

Введение Экспертные системы появились в 70-х годах, как компьютерные программы, способные к работе как человек-эксперт в определенной предметной области. Одним из важнейших качеств экспертов является их способность выражать свои знания и объяснить свою позицию. Эта способность особенно важна в экспертных системах, не только для анализа работы системы во время проектирования, но и для обоснования результата предложенного системой во время ее непосредственной работы. Действительно, эксперимент, проведенный с использованием экспертной системы MYCIN показал, что врачи неохотно соглашались с рекомендациями системы, если не понимают, как оно было получено [1]. Способность объяснения важна в экспертной системе по нескольким причинам. Самая общая причина – объяснения необходимы, чтобы оправдать рекомендации, данные системой. Объяснения также помогают пользователю признавать ограничения экспертной системы. Нечес, Мур и др. [2] отметили, что рекомендации системы могут смутить пользователей, которые неуверенны в возможностях системы. Средство объяснения позволяет пользователю исследовать рассуждение системы, чтобы определить, рассматривала ли система все факты и правила и рассуждала ли в соответствии с ними [3]. Объяснения являются эффективными, если они основаны на ожидаемом уровне знаний получателя объяснения. Например, объяснение о том, как работает автомобильный двигатель, будет существенно отличаться для студента-музыканта и студента-инженера. Адаптация системы и пользователя

Одной из ключевых особенностей эффективного объяснения является способность системы взаимодействовать с каждым пользователем с учетом его конкретных потребностей и ожиданий, которые должны зависеть от знаний пользователя. В настоящее время объяснительный компонент экспертной системы не учитывает уровень знаний. Следовательно, объяснения генерируются для всех пользователей одинаково, например, объяснение предназначено для новичков, в то время как с системой работает пользователь, являющийся экспертом. Поэтому для улучшения объяснения необходимо представлять знания пользователя о предметной области, в которой работает экспертная система. Здесь можно выделить методы, основанные на статической модели пользователя. Такие модели не учитывают того, что знания увеличиваются по мере взаимодействия с системой. Альтернативой предыдущего метода является использование динамической модели пользователя, которая явно представляет знания и фиксирует изменения в знаниях пользователя происходящие в период работы с экспертной системой. Эта модель помогает адаптировать модуль объяснения к знаниям пользователя в каждый момент времени. Кроме представления знаний пользователя о предметной области, необходимо учитывать уровень детализации объяснений, поскольку он тесно связан с пользовательскими знаниями. Например, это можно реализовать следующим образом: экспертная система присваивает значение

коэффициента для каждого элемента знаний пользователя и значение порога. Тогда модуль объяснения будет отображать только те элементы, которые превышают некоторый указанный порог. Путем снижения порога, пользователь может получить более подробное объяснение и наоборот. Этот способ может предложить различные уровни детализации, не используя модель пользователя. Однако вместо того, чтобы иметь одинаковые значения коэффициента для каждого элемента, было бы целесообразно, чтобы значения менялись динамически в зависимости от уровня пользователя. Таким образом, объяснение включает в себя задачу нахождения информации из доступных источников знаний и организацию этих знаний в согласованный и последовательный текст. Важным требованием формирования такого объяснения является возможность варьировать содержания объяснение в соответствии с опытом и знаниями пользователя, регулируя уровень детализации базовых знаний в предметной области. Для достижения требуемой корректировки знаний в предметной области, необходимо разработать способ, позволяющий производить систематические изменения в представлении знаний. Однако при моделировании предположений об уровне знаний пользователя можно столкнуться с некоторой неопределенностью, которая связана со следующими факторами:

- неполнота данных при определении уровня знаний пользователя и его интересах при работе с системой;
- ошибки в указании того, какой уровень детализации можно считать достаточным для удовлетворения запроса пользователя;
- сложности в сопоставлении всех возможных уровней на всевозможные уровни детализации для имеющейся модели.

Таким образом, должен быть механизм способный обрабатывать эти вопросы. Байесовские сети являются эффективным средством решения проблемы неопределенности и, следовательно, могут быть использованы для решения задачи генерации объяснений. Байесовские сети Байесовская сеть представляет собой ориентированный ациклический граф, в котором узлы представляют переменные объектов и связи между ними. Узлы имеют маркировку условных вероятностей, которые дают оценки прочности зависимостей между значениями переменных. Поэтому, байесовские сети – это компактное, локализованное представление вероятностной модели, использующее качественное графическое описание причинно-следственных связей между субъектами и количественное измерение силы этих связей. С математической точки зрения байесовская сеть — это модель для представления вероятностных зависимостей, а также отсутствия этих зависимостей. При этом связь $A \rightarrow B$ является причиной, когда событие A является причиной возникновения B , влияет на значение, принятое B . Для того чтобы применить байесовские сети, необходимо решение следующих задач:

- структурирование сети на основе качественной информации причинных воздействий между выбранными параметрами, т.е. определение узлов сети;
- выявление вероятностной информации, чтобы комментировать ссылки сети, т.е.

принятия решения об априорных вероятностях, связанных с сетью; • выбор способа доказательства и сбора результатов. Входные данные для построения модели:

- Структурное описание системы: определить ее компоненты и взаимосвязи между ними. Это делается до начала работы и остается неизменным в течение всего процесса объяснения.
- Библиотека образцов-фрагментов, представленных на различных уровнях детализации. Каждый фрагмент модели, представляет собой определенный аспект структурной и поведенческой информации данного компонента системы. Образцовые фрагменты для различных представлений компонента организованы в класс предположений для данного компонента. Фрагменты в пределах класса предположения имеют отличающиеся и часто взаимно противоречащие условия, и таким образом только один фрагмент от класса данного компонента может использоваться, чтобы представлять компонент в системной модели.
- Требуемый уровень детализации для под-множества модели фрагментов, необходимые для построения модели системы.
- Набор объектов (элементов или процессов), полученных от запросов пользователя. Структурное описание системы используется для построения байесовской сети. Каждый узел сети выступает за выбор или отказ от конкретной модели фрагмента и может принимать значения из множества {Да, Нет}. Ребра графа байесовской сети представляют отношения «причина-следствие» и определяют описание системы: если компонент А имеет свой выход и соединен с входом компонента В, то устанавливается связи от модели компонента А по отношению к В. Учитывая структуру байесовской сети, каждый узел сети приобретает значение «да» или «нет», когда узлам-родителям присваиваются значения. При формировании логического правила, необходимо учитывать факт того является ли узел корневым узлом, т.е. тем, у которого нет родительских узлов. Фрагмент модели выбирается исходя из следующих двух важных факторов: сила причинного влияния, которое фрагмент получает от выбранных родителей и соответствия фрагмента относительно уровня детализации этих родительских узлов. Это можно сформулировать следующим образом: каждый некорневой узел MF_{ji} , где $j = 1, \dots, L$, представляет выбор или отказ компонента модели j от компонента C_i , $i = 1, \dots, M$, с родительским узлом U_{jk} , $k = 1, \dots, N$, априорная вероятность выбора соответствующего фрагмента, когда некоторые из родительских узлов принимают значение «да» (остальные родительские узлы никакого значения не принимают) определяется по формуле: где P располагается среди родителей, которые приняли значение «да» и $\sum$ - оценка суммы причинно-следственных связей влияния родителей U_{jt} на компонент MF_j , когда родитель U_{jt} принимает значение " да". Пусть n число тех переменных, которые влияют на переменные MF_{ji} и N общее число переменных, определенных в U_{jt} , тогда рассчитывается как отношение [4]. Для назначения априорных вероятностей для корневого узла, при условии, что все компоненты изначально могут быть

определены, можно использовать следующее логическое правило: для всех корневых узлов $RM\ Fri$, где $r=1, \dots, R$, компонента C_i , $i=1, \dots, S$, априорная вероятность вычисляется. Вероятности для остальных узлов можно рассчитать с использованием предыдущих значений вероятностей корневых узлов сети. Это может быть сделано с помощью принципа максимума энтропии, с учетом ограничений, что сумма всех вероятностей значений узлов должна быть равна единице и предполагая равномерное распределение этих вероятностей [4].

Учитывая структуру и априорные вероятности, рассуждения осуществляются на основании доказательства того какие компоненты являются актуальными для пользователя и какую модель механизма рассуждения использовать для текущего пользователя. Процесс рассуждения помогает с решением того какие фрагменты должны быть выбраны. Доказательство предполагает выбор обозначенных фрагментов с вероятностью равной единице. Чтобы определить вероятности узлов, для которых не существует доказательства, используется стандартный метод вывода сети Байеса. Когда процесс обоснования завершается, рассчитываются вероятности каждого фрагмента модели.

Байесовские рассуждения определяют самое вероятное представление каждого отдельного компонента относительно знаний и опыта пользователя. Несмотря на то, что данный процесс имеет математическое обоснование, необходимо удостовериться, что только один фрагмент выбран для каждого компонента и системная модель, создана комбинируя выбранные фрагменты модели через их взаимосвязи. Отметим, что нет полной гарантии того, что представление объяснений будет соответствовать информационным потребностям пользователя. Чтобы изменить решения, принятые выводом Байеса необходимы дополнительные ограничения, которые бы также устранили включение фрагментов, которые могут привести к несоответствующей модели. Эти ограничения должны быть применены всякий раз, когда образцовый фрагмент рассматривают для выбора согласно рассуждениям сети. Такая частично сформулированная модель проверяется на соответствие, и если не все ограничения удовлетворены, от образцового фрагмента отказываются. В этом случае следующий самый вероятный образцовый фрагмент (того же самого компонента) рассматривают как кандидата на выбор. От фрагмента нельзя отказаться, как только он прошел тест ограничений соответствия. Выводы

Предложен подход для разработки моделей для поддержки задачи объяснения, использующей для рассуждения байесовскую сеть. Сеть построена на основе структурного описания системы, чтобы облегчить выбор подходящей модели. Изначально выбирают фрагменты для некоторых компонентов, основанные на интересах и уровне знаний пользователя.