

12. Лекция: Эволюционное моделирование и генетические алгоритмы

Рассматриваются основные понятия и принципы эволюционного моделирования систем, а также генетических алгоритмов - адекватного аппарата его проведения.

Цель лекции: ввести в суть проблемы, сформулировать основные положения и принципы, цели эволюционного моделирования и дать общее понятие о генетических алгоритмах и их возможностях в эволюционном моделировании.

Потребность в прогнозе и адекватной оценке последствий осуществляемых человеком мероприятий (особенно негативных) приводит к необходимости моделирования динамики изменения основных параметров системы, динамики взаимодействия открытой системы с его окружением (ресурсы, потенциал, условия, технологии и т.д.), с которым осуществляется обмен ресурсами в условиях враждебных, конкурентных, кооперативных или же безразличных взаимоотношений. Здесь необходимы системный подход, эффективные методы и критерии оценки адекватности моделей, которые направлены не только (не столько) на максимизацию критериев типа "прибыль", "рентабельность", но и на оптимизацию отношений с окружающей средой. Если критерии первого типа важны, например, для кратко- и среднесрочного прогнозирования и тактического администрирования, то второго типа - для средне- и долгосрочного прогноза, для стратегического администрирования. При этом необходимо выделить и изучить достаточно полную и информативную систему параметров исследуемой системы и его окружения, разработать методику введения мер информативности и близости состояний системы. Важно отметить, что при этом некоторые критерии и меры могут часто конфликтовать друг с другом.

Многие такие социально-экономические системы можно описывать с единых позиций, средствами и методами единой теории - эволюционной.

При *эволюционном моделировании* процесс моделирования сложной социально-экономической системы сводится к созданию модели его эволюции или к поиску допустимых состояний системы, к процедуре (алгоритму) отслеживания множества допустимых состояний (траекторий). При этом актуализируются такие атрибуты биологической эволюционной динамики (в скобках даны возможные социально-экономические интерпретации этих атрибутов для *эволюционного моделирования*) как, например:

1. сообщество (корпорация, корпоративные объекты, субъекты, окружение);
2. видовое разнообразие и распределение в экологической нише (типы распределения ресурсов, структура связей в данной корпорации);

3. экологическая ниша (сфера влияния и функционирования, эволюции на рынке, в бизнесе);
4. рождаемость и смертность (производство и разрушение);
5. изменчивость (экономической обстановки, ресурсов);
6. конкурентные взаимоотношения (рыночные отношения);
7. память (способность к циклам воспроизводства);
8. естественный отбор (штрафные и поощрительные меры);
9. наследственность (производственные циклы и их предыстория);
10. регуляция (инвестиции);
11. самоорганизация и стремление системы в процессе эволюции максимизировать контакт с окружением в целях самоорганизации, возврата на траекторию устойчивого развития и другие.

При исследовании эволюции системы необходима ее декомпозиция на подсистемы с целью обеспечения:

1. эффективного взаимодействия с окружением;
2. оптимального обмена определяющими материальными, энергетическими, информационными, организационными ресурсами с подсистемами;
3. эволюционируемости системы в условиях динамической смены и переупорядочивания целей, структурной *активности* и сложности системы;
4. управляемости системы, идентификации управляющей подсистемы и эффективных связей с подсистемами системы, обратной связи.

Пусть имеется некоторая система S с N подсистемами. Для каждой i -й подсистемы определим вектор $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, \dots, x_{n_i}^{(i)})$ основных параметров (т.е. параметров, без которых нельзя описать и изучить функционирование подсистемы в соответствии с целями и доступными ресурсами системы) и функцию $s^{(i)} = s(x^{(i)})$, которую назовем функцией *активности* или просто *активностью* этой подсистемы.

Пример. В бизнес-процессах это понятие близко к понятию деловой *активности*.

Для всей системы определены вектор состояния системы x и *активность* системы $s(x)$, а также понятие общего потенциала системы.

Пример. Потенциал *активности* может быть определен аналогично биологическому потенциалу популяции, например, с помощью интеграла от *активности* на задаваемом временном промежутке моделирования.

Эти функции отражают интенсивность процессов как в подсистемах, так и в системе в целом.

Важными для задач моделирования являются три значения $s^{(i)}_{\max}$, $s^{(i)}_{\min}$, $s^{(i)}_{\text{opt}}$ - максимальные, минимальные и оптимальные значения *активности* i -й подсистемы, а также аналогичные значения для всей системы (s_{\max} , s_{\min} , s_{opt}). В качестве показателя экономического состояния можно брать также отношение значения этого показателя к его

нормированному значению, а для комплексного учета влияния параметров на состояние системы можно использовать аналоги меры информационной близости, например, по К. Шеннону.

Если дана открытая экономическая система (процесс), а H_0, H_1 - энтропия системы в начальном и конечном состояниях процесса, то мера информации определяется как разность вида:

$$\Delta H = H_0 - H_1.$$

Уменьшение ΔH свидетельствует о приближении системы к состоянию статического равновесия (при доступных ресурсах), а увеличение - об удалении. Величина ΔH - количество информации, необходимой для перехода от одного уровня организации системы к другой (при $\Delta H > 0$ - более высокой, при $\Delta H < 0$ - более низкой организации).

Возможен подход и с использованием меры по Н. Моисееву. Пусть дана некоторая управляемая система, о состояниях которой известны лишь некоторые оценки - нижняя s_{\min} и верхняя s_{\max} . Известна целевая функция управления $F(s(t), u(t))$, где $s(t)$ - состояние системы в момент времени t , а $u(t)$ - управление из некоторого множества допустимых управлений, причем считаем, что достижимо u_{opt} - некоторое оптимальное управление из пространства U , $t_0 < t < T$, $s_{\min} \leq s \leq s_{\max}$. Мера успешности принятия решения:

$$H = | (F_{\max} - F_{\min}) / (F_{\max} + F_{\min}) |,$$

$$F_{\max} = \max F(u_{\text{opt}}, s_{\max}), \quad F_{\min} = \min F(u_{\text{opt}}, s_{\min}),$$

$$t \in [t_0; T], \quad s \in [s_{\min}; s_{\max}].$$

Увеличение H свидетельствует об успешности управления системой (успешности принятого управляющего решения).

Активности подсистем прямо или опосредованно взаимодействуют с помощью системной *активности* $s(x)$, например, по простой схеме вида

$$\begin{cases} \frac{ds(t)}{dt} = \sum_{i=1}^n \varphi^{(i)}(s; s^{(i)}, t), \\ \frac{ds^{(i)}(t)}{dt} = \psi^{(i)}(s^{(i)}; s, t) \end{cases}$$

Функции $\varphi^{(i)}$, $\psi^{(i)}$ должны отражать эволюционируемость системы, в частности, удовлетворять условиям:

периодичности, цикличности, например:

1. $(\exists 0 < T < \infty, \forall t: \varphi^{(i)}(s; s^{(i)}, t) = \varphi^{(i)}(s; s^{(i)}, t+T),$
2. $\psi^{(i)}(s; s^{(i)}, t) = \psi^{(i)}(s; s^{(i)}, t+T);$
3. затухания при снижении *активности*, например:
4. $(s(x) \rightarrow 0 \quad \forall i=1, 2, \dots, n) \Rightarrow (\varphi^{(i)} \rightarrow 0, \psi^{(i)} \rightarrow 0);$

5. равновесности и стационарности: выбор (определение) функции $\varphi^{(i)}$, $\Psi^{(i)}$ осуществляется таким образом, чтобы система имела точки равновесного состояния, а $s_{opt}^{(i)}$, S_{opt} достигались в стационарных точках $x_{opt}^{(i)}$, X_{opt} для малых промежутков времени; в больших промежутках времени система может (в соответствии с теорией катастроф) вести себя хаотично, самопроизвольно порождая регулярные, упорядоченные, циклические взаимодействия (детерминированный хаос).

Взаимные активности $\Psi^{(ij)}(s; s^{(i)}, s^{(j)}, t)$ подсистем i и j мы не учитываем. В качестве функции $\varphi^{(i)}$, $\Psi^{(i)}$ могут быть эффективно использованы производственные функции типа Кобба-Дугласа:

$$\varphi(s) = \prod_{i=1}^n \left(\frac{s(t) - s_{\max}}{s_{opt} - s_{\min}} \right)^{\alpha_i(t)} \left(\frac{s_{\max} - s(t)}{s_{\max} - s_{opt}} \right)^{-\alpha_i(t) \frac{s_{\max} - s_{opt}}{s_{opt} - s_{\min}}}$$

В таких функциях важен параметр α_i , отражающий степень саморегуляции, адаптации системы. Как правило, его нужно идентифицировать.

Функционирование системы удовлетворяет на каждом временном интервале $(t; t+\tau)$ ограничениям вида

$$\int_t^{t+\tau} s(z) dz \leq K_\tau,$$

$$\int_t^{t+\tau} s^{(i)}(z) dz \leq K_\tau^{(i)},$$

$$\sum_{i=0}^n \int_t^{t+\tau} s^{(i)}(z) dz \leq M_\tau$$

При этом отметим, что выполнение для $\tau > 0$ одного из двух условий

$$\int_t^{t+\tau} s^{(i)}(z) dz > K_\tau^{(i)}$$

$$\sum_{i=0}^n \int_t^{t+\tau} s^{(i)}(z) dz > M_\tau$$

приводит к разрушению (катастрофе) системы.

Пример. Пусть имеется некоторая социально-экономическая среда, которая возобновляет с коэффициентом возобновления $\alpha(\tau, t, x)$ ($0 < t < T$, $0 < x < 1$, $0 < \tau < T$) свои ресурсы. Этот коэффициент зависит, в общем случае, от мощности среды (ее ресурсоемкости, ресурсообеспеченности). Рассмотрим простую гипотезу: $\alpha(\tau, t, x) = \alpha_0 + \alpha_1 x$, и чем больше ресурсов - тем больше темп их возобновления. Можно записать непрерывную эволюционную модель (a - коэффициент естественного прироста ресурсов, b - их убыли):

$$x'(\tau) = -\left[a(\tau) + \int_0^{\tau} b(\tau, s)x(s)ds\right]x(\tau),$$

$$x(0) = \int_0^1 [\alpha_0(\tau) + \alpha_1(\tau)x(\tau)]x(\tau)d\tau.$$

Обозначим $\alpha(\tau) = \alpha_0(\tau) + \alpha_1(\tau)x(\tau) > 0$. Тогда

$$x(0) = \int_0^T \alpha(\tau, x)x(\tau)d\tau.$$

Задача всегда имеет решение $x \equiv 0$. Тогда эволюционный потенциал системы можно определить как величину:

$$\lambda = \int_0^T \alpha(\tau) e^{-\int_0^{\tau} a(s)ds} d\tau$$

Чем выше темп α - тем выше λ , чем меньше α - тем ниже λ . Каким бы хорошим ни было состояние ресурсов в начальный момент, они неизменно будут истощаться, если потенциал системы меньше 1.

Пример. Пусть u_{\max} - максимальный уровень синтаксических ошибок в программе P , $u(t)$ - их оставшееся количество к моменту времени t . Исходя из простейшей эволюционной модели $du/dt + \lambda u_{\max} = 0$, $u(t_0) = u_0$, можно заключить, что уровень ошибок убывает при $\lambda(c - t_0) \neq -1$ ($t_0 < c < T$) по закону: $u(t) = u_0(1 + \lambda(c - t))/(1 + \lambda(c - t_0))$. Если задать дополнительно $u(t^*) = u_*$, (u_{\max} - неизвестная величина, $t^* \neq t_0$), то закон изменения уровня ошибок находится однозначно, так как: $c = (u_* t_0 - u_0 t^*) / (\lambda u_* - \lambda u_0) - 1/\lambda$.

Отметим, что если ds/dt - общее изменение энтропии системы при воздействии на систему, ds_1/dt - изменение энтропии за счет необратимых изменений структуры, потоков внутри системы (рассматриваемой как открытая система), ds_2/dt - изменение энтропии за счет усилий по улучшению обстановки (например, экономической, экологической, социальной), то справедливо уравнение И. Пригожина:

$$ds/dt = ds_1/dt + ds_2/dt.$$

При *эволюционном моделировании* социально-экономических систем полезно использовать и классические математические модели, и неклассические, в частности, учитывающие пространственную структуру системы (например, клеточные автоматы и фракталы), структуру и иерархию подсистем (например, графы и структуры данных), опыт и интуицию (например, эвристические, экспертные процедуры).

Пример. Пусть дана некоторая экологическая система Ω , в которой имеются точки загрязнения (выбросов загрязнителей) x_i , $i=1, 2, \dots, n$. Каждый загрязнитель x_i загрязняет последовательно экосистему в промежутке времени $(t_{i-1}; t_i]$, $t_i=t_i-t_{i-1}$. Каждый загрязнитель может оказать воздействие на *активность* другого загрязнителя (например, уменьшить, нейтрализовать или усилить по известному эффекту суммирования воздействия загрязнителей). Силу (меру) такого влияния можно определить через r_{ij} , $R=\{r_{ij}: i=1,2,\dots, n-1; j=2,3,\dots, n\}$.

Структура задаётся графом: вершины - загрязнители, ребра - меры.

Найдём подстановку минимизирующую функционал вида:

$$F = \sum_{i=1}^{n-1} \sum_{j=2}^n r_{ij} = \min$$

где F - суммарное загрязнение системы с данной структурой S .

Чем быстрее (медленнее) будет произведен учёт загрязнения в точке x_i , тем быстрее (медленнее) осуществимы социо-экономические мероприятия по его нейтрализации (усилению воздействия). Чем меньше будет загрязнителей до загрязнителя x_i , тем меньше будет загрязнение среды.

В качестве меры r_{ij} может быть взята мера, учитывающая как время начала воздействия загрязнителей (предшествующих данной x_j), так и число, а также интенсивность этих загрязнителей:

$$s_{ij} = v_{ij} \left[h_j \sum_{k=i+1}^{j-1} \tau_k + (1 + h_j)(j - i) \right]$$

где v_{ij} - весовой коэффициент, определяющий степень влияния загрязнителя x_i на загрязнитель x_j (эффект суммирования), h_j - весовой коэффициент, учитывающий удельную интенсивность действия загрязнителя x_j или интервал τ_i , в течение которого уменьшается интенсивность (концентрация) загрязнителя. Весовые коэффициенты устанавливаются экспертно или экспериментально.

Принцип *эволюционного моделирования* предполагает необходимость и эффективность использования методов и технологии искусственного интеллекта, в частности, экспертных систем.

Основная трудность при построении и использовании эволюционных моделей: в Природе и Познании, в которых эти модели и цели явно или неявно существуют, результаты функционирования системы и достижения цели прослеживаемы часто лишь по прошествии длительного периода времени, хотя в Обществе и Экономике Человек стремится получить результаты в соответствии с целью явно и быстро, с минимальными затратами Ресурсов.

Адекватным средством реализации процедур *эволюционного моделирования* являются *генетические алгоритмы*.

Идея *генетических алгоритмов* "подсмотрена" у систем живой природы, у систем, эволюция которых разворачивается в сложных системах достаточно быстро.

Генетический алгоритм - это алгоритм, основанный на имитации генетических процедур развития популяции в соответствии с принципами эволюционной динамики, приведенными выше. Часто используется для решения задач оптимизации (многокритериальной), поиска, управления.

Данные алгоритмы адаптивны, развивают решения, развиваются сами. Особенность этих алгоритмов - их успешное использование при решении NP-сложных проблем (проблем, для которых невозможно построить алгоритм с полиномиально возрастающей алгоритмической сложностью).

Пример. Рассмотрим задачу безусловной целочисленной оптимизации (размещения): найти максимум $f(i)$, i - набор из n нулей и единиц, например, при $n=5$, $i=(1,0,0,1,0)$. Это очень сложная комбинаторная задача для обычных, "негенетических" алгоритмов. *Генетический алгоритм* может быть построен следующей укрупненной процедурой:

1 - генерируем начальную популяцию (набор допустимых решений задачи) - $I_0=(i_1, i_2, \dots, i_n)$, $i_j \in \{0,1\}$ и определяем некоторый критерий достижения "хорошего" решения, критерий останова α , процедуру СЕЛЕКЦИЯ, процедуру СКРЕЩИВАНИЕ, процедуру МУТАЦИЯ и процедуру обновления популяции ОБНОВИТЬ;

2 - $k:=0$, $f_0:=\max\{f(i), i \in I_0\}$;

3 - нц пока не(α)

- a. с помощью вероятностного оператора (селекции) выбираем два допустимых решения (родителей) i_1, i_2 из выбранной популяции (вызов процедуры СЕЛЕКЦИЯ);
- b. по этим родителям строим новое решение (вызов процедуры СКРЕЩИВАНИЕ) и получаем новое решение i ;
- c. модифицируем это решение (вызов процедуры МУТАЦИЯ);
- d. если $f_0 < f(i)$ то $f_0 := f(i)$;
- e. обновляем популяцию (вызов процедуры ОБНОВИТЬ);

$$f . k := k + 1$$

4 - кц

Указанные процедуры определяются с использованием аналогичных процедур живой природы (на том уровне знаний о них, что мы имеем). Например, процедура СЕЛЕКЦИЯ может из случайных элементов популяции выбирать элемент с наибольшим значением $f(i)$. Процедура СКРЕЩИВАНИЕ (кроссовер) может по векторам i_1, i_2 строить вектор i , присваивая с вероятностью 0,5 соответствующую координату каждого из этих векторов-родителей. Это самая простая процедура. Используют и более сложные процедуры, реализующие более полные аналоги генетических механизмов. Процедура МУТАЦИЯ также может быть простой или сложной. Например, простая процедура с задаваемой вероятностью для каждого вектора меняет его координаты на противоположные (0 на 1, и наоборот). Процедура ОБНОВИТЬ заключается в обновлении всех элементов популяции в соответствии с указанными процедурами.

Пример. Работу банка можно моделировать на основе *генетических алгоритмов*. С их помощью можно выбирать оптимальные банковские проценты (вкладов, кредитов) некоторого банка в условиях конкуренции с тем, чтобы привлечь больше клиентов (средств). Тот банк, который сможет привлечь больше вкладов, клиентов и средств, и выработает более привлекательную стратегию поведения (эволюции) - тот и выживет в условиях естественного отбора. Филиалы такого банка (гены) будут лучше приспособляться и укрепляться в экономической нише, а, возможно, и увеличиваться с каждым новым поколением. Каждый филиал банка (индивид популяции) может быть оценен мерой его приспособленности. В основе таких мер могут лежать различные критерии, например, аналог экономического потенциала - рейтинг надежности банка или соотношение привлеченных и собственных средств банка. Такая оценка эквивалентна оценке того, насколько эффективен организм при конкуренции за ресурсы, т.е. его выживаемости, биологическому потенциалу. При этом особи (филиалы) могут приводить к появлению потомства (новых банков, получаемых в результате слияния или распада), сочетающего те или иные (экономические) характеристики родителей. Например, если один банк имел качественную политику кредитования, а другой - эффективную инвестиционную политику, то новый банк может приобрести и то, и другое. Наименее приспособленные особи (филиалы) совсем могут исчезнуть в результате эволюции. Таким образом, отрабатывается генетическая процедура воспроизводства новых банков (нового поколения), более приспособленных и способных к выживанию в процессе эволюции банковской системы. Эта политика со временем пронизывает всю банковскую "популяцию", обеспечивая достижение цели - появления эффективно работающей, надежной и устойчивой банковской системы. Приведем соответствующий *генетический алгоритм* (укрупненный и упрощенный):

```
алг ГЕНЕТИЧЕСКИЙ_АЛГОРИТМ_БАНКОВСКОЙ_СИСТЕМЫ
  ввод Начальная структура банка (начальная популяция);
  СТРУКТУРА | процедура оценки структуры по приспособлению
```

```

Стоп:=0      | флаг для завершения эволюционного процесса
нц пока (Стоп=0)
    СЕЛЕКЦИЯ | процедура генетического отбора нового поколения
нц пока (МЕРА) | цикл воспроизводства с критерием МЕРА
    | мерой эффективности банковской системы
    РОДИТЕЛИ | процедура выбора двух структур (филиалов)
    | объединяемых (скрещиваемых) на новом шаге
    ОБЪЕДИНЕНИЕ | процедура образования (объединения)
    | нового банка (филиала)
    ОЦЕНКА | процедура оценки устойчивости нового банка,
    | образования (рейтинга, устойчивости)
    ВКЛЮЧЕНИЕ | процедура включения (не включения) в новое
    | поколение (в банковскую систему)
кц
МУТАЦИЯ | процедура эволюции (мутации) нового поколения
если (ПРОЦЕСС) | проверка функционала завершаемости эволюции
то Стоп:=1
кц
кон.

```

Мы не конкретизируем структуру процедур СЕЛЕКЦИЯ, МЕРА, РОДИТЕЛИ, ОБЪЕДИНЕНИЕ, ОЦЕНКА, ВКЛЮЧЕНИЕ, МУТАЦИЯ, ПРОЦЕСС, хотя даже на интуитивном уровне ясно, что в этом алгоритме они играют решающую роль для эволюционного процесса. Не менее важен и правильный (эффективный) выбор структуры, а также представления (описания) этой структуры. Часто ее выбирают по аналогии со структурой хромосом, например, в виде битовых строк. Каждая строка (хромосома) представляет собой конкатенацию ряда подстрок (генная комбинация). Гены располагаются в различных позициях строки (локусах хромосомы). Они могут принимать некоторые значения (аллели), например, для битового представления - 0 и 1. Структура данных в *генетическом алгоритме* (генотип) отражает генетическую модель особи. Окружающая среда, окружение определяется вектором в пространстве параметров и соответствует термину "фенотип". Мера качества (процедура МЕРА) структуры часто определяется целевой функцией (приспособленности). Для каждого нового поколения *генетический алгоритм* осуществляет отбор пропорционально приспособленности (процедура ОТБОР), модификацию (процедуры РОДИТЕЛИ, ОБЪЕДИНЕНИЕ, ВКЛЮЧЕНИЕ) и мутацию (процедура МУТАЦИЯ). Например, в процедуре ОТБОР каждой структуре ставится в соответствие отношение ее приспособленности к суммарной приспособленности популяции и затем происходит отбор (с замещением) всех особей для дальнейшей генетической обработки в соответствии с этой величиной. Размер отбираемой комбинации можно брать пропорциональным приспособляемости, и поэтому особи (кластеры) с более высокой приспособленностью с большей вероятностью будут чаще выбираться, чем особи с низкой приспособленностью. После отбора выбранные особи подвергаются кроссоверу (рекомбинации), т.е. разбиваются на пары. Для каждой пары может применяться кроссовер. Неизменные особи переходят к стадии мутации. Если кроссовер происходит, полученные потомки заменяют собой родителей и переходят к мутации.

Хотя *генетические алгоритмы* и могут быть использованы для решения задач, которые, видимо, нельзя решать другими методами, они не гарантируют нахождение

оптимального решения (по крайней мере, - за приемлемое время; полиномиальные оценки здесь часто неприменимы). Здесь более уместны критерии типа "достаточно хорошо и достаточно быстро". Главное же преимущество в другом: они позволяют решать сложные задачи, для которых не разработаны пока устойчивые и приемлемые методы, особенно на этапе формализации и структурирования системы, в когнитивных системах. *Генетические алгоритмы* эффективны в комбинации с другими классическими алгоритмами, эвристическими процедурами, а также в тех случаях, когда о множестве решений есть некоторая дополнительная информация, позволяющая настраивать параметры модели, корректировать критерии отбора, эволюции.

Вопросы для самоконтроля

1. Что такое *эволюционное моделирование*? Каковы критерии эффективности при *эволюционном моделировании*? Для какого типа прогнозирования (по длительности) используется и является эффективным *эволюционное моделирование*?
2. Что такое *генетический алгоритм*?
3. Каковы основные общие и различные свойства генетических и "не генетических" алгоритмов?

Задачи и упражнения

1. Привести одну экологическую или экономическую эволюционирующую систему и сформулировать основные принципы и понятия для постановки задачи *эволюционного моделирования* этой системы.
2. На примере некоторой системы показать, как можно осуществить её декомпозицию с целью её *эволюционного моделирования*. Указать приоритеты декомпозиции. Привести для задачи некоторый способ (описание) *активности* системы, а также функции, по которым можно определять эволюционируемость системы.
3. Описать укрупненный *генетический алгоритм* эволюции некоторого предприятия (некоторых предприятий).

Темы научных исследований и рефератов

1. *Эволюционное моделирование* - особенности, значение, приложения.
2. *Генетические алгоритмы* - особенности, значение, применение.
3. Имитационное *эволюционное моделирование* плохо структурируемых, плохо формализуемых систем с помощью *генетических алгоритмов*.