

УДК 681.51

Вариант динамического ДСМ-метода для обучения мобильного робота

Т.А. Волкова (*wolf39@mail.ru*)
РГГУ, Москва

В работе описывается создание адаптивного робота, обучение которого реализовано с помощью динамического ДСМ-метода. Приведена модификация ДСМ-метода, в которой примеры после порождения из них гипотез могут быть удалены из обучающей выборки. Описаны эксперименты с роботом, который был построен в Лаборатории робототехники и искусственного интеллекта Политехнического музея.

Введение

В настоящий момент наблюдается активное внедрение методов искусственного интеллекта в робототехнику в таких областях, как распознавание изображений, планирование, коллективное поведение.

В работе [Добрынин, Карпов, 2006] приведена классификация роботов:

1. Программные (действующие по заранее заданной программе)
2. Адаптивные роботы (способные изменять поведение под воздействием окружающей среды)
3. Интеллектуальные роботы (способные выполнять задачу, поставленную в общем виде)

Из самых известных адаптивных роботов современности стоит упомянуть известных шагающих роботов Big Dog и Little Dog компании Boston Dynamics. Обзор современного состояния интеллектуальных роботов изложен в статье [Добрынин, 2006b]

Особую роль в создании адаптивных и интеллектуальных роботов играют методы машинного обучения. Одним из наиболее развитых методов машинного обучения является ДСМ-метод.

1. Постановка задачи обучения

Робот должен обучаться, чтобы совершенствовать своё поведение с накоплением опыта. Робот существует в среде, все объекты которой имеют описание (пусть X - множество всех описаний объектов) и принадлежат к некоторым классам из множества Y . Робот может сам получать факты из окружающей среды, поэтому обучение должно быть

индуктивным, то есть основанным на выявлении закономерностей в эмпирических данных. Задачей будет являться построение *индуктивного классификатора* $h: X \rightarrow Y$ по модели «стимул-реакция» (выработка условного рефлекса).

В качестве *обучающей выборки* $X^m = (x_1, y_1), \dots, (x_m, y_m)$ возьмем множество пар «обстановка – действие». Под *обстановкой* понимается состояние окружающей среды, зафиксированное рецепторами робота – освещенность, цвет поверхности, наличие препятствия, угол по отношению к некоторому ориентиру (маяку). Под *действием* будем понимать влияние эффикторов робота на окружающую среду – вращение колес, свет, звук. Возможны более сложные действия, состоящие из простых: например, действие «Атака» является комбинацией езды вперед и света фонариком, а «Защита» - комбинацией езды назад и звука.

Обучение будет осуществляться на *примерах* (положительных и отрицательных), которые выдает учитель. Формирование положительного примера назовём *поощрением*, отрицательного примера – *наказанием*. Обобщение нескольких примеров (только положительных или только отрицательных) назовём *гипотезой*.

Фазы *обучения* (обобщения, получения гипотез) и *действия* (применения гипотез) повторяются циклически. Кроме того, должен существовать некоторый критерий, показывающий, что обучение достаточно и теперь можно перейти к фазе работы.

Обучение является *интерактивным*, то есть сигналы поощрения и наказания выдаются непосредственно в ходе работы системы.

2. ДСМ-метод

В данной работе в качестве метода обучения используется ДСМ-метод автоматического порождения гипотез. Данный метод обобщает в гипотезах информацию, полученную из обучающей выборки (шаг индукции), затем применяет эти гипотезы для классификации неизвестных объектов (шаг аналогии), а также имеет критерий достаточного основания правдоподобного вывода (шаг абдукции).

Классический ДСМ-метод использует внешние истинностные значения t и f , а также множество внутренних истинностных значений $V = \{+, -, 0, t\}$.

Обучающая выборка состоит из множества объектов, множество имён которых образует множество O . Имена исследуемых свойств образуют множество P . Для описания того, какими свойствами какой объект обладает, мы можем представить обучающую выборку с помощью матрицы $F: O \times P \rightarrow V$. Если $F(o,p) = +$, то говорят, что объект o обладает свойством p . Аналогично для $F(o,p) = -$. Если $F(o,p) = 0$, то это значит, что

сведения об обладании o свойством p противоречивы. Случай $F(o,p)=\perp$ соответствует объекту, о котором неизвестно, обладает ли он свойством p .

Совокупность значений признаков образует *фрагмент* s объекта o . На множестве всех фрагментов должна быть задана бинарная операция сходства, обладающая свойствами коммутативности, ассоциативности и идемпотентности. Вычисляются пересечения фрагментов всех подмножеств объектов. Все такие пересечения образуют множество потенциальных причин C .

На шаге *индукции* из потенциальных причин для разных свойств порождается матрица $H: C \times P \rightarrow V$. Если $H(c,p)=+$, то говорят, что фрагмент s является причиной наличия свойства p . Если $H(c,p)=-$, то говорят, что фрагмент s является причиной отсутствия свойства p . $H(c,p)=0$ означает, что сведения о s как о причине свойства p противоречивы. Случай $H(c,p)=\perp$ соответствует отсутствию информации.

На шаге *аналогии* с использованием причин из матрицы H формируется новая матрица $F': O \times P \rightarrow V$, в которой t -значения исходной матрицы F могут быть заменены на $\{+, -, 0\}$.

Также с помощью матрицы H могут быть объяснены все значения в матрице F . Эта процедура называется *абдуктивным объяснением* и служит критерием достаточного основания для правдоподобного вывода. С другой стороны, процедура абдукции может быть использована для пополнения обучающей выборки: "уникальные" примеры, которые не объясняются абдукцией, могут быть дополнены новыми примерами.

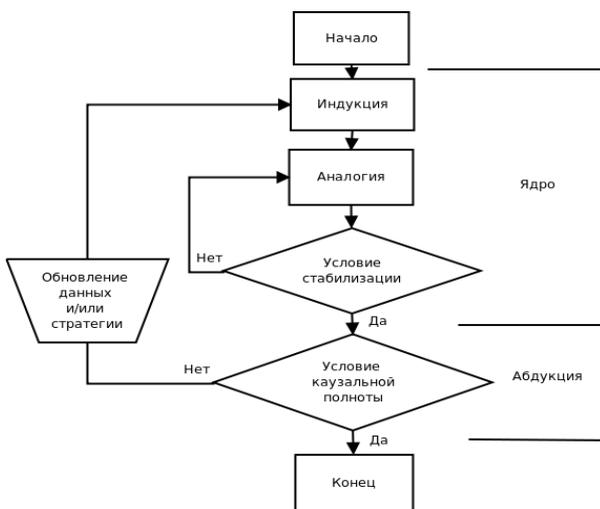


Рис. 1. Блок-схема классического ДСМ-метода.

ДСМ-метод предназначен для поиска причинно-следственных зависимостей в базе фактов предметной области и является качественным, а не количественным методом анализа данных. В настоящее время ДСМ-метод успешно применяется в фармакологии, медицине, социологии, почерковедении. Подробные сведения о ДСМ-методе, его теории и практическом применении изложены в сборниках статей [Аншаков, Фабрикантова, 2009] [Панкратова, Финн, 2009].

В отличие от других методов, ДСМ обеспечивает быстрое обучение. Так, в большинстве широко применяемых методов (нейронные сети, автоматы, генетические алгоритмы) для обучения какому-либо рефлексу роботу требуется как минимум 10-12 тактов обучения (для сравнения, 1-2 такта требуется в ДСМ-методе). ДСМ выдаёт полученные правила в явном виде, что облегчает интерпретацию результатов обучения. Из недостатков ДСМ-метода отмечается неприменимость в условиях противоречий в обучающих примерах. Идеи о том, как данная проблема могла бы быть решена, содержатся и в настоящей работе.

3. Динамический ДСМ-метод

Существующие ДСМ-системы работают с заранее заданным набором примеров. Примеры могут добавляться экспертом, если оказывается, что их недостаточно, но в этом случае каждый раз происходит перезапуск системы. Для работа такой вариант неприемлем из-за большой вычислительной сложности.

В работе [Добрынин, 2006а] была предложена модель *динамического ДСМ-метода*, в котором примеры могут добавляться в процессе функционирования системы. «В отличие от классического ДСМ метода, который работает с замкнутым множеством исходных примеров и заранее определенными их свойствами, динамический ДСМ метод позволяет работать в открытой среде с неизвестным заранее количеством примеров».

Такая особенность динамического ДСМ-метода накладывает определённые требования на алгоритм поиска пересечений. Очевидно, что он должен быть *пошаговым*, а не *пакетным* - множество пересечений не должно полностью пересчитываться каждый раз. Исходя из этого, в данной работе был выбран алгоритм Норриса с некоторыми модификациями [Kuznetsov, Obiedkov, 2002]. Ключевой проблемой метода Норриса является определение каноничности порождения пересечения: пересечение объектов называется *каноничным*, если оно порождается из родителей в порядке возрастания их номеров. Это необходимо для того, чтобы не порождать одно и то же пересечение многократно.

Самым же существенным отличием динамического ДСМ-метода от классического является отсутствие *критерия завершения*. В классическом ДСМ-методе им является достижение насыщения - невозможность

породить новые гипотезы. В динамическом варианте база фактов постоянно расширяется, и требуется сформулировать другое условие. Им может стать абдукция – познавательная процедура ДСМ-метода.

4. Модельная задача

В мире робота существуют объекты, которые он учится классифицировать. Например, объектами могут быть совокупности параметров окружающей среды в некоторые моменты времени.

Для обучения робота была выбрана несложная модельная задача, которую можно было решить с имеющимся набором датчиков и исполнительных механизмов. Учитель, в роли которого выступает человек, обучает робота правилам вида «если - то» - например, «если темно и не видно маяка – входи в режим защиты».

Признаки:

1. «Робот на черном/на белом»
2. «Темно/светло»
3. «Препятствие впереди/нет препятствия»
4. «Видит маяк под углом а»

Свойства (сложные действия)

1. Атака (Ехать вперед и светить фонариком)
2. Защита (Ехать назад и пищать)
3. Поиск (Повернуться влево на 90 градусов)

Для применения ДСМ-метода необходимо задать операции сходства на фрагментах и на множествах свойств. Обычно фрагменты и свойства представляются битовыми строками. Поэтому операция сходства соответствует побитовому умножению (рис. 3).

$$\begin{array}{r}
 \begin{array}{|c|c|c|c|} \hline 1 & 0 & 1 & 0 \\ \hline \end{array} & \begin{array}{|c|c|} \hline + & t \\ \hline \end{array} \\
 \text{⤵} & \\
 \begin{array}{|c|c|c|c|} \hline 0 & 1 & 1 & 0 \\ \hline \end{array} & \begin{array}{|c|c|} \hline + & + \\ \hline \end{array} \\
 \hline
 \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 0 \\ \hline \end{array} & \begin{array}{|c|c|} \hline + & t \\ \hline \end{array}
 \end{array}$$

Рис.2. Операция пересечения

5. Программная реализация обучения

В статье [Добрынин, Карпов, 2006] вводится реализация схемы поощрения и наказания, при которой после окончания действия робот ждет оценки учителя некоторое время. В настоящей работе использован

тот же принцип: процесс жизнедеятельности робота разделён на *фазы* действия и ожидания.

В фазе действия робот фиксирует обстановку и выбирает правило, которое подходит для данного случая; если такое правило не находится, то робот совершает случайное действие. Правильный поступок робота учитель поощряет. Если же робот ошибается, то учитель наказывает его (робот получает отрицательный пример). Возможна реализация, при которой учитель предъявляет положительный пример (говорит, как нужно поступать в этой ситуации). Это нужно для того, чтобы робот с самого начала обучения мог набрать достаточное количество положительных примеров.



Рис.3. Схема обучения робота

6. Постановка задачи «забывания»

Максимальное число примеров экспоненциально зависит от количества признаков, поэтому возникает проблема хранения примеров в памяти робота. "Забывание примеров" существенно и по другой причине. Хотелось бы, чтобы с течением времени старая обучающая выборка не участвовала в порождении новых гипотез.

Замечено, что большинство примеров совпадает с пересечениями, у которых 1 родитель. Совместно с Д.В. Виноградовым (ВИНИТИ РАН) был разработан вариант ДСМ-метода, где примеры сохраняются как пересечения с одним родителем. После этого примеры могут быть забыты.

Для проверки каноничности итерация по примерам заменяется итерацией по пересечениям (недостаток - потенциально высокая вычислительная сложность). Однако, имеется выигрыш по объему занимаемой памяти: в классическом ДСМ в худшем случае мы имеем 2^n пересечений и 2^n примеров, тогда как в настоящей реализации - только 2^n пересечений, среди которых есть и примеры.

Таким образом, данный вариант можно использовать, когда признаков мало, а примеров много. У робота примеры непрерывно поступают, поэтому использование такой модификации оправданно.

Некоторые пересечения «скрывают» примеры: их фрагмент совпадает с фрагментом примера. Такие пересечения должны рассматриваться наравне с пересечениями с одним родителем. Обычно число пересечений со «скрытыми родителями» не превосходит трети от всех.

Была разработана модификация метода Норриса. В ней на стадии проверки на относительную и абсолютную каноничность итерация по предыдущим примерам заменяется итерацией по пересечениям со «скрытыми родителями». При добавлении нового пересечения или модификации старого проверяется, совпадают ли фрагмент текущего примера и фрагмент текущего пересечения. В этом случае у пересечения ставится флаг "скрытые родители". Такой же флаг ставится и при добавлении одноэлементного пересечения.

7. Учёт числа примеров

Робот живет в изменяющейся среде (по терминологии Д.А. Поспелова, мире с "переключающимися средами"). Он должен быть способен адаптироваться к изменениям - переобучаться при поступлении сигнала поощрения или наказания, причем максимально быстро.

Но есть и в некотором смысле противоположное утверждение: робот должен быть устойчивым к шумам и выбросам в данных. Это накладывает ограничения на используемые методы обучения: они должны быть устойчивы к одиночным (редким) ошибкам, таким как ложное срабатывание одного из датчиков или недосмотр учителя.

В классическом ДСМ-методе существует стратегия "запрет на контрпример", в которой гипотеза уничтожается, если в обучающей выборке находится более чем n контрпримеров (порог n задается экспертом эмпирически). Контрпримером называется пример, который объясняется гипотезой, но имеет отличный от нее знак. "Запрет на контрпример" позволяет избавляться от противоречивых гипотез; в нашем случае такой запрет необходим, чтобы робот мог переучиваться – исключать старые и неактуальные гипотезы и приобретать новые.

Такой вопрос не стоял в задаче с обучением робота езде по линии [Добрынин, Карпов, 2006], так как в той модели не было отрицательных примеров, а следовательно гипотезы не могли уничтожаться.

В существующих ДСМ-системах новый пример проверяется на уникальность. Повторные примеры не рассматриваются, так как они не интересуют эксперта. В случае с роботом повторные примеры стоит учитывать: чем больше встретилось примеров, удовлетворяющих

некоторой гипотезе, тем больше следует доверять гипотезе. Примеры приходят роботу постоянно, поэтому часто встречаются дубликаты.

Если использовать "запрет на контрпример", остается открытым вопрос о том, как назначать пороги для отсеивания контрпримеров: назначение их постоянными не представляется разумным.

8. Эксперименты

Эксперименты проводились на мобильном роботе "Навигатор", который представляет собой автономную тележку с микроконтроллером, рецепторами и эффекторами. Интеллектуальная система находится на компьютере и передает управляющие сигналы роботу по Bluetooth. Архитектура всего программного комплекса изображена на рис. 4.

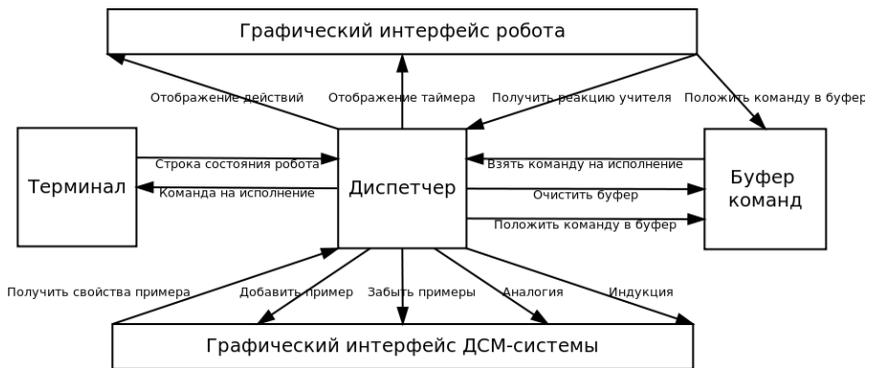


Рис.4. Архитектура системы

В результате эксперимента робот успешно обучился нескольким стратегиям поведения. Кроме того, были проведены эксперименты по адаптации робота в искусственно измененной ситуации.

Вот пример обучения. Мы хотим научить робота, как вести себя при встрече с препятствием:

1) Сперва научим его поведению в случае отсутствия препятствия. Выдадим 2 положительных примера: "если темно и нет препятствия – атака" и "если светло и нет препятствия – атака". Робот породит гипотезу "если нет препятствия – атака".

2) Аналогично робот породит гипотезы "если темно и есть препятствие – защита" и "если светло и есть препятствие – поиск".

3) Теперь допустим, что среда изменилась и робот должен при отсутствии препятствия всегда входить в режим поиска. Выдадим роботу хотя бы один контрпример: "если темно и нет препятствия – не атака", и полученная ранее гипотеза "если нет препятствия – атака" будет

уничтожена. Теперь, как и в пункте 1, робот может быть научен новому правилу поведения в данной ситуации.

Пример хорошо иллюстрирует описанную выше проблему: гипотеза, подразумеваемая учителем и порожденная из двух примеров, уничтожается одним-единственным контрпримером.

Заключение

Проведенные эксперименты показывают, что простой перенос существующих ДСМ-систем был бы неэффективен из-за особенностей робота как системы реального времени.

Результаты работы могут быть использованы для наглядной демонстрации ДСМ-обучения в экспозиции Политехнического музея.

В ближайшей перспективе предполагается разработать схему оценивания гипотез и реализовать специфическую для робота абдукцию.

В дальнейшем мы планируем применить и другие методы машинного обучения для адаптации робота.

Список литературы

- [Аншаков, Фабрикантова, 2009] ДСМ-метод автоматического порождения гипотез: Логические и эпистемологические основания. Сост. О.М. Аншаков, Е.Ф. Фабрикантова; Под общ. ред. О.М. Аншакова. – М.: Книжный дом «Либроком», 2009.
- [Добрынин, 2006а] Динамический ДСМ-метод в задаче управления интеллектуальным роботом. Десятая национальная конференция по искусственному интеллекту с международным участием КИИ-2006 (25-28 сентября 2006 г., Обнинск): Труды конференции. В 3-т. Т.2. - М: Физматлит, 2006. - 310 с.
- [Добрынин, 2006б] Добрынин Д.А. Интеллектуальные роботы вчера, сегодня, завтра //X национальная конференция по искусственному интеллекту с международным участием КИИ-2006 (25-28 сентября 2006 г., Обнинск): Труды конференции. В 3-т. Т.2. М:Физматлит, 2006
- [Добрынин, Карпов 2006] Добрынин Д.А., Карпов В.Э. Моделирование некоторых форм адаптивного поведения интеллектуальных роботов. Информационные технологии и вычислительные системы №2, 2006 с.45-56
- [Панкратова, Финн, 2009] Автоматическое порождение гипотез в интеллектуальных системах. Сост. Е.С. Панкратова, В.К. Финн; Под общ. ред. В.К.Финна. Предисл. Ю.М. Арского – М.: Книжный дом «Либроком», 2009.
- [Kuznetsov, Obiedkov, 2002] S.O. Kuznetsov and S.A. Obiedkov, [Comparing Performance of Algorithms for Generating Concept Lattices](#), Journal of Experimental and Theoretical Artificial Intelligence, vol. 14 (2002), pp. 189-216. (<http://new.hse.ru/C7/C18/Kuznetsov/default.aspx>)