of documents.
2) The readability scores are not reliable indicators when the document under evaluation has less than 200 words [14]. However, many of the requirements statements have less than 50 words.
3) Although Coh-Metrix attempts to measure the cohesion of text, the definition of cohesion used by Coh-Metrix is different from the definition of cohesion used in Computer Science, and there are no coupling metrics in Coh-Metrix.
4) Coh-Metrix does not have a single metric to represent the size, cohesion, or coupling complexity. Coh-Metrix includes more than 50 metrics to measure very specific aspects of texts. No composite metric that combines those specific aspects of a document has been proposed.
5) Coh-Metrix attempts to measure the cohesion of texts. Future work of Coh-Metrix may address comprehension, or understandability. However, Coh-Metrix will never address the issue of testability and many other goodness properties.

4. NP CHUNK BASED COMPLEXITY METRICS

A major issue of a measurement program is "what to measure," and it is one of the most critical issues to the measurement research.

Because humans tend to read and speak texts one chunk at a time, Abney proposed using what is called *chunks* as the basic language parsing unit. There are several categories of chunks similar to the traditional categories of phrases. For example, there are Noun Phrase (NP) chunks, Verb Phrase (VP) chunks, Prepositional Phrase (PP) chunks, … etc [1]. This research [4] focuses on NP chunks and ignores other types of chunks.

Three Core Metrics

It is believed that a small subset of existing metrics can enable parsimonious evaluation, prediction and control of software complexity [9]. This research [4] hence proposes three types of complexity metrics, NPC-Count, NPC-Cohesion, and NPC-Coupling, for measuring the complexity of requirements in a requirements document.

Size counts are the oldest method of measuring complexity. For software design and coding, the most popular size count is Line of Code (LOC). The wide acceptance of LOC as a complexity metric is due to its simplicity, ease of application, inertia of tradition, absence of alternative size metrics, and its intuitive appeal [5, 11]. Based upon the above reasons, two distinct metrics (NPC-Sentence and NPC-Req) are developed to count the NP chunks of a text, and these two metrics are collectively named as NPC-Count.

Darcy and Kemerer believe that cohesion and coupling are effective metrics and they can represent the essential complexity measures for the general software design tasks [3]. Hence, NPC-Cohesion and NPC-Coupling are chosen in this research [4] to represent the complexity of requirements. To assist the identification of low quality requirements, a composite metric (NPC-Composite) that combine cohesion and coupling measures is also proposed and studied in the research.

Sentence/Requirements Statement Level Complexity

The sentence level complexity metric, or NPC-Sentence, can be calculated as follows. For each NP chunk, the occurrence count in a sentence is divided by the total occurrence counts in all sentences. Then all the frequency distributions of the NP chunks in the sentence are added together to form the final complexity value.

The requirements statement level complexity metric, or NPC-Req, is the aggregation of NPC-Sentence of the component sentences.

Intra-Section Level Complexity

The proposed NPC-Cohesion metric is a normalized cluster size that can be calculated using the sum of all cluster sizes in a requirement section divided by the size of the requirements section. Here a cluster is defined as the collection of adjacent sentences in a requirements section that shares the same NP chunks. For example, if sentence 1 contains NP chunk A, sentence 2 contains NP chunk A and B, and sentence 3 contains NP chunk B, then the three sentences form a single cluster.

Inter-Section Level Complexity

The proposed NPC-Coupling metric value is the sum of the spatial distances between its internal and external NP chunks. If an NP chunk belongs to a cluster, then the centroid of the cluster is used to calculate its distance to the external NP chunks.

5. EXPIRICAL CASE STUDY

Case Study Methodology

The case study methodology [22] is an empirical research strategy commonly used in psychology, sociology, political science, social work, business, community planning, and economics. The case study methodology adopted here consists of five components, which form a logic plan for the research design of case studies.

1) A study's questions
2) Study propositions, or hypotheses
3) Unit(s) of analysis
4) The logic linking of the data to the propositions
5) The criteria for interpreting the finding

There are three types of case studies: exploratory, descriptive, and explanatory. In a nutshell, an exploratory case study is either used to define the questions and hypotheses of a subsequent case study or to determine the feasibility of a subsequent research, i.e., an explanatory case study. A descriptive case study presents a complete description of a phenomenon. An explanatory case study explains the cause-effect relationships indicated in the research question [22].

Two Phased Case Study

The goal of the research [4] is to answer the constraint question about identifying low quality requirements statements in a requirements document with the specified constraints. This question can be divided into two sub-questions.

Q1. Can NP chunk based complexity metrics be more effective than the term based complexity metrics in terms of measuring requirements content goodness properties?

Q2. How and why NP chunk based complexity metrics measure the content goodness properties of requirements statements?

The two phases of the case study, exploratory in phase I and explanatory in phase II, are designed to answer the two sub-questions, respectively. The first sub-question and hence the phase I case study is an evaluation of NP chunk based complexity metrics. If NP chunk based complexity metrics cannot produce consistent results as the term based complexity metrics proposed by [18], there is no reason to perform further study on the research question. Ideally, the NP chunk based complexity metrics should be more effective than the term based complexity metrics; otherwise, the research provides little contribution to the research question.

The second sub-question and the corresponding phase II case study assume the phase I case study has positive outcomes. The phase II study then explains how and why the NP chunk based complexity metrics work. Evidence and findings to support the proposed metrics are presented one by one in this case study.

Exploratory Case Study – Phase I

The five components of the phase I case study are described as follows.

Study Question: The study question is "Can NP chunk based complexity metrics be more effective than the term based complexity metrics in terms of measuring requirements content goodness properties?"

Study Propositions, or Hypotheses: The purpose of the phase I case study is to determine whether the NP chunk based complexity metrics can measure content goodness properties of requirements documents. The term based complexity metrics published by Ricker [18] reveal positive correlation to understandability, one of the content goodness properties of requirements documents.

The derived specific study hypotheses/propositions are as follows.

H1. Consistency: The NP chunk based complexity metrics are consistent with the term based complexity metrics.

Ricker proposed three term based complexity metrics: context, cohesion, and coupling, for requirements statements. However, the published metric values, or measures in [18] focus mainly on the cohesion and coupling metrics. The only metrics that can be compared against are cohesion and coupling metrics. Hence, the above proposition is divided into the two sub-propositions: one for cohesion and the other for coupling.

P1.1. Cohesion: NPC-Cohesion, the NP chunk based cohesion metric, is consistent with the term based cohesion metric.

P1.2. Coupling: NPC-Coupling, the NP chunk based coupling metric is consistent with the term based coupling metric.

For simplicity reason, the degree of consistency for the above two propositions can be categorized into three ordinal values: strongly consistent, somewhat consistent, and cannot-determine. The degree of consistency must be strong in order to claim the two sets of metrics are consistent to each other.

P2. Sensitivity/Accuracy: The NP chunk based complexity metrics are either more sensitive or more accurate than the term based complexity metrics.

The degree of sensitivity or accuracy can also be categorized into three ordinal values: strongly sensitive/accurate, somewhat sensitive/accurate, and cannot-determine. The degree of sensitivity/accuracy must be strong to claim the proposed metrics are more sensitive or accurate than Ricker's metrics.

P3. Additional Information: The NP chunk based complexity metrics can provide additional information on the requirements content goodness properties than the term based complexity metrics.

The linking of derived data to the above proposition can be categorized into two ordinal values: "yes" (it provides additional information) and "no" (it does not provide additional information).

Unit(s) of Analysis: The unit of analysis for the phase I case study is a requirements document of a Federal Aviation Agency (FAA) project available in [18].

The Logic Linking of the Data to the Propositions Criteria for Interpreting the Findings: The logic linking of the data to the propositions represents the first data analysis step in the case study design, which can be divided into two sub-steps: cohesion and coupling. The second data analysis is to interpret the findings using the evaluation criteria stated above.

Cohesion Metrics: Based upon the proposed NPC-Cohesion metric defined previously, the NPC-Cohesion measures and the cohesion measures published in [18] are consistent with each other except in one section – section 11 of the FAA requirements document.

The mismatch between the two cohesion metrics can be explained as follows. Section 11 of the FAA requirements document consists of two sentences. By closely examining the two sentences, it was found that there are no common NP chunks between the two sentences. This is why the NPC-Cohesion metric gives a low cohesion measure for the above requirements section. On the other hand, Ricker uses terms to measure the cohesion of the section, and the word "outputs" appears in the first sentence as a noun, while the word "output" appears in the second sentence as a verb. Ricker's algorithm does not consider syntactic categories and hence links the two sentences.

It is believed that a word in different forms, i.e., verbs and nouns, in different sentences should not always be considered as cohesive, since the two words in the two forms can refer to two totally different objects. By closely examining the two sentences, it can be found that the word "output" in the two sentences indeed refers to two different things or two different concepts. Hence, the proposed cohesion metrics is more effective.

As indicated above, the evaluation criterion for the cohesion proposition (P1.1) is whether the two sets of metrics are strongly consistent with each other. Since there is only one mismatch and the mismatch can be explained, the degree of consistency is strong. For the additional information proposition (P3), the evaluation criterion for linking the data to the proposition is whether the NP chunks based complexity metrics can provide additional information. Since the NP-Cohesion metric can differentiate word classes, the NP-Cohesion metric does provide additional information. It is hence concluded that NPC-Cohesion supports proposition P1.1 and P3.

Coupling Metrics: The coupling measures based on the NPC-Coupling metric are consistent with the coupling measures in [18] except in one section – section 4 of the requirements document.

The discrepancy between the two coupling metrics can be explained as follows. Section 3 is titled as "routing processing", and Section 4 is titled as "additional routing processing." Since the fourth section is a supplement to the third section, its coupling in Ricker's method is very high.

On the other hand, the coupling value for section 4 is low for the NPC-Coupling metric because the spatial distance between the two sections is low. In other words, the effect of spatial distance is counted in the NP-Coupling metric, while Ricker's method does not consider the spatial distance.

Since there is only one mismatch and the mismatch can be explained, the degree of consistency is strong. For the additional information proposition (P3), the evaluation criterion for linking the data to the "additional information" proposition is whether the NP chunks based complexity metrics can provide additional information. Since the NPC-Coupling metric can measure spatial distance, the NPC-Coupling metric does provide additional information. It is hence concluded that NPC-Coupling supports proposition P1.2 and P3.

Sensitivity/Accuracy: The NPC-Cohesion metrics are relative measures. They are normalized and fall in the range of 0 to 1. Comparing such relative measures derived from different requirements documents is not logical. In other words, it is not appropriate to compare the sensitivity and accuracy of the NPC-Cohesion metrics with Ricker's metrics.

Although the NPC-Coupling metrics are based upon spatial distance between NP chunks and they are not normalized, comparing it with Ricker's metric which uses different units of measurement does not seem to be logical either. All in all, the evaluation for the derived data from the case study and the P2 proposition results in the "cannot-determine" ordinal value.

Summary: Based upon the above analysis, it can be concluded that the derived data from the case study met the evaluation criteria for the consistency hypothesis (H1) and additional information proposition (P3). On the other hand, no evidence supports the opposite argument. Hence, the phase I study question is asserted. It is clear that the NP chunk based complexity metrics are more effective than the term based complexity metrics.

Explanatory Case Study – Phase II

Study Question: The phase II study question is "How and why NP chunk based complexity metrics measure the content goodness properties of requirements statements?"

P4. NP Chunk Counts: The NP-Count, as a simple form of complexity metric, can measure the content goodness properties of the requirements statements.

P5. Cohesion: NP chunk based cohesion complexity metrics such as the NPC-Cohesion metric can measure the content goodness properties of the requirements statements.

P6. Coupling: NP chunk based coupling complexity metrics such as the NPC-Coupling metric can measure the content goodness properties of the requirements statements.

The evaluation criteria for the linking of derived data from the case study to the above three propositions is whether the linking can explain the cause-effect relationship. For simplicity reason, the cause-effect relationship can be categorized into three ordinal values: strong, medium, and weak/no cause-effect relationship.

Unit(s) of Analysis: In the phase II research design, the unit of analysis is also a requirements document. Two sources of requirements documents are used for the case study: (1) four versions of the Interactive Matching and Geocoding System II (IMAGS II) requirements documents for U. S. Bureau of Census and (2) the FAA requirements document used in the phase I study.

The Logic Linking of the Data to the Propositions Criteria for Interpreting the Findings: In this section the three major categories of metrics, sentence/requirements complexity metrics, cohesion metrics, and coupling metrics, are discussed separately.

NPC-Sentence (Sentence Level Complexity Metrics): The NPC-Sentence metric proposed in a previous section is basically a way to count the NP chunks, and it can be used to identify complex requirements.

Section 3.4 of the IMAGS II requirements document is used to illustrate how the NPC-Sentence metric works. The complexity measures are first obtained from Section 3.4 of both the version 2 and version 3 of the requirements documents. The complexity measures are then compared between the two versions of the requirements. The NPC-Sentence measures of Section 3.4 of the version 2 requirements document show that sentence 10, 11, and 12 have high degree of complexity. Subsequent iteration of requirements review indeed identified those three sentences as "difficult to understand". A new set of sentences were then developed in the version 3 of the requirements document. The comparison of the two versions of the requirements section shows that the complexity measures of the three sentences are improved in the new version of the requirements document.

NPC-Req (Requirements Level Complexity Metrics): Another section of the IMAGS II requirements document is used to illustrate the capability of the NP chunk complexity metrics at the requirements level.

The NP-Req metric values are compared between two versions of Section 3.2 of the IMAGS II requirements document: version 2 and 3. During the third iteration of the requirements gathering phase, four modifications are made, and the version 2 NPC-Sentence measures do not show clearly which sentences should be improved or re-written.

On the other hand, the version 2 NPC-Req metrics show that two requirements are the most complex requirements in the section. This coincides two of the modifications shown on the version 3 requirements document.

Based upon the above analysis, it is clear that NP chunk counts can measure the complexity of requirements statements and hence show the strong cause-effect relationship to the content goodness properties of requirements statements. In other words, the proposition P4 is supported.

Cohesion and Coupling Metrics: The cohesion measures for the version 1 to version 4 of IMAGS II requirements documents can be illustrated by the differences between the two adjacent versions of IMAGS II requirements documents. The results are three sets of measures, and they reveal that the iteration from version 1 to version 2 and from version 2 to version 3 have substantial changes. On the other hand, the iteration from version 3 to version 4 is bounded in a relatively narrow range.

Similar to the cohesion metrics, the coupling metrics also show that the iteration from version 3 to version 4 of the requirements stays in a relatively narrow range.

In addition to NPC-Cohesion and NPC-Coupling, NPC-Composite is used in the study. For the IMAGS II project, NPC-Composite shows that Section 30 is the worst requirements section. After examining the requirements document, it is found that Section 30 is for reports. Reports requirements typically reference all other sections and are independent of each other. The next low quality requirements sections are Section 3, 6, and 33. Section 3 is the requirements for the overall operations, which includes multiple requirements for suspend and shutdown some operations but leaves others operational. Section 3 indeed contains complicated requirements. Section 6 discusses the address import, and Section 33 provides performance related requirements. These two sections do not seem to be complicated.

For the FAA project, the most complicated requirements section indicated by NPC-Composite is Section 13, which has the highest number of sentences and the cohesion value for Section 13 is zero. This section is indeed complicated. The next set of low quality requirements sections are Section 6, 19, and 22. Although the NP chunk count of Section 6 is not the highest, the coupling value is the highest. The problem with Section 19 and 22 is evident by their zero cohesion value.

The Phase I case study gives evidence that the proposed NPC-Cohesion and NPC-Coupling metrics are consistent with Ricker's metrics which are correlated to understandability, a content goodness property, of requirements statements. This section again reports evidence that the NPC-Composite metric has a strong cause-effect relationship to the content goodness properties of requirements statements. In other words, NPC-Cohesion supports the proposition P5, and NPC-Coupling supports the proposition P6.

Summary: Based upon the evidence discussed above, the hypothesis that the proposed complexity metrics can identify low quality requirements statements in a requirements document can be asserted.

6. SUMMARY OF CONTRIBUTIONS

This research [4] made two contributions: (1) the invention of a suite of complexity metrics to measure the content goodness properties of requirements documents and (2) the empirical case study to evaluate the invented suite of complexity metrics.

The invented complexity metrics are researched and developed to identify low quality requirements in requirements documents. These metrics are based on the NP chunks in requirements documents. In the empirical two phased case study, it is demonstrated that the proposed metrics can measure the content goodness properties of requirements statements.

The research demonstrates the feasibility of using NP chunks as the elements of measurement for complexity metrics. In addition the invented suite of complexity metrics provides requirements engineers and managers with a tool to measure the quality of the requirements statements. These metrics can be use to identify low quality requirements. They can also be used to identify requirements and requirements sections that may require more rigorous testing. Potential flaws and risks can be reduced and dealt with earlier in the software development cycle.

At a minimum, these metrics should lay the groundwork for automated measures of requirements documents. Because those metrics are constructed by programs, they are easy to collect, a vital characteristics for a successful measurement program [16].

7. REFERENCES

[1] S. Abney, "Parsing By Chunks". Robert Berwick and Steven Abney and Carol Tenny (eds), **Principle-Based Parsing**, Kluwer Academic Publishers, 1991.

[2] S. Cant, D. R. Jeffery, and B. Henderson-Sellers, "A conceptual model of cognitive complexity of elements of the programming process", **Information and Software Technology**, 37(7), 351-362. 1995.

[3] D.P. Darcy and C.F. Kemerer, **Software Complexity: Toward a Unified Theory of Coupling and Cohesion**, MIS Research Center, Carlson School of Management, University of Minnesota, February 8, 2002.

[4] C.Y. Din, **Requirements Content Goodness and Complexity Measurement Based On NP Chunks**, Doctoral Dissertation, George Mason University, Fairfax, VA, 2007. Reprinted by VDM Verlag Dr. Muller, 2008.

[5] H.E. Dunsmore, "Software Metrics: An Overview of an Evolving Methodology", **Information Processing and Management**, 20(1-2), 183-192, 1984.

[6] N.E. Fenton and M. Neil, "Software metrics: roadmap" **Proceedings of the International Conference on Software Engineering (ICSE)**, 357-370, 2000.

[7] A.C. Graesser, D.S. McNamara, M.M. Louwerse, and Z. Cai "Coh-Metrix: Analysis of Text on Cohesion and Language", **Behavioral Research Methods, Instruments, and Computers**, 36(2), 193-202, 2004.

[8] B. Henderson-Sellers, **Object-Oriented Metrics – Measures of Complexity**, Prentice Hall PTR, New Jersey, 1996.

[9] C.F. Kemerer, "Progress, Obstacles, and Opportunities in Software Engineering Economics", **Communications of ACM**, 41, 63-66, 1998.

[10] T. Klemola, “A Cognitive Model for Complexity Metrics”, **4th International ECOOP Workshop on Quantitative Approaches in Object-Oriented Software Engineering**, June 13, 2000.

[11] S. Lee, **Proxy Viewpoints Model-Based Requirements Discovery**, Doctoral Dissertation, George Mason University, 2003.

[12] A.V. Levitin, “How to Measure Size, and How Not to”, **Proc. Tenth COMPSAC 1986**, Chicago, Oct 8-10, 1986, IEEE Computer Society Press, Washington DC, 314-318, 1986.

[13] H.F. Li and W.K. Cheung, “An empirical study of software metrics”, **IEEE Transactions on Software Engineering**, 13(6), 697-708, 1987.

[14] D.S. McNamara, “Reading both high and low coherence texts: Effects of text sequence and prior knowledge”, **Canadian Journal of Experimental Psychology**, 55, 51-62, 2001.

[15] D.S. McNamara, E. Kintsch, N.B. Songer, and W. Kintsch, “Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text”, **Cognition and Instruction**, 14, 1-43, 1996.

[16] S.L. Pfleeger, “Lessons learned in building a corporate metrics program”, **IEEE Software**, 10(3), 67-74, 1993.

[17] S. Purao and V. Vaishnavi, “Product Metrics for Object-Oriented Systems”, **ACM Computing Surveys**, 35(2), 191-221, 2003.

[18] M. Ricker, **Requirements Specification Understandability Evaluation with Cohesion, Context, and Coupling**, Doctoral Dissertation, George Mason University, Fairfax, VA, 1995.

[19] R.E. Schneider, **Process for building a more effective set of requirement goodness properties**, Doctoral Dissertation, George Mason University, Fairfax, VA, 2002.

[20] I. Sommerville, **Software Engineering**: update, 8th Edition, International Computer Science, Addison Wesley, 2006.

[21] W.M. Wilson, L.H. Rosenberg, and L.E. Hyatt, “Automated Quality Analysis of Natural Language Requirement Specifications”, **Fourteenth Annual Pacific Northwest Software Quality Conference**, October, 1996.

[22] R.K. Yin, **Case study research**, 3rd edition. Thousand Oaks, CA: Sage Publications, 2003.