

О ВОЗМОЖНОСТИ ПОСТРОЕНИЯ ИНФОРМАЦИОННО-ЭКСПЕРТНОЙ СИСТЕМЫ ДЛЯ АСТРОМЕТРИИ

М.В.Воротков, В.Л.Горшков, Н.О.Миллер

Изложены основные принципы и методы построения базы знаний для астрометрии. Описана структура программной реализации рабочей среды базы знаний. Для демонстрации возможностей разрабатываемой рабочей среды использованы ряды параметров ориентации Земли.

ВВЕДЕНИЕ

Роль классической наблюдательной астрометрии в той её части, которая занималась определением **параметров ориентации Земли (ПОЗ)**, следует считать в основном завершённой. Это относится как к проблеме определения ПОЗ, так и к тесно с ней связанной проблеме уточнения опорной системы координат с помощью классических астрометрических средств наблюдений. Обе эти задачи решены с точностями, на два порядка превосходящими классическую службу ПОЗ: с одной стороны, современными техническими средствами определения ПОЗ (РСДБ, ЛЛС/Л, ГЛОНАСС/GPS) и высокоточной системой координат звезд в системе ICRS (Hipparcos) с другой. Субмиллисекундная точность мониторинга ПОЗ позволила значительно улучшить модель вращения Земли. Однако имеется ряд проблем, связанных с нестыковкой принятой модели с наблюдениями. В частности, чандлеровское движение полюса с амплитудой, во много раз превышающей уровень современной точности определения ПОЗ, не укладывается в существующую модель вращения Земли и допускает различную физическую интерпретацию [1], [2], [3].

Это одна из многих научных проблем, в которых требуется нестандартное, эвристическое обобщение неоднозначной, неполной, а иногда и противоречивой информации. Для решения подобных задач создаются специальные **базы данных (БД)**, **базы знаний (БЗ)** и информационно-экспертные системы (ЭС) [4]. К сожалению, сильная специализированность этих систем вынуждает пользователей тратить значительные усилия на их адаптацию, часто безуспешную, или создавать свою специализированную систему.

В данной работе сделана попытка сформулировать и частично реализовать некоторые принципы построения **эвристической базы знаний (ЭБЗ)**. Условимся в дальнейшем понимать под ЭБЗ рабочую среду, помогающую сопоставлять элементы информации в любом их представлении и создавать на их основе обобщающие структуры. ЭБЗ не берет на себя функции ЭС, а лишь обеспечивает наиболее продуктивное творческое участие исследователя в анализе информации. Это средство должно обеспечить восприятие материала как некоторой целостности, иерархически организованного объекта, что дает возможность увидеть новые его стороны или существование недостающих элементов. Образно говоря, предлагается разработать средство, которое позволит заменить тип взаимодействия с компьютером с отношения «человек и калькулятор» на отношение «человек и музыкальный инструмент».

ОБОСНОВАНИЕ ВОЗМОЖНОСТИ СОЗДАНИЯ ЭБЗ.

Хорошо известны трудности, с которыми сталкивается исследователь при осмыслении больших объёмов информации с целью поиска на её основе новых закономерностей даже с помощью компьютерных средств. Так, обычный человек не в состоянии одновременно оперировать более чем 8-10 объектами. Трудности возникают также при необходимости оперировать неполной или недоопределённой информацией. Целостному, синтезирующему восприятию данных не способствует укоренившийся последовательный, древоподобный характер построения логических конструкций. Формальной методики обобщений не существует, это творческий по своей природе процесс.

Отсюда возникают и технические трудности в разработке универсальных средств эвристического поиска. Существующие средства для создания ЭС немногочисленны, ориентированы, как правило, на рабочие станции, очень дорогостоящи и, вследствие этого, малодоступны для индивидуальных пользователей. Для примера можно назвать графическую объектно-ориентированную среду для создания ЭС корпорации Gensym «G2» [5] и объектно-реляционную СУБД Международной Компьютерной Ассоциации «JASMINE». Создание же «с нуля» мощной среды проектирования кажется невыполнимой задачей для небольшой группы разработчиков и, тем более, отдельных пользователей, которые хотят решать свои конкретные задачи, а не заниматься программированием.

Ответом на сформулированные проблемы может быть развитие следующих программно реализуемых предложений:

- Использование визуализированного, объектно-ориентированного подхода, существенной составляющей которого является требование к художественности используемых средств. Представление отношений между объектами осуществляется средствами цветопередачи, формой, компоновкой и т.п. Использование нелогических способов восприятия в сочетании с формальной привязкой элементов БД к указанным атрибутам графических объектов и дальнейшей их программной обработкой решает проблему идентификации информации при неполном знании [4] (нет необходимости называть - достаточно указать, разместить или обозначить). Существенным моментом при этом является то, что художественные атрибуты в самой своей природе несут некоторый естественный принцип систематизации.

- Использование специального динамического структурирования материала позволяет оперировать им в целом, имея при этом в одновременном рассмотрении небольшое количество структурных единиц информации.

- Следующее предложение требует небольшого пояснения. Многие естественные объекты и явления демонстрируют иерархически организованные самоподобные структуры. Как следствие, наши представления о них должны обладать аналогичным структурным принципом, связывающим объекты в понятийном пространстве [5]. Назовем его принципом естественного структурирования, а таким образом организованные **системы естественно структурированными (ЕСС)**. Этот принцип может быть использован для выявления недостающих элементов и связей, а также как источник дополнительной информации при неполном знании об объекте.

Принцип ЕСС возник из попыток отобразить и воспроизвести иерархию и соотношение объектов в максимально возможной полноте. Этот принцип интуитивно хорошо чувствуется. Можно было бы назвать его принципом фрактальности, если бы не перегруженность этого понятия в современной научной терминологии. ЕСС является основой, которая делает практически реализуемым описываемый подход к созданию ЭБЗ.

Естественное структурирование в системе реализуется благодаря следующим условиям:

1. Наличие одного или нескольких параметров, потенциально позволяющих установить иерархические отношения между элементами.

2. Наличие универсальных алгоритмов, каждый из которых действует на нескольких смежных иерархических уровнях.

3. Применение этих алгоритмов на одном из уровней в определённой степени параметризует их последующее применение на других уровнях.

4. Степень этого влияния быстро убывает при иерархическом удалении от области его применения.

Два последних положения призваны обеспечивать гибкость и устойчивость системы.

Эти условия можно рассматривать как попытку сформулировать представление о механизме, порождающем природные самоподобные иерархические структуры. (Проявление некоторого процесса в одном пространственно-временном масштабе влияет на его проявление в другом).

Реализация вышеизложенных условий приводит к возникновению сложной многоуровневой конструкции, наиболее существенной особенностью которой является то, что «заселённость» уровней и подуровней не равномерна, а плотность связей между элементами убывает в зависимости от иерархического удаления последних друг от друга.

Трудоемкость технического осуществления программного проекта в значительной степени определяется его универсальностью. Создать узкоспециализированное приложение несложно, но его адаптивность мала. Однако можно совместить этапы проектирования, разработки и даже эксплуатации ЭБЗ таким образом, чтобы некоторое специализированное приложение могло оперативно развиваться в процессе эксплуатации в любых других направлениях. Для обеспечения такой возможности в создаваемой структуре:

- каждый элемент должен допускать коррекцию вне зависимости от своего иерархического положения (этому требованию удовлетворяет классическая объектно-ориентированная модель),

- каждая функция или задача также должна иметь возможность независимой коррекции (это условие выполнить значительно сложнее, т.к. в общем случае функционирование системы может быть циклично и рекурсивно),

- устанавливается «мягкое» подчинение между иерархическими структурами, при котором модификация старших структур полуавтоматически корректирует функционирование младших.

Предлагаемый подход реализует интерфейсно развивающуюся систему. Его элементы становятся характерными для современных популярных программных продуктов.

ПРОЕКТ ЭБЗ «BISER»

Практическая реализация данного подхода осуществляется в течение нескольких лет небольшой группой разработчиков. Потребность в эвристической систематизации и предоставлении большого объема слабо формализованной информации побудила искать нестандартные пути создания адекватных данной задаче средств.

DOS- приложение

Первым законченным этапом разработки можно считать создание интегральной среды, воплотившей в себе описываемые выше идеи как изолированное DOS - приложение (в настоящее время разрабатывается Windows - приложение). Написанный на языке C++ пакет имеет структуру, удовлетворяющую требованиям ЕСС. Несмотря на использование стандартных средств объектно-ориентированного программирования, в структуре пакета реализованы перекрестные и рекурсивные соотношения между объектами. Это обеспечивается иерархической структурой пакета, в которой библиотеке классов отводится лишь один уровень.

Функционально пакет представляет реляционную БД, графическую среду с редактором векторного и растрового формата и среду визуального программирования с многоуровневым языком-интерпретатором.

Рабочее пространство BISER - условно-бесконечное поле со стандартными возможностями масштабируемого просмотра. В любом месте этого поля могут быть размещены точки. С каждой точкой может быть связан объект. Совокупность размещенных объектов образуют схему. Объектами являются векторные примитивы, растровые элементы, элементы встроенной базы данных или просто текст, выполняемые модули, а также связи между объектами. Объектом может быть также любое количество объектов, любое количество их повторений и схема в целом. Объект может быть создан как интерфейсно (манипуляции «мышью», копирование), так и написанием кода. Любая схема может быть прорисована и /или выполнена как последовательность исполняемых объектов этой схемы. Разделение на прорисовку и выполнение условно, так как функции прорисовки объектов абсолютно равноправны с остальными функциями.

Язык программирования имеет пять уровней, три верхних из которых реализованы как интерфейс среды, а остальные два написаны на C++ и требуют компиляции. Отсутствие четких границ между интерфейсом и языком программирования, средой и приложением, пользовательской частью и внешним программным продуктом дает возможность, сохраняя целостность системы, развивать её в любом желаемом направлении .

За время работы над проектом произошли сильные изменения, касающиеся как доступного программного обеспечения, так и идеологии программирования в целом. Стало ясно, что описанный этап разработки следует рассматривать как эксперимент в реализации вышеизложенных идей.

Моделирование рабочей среды средствами Microsoft Office и Corel

Проверкой универсальности найденных принципов и методики стал эксперимент по их реализации средствами EXCEL, WORD, CORELDRAW. Была создана рабочая среда, использующая рабочие пространства EXCEL и CORELDRAW. Встроенные диалекты Visual Basic позволили создать надстройку развиваемого языка (также визуального) следующего уровня. Несмотря на скромные вычислительные возможности, слабое СУБД и громоздкие форматы документов, средства Microsoft Office выгодны мощным объектно-ориентированным графическим редактором, наличием авторекодера макрокоманд и, главное, дружественным интерфейсом. В этой среде начато формирование проблемной и вычислительной областей для астрометрической базы знаний (АБЗ).

Вариант структурного построения ЭБЗ

Данный раздел можно рассматривать как расширенную подпись к рисунку 1, созданному, кстати, в среде BISER, также как и вся структура данной статьи. Разделение BISER на программную среду и ее приложение весьма искусственно. Поэтому развитие рабочей среды в соответствии с сильно различающимися между собой задачами может привести к расслоению одной среды на несколько, имеющих в дальнейшем самостоятельное развитие. Каждая из дочерних сред может также расслоиться или, наоборот, слиться с другой в процессе решения новых проблем. Описываемый процесс представляет формирование следующего иерархического уровня и отражается в пополнении утилитарной БЗ. Этот процесс аналогичен манипуляциям с документами в среде Microsoft Office, но удельный вес изменений может быть столь значителен, что естественно образование самостоятельного продукта.

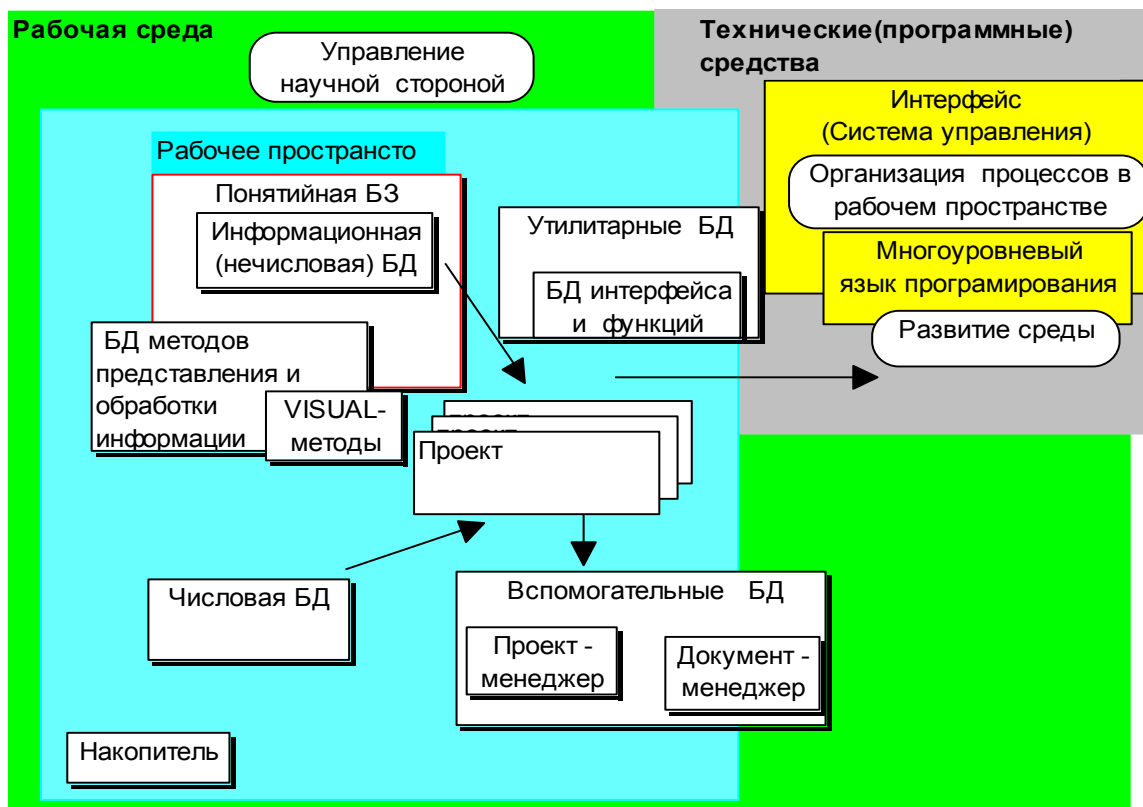


Рис.1 Общая структурная схема рабочей среды BISER.

Рабочая среда включает следующие обязательные элементы.

Рабочее пространство

Поле, на котором происходит манипуляция с объектами на экране. В простейшем случае это равномерно масштабируемое двумерное пространство, дополненное структурой ячеек или строк. Возможны более изощренные формы рабочего пространства: неравномерно масштабируемое, связка зависимо масштабируемых пространств. Строго говоря, к рабочему пространству предъявляется невыполнимое в общем случае требование - быть адекватной областью понятийного пространства. В реальных условиях акцент делается на дополнение топологических средств изобразительными и динамическими.

Проект

Проблемно-ориентированный компонент (совокупность всего, что имеет отношение к данной конкретной задаче). Как наиболее яркий носитель иерархической вложенности и самоподобия, проект может автономно содержать все перечисляемые ниже основные проблемные компоненты рабочей среды. С одной стороны, он может содержать другие более мелкие проекты, а с другой стороны, границы его могут быть размыты (один и тот же проблемный компонент присутствует в разных проектах).

Понятийная БЗ

Является центральным компонентом рабочей среды, ключом к получению числовой и нечисловой информации, обеспечивает планирование, восприятие проблемы в целом, эвристический поиск. Она тесно связана с информационной БД, которая содержит текстовую информацию разного уровня и объема (комментарии, справки, статьи), а также графические и специальные форматы.

Числовая БД

Ограниченно визуализирована. Может включать помимо числовых данных средства их создания.

БД средств представления и обработки информации

Банк исполнимых программ и функций обработки, дополненный систематизирующей их понятийной структурой. Особое место здесь занимают визуальные методы, средства компьютерной графики и ее динамические возможности. В какой-то степени эти методы входят в пользовательскую часть развиваемого интерфейса среды и создаются по мере надобности.

Вспомогательные БД

Отображают процесс работы над проблемой и обеспечивают доступ к предыдущим этапам или родственным проектам. Включают БД интерфейса и функций - постоянно изменяемый и пополняемый компонент утилитарной БД, помогающий пользователю ориентироваться в избыточном многообразии средств,

Утилитарные БД

Отображают процесс формирования рабочей среды и не имеют прямого отношения к проблемной области. Наиболее существенное место среди них занимают :

- ◆ проект-менеджер - средство планирования и отображения процесса работы над проблемой,
- ◆ документ-менеджер - обеспечивает поиск и доступ к структурным компонентам.

Накопитель

Средство, позволяющее временно разместить и обработать (систематизировать) блок информации, для которого пока непонятно или труднодоступно место в общей структуре БЗ. Присутствует практически на всех уровнях описываемой конструкции и допускает вложенность.

Технические (программные) средства

Основную массу технических возможностей предоставляют общедоступные базовые программные средства. Кроме того, в набор средств могут входить любые специальные продукты. Однако ядром среды служит компактный пакет специальных средств, несущих идеологию ЕСС. Это ядро развивается отдельно для каждой рабочей среды.

Интерфейс (Система управления)

Состоит из интерфейса используемых стандартных пакетов, универсальных средств, разработанных для данной рабочей среды, создаваемых пользователем элементов управления.

Многоуровневый язык программирования

Специальная многоуровневая организация языка программирования:

- нижний уровень языка определяется языком базовых программных средств;
- промежуточные уровни языка создаются в соответствии с динамическими требованиями ЕСС, включаются в ядро среды, что однако не является обязательным, поскольку формируемый язык может быть узко специализированным;
- верхний уровень языка плавно переходит в интерфейс.

Демонстрационный пример.

Работа над АБЗ находится в начальной стадии систематизации информации. Тем не менее, уже на этой стадии оказалась возможной достаточно эффективная поисковая научная работа.

В качестве демонстрационного примера работы над АБЗ представим узкий срез проблемы интерпретации данных параметров ориентации Земли, имеющий некоторую эвристическую окраску. Единственный источник информации для выводов - временные ряды ПОЗ [6]. Результат гармонического анализа этих рядов допускает три типа физической интерпретации.

А. (Наиболее популярная точка зрения) Наблюдаемая картина - суперпозиция двух главных физических компонент: годовой P_1 и чандлеровской P_4 (1,2 года). Природа первой в общих чертах очевидна, в то время как чандлеровское движение полюса вот уже почти век остается без убедительной физической интерпретации, как уже отмечалось во введении.

В. Период 1,2 года - порождение гармонического анализа, а физически существует шестилетний процесс P_6 , модулирующий годовой. Строгое соответствие с наблюдаемыми данными получается в предположение о суперпозиции двух амплитудных модуляций, а именно:

$$\begin{array}{ccccccc} P_1 & & P_4 & & P_1 & & P_6 \\ \sin(2\pi t) & + & \sin(2\pi t/1.2) & = & \sin(2\pi t) & + & \sin(2\pi t) \sin(2\pi t/6+f) - \sin(2\pi t+f) \sin(2\pi t/6) \end{array}$$

где $f = \pi/2$, t - тропический год. Параметр f может быть интерпретирован как фазовое запаздывание. Отсутствие очевидного природного фактора, задающего шестилетний период, и соответствующего механизма модуляции годового периода препятствуют развитию этого подхода.

С. Реально существуют все три периодических процесса - годовой, чандлеровский и модулирующий их шестилетний. На первый взгляд, эта точка зрения только наследует трудности двух предыдущих. Но ее рассмотрение приводит к продуктивным обобщениям.

В сложных системах суперпозиция компонент может подавлять периодичность. Параметризирующие же отношения в сложной системе, такие как фазовая, частотная или амплитудная модуляция, при наличии синхронизации, наоборот, могут способствовать усилению периодичности. Например, взаимная модуляция двух- и трехгодичных периодических процессов порождает шестилетний, а трех- и четырехлетних - двенадцатилетний.

Можно предложить механизм, способный породить, в частности, устойчивый шестилетний период. Предположим, что в системе присутствует строго периодический процесс P_i с периодом T и другой процесс P_a с периодом T_1 , находящийся в автоколебательном режиме. Если параметры последнего имеют фазовую зависимость от первого процесса (например, сброс приурочен к определенной его фазе), то может возникнуть синхронизация процессов $T_1 = nT$, где $n=1,2,3,\dots$. Ясно, что в реальности наиболее устойчивы к изменению параметров периоды с малыми значениями n ($2T$, $3T$ и $4T$). На рис.2 представлена модель такой синхронизации для $n=3$ с помощью рекурсивной функции:

$$X_i = X_{i-1} - \left(\left(\frac{1}{T_x} \sum_{j=2}^{T_x} X_{i-j} \right)^{\frac{1}{3}} + v_i \right) * \left(2 - A \sin^2(2\pi T_i) \right)^{-1},$$

соответствующей кумулятивному процессу, приращение которого состоит из «памяти» на глубину T_x и нормально распределенной случайной компоненты v , находящейся в фазовой зависимости от синхронизирующего процесса с периодом T .

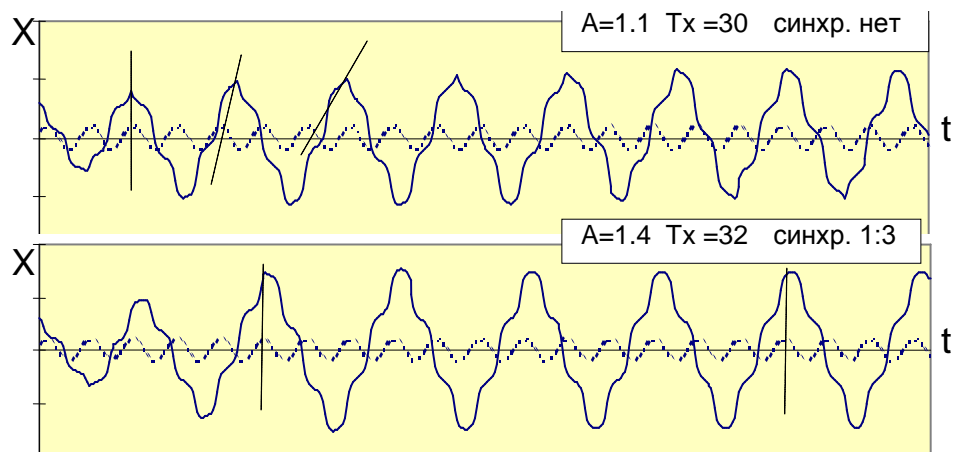


Рис.2

Как уже указывалось выше, взаимная модуляция наиболее устойчивых периодов ($n=2,3,4$), синхронизируемых годовым, порождает 6-летний и 12-летний периоды.

Кроме того, предполагается возможность проявления механизма фазовой синхронизации порожденных процессов. Имеется в виду следующее. Если в системе присутствуют несколько порожденных процессов, имеющих период nT , то они могут иметь фазовое расхождение kT , где

k и n - целые, $k < n$. В этом случае, при наличии причинной связи между этими процессами, могут начать работать достаточно естественные механизмы выравнивания фаз ($k = 0$). Таким образом, в системе устанавливается общий ритм. Если вследствие флуктуаций один из периодических процессов системы “выбивается” из общего ритма, то это не приводит к сбою периодичности системы в целом, однако параметры самого процесса меняются. Элементы предложенного механизма известны в биосферных системах, где сезонная (фазовая) асимметрия может быть ярко выражена.

Рассмотренные типы взаимоотношений между компонентами представлены на рисунке 3.

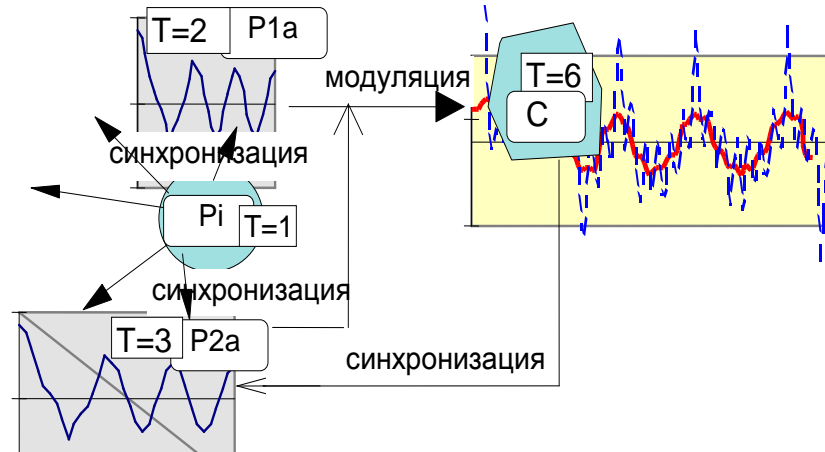


Рис.3

Представим модель следующим рекурсивным соотношением:

$$A_{k,i} = A_{k,i-1} + Ky_{k,i}(a_k + v_k) \sin^2(2\pi T)$$

$$\dots\dots y_{k,i-1} > 0 \wedge A_{k,i-1} > Y1_k - \left| \sum_{j=1}^n C_{i-j} \right| \rightarrow y_{k,i} = -1$$

$$\dots\dots y_{k,i-1} > 0 \wedge A_{k,i-1} \leq Y1_k + \left| \sum_{j=1}^n C_{i-j} \right| \rightarrow y_{k,i} = 1$$

$$\dots\dots y_{k,i-1} \leq 0 \wedge A_{k,i-1} < Y0_k - \left| \sum_{j=1}^n C_{i-j} \right| \rightarrow y_i = 1$$

$$\dots\dots y_{k,i-1} > 0 \wedge A_{k,i-1} \geq Y0_k + \left| \sum_{j=1}^n C_{i-j} \right| \rightarrow y_i = -1$$

$$C_i = A_{1,i} A_{2,i} M$$

где $K = 1$; $a_1 = a_2 = 0.5$; $Y1_1 = Y1_2 = 3$; $Y0_1 = Y0_2 = -3$; $M = 0.1$; $k = 1,2$; $i = 1,2,3 \dots$,

а v_k - рандомизированная величина

На рис. 4 показаны результаты моделирования отображенных выше соотношений.

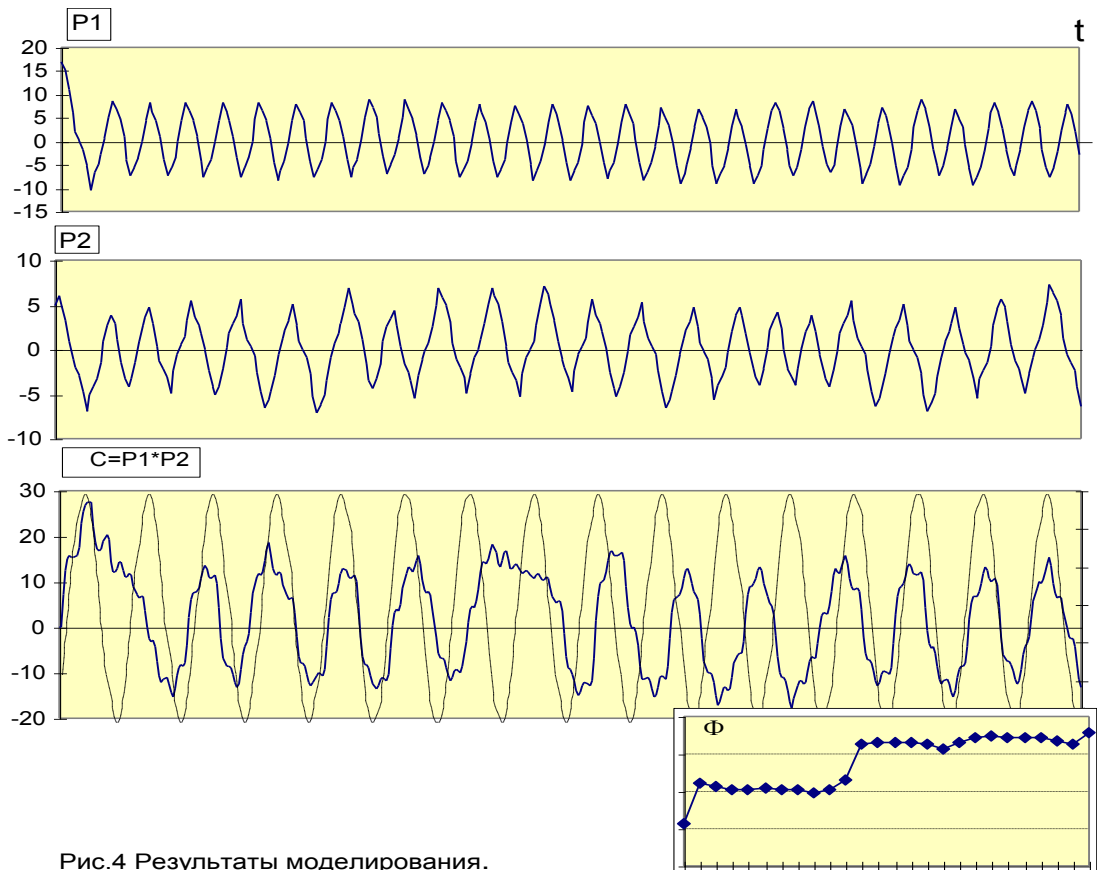


Рис.4 Результаты моделирования.

Два релаксационных пороговых процесса $P1_a$ и $P2_a$ при синхронизации процессом P_i с периодом T становятся «псевдопериодическими» ($T_1=2$ и $T_2=3$). Нормально распределенные случайные компоненты порождают «пропуски» ($T_2=4$). Их взаимная модуляция порождает процесс C со средним периодом $T_c=6$. Обратное влияние C на $P1_a$ и $P2_a$ приводит к появлению промежутков времени, на которых фаза и период (при условности этих понятий для «псевдопериодических» процессов) T_c постоянны и разделены моментами быстрого изменения фазы на величину πn ($n=1,2,3,\dots$), что показано в нижней рамке рисунка 4.

Представленная модель не претендует на моделирование реальных природных процессов, а лишь иллюстрирует вышеприведенные соображения.

В свете сказанного, вопрос о реальности шести- или двенадцатилетнего периода может быть сведен к выявлению конкретных физических механизмов обеспечивающих синхронизирующую параметризацию процессов. Рис. 5 демонстрирует на реальном наблюдательном материале, связанном с движением полюса, предложенный механизм синхронизации. На этом рисунке приведена фаза движения полюса после снятия линейного тренда (векового движения полюса) и строго периодического движения, принятого равным 5 оборотам за 6 лет, то есть приведены угловые невязки в движении полюса относительно её равномерного движения. Данная равномерная угловая скорость движения полюса является приближенно средней за рассматриваемый период. Видно совпадение предложенной модели с реальным восстановлением фазы ($2\pi n$) в движении полюса в 20-х годах.

Даже если рассматриваемые механизмы не вносят существенного вклада в движение полюса, их обнаружение может представлять самостоятельный интерес, при изучении геосферных систем.

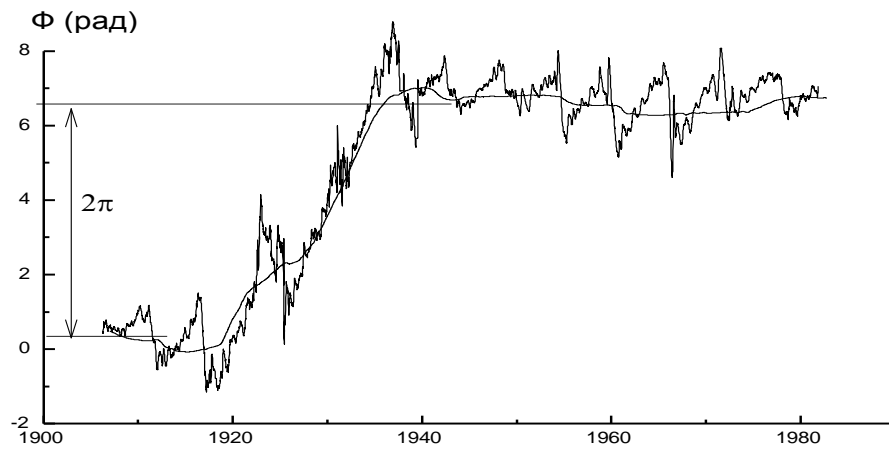


Рис.5 Фаза движения полюса.

В заключение отметим, что демонстрационный пример не претендует на строгость. Работа в этом направлении будет продолжена. Этот пример лишь иллюстрирует метод построения астрометрической базы знаний, к формированию которой авторы приглашают присоединиться всех заинтересованных.

ЛИТЕРАТУРА:

- [1] **Ю.Н.Авсюк**, «Приливные силы и природные процессы», ОИФЗ РАН, М., 1996, с.188.
- [2] **Г.С.Курбасова, Л.В.Рыхлова**, «Чандлеровское движение полюса Земли в системе Земля-Луна», АЖ, 1995, т.72, № 6, с. 945-950.
- [3] **Н.С.Сидоренков**, «Влияние Эль-Ниньо – южного колебания на возбуждение Чандлерова движения полюса», в кн: «Современные проблемы и методы астрометрии и геодинамики», ИПА РАН, СПб, 1996, с. 322-327.
- [4] **А.В. Алексеев, А.Н.Борисов и др.**, «Интеллектуальные системы принятия проектных решений», 1997, Рига, «Зинатне», 319 с.
- [5] **Э.В.Попов, И.Б.Фоминых и др.**, «Статические и динамические экспертные системы», 1996, М., «Финансы и Статистика», 320 с.
- [6] **J.Vondrak, I.Pesek, C.Ron, A.Сепек**, «Earth orientation parameters in 1899.7-1992.0 in the ICRS based on the Hipparcos reference frame», Publ. Astron. Inst. Acad. Sci. Czech. R. №87, 1998.

ON POSSIBILITY OF EXPERT SYSTEM CONSTRUCTION FOR ASTROMETRY

M.Vorotkov, V.Gorshkov, N.Miller

SUMMARY

The first principles and methods for construction of expert system for astrometry are given. The using of image characters for visualization and of fractal structure in object oriented approach are the main means for advanced expert system BISER.