

Лекция 1

Слайд 1

Название лекции

Математическое моделирование динамики молекулярно-генетических систем

Слайд 2

Цели лекции

- Рассмотреть общие понятия:

Что такое модель и для чего она нужна? Что такое моделирование?

Приемы и способы моделирования. Процесс построения модели.

Генные сети как объект моделирования, их специфика и особенности

- Методы моделирования динамики генных сетей и их классификация
- Обобщенный химико-кинетический метод моделирования

- Формальное описание математических и компьютерных моделей генных сетей

Автоматическое конструирование математических моделей на основе информации из баз данных

- Проблемы математического моделирования генных сетей
- Верификация параметров математических моделей генных сетей (решение обратной задачи)

Слайд 3

Что такое модель?

Модель, в широком смысле любой образ, аналог (мысленный или условный: изображение, описание, схема, чертеж, график, план, карта и т. п.) какого-либо объекта, процесса или явления («оригинала» данной модели), используемый в качестве его «заместителя», «представителя» [Энциклопедия Кирилла и Мефодия].

Слайд 4

Что такое моделирование?

Моделирование - исследование каких-либо явлений, процессов или систем объектов путем построения и изучения их моделей; использование моделей для определения или уточнения характеристик и рационализации способов построения вновь конструируемых объектов.

Моделирование одна из основных категорий теории познания: на идее моделирования по существу базируется любой метод научного исследования как теоретический (при котором используются различного рода знаковые, абстрактные модели), так и экспериментальный (использующий предметные модели) [Энциклопедия Кирилла и Мефодия].

Слайд 5

Для понимания, как устроен изучаемый объект: какова его структура, основные свойства, законы развития и взаимодействия с окружающим миром;

Для того чтобы научиться управлять объектом (или процессом) и определить наилучшие способы управления при заданных целях и критериях;

Для того чтобы прогнозировать прямые и косвенные последствия реализации заданных способов и форм воздействия на объект.

[Горстко А.Б. Познакомьтесь с математическим моделированием, 1991 г.]

Слайд 6

Приемы и способы моделирования

Моделирование можно разделить на материальное (предметное) и идеальное. Под материальным моделированием понимаются способы моделирования, при которых

исследование ведется на основе модели, воспроизводящей основные геометрические, физические, динамические и функциональные характеристики изучаемого объекта. Идеальное моделирование носит теоретический характер и основано не на материальной аналогии объекта и модели, а на аналогии идеальной, мыслимой.

Материальное моделирование, в свою очередь, можно подразделить на физическое и аналоговое. При физическом моделировании реальному объекту противопоставляется его увеличенная или уменьшенная копия (например, макет здания; модель самолета; планетарий и т.д.). Аналоговое моделирование основано на аналогии процессов и явлений, имеющих различную физическую природу, но одинаково описываемых формально (например, изучение механических колебаний с помощью электрической схемы и т.д.).

Идеальное моделирование, в свою очередь, можно подразделить на интуитивное и знаковое. Интуитивное моделирование основано на интуитивном представлении об объекте исследования (например, жизненный опыт - интуитивная модель окружающего мира). Знаковое моделирование использует в качестве моделей знаковые преобразования какого-либо вида: схемы, графики, чертежи, формулы, наборы символов и т.д., а также включает совокупность законов, по которым можно оперировать с выбранными знаковыми образованиями и их элементами.

[Горстко А.Б. Познакомьтесь с математическим моделированием, 1991 г.]

Слайд 7

В процессе построения моделей:

1. Формируются основные вопросы о поведении системы, ответы на которые мы хотим получить с помощью модели.
2. Из множества законов, управляющих поведением системы, учитываются те, влияние которых существенно при поиске ответов на поставленные вопросы.
3. В дополнение к этим законам, если необходимо, для системы в целом или отдельных ее частей формулируются определенные гипотезы о функционировании. Как правило, эти гипотезы правдоподобны в том смысле, что могут быть приведены некоторые теоретические доводы в пользу их принятия.
4. Гипотезы так же, как и законы, выражаются в форме определенных математических соотношений, которые объединяются в некоторое формальное описание модели.

[Горстко А.Б. Познакомьтесь с математическим моделированием, 1991 г.]

Слайд 8

Функционирование живых организмов контролируется информацией, закодированной в их геномах. Генетическая информация считывается и преобразуется с помощью фундаментальных генетических процессов: репликации, транскрипции, трансляции и др. Синтезируемые на их основе генетические макромолекулы (ДНК, РНК и белки) взаимодействуют друг с другом, обеспечивая функционирование генных сетей.

Слайд 9

Актуальность проблемы

Одной из центральных проблем постгеномной молекулярной биологии и генетики является исследование принципов организации, функционирования и эволюции генных сетей - сложных молекулярно-генетических систем, обеспечивающих формирование фенотипических (молекулярно-генетических, биохимических, физиологических, морфологических и т.д.) характеристик живых организмов на основе информации, закодированной в их геномах.

Слайд 10

В настоящее время структурно-функциональная организация и динамика генных сетей всесторонне исследуется. При этом, можно выделить следующие основные направления исследования в данной области: геномика, транскриптомика, протеомика, функциональная

геномика/протеомика, метаболомика и системная биология, которая интегрирует все вышеперечисленные направления.

Слайд 11

Произшедшая в последние десятилетия революция в методах секвенирования геномов и разработка новейших технологий (генные чипы, протеомные чипы и т.д.), позволяющих в одном эксперименте изучать динамику функционирования многих десятков тысяч генов, продуктов их экспрессии, низкомолекулярных соединений привели к возникновению в науке о жизни качественно новой ситуации, которая выдвинула в число наиболее актуальных фундаментальную проблему осмысления беспрецедентно огромных объемов экспериментальных данных о функционировании молекулярно-генетических систем, накопление которых идет с все возрастающей скоростью. Решение данной проблемы принципиально невозможно без использования современных информационных технологий и эффективных математических и компьютерных методов анализа данных и моделирования биологических систем и процессов. Именно поэтому теоретическое и компьютерное исследование молекулярно-генетических систем, приобретает в настоящее время фундаментальное и первоочередное значение. В связи с этим в последнее время в мире наблюдается резкое повышение интереса к исследованию экспрессии генов, реконструкции генных сетей и развитию методов моделирования их динамики. Создано огромное число специализированных баз данных, в которых накапливается информация по структурно-функциональной организации молекулярно-генетических систем.

Слайд 12 и 13

Среди них можно выделить базы данных по регуляторным и метаболическим путям:

KEGG: Regulatory Pathways (<http://www.genome.ad.jp/kegg/regulation.html>)

SPAD: Signal Transduction (<http://www.grt.kyushu-u.ac.jp/spad/>)

CSNDB: Cell Signaling Networks (<http://geo.nihs.go.jp/csndb/>)

MIPS: Yeast Pathways (<http://www.mips.biochem.mpg.de/proj/yeast/pathways/index.html>)

Interactive Fly: Drosophila Genes (<http://sdb.bio.purdue.edu/fly/aimain/1aahome.htm>)

Wnt Signaling Pathway (<http://www.stanford.edu/~rnusse/wntwindow.html>)

TRANSPATH: Gene Regulatory Pathways (<http://transpath.gbf.de/>)

KEGG: Metabolic Pathways (<http://www.genome.ad.jp/kegg/metabolism.html>)

EMP: Enzymes and Metabolic Pathways (<http://emp.mcs.anl.gov/>)

WIT: Metabolic Reconstruction (<http://wit.mcs.anl.gov/WIT2/>)

UM-BBD: Microbial Biocatalysis/Biodegradation (<http://umbbd.ahc.umn.edu/>)

BioCyc: Procarotic Genes and Metabolism (<http://biocyc.org/>)

MetaCyc: Metabolism from 150 species (<http://metacyc.org/>)

SoyBase: Soybean Metabolism (<http://cgsc.biology.yale.edu/metab.htm>)

Boehringer Mannheim: Biochemical Pathways (<http://www.expasy.ch/cgi-bin/search-biochem-index>)

IUBMB-Nicholson Minimaps (<http://www.tcd.ie/Biochemistry/IUBMB-Nicholson/>)

и т.д. (см. слайды 12 и 13)

Слайд 14

Генные сети – сверхбольшие системы со сверхсложной структурно-функциональной организацией и динамическими свойствами.

Анализ огромных объемов экспериментальных данных, отражающих сложные процессы функционирования молекулярно-генетических систем, исследование структурной организации, молекулярных механизмов функционирования, закономерностей эволюции генных сетей, оценки влияния мутаций на функцию генных сетей, реконструкции генных сетей на основе экспериментальных данных и т.д., принципиально невозможен без

использования современных информационных технологий и эффективных математических методов анализа данных и моделирования биологических систем и процессов.

Слайд 15

На данном слайде представлен пример сложно устроенной генной сети macrophage activation by the IFN-gamma and lipopolysaccharides. Из рисунка можно видеть насколько запутанной может быть структурно-функциональная организация генной сети.

Слайд 16

В генных сетях можно выделить 5 основных элементов, которые их составляют:

1. Ядро генной сети - группа координировано работающих генов.
2. РНК и белки, кодируемые этими генами.
3. Низкомолекулярные компоненты, осуществляющие переключение функции генных сетей в ответ на внешние воздействия (гормоны и другие сигнальные молекулы), энергетические компоненты, различные метаболиты и т.д.
4. Пути передачи сигналов, обеспечивающие активацию или подавление экспрессии генов.
5. Отрицательные и положительные обратные связи, стабилизирующие параметры генной сети на определенном уровне или, напротив, отклоняющие их от исходного значения, обеспечивая переход к новому функциональному состоянию.

Слайд 17

В чем состоит специфика объектов моделирования? Рассматриваемые системы могут состоять из сотен и тысяч компонентов с очень сложной, многоуровневой организацией. В свою очередь это находит отражение в ее и имеющей свою иерархию структурно-функциональной организацией. Можно выделить несколько уровней организаций генных сетей, которые, в свою очередь, учитываются при моделировании этих объектов. Самый нижним уровнем является биохимический уровень. Здесь рассматривается огромная сеть биохимических процессов, происходящих в клетке или организме. Следующим уровнем является генетический. Здесь уже учитывается расположение и ориентация генов и регуляторных элементов, полиаллельность и т.д. Затем можно выделить клеточный уровень, где учитываются процессы активного и пассивного переноса, компартментализация и т.д. Также можно выделить более высокие уровни организации –органый, тканевой, организменный, популяционный и т.д. уровни. В большинстве случаев, объекты моделирования обладают сложным нелинейным поведением в силу наличия в них отрицательных и положительных обратных связей

Слайд 18

Биологические системы является самовоспроизводящимися динамическими системами.

Особенностью функционирования биологических систем во времени является то, что у них, как правило, в процессе функционирования не остаются постоянными параметры пространственной организации, а также могут меняться характеристики генетического строения.

Факторы изменения пространственной организации строения биологической системы: рост организмов, регенерация органов и тканей, деление и гибель клеток, органов, размножение организмов, метаморфоз и т.д.

Факторы изменения генетического строения биологической системы:

репликация, рекомбинация, мутации, делеции, вставки, деление клеткок, слияние половых клеток, трансгенез, привнесение новых геномов или частей генома путем внедрение в биологическую систему других биологических систем, как то вирусов, бактерий, паразитов, параллельный перенос генетической информации и т.д.

Слайд 19

На сегодняшний день развито значительное количество методов моделирования динамики генных сетей и существуют различные их классификации.

Например, методы можно разделить на:

- непрерывные методы
- дискретные методы
- гибридные методы

с привязкой к различным видам формализации времени при моделировании биологических процессов.

Слайд 20

Классификация методов моделирования динамики генных сетей с привязкой к биологическому процессу:

- детерминистический подход (подходы, основанные на применении дифференциальных и интегро-дифференциальных уравнений)
- стохастический подход
- логический подход

Слайд 21

Классификация методов моделирования динамики генных сетей, основанная на типах формализации молекулярно-генетических процессов:

- Булевы сети [Weisbuch, 1986; Kauffman, 1993; Somogyi and Sniegoski, 1996; Huang, 1999; Shmulevich & Kauffman, 2004] и вероятностные Булевы сети [Shmulevich et al., 2002a; Shmulevich et al., 2002b];
- обобщенные логические сети [Thomas & d'Ari, 1990; Thieffry, Thomas, 1995, Thomas et al., 1995; Sanchez et al., 1997; Mendoza et al., 1999];
- Байесовские сети [Friedman et al., 2000; Sachs et al., 2002] и динамические Байесовские сети [Ong et al., 2002; Kim et al., 2004];
- сети Петри [Reddy et al., 1993; Hofstadt & Meineke, 1995, Peleg et al., 2002; Moore & Hahn, 2003], стохастические сети Петри [Goss, Peccoud, 1998], гибридные сети Петри [Matsuno et al., 2000], функциональные гибридные сети Петри [Doi et al., 2004];
- нейронные сети и их обобщения [Taylor, Beiu, 1996; Hanai & Honda H, 2004];
- обобщенный пороговый метод моделирования [Tchuraev, 1991, Prokudina et al., 1991; Tchuraev & Galimzyanov, 2001; Tchuraev & Galimzyanov, 2003];
- подходы, основанные на применении обыкновенных нелинейных дифференциальных уравнений [Bliss et al., 1982; Goldbeter, 1995; Mestl et al., 1995; Endy et al., 1997; Koh et al., 1998; Carrier & Keasling, 1999; Smolen et al., 2000; Chen et al, 2000, 2004; Novak et al, 2001; de Jong, 2002];
- подходы, основанные на применении обыкновенных нелинейных дифференциальных уравнений с запаздывающим аргументом [Landahl, 1969; Mahaffy, 1984; MacDonald, 1989; Smolen et al., 2000, Лихошвай и соавт., 2004];
- обобщенный химико-кинетический метод моделирования [Likhoshvai et al., 2000; Ratushny et al., 2000a; Ratushny et al., 2000b; Лихошвай и соавт., 2001; Nedosekina et al., 2002; Ратушный и соавт., 2003; Ratushny et al., 2003];
- подходы, основанные на применении дифференциальных уравнениях в частных производных (модели пространственно распределенных систем) [Turing, 1951; Meinhardt, 1977, 1982, 2000, 2004; Bunow, 1980; Gierer, 1981; Britton, 1986; Kauffman, 1993; Mjolsness et al, 1991; Reinitz et al, 1995; Reinitz & Sharp, 1995; Sanchez et al, 1997];
- стохастическое моделирование [Gillespie, 1977; McAdams & Arkin, 1997; Gibson & Bruck, 2000; Gillespie, 2001; Kepler & Elston, 2001; Rao & Arkin, 2003; Rathinam et al., 2003, Turner et al., 2004; Vladimirov & Likhoshvai, 2004];

Слайд 22

Графическое моделирование – это один из способов моделирования структурно-функциональной организации генных сетей. На слайде представлены несколько примеров орграфов простых систем. Орграф G состоит из конечного множества V вершин и набора E упорядоченных пар различных вершин. Любая такая пара (i, j) называется дугой, или ориентированным ребром, и обычно обозначается ij .

Слайд 23

Графические модели генных сетей строятся на основе информации, хранящейся в базах данных и базах знаний, но они также могут быть реконструированы на основе данных по экспрессии генов посредством подходов *inductive* либо *reverse engineering*.

Для выявления паттернов экспрессии (*similar temporal expression patterns*) применяются различные алгоритмы кластеризации экспрессионных данных.

Для дополнительной верификации структуры генной сети и выявления мнимых регуляторных связей применяется ряд различных методик:

- анализ временных запаздываний (*time lags*);
- проведение экспериментов по возмущению системы (*perturbation experiments*);
- картирование коэкспрессирующихся генов на выполнение определенной функции в организме (*mapping of co-expressed genes to pathway functions*);
- скрининг промоторных областей коэкспрессирующихся генов для выявления одинаковых регуляторных районов.

Слайд 24

Булевы сети

Модели булевых сетей (*Boolean networks*) основаны на булевой логике. Предполагается, что каждый ген может находиться в двух состояниях: “экспрессируется”, либо “не экспрессируется”. Для простоты, эти состояния обозначаются “1” и “0”, соответственно. Состояние всей модели описывается списком экспрессирующихся и не экспрессирующихся генов.

Слайд 25

На данном слайде приведен пример булевой функции для трехгенной системы (XYZ):

$$X(\text{next}) = X(\text{current}) \text{ AND } Y(\text{current})$$

$$Y(\text{next}) = X(\text{current}) \text{ OR } Y(\text{current})$$

$$Z(\text{next}) = X(\text{current}) \text{ OR } \{[\text{NOT } Y(\text{current})] \text{ AND } Z(\text{current})\}$$

Слайд 26

В нашем институте развивается оригинальный обобщенный химико-кинетический метод моделирования (*ОХКММ*) который ориентирован на формализованное, в первую очередь портретное, описание функционирования произвольных биосистем.

Слайд 27

В рамках *ОХКММ* доступны и развиваются средства, которые позволяют эффективно с минимальными для пользователя затратами, моделировать такие структурные и функциональные особенности генных сетей, как:

- Цис- и транс- эффекты;
- Взаимное расположение и ориентацию генов и их регуляторных сайтов;
- Матричность фундаментальных процессов (репликация, транскрипция, трансляция);
- Полиплоидия;
- Генетические перестройки, рекомбинация, кроссинговер;
- Мутации;

- Многокомпарментность;
- Многовариантность;
- и т.д.

Как это делается, будет рассказано в одной из следующих лекций.

Слайд 28

На данном слайде проиллюстрирована методология моделирования в рамках ОХКММ, которая заключается в очень простом и можно сказать фундаментальном принципе (блочном принципе). Согласно этому принципу моделируемая биологическая система расчленяется на элементарные подсистемы, каждая из которых описывается изолированно от других. Описание элементарных подсистем проводится в терминах элементарных процессов. Элементарные процессы описываются на основе формальных блоков. Затем строятся модели элементарных подсистем, на основе которых конструируется математическая модель исходной биологической системы и ее различных модификаций.

Слайд 29

Формализация в рамках ОХКММ осуществляется на основе блочного принципа

- Описание элементарных подсистем проводится в терминах элементарных процессов.
- Основной принцип локальная независимость процессов.
- Элементарные процессы описываются на основе формальных блоков.
- Формальные блоки однозначно характеризуются (1) упорядоченным списком формальных динамических переменных (X), (2) упорядоченным списком формальных параметров (P) и (3) законом преобразования информации (F).

Слайд 30 31

На этом слайде приведены некоторые простые примеры формальных блоков, таких как бимолекулярная обратимая реакция, мономолекулярная необратимая реакция, обобщенная реакция Михаэлиса-Ментен. X и P с чертой – это векторы динамических переменных и параметров, соответственно. F – закон преобразования информации (в данном случае закон действующих масс).4.

Слайд 32

5. Регуляторные процессы

На данном слайде представлена одна из схем формализации регуляторных процессов, в которой выделяются:

n_1 – активаторов (A_i) и m_1 – ингибиторов (I_i)

по аддитивному механизму;

n_2 – активаторов (A_j) и m_2 – ингибиторов (I_j)

по мультипликативному механизму

Слайд 33

На данном слайде приведены примеры формальных блоков иной природы. 4-ый блок – блок суммирования (арифметический блок). 5-ый и 6-ой – это дискретные формальные блоки, пороговый и стохастические автоматы, соответственно. В общем случае формальные блоки могут быть любой природы.

Слайд 34

На данном слайде представлены примеры описания элементарных событий генной сети регуляции биосинтеза холестерина в клетке в рамках ОХКММ. Слева представлена данная генная сеть из базы данных GeneNet. Выборочно можно рассмотреть описание нескольких биохимических процессов. Этот процесс, переход одного вещества в другое, может быть

описан как мономолекулярная реакция, с соответствующим видом правой части, этот, переработка определенного субстрата в определенный продукт, как ферментативная, а этот, димеризация транскрипционного фактора, как бимолекулярная и т.д.

Данный вид описания относится к наиболее нижнему уровню организации генных сетей – биохимическому. (Все модели в конечном итоге представляются в виде списков элементарных процессов, которые затем могут использоваться различными расчетными программами)

Слайд 35

На данном слайде в качестве иллюстрации фрагмент системы дифференциальных уравнений, как конкретная реализация математической модели генной сети регуляции биосинтеза холестерина в клетке. В этой системе представлен суммарный закон изменения концентрации холестерина в клетке, рецепторов липопротеинов низкой плотности и т.д. В таблице представлены константы, используемые в данном фрагменте системы ОДУ, их значения и размерности.

Слайд 36

На данном слайде приведены способы представления генных сетей в виде двудольных орграфов на примерах конститутивного синтеза, мономолекулярной реакции, бимолекулярной реакции, ферментативной реакции.

Слайд 37

На данном слайде приведен двудольный граф математической модели гаплоидной генной сети регуляции биосинтеза холестерина в клетке, которая описывает динамику 39 переменных, участвующих в 74 процессах. Сплошная стрелка означает участие в процессе с ненулевой стехиометрией, пунктирная – с нулевой. 1- конститутивный синтез, 2 - утилизация, либо деградация, 3 - мономолекулярная реакция, 4 - бимолекулярная реакция, 5 – ферментативная реакция.

Слайд 38

Создание портретных математических моделей функционирования живых клеток, детально и с исчерпывающей полнотой описывающих процессы внутриклеточного метаболизма и его генетической регуляции - одна из ключевых проблем системной компьютерной биологии. Ее называют также проблемой *электронной клетки*. В настоящее время исследователи очень далеки от возможности построения полных математических моделей эукариотических клеток. Что же касается клеток прокариотических, создание их полных математических моделей – дело самого ближайшего будущего.

В настоящей лекции мы рассмотрим некоторые моменты, касающиеся возможности автоматического конструирования компьютерных динамических моделей метаболических путей и их генетической регуляции на основе информации, хранящейся в различных базах данных. Автоматическое конструирование математических моделей генных сетей и дальнейшей их адаптации может стать основой (первым шагом) для создания электронной клетки.

На данном слайде приведен скриншот одного из метаболического пути (цикла трикарбоновых кислот), представленного в базе данных KEGG (<http://www.genome.ad.jp/>).

Слайд 39

На данном слайде приведен двудольный граф математической модели цикла трикарбоновых кислот, сгенерированной на основе графа связей, представленного выше. Модель описывает динамику 73 переменных, участвующих в 139 процессах. Сплошная стрелка означает участие в процессе с ненулевой стехиометрией, пунктирная – с нулевой.

Конечно же, данный вид формализации является лишь первым приближением, который в дальнейшем может быть уточнен, с использованием различных алгоритмов верификации моделей на основе имеющихся количественных экспериментальных данных.

Слайд 40

С помощью данной технологии были построены модели для 101 метаболических путей (см. слайд) для организма *Escherichia coli* K-12. В базе данных KEGG (<http://www.genome.ad.jp/>) представлена информация для более чем 100 организмов.

Слайд 41 и 42

На данном слайде приведен двудольный граф математической модели общего метаболизма, представленного базе данных KEGG (<http://www.genome.ad.jp/>), с возможностью спецификации на различные организмы (т.е. устанавливать фильтры на ферментативные реакции, которые не идут в тех или иных организмах). Модель описывает динамику 4036 переменных, участвующих в 3973 процессах. Сплошная стрелка означает участие в процессе с ненулевой стехиометрией, пунктирная – с нулевой.

Слайд 43 и 44

Автоматическое конструирование компьютерных динамических моделей регуляции экспрессии генов также возможно. Например, моделирование регуляции инициации транскрипции на основе информации о взаиморасположении регуляторных участков в промоторных областях.

На данных слайдах приведены несколько примеров оперонов *Escherichia coli* K-12, структура которых представлена в базе данных EcoCyc (<http://www.ecocyc.org>).

Слайд 45

Также возможно конструирование компьютерных динамических моделей формирования ферментативных комплексов на основе графа взаимодействия отдельных субъединиц.

На данных слайдах приведены примеры образования различных ферментативных комплексов из белковых субъединиц, кодируемых генами *gcvH*, *gcvT*, *gcvP*, *lpd*, *sucA*, *sucB*, *aceF*, *aceE* в организме *Escherichia coli* K-12. Данные взяты из БД EcoCyc (<http://www.ecocyc.org>). Компьютерная динамическая модель, описывающая данную схему регуляции формирования ферментативных комплексов включает в себя 104 процесса

Слайд 46

На данном слайде приведен двудольный граф компьютерной динамической модели генной сети регуляции цикла трикарбоновых кислот в организме *Escherichia coli* K-12. Модель описывает функционирование 757 переменных и различных состояний системы, в системе учитываются 1882 процесса. Сплошные стрелки указывают на участие в процессе с ненулевой стехиометрией, пунктирные – с нулевой.

Слайд 47

Необходимо сказать несколько слов о проблемах, которые возникают при моделировании динамики генных сетей. Они разделяются на два вида. Первые - это математические проблемы, которые возникают на стадии построения, расчета и адаптации моделей. К ним относятся такие проблемы как: решение систем дифференциальных уравнений; также то, что модели часто состоят частично из системы ДУ и дискретной части; решение обратной задачи (адаптация и определение параметров модели) и т.д. Вторые – это методологические проблемы, которые возникают на стадии работы с моделями (собственно для чего строятся модели). К ним относятся такие проблемы как: исследование устойчивости функционирования систем; решение задач управления; исследование эволюции моделируемых биологических систем и т.д.

Этим проблемам будет посвящено несколько следующих лекций, в которых более детально будут рассмотрены данные вопросы.

Слайд 48

Одной из наиболее важных и актуальных проблем при построении математических моделей генных сетей является верификация параметров модели. Здесь можно выделить основные моменты. Это пробелы в знании структурно-функциональной организации некоторых частей генных сетей (посредники, механизмы протекания процессов и т.д.). Нехватка количественных данных, пригодных для адаптации конкретной модели. В базах данных в основном накапливаются качественная описательная информация (информация о структурно-функциональной организации генной сети, механизмы различных процессов, организация регуляторных участков генов и т.д.), а также количественная информация статического характера (константы ферментативных реакций, массы белков, и т.д.). Рассредоточение кинетических данных в огромном количестве научных публикаций. Для данной области знаний характерна слабоструктурированность данных. Гетерогенность данных (наборы экспериментальных динамических данных получаются в своих конкретных условиях, при разных экспериментальных воздействиях в различные моменты времени и т.д.).

Слайд 49

На данном слайде приведен пример получения либо оценки статических данных модели. Например, константы многих ферментативных реакций доступны в базах данных, таких как WIT, BRENDA и т.д. Могут быть оценены константы таких фундаментальных процессов как репликация, транскрипция, трансляция и т.д. Также могут быть оценены константы, характеризующие взаимодействие генной сети с ее клеточным и организменным окружением, исходя из равновесных концентраций; времен жизни или полураспада компонентов системы; интегральных характеристик и т.д.

Слайд 50

- Математическая модель генной сети характеризуется набором констант k_1, \dots, k_m в системе дифференциальных уравнений
- Как правило, экспериментально измеренные значения известны только для ограниченного числа констант
- Значения остальных констант определяются на основе численного эксперимента

Слайд 51

На данном слайде приведен элемент базы данных по количественным характеристикам, которая разрабатывается в нашем институте.

Слайд 52

Верификация структуры математических моделей, а также их параметров может проводиться на основе различных оптимизационных алгоритмов. Одним из самых распространенных из них является генетический алгоритм, блок схема которого приведена на данном слайде. Об этом алгоритме более подробно будет вам рассказано в одной из следующих лекций.

Слайд 52

В общем случае, после верификации структуры математической модели проводится поиск значений ее параметров k_1, \dots, k_m , которые обеспечивают максимальное соответствие между рассчитанным и наблюдаемым динамическим поведением генной сети для большого количества экспериментов, при этом обычно оптимизируются различного сорта функционалы, 2 примера которых приведены на слайде.

Слайд 53

На данном слайде приведена блок-схема процедуры верификации параметров модели методом численного эксперимента с использованием технологии построения сценариев. На сегодняшний день в нашем институте разрабатывается база данных по количественным характеристикам генных сетей. В БД входят статические и динамические данные. На основе статических данных строится модель генной сети. Учет специфики динамических данных выражается в необходимости настройки модели на каждый конкретный эксперимент. Модель, настроенная на конкретный эксперимент, называется сценарием. Учет большое количество экспериментов приводит к возникновению пакета сценариев, который используется для адаптации модели генной сети как единое целое. Процесс адаптации организован как замкнутый цикл, в котором важнейшую часть занимает экспертная оценка. При появлении новых данных, несогласующихся с моделью, эксперт может принять решение, например, об изменении структуры модели.