Improving Interpretability in Symbolic Regression Models Using Multi-objective GP and the Transient Terminal Set

Asher Stout, 300432820

December 14, 2020

1 Overview of Work & Research

Work began on the project with research into previous approaches for improving the interpretability of Symbolic Regression models, and ultimately Multi-objective GP (MOGP) was selected as an area for further study. Several relevant papers defined the interpretability of a solution as the depth or size of its tree, which may not accurately correlate to its computational complexity. Thus, it was decided that a new measure of complexity be developed for use in MOGP. In addition, the entirely random nature of mutation operators in MOGP was identified as place for potential innovation for improving Symbolic Regression model interpretability. This resulted in a new approach being developed called the **Transient Terminal Set**, which is described in further detail in **Section 2**.

Aside from research, extensive time was invested into learning the tools required for the completion of the project objectives - particularly Python, DEAP, and LaTeX. This involved practical applications in the form of tutorials and implementing a standard MOGP and culminated in the formal definition of the Transient Terminal Set algorithm (see **Section 2**) and its implementation using DEAP (see **Section 3**).

The total hours spent on the project, as of 15 December 2020, were **130**.

2 Transient Terminal Set

3 Experimentation Results

Note: The code for this experiment can be accessed at `https://github.com/VeryEager/transient-terminal-gp`

4 Intended Future Work