

О ФУНКЦИОНАЛЬНОМ ПРОЕКТИРОВАНИИ БОЛЬШИХ СИСТЕМ УПРАВЛЕНИЯ

В.К. Быховский

Введение

Несмотря на то, что в настоящее время в больших масштабах ведется разработка автоматизированных систем управления (АСУ), теория и методы проектирования АСУ развиты недостаточно. Это находит отражение в интуитивном и кустарном характере проектирования, низкой экономичности разрабатываемых АСУ, больших трудностях при внедрении и сопровождении систем у потребителей и т.д.

Обычно при проектировании АСУ исходят из необходимости реализации такой управляющей структуры (программы управления), которая была бы адекватна некоторому комплексу Г функциональных требований, предъявляемых извне. Логическую структуру этого комплекса Г обычно определяют на стадии предварительного анализа условий функционирования объекта управления.

Однако в условиях конкретного предприятия (АСУП) комплекс Г не является фиксированным — происходит корректировка планов, изменяется номенклатура продукции, происходит техническая переориентация предприятия, изменяются связи данного предприятия с другими предприятиями (поставщиками и потребителями) и т.д.

Это означает, что внешние требования к комплексу Г зависят от времени; зафиксировать Г - означает заморозить обстановку внутри и вне предприятия, что невозможно.

Таким образом, если программа управления спроектирована в расчете на "жесткие" требования Γ_0 , то из-за неизбежных изменений Γ_0 требуется непрерывная корректировка управляющей программы для обеспечения адекватности этой программы новым требованиям $\Gamma_0 + \Gamma$. Это отвлекает силы предприятия - разработчики АСУ становятся несбалансированной ("заплаты"), неэкономичной по использованию ресурсов и т.д.

В настоящей работе предпринимается попытка выяснить происхождение указанных трудностей и наметить возможные подходы к проектированию больших АСУ, основанные на использовании общих концепций, звезденных в теории систем Биром, Энби (см., например, [1,2]) и Лапуновым [3].

Концепция вырождения. Управление как снятие вырождения

Рассмотрим систему С (рис. I), на которую в момент времени t накладывается комплекс ограничений Г. Ограничения включают физические, юридические, финансовые, политические, социальные и другие аспекты, то есть могут иметь весьма общую природу. Так, в транспортных системах управления роль Г играют "Единые правила движения", система ограничительных знаков и т.д. В экономических системах роль Г играют юридические законы (ограничения), приказы и другие документы, обязательные для исполнения. В механических управляющих системах ограничения носят непосредственный физический характер и выделяют из ряда возможных траекторий движения некоторые допустимые ограничения траектории. Вообще говоря, ограничения Г задаются в пространстве, в котором определена количественная мера информации [1]. Качественно комплекс Г можно охарактеризовать некоторой мерой разнообразия.

Система С, в свою очередь, выполнена из материала, состояние которого не может быть однозначно задано с помощью одного или нескольких параметров. Таким материалом может быть, например, память ЭВМ, которая состоит из бинарных элементов (ферритов). Состояние памяти не может быть задано указанием только

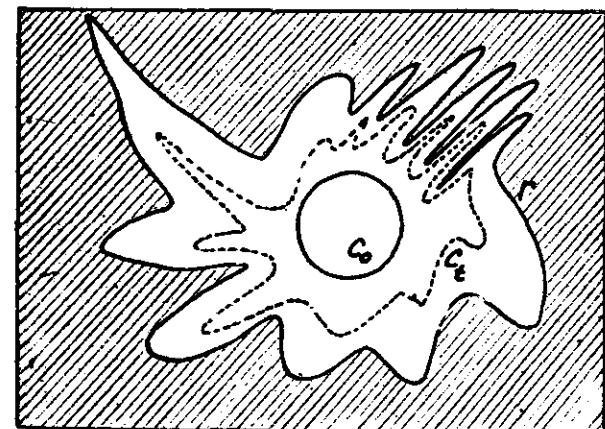


Рис. I. Схематическое представление процесса снятия вырождения в системе С при наложении ограничений Г.

S_0 - система С в начальный момент времени (наложение ограничений); S_t - система С в ходе снятия вырождения (текущий момент времени).

числа ферритов – для однозначного задания состояния всей памяти следует указать состояние каждого отдельного элемента (то есть задать совокупность N двоичных чисел, где N – число элементов памяти). Такие среди можно назвать вырожденными (в физическом смысле). Степень вырождения измерима. Так, для бинарных памяти ЭВМ степень вырождения φ дается соотношением $\varphi = 2^{-N}$, для памяти на многоустойчивых элементах $\varphi = k^{-N}$, где k – число устойчивых состояний элемента памяти.

Приведение (установка) памяти в заданное состояние достигается путем установки каждого элемента в определенное состояние, то есть путем определенной настройки (программирования) памяти. Если память состоит из N элементов (для памяти III поколения $N = 10^6 - 10^8$), то поэлементная установка памяти представляет собой усложненную операцию. Как известно, установка памяти выполняется человеком-программистом путем ввода в память некоторой заранее составленной программы. Настроенная указанным образом память содержит набор "слов" или сигналов, используемых для управления исполнительным информационно-обрабатываемым оборудованием (схемами сложения, ввода-вывода и т.д.).

Процесс настройки памяти можно трактовать как процесс снятия вырождения под влиянием наложения комплекса ограничений G на вырожденную систему. В ходе этого процесса система переводится из начального состояния C_0 (когда ограничений не было) в конечное состояние C_1 , полностью согласованное (адаптированное) с наложенными ограничениями G . Но всякий процесс перевода системы из одного состояния в другое возможен лишь под влиянием надлежащих управляющих воздействий. Иными словами, снятие вырождения в системе под действием ограничений G можно трактовать как управление системой со стороны комплекса ограничений G (G -управление системой или G -настройка памяти).

Другой аспект снятия вырождения G -ограничениями связан с уменьшением первоначальной неопределенности в результате приема информации G . Этот процесс устранения неопределенности (снятие вырождения) естественно трактовать как процесс принятия решения системой под влиянием полученной информации G . Получив информацию, система делает выбор. Этот подход полностью согласуется с концепцией Эшиби [2]: "Нам до конца часто приходится слышать, что машина не способна к выбору, но справедливо как раз обратное: каждая машина, стремящаяся к равновесию, совер-

шает соответствующий акт выбора". Разумеется, если наложены ограничения достаточно однородные, а вырождение, подлежащее снятие (устранению) велико, то выделенная в результате снятия вырождения структура все еще может обладать некоторой неопределенностью (вырождением), то есть принятое решение не будет строго однозначным (будет содержать элементы неопределенности).

Количественное соотношение между степенью снятого вырождения $\Delta \varphi$ и разнообразием (объемом информации), содержащимся в G -ограничениях, дается 10-й теоремой Шеннона: устранившая мера неопределенности (степень снятого вырождения) $\Delta \varphi$ не может превышать пропускной способности G канала передачи информации, под действием которой устраняется неопределенность, то есть $\Delta \varphi \leq G$. Более наглядно тот же самый результат формулируется законом необходимого разнообразия Эшиби: выбор из ряда возможностей, то есть снятие вырождения не производится даром – платя за выбор (за принятие решения) заключается в подаче надлежащего объема ограничивающей информации G . Этот закон подводит объективную основу под алгоритмы принятия решения, указывая, что выбор из многих альтернатив требует получения надлежащего количества информации.

Как видно, концепции вырождения и снятия вырождения G -ограничениями играют важную роль при анализе систем управления с точки зрения их устойчивости относительно изменений внешних условий на объекте управления. В существующих АСУ II-го поколения запас вырождения определяется объемом памяти, снятие вырождения осуществляется путем настройки памяти программой, составленной человеком-программистом и отражающей логическую структуру G -ограничений. Иными словами, в системах управления II-го поколения настройка (программирование) памяти осуществляется не автоматически – путем прямой передачи внешней обстановки G через надлежащие устройства ввода-вывода и каналы связи в память, а с участием человека-программиста, воспринимающего G -ограничения. Программист производит пробную настройку памяти, а затем на стадии интерпретации и исполнения управляющих сигналов из памяти оценивает (на основании критерия качества) соответствие отклика системы G -ограничениям и, следовательно, пригодность пробной настройки. В случае недостаточного соответствия настройка корректируется и проба повторяется до достижения

заданного соответствия между откликом системы (определенным настройкой памяти) и внешними Γ -ограничителями.

Важность 10-й теоремы Шеннона или эквивалентного ей закона необходимого разнообразия Энби для анализа процесса настройки (программирования) памяти состоит в указании на своеобразный закон сохранения: для выбора настройки, согласованной с Γ -ограничениями, должна быть переработана информация объемом Γ .

Таким образом, для каждого из комплексов ограничений $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ существует эквивалентная настройка памяти (состояние памяти) или программа $P_u (u = 1, 2, \dots, n)$. Для проектировщиков систем это означает, что если комплекс ограничений обладает разнообразием Γ (скажем, $\Gamma = 10^{10}$ бит), то для формирования соответствующей управляющей программы (для соответствующей настройки памяти) должна быть переработана человеком-программистом информация объемом Γ (или эта информация должна быть непосредственно передана в память для снятия надлежащего вырождения).

Критерий качества управления, уравнения адаптации и основная проблема проектирования систем

Мы видели, что процесс снятия вырождения в системе C при наложении Γ -ограничений можно трактовать как управление системой с помощью Γ -ограничений, при котором система переводится из некоторого начального состояния C_0 в определенное конечное состояние C_1 , согласованное с наложенными Γ -ограничениями. Обычно существует не единственный путь перевода $C_0 - C_1$. Для сравнения различных путей перевода (различных управлений) должна быть выработана мера M , характеризующая каждое управление (каждую траекторию управления). Такая мера определяется, разумеется, на классе возможных траекторий. С помощью меры M из всех траекторий (принадлежащих классу K) можно выделить те, которые обладают определенными свойствами. Так, соотношение

$$M = \max(M_\tau), \tau \in K, \quad (I)$$

выделяет из класса K траекторию $\tau = \tau_0$, на которой мера M до-

стигает максимума. Разумеется, такое соотношение полезно при фактическом отыскании оптимальной траектории, так как позволяет сравнивать любые две траектории и указать, какая из них ближе к оптимальной.

Хотя 10-я теорема или закон Энби и акцентируют внимание на важнейшем ограничении при проектировании систем, в них не содержится ответа на вопрос, какой должна быть траектория управления настройки памяти (снятие вырождения), вдоль которой построение программы, согласованной с Γ -ограничениями осуществляется, например, за кратчайшее время. Закон Энби устанавливает только характеристики конечного состояния настройки памяти в процессе управления настройкой.

Практически, однако, при проектировании больших АСУ, как правило, исходят из некоторой пробной программы управления, и основной вопрос проектирования АСУ состоит в определении траектории приведения пробной программы в состояние, соответствующее (в пределах заданной меры) наложенным Γ -ограничениям. Проектированию АСУ нужно иметь инструмент, с помощью которого можно сравнивать две траектории и определить, какая ближе к оптимальной траектории синтеза управляющей программы (настройки памяти):

Метод отыскания оптимальной траектории настройки памяти так или иначе должен быть основан на использовании надлежащего критерия качества управления. Если критерий установлен, общепринятый прием отыскания оптимальной траектории состоит в построении уравнений минимизации для данного критерия.

Построим критерий качества управления настройкой памяти под Γ -ограничениями. С каждой траекторией управления можно связать скорость (или время) приведения памяти в состояние, соответствующее Γ -ограничениям. Потребуем, чтобы эта скорость была максимальна (и, следовательно, время минимально). Тогда траектория управления, обеспечивающая быстрейшее выведение системы в Γ -состояния, обладает свойством:

$$\frac{dc}{dt} = \max_{\gamma \in K} \left(\frac{dc}{dt} \right)_\gamma, c(t) = \int d\vec{x} f(\vec{x}, t) j(\vec{x}, \gamma) = \langle j \vec{j} \rangle, \quad (2)$$

где $(dc/dt)_\gamma$ — скорость сближения текущего и оптимального состояний системы на траектории γ , $c(t)$ — мера близости (см. ниже) двух состояний (памяти) и K — класс возможных траекто-

рый управления.

Максимальная скорость настройки памяти возможна лишь при условии, если в канале коррекции, по которому передается Γ -информация, отсутствуют шумы и помехи. Однако практически канальная коррекция не идеальна — в передаче Γ -информации принимают участие люди, промежуточные автоматические системы и т.д. Поэтому траектории управления настройкой могут быть самыми различными.

Дадим наглядную интерпретацию критерия максимальной скорости настройки памяти в Γ -состояние. Пусть вырожденная система, подлежащая настройке, задана однородным (ламинарным) потоком электронов. Такой параллельный однородный поток протекает через регистр дешифратора блока управления, на который устанавливается очередная команда a , моделирующая этот поток. Мы рассмотрим соответствующую непрерывную модель. При наложении распределенных Γ -ограничений (модель состояния регистра с командой на нем) постепенно нарастает деформация (неустойчивость) электронных траекторий, протекающих около участка наложения ограничений $\Gamma(x)$. Эта неустойчивость может разрешиться лишь в результате излома деформированной траектории и возникновения петли стабилизации (петли обратной связи). Стабилизация обеспечивается путем доставки в участок неустойчивости возвратного потока, задержанного по фазе так, чтобы прямой и возвратный потоки смешивались в противофазе (деструктивная интерференция потоков). Условием смешивания в противофазе определяется длина петли обратной связи, то есть осуществляется квантование траекторий. Если появление одной петли обратной связи не компенсирует в полной мере неустойчивость, протекает дальнейший процесс деформации траектории с образованием многократно вложенных петель стабилизации, при этом сами петли распределяются вдоль линий приложения Γ -ограничений. Таким образом, в вырожденной среде возникает динамическая структура потока $\tilde{f}(\vec{x}, t)$ или динамический граф \tilde{f} , согласованный с Γ -ограничениями. Этот граф и является прообразом управляющей структуры, возникающей в вырожденной среде при снятии вырождения Γ -ограничениями. При всяком изменении Γ -ограниченный граф $\tilde{f}(\vec{x}, t)$ претерпевает деформацию и переходит в новый граф, соответствующий новым Γ -ограничениям.

Критерий (2) максимальной скорости настройки выделяет из

множества траекторий настройки оптимальную (по быстродействию) траекторию, и соотношение (2) в сущности представляет собой неклассическое уравнение в частных производных для отыскания оптимальной траектории (оптимальной программы управления, полностью согласованной с наложенными Γ -ограничениями). Мерой близости $C(t)$ может быть, в частности, интеграл перекрывания двух таких структур — текущей \tilde{f} и оптимальной f^* , то есть величина вида

$$\langle \tilde{f} \tilde{f}^* \rangle = \int d\vec{x} \tilde{f}(\vec{x}, t) \tilde{f}^*(\vec{x}, -). \quad (3)$$

Для распределенных систем управления уравнения движения динамического графа $\tilde{f}(\vec{x}, t)$ вытекают из уравнения Шредингера, описывающего состояние квантовой (когерентной) динамической системы, и имеют следующий вид:

$$\begin{aligned} \partial \tilde{f} / \partial t &= -(1/\rho) \tilde{f} \nabla \tilde{f} - (e\rho/m) \nabla \Gamma - (\tilde{f} \nabla) (\tilde{f} / \rho) - (\hbar^2/2m) \rho \nabla (\rho \nabla \tilde{f}^2), \\ \partial \rho / \partial t &= -\nabla \tilde{f} \cdot \tilde{f}^* \Big|_{t=0} = \tilde{f}(\vec{x}, 0) \cdot \rho(\vec{x}, t) \Big|_{t=0} = \rho(\vec{x}, 0). \end{aligned} \quad (4)$$

Здесь \hbar — постоянная Планка, m — масса и e — заряд токовой частицы, $\Gamma(\vec{x})$ — внешние распределенные ограничения (возмущения) $\tilde{f}(\vec{x}, 0)$ и $\rho(\vec{x}, 0)$ — начальные распределения токов и плотностей, соответственно.

Динамический граф $\tilde{f}(\vec{x}, t)$ задается векторной функцией точки, определенной во всем пространстве в каждый момент времени. Плотность потока в точке \vec{x} дается функцией $\rho(\vec{x}, t)$. Динамический граф фактически представляет собой диаграмму потока (flow-chart), то есть блок-схему управляющей программы, и уравнения (4) показывают, как, исходя из некоторой пробной программы управления (пробной блок-схемы), получить программу (блок-схему), полностью согласованную с наложенными Γ -ограничениями. Уравнения (4) показывают, как в вырожденной среде в результате наложения достаточно разнообразных Γ -ограничений возникают модные иерархические (многократно вложенные) графы — программы, адекватные наложенным Γ -ограничениям.

Таким образом, с помощью управляющих ограничений Γ в системе C среди множества вырожденных структур выделяется подмножество структур (в пределе — одна структура), адекватное наложению.

женным Γ -ограничениями. Важно, что этот процесс снятия вырождения (устранения неопределенности, принятия решения) имеет распределенный параллельный характер в отличие от аналогичных процессов, протекающих в АСУ I-го поколения, в которых снятие вырождения (отбор, сравнение структур) осуществляется последовательно в результате постепенного перебора возможных структур в последовательных процессорах II-го поколения.

Параллельная память обладает еще одним важным преимуществом по сравнению с АСУ I-го поколения. Позлементные памяти II-го поколения не допускают наложения Γ -отображений (управляющих программ) в памяти, поэтому скорректированная программа должна переписываться в свободную память, даже если корректировка незначительна. Это означает, что при внесении ряда изменений возникает множество не связанных между собой вариантов программы управления. Еще одно усложняющее обстоятельство связано с нелокализованным характером Γ -ограничений; требуемая перестройка значительного участка памяти элемент за элементом с сохранением надлежащих связей между ними оказывается весьма трудной.

В параллельных системах памяти перестройка (перепрограммирование) в результате изменения Γ -ограничений осуществляется распределенным воздействием $\Delta\Gamma$. При этом существенно, что скорректированная программа накладывается на все прежние программы (единая накопительная среда), так что в конечном итоге, если изменения $\Delta\Gamma$ носят "пульсирующий" характер (с сохранением некоторого общего ядра, что характерно для практических условий, обладающих определенной устойчивостью), то среди множества наложенных между собой программ вычленяется некоторая обобщенная программа, адекватная некоторым упредленным Γ -ограничениям. Это свойство параллельных систем накопления и обработки информации может играть решающую роль при формировании обобщенных (централизованных) концепций или программ.

Применения к синтезу АСУ. Язык как система управления

Выделение представления о вырожденных системах, о снятии вырождения при наложении Γ -ограничений, о скорости это-

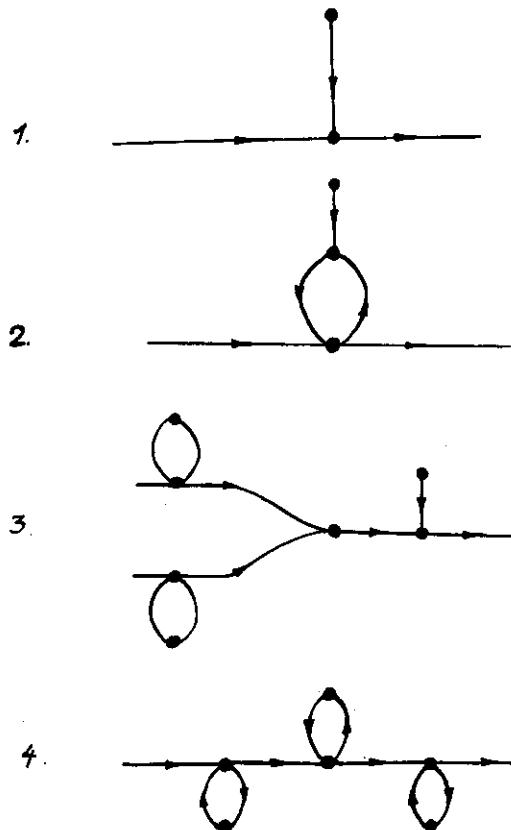


Рис. 2. Примеры динамических графов систем управления:
 1 - прямое управление исполнительной системой;
 2 - управление исполнительной системой, содержащей обратную связь;
 3 - управление параллельным выполнением двух циклических процессов с последующими синхронизациями результата и прямым управлением обработкой результата синхронизации;
 4 - управление последовательным выполнением трех циклических процессов с передачей результата предыдущего цикла на вход последующего.

го снятия и о динамических программах-графах (рис.2), возникающих в вырожденных системах в ответ на наложение Γ -ограничений, могут быть применены для синтеза АСУ различного функционального назначения.

Основной интерес среди таких АСУ представляют:

системы(программы) управления автоматическим оборудованием;

системы (программы) управления человеко-машинными системами и чисто человеческими коллективами.

Заметим, что, как известно, понятия управляющей системы и управляемого объекта являются относительными. Это означает, что в качестве Γ -ограничений могут выступать динамические программы-графы - в этом случае изложенный аппарат может быть использован для анализа динамических игр между программами.

Итак, основные выводы, важные при проектировании АСУ, можно сформулировать следующим образом:

1) АСУ должны проектироваться с запасом вырождения (памяти), объем которого определяется информативностью Γ -ограничений; при снятии вырождения Γ -ограничениями должен оставаться некоторый запас вырождения - ресурс устойчивости относительно дополнительных воздействий;

2) АСУ должны проектироваться с "неплотной" памятью параллельного доступа - лишь в этом случае достигается устойчивость программы относительно изменений Γ -ограничений; неплотность означает возможность взаимного наложения программ в памяти;

3) АСУ должны проектироваться с автоматизированными средствами настройки (программирования) памяти, основанными на наличии надлежащих устройств ввода/отображений в память Γ -ограничений;

4) система интерпретации и исполнения управляющих сигналов проектируемой АСУ должна обладать средствами настройки (запасом вырождения), управляемыми критериями качества функционирования всей АСУ.

Требование "неплотности" и параллельности памяти тесно связано с требованием автоматической настройки памяти и требует отдельного обсуждения. Если необходимость некоторого (вообще, неизвестного) запаса памяти иногда принимается во внимание при проектировании АСУ II-го поколения, то автоматизированными механизмами настройки памяти Γ -ограничениями полностью преоб-

регают. Однако, поскольку проблема настройки памяти Γ -ограничениями объективно существует, автоматические системы настройки памяти реализуются на программном уровне. Это не что иное, как системы математического обеспечения - автоматизированные системы синтеза программ управления, адекватных наложенным Γ -ограничениям. Эти системы воспринимают Γ -ограничения в терминах своего входного языка и формируют на выходе программы, настраивающие память и соответствующие Γ -ограничениям.

Таким образом, необходимость передачи (для настройки памяти) внешней информации объемом Γ приводит к задаче создания канала связи между Γ -ситуацией и памятью АСУ. Роль такого канала играет пока математическое обеспечение АСУ, то есть автоматизированная система синтеза программ управления по Γ -ограничениям. Поскольку в больших АСУ объем Γ -информации весьма значительный, а автоматизированные средства передачи этой информации в память (на уровне оборудования) отсутствуют или развиты очень слабо, то даже при наличии надлежащего запаса вырождения (памяти) в АСУ синтез адекватных Γ -ограничениям программ управления может быть обеспечен только при наличии мощных систем математического обеспечения (МО) - систем автоматического синтеза программ. Именно поэтому с ростом логической мощности АСУ, то есть с ростом информативности Γ -ограничений, учитываемых в проектируемых АСУ, нагрузка на канал передачи Γ -информации так сильно возрастает. Поскольку не принимается никаких мер по созданию прямого канала передачи Γ -информации на уровне оборудования, почти вся нагрузка ложится на системы МО.

При наличии автоматических каналов передачи Γ -информации в память АСУ, то есть при наличии эффективных систем ввода-отображения Γ -ситуации, роль программного МО оказалась бы несравненно меньшей, а экономичность и надежность АСУ несравненно выше существующих. Иными словами, современное гипертрофированное МО - результат неспособности в настоящее время проектировать эффективные системы памяти и каналы прямого параллельного взаимодействия между Γ -ситуацией и памятью АСУ. Таким образом, разработка эффективных систем взаимодействия ситуации I - машина - ситуация 2, частным случаем которой является проблема взаимодействия человек-машина, является одной из основ-

ных задач проектирования больших АСУ.

Этот подход предполагает рассмотрение МО как канала передачи Γ -информации в память АСУ; на входе канала — Γ -информация, на выходе — программа, настраивающая память в соответствии со структурой Γ -информации. Таким образом, МО заменяет механизм прямой передачи Γ -информации в память и может быть рассмотрено как часть вырожденной системы, непосредственно контактирующая с Γ -информацией. Это означает, что для проектирования МО, то есть для проектирования входного языка системы автоматического синтеза программ настройки (программ управления), может быть базой изменений применен подход, изложенный выше и основанный на концепции формирования динамического графа-языка под влиянием наложенных на вырожденную систему Γ -ограничений.

Итак, под каждую Γ -ситуацию может быть сформирован адекватный граф-язык (вход в систему автоматического синтеза программ), испытывающий надлежащие деформации при изменении Γ -ситуации. Если (подобно предыдущему) в реализации Γ -ситуации вычленяется обобщенное ядро, аналогичное ядро вычленяется и в языке-графе, адекватном Γ -ограничениям; в общем случае при резком изменении Γ -ситуации граф языка претерпевает надлежащие деформации, включающие как расширение (появление новых петель стабилизации), так и его сужение (исчезновение ненужных петель).

Если применить эти качественные результаты к традиционным системам МО, можно получить следующие выводы:

1) существующие системы автоматического синтеза программ (системы МО) должны быть дополнены механизмами расширения и сужения средств языка в зависимости от изменений Γ -ограничений, налагаемых на проектируемую АСУ;

2) существующие системы МО должны быть снабжены эффективными средствами контакта с Γ -ситуацией (визуальный и звуковой вход в системы программирования), обеспечивающими более прямую передачу Γ -информации в память АСУ (с разгрузкой человека-программиста как промежуточной системы между Γ -ситуацией и памятью АСУ);

3) исполнительные средства системы МО должны иметь механизмы настройки, управляемые критерием качества функционирова-

ния АСУ (то есть операторы и операнды языка МО должны обладать средствами адаптации, настраиваемыми критерием качества работы всей АСУ).

В дальнейшем помимо учета этих качественных соображений основное внимание при разработке систем МО должно быть уделено методам проектирования, вытекающим из установленного критерия качества и соответствующих уравнений движения.

Л и т е р а т у р а

1. БИР С. На пути к кибернетическому предприятию, —"Принципы самоорганизации", М., "Мир", 1966, стр. 48-130.

2. ЭШБИ У.Р. Принципы самоорганизации. —"Принципы самоорганизации", М., "Мир", 1966, стр. 314-343.

3. ЛЯПУНОВ А.А., ЯБЛОНСКИЙ С.В. Теоретические проблемы кибернетики. —"Проблемы кибернетики", 1963, вып. 9, стр. 5-22