



The 6th International Conference
«Distributed Computing and Grid-technologies in Science and Education»

Естественные модели параллельных вычислений

Ершов Н.М., Попова Н.Н.

erhovnm@gmail.com popova@cs.msu.su

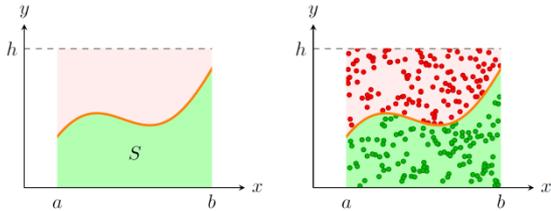
*Московский государственный университет им. М. В. Ломоносова,
факультет вычислительной математики и кибернетики*

Естественные вычисления = {
информатика
математика
естественные науки

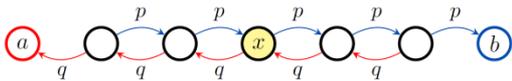
- Содержание курса: классические разделы (клеточные автоматы, нейронные сети) и современные подходы (ДНК-вычисления, методы роевого интеллекта).
- Каждая из рассмотренных в курсе моделей обладает значительным внутренним параллелизмом.
- Рассматриваются вопросы массивно-параллельной реализации моделей.
 - **Практический** аспект: большое число разнообразных приложений – моделирование, интеллектуальный анализ данных, распознавание образов, управление.
 - **Педагогический** аспект: большое число паттернов параллельного программирования.
 - **Теоретический** аспект: параллельные алгоритмические модели.

1. Методы Монте-Карло

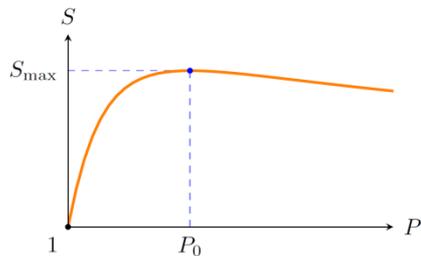
Методы Монте-Карло, при всей их универсальности, обладают существенным недостатком – скорость их сходимости к точному решению является очень низкой. Это означает, что для получения решений с высокой точностью требуется проведение очень **большого** числа экспериментов. Положительным фактором в этой ситуации является то, что методы Монте-Карло легко и эффективно **распараллеливаются**.



Подробно рассматриваются задача вычисления определенных интегралов и одномерная модель случайного блуждания (задача о разорении игрока).



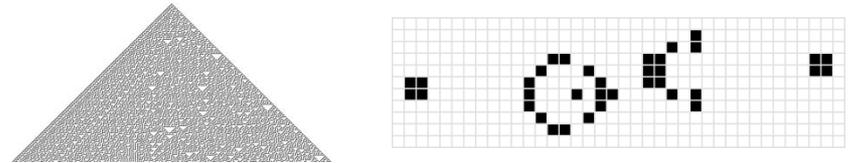
Высокая эффективность распараллеливания алгоритмов Монте-Карло основана на таких их свойствах, как независимость отдельных испытаний и большое число этих испытаний. Первое свойство позволяет легко распараллелить основную вычислительную часть алгоритма, а второе гарантирует, что коммуникация на этапе сбора статистики будет требовать существенно меньшего времени по сравнению с вычислениями.



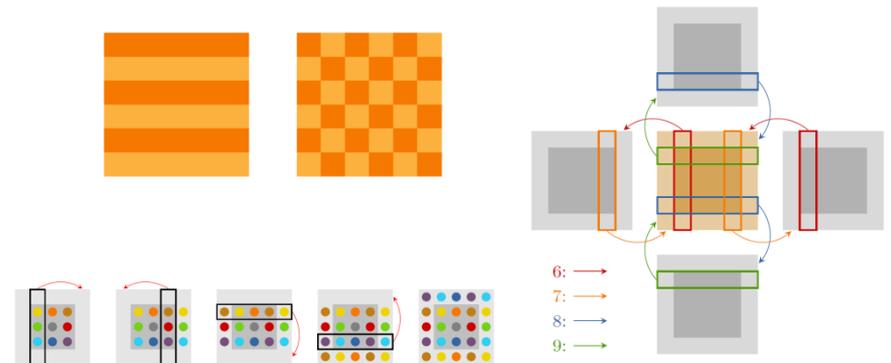
2. Клеточные автоматы

Клеточные автоматы были впервые введены в рассмотрение Джоном фон Нейманом в 1940-х годах с целью моделирования процессов самовоспроизведения в биологических системах. Несмотря на свою столь продолжительную историю, клеточные автоматы остаются одним из самых распространенных средств дискретного моделирования.

Клеточные автоматы обладают высокой степенью внутреннего параллелизма, что делает их весьма привлекательным объектом приложения для современных суперкомпьютерных систем. Еще одним интересным свойством клеточных автоматов с точки зрения параллельных вычислений является то, что многие их конфигурации являются алгоритмически универсальными, а это значит, что клеточные автоматы могут использоваться и в качестве теоретической модели распределенных вычислений.

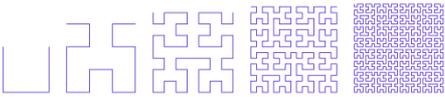


Подробно рассматриваются вопросы распараллеливания одномерных элементарных клеточных автоматов и двумерных автоматов (на примере игры «Жизнь»).

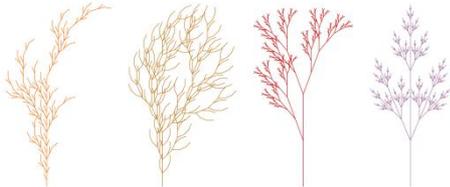


3. Системы Линденмайера

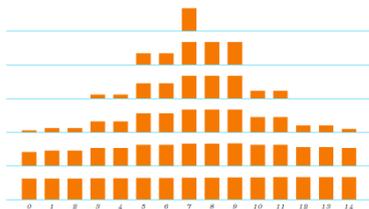
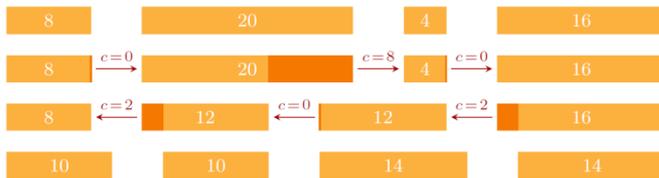
Системы Линденмайера (L-системы) являются частным случаем строковых перезаписывающих систем, и некоторым специальным вариантом формальных грамматик. От последних L-системы отличаются **параллельным** характером применения правил.

$$\begin{cases} A \rightarrow +BX - AXA - XB+ \\ B \rightarrow -AX + BXB + XA- \end{cases}$$


L-системы были разработаны венгерским биологом Линденмайером в 1968 году. Областью применения L-систем является моделирование процессов роста и формообразования в различных биологических системах.

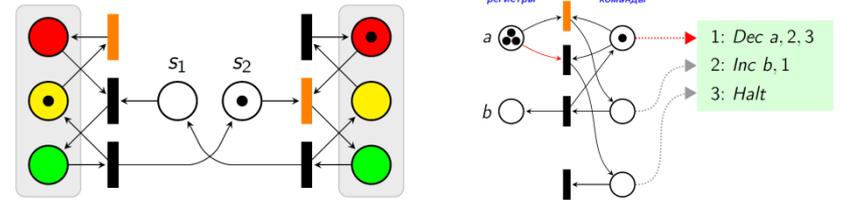


L-системы, несмотря на значительную степень внутреннего параллелизма, обладают рядом свойств, затрудняющих их эффективное распараллеливание. В частности, при параллельной реализации L-систем возникает необходимость в разработке механизма **динамического** выравнивания нагрузки.

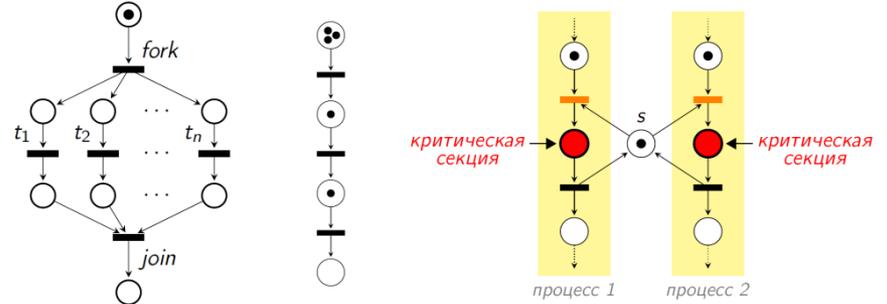


4. Сети Петри

Сети Петри представляют собой простое и удобное средство для моделирования разнообразных распределенных систем и процессов. Эта модель была придумана немецким ученым Карлом Петри в 1939 году для описания химических процессов.



Сети Петри по своему определению обладают встроенным параллелизмом, поэтому они традиционно используются для моделирования разнообразных параллельных процессов, как искусственных (например, вычислительные сети и системы), так и естественных (например, сети химических реакций).

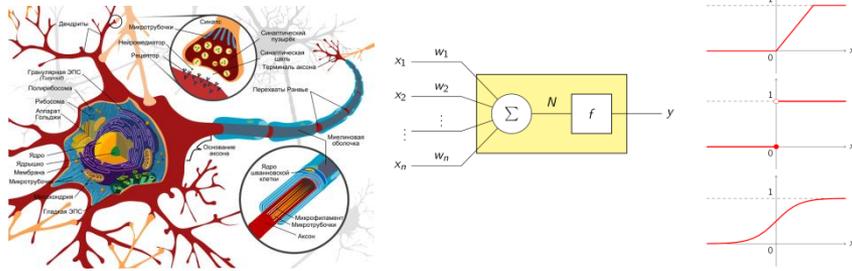


Это делает сети Петри с одной стороны адекватным инструментом анализа параллельных процессов, с другой - удобным объектом для реализации на параллельных вычислительных системах.

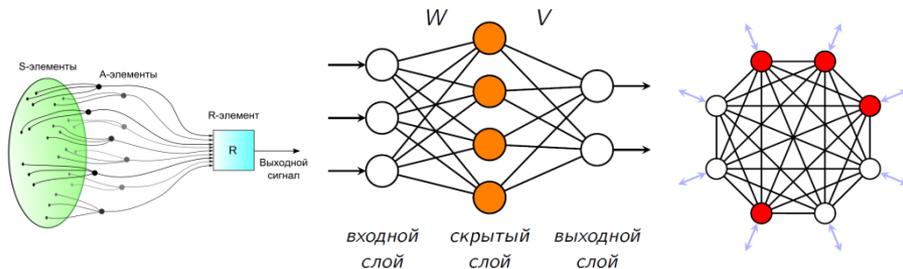


5. Искусственные нейронные сети

Под искусственной нейронной сетью понимается математическая модель (а также ее программная или аппаратная реализация), построенная по принципу организации и функционирования биологических нейронных сетей – нервных систем живых организмов. Понятие нейронной сети возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. Впоследствии модели нейронных сетей стали использовать и в практических целях: в задачах прогнозирования, распознавания образов, управления и т. д.

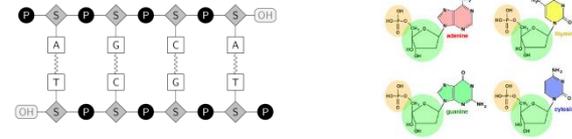


Искусственная нейронная сеть представляет собой ориентированный граф, вершинами которого являются отдельные нейроны. Каждая дуга такого графа соответствует отдельному синапсу и помечена силой этого синапса. Процесс настройки синаптических весов принято называть **обучением** нейронной сети. В зависимости от цели обучения различают несколько его типов: обучение с учителем, без учителя; с подкреплением.



6. ДНК-вычисления

ДНК-вычисления являются частным случаем молекулярных вычислений, основанных на массированном использовании молекул ДНК. Ключевой особенностью такой модели вычислений является то, что все действия в ней выполняются одновременно над очень большим числом молекул, что позволяет достичь **полиномиального** времени при решении **переборных** задач.



История ДНК-вычислений началась в 1994 году с работы Леонарда Адлемана, который в течение семи дней решил NP-полную задачу поиска гамильтонова пути в графе с семью вершинами, выполняя различные манипуляции над молекулами ДНК.

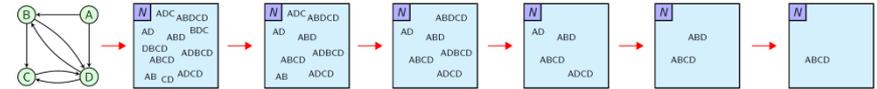
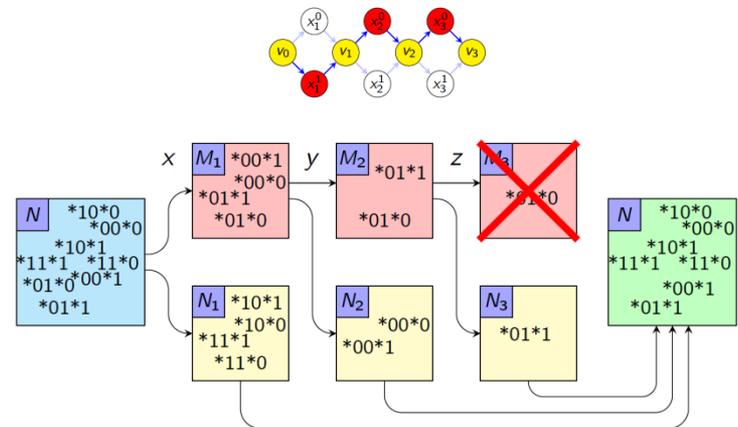
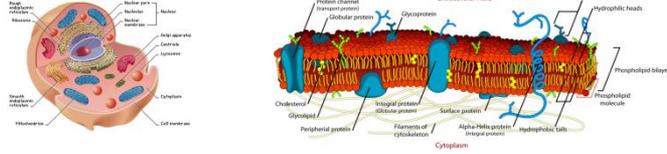


Схема Липтона применяется для решения задач, в которых требуется выполнять перебор всех двоичных последовательностей заданной длины (например, задача о сумме элементов подмножества или задача о выполнимости булевой формулы).

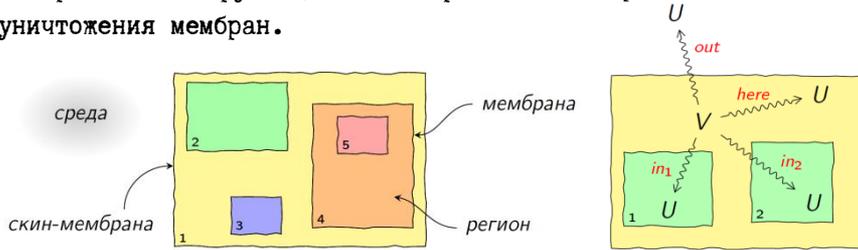


7. Мембранные системы

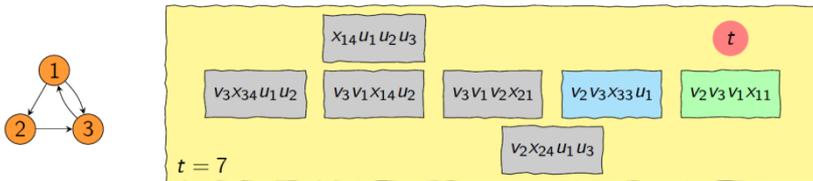
Мембранные системы (или Р-системы) являются вычислительной моделью, основанной на простейшей аналогии с биологическими клетками.



В центре данной теории лежит понятие мембраны – некоторого контейнера, содержащего в себе неупорядоченный набор объектов (символов). Мембраны могут вкладываться друг в друга, образуя мембранную структуру. Для содержимого мембран определяются правила эволюции, правила обмена информацией мембраны с окружающей ее средой, и правила деления и уничтожения мембран.

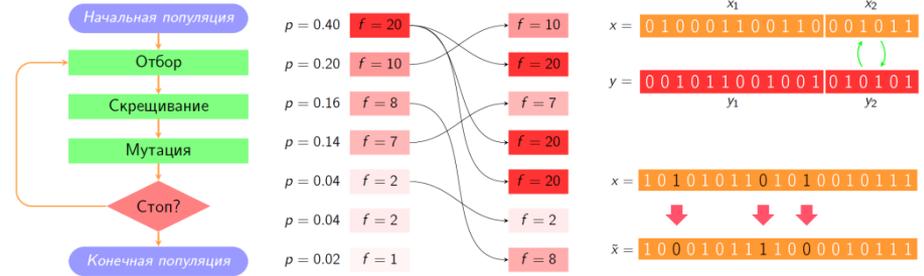


Мембранные системы могут рассматриваться как параллельная алгоритмическая модель. Особенностью таких систем является то, что степень их параллелизма оказывается динамической величиной. Это свойство можно использовать для построения Р-систем, решающих различные NP-полные задачи за полиномиальное время.

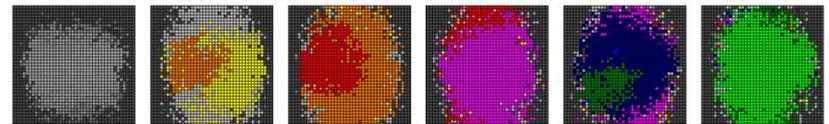
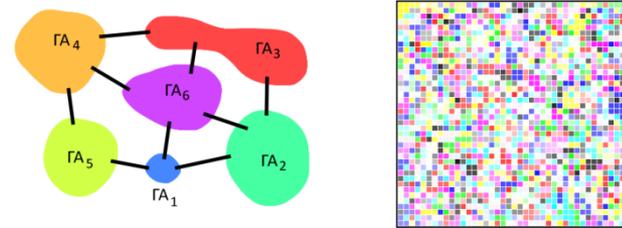


8. Генетические алгоритмы

Работа генетических алгоритмов основана на моделировании эволюционных процессов с использованием генетических механизмов. Рассматривается набор (популяция) сразу нескольких возможных решений задачи. Решения «борются» за существование (выживают в этом процессе отбора наиболее приспособленные), обмениваются информацией (через механизм скрещивания). Мутации применяются для поддержания генетического разнообразия популяции.



Генетические алгоритмы в принципе обладают хорошим потенциалом для параллельной реализации, т. к. представляют собой совокупность отдельных объектов (решений), которые могут обрабатываться более менее независимо друг от друга. Две распространенные параллельные схемы: островная модель и клеточные генетические алгоритмы.

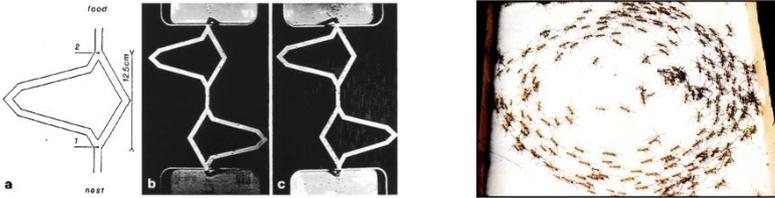


9. Муравьиные алгоритмы

Муравьиные алгоритмы – это семейство приближенных алгоритмов для решения различных сложных оптимизационных задач. Идея этих алгоритмов основана на моделировании поведения муравьиной колонии, выполняющей поиск пути от муравейника к источнику пищи.

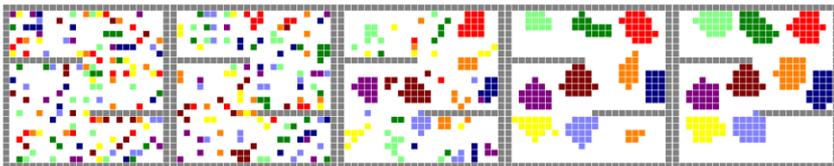


При поиске путей к источникам пищи муравьи помечают пройденный путь специальным веществом – феромоном. Остальные муравьи, попадая на такой путь, с большой вероятностью начинают двигаться по этому пути.



Первой задачей, для которой был построен муравьиный алгоритм, была задача коммивояжера. На основе муравьиного алгоритма для задачи коммивояжера была разработана более общая стратегия решения задач комбинаторной оптимизации, получившая название метаэвристики муравьиной колонии.

Роевая робототехника представляет собой новый подход к построению распределенных робототехнических систем, которые состоят из большого числа просто устроенных физических роботов. Желаемое коллективное поведение в таких системах возникает при взаимодействии роботов между собой и со средой.



10. Роевая оптимизация

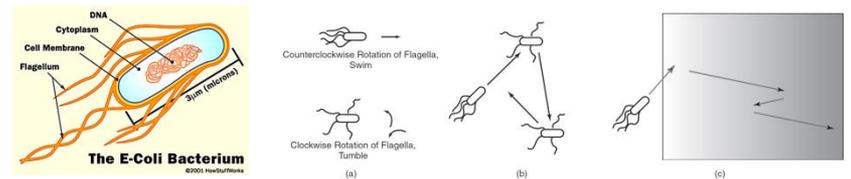
Под роевой оптимизацией понимается класс алгоритмов, направленных на решение сложных оптимизационных задач, работа которых основана на моделировании **коллективного** поведения различных колоний живых организмов.



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.

Модель, положенная в основу метода роя частиц, была получена упрощением модели Рейнолдса – отдельные особи популяции стали представляться объектами, не имеющими размера, но обладающими некоторой скоростью. Эти объекты стали называться частицами, а их популяция – роем.

Алгоритм бактериального поиска, разработанный Кевином Пассино, основан на моделировании процесса поиска питательных веществ бактериями наиболее изученного вида *E. coli*.



Алгоритм оптимизации, моделирующий процесс поиска пчелами мест с высоким содержанием нектара, разработан в 2005 году Фамом. Базовый вариант этого алгоритма является комбинацией алгоритмов локального и случайного поиска.

- Курс «Естественные модели параллельных вычислений» читается с 2010 года студентам факультета ВМК МГУ, а также магистрантам кафедры прикладной математики и информатики университета «Дубна».
- Изложение теоретического материала курса сопровождается рассмотрением возможных схем распараллеливания вычислений, а в практической части курса предполагается выполнение студентами программной реализации рассматриваемых моделей с использованием технологии MPI и проведение численного исследования эффективности выбранных схем распараллеливания вычислений.
- Материалы курса, в том числе презентации лекций, теоретические задания (тесты) и практические задания, доступны в сети Интернет. Теоретические задания принимаются с использованием онлайн сервисов Google Drive.



<http://naturalmodels.blogspot.ru>



<http://mmroi.blogspot.ru/>