Abstract

We present a self-learning approach for synthesizing programs from integer sequences. Our method relies on a tree search guided by a learned policy. Our system is tested on the On-Line Encyclopedia of Integer Sequences. There, it discovers, on its own, solutions for 27987 sequences starting from basic operators and without human-written training examples.

1 Introduction

The search for abstract patterns is one of the principal occupations of mathematicians. The discovery of similar patterns across different mathematical fields often leads to surprising connections. Probably the most famous example of such an unexpected connection in mathematics is the Taniyama–Shimura conjecture proved in 2001 (Breuil et al. 2001). It relates elliptic curves over the field of rational numbers with a special kind of complex analytical functions known as modular forms. This conjecture became especially famous because a restricted version of it implies Fermat’s last theorem (Wiles 1995). The connections found by the system described in this paper are more modest. For instance, it has created formulas for testing prime numbers based both on Fermat’s little theorem\(^1\) (stated in 1640) and Wilson’s theorem\(^2\) (stated in 1770). When a symbolic representation (e.g. formula) describing a pattern is conjectured and preferably proven, a mathematician can start reasoning and deriving additional facts about the theory in which the pattern occurs.

Integer sequences are a very common kind of mathematical patterns. A compilation of such sequences is available in the On-Line Encyclopedia of Integer Sequences (OEIS) (Sloane 2007). Our objective in this project is to create a system that can discover, on its own, programs for the OEIS sequences. Such programs will be chosen to be small since usually the best explanations for a particular phenomenon are the simplest ones. This principle is known as Occam’s razor or the law of parsimony. It has been one of the most important heuristics guiding scientific research in general. In a machine learning setting, this can be seen as a form of regularization. A mathematical proof of this principle relying on Bayesian reasoning and assumptions about the computability of our universe is provided in Solomonoff’s theory of inductive inference (Solomonoff 1964).

The programs produced by our system are readable, making it possible for mathematicians to gain insight into the nature of the patterns by analyzing the programs.\(^1\) We believe that in the future, such systems will assist mathematicians during their conjecture-making process when investigating open problems.

1.1 Overview

Our approach is to synthesize programs for the OEIS sequences (Section 2) in a simple domain-specific language (Section 3) in many iterations of the following self-learning loop (Figure 1). Each iteration (generation) of the loop consists of three phases: a generating phase (Section 4), a testing phase (Section 5) and a training phase (Section 6). Initially, multiple searches are randomly building programs that generate integer sequences and checking if they are in the OEIS. Then, for each generated OEIS sequence we select the smallest program that generates it. From those solutions, a tree neural network is trained to predict what the right building action is, given a target sequence and a partially built program. The next searches are then guided by the statistical correlations learned by the network, typically producing further solutions. We describe the experiments in Section 7 and analyze some of the solutions in Section 8. Section 9 discusses related work.

\(^1\)See our web interface (Gauthier and Urban 2022b).
As of March 2022, the number of OEIS entries reached 10,000. The interpretations of operators: cond, loop, loop2, compr are inductively constructed. Higher-order arguments of looping operators can depend on the allocated time and memory. The two other looping operators loop and loop2 can also be defined with recurrent relations:

\begin{align*}
\text{loop}(f, a, b) &:= u_n \text{ where } u_0 = b, u_n = f(u_{n-1}, n) \\
\text{loop2}(f, g, a, b, c) &:= u_n \text{ where } (u_0, v_0) = (b, c) \text{ and } \\
&\quad (u_n, v_n) = (f(u_{n-1}, v_{n-1}), g(u_{n-1}, v_{n-1}))
\end{align*}

The size of a program is measured by counting the number of operators composing it. This number can be obtained by counting the number of tokens in the expression ignoring binders (\(\lambda(x, y)\)), commas and parentheses. For example, the program \(\text{loop}(\lambda(x, y), x \times y, x, 1)\) has size 6.

### Generating Programs

To generate programs, we rely on multiple searches. Each of them targets a randomly chosen OEIS sequence \(s\) and constructs a set of programs intended to generate \(s\). First, we describe our bottom-up process for constructing a single program. Then, we explain how this method can be extended to synthesize multiple programs at once by sharing construction nodes in a policy-guided tree search.

#### Program Construction

Programs are built following their reverse polish notation. Starting from an empty stack, a learned policy is used to repeatedly choose the next operator to push on top of a stack. This policy is computed by a tree neural network (Section 6) based on the target sequence \(s\) and the current stack \(k\). For instance, the program \(\text{loop}(\lambda(x, y), x \times y, x, 1)\) can be built from the empty stack by the following sequence of actions:

\[
[\text{[]} \xrightarrow{\text{loop}} [x] \xrightarrow{\text{cond}} [x, y] \xrightarrow{\text{loop}} [x \times y] \xrightarrow{\text{loop}} [x \times y, x, 1]
\]

The stack here starts as an empty list, growing up to length 3 after the fifth action. Then the final \text{loop} action reduces it (\(y_0\), i.e., these stand for \(a(x_0, y_0), a(x_0, y_0) - 1(x_0, y_0), \text{and } cond(a(x_0, y_0), b(x_0, y_0), c(x_0, y_0))\)).
As an example, the program loop we evaluate (test) them on an initial segment of natural numbers (OEIS sequence A142

ters (or \(p\)) of those where \(x \mod y\) is run until a given timeout and we collect all the programs generated during the search for the next phase (see Section 5).

to a stack of length 1 containing the final program. Note that the higher-order actions such as loop include the creation of \(\lambda\)-terms (here \(\lambda(x,y). x \times y\)) as the appropriate functional arguments.

Tree search The search process is initialized with an empty stack at the root of our search tree (Figure 3). Starting from this root, actions are selected according to the action probabilities returned by the policy predictor (Section 6) for the current node until an action whose output does not yet appear in the tree is selected. Then, a node is created containing the result of that action. After the creation of a new node, the selection process is restarted from the root. The search is run until a given timeout and we collect all the programs generated during the search for the next phase (see Section 5).

To speed up the search, we cache embeddings (see Section 6) for all sub-trees in the target sequence and in the program stack. An illustration of the gradual extension of the search tree is given in Figure 3.

5 Testing Programs

To see if the synthesized programs produce OEIS sequences, we evaluate (test) them on an initial segment of natural numbers. To do that, we first remove all (potentially) binary programs, i.e., those where \(y\) occurs free (not under \(\lambda\)), from the set of programs collected during the search(es). Then, for each remaining program \(p\), we generate its sequence \(p(0), p(1), p(2), \ldots\). Formally, we say that a program \(p\) covers (or is a solution for) a finite sequence \((s_x)_{0 \leq x \leq n}\) if and only if:

\[
\forall x \in \mathbb{Z}. 0 \leq x \leq n \Rightarrow p(x) = s_x
\]

As an example, the program \(\text{loop}(\lambda(x,y). x \times y, x, 1)\) (i.e. \(x_0 = 1\) and \(x_n = u_{n-1} \times n\)) covers any finite sequence of factorial numbers (OEIS sequence A142\(^3\)).

Finally, for each OEIS sequence \(s\), we select the shortest program \(p\) among its solutions. If there are multiple programs with the smallest size, we break the ties using a fixed total order. After that, training examples are extracted from such selected sequence-solution pairs, see Section 6.

Note that during the search targeting a particular sequence \(s\), programs for many other OEIS sequences may be generated. These programs can be considered as positive examples together with their respective OEIS sequences. This process is a form of hindsight experience replay (Andrychowicz et al. 2017) where the failed attempts that lead to other OEIS targets are used for training.

6 Training from Solutions

A tree neural network (TNN) (Goller and Küchler 1996) is used as our policy predictor since its dynamic structure naturally matches the tree structure of our programs. This machine learning model is both sufficiently fast for our purpose and has been shown to perform well on several arithmetical tasks (Gauthier 2020), outperforming several other models including the NMT (Luong, Pham, and Manning 2015) recurrent neural toolkit.

Definition A TNN projects labeled trees into an embedding space \(\mathbb{R}^d\). Given a tree \(t = f(t_1, \ldots, t_n)\), an embedding function \(e : \text{Tree} \mapsto \mathbb{R}^d\) can be recursively defined by:

\[
e(f(t_1, \ldots, t_n)) = \text{def} N_f(e(t_1), \ldots, e(t_n))
\]

where \(N_f\) is a function from \(\mathbb{R}^{n \times d}\) to \(\mathbb{R}^d\) computed by the neural network building block associated with the operator \(f\).

Training examples The basis for creating the set of the training examples are the covered OEIS sequences and their shortest discovered programs. In general, each sequence \(s\) and its program \(p\) generate multiple training examples corresponding to the states and actions involved in building \(p\).

In more detail, each such training example consists of a pair \((\text{head}(k, s), \rightarrow_a)\). The term \(\text{head}(k, s)\) joins a stack \(k\) and a sequence \(s\) into a single tree with the help of an additional operator \(\text{head}\). And \(\rightarrow_a\) is the action required on \(k\) to progress towards the construction of \(p\). From the solution \(\text{loop}(\lambda(x,y). x \times y, x, 1)\) (Figure 4) we can extract one training example per action leading to its creation. One such example is \(((\text{head}([x \times y, x], 1, 1, 2, 6, \ldots)), \rightarrow_1)\).

Our TNN is actually predicting a policy distribution over all the actions. So, the action \(\rightarrow_1\) is projected to \(\mathbb{R}^d\) using a one-hot embedding. This essentially tells the network that every other action on the same stack and for the same target sequence is a negative example. The computation flow for the forward pass of this training example is depicted in Figure 4.

Training The TNN then learns to predict the right action in each situation by following the batch gradient descent algorithm (Li et al. 2014). Each batch consists of all the examples derived from a particular sequence-solution pair. To speed up the training within a batch, subprograms as well as the numbers in the common sequence are shared in the computation tree.

Noise To encourage exploration, half of the searches are run with some noise added to the (normalized) policy distribution \(q\) returned by the TNN. The noisy policy distribution

---

\(^3\)The sequence is at https://oeis.org/A000142. Its program can be found with our web interface (Gauthier and Urban 2022b) and input 1, 2, 6, 24 at http://grid01.circ.cvut.cz/~thibault/cgi-bin/qsynt_v3.cgi?Formula=1,1,2,6,24.
Embedding integer sequences  Integers are embedded in base 10 with the help of a digit concatenation operator (::d) operator and a unary minus (−u) operator. For example, the number -159 is represented by the tree \(-u(9::d(5::d1))\).

In order to speed up training, we only embed the first 16 elements of the targeted sequence. Moreover, integers larger than 10^6 are replaced by a single embedding given by a new nullary operator big. For integers smaller than \(-10^6\), the embedding of the tree \(-u(big)\) is computed. With these restrictions, 91.6% of OEIS sequences can in theory be uniquely identified by their embeddings.

7 Experiments
Our experiments are performed on the full OEIS. The target sequence for each search is chosen uniformly at random from this dataset.

Each of these experiments is run on a server with 32 hyperthreading Intel(R) Xeon(R) CPU E5-2698 v3 @ 2.30GHz, 256 GB of memory, and no GPU cards. The operating system of the server is Ubuntu 20.4, GNU/Linux 5.4.0-40-generic x86_64.

The code for our project is publicly available in our repository (Gauthier and Urban 2022a). The repository contains a full list of all the solutions found by our algorithm during the main self-learning run. A web interface is provided for demonstration purposes (Gauthier and Urban 2022b). In the following, we present the chosen hyperparameters for each phase.

7.1 Hyperparameters

Search parameters  Each search phase is run in parallel on 16 cores targeting a total of 160 targets. On each target, a search is run for 10 minutes (2 minutes for side experiments – cf. Section 7.2).

Test parameters  A program \(p\) is tested (evaluated) with an initial time limit of 50 microseconds. Additional 50 microseconds are added to the timeout each time \(p\) generates a new term. This means that a program \(p\) times out if it takes more than a millisecond (= 20 \times 50 microseconds) to generate the first 20 terms. The execution also stops when an integer with an absolute value larger than 10^{285} is produced. This bound was chosen to be larger than the largest number (in absolute value) in the OEIS, which is approximately equal to 4.685 \times 10^{284}.

Training parameters  The TNN embedding dimension \(d\) is chosen to be 64. Each neural network block consists of two fully-connected layers with \text{tanh} activation functions. To get the best performance, our TNNs are implemented in C with the help of the Intel MKL library (Wang et al. 2014). The values for these parameters were optimized over multiple self-learning runs. The variation in the number of solutions with respect to multiple parameter values was taken into account to select the best parameters. These experiments were typically stopped before the 5th generation and ran with a timeout of 2 minutes per search.

7.2 Side Experiments
We have identified four sources of randomness that can affect the results of our experiments:

- the random selection of the target sequences,
- the initialization of the TNN with the random weights,
- the random noise added on top of the policy in half of the searches,
- and the fluctuations in the efficiency of the server (interacting with the time limits).

In our side experiment (E1), we measure the degree to which they influence our results. In particular, we run a self-learning run for 5 generations three times with the same parameters. The differences in the numbers of solutions are reported in Table 1. The worst performing run gives 4% fewer solutions than the best performing run, demonstrating the robustness of the system. In our side experiment (E2), we evaluate the effect of selecting random solutions instead of the smallest ones during the test phase. This experiment is also run for 5 generations and its results are included in Table 1 for a comparison. We observe a decrease in the number of solutions by about 10 percent when compared with the worst performance of the default selection strategy.

Thus, we can say that selecting the smallest solutions instead of the random ones for training helps finding new solutions in later generations. This experiment also shows how efficient this additional application of Occam’s razor is in
Table 1: Side experiments: solutions at each generation.

<table>
<thead>
<tr>
<th>Generation</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment (E1)</td>
<td>280</td>
<td>4317</td>
<td>6970</td>
<td>8724</td>
<td>10139</td>
</tr>
<tr>
<td></td>
<td>259</td>
<td>4624</td>
<td>7244</td>
<td>8771</td>
<td>9767</td>
</tr>
<tr>
<td>Experiment (E2)</td>
<td>310</td>
<td>4905</td>
<td>7275</td>
<td>8570</td>
<td>9747</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solutions</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

Figure 5: Solutions found after $x$-th generation

Figure 6: Number of solutions with size $x$

7.3 Full-Scale Self-Learning Run

Finally, we perform a full-scale self-learning run with a larger timeout of 10 minutes per search and for a larger number of generations. The cumulative numbers of solutions found after each generation are shown in Table 2 and plotted in Figure 5.

The unguided generation 0 discovers only 993 solutions. After learning from them, the system discovers additional 8438 solutions in generation 1. There are two reasons for this effect. First, the learning globally biases the search towards solutions of the sequences that are in the OEIS. Second, after the first learning, the TNN takes into account which sequence is being targeted. Therefore each TNN-guided search explores a different part of the program space.

At generation 25, 187 new solutions are constructed which is still more than the 160 targeted sequences. Thus, it is clear that hindsight experience replay is crucial in our setting. Note that due to the bottom-up nature of our tree search, smaller solutions are already more likely to be constructed first. Therefore the randomly selected solutions are already relatively small. In other words, our explicit application of Occam’s razor (selecting the smallest solutions) is combined with an implicit bias towards Occam’s razor given by the nature of our search.

General evaluation metrics To evaluate progress relative to other future competing works, we propose to judge the quality of a synthesis system based on four criteria:

- The number of sequences fully generated: 27987.
- The average size of the solutions (smaller is better): 17.4.
- The simplicity of the language: 13 basic operators.
- The computation power required: less than 100 CPU days (no GPUs were used).

We are not aware of any other self-learning experiment on the OEIS that provides all this information for comparison. Further discussion of related works is provided in Section 9. The distribution of programs according to their sizes is presented in Figure 6. The long tail of this distribution is not shown in this figure. The largest solution has size 455.

Generalization to larger inputs For each sequence $seq$ in the OEIS, we denote by $n_{seq}$ the number of terms in the standard OEIS dataset used in our experiments. The frequency of each sequence length in this dataset is given in Figure 2. To test the generalization of our solutions to larger inputs, we look for additional terms for each sequence. For many OEIS sequences, such information can be found in b-files provided by the OEIS website (Sloane 2007). We verify further 12639 solutions that have an extra 100 terms in their corresponding b-file with a maximum absolute value of $10^{285}$ for each term. The proportion of sequences generated by those solutions, that covers $x$ additional terms, is plotted in Figure 7. More than 92% of the generated sequences match their intended counterpart on the first $n_{seq} + 100$ terms. This result confirms the premise of our introduction that small programs provide good explanations for the particular case of mathematically interesting integer sequences.

Testing alone however can not guarantee that our synthesized programs produce their corresponding OEIS sequences, as defined by their English descriptions, for all inputs $x \in \mathbb{N}$. In the following analysis (Section 8), we will show that this is the case for a few selected sequences.

8 Analysis of the Solutions

To understand the abilities of our system, we analyze here some of the programs discovered by the full-scale self-learning run (Section 7.3).
As a measure of how important a program $p$ is for the construction of other programs, we also compute the number of occurrences of $p$ in other programs (with multiple occurrences per program counted only once). This number is below indicated in brackets between the OEIS number (A-number) and its OEIS description.

The programs can also be searched for and analyzed using our web interface (Gauthier and Urban 2022b). Its typical use is to enter the initial terms of a sequence, adjust the search parameters, explore the best matching OEIS sequences, and present the programs found both in our language and after a translation to Python. The Brython\footnote{https://brython.info/} interactive editor can then be used to explore, modify and immediately run the discovered programs. For some of the solutions, we additionally write their expressions using recurrence relations (in the next bullet after their native formulations) in order to facilitate their analysis.

1. **Solution for A1146** [60 occurrences], $s_x = 2^{2^x}$.
   - loop$(\lambda(x,y), x \times x, x, 2)$
   - $u_x$ where $u_0 = 2$, $u_n = u_{n-1} \times u_{n-1}$

   Unexpectedly, only a single loop was needed to create this fast increasing function.

2. **Solution for A45** [134 occurrences], the Fibonacci sequence:
   - loop$2(\lambda(x,y), x + y, \lambda(x,y), x, x, 0, 1)$
   - $u_x$ where $u_0 = 0$, $u_1 = 1$, $u_n = u_{n-1} + u_{n-2}, v_n = u_n$

   This implements an efficient linear algorithm, whereas the natural implementation of the recurrence relation $u_n = u_{n-1} + u_{n-2}$ is exponential.

3. **Solution for A312** [74 occurrences], $s_x = x^x$:
   - loop$2(\lambda(x', y'), x' \times y', \lambda(x', y'), y', x, 1, x)$
   - $u_x$ where $u_0 = 1$, $u_1 = x$, $u_n = u_{n-1} \times v_{n-1}, v_n = v_{n-1}$

   This example illustrates the use of the local variable $y'$ as a storage for the top level variable $x$. The value of $y'$ (or $v_n$) is constant throughout the loop and equal to $x$.

4. **Solution for A108** , the Catalan numbers $\left(\frac{2^x}{x+1}\right)$:
   - loop$(\lambda(x,y), 2 \times (x - x \text{ div } y + y), x, 1) \text{ div } (1 + x)$
   - $u_x \text{ div } (1 + x)$ where $u_0 = 1$ and $u_n = 2 \times (u_{n-1} - u_{n-1} \text{ div } n + u_{n-1})$

   We prove by recurrence on $x$ that this solution generates the Catalan numbers. Since the denominators are equal, we only need to prove that $u_x = \left(\frac{2^x}{x+1}\right)$. The base case is true since $1 = \left(\frac{2^0}{0+1}\right)$. For the inductive case, we assume that $u_x = \left(\frac{2^x}{x+1}\right)$ and show that:

   $$u_{x+1} = 2(u_x - \frac{u_x}{x+1} + u_x) = \frac{2(2x+1)}{x+1}u_x =$$

   $$\frac{2(2x+1)}{x+1}\frac{(2x)!}{x!^2} = 2(x+1)\frac{(2x+1)!}{(x+1)!^2} =$$

   $$\frac{2(x+1)}{2x+2} = \frac{2(x+1)}{x+1} = 2 \left(\frac{2^x}{x+1}\right)$$

5. **Solution for A10051** [28 occurrences], the characteristic function of the primes:
   - $(\lambda(x,y), x \times y, x, x) \text{ mod } (1 + x)$\text{ mod } 2$
   - $(u_x \text{ mod } (1 + x)) \text{ mod } 2$ where $u_0 = x, u_n = u_{n-1} \times n$

   We can now prove the following conjecture relating the discovered and reformulated formula with the prime characteristic function $1_p$:

   $$\forall x \in \mathbb{N}, (\text{mod } (1 + x)) \text{ mod } 2 = 1_p(x)$$

   This conjecture is a variation of Wilson’s theorem and its proof reveals how the generated formula is structured. We reason modulo $1 + x$, saying that $a \equiv b$ if $a$ and $b$ are equal modulo $1 + x$. We prove the conjecture by considering four cases:

   - If $1 + x$ is prime, then every non-zero element has an inverse in the $\mathbb{Z}_{1+x}$, and the only elements that are their own inverses are $1$ and $-1 \equiv x \mod 1 + x$. Indeed, a field has a maximum of two solutions for the equation $X^2 \equiv 1 \mod 1 + x$ according to the fundamental theorem of algebra. Thus, the elements of the product $x!$ can be regrouped into pairs of inverses except for $x$, hence $x \times x! \equiv x \times x \equiv 1 + x$.

   - If $1 + x$ is not prime and is not the square of a prime, then $x!$ divides $1 + x$ since $1 + x$ has two distinct proper divisors.

   - If $1 + x = p^2$ where $p$ is an odd prime, then $(p^2 - 1)!$ divides $p^2$ since $p$ and $2p$ appear in $(p^2 - 1)!$.

   - If $1 + x = 4$, then $3 \times 6 \equiv 2$. Only in this case, the final modulo 2 operation is necessary.

From this primality test, the operator \texttt{compr} can construct the set of prime numbers. However, this construction fails to pass the testing phase because it is too inefficient. The creation of an efficient prime number generator is an open problem at the intersection of mathematics and computer science that we were not expecting to solve. In a previous run, a slightly more efficient but inaccurate way of...
generating prime numbers was re-discovered by synthesizing the Fermat primality test in base 2. The generated sequence of pseudoprime values deviates from the prime number sequence. Indeed, the number 341 passes Fermat’s primality test but is not prime.

6. **Solution for A46859**, \(1 \times 3 \times 7 \times 61\), “simplified” Ackermann:

\[
\text{loop}(\lambda(x, y). \text{loop2}(\lambda(x, y). x \times y, x, x, 2), 2) - x, x, 1)
\]

Our pseudo-solution is primitive recursive while Ackermann is not. The next number this solution generates is 13114940639623 shy of the expected \(2^{2^{34}} - 3\) too large to be included in the OEIS. Generally, our system is likely to give the “wrong” solution when sequences have very few terms.

7. **Solution for A272298**, \(x = x^2 + 324\):

\[
\text{loop}(\lambda(x, y). x \times x, 2, 1 + 1 + 1 + \ldots + 1)
\]

This is the longest solution generated by our system. It has size 654 as the number 324 is constructed by repeatedly adding 1. A shorter way of expressing the constant 324 is found before the end of our self-learning run:

\[
\text{loop}(\lambda(x, y).x \times x, 2, 1 + 2) \times (2 + 2) = 3^4 \times 4
\]

8. **Solution for A66298**, \(s_x = \text{googol} \ (\mod \ x)\):

\[
\text{loop}(\lambda(x, y). \text{loop}(\lambda(x, y). (2 + (2 + 2)) \times x, x, 1), 2, 2) \mod (1 + x)
\]

It contains the largest constant \(10^{100}\) used in our solutions. After recognizing the sub-expression for \(10^2\), this program can be rewritten as \(10^{10^2} \mod (1 + x)\). This program uses \(1 + x\) instead of \(x\) since the sequence starts at \(s_1\) in the OEIS.

9. **Solution for A195**, \(s_x = \text{floor}(\log(x))\):

\[
((\text{loop}(\lambda(x, y). (y \div \text{loop2}(\lambda(x, y). x + y, \lambda(x, y). x, x, 1, 1))
+ x, 1 + x, 1 - 1) \mod (1 + x)) \div 2
\]

Among the 133 solutions found at generation 18, this is the smallest. It has size 25. Further work is required to check if the last two presented solutions are correct. In the future, we will aim to prove that the intended solution and the synthesized one are equal with the help of interactive and automatic theorem provers. Solutions discovered during the full self-learning run are available for further analysis in our repository (Gauthier and Urban 2022a).

9 **Related Work**

The closest related recent work is (d’Ascoli et al. 2022), done in parallel to our project. The goal there is similar to ours but their general approach is focused on training a single model using supervised learning techniques on synthetic data. In contrast, our approach is based on reinforcement learning:

\[\text{https://oeis.org/A015919}\]

\[\text{https://oeis.org/A066298}\]

\[\text{https://oeis.org/A046859}\]

\[\text{https://oeis.org/A000195}\]

\[\text{https://oeis.org/A015919}\]

we start from scratch and keep training new models as we discover new OEIS solutions.

The programs generated in (d’Ascoli et al. 2022) are recurrence relations defined by analytical formulas. Our language seems to be more expressive. For example, our language can use the functions it defines by recurrence using loop and loop2 as subprograms and construct nested loops. To fit the Transformer architecture, (d’Ascoli et al. 2022) represent programs as a sequence of tokens, whereas our tree neural network does not need such transformations. The performance of the model in (d’Ascoli et al. 2022) is investigated on real number sequences, whereas our work focuses only on integer sequences. Overall, it is hard to directly compare the performance of the two systems. Our result is the number of OEIS sequences found by targeting the whole encyclopedia, whereas (d’Ascoli et al. 2022) report the test accuracy only on 10,000 easy OEIS sequences.

The inspiration for many parameters of our self-learning system DreamCoder (Ellis et al. 2021) has demonstrated self-improvement from scratch in various programming tasks. Its main contribution is the use of definitions that compress existing solutions and facilitate building new solutions on top of the existing ones. As seen from the experiments in (Ellis et al. 2021), adding definitions typically works up to a point. After that point, the extra actions are hurting significantly the chance of constructing a program that does not need them.

10 **Conclusion**

Our system has created from scratch programs that generate the full list of available terms in the OEIS for 27987 sequences. Based on *Occam’s razor*, we have argued that producing shorter programs is more likely to generate better explanations for particular sequences. We have also shown that the solutions discovered are correct for some famous sequences. And we have observed that preferring shorter programs during the training increases the performance of the system.

In the future, we would like to create a benchmark of theorem proving problems from the solutions found in our experiments. There, automatic theorem provers would be tasked to prove that a particular intended definition for a sequence (e.g. the prime characteristic function \(1_P(x)\)) is equivalent to the program that we have discovered (e.g. \((x \times x! \mod (1 + x)) \mod 2\)). We believe that such a non-synthetic benchmark would contribute greatly to the development of automated inductive reasoning.
Acknowledgments

This work was partially supported by the CTU Global Postdoc funding scheme (TG), Czech Science Foundation project 20-06390Y (TG), Czech MEYS under the ERC CZ project POSTMAN no. LL1902 (TG), Amazon Research Awards (TG, JU), EU ICT-48 2020 project TAILOR no. 952215 (JU), and the European Regional Development Fund under the Czech project AI&Reasoning no. CZ.02.1.01/0.0/0.0/15_003/0000466 (JU).

References


