Адаптация размера популяции в $(1 + (\lambda, \lambda))$ ГА при помощи модифицированного правила одной пятой [1]

Басин А. О., аспирант кафедры компьютерных технологий Университета ИТМО, anton.bassin@gmail.com

Буздалов М. В., к. т. н., доцент кафедры компьютерных технологий Университета ИТМО, mbuzdalov@gmail.com

Аннотация

Автоадаптация параметров может значительно улучшить производительность эволюционных алгоритмов. Одним из ярких примеров является генетический алгоритм $(1 + (\lambda, \lambda))$, в котором адаптация размера популяции позволяет достичь линейного времени работы на задаче OneMax. Однако для других задач использование адаптации может привести к снижению производительности, по сравнению со статическим выбором параметров. В частности, правило одной пятой может слишком быстро увеличивать численность популяции в задачах с низкой корреляцией между значением приспособленности особи и расстоянием до глобального оптимума.

Мы предлагаем модификацию правила одной пятой, которая имеет менее негативное влияние в сценариях, когда исходное правило снижает производительность. Наша модификация имеет хорошую производительность на OneMax как теоретически, так и на практике. Новый подход также показывает лучшие результаты на линейных функциях со случайными весами и на задаче максимальной выполнимости булевых формул MAX-SAT.

Введение

Генетический алгоритм $(1 + (\lambda, \lambda))$ [1] является ярким примером успешного применения автонастройки параметров. Несмотря на первые успехи в применении данного алгоритма для решения отличных от OneMax задач,

[1] Данное исследование было поддержано Российским Научным Фондом, договор № 17-71-20178.

таких как линейные функции со случайными весами [1, 2], Royal Road функции [1] или задач MAX-SAT [4] (алгоритм показывает хорошую производительность на практике и в теории [2]), этот генетический алгоритм довольно медленно осваивает другие территории. Объяснение этому, возможно, состоит в том, что этот метод плохо справляется с задачами, в которых корреляция между значением приспособленности особи и расстоянием до глобального оптимума является низкой.

Исследование данной проблемы мы начинаем с анализа ландшафта для нескольких типов задач. Затем предлагаем модификацию правила одной пятой, которая замедляет (дез)адаптацию. Далее проводим эксперименты по решению тестовых задач и подтверждаем, что использование нашей модификации правила адаптации позволяет смягчить последствия неправильного поведения исходного алгоритма.

Расширенная версия данной статьи[2] содержит теоретическое доказательство линейного времени выполнения предлагаемого алгоритма на задаче OneMax. Данная работа также показывает, что если $(1 + (\lambda, \lambda))$ ГА выбирает оптимальные значения λ из следующего раздела, то он превосходит существующие варианты выбора параметров на *всех* тестовых задачах.

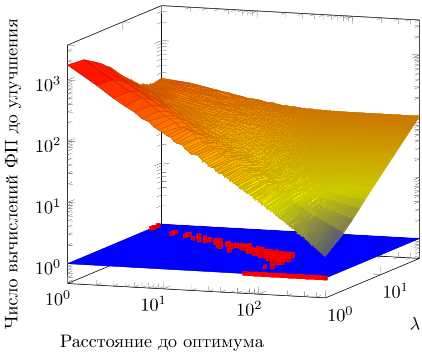
Анализ ландшафта оптимального размера популяции

На Рис. 1a–1e мы измерили влияние выбора значений λ, зависящих от расстояния Хэмминга до оптимума, для задач OneMax, $\textsc{LinInt}_2$, $\textsc{LinInt}_5$, $\textsc{LinInt}_n$ и MAX-SAT [3] с логарифмической плотностью. Размерность задач была взята равной $n = 10^3$. Для всех $\lambda_0 \in [1..50]$ мы 10^3 раз запускали $(1 + (\lambda, \lambda))$ ГА с фиксированным параметром $\lambda = \lambda_0$ для каждой отдельной задачи и различных значений λ_0. Для всех расстояний до оптимума $1 \le d \le 500$ мы рассматривали общее число вычислений функции приспособленности $E(d, \lambda)$, произведенных при данном значении d, и общее число возникновения событий нахождения более хорошего решения $I(d, \lambda)$.

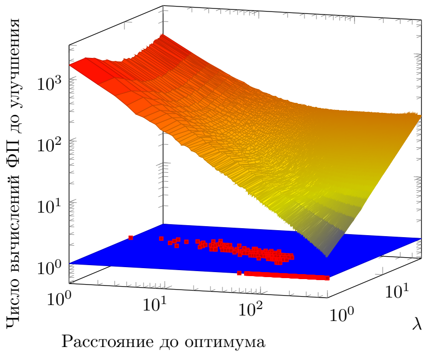
Верхние поверхности на рисунках отображают приближенное *ожидаемое число вычислений функции приспособленности до улучшения особи* $= E(d, \lambda)/I(d, \lambda)$. Красные квадраты на нижней поверхности показывают около-оптимальные значения λ, лежащие в пределах 2% разницы от экспериментально полученного оптимального значения.

Рис. 1a в целом подтверждает известный факт о том, что $\lambda \approx \sqrt{n/(n - f(x))}$ является около-оптимальным выбором размера популяции

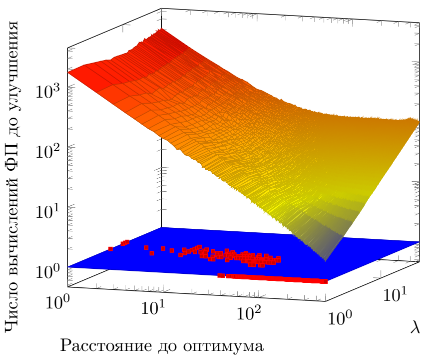
[2] http://arxiv.org/abs/1904.07284



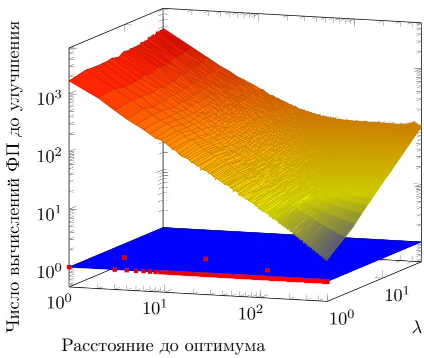
(a) OneMax, $n = 10^3$



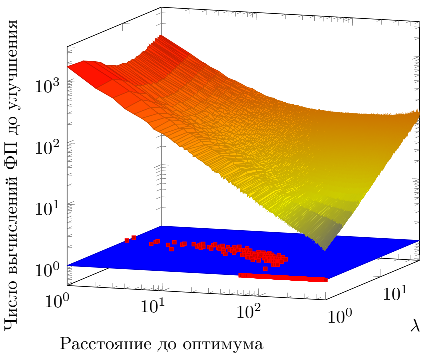
(b) LinInt$_2$, $n = 10^3$



(c) LinInt$_5$, $n = 10^3$

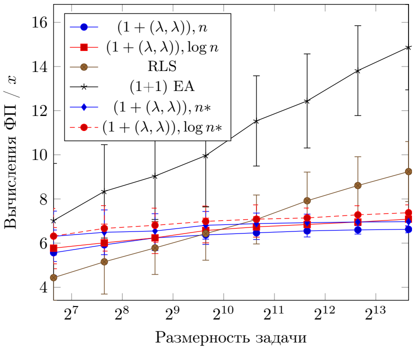


(d) LinInt$_n$, $n = 10^3$

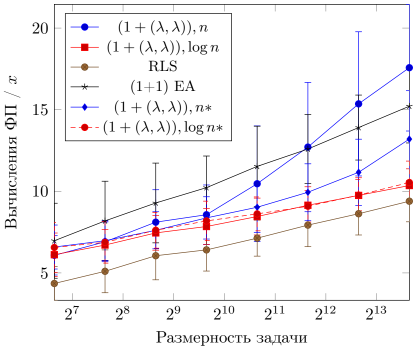


(e) MAX-SAT, $n = 10^3$, $m = 4n \log n$

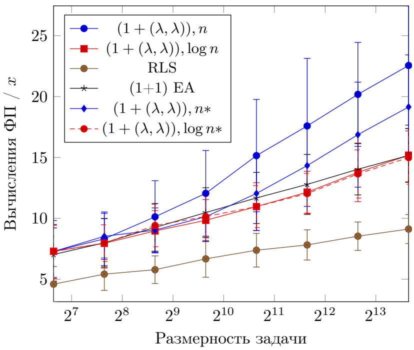
Рис. 1: Графики производительности $(1 + (\lambda, \lambda))$ ГА при различных $\lambda(d)$.



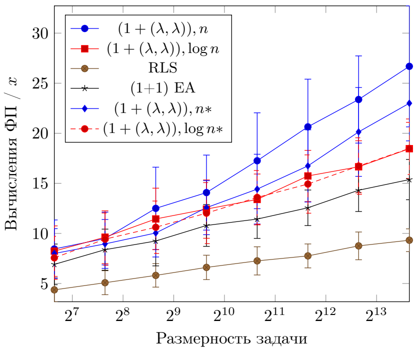
(a) OneMax



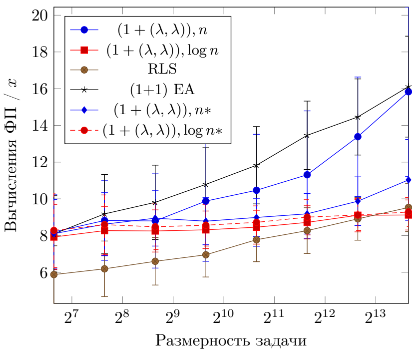
(b) $\textsc{LinInt}_2$



(c) $\textsc{LinInt}_5$



(d) $\textsc{LinInt}_n$



(e) MAX-SAT

Рис. 2: Время выполнения на различных функциях для $(1 + (\lambda, \lambda))$ ГА на задаче OneMax. Для задач LinInt$_2$ и LinInt$_5$ в логарифмических осях тенденция все еще сохраняет характер, приближенный к линейному. Верхняя поверхность может указывать на то, что правило одной пятой не успевает поддерживать λ в хорошей кондиции. На пограничном случае задачи LinInt$_n$ оптимальные λ сконцентрированы около $\lambda = 1$, и ландшафт выглядит слишком сложным для $(1 + (\lambda, \lambda))$ ГА. Кривая оптимальных λ кажется более изогнутой для MAX-SAT по сравнению с задачей OneMax. Для более больших расстояний значения размера популяции схожи с наблюдаемыми при OneMax, но они становятся непредсказуемыми при движении в сторону глобального оптимума.

Модификация правила одной пятой

Алгоритм 1 $(1 + (\lambda, \lambda))$ ГА, модифицированная автонастройка $\lambda \le \overline{\lambda}$

$F \leftarrow \text{const} \in (1; 2)$ ▷ Скорость адаптации
$U \leftarrow 5$ ▷ Правило “1/5-ой”
$B \leftarrow 0$ ▷ Число “плохих” итерации
$\Delta \leftarrow 10$, $\lambda_+ \leftarrow 1$ ▷ Диапазон роста λ, база λ
$x \leftarrow$ UniformRandom$(\{0, 1\}^n)$
for $t \leftarrow 1, 2, 3, \ldots$ **do**
  $p \leftarrow \lambda/n$, $c \leftarrow 1/\lambda$, $\lambda' \leftarrow [\lambda]$, $\ell \sim \mathcal{B}(n, p)$
  for $i \in [1..\lambda']$ **do** ▷ Фаза 1: Мутация
    $x^{(i)} \leftarrow$ Mutate(x, ℓ)
  $x' \leftarrow$ UniformRandom$(\arg\max_{x^{(i)}} f)$
  for $i \in [1..\lambda']$ **do** ▷ Фаза 2: Скрещивание
    $y^{(i)} \leftarrow$ Crossover(x, x', c)
  $y \leftarrow$ UniformRandom$(\arg\max_{y^{(i)}} f)$
  if $f(y) > f(x)$ **then** ▷ Селекция и адаптация
    $x \leftarrow y$, $\lambda \leftarrow \max\{\lambda/F, 1\}$
    $\lambda_0 \leftarrow \lambda$, $B \leftarrow 0$, $\Delta \leftarrow 10$
  else
    if $f(y) = f(x)$ **then** $x \leftarrow y$
    if $(B \leftarrow B + 1) = \Delta$ **then** $B \leftarrow 0$, $\Delta \leftarrow \Delta + 1$
    $\lambda \leftarrow \min\{\lambda_0 F^{B/(U-1)}, \overline{\lambda}\}$

Основная идея предлагаемой модификации (Алгоритм 1) правила автонастройки $(1 + (\lambda, \lambda))$ ГА заключается в том, чтобы запретить непрерывный рост λ при длительном периоде неудачных запусков, но сохранить возможность увеличения числа популяции до произвольно высоких значений в несколько шагов при необходимости. Также увеличивается шанс на выполнение итераций ГА с довольно низким значением λ, что может быть полезно для определенных видов задач. Данная модификация приближенно возводит в квадрат число итераций, необходимых для достижения некоторого отдаленного значения λ, по сравнению с классическим правилом одной пятой. В результате, даже если улучшения значения приспособленности нет, особенности нашей модификации правила адаптации λ позволяют произвести большее число итераций с около-оптимальными значениями λ в случаях, когда в исходном $(1 + (\lambda, \lambda))$ ГА число размера популяции превышает оптимальное значение для текущего расстояния до глобального оптимума.

Эксперименты

Мы оценивали исходный $(1 + (\lambda, \lambda))$ ГА с верхними ограничениями на размер популяции $\overline{\lambda} = n$ и $\overline{\lambda} = 2 \log n$, тот же алгоритм с модифицированной адаптацией, $(1 + 1)$ эволюционный алгоритм со стандартной побитовой мутацией и случайный локальный поиск (RLS). В экспериментах используются пять типов задач из раздела “Анализ ландшафта оптимального размера популяции”. Каждый алгоритм запускался 100 раз для каждой задачи размерности $n \in \{100, 200, 400, 800, 1600, 3200, 6400, 12800\}$.

Результаты представлены на Рис. 2a–2e. При решении задачи OneMax все методы показывают хороший результат. Предложенный алгоритм уступает исходному примерно 10%, однако динамика производительности остается линейной. На $\textsc{LinInt}_2$ версии алгоритмов с логарифмическим ограничением уже ведут себя не линейно, но лучше, чем неограниченные версии. Модифицированный ГА без ограничений находится ниже $(1 + 1)$ ЭА, поэтому можно утверждать, что, по крайней мере, константа во времени выполнения для него меньше по сравнению с исходным ГА. Для задач $\textsc{LinInt}_5$, $\textsc{LinInt}_n$ общая тенденция для неограниченных алгоритмов состоит в том, что график времени их выполнения растет выше соответствующих значения для $(1 + 1)$ ЭА. Но модифицированная стратегия всегда показывает лучший результат. Наклон этих графиков позволяет предположить, что время выполнения аппроксимируется значением $\Theta(n(\log n)^2)$. На графике для MAX-SAT число вычислений функций приспособленности для исходной неограниченной версии алгоритма также растет гораздо быстрее, чем для модифицированной.

Заключение

Мы предложили модификацию правила одной пятой для автонастройки параметра λ в $(1+(\lambda, \lambda))$ ГА. Она нацелена на снижение влияния нежелательных эффектов, которые приводят к упадку производительности на задачах с низкой корреляцией между значением приспособленности и расстоянием до глобального оптимума. На задаче OneMax предлагаемая стратегия сохраняет линейное время работы в теории и на практике. В трудных для исходной схемы автонастройки параметра λ случаях, мы смогли добиться стабильных улучшений, по сравнению с классическим $(1 + (\lambda, \lambda))$ ГА, для всех проблемных функций.

Литература

[1] Benjamin Doerr, Carola Doerr, and Franziska Ebel. From black-box complexity to designing new genetic algorithms. // Theoretical Computer Science, Vol. 567, 87–104, 2015.

[2] Maxim Buzdalov and Benjamin Doerr. Runtime Analysis of the $(1 + (\lambda, \lambda))$ Genetic Algorithm on Random Satisfiable 3-CNF Formulas. // Proceedings of Genetic and Evolutionary Computation Conference, 1343–1350, 2017.

[3] Michael R. Garey and David S. Johnson. Computers and Intractability: A Guide to the Theory of NP-Completeness. // W. H. Freeman & Co., New York, NY, USA, 1979.

[4] B. W. Goldman and W. F. Punch. Parameter-less Population Pyramid. // Proceedings of Genetic and Evolutionary Computation Conference, 785–792, 2014.