Master thesis specification

Web elements and events identification algorithms for e-learning solutions

Context

Lemon Learning [1] integrates interactive guides into web pages and SaaS. Those guides are made to bring the user step by step to realize a precise task, by indicating what actions has to be made through help tooltips.

The Lemon guides are generated by an author tool developed internally. This tool is aimed at users with no particular computer experience.
In order for the guides to perform well, the Lemon tool must be able to

- Identify the element on which the tooltip points
- Identify an event for passing to the next step

The project consists in identifying in an unique and robust way an element in a web page.

The Lemon module aims at being integrated onto all SaaS platforms, it must then be able to identify elements and events in a precise and unique way on SaaS of very different structure, with variable user contexts, where the DOM does not always have the same structure.
This identification must be automatic, because the user is not supposed to have any computer science knowledge.

To solve this problem, Lemon Learning needs an algorithm generating one or several paths having the following characteristics:

- Accurate enough to identify the wanted element and not another one
- Flexible/robust enough to be able to adapt to different user contexts.

For this, two approaches can be taken:

- Creation of a unique algorithm, allowing to create a path in all the pages, by using the HTML attributes of the elements.
- On each platform, an algorithm generating different paths is created by machine learning.

Research in the area:

There are several small open source projects tackling this problem, but most of them are not sturdy enough to be applied to our use (constant evolution of the SaaS and user contexts).
Some use the most frequent HTML/CSS attributes (id, class, titles), other are configurable to use predefined attributes.
Some are completely automated, others need a step by step construction by the user.
Some use like Lemon Learning the Sizzle library (included in jQuery), which allows the transversal selection of DOM elements (jQuery selectors).

Lemon Learning has been developing its edition tool for 1 year, and has continually done research on the topic, but the results are insufficient to this day.
The system is based on the creation of several alternative paths evaluated one after the other.
Multiple leads are considered, for example using applied CSS rules, using elements possessing and eventHandler, using an system of points per attribute, or using machine learning.
Concerning event detection, Lemon Learning has not done a lot of research in the area, and the state of the art concerning it is very limited.

Methods

The thesis will consist in finding additional literature and using the already found one to decide on which approach will be the most efficient for the identification problem. This means a lot of research on internet based on what Lemon Learning has already found and is considering but also additional personal initiatives.

The next step will be to create a complete test process to test the efficiency and robustness of the algorithms. An 80% robustness will mean that for 100 displays of a page, for different users and at different times we will have found the right element 80 times. This will involve coding mainly in javascript and choosing the right testing framework.

Then the implementation of one or several algorithms for path generation (be it unique or by machine learning). Also done in javascript with probably a heavy use of jQuery.

Finally the creation of an alternative version of the Lemon creation tool.

Preliminary time plan

Week	Moment
1-3	Literature research, list and analyse the different possibilities and algorithms for the path creation.
4-5	Analyse in depth the one or two chosen methods and the possible test processes. Create the report architecture.
6-9	Create and use the test processes. Write the method part of the report.
9-13	Implement the algorithms for path generation.
13-16	Create the alternate version of the Lemon creation tool.
16-18	Fix bugs and rerun tests.
19-20	Finish report and prepare presentation.

Reviewer

Lars Oestreicher

Supervisor

Olivier Praseuth

Relevant courses

User Interface Programming II
E-commerce project
Advanced Interaction Design
Artificial Intelligence
Algorithms and Data Structures II
Project CS

References

[1] Lemon Learning. Accessed 28 November 2016, [Online]. Available: http://lemonlearning.fr/en/.

Preliminary time plan

Week	Moment
1-3	Literature research, list and analyse the different possibilities and algorithms for the path creation.
4-5	Analyse in depth the one or two chosen methods and the possible test processes. Create the report architecture.
6-9	Create and use the test processes. Write the method part of the report.
9-13	Implement the algorithms for path generation.
13-16	Create the alternate version of the Lemon creation tool.
16-18	Fix bugs and rerun tests.
19-20	Finish report and prepare presentation.

Reviewer

Supervisor

Olivier Praseuth

Relevant courses

User Interface Programming II
E-commerce project
Advanced Interaction Design
Artificial Intelligence
Algorithms and Data Structures II
Project CS

References

[1] Lemon Learning. Accessed 28 November 2016, [Online]. Available: http://lemonlearning.fr/en/.