Dynamic Algorithm Portfolios

Matteo Gagliolo¹ and Jürgen Schmidhuber¹,²

¹IDSIA, Galleria 2, 6928 Manno-Lugano, Switzerland
²TU Munich, Boltzmannstr. 3, 85748 Garching, München, Germany
{matteo,juergen}@idsia.ch

Abstract. Traditional Meta-Learning requires long training times, and is often focused on optimizing performance quality, neglecting computational complexity. Algorithm Portfolios are more robust, but present similar limitations. We reformulate algorithm selection as a time allocation problem: all candidate algorithms are run in parallel, and their relative priorities are continually updated based on runtime information, with the aim of minimizing the time to reach a desired performance level. Each algorithm’s priority is set based on its current time to solution, estimated according to a parametric model that is trained and used while solving a sequence of problems, gradually increasing its impact on the priority attribution. We present new experimental results for this recent approach.

1 Motivation

Most solvable AI problems can be addressed by more than one algorithm; most AI algorithms feature a number of parameters that have to be set. Both choices can dramatically affect the quality of the obtained solution, and the time spent obtaining it. Algorithm Selection, or Meta-Learning, techniques [1,2] typically address these questions by solving a large number of problems with each of the available algorithms, in order to learn a mapping from (problem,algorithm) pairs to expected performance. The obtained mapping is later used to select and run, for each new problem, only the algorithm that is expected to give the best results.

This approach, tough being preferable to the far more popular “trial and error”, poses a number of problems. It presumes that such a mapping can be learned at all, i.e. that the actual algorithm performance on a given problem will be predictable with enough precision before even starting the algorithm — often not the case with stochastic algorithms, whose performance can exhibit large fluctuations among different runs. It also assumes problem instances met during the training phase to be statistically representative of successive ones. For these reasons, there usually is no way to detect a relevant discrepancy between expected and actual performance of the chosen algorithm. It also neglects computational complexity issues: ranking between algorithms is often based solely on the expected quality of the performance, and the time spent during the training phase is not even considered, although it can be large enough to cancel any practical advantage of algorithm selection.¹

The Algorithm Portfolio paradigm [4,5] consists in selecting a subset of the available algorithms, to be run in parallel, with the same priority, until the fastest one solves

¹ E.g. 3 years of CPU time in [3].
the problem. This simple scheme is more robust, as it’s less likely that performance estimates will be wrong for all selected algorithms, but it also involves an additional overhead, due to the “brute force” parallel execution of all candidate solvers. In our view, a crucial weakness of these approaches is that they don’t exploit any feedback from the actual execution of the chosen algorithms. We try to move a step in this direction, introducing Dynamic Algorithm Portfolios. Instead of first choosing a portfolio then running it, we iteratively allocate a time slice, sharing it among all the available algorithms, and update the relative priorities of the algorithms, based on their current state, in order to favor the most promising ones. Instead of basing the priority attribution on performance quality, we fix a target performance, and try to minimize the time to reach it. To this aim, we search for a mapping from (problem,algorithm,current algorithm state) triplets to expected time to reach the desired performance quality. To further reduce computational complexity, we focus on lifelong-learning techniques that drop the artificial boundary between training and usage, exploiting the mapping during training, and including training time in performance evaluation.

In [6] we termed this approach Adaptive Online Time Allocation (AOTA), and introduced an example of a fixed heuristic mapping; in [7] we proposed a method to learn the mapping while solving a problem sequence. In the following we briefly present some related work (Sect. 2); describe the AOTA framework (Sect. 3) and its current instantiations (Sect. 4). Sect. 5 reports new experimental results. Sect. 6 closes the article, fixing some appointments with future work.

2 Previous work

A number of interesting “dynamic” exceptions to the static algorithm selection paradigm can be met in literature (see the techrep version of [6] for a more exhaustive bibliography). In [8], algorithm recommendation is based on the performance of the candidate algorithms during a predefined amount of time, called the observational horizon. In anytime algorithm monitoring [9], the dynamic performance profile of a planning technique is updated according to its performance, in order to stop the planning phase when further improvements in the actions planned are not worth the time spent in evaluating them. The “Parameterless GA” [10] is a fixed heuristic time allocation technique for Genetic Algorithms. In [12], a system solves function inversion and time-limited optimization problems by searching in a space of problem solving techniques, allocating time to them according to their probabilities, and updating the probabilities according to positive and negative results on a sequence of problems. In a Reinforcement Learning [13] setting, algorithm selection can be formulated as a Markov Decision Process: in [14], the algorithm set includes sequences of recursive algorithms, formed dynamically at run-time solving a sequential decision problem, and a variation of Q-learning is used to find an online algorithm selection policy; in [15], a set of deterministic algorithms is considered, and, under some limitations, static and dynamic algorithm selection techniques based on dynamic programming are presented. Success Story algorithms [16] can undo policy modifications that did not improve the reward rate.
Other research areas that can instead be related to “static” algorithm selection include landmarking [17], anytime algorithm scheduling [18], time limited planning [19], bandit problem solving [20], racing [21], search in program space [22].

3 AOTA framework

Consider a sequence $B$ of $m$ problem instances $b_1, b_2, \ldots, b_m$, roughly sorted in increasing order of difficulty, and featuring precise stopping criteria (e.g. search problems in which the solution is known to exist and can be recognized; optimization problems in which a reachable target value for performance is given); and a set $A$ of $n$ algorithms $a_1, a_2, \ldots, a_n$, that can be applied to the solution of the problems in $B$, paused and resumed at any time, and queried, at a negligible cost, for state information $d \in \mathbb{R}^d$ related to their progress in solving the current instance. We aim at minimizing the time to solve the whole problem sequence $B$. To describe the state of a Dynamic Algorithm Portfolio (DAP), let $t_i$ be the time already spent on $a_i$, $\tau_i$ the current estimate of the time still needed by $a_i$ to solve the current problem, $x_i$ a feature vector, possibly including information about the current problem instance, the algorithm $a_i$ itself (e.g. its kind, the values of its parameters), and its current state $d_i$; $H_i = \{(x_i^{(r)}, t_i^{(r)}), r = 0, \ldots, b_i\}$ a set of collected samples of these pairs, $f_\tau$ a model that maps histories $H_i$ to estimated $\tau_i$. If the model $f_\tau$ was precise enough, we would not need to run more than one algorithm, the $a_i$ that is mapped to a lower $\tau_i$ before its start ($t_i = 0$): it is instead more realistic to assume that the model’s estimates are rough, but can be improved by collecting more data in $H_i$, i.e. by getting more run-time feedback on the actual performance of $a_i$ on current problem instance. We then introduce a set of nonnegative scalars $P_A = \{p_1, \ldots, p_n\}, p_i \geq 0, \sum_{i=1}^n p_i = 1$, that represent the current bias of the portfolio, slice machine time with a small interval $\Delta t$, and iteratively share each time slice between the algorithms proportionally to the current bias; before each iteration, the bias is updated according to a function $f_P$ of $\{\tau_i\}$, that obviously gives more time to expected faster $a_i$ (i.e. the ones with a low $\tau_i$); after a share $p_i \Delta t$ has expired, $\tau_i$ is updated based on current $H_i$. In intra-problem AOTA, the predictive model $f_\tau$ is fixed; in inter-problem AOTA, $f_\tau$ itself is adaptive, and gets updated after each problem’s solution. A pseudocode for inter-problem AOTA would then be:

```
For each problem $b_k$
    initialize $\{\tau_i\}$
    While ($b_k$ not solved)
        update $P_A = f_P(\{\tau_i\})$
        For each algorithm $a_i$
            run $a_i$ for $p_i \Delta t$
            update $H_i = H_i \cup (x_i, t_i)$
            update $\tau_i = f_\tau(H_i)$
        End
    End
    update $f_\tau$ based on $H$
End
```
For $f_P$, one reasonable heuristic, that gave good results, consists in assigning $1/2$ of the current time slice to the expected fastest algorithm (i.e. the one with lowest $\tau$), $1/4$ to the second fastest, and so on. This heuristic cannot be directly applied to inter-problem AOTAs, though, as the model would obviously be unreliable during the first problems of the sequence. In this case it would be better to start the problem sequence with a “brute force” $f_P (p_i = 1/n)$, and vary it gradually towards the above described “ranking” $f_P (p_i = 2^{-r_i}, r_i$ being the current rank of $a_i$ based on $\{\tau_i\}$).

4 Example AOTAs

In [6] we presented a fixed heuristic $f_\tau$. We considered algorithms with a scalar state $x$, that had to reach a target value: $H_i$ in this case is a simple learning curve. Through a shifting window linear regression, we extrapolated for each $i$ the time $t_{i,sol}$ at which the current learning curve $H_i$ would reach the target value, in order to estimate the time to solution $\tau_i = t_{i,sol} - t_i$. Even though the estimates were obviously optimistic, they were updated so often that the overall performance of the intra-problem AOTA was remarkably good; its obvious limitations were that it required some prior knowledge about the algorithms, and a simple relationship between the learning curve and the time to solution.

What if we instead want to learn a potentially complex mapping $f_\tau$ from scratch? For a successful algorithm $a_i$ that solved the problem at time $t_{i}^{(h_i)}$, we can a posteriori evaluate the correct $\tau_i^{(r)} = t_{i}^{(h_i)} - t_{i}^{(r)}$ for each pair $(x_{i}^{(r)}, t_{i}^{(r)})$ in $H_i$. In a first tentative experiment, that led to poor results, these values were used as targets to learn a regression from pairs $(x, t)$ to residual time values $\tau$. The main problem with this approach is which $\tau$ values to choose as targets for the unsuccessful algorithms.

The alternative we presented in [7] is inspired by censored sampling for lifetime distribution estimation [23], and consists in learning a parametric model $g(\tau|x_i, t_i; w)$ of the conditional probability density function (pdf) of the residual time $\tau$. To see how the model can be trained, imagine we continue running the portfolio for a while after the first algorithm solves the current task, such that we end up having one or more successful algorithms $a_i$, for whose $H_i$ the correct targets $\tau_i^{(r)}$ can be evaluated as above. Assuming each $\tau_i^{(r)}$ to be the outcome of an independent experiment, including $t$ in $x$ to ease notation, if $p(x)$ is the (unknown) pdf of the $x_i^{(r)}$ we can write the likelihood of $H_i$ as

$$L_{\text{succ}}(H_i) = \prod_{r=0}^{h_i - 1} g(\tau_i^{(r)}|x_i^{(r)}; w)p(x_i^{(r)})$$ (1)

For the unsuccessful algorithms, the final time value $t_{i}^{(h_i)}$ recorded in $H_i$ is a lower bound on the unknown, and possibly infinite, time to solve the problem, and so are the $\tau_i^{(r)}$, so to obtain the likelihood we have to integrate (1)

$$L_{\text{fail}}(H_i) = \prod_{r=0}^{h_i - 1} (1 - G(\tau_i^{(r)}|x_i^{(r)}; w))p(x_i^{(r)})$$ (2)
where \( G(\tau|x; w) = \int_0^\tau g(\xi|x; w) d\xi \) is the conditional cumulative distribution function (cdf) corresponding to \( g \).

We can then search the value of \( w \) that maximizes \( \Sigma(H) = \prod_i \Sigma(H_i) \), or, in a Bayesian approach, maximize the posterior \( p(w|H) \propto \Sigma(H|w)p(w) \). Note that in both cases the logarithm of these quantities can be maximized, and terms not in \( w \) can be dropped.

To prevent overfitting, and force the model to have a realistic shape, we can use some known parametric lifetime model, such as a Weibull distribution [23], with pdf
\[
g(\tau|x, t; w) = \lambda^\beta \tau^{\beta-1} e^{-\lambda \tau^\beta},
\]
and express the dependency on \( x \) and \( w \) in its two parameters \( \lambda = \lambda(x; w) \), \( \beta = \beta(x; w) \): these can in turn be the two outputs of a more complex parametric model, with input \( x \), and parameter \( w \), whose value can be optimized by gradient descent, minimizing the negative logarithm of the resulting \( \Sigma(H) \), in a fashion that is common for modeling conditional distributions (see e.g. [24], par 6.4).

One advantage of this approach is that it fully exploits the state history information gathered, as it allows to learn from the unsuccessful algorithms as well.

5 Experiments

We present experimental results with the same \( A \) and \( B \) as in [7]: \( A \) a set of 76 simple generational Genetic Algorithms [11], differing in population size \((2^i, i = 1 \ldots 19)\), mutation rate \((0 \text{ or } 0.7/L, L \text{ being the genome length})\) and crossover operator (uniform or one-point, with rate 0.5 in both cases); \( B \) a sequence of artificial deceptive problems, such as the “trap” described in [10], consisting of \( n \) copies of an \( m \)-bit trap function: each \( m \)-bit block of a bitstring of length \( nm \) gives a fitness contribution of \( m \) if all its bits are \( 1 \), and of \( m - q \) if \( q < m \) bits are \( 1 \). This time we generated a sequence of 24 problems, varying the genome length from 20 to 96 and the size \( m \) of the deceptive block from 2 to 4. The problems were sorted based on a rough estimate of their difficulty, from 10 blocks of size \( 2 \) to 24 blocks of size \( 4 \). The feature vectors \( x \) included two problem features (genome length and block size), the algorithm parameters, the current best and average fitness values and the time spent, together with their last variations, for a total of 11 inputs.

To model \( g \) we used an Extreme Value distribution on the logarithms of time values,\(^2\) with parameters \( \eta(x; w) \) and \( \delta(x; w) \) being quadratic expansions of \( x \) of the form
\[
w(x) = w_0 + \sum_i w_i x_i + \sum_{i,j} w_{i,j} x_i x_j.
\]
The weights were obtained by maximizing the Bayesian posterior described in sect. 4, using a Gaussian prior with \( \sigma = 1 \).

As for \( f_P \), \( p_i \) was set proportionally to \((\frac{m+1}{m+1})^{-r_i} \), \( r_i \) being the current rank of \( a_i \) in order of increasing \( r_i \), \( j \) the index of current task, \( m \) the total number of tasks. In this way the distribution of time is uniform during the first task (when the model is still untrained), and tends through the task sequence to the “ranking” \( f_P \) described in Sect. 3 \( (p_i = 2^{-r_i}) \), in which the expected fastest solver gets half of the current time slice. After the solution of each task, a further \( 20\% \) of the time spent is allocated to the remaining algorithms, gathering more data in their histories in order to improve the model. Results were similar to the ones obtained in [7] with a slightly different AOTA, employing a Neural Network as parametric model.

\(^2\) If \( \tau \) is Weibull distributed, \( l = \log \tau \) has Extreme Value distribution.
Fig. 1. The cumulative time spent on the sequence of tasks by the adaptive $f_r$ method, labeled L2-AOTA, compared with the fixed $f_r$, AOTA$_{ra}$, the best algorithm in the set (BEST) and the brute-force approach (BRUTE). Time is measured in fitness function evaluations, values shown are upper 95% confidence limits calculated on 20 runs.

In Fig. 1 we compare the adaptive $f_r$ approach, labeled L2-AOTA, to the fixed $f_r$, AOTA$_{ra}$ from [6], this time with a ranking $f_P$, which gave slightly better results than the “greedy” $f_P$ of the original article. The adaptive model was able to learn the bits of a-priori knowledge that we had to pre-wire in the fixed heuristic AOTA (for example, the fact that increments in the average fitness are an indicator of potentially good performance) and it eventually outperformed it. Also shown are the performances of the (possibly different at each run) fastest solver of the set, labeled BEST, which would be the performance of an ideal algorithm selection with “foresight” of the correct $\tau_i$ values at $t_i = 0$; and the estimated performance of a brute force approach, i.e. running all the algorithms in parallel until one solves the problem, which leaves the figure and completes the task sequence at time $3.3 \times 10^7$. The average overhead with respect to BEST was about 11 for AOTA$_{ra}$ and 8 for L2-AOTA. This latter would be e.g. the performance of an already trained “static” Algorithm Portfolio that picked, for each problem, 8 of the 76 algorithms, always including the fastest; to fairly compare with such a technique, though, we should also consider its additional training time.
6 Conclusions and future work

The main purpose of this work was to present a new paradigm for algorithm selection and meta-learning, the Dynamic Algorithm Portfolio, and describe two examples for it, introduced in [6,7]. We also extended the experiments in [7] to a simpler parametric model $f_\tau$. The particular bias update strategy $f_P$ adopted allowed to use the model while training it, gradually increasing its impact on the time allocation as more problems were solved, and more historic data gathered, and practically saving computation time by limiting the learning process to the “interesting” parts of the training space.

We advocate the use of Dynamic Algorithm Portfolios with sets of computationally heavy algorithms. For faster algorithms, a more refined approach should also take into account the cost of updating the model. The models used here and in [7] were trained on all historic data gathered so far, with a “batch learning” approach: for longer problem sequences, an online method would obviously be preferable.

The adaptive $f_\tau$ method still requires some prior knowledge, as the choice of the algorithms in $A$ and of the features to include in $x$, as well as the sorting of the problems in $B$, are all left to the user. Another limitation is that the user has to set a target performance quality for each problem. While this might be natural for certain problems, in many cases it would be preferable to set instead a parameter quantifying how much extra time should be spent relative to the improvement in the solution. This would require learning a more complex model of performance profile, as in [9].

In future work we plan to address these limitations; ongoing experiments focus on alternative parametric models, and different algorithm set/problem sequence combinations.

Acknowledgements. This work was supported by SNF grant 16R1GSMLR1.

References