Лекция 3 :: Метод роя частиц

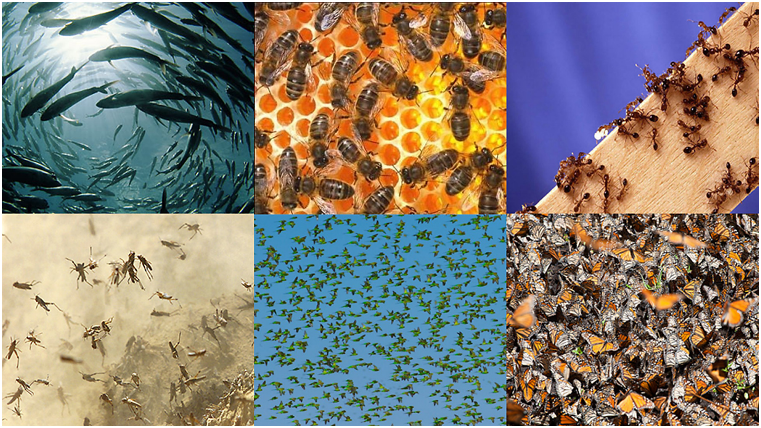
Ершов Н.М.

ershovnm@gmail.com

Роевая оптимизация и роевой интеллект

- Под *роевой оптимизацией* понимается класс алгоритмов, направленных на решение сложных оптимизационных задач (дискретная оптимизация, многомерная оптимизация и многокритериальная оптимизация), работа которых основана на моделировании коллективного поведения различных колоний живых организмов.
- Системы *роевого интеллекта* (swarm intelligence) представляют собой модели разнообразных типов колллективного поведения в колониях различных живых организмов — стаях птиц (bird flocks), косяках рыб (fish schools), колониях муравьев и т. п.

Роевой интеллект



Модель Рейнолдса

- Отличительной особенностью роевых моделей является то, что каждый отдельный организм в этой модели выполняет очень простые действия, подчиненные своей локальной цели, взаимодействуя с ограниченным числом других организмов в рассматриваемой колонии.
- Работы в области роевого интеллекта были инспирированы исследованиями разнообразных реальных колоний. Например, метод роя частиц основан на исследовании и моделировании коллективного поведения стай птиц, проведенного Рейнолдсом.
- Особенностью поведения птичьих стай является то, что в них нет никакого единого центра управления, тем не менее вся стая демонстрирует весьма цельное поведение.

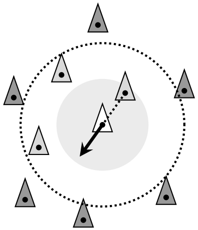
Птичьи стаи

Птичьи стаи

Модель Рейнолдса

- Рейнолдсом были сформулированы три простых принципа, которых должна придерживаться каждая отдельная птица:

  1) Каждая птица старается не приближаться к другим птицам на расстояние меньшее некоторой заданной величины.

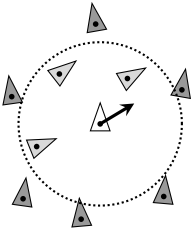


Этот принцип предназначен для того, чтобы избежать столкновений среди птиц.

Модель Рейнолдса

- Рейнолдсом были сформулированы три простых принципа, которых должна придерживаться каждая отдельная птица:

  2) Каждая птица старается выбрать свой вектор скорости наиболее близким к среднему вектору скорости среди всех птиц в своей локальной окрестности.

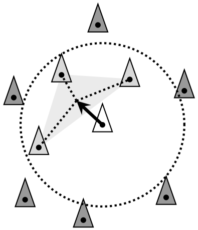


Этот принцип координирует скорость и направление полета.

Модель Рейнолдса

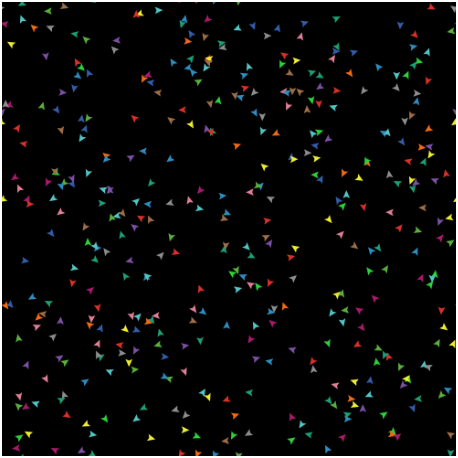
- Рейнолдсом были сформулированы три простых принципа, которых должна придерживаться каждая отдельная птица:

3) Каждая птица старается расположиться в геометрическом центре масс своей локальной окрестности.

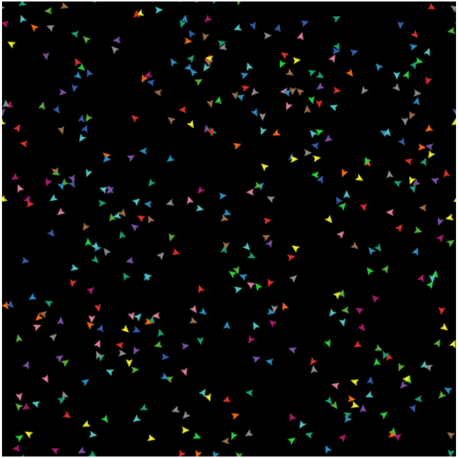


Этот принцип заставляет каждую птицу не отрываться от стаи (а в идеале, расположится внутри стаи).

Пример (NetLogo)



Пример (NetLogo)



Stanley and Stella in: Breaking the Ice

Коллективная оптимизация

- Переход от моделирования коллективного поведения к коллективной оптимизации основан на следующей биологической идее: организмы объединяются в колонии для улучшения своих условий существования — каждый организм в колонии в среднем имеет больше шансов на выживание в борьбе с хищниками, колония может более эффективно производить поиск, обработку и хранение пищи по сравнению с отдельными особями и т. д.
- Другими словами любая колония организмов в течение всего времени своего существования с той или иной степенью эффективности решает различные оптимизационные задачи, чаще всего — многокритериальные (например, максимизация количества пищи с одновременной минимизацией потерь от хищников).

Коллективная оптимизация

- Переход от моделирования коллективного поведения к коллективной оптимизации основан на следующей биологической идее: организмы объединяются в колонии для улучшения своих условий существования — каждый организм в колонии в среднем имеет больше шансов на выживание в борьбе с хищниками, колония может более эффективно производить поиск, обработку и хранение пищи по сравнению с отдельными особями и т. д.
- Другими словами любая колония организмов в течение всего времени своего существования с той или иной степенью эффективности решает различные оптимизационные задачи, чаще всего — многокритериальные (например, максимизация количества пищи с одновременной минимизацией потерь от хищников).

- Указанные соображения и легли в основу построения разнообразных математических методов оптимизации.
- В этой и следующих двух лекциях мы рассмотрим четыре классических алгоритма роевой оптимизации — метод роя частиц, муравьиные алгоритмы, алгоритмы бактериального поиска и пчелиные алгоритмы.
- В ряде работ по роевой оптимизации принято относить к системам роевого интеллекта и генетические алгоритмы, которые при некоторой интерпретации могут рассматриваться, как колонии неких простых организмов, взаимодействие между которыми осуществляется с помощью операторов скрещивания и отбора.

Метод роя частиц

- Метод роя частиц был разработан Джеймсом Кеннеди и Расселом Эберхартом в 1995 году.
- Модель, положенная в основу этого метода, была получена упрощением модели Рейнолдса, в результате чего отдельные особи популяции стали представляться объектами, не имеющими размера, но обладающими некоторой скоростью.
- В силу крайней схожести с материальными частицами получившиеся простые объекты стали называться *частицами*, а их популяция — *роем*.

Задача оптимизации

- Метод роя частиц предназначен для решения задач многомерной непрерывной оптимизации и основан на моделировании социального поведения колоний животных, выполняющих коллективный поиск мест с наилучшими условиями существования.
- Пусть задана функция $f : R^n \to R$, требуется найти глобальный максимум этой функции, т. е. точку x_0 такую, что $f(x_0) \geq f(x)$ для любого $x \in R^n$.
- Суть подхода, на котором основан метод роя частиц, заключается в том, что глобальный максимум функции f ищется с помощью системы (роя), состоящей из m частиц.
- Частицы выполняют поиск, перемещаясь по пространству решений R^n.

Траектории частиц

- Положение i-ой частицы задается вектором $x_i \in R^n$, значение $f(x_i)$ определяет функцию качества этой частицы в текущий момент времени.
- Каждая частица в рое обладает своей собственной скоростью $v_i \in R^n$, которая определяет, как изменяются координаты частицы со временем:

$$x_i \leftarrow x_i + \tau v_i,$$

  где τ — некоторая единица измерения скорости (продолжительность одного такта работы алгоритма, например, можно положить $\tau = 1$).
- Последовательность (дискретная) положений i-ой частицы будем называть ее *траекторией*.

Обновление скорости частицы

- Ключевая особенность метода роя частиц заключается в способе обновления скорости отдельных частиц, которое выполняется по формуле

$$v_i \leftarrow v_i + \alpha(p_i - x_i) + \beta(g - x_i).$$

- Первое слагаемое в этой формуле представляет собой *инерцию* частицы.

Обновление скорости частицы

- Ключевая особенность метода роя частиц заключается в способе обновления скорости отдельных частиц, которое выполняется по формуле

$$v_i \leftarrow v_i + \alpha(p_i - x_i) + \beta(g - x_i).$$

- Вектор p_i (второе слагаемое) служит простейшей моделью *индивидуальной памяти* — его координаты равны координатам лучшей точки траектории i-ой частицы за все время ее существования.
- Говорят, что второе слагаемое реализует *принцип простой ностальгии* — каждая частица «хочет» вернуться в ту точку, где ею было достигнуто лучшее значение функции f.

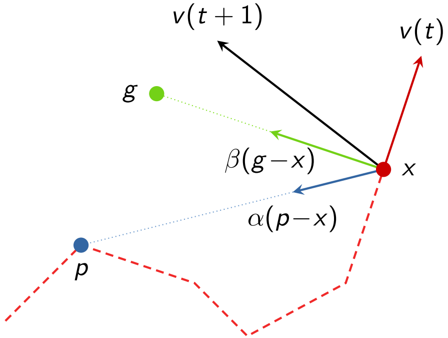
Обновление скорости частицы

- Ключевая особенность метода роя частиц заключается в способе обновления скорости отдельных частиц, которое выполняется по формуле

$$v_i \leftarrow v_i + \alpha(p_i - x_i) + \beta(g - x_i).$$

- Вектор g (третье слагаемое) соответствует лучшей точке, обнаруженной за время своего существования *всем роем*, т. е. представляет собою некую *коллективную память* роя.
- Это слагаемое определяет простую схему социального взаимодействия между отдельными частицами роя.

Обновление скорости



Обновление скорости частицы

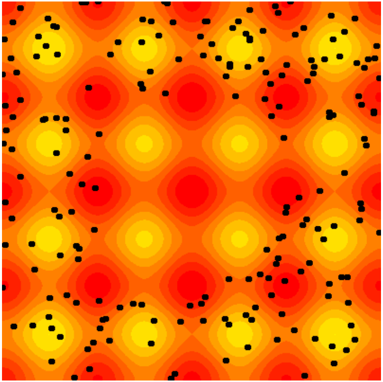
- Другими словами, изменение скорости каждой частицы (ускорение) определяется как взвешенная сумма двух векторов, первый из которых направлен на лучшую точку, обнаруженную данной частицей, а второй — на лучшую точку, обнаруженную всем роем.
- Коэффициенты α и β могут выбираться из разных соображений.
- Численные эксперименты показали, что лучшей является вероятностная схема — либо оба коэффициента выбираются случайным образом из диапазона $[0, 1]$, либо значение α выбирается случайным образом из этого диапазона, а значение β полагается равным $1 - \alpha$.

Формальная схема метода

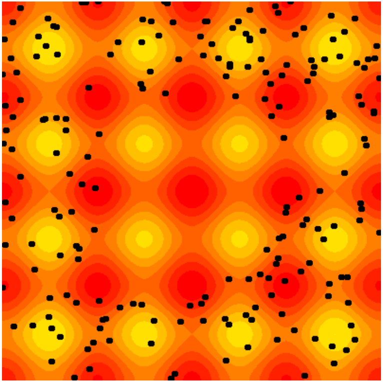
$\textsc{ParticleSwarmOptimization}(f, n, m)$

Создание случайного роя из m частиц
Вычисление векторов p_i и g
Случайная инициализация скоростей частиц
while Не выполнен критерий останова **do**
    for $i \in [1 \ldots m]$ **do**
        $\alpha, \beta \leftarrow \textsc{Random}(0, 1)$
        $v_i \leftarrow v_i + \alpha(p_i - x_i) + \beta(g - x_i)$
        $x_i \leftarrow x_i + \tau v_i$
        if $f(x_i) > f(p_i)$ **then**
            $p_i \leftarrow x_i$
            if $f(x_i) > f(g)$ **then**
                $g \leftarrow x_i$
return g

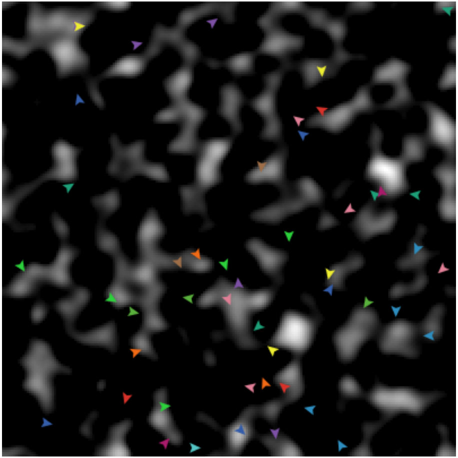
Пример 1



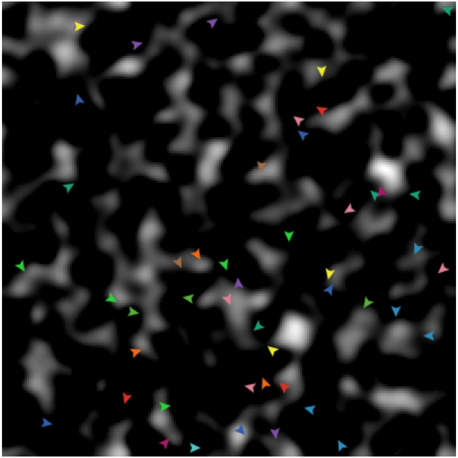
Пример 1



Пример 2 (NetLogo)



Пример 2 (NetLogo)



Особенности метода роя частиц

- Метод роя частиц является, по сути, метаэвристикой, хорошо зарекомендовавшей себя при решении различных оптимизационных задач.
- Его отличительной особенностью от многих других методов является то, что для метода роя частиц необходимо уметь вычислять только значение оптимизируемой функции, но не ее градиент.
- Т.е. функция, подлежащая оптимизации, не обязана быть дифференцируемой, более того, она может быть разрывной, зашумленной и т.п.

Особенности метода роя частиц

- С другой стороны, в силу того, что метод оперирует понятием скорости частиц, необходимым условием его применимости является непрерывность области определения функции.
- В частности, это означает, что данный метод неприменим напрямую к задачам дискретной оптимизации.

Вариации метода

- Большая часть предложенных вариаций базового алгоритма Кеннеди и Эберхарта касается модификации формулы скорости.
- Основная проблема, связанная с применением этой формулы заключается в том, что возникающие при этом скорости частиц мог быть сколь угодно большими, это привносит некоторую неустойчивость в работу метода.
- Простейший способ ограничить скорости частиц заключается в прямом запрете на превышение некоторой максимальной скорости $v_{\max}$.

Вариации метода

- Более естественный способ ограничения модуля скорости заключается в введении некоторого «сопротивления среды», задаваемого положительным коэффициентом $\gamma < 1$:

$$v_i \leftarrow \gamma v_i + \alpha(p_i - x_i) + \beta(g - x_i).$$

Вариации метода

- Другие, менее значительные изменения базового алгоритма касаются следующих моментов:
  - способ начального распределения частиц по пространству решений;
  - способ выбора начальной скорости частиц;
  - схема вычисления коэффициентов α и β;
  - обновление вектора g только после вычисления новых координат всеми частицами роя;
  - использование в формуле скорости вектора g, вычисленного не для всего роя (глобальная операция), а только для некоторого подмножества i-ой частицы.

Метод роя частиц и дискретная оптимизация

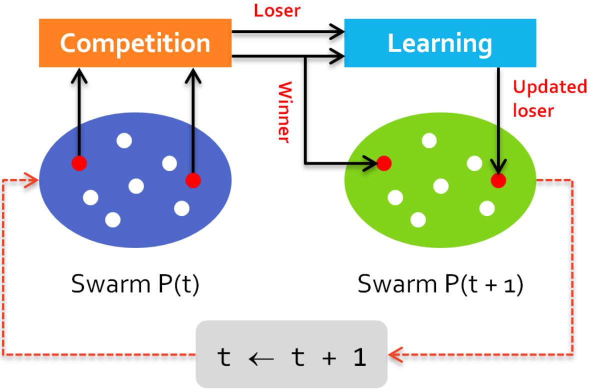
- Проблемой является то, что в дискретном пространстве решений (например, на множестве двоичных последовательностей) понятие скорости частицы полностью теряет свой смысл.
- Один из вариантов решения этой проблемы, предложенный самим Эберхартом, заключается в том, что k-ая компонента скорости частицы, ограниченная интервалом $[0, 1]$ трактуется, как вероятность изменения k-ой координаты этой частицы — чем больше скорость, тем чаще меняется координата (например, 0 на 1 и наоборот).
- Такой прием позволяет отделить пространство решений от пространства скоростей и, в частности, эффективно решать задачи оптимизации в дискретных пространствах решений.

Метод роя частиц и дискретная оптимизация

- Если пространство решений представляет собой множество перестановок из n элементов (например, в задаче коммивояжера), то для адаптации метода роя частиц можно использовать следующий прием.
- Рой частиц «живет» в обычном непререрывном пространстве $\mathbb{R}^n$. Координаты i-ой частицы $(x_1, x_2, \ldots, x_n)$ преобразуются в перестановку по следующему алгоритму: упорядочиваем координаты x_k по возрастанию, номера индексов в упорядоченной последовательности и образуют искомую перестановку.
- Например, вектор $(2, 7, 11, 3, 1)$ будет соответствовать перестановке $\langle 5, 1, 4, 2, 3 \rangle$.

Алгоритм CSO

- Относительно недавняя вариация алгоритма PSO — алгоритм CSO (Competetive Swarm Optimizer).
- Главное отличие — способ обновления скорости частицы.



Псевдокод CSO

$t = 0$;
randomly initialize $P(0)$;
while *terminal condition* is not satisfied **do**
  calculate the fitness of all particles in $P(t)$;
  $U = P(t)$, $P(t+1) = \emptyset$;
  while $U \neq \emptyset$ **do**
    randomly choose two particles $X_1(t)$, $X_2(t)$ from U;
    if $f(X_1(t)) \leq f(X_2(t))$ **then**
      $X_w(t) = X_1(t)$, $X_l(t) = X_2(t)$;
    else
      $X_w(t) = X_2(t)$, $X_l(t) = X_1(t)$;
    end if
    add $X_w(t)$ into $P(t+1)$;
    update $X_l(t)$ using (6) and (7);
    add the updated $X_l(t+1)$ to $P(t+1)$;
    remove $X_1(t)$, $X_2(t)$ from U;
  end while
  $t = t + 1$;
end while

$$\begin{aligned} V_{l,k}(t+1) = {} & R_1(k,t)V_{l,k}(t) \\ & + R_2(k,t)(X_{w,k}(t) - X_{l,k}(t)) \\ & + \varphi R_3(k,t)(\bar{X}_k(t) - X_{l,k}(t)). \end{aligned} \tag{6}$$

$$X_{l,k}(t+1) = X_{l,k}(t) + V_{l,k}(t+1), \tag{7}$$

Библиография

1. C. W. Reynolds, *Flocks, herds and schools: a distributed behavioral model*, Computer Graphics, 21, 4, pp. 25-34.
2. https://www.youtube.com/watch?v=3bTqWsVqyzE — Stanley and Stella in: Breaking the Ice.
3. J. Kennedy, R. C. Eberhart, *Particle swarm optimization*, Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ, pp. 1942–1948, 1995.
4. http://emlab.utep.edu/pdfs/Poster_PSO.pdf — EM Lab Poster on Particle Swarm Optimization.
5. R. Cheng and Y. Jin, *A Competitive Swarm Optimizer for Large Scale Optimization*, in IEEE Transactions on Cybernetics, vol. 45, no. 2, pp. 191-204, Feb. 2015.

Спасибо за внимание!