HUNG PHAN

(515) 735 6030 hungphd@iastate.edu http://www.public.iastate.edu/~hungphd/

908 Douglas Ave. Unit 12, Ames, IA 50010, USA

KEY WORDS OF INTEREST

Natural Language Processing, Machine Translation, Syntax Directed Translation, Mining Software Repository, Machine Learning, Software Engineering, AI.

EDUCATION

Ph.D. in Department of Computer Science, Iowa State University, USA. 7/2015-present

B.Eng. in Software Engineering, Hanoi University of Science and Technology, Vietnam. 8/2007-6/2012

PROJECTS

NLPL - Combining Natural Language with Programming Language 1/2018-Ongoing

- Provide a tool to realize the idea of Literate Programming, which is introduced by Donald Knuth in 1990s. This tool, NLPL, will help developers to write the natural language inside the environment of Java source code.
- Design an algorithm to analyse the Comments in the source code and produce the translation related to the NLP parse tree at terminal and non-terminal levels, called **Syntax Directed Translation**.
- Verifying the validity of developers' comments by standard english grammars and the environment of surrounding code.

StatType - Inferring Fully Qualified Name for Online Forums Code Snippets 1/2017-2/2018

- Implementing an AST Parser to convert from code to tokens with the extraction of Fully Qualified Name (FQN).
- Building a parallel corpus from 1000 Java github projects, with the source are tokens from code without FQN and the target are tokens with FQN. Get an Machine Translation model from this large scale corpus.

- Design an algorithm for inferring StackOverflow Code Snippets' FQN based on Phrase-based SMT.

Inferring between Documentation and Implementation in Java 8/2016-5/2017

- Building a parallel corpus, in which one language is the condition of exception for each method in java doc, the other language is the correspoding condition of Throw statements in Implementation in large scale corpus..
- Implementing a Phrase-based Statistical Machine Translation model to learn the inference from Javadoc to Implementation and vice-versa.
- Applying the inference problem in automatically **generate new Javadoc** and **check the inconsistency** between Documentation and Implementation of Behavior Exceptional.

JV2CS - Statistical Migration of API usages 8/2015-6/2016

- Mining pairwise mapping between APIs in Java and CSharp, learning the semantic relation between APIs.
- Implementing a Machine Translation (MT) model, by analyzing large scale source code repository in Java, producing an parallel corpus contains API sequences of Java and C# and applying Phrase-based SMT to build an MT model.
- Designing a method for optimizing original MT model, by adding pairwise mapping between APIs learned in large-scale corpus.

PROFESSIONAL SKILLS

- Deep knowledge about Data Structure and Algorithm.
- Deep knowledge about Database and Text Processing.
- Experiences in Java, C#, C/C++, Python, SQL.
- Matlab, Eclipse, VS, Git.
- JDT, Soot, Z3.
- Phrasal, Tensorflow.

SELECTED PUBLICATIONS

Statistical Learning of API Fully Qualified Names in Code Snippets of Online Forums. Hung PHAN, Hoan Nguyen, Ngoc Tran, Linh Truong, Anh Nguyen and Tien Nguyen. ICSE 2018.

A Deep Neural Network Language Model with Contexts for Source Code . Anh Tuan Nguyen, Trong Van Nguyen, **Hung Dang Phan** and Tien N. Nguyen. SANER 2018

Statistical Learning for Inference between Implementations and Documentation . Hung PHAN, Hoan Anh Nguyen, Tien N. Nguyen and Hridesh Rajan. **ICSE-NIER 2017**.
Statistical migration of API usages. Hung Dang Phan, Anh Tuan Nguyen, Trong Duc Nguyen and Tien N. Nguyen. **ICSE-Demo 2017**.
Exploring API Embedding for API Usages and Applications. Trong Duc Nguyen, Anh Tuan Nguyen, **Hung Dang Phan**, and Tien N. Nguyen. **ICSE 2017**.