Advancing Petri Net Theory: Deeper understanding to Classical Problems and Integration into Workflow Applications

Petri nets, developed over 50 years ago, serve as a powerful mathematical modeling tool, providing insights into the dynamics of concurrent systems. Petri nets can be conceptualized as graphs with tokens that move around, vanish, or are created according to specific rules. The most fundamental problem for Petri nets is known as reachability. Intuitively, we inquire whether there is a way to move tokens in such a manner that we can transition from the initial configuration to the final configuration. Only recently, the worst-case complexity of the reachability problem has been established. This result has been widely celebrated. Despite this, we still lack an understanding of how the graph's topology influences the complexity of the reachability problem. Many studies have been conducted, revealing various topologies that make the reachability problem easy. However, the robustness of these subclasses remains unclear. One of the primary objectives of this project is to address this knowledge gap.

One may wonder why this research is relevant to the world. As previously mentioned, Petri nets and their extensions are among the most successful tools for modelling concurrent systems. They play a crucial role in Business Process Management, serving as a cornerstone for many methods and tools. In this context, tokens represent people, objects, or documents that change their state when actions are performed. Common questions posed by business analysts include:

1. Is it possible that, in some rare cases, documents circulate endlessly in the office without reaching the director's desk?
2. Is there a node in the workflow through which almost all cases pass, limiting the efficiency of case processing?
3. What is the average time for processing a case through the workflow?
4. Why do 30% of cases end in failure?
5. and so on.

In business process analysis, analysts do not rely solely on models; they also have access to *event logs*, which are extensive files containing sequences of recorded events. Event logs can be used to augment the model, they can be replayed to observe bottlenecks and the actual performance of the system. However, in the complex reality, some cases may not fit the model. These undesired behaviors could result from errors in the log, mistakes in the model, or a combination of both. Errors in the log may result from faulty sensors, mistakes in the logging system, or human errors, as in many cases, part of logs is created by humans, for example, patient data from hospitals. The inaccurate models can be created by experts or generated by some algorithm, as the best possible explanation of the so far observed logs.

Various approaches to locating errors have been proposed, with the most successful among them based on the concept of alignment. To align two words, we seek the minimal number of edit operations required to transform one word into another. The current consensus is that the fundamental operations include removing a letter, adding a letter, and exchanging letters. Interestingly, the same concept is used in spellchecking, where the goal is to align a word from the text (a sequence of letters) with words from a dictionary.

In the workflow environment, we align the model that may represent an infinite number of possible behaviours with a sequence from the log. Thus, the alignment problem here is more complex than alignment in spell-checking, but algorithms do not need to provide an immediate answer.

Many methods are used to solve the alignment problem, and none of them is perfect. Surprisingly, some of these methods rely on techniques used to decide reachability for Petri nets. In this project, we will also fine-tune known novel techniques for reachability and apply them to the alignment problem. Progress in this area will impact tool performance and reduce the cost of business process improvement, ultimately improving the quality of our lives.