

ВЫСОКОУРОВНЕВОЕ ПРОГРАММИРОВАНИЕ И СРЕДА МАТЛАВ. ПОДВОДНЫЕ КАМНИ ИНТЕГРАЦИИ

Чудновский М.М.

Научный руководитель к. т. н., доцент Русанова О.А.

Сибирский федеральный университет

Развитие информационных технологий и производительности современных ЭВМ идет в очень высоком темпе. Усложняются технологии, увеличивается вычислительная мощность компьютеров как домашнего, так и промышленного применения. Исходя из этих предпосылок, происходит расширение круга возможных задач, решение которых возлагается на ЭВМ. Именно этим обеспечивается появление специализированных информационных систем, или отдельных приложений, для обеспечения сбора, обработки и предоставления необходимых данных в контексте решаемой проблемы или автоматизируемого бизнес-процесса. Жизненный цикл таких систем может быть различным и зависит от сферы применения, а также масштаба реализации. Однако в динамике построения подобных систем можно выделить некоторые общие черты – от поколения к поколению усложняются алгоритмы работы с информационными потоками в системе, а также возрастает объем данных, которые анализируются и обрабатываются ИС.

Современные информационные системы в общем случае имеют модульное построение, причем реализация отдельных составных частей не всегда осуществляется одним разработчиком. К примеру, реализация серверов баз данных и СУБД, посредством которых строится хранилище данных во многих ИС, как правило, осуществляется на основе уже разработанных, и активно развиваемых технологий компаний Microsoft, IBM, ORACLE и других. Модульная структура обеспечивает много преимуществ, таких как: гибкость разработки, масштабируемость, эффективное использование ресурсов, конвейерная обработка данных. Исходя из этого, процесс обработки данных по сложным и ресурсоемким алгоритмам целесообразно выносить в отдельный структурированный блок. Данное решение позволяет оптимизировать работу информационной системы в целом. Также, реализация подобного блока вычислений может осуществляться сторонним разработчиком, в случае параллельной разработки информационной системы. Возможна ситуация, когда часть подсистем и вычислительных блоков уже реализована. В данном случае возникает проблема компиляции данных структурных единиц в единое целое, либо эффективное внедрение в уже существующую информационную систему.

Как уже отмечалось, современные информационные системы решают широкий спектр задач. И, в зависимости от предметной области, зачастую требуется реализация методов обработки данных на основе сложных математических алгоритмов, либо необходимо моделирование различных систем или объектов. Ярким примером могут служить системы компьютерного зрения, системы искусственного интеллекта, системы синтеза и распознавания речи.

В зависимости от интегрированной среды разработки (IDE) информационной системы, платформы ее развертывания и некоторых прочих условий (например, требований заказчика или ограниченность ресурсов на этапе разработки) вопрос реализации функциональных блоков сложной обработки данных может решаться несколькими способами:

- полная реализация средствами IDE;
- использование библиотек исходного кода, обеспечивающих требуемый функционал;
- использование технологий взаимодействия объектов (COM);
- динамически подключаемые модули.

Каждый из описанных методов имеет свои преимущества и недостатки, поэтому применение того или иного решения всецело зависит от контекста решаемой задачи.

Наиболее известной средой математического программирования является продукт MATLAB (актуальная версия r2012b) компании MathWorks. Данное решение используется для модельно-ориентированного проектирования (имитационного моделирования), разработки систем управления (САУ, АСУ ТП), систем связи (в том числе ВЧ) и цифровой обработки сигналов (ЦОС, DSP), генерации кода для ПЛИС и микроконтроллеров, анализа и сбора данных, а также для иных вычислительных задач. Исходя из широкой сферы применения, и, учитывая свободно-распространяемую среду выполнения компонентов MATLAB MCR, задача интеграции функциональных блоков, реализованных в среде MATLAB в конечное приложение, или информационную систему становится актуальной. Решение подобной задачи может строиться на разных принципах взаимодействия, в зависимости от платформы разработки конечного приложения.

Современные IDE, а также среды выполнения приложений представлены очень широко, и поэтому рассмотреть взаимодействие MATLAB с каждой из них не представляется возможным. Одной из самых популярных платформ для разработки приложений является решений компании Microsoft - .net framework. Данная унифицированная среда разработки используется программистами по всему миру, и существует множество решений различного масштаба, построенных на данном «фреймворке». Использование данной платформы предоставляет широкие возможности для взаимодействия разрабатываемого приложения с операционной системой, а также набором периферийных устройств на низком уровне. Резюмируя, можно с уверенностью говорить о том, что .net обеспечивает большую интеграцию приложения с ОС по сравнению с внутренними средствами MATLAB (например, MATLAB Gui tools), что еще раз подводит к выводу о целесообразности использования компонентов MATLAB при построении ИС и приложений на базе MS .NET Framework.

Для использования компонентов MATLAB в проектируемом приложении¹ наиболее оптимально применять модели динамически подключаемых компонентов, DLL. Компания MathWorks в своем решении предусматривает такую возможность, процесс создания динамических библиотек для различных платформ (в том числе и .NET) реализуется посредством специального компонента MATLAB – Deploy toolbox. Для построения библиотек DLL необходимо предварительно настроить компоновщик MATLAB, который на основе среды MCR и установленного в системе компилятора языка c++², позволяет создавать отдельные приложения MATLAB или компоненты DLL.

Реализация динамических библиотек MATLAB осуществляется на основе функций, написанных на одноименном языке и оформленных в отдельные m-файлы. Результатом работы deploy toolbox в рассматриваемом случае является файл *.dll, который реализует необходимый класс для .NET. Методами данного класса и обеспечивается функционирование функций MATLAB (функции могут реализовывать практически весь потенциал MATLAB в зависимости от контекста задачи). Для использования библиотеки dll, реализованной посредством MATLAB deploy toolbox необходимо сделать ссылку на данный модуль из проекта .NET³, а также подключить модуль MWArray.dll, который отвечает за обмен информацией между компонентом MATLAB и средой .NET.

При разработке ИС или конечных приложений на языке C# для .NET необходимо также использовать соответствующие пространства имен, для использования подключенных библиотек к проекту. Для корректной интеграции необходимы:

¹ Здесь и далее подразумевается, что приложение проектируется на базе актуальной версии MS .NET Framework

² Важно ответить, что могут применяться различные компиляторы – MinGW, Microsoft Visual Studio и другие.

³ В данной статье рассматривается язык программирования C# для платформы .NET, однако это не имеет принципиального значения, так как принцип интеграции един для всех языков среды .NET.

- пространство имен, одноименное с DLL модулем, который реализует требуемые m-функции;
- пространство имен MathWorks.MATLAB.NET.Utility для корректной интеграции;
- пространство имен MathWorks.MATLAB.NET.Arrays для формирования инфопотоков между модулями системы.

Соблюдение описанных требований обеспечивает полноценное взаимодействие MATLAB и .NET, однако как в любом процессе интеграции возможны нештатные ситуации, которые могут привести к некорректной работе системы в целом. По информации от компании MathWorks, существует несколько компонентов MATLAB, которые не могут быть реализованы в качестве самостоятельных приложений, а значит и в виде динамических модулей для .NET. Данное ограничение обозначивает себя необычным образом, а именно:

- на этапе разработки функции MATLAB проблем не возникает, т.к. ограничения не вступают в силу;
- на этапе компоновки разработанной функции также не возникает проблем, так как с точки зрения синтаксиса MATLAB – программа корректна;
- на этапе подключения модуля в проект .NET проблем также не возникает, так как на этапе компоновки DLL ошибок не возникло;
- на этапе выполнения модуля (не путать с выполнением программы) возникает критическая ошибка, так как среда MCR не поддерживает ряд компонентов MATLAB.

Примером подобной ситуации может информационная система голосовой верификации, реализованная на языке C#, базирующаяся на нейросетевом подходе. При подобной реализации перцептрон, как основное вычислительное ядро системы выносится в отдельный динамический модуль и реализуется средствами языка MATLAB. Согласно документации MathWorks, все функции реализации нейросетей (командный аналог Neural Networks Toolbox) поддерживаются средой MCR, и значит подлежат компоновке в самостоятельное приложение или динамический модуль.

Моделирование созданной и обученной нейросети MATLAB происходит посредством перегруженной функции `sim(object)` одноименного языка, где *object* - объект моделирования. Данная функция является общей для большинства объектов MATLAB и реализует модели их функционирования в зависимости от типа и структуры рассматриваемых объектов. Использование данного метода не поддерживается средой MCR, и как следствие ограничивает использование многих моделей MATLAB во внешних системах. Следствием этого является то, что процесс создания и обучения нейросети (скомпонованной в отдельный динамический модуль) выполняется корректно, но процесс моделирования полученных объектов вызывает критическую ошибку всей системы. Подобная проблема ставит под сомнение возможность создания системы верификации на основе интеграции Matlab и C# и нейросетевого подхода.

К сожалению, на данный момент официального решения описанной ошибки не существует, более того, реализация его компанией MathWorks не планируется. Однако, если рассмотреть структуру нейросети, можно сделать утверждать, что данный объект можно декомпозировать по слоям⁴.

При таком подходе, задача моделирования нейронной сети разделяется на задачи моделирования ее отдельных компонентов (слоев) и их взаимодействия. В данном случае становится возможным использование метода MATLAB `tansig(X)`⁵, который обеспечивает функционирование одного сегмента (слоя) нейронной сети и отлично поддерживается средой MCR. В данном случае необходимы параметры нейросети (передаточные функции нейронов, весовые коэффициенты), построение которых обеспечивается встроенными

⁴ Имеются ввиду S-элементы, A-элементы и R-элементы.

⁵ В данном случае X – выражение, описывающее соответствующий слой нейронной сети.

функциями MATLAB для обучения нейросетей, и которые также отлично поддерживаются средой MCR. Данные факты позволяют сделать вывод, что функция `sim` (object) трансформируется в выражение `tansig(S)+tansig(A)+tansig(R)`, где S,A,R сегменты нейросети для рецепторов, ассоциативных и реагирующих элементов. Таким образом, получается m-функция, которая полностью поддерживается средой MCR. Данные для нее формируются m-функциями, которые также выполняются без ошибок в MCR. Исходя из этого, получается готовый динамический модуль Matlab, реализующий нейронную сеть и без ошибок функционирующий в среде MCR, что является решением обозначенной ошибки интеграции.

Выводы. Средства пакета MATLAB являются эффективным инструментом для моделирования нейронных сетей различной структуры. Это позволяет использовать данные модели в научной или производственной практике. Нейронная сеть, реализованная пакетом компании MathWorks, может формироваться в виде отдельного функционального блока, для интеграции в более сложные информационные системы. При такой реализации, неизбежно возникает ошибка, ввиду того что, на данный момент в среде MCR не возможно моделирование объектов при помощи встроенной функции Matlab. Необходимо отметить, что все функции для создания и обучения нейросети функционируют должным образом и не вызывают ошибок среды. Исходя из этого, можно решить данную проблему интеграции путем декомпозиции исходной нейронной сети на составляющие. Ввиду того, что процессы создания и обучение нейронной сети не вызывают ошибок, можно моделировать поведение отдельных сегментов нейросети, используя данные, полученные в результате выполнения этих процессов. Данное решение позволяет обойти ограничения среды MATLAB и интегрировать воссозданную нейронную сеть в любую целевую систему. Следует заметить, что принципиально такой подход может быть применим не только к средству MATLAB NNtoolbox, но и к другим объектам данного программного продукта, которые моделируются стандартными способами.

В заключении можно сказать, даже при наличии формализованных интерфейсов интеграции возможно появление тех ситуаций, когда некоторые ограничения конечных систем приводят к некорректной работе интерфейса. В данном случае важны принципы функционирования отдельных структурных блоков, которые реализуются на базе данных систем. В конечном итоге очень часто существует возможность преобразовать некоторые элементы систем так, что становится возможным обойти ограничения интеграции.

ЛИТЕРАТУРА:

1. C# и платформа .NET - Э. Троелсен. СПб: Питер Пресс, 2010. - 769 с.
2. MATLAB обработка сигналов и изображений. - Специальный справочник Дьяконов В., Абраменкова И. СПб.: Питер, 2002. - 608 с.