This research project is based on formally proving mathematical properties of regular languages from the theory of automata. In the literature, automata theory first appeared as an interdisciplinary field, comprising notions from mathematics, computer science and electrical engineering. This field is developed on the model of "abstract machines" called automatas, that are able to compute values given an input represented by words. Also, different variants of them (Turing Machines) are extensively used in the decision whether a problem can be solved in reasonable amount of time (NP-completeness problems).

The main goal of the project is to develop a mathematical formalization of properties from regular languages by using the Coq dependently-typed proof assistant. These notions are highly relevant to pure mathematics, artificial intelligence, linguistics, etc.

A *language* is defined to be a set of words. Words are formed as lists of symbols; a collection of finite symbols define an alphabet. We will give an overview of the construction of finite sets in type theory, by defining products, coproducts, existence operator dealing with finite sets, isomorphisms between structures.etc.

The *automaton* is the machine on which *words* are given as input. There exist several properties of whether an input is accepted , this model giving birth to interesting mathematical results, some of them being tackled in this project.

A *regular language* is formed by the words that are recognized by an automaton. In our model, we develop deterministic and non-deterministic automatas and reason about them. Regular languages can also be expressed by a special kind of formulas, called *regular expressions*. We want to show that there exists an equivalence of finite automata and regular expressions. Also, we may represent the closure properties of regular languages (substitution, homomorphism, quotients).

We will prove a property used in showing that a language is not regular, i.e the pumping lemma. Another useful result that we may represent in our formalization is the Myhill=Nerode theorem. This represents a necessary and sufficient condition for a language to be regular. Also, it implies that there exists a unique minimal DFA with minimum number of states. From this, we can use it in order to verify the table-filling algorithm from the automata minimization process.

Based on how long the work will take, we will like to prove some mathematical properties from 2-way finite automata. [to be decided]

So far (completed)

1) Languages and words : introduction, union, concatenation, properties , kleene star operations, etc.
2) Finite sets construction – coproducts, products, associated laws, existence operator for finite sets
3) Automata definition: DFA , NFA, extended transition functions
4) Conversion from an Nfa to dfa (subset construction), and viceversa
5) Implementation of converting a regular expression to NFA (definitions, lemmas)

6) Pumping lemma – definition, proving of properties relevant to the decomposition of the words, proving lemma using a DFA and then, using the equivalence between the language of regular expressions, we may state that the lemma is valid.

In progress (partially completed):

- (a)finish correctness lemma for proving the correctness of converting a regular expression to an NFA under a star operator form

- b)finish the general correctness lemma (using a) that given a regular expression, its language is equivalent to the language of the NFA

-using b), state that the language of a DFA needed in our auxiliary pumping lemma is equivalent to the one of the regular expression, i.e the language is regular.

- attempted work on Myhill Nerode theorem, implemented some properties (right invariant, finite equivalence classes, etc)

To do:

- Verification of the table filling algorithm
- Language equivalence between regular expressions is decidable
- Derivatives of regular expressions [to be decided]
- Verify properties on 2-DFA [to be decided]
- Buchi automatons
- Ramseyan factorizations on words (?)
- Use category theory to model some of the work done [to be decided]