



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

itee^{PhD}
information technology
electrical engineering



DIE
TI

UNI
NA

Luigi Libero Lucio Starace

Web application model inference to support model-based testing

Tutor: Prof. Sergio Di Martino
Cycle: XXXV

co-Tutor: Prof. Adriano Peron
Year: 2020/2021

My Background

- M.Sc. Degree in Computer Science
 - Thesis: *“Model-based testing and Model Checking for Safety-critical Hierarchical Systems”*
- Ph.D. fellowship funded by *NetCom Group S.p.A.*

Research Field of Interest

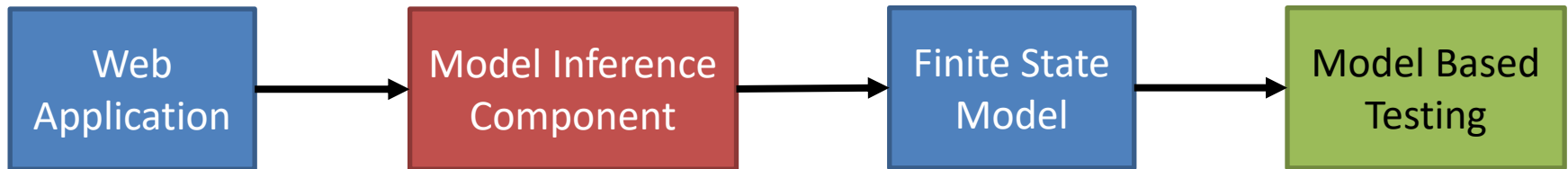
Main field of interest is **software verification**:

- Model-Based Testing for web applications;
- Regression Test Prioritization;
- Empirical studies on Software Testing;

Main Research Activity: Overview

Model Inference for Web Applications

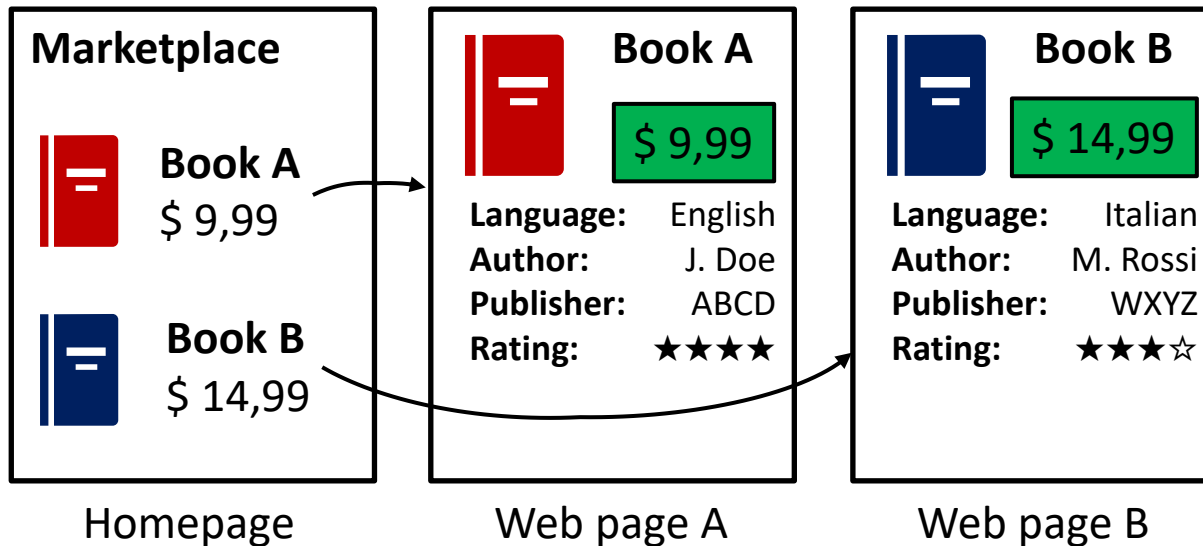
- To support model-based testing



Goal: improve the web application testing process

Contribution: define near-duplicate detection techniques to improve the quality of models inferred by crawling

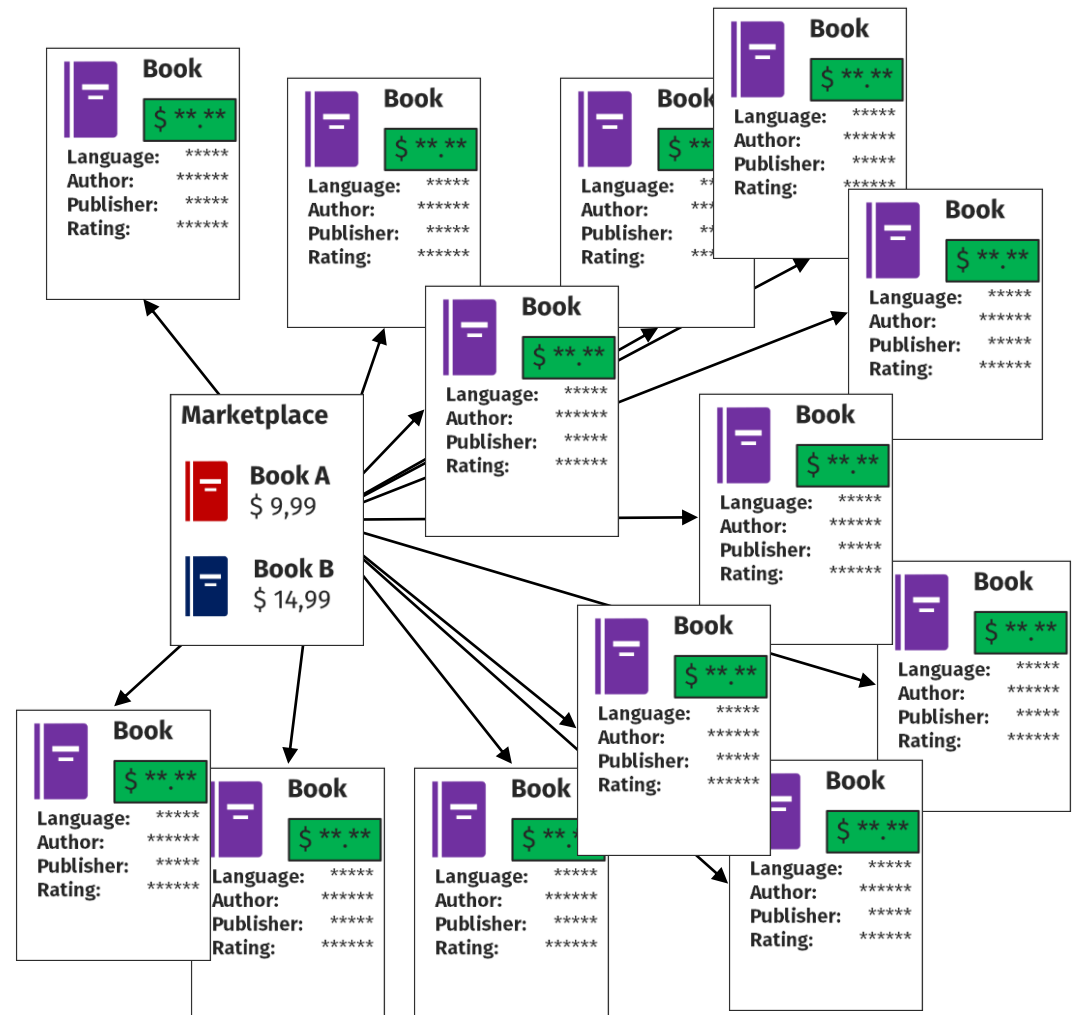
The near-duplicates problem



- Near-duplicates are states corresponding to web pages that are different, but nonetheless represent the same functionality

The near-duplicates problem

- Near-duplicates have a negative impact on the **precision** and **completeness** of the models;
- Limiting the application of model-based testing techniques



Proposed solution

Currently investigating a solution based on **Tree Kernel** functions:

- Largely used in NLP
- Measure similarity between tree-structured objects
- Idea: measure similarity between the tree-structured DOM representation of web pages

Validation

- Compare the models inferred using the proposed technique with gold standard models provided by a massive dataset made available in a recently-presented ICSE paper [1].
- Use the near-duplicate detection techniques benchmarked in [1] as a baseline.

[1] Yandrapally et al. “Near-duplicate detection in web app model inference.” Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering. 2020.

Products: Journal Papers

International Journal of Geographical Information Science (IJGIS)

Vehicular crowd-sensing: a parametric routing algorithm to increase spatio-temporal road network coverage

Asprone, D., Di Martino, S., Festa, P., **Starace, L. L. L.** (2021)

Status: **published**

IEEE Internet of Things Journal (JIOT)

A Budget-aware Decentralized Incentivization Solution for Vehicular Crowd-Sensing

Di Martino, **Starace, L. L. L.** (2021)

Status: **under revision**

Products: Conference Papers

15th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2021)

Web Application Testing: Using Tree Kernels to Detect Near-duplicate States in Automated Model Inference

Corazza, A., Di Martino, S., Peron, A., **Starace, L. L. L.** (2021)

Status: [accepted and published in the conference proceedings](#)

14th International Conference on the Quality of Information and Communications Technology (QUATIC 2021)

Expressing structural temporal properties of safety critical hierarchical systems

Benerecetti, M., Mogavero, F., Peron, A., **Starace, L. L. L.** (2021)

Status: [accepted and published in the conference proceedings](#)

24th Euro Working Group on Transportation Meeting (EWGT 2021)

Vehicular Crowd-Sensing on Complex Urban Road Networks: A Case Study in the City of Porto

Di Martino, S., **Starace, L. L. L.** (2021)

Status: [accepted \(to appear in the conference proceedings\)](#)

Summary of Study Activities

	Courses	Seminars	Research	Tutorship
Year 2 ECTS	11,5	10,8	41	1,6
Expected Year 2 ECTS range	10 – 20	5 – 10	30 – 45	0 – 1,6

Among others, I took the following courses:

- Statistical data analysis for science and engineering research (4 ECTS);
- Data Science for Patient Records Analysis (2.5 ECTS)

Selected attended conferences and events:

- *Empirical Software Engineering and Measurement (ESEM 2021).*
- *Doctoral Symposium at the International Symposium on Software Testing and Analysis (ISSTA 2021).*

Tutorship Activities

Carried out tutorship activities in the following courses at UniNA:

- *Ingegneria del Software* (B.Sc. In Computer Science)
- *Object Orientation* (B.Sc. In Computer Science)
- *Software Project Management and Evolution* (M.Sc. In Computer Science)

Tutored 7 B.Sc. and 3 M.Sc. students in Computer Science working on their internal curricular internship and theses.

Plans for the third year

- Spend two additional months at the Università della Svizzera Italiana, Lugano, CH.
- Continue working on validating the proposed TK-based approach
- Investigate the feasibility of a deep-learning based solution
- Submit at least one article to a top-journal in software engineering with the results achieved so far.

Thesis topic

Generalize the proposed near-duplicate detection approach and define an empirical framework to support both:

- Definition of novel near-duplicate detection techniques (possibly by combining existing approaches)
- Empirical assessment of their effectiveness

Thank you for your attention