LOGICAL CONNECTIONS AMONG FOUR CRITICAL ELEMENTS OF YOUR DISSERTATION

YANZHEN QU, PH.D.
AGENDA

Introduction
Problem Statement
Hypothesis Statement
Research Questions
Research Design
Conclusion
Q&A
Three critical elements of your dissertation are: problem statement, hypothesis statement, and research questions. Then you also need to complete the fourth critical element: the research design to come out a set of the experiments of data collection in order to answer your research questions quantitatively. In this talk we will explore the logical connections between the elements that we just mentioned, and show how you can utilize these logical connections to step by step to generate some of the most important content of your dissertation so that the quality and the productivity of your dissertation can be closely controlled.

Three main learning objectives are

• Learn how to define one sentence problem statement, one sentence hypothesis statement and one sentence contribution statement.

• Learn how to derive a set of research questions based on the problem statement and hypothesis statement.

• Learn how to use research questions to guide the research design so that your data collection process is effective and efficient.
Congratulations!
You are now on the most challenging “long match”!

THE EVOLUTION OF INTELLECTUAL FREEDOM

Before Grad School  Grad Student  Assistant Professor  Tenured Professor  Emeritus Professor

“I’m going to research whatever I want!”
“I’m going to research whatever my professor wants!”
“I’m going to research whatever my tenure committee wants!”
“I’m going to research whatever my grant committee wants!”
“I’m going to research whatever I want!”

“Research in Peace”

WWW.PHDCOMICS.COM
INTRODUCTION

Elements of a Dissertation or an Academic Paper
1. Title
2. Abstract
3. Keywords
4. Introduction
5. Problem Statement
6. Hypothesis Statement
7. Research Questions
8. Related Works
9. Research Design/Methodology
10. Experiments & Data Analysis
11. Conclusion
12. Reference

Where is the best place to start? Why?
INTRODUCTION

Phase 1:
Problem Statement
Hypothesis Statement
Research Questions

Phase 2:
Introduction
Related Works
References

Phase 3:
Research Design/Methodology
Experiments & Data Analysis
Conclusion

Phase 4:
Abstract
Keywords
Title

Time
Problem Statement

The “One Sentence Problem Statement” is a Statement having only one sentence with two components:

- The Effect or Syndrome of the problem;
- The Root Cause of the problem.

Example:

“Current malware code-based signature-detection methodologies are ineffective in detecting advanced malware due to the variant nature attributed to polymorphic and metamorphic capabilities, causing difficulty in identifying strings that can be used for signature creation.”

“Current malware code-based signature-detection methodologies are ineffective in detecting advanced malware” is the Problem Effect or Symptom

“the variant nature attributed to polymorphic and metamorphic capabilities, causing difficulty in identifying strings that can be used for signature creation” is the Problem Root Cause
Hypothesis Statement

The predicated relationship among the variables being investigated [1].

The “One Sentence Hypothesis Statement” is a Statement having only one sentence with the following characteristics:

- Driven by the Problem Statement such that it is an assumed solution to address the root cause of the problem;
- Purposeful such that it has something new to the Body of Knowledge;
- Testable such that it, combined with the problem statement, can derive some causation relationship type Research Questions, each of which is testable.

Example:

“If we create an aggregated signature by first extending the signature searching scope beyond malware code, and then choosing the statistically significant invariants through logistic regression, we will be able to effectively detect advanced malware with this new signature-based detection methodology.”

“an aggregated signature by first extending the signature searching scope beyond malware code, and then choosing the statistically significant invariants through logistic regression” is a new solution proposed to address the root cause of the problem - “the variant nature attributed to polymorphic and metamorphic capabilities”. This new solution is also the new contribution to the Body of Knowledge at the moment it was proposed.
How to form your Research Questions?

Create Causation Relationship Research Questions based on problem and hypothesis statements:

• Cause (Independent Variable): The factor makes something else exist or change
• Effect (Dependent Variable): The difference between prior and post change due to manipulating Independent Variables

Tip:
• This type of research questions allow you to see the expected results ahead of time.
• IVs and DVs represent key concepts of the solution which can be abstracted out from problem statement (get IVs) and hypothesis statement (get DVs).
• Focus on “How” not on “Why”
Example

Problem Statement
“Current malware code-based signature-detection methodologies are ineffective in detecting advanced malware due to the variant nature attributed to polymorphic and metamorphic capabilities, causing difficulty in identifying strings that can be used for signature creation.”

Advanced malware is a new type of malware with many variants -> Malware Type <- Cause/Independent Variable

Hypothesis Statement
“If we create an aggregated signature by first extending the signature searching scope beyond malware code, and then choosing the statistically significant invariants through logistic regression, we will be able to effectively detect advanced malware with this new signature-based detection methodology.”

“Effectively detect advanced malware” -> Detection Rate <- Effect/Dependent Variable
Example of Research Question:

“How will the detection rate of the behavior-based signature change against various types of malware such as known malware, zero-day variant malware and zero-day malware?”

**Independent Variable:** Malware Type

**Dependent Variable:** Detection Rate
Based on your Research Question

Example of Research Question:

“How will the detection rate of the behavior-based signature change against various types of malware such as known malware, zero-day variant malware and zero-day malware?”

Independent Variable: Malware Type
Dependent Variable: Detection Rate

Part I: Experiments Design and Data Collection

Input Data Set: Sort your experiments based on the Malware Type. That is you will make experiments classified such that all the malware samples belonged to the same type will be considered as the same experiment. In this example, we have three types of malware: (a) Malware Family, (b) Zero Day Malware Variants, and (c) True Zero Day Malware.

Data Processing Procedure/Algorithm: what you will do to each type of input Data Set so that you can get the result (or the output) of DV?

Any Theoretical Model used to assist the Data Processing can also be presented here.

Output Data Set: The results of the Data Processing Procedure/Algorithm.

Part II: Data Analysis

Explain the meaning of the Experiments generated output data

“How will the detection rate of the behavior-based signature change against various types of malware such as known malware, zero-day variant malware and zero-day malware?”
CONCLUSION

How can I find my boat (research problem)?

Wherever you find a bad experience

Start from your current class and your current term paper.
CONCLUSION

Publish your work at a peer-reviewed publication or conference.

“Yeah, but good luck getting it peer-reviewed.”