

УДК 004.942:009

Хлопотов Максим Валерьевич

ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет
информационных технологий, механики и оптики»
Санкт-Петербург, Россия¹

Старший преподаватель кафедры интеллектуальных технологий в гуманитарной сфере
khlopotov@niuitmo.ru

Применение байесовской сети при построении моделей обучающихся для оценки уровня сформированности компетенций

¹ 197101, г. Санкт-Петербург, Кронверкский пр., д.49, (812)2328645

Аннотация. Модель студента является одной из базовых компонент интеллектуальных компьютерных систем обучения. Модель, в общем случае, отражает знания обучающего или системы о студентах. К таким знаниям необходимо отнести факты о его предметных знаниях и умениях студента, о степени сформированности его компетенций. Существует три вида моделей обучающегося: стартовая, эталонная и текущая. Стартовая модель фиксирует характеристики обучающегося к началу процесса обучения. Цель обучения задает эталонная модель, получаемая из нормативных документов. Такую эталонную модель принято называть нормативной. При этом в качестве эталонной модели может также выступать модель, полученная в результате анализа потребностей работодателей. Текущая модель характеризует свойства обучающегося на разных этапах обучения с заранее заданной частотой. Текущая модель предназначена для слежения за учебным прогрессом конкретного обучающегося. Механизмами построения текущей модели являются диагностика и тестирование.

Байесовские сети представляют собой удобный инструмент для описания достаточно сложных процессов и событий с неопределенностями. Для описания байесовской сети необходимо определить структуру графа и параметры каждого узла. Эта информация может быть получена непосредственно из данных или из экспертных оценок. Структура байесовской сети отражает структуру компетенций студентов, и является инструментом, с помощью которого можно выносить суждения и оценки относительно уровня готовности студентов к профессиональной деятельности. В статье описан процесс построения байесовской сети для моделирования обучающихся. Основные этапы этого процесса: идентификация переменных, определение структуры, определение параметров. После того, как байесовская сеть сконструирована, она готова для того, чтобы с её помощью можно было проводить вычисления. После поступления некоторого количества свидетельств, могут быть вычислены апостериорные вероятности. В байесовской сети, построенной для моделирования обучающихся, существует два основных способа вывода: диагностические рассуждения и предсказательные рассуждения. В диагностических рассуждениях оцениваются значения вероятностей для целевых переменных, определяются наиболее вероятные состояния.

Основная цель компетентностного моделирования – построить такую байесовскую сеть, в которой все составные компетенции являются родительскими вершинами для компетенций, а все простые компетенции являются родителями для компетентностно-ориентированных заданий. В связи с этим, отметим свойства отношений между компетенциями важные для построения компетентностной модели. К таким свойствам относятся: иерархичность, эквивалентность, агрегация, общность и условность. В статье описана общую схему работы со списком компетенций, которые сформулированы в стандарте направления подготовки. Рассмотрены примеры оценки уровня сформированности компетенции.

Ключевые слова: моделирование обучающихся; интеллектуальный анализ образовательных данных; байесовская сеть; оценка компетенций; компетентностная модель; вероятностные рассуждения; язык программирования R.

Модель студента является одной из базовых компонент интеллектуальных компьютерных систем обучения. Она содержит достаточно полную информацию об обучающемся: уровень его знаний, умений и навыков, способность к обучению, способность выполнения заданий (умеет ли он использовать полученную информацию), личностные характеристики (тип, ориентация) и другие параметры. Модель студента динамична, т.е. изменяется в процессе прохождения курса, в ходе работы с системой [4].

Первые модели обучающихся были описаны в работах П.Л. Брусиловского [3], В.А. Петрушина [8] и др. В эти работы было показано, что для поддержки обучения необходимы знания о предмете обучения, о стратегиях и методах обучения, знания об обучающихся. Было предложено большое число подходов, конкретных моделей и формализмов для представлений моделей, которые используются для организации процесса обучения.

Так как в последнее время в научной и педагогической литературе вместо термина «обучаемый» принято использовать термин «обучаемых», в рамках настоящего исследования вместо термина «модель обучаемого» используется термин «модель обучающегося».

Понятие «модель обучающегося» возникло в компьютерных технологиях обучения в связи с необходимостью формализовать представления об обучающемся [23]. Применение таких технологий предполагает использование в учебном процессе информационной системы. Модель обучающегося, в общем случае, отражает знания обучающего или информационной системы о студентах. Очевидно, что, прежде всего, к таким знаниям необходимо отнести факты о его предметных знаниях и умениях студента, о степени сформированности его компетенций. Также очевидно, что модель может формироваться из характеристик обучающегося, измеряемых во время обучения. Кроме того, модель может рассматриваться как отражение результатов использования конкретных обучающих воздействий [9].

Авторы [1] предлагают классификацию, в соответствии с которой, модели обучающегося делятся на две основные группы: *фиксирующие* и *имитационные*. Также предлагается классифицировать модели по способу представления знаний: декларативные, процедурные и распределенные [6]. Наибольшее распространение на практике получили оверлейные модели, которые относятся к фиксирующим и декларативным моделям.

При построении *оверлейной модели* используется предположение об одинаковой структуре знаний студента и информационной обучающей системы. Второе предположение заключается в том, что знания (компетенции) обучающегося являются подмножеством знаний (компетенций) системы [1].

Выделим следующие характеристики обучающегося, которые могут быть учтены в параметрах модели:

1. уровень сформированности компетенции;
2. мотивация;
3. психологические характеристики;
4. стиль изучения материала;
5. скорость усвоения;
6. внимание и способность к обучению.

Рассмотрим три вида моделей обучающегося: *стартовую*, *эталонную* и *текущую*. Стартовая модель фиксирует характеристики обучающегося к началу процесса обучения. К таким характеристикам, прежде всего, относятся его начальные знания и умения. Цель обучения задает эталонная модель, получаемая из нормативных документов, в частности из федерального государственного образовательного стандарта высшего профессионального образования (ФГОС ВПО). Такую эталонную модель принято называть *нормативной*. При этом в качестве эталонной модели может также выступать модель, полученная в результате анализа

потребностей работодателей. *Текущая модель* характеризует свойства обучающегося на разных этапах обучения с заранее заданной частотой. Например, в конце каждого модуля или семестра. Текущая модель является динамической и технологически наиболее сложной. Текущая модель предназначена для слежения за учебным прогрессом конкретного обучающегося. Механизмами построения текущей модели являются *диагностика* и *тестирование*. Общий подход к построению текущей модели обучающегося развит в [1].

Байесовская сеть для моделирования обучающихся

Байесовская сеть (доверия) — это ациклический ориентированный граф, в котором каждая вершина (*узел сети*) представляет n -значную переменную, дуги обозначают существование непосредственных причинно-следственных зависимостей между соединенными переменными, а сила этих зависимостей количественно выражается в виде условных вероятностей, сопоставленных каждой из переменных [25].

Байесовские сети являются одним из видов вероятностных графических моделей. Строгое формальное определение и теория байесовских сетей доверия построены и развиты в трудах [25, 18, 10, 11].

Байесовские сети представляют собой удобный инструмент для описания достаточно сложных процессов и событий с неопределенностями. Основной идеей построения сети является разложение сложной системы на простые элементы. Для объединения отдельных элементов в систему используется математический аппарат теории вероятностей. Такой подход обеспечивает возможность строить модели с множеством взаимодействующих переменных для последующей разработки эффективных алгоритмов обработки данных и принятия решений. С математической точки зрения байесовская сеть — это модель для представления вероятностных зависимостей, а также отсутствия этих зависимостей [2].

Для описания байесовской сети необходимо определить структуру графа и параметры каждого узла. Эта информация может быть получена непосредственно из данных или из экспертных оценок. Такая процедура называется обучение байесовской сети [24].

Как отмечено в [19] байесовская сеть – это распространённый выбор исследователей для описания нечёткой связи между достижениями студента и их компетенциями во многих исследовательских проектах. Модели на основе байесовских сетей уже с конца 90-х годов прошлого века достаточно активно используются при разработке компьютерных обучающих средств [28], особенно иностранными исследователями [27, 22, 15, 26, 17, 14].

Структура байесовской сети отражает структуру знаний студентов, и является инструментом, с помощью которого можно выносить суждения и оценки относительно уровня готовности студентов, а также принимать решения [19].

В работе [20] байесовские подходы к построению моделей студентов классифицированы на три типа. Первый тип моделей, в которых эксперты определяют структуру сети, а также начальные и условные вероятности. Второй тип – модели ориентированные на максимизацию эффективности за счёт ограничения структуры сети. Третий тип – модели, основанные на данных, в которых используются данные из предыдущих экспериментов для генерации структуры сети и значений вероятностей.

Привлекательность байесовских моделей заключается в их высокой производительности, а также в интуитивно понятном представлении в виде графа [16].

В работе [22] рассматривается задача моделирования знаний при адаптивном тестировании студентов по заданной дисциплине. Структура учебного курса предполагает разбиение дисциплины на главы, а каждой из глав, в свою очередь, соответствует набор

понятий. Тестирование включает набор тестовых заданий, каждое из которых может требовать владения одним или несколькими понятиями. В свою очередь, владение каждым из понятий может быть необходимым для выполнения одного или нескольких тестовых заданий. В этой работе используется байесовская сеть с бинарными переменными, сопоставленными дисциплинам, темам, концептам и вопросам (заданиям). Условные вероятности для переменных задаются преподавателем.

В работе [27] отмечается, что измерение уровня компетенции студентов с помощью их ответов на тестовые задания – это типичная задача вероятностных рассуждений. Два наиболее частых случая, в виду которых возникает неопределённость, называются в иностранной литературе терминами *slip (промах)* и *guess (отгадка)*. Студенты могут случайно ответить неправильно на вопрос, ответ на который они знают – это ситуация называется промах. Также студенты могут случайно угадать правильный ответ или списать задание. Такой случай называется отгадка.

Рассмотрим байесовскую сеть, в которой компетенция проверяется одним заданием. Соответствующая сеть представлена на рисунке 1. В вершине A – информация о сформированности компетенции, а в вершине $a1$ – результат выполнения задания. Для упрощения рассмотрим случай, когда в вершине A переменная может принимать значение «сформирована» или «не сформирована». При этом задание оценивается так: либо «выполнено», либо «не выполнено».



Рисунок 1. Байесовская сеть, в которой компетенция A проверяется одним заданием $a1$
(разработано автором)

Предположим, дано следующее:

1. Условная вероятность правильного ответа на задание, при условии, что компетенция сформирована – 0,95 (с учётом возможности промаха)

$$p(a1 = \text{«выполнено»} \mid A = \text{«сформирована»}) = 0,95$$

2. Условная вероятность того, что задание выполнено верно, при том, что компетенция не сформирована – 0,2.

$$p(a1 = \text{«выполнено»} \mid A = \text{«не сформирована»}) = 0,2$$

Рассмотрим пример, который иллюстрирует основу использования байесовской сети для оценки сформированности компетенции.

Пример 1

Пусть начальная (априорная) вероятность того, что компетенция сформирована равняется 0,5.

$$p(A = \text{«сформирована»}) = 0,5$$

Это означает, что у нас нет никаких предположений относительно сформированности компетенции, т.е. соответствует состоянию полной неопределённости. Отсюда мы можем посчитать вероятности для узла $a1$, а именно $p(a1 = \text{«выполнено»})$ и $p(a1 = \text{«не выполнено»})$.

Очевидно, что

$$p(a1 = \text{«выполнено»}) + p(a1 = \text{«не выполнено»}) = 1$$

При этом

$$p(a1 = \text{«выполнено»}) = 0,575$$

$$p(a1 = \text{«не выполнено»}) = 1 - p(a1 = \text{«выполнено»}) = 0,425$$

Если студент выполняет задание верно, т.е. мы получаем свидетельство $a1 = \text{«выполнено»}$. В этом случае может быть рассчитана апостериорная вероятность сформированности у студента компетенции, т.е. рассчитывается значение $p(A = \text{«сформирована»} | a1 = \text{«выполнено»})$.

$$p(A = \text{«сформирована»} | a1 = \text{«выполнено»}) = 0,826$$

Вероятность 0,826 довольно высока, и гораздо выше, чем изначальная вероятность 0,5. Однако, эта вероятность пока ещё далека от 1 в виду того, что существует возможность отгадки.

Рассмотрим теперь случай, в котором студент выполняет задание неверно, т.е. мы получаем свидетельство $a1 = \text{«не выполнено»}$. В этом случае также может быть рассчитана апостериорная вероятность сформированности у студента компетенции, т.е. рассчитывается значение $p(A = \text{«сформирована»} | a1 = \text{«не выполнено»})$.

$$p(A = \text{«сформирована»} | a1 = \text{«не выполнено»}) = 0,059$$

Таким образом, остаётся небольшая возможность того, что компетенция всё-таки сформирована, т.е. произошёл промах.

$$\text{В случае если, например, } p(A = \text{«сформирована»}) = 0,9$$

Если задание выполнено верно студентом, уверенность в том, что компетенция сформирована очень близка к 1.

$$p(A = \text{«сформирована»} | a1 = \text{«выполнено»}) = 0,977$$

В случае неверно выполненного задания, существует довольно высокая вероятность того, что компетенция не сформирована.

$$p(A = \text{«сформирована»} | a1 = \text{«не выполнено»}) = 0,36 \blacksquare$$

Из примера 1 можно сделать следующие выводы:

1. если компетенция проверяется одним заданием, априорная вероятность сформированности компетенции существенно влияет на апостериорную. В связи с этим необходимо увеличивать количество заданий;

2. для того, чтобы суждения о сформированности компетенции были более достоверными, т.е. вероятность приближалась к 0 или к 1, необходимо снижать вероятность промаха и отгадки.

При построении байесовской сети для моделирования обучающихся важно учитывать следующие свойства:

1. байесовская сеть – это ориентированный ациклический граф. Вершины байесовской сети попарно соединяются между собой ориентированными рёбрами. При этом в графе отсутствуют циклы, состоящие из ориентированных рёбер;
2. вершинами графа являются переменные из модели обучающегося, т.е. некоторые признаки обучающегося, например, уровни сформированности компетенций и оценки за выполненные компетентностно-ориентированные задания. Эти признаки могут быть непрерывными или дискретными. При моделировании обучающихся чаще используются дискретные данные, т.к. при преобразовании образовательных данных они практически всего при необходимости проходят дискретизацию;
3. для любых двух вершин байесовской сети X и Y , если ребро направлено от X к Y , то X называется *родительской* по отношению к Y . При этом у любой вершины могут быть несколько родительских вершин. В свою очередь родительская вершина может являться таковой для одной или нескольких вершин;
4. все вершины, у которых есть родительские вершины, определяются таблицей (или функцией) условных вероятностей. Это позволяет формализовать рассуждения о сформированности компетенции следующим образом. Если студент владеет компетенцией, то он выполнит соответствующее этой компетенции задание. Таким образом, может быть задана условная вероятность для выполнения задания, при условии, что компетенция сформирована. Если $parents(X)$ – это множество родительских вершин для вершины X в байесовской сети, то X характеризуется распределением условных вероятностей $P(X | parents(X))$, которое количественно оценивает влияние родительских вершин на вершину X .

Если задана байесовская сеть со случайными элементами $X = \{ X_1, \dots, X_n \}$, тогда их совместное распределение определяется формулой:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | parents(X_i)) \quad (1)$$

Формула (1) позволяет вычислить маргинальную вероятность любой комбинации значений переменных набора X . Кроме того, при поступлении свидетельства о наборе значений, полученных некоторыми переменными сети, можно вычислить условную вероятность любой комбинации значений других переменных в сети, используя теорему Байеса.

5. Для вершин, не имеющих родительских вершин, вероятности являются *безусловными* (*маргинальными*). Другими словами, если у вершины X нет родителей, то распределение вероятностей в ней является *безусловным*, в противном случае – *условным*. Вероятность определяется по формуле:

$$P(X_i | \text{parents}(X_i)) = P(X_i) \quad (2)$$

6. Если значение в узле получено в результате опыта, то саму вершину называют *свидетелем*, а результат такого опыта – *свидетельством*.

Таким образом, отношения в правильно построенной байесовской сети определяют условную зависимость переменных.

Опишем процесс построения байесовской сети для моделирования обучающихся. Основные этапы этого процесса: идентификация переменных, определение структуры, определение параметров.

Моделирование должно начинаться с *идентификации переменных*, которые относятся к моделируемой предметной области. Переменные можно разделить на четыре класса в соответствии с их ролью в модели: целевые, свидетельства, факторы, вспомогательные. Рассмотрим эти классы переменных подробнее.

1. *Целевые переменные*. Эти переменные используются для моделирования того, что представляет интерес. Как правило, целевые переменные отражают скрытые (латентные) характеристики. Это означает, что нет возможности измерить их напрямую. В системах технической диагностики целевыми переменными называют неисправности. Примером целевой переменной в образовании является понимание студентом какого-либо понятия. Это нельзя измерить напрямую, а только с помощью, например, теста или экзамена. Уровень сформированности компетенции – это целевая переменная.

2. *Свидетельства*. Переменные свидетельства иначе называют переменными наблюдения. Они используются для того, чтобы предоставить информацию относительно целевых переменных. В моделировании обучающихся свидетельствами могут быть действия пользователя. Причём это могут быть совершенно разные уровни действий – от нажатия на кнопку мыши до выполнения компетентностно-ориентированного задания.

3. *Факторы*. Это переменные, которые моделируют источники влияния на целевую переменную. Их также называют контекстными переменными. Факторы делят на четыре категории в соответствии с их влиянием на переменную: промодутеры, замедлители, требования, исключения. В статье [Millan, 2010] даны описания этих категорий факторов. *Промодутеры* имеют положительную корреляцию с целевой переменной и способствуют тому, чтобы характеристика проявилась. *Замедлители* действуют наоборот и имеют отрицательную корреляцию. *Требования* являются обязательными для того, чтобы связанная характеристика проявилась. *Исключения* сводят вероятность возникновения связанной характеристики до нуля.

Включение слишком большого числа факторов в модель является ошибкой. Это делает модель менее прозрачной и понятной. Простые модели проще оценивать, понимать и интерпретировать. Однако это не означает, что нужно всегда сокращать модель и избавляться от контекстных переменных. Если переменные предоставляют важную и полезную информацию, они должны быть включены в итоговую модель. Примерами факторов в байесовских сетях для моделирования обучающихся являются мотивация студентов, их личностные качества, квалификация преподавателей и т.д.

4. *Вспомогательные переменные*. Эти переменные используются для удобства. Например, если у узла есть много родительских узлов, промежуточные вспомогательные переменные могут быть использованы для их группировки. За счёт этого структура сети упрощается, а количество параметров уменьшается.

Кроме деления переменных в соответствии с их ролью, существует деление в зависимости от природы измерений. В этой связи выделяют дискретные и непрерывные

переменные. Особый вид дискретных переменных, часто используется в байесовских сетях для моделирования обучающихся, это переменные с двумя состояниями. Такие переменные называют бинарными. Непрерывные переменные чаще всего дискретизируются.

После определения переменных следующий шаг в построении модели – это *определение структуры*. Структура сети определяется расстановкой рёбер между узлами-переменными. Как уже упоминалось выше, в байесовской сети рёбра являются направленными. Изменение направления ребра имеет значение. При этом смысл ребра заключается в том, что переменная в исходной вершине оказывает непосредственное влияние на переменную в целевой. Таким образом, случайные события соединяются причинно-следственными связями. В связи с этим байесовские сети иногда называют причинно-следственными (causal). Однако, с математической точки зрения байесовские сети не обязательно говорят о причинно-следственной связи между переменными. Часто говорят также о диагностической связи между узлами в сети. Структура байесовской сети может быть получена непосредственно из данных или из экспертных оценок.

Итоговый шаг в построении модели – *определение параметров*. Для этого нужно задать априорные распределения для узлов, у которых нет родителей (корневые узлы), а также распределения условных вероятностей для всех остальных узлов байесовской сети. Как и в случае с определением структуры, параметры могут быть заданы экспертом, либо получены из данных. Также существует возможность комбинировать оба этих подхода.

Отметим, что априорное распределение вероятностей – это безусловное распределение, которое не изменяется, какие бы ни были получены свидетельства.

Отметим, что каждая переменная, в общем случае, может принимать неограниченное число значений. На большое число значений в значительной степени усложняет модель, так как в этом случае требуется задавать огромное количество параметров для распределений условных вероятностей. В связи с этим часто в байесовских сетях используют бинарные переменные.

После того, как байесовская сеть сконструирована, она готова для того, чтобы с её помощью можно было проводить вычисления. После поступление некоторого количества свидетельств (наблюдений), могут быть вычислены апостериорные вероятности. В байесовской сети, построенной для моделирования обучающихся, существует два основных способа вывода: диагностические рассуждения и предсказательные рассуждения. В диагностических рассуждениях оцениваются значения вероятностей для целевых переменных, определяются наиболее вероятные состояния.

Оценка сформированности компетенции

Переход высшего профессионального образования на ФГОС ВПО третьего поколения ставит перед вузами ряд задач. Реализация целей и задач ООП, результатом освоения которых являются компетенции выпускников, актуализирует проблему формирования и оценивания уровней сформированности компетенций на различных стадиях обучения студентов [5].

Под *компетенцией* будем понимать способность индивидуума применять знания, умения и навыки в профессиональной деятельности. Показателем сформированности компетенции является готовность обучающегося к профессиональной деятельности [13].

Принципиально важно, что гарантированное обеспечение сформированности профессиональной компетенции зависит от чёткости и однозначности формулировки исходной компетенции. Она, в свою очередь, однозначно определяется видом профессиональной деятельности и её областью, которая и выступает в качестве её объекта.

Детализация компетенций позволяет, с одной стороны, определить их иерархию, а с другой – сформулировать для каждого уровня результат обучения. В формулировках каждой компетенции выражаются ожидания того, какую профессиональную деятельность выпускник освоит и в качестве какого специалиста он сможет предложить себя на рынке труда.

Компетентностная модель специалиста представляет собой совокупность учебно-методической документации, регулирующей цели обучения и воспитания, а также задачи и ожидаемые результаты подготовки выпускников. Т.е. компетентностная модель специалиста является эталонной моделью. *Компетентностная модель выпускника* формируется в ходе обучения. Т.е. компетентностная модель выпускника является текущей моделью. Полученная модель является ожидаемым результатом обучения и закладывается в основу разработки учебного плана и рабочих программ.

Компетентностная модель специалиста разрабатывается выпускающей кафедрой на основе Федерального государственного образовательного стандарта (ФГОС) и во взаимодействии с работодателями.

Сотрудники выпускающей кафедры детализируют указанные в ФГОСе профессиональные компетенции, учитывая потребности рынка, вуза и студентов. Полученная модель является ожидаемым результатом обучения и закладывается в основу разработки учебного плана и рабочих программ.

Построение компетентностной модели специалиста должно предваряться созданием компетентностной модели для направления подготовки, содержащей всё множество формируемых компетенций, а также связи между ними. Фактически модели специалиста и выпускника представляет собою выборки этой исходной модели.

Сравнение моделей специалиста и выпускника позволит найти разницу между текущей и эталонной моделями. На основе полученных результатов делаются первичные выводы о степени соответствия выпускников требованиям рынка труда, а также разрабатывается комплекс мер по уменьшению или полному устранению разницы между моделями.

В процессе диагностики выявляется фактический результат и степень достижения ожидаемого результата. Для выявления степени достижения необходимо использовать параметры: идентификатор и название компетенции и уровень освоения компетенции. В случае с компетентностной моделью специалиста требуемый уровень освоения также определяется кафедрой.

В качестве показателя сформированности компетенции выступает готовность выполнять профессиональную деятельность [13].

Готовности проявляются при выполнении студентами компетентностно-ориентированных заданий. При этом особое внимание в структуре готовности следует уделять мотивационному компоненту, так как именно он влияет на эффективность формирования других компонентов готовности, а также предопределяет выбор студентом дальнейшего образовательного маршрута (магистратура, аспирантура).

Компетентностно-ориентированные задания (КОЗ) создаются для проверки уровня сформированности компетенций. Конструкция КОЗ, подчиняется определённым требованиям, обусловленным тем, что такое задание организует деятельность студента, а не воспроизведение им учебной информации [5].

Результаты обучения формулируются преподавателями во взаимодействии с работодателями в виде компетенций, интегрирующих в себе знания, понимание, умения, опыт и ценностные ориентации, требуемых от специалистов.

Дескрипторы уровней в обобщённом виде описывают результаты обучения для различных уровней сформированности компетенций как перечень описаний общих характеристик, ожидаемых и требуемых от студентов результатов обучения. Формулирование дескрипторов уровней компетенций – задача, решаемая разработчиками компетентностной модели студента по основной образовательной программе. Уровень сформированности компетенции является латентным признаком объекта обучения.

Использование дескриптора приводит к разбиению требований к уровням освоения компетенций на отдельные части. Выделим три основных уровня освоения компетенций: начальный, базовый, углублённый. Начальный – знания и умения, необходимые и достаточные для понимания, постановки проблемы и её практического решения. Базовый – знания и умения, необходимые и достаточные для применения эффективных алгоритмических методов и моделей при решении типовых задач. Углублённый – знания и умения, необходимые и достаточные для применения эффективных методов и моделей при решении типовых и профессиональных задач [5]. Для построения байесовской сети добавим также нулевой уровень сформированности компетенции.

Для подтверждения достижения необходимо оценивание и сопоставление ожидаемого и достигнутого результата образования по уровням освоения [7].

При построении компетентностной модели на основе байесовской сети, в вершины записывается вероятность сформированности компетенции на каждом из уровней. При этом, очевидно, должно быть соблюдено условие, при котором сумма всех вероятностей равняется 1. Безусловные априорные вероятности либо назначаются экспертом, например, преподавателем, который формирует соответствующую компетенцию, либо вычисляются из образовательных данных, если таковые данные имеются. Например, если для большой группы известно уровни сформированности компетенции для большого количества студентов, освоивших образовательную программу. На априорную вероятность могут повлиять личностные качества студентов, такие как мотивация, психологические характеристики, стиль изучения материала, скорость усвоения, внимание и способность к обучению. Эти компоненты модели могут быть вершинами байесовской сети. При этом такие вершины являются для компетенций родительскими. При наличии таких вершин, они должны являться свидетелями. В этом случае, для компетенций определяется условная априорная вероятность.

Также назначается экспертом или вычисляется из данных условная вероятность для узлов сети, которые для которых компетенция является родителем. Такими узлами могут быть дисциплины или отдельные темы, компетентностно-ориентированные задания. Собственно, результаты выполнения компетентностно-ориентированных заданий становятся свидетельствами, по которым впоследствии вычисляется апостериорное значение для вероятности сформированности компетенции на заданном уровне.

В байесовской сети, построенной в качестве компетентностной модели, компетентностно-ориентированные задания связываются рёбрами с теми компетенциями, сформированность которых они проверяют. Простейший пример такой связи показан на рисунке 2. Здесь A – это компетенция, а $a1$ и $a2$ – это компетентностно-ориентированные задания, связанные с компетенцией A .

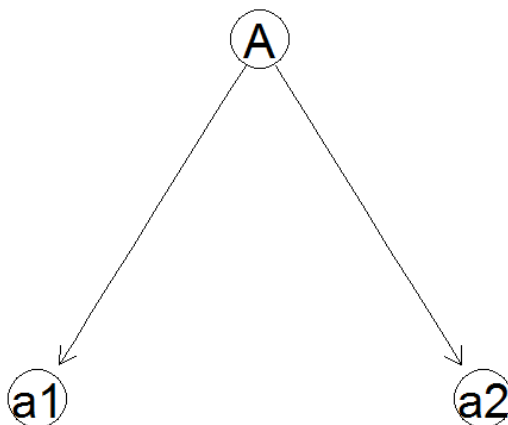


Рисунок 2. Простейшая связь компетенции и двух компетентностно-ориентированных заданий (разработано автором)

С одной компетенцией могут быть связаны несколько компетентностно-ориентированных заданий. При этом количество таких заданий должно быть не меньше, чем количество заранее заданных в дескрипторе уровней компетенции.

Для упрощения байесовской сети в настоящей работе накладывается следующее ограничение: одно компетентностно-ориентированное задание проверяет только одну компетенцию. У задания есть уровень сложности и уровень важности. Эти уровни задаются авторами заданий. От значения уровней сложности и важности зависит условная вероятность в вершине с компетентностно-ориентированным заданием.

Свидетельства могут иметь разный вид в зависимости от применяемой шкалы оценок. Например, если автор задания учитывает только сам факт правильного выполнения, то свидетельства «задание выполнено» или «задание не выполнено» образуют полную группу всех возможных свидетельств. В случае, если преподаватель выставляет за выполненное задание балл, свидетельствами могут быть оценки, например: «2», «5», «80», «В», и т.п. Чем больше значений могут принимать переменные в вершинах байесовской сети, тем сложнее становится настройка её параметров. Ведь, например, при трёх возможных уровнях сформированности компетенций и при четырёхбалльной системе оценки (2, 3, 4, 5) для каждого задания нужно прописать заранее 12 значений для условной вероятности.

С компетентностно-ориентированными заданиями в байесовской сети связываются только простые (элементарные) компетенции. Однако простые компетенции должны быть получены на основе анализа ФГОС ВПО, в котором перечислены общекультурные и профессиональные компетенции, которые должны быть сформированы у выпускников направления подготовки.

Основная цель компетентностного моделирования – построить такую байесовскую сеть, в которой все составные компетенции являются родительскими вершинами для компетенций, а все простые компетенции являются родителями для компетентностно-ориентированных заданий.

В связи с этим, отметим свойства отношений между компетенциями важные для построения компетентностной модели. К таким свойствам относятся: иерархичность, эквивалентность, агрегация, общность и условность.

Иерархичность означает, что компетенция, сформулированная в ФГОС ВПО, является составной и может быть декомпозировано на другие составные или простые компетенции, которые в свою очередь будут проверяться компетентностно-ориентированными заданиями. Пример составной компетенции представлен на рисунке 3. Здесь компетенция А является составной компетенцией и может быть описана в виде декомпозиции. При этом компетенции более низкого уровня также могут являться составными. Декомпозиция проводится до момента, когда не останется не декомпозированных составных компетенций.

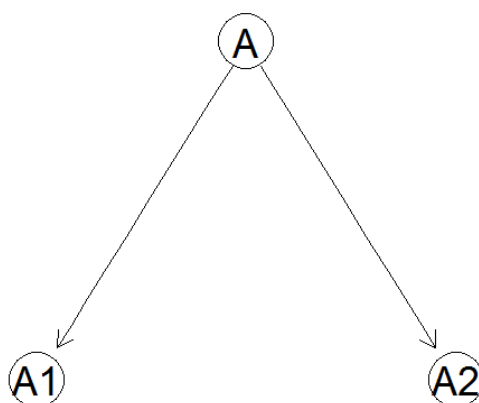


Рисунок 3. Фрагмент байесовской сети с декомпозицией составной компетенции А на А1 и А2 (разработано автором)

Эквивалентность означает, что две или более разных компетенций в ФГОС ВПО или в требованиях работодателей, на самом деле, означают одно и то же, т.е. требуют готовность к одинаковой деятельности. В байесовской сети такое отношение между компетенциями порождает сходящуюся связь, представленную на рисунке 4. Здесь компетенции А и В эквивалентны. Вместо них будет использована вершина АВ1. Именно для АВ1 первоначально будет посчитана вероятность на основе полученных свидетельств.

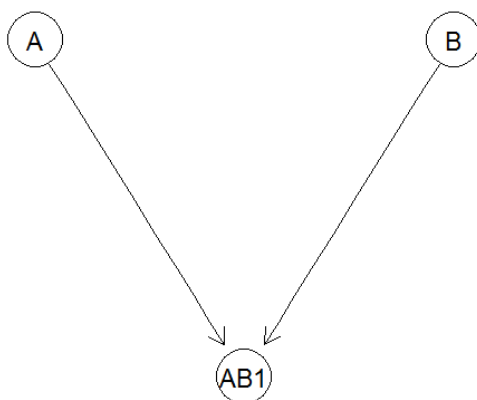


Рисунок 4. Фрагмент байесовской сети с эквивалентными компетенциями *A* и *B* (разработано автором)

Агрегация означает, что одна или более компетенция в ФГОС ВПО являются частью другой компетенции. В байесовской сети такое отношение между компетенциями порождает последовательную связь, представленную на рисунке 5. Здесь компетенция *B* является частью компетенции *A*. При этом обе эти компетенции, в свою очередь, могут быть декомпозированы до простых компетенций.

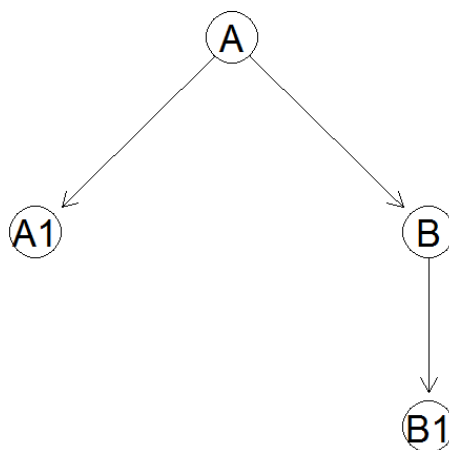


Рисунок 5. Фрагмент байесовской сети, в которой компетенция *B* является частью компетенции *A* (разработано автором)

Общность означает, что у двух или более компетенций в ФГОС ВПО имеется общая часть. Следовательно, каждая из этих компетенций является составной и к ним применимо свойство иерархичности. В байесовской сети такое отношение между компетенциями можно представить следующим образом (см. рисунок 6). Здесь *A* и *B* связаны отношением общности через общую составную часть в виде компетенции *AB1*. *A1* и *B1* в свою очередь являются независимыми составными частями соответственно составных компетенций *A* и *B*.

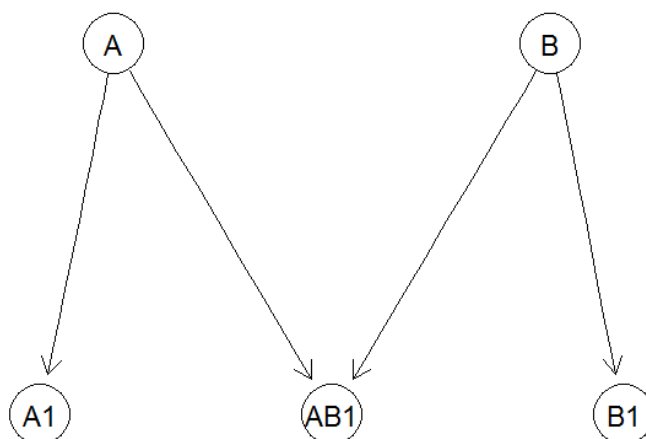


Рисунок 6. Фрагмент байесовской сети с отношением общности между компетенциями *A* и *B* (разработано автором)

Условность означает, что для освоения компетенции необходимо, чтобы у студента была сформирована другая компетенция. Такое отношение задаётся последовательной связью. На рисунке 7 показана условная зависимость компетенции *B* от компетенции *A*. Это значит, что для освоения компетенции *B* необходимо, чтобы у студента была сформирована на определённом уровне компетенция *A*.

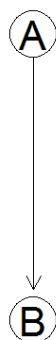


Рисунок 7. Фрагмент байесовской сети, в которой компетенция *B* условно зависит от компетенции *A* (разработано автором)

Опишем общую схему работы со списком компетенций, которые сформулированы в ФГОС ВПО направления подготовки.

Алгоритм 1. Построение структуры байесовской сети для компетентностной модели направления подготовки

На первом шаге в этом списке выявляются эквивалентные компетенции, если они есть. Эквивалентные компетенции заменяются простой или составной компетенцией.

На следующем шаге в списке компетенций выявляются агрегации. Если какая-то компетенция является частью другой, между ними устанавливается связь. Другой детализации не проводится.

Затем выявляются составные компетенции. Каждая составная компетенция детализируется путём выделения двух или более составляющих. После детализации всех составных компетенций получены все переменные составляющие структуру байесовской сети. Однако выявлены ещё не все связи.

Если среди составляющих могут быть обнаружены эквивалентные компетенции, значит, имеет место общность двух или более компетенций. Эта связь устанавливается между родителями эквивалентных компетенций.

Вслед за этим переходят к завершающему этапу построения структуры байесовской сети. На этой этапе между компетенциями устанавливается условная связь.

В результате получаем компетентностную модель, которая может быть положена в основу оценки сформированности компетенций.

Выше был рассмотрен пример 1, в котором показано, как компетентностно-ориентированные задания связываются с компетенцией. Однако в этом примере с компетенцией было связано лишь одно задание. Рассмотрим более реалистичный пример.

Пример 2

Рассмотрим байесовскую сеть, в которой компетенция проверяется тремя заданиями. Соответствующая сеть представлена на рисунке 8.

В вершине A – фиксируется информация о сформированности компетенции, а в вершинах $a1$, $a2$, $a3$ – соответственно результаты выполнения заданий. В вершине A переменная может принимать значения 0, 1, 2 и 3 в соответствии с возможными уровнями сформированности компетенции. Эти уровни должны быть описаны в дескрипторы компетенции. При этом задания оцениваются по пятибалльной системе: 1, 2, 3, 4 и 5.

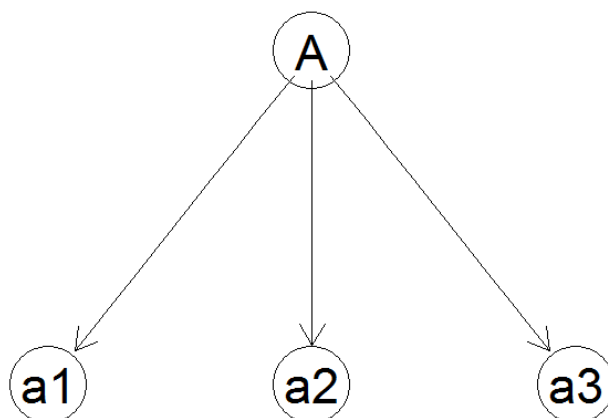


Рисунок 8. Байесовская сеть с одной компетенцией A , которая проверяется тремя заданиями (разработано автором)

Пусть начальная (априорная) вероятность того, что компетенция сформирована равна для всех четырёх уровней.

$$p(A = \langle 0 \rangle) = p(A = \langle 1 \rangle) = p(A = \langle 2 \rangle) = p(A = \langle 3 \rangle) = 0,25$$

Это означает, что у нас нет никаких предположений относительно сформированности компетенции.

Определим распределения условной вероятности для узлов, в которых записаны свидетельства. Пусть задание a_1 проверяет сформированность компетенции на уровне 1, задание a_2 – на уровне 2, задание a_3 – на уровне 3, соответственно.

Для того чтобы более точно задать условные вероятности необходимо воспользоваться данными, полученными из учебного процесса. Но в этом примере зададим вероятности экспертно, исходя из следующих предпосылок:

1. студенты, с компетенцией сформированной на более высоком уровне, будет чаще выполнять задания на высокие баллы (5 и 4), чем студенты с более низким уровнем сформированности компетенции;
2. вероятность отгадки одинакова для всех заданий, вне зависимости от уровня сложности;
3. вероятность выполнения сложного задания ниже, чем вероятность выполнения более лёгкого задания;
4. задание, за которое студентом получено 1 или 2 балла, считается невыполненным.

Для удобства представим условные вероятности в виде таблиц (см. таблицы 1-3).

Таблица 1

Условные вероятности для задания a1

Условная вероятность (%)	Уровень сформированности компетенции	0	1	2	3
Оценка за компетентностно-ориентированное задание	1	50	2	2	2
	2	30	3	3	3
	3	15	15	10	5
	4	3	20	15	10
	5	2	60	70	80

Таблица 2

Условные вероятности для задания a2

Условная вероятность (%)	Уровень сформированности компетенции	0	1	2	3
Оценка за компетентностно-ориентированное задание	1	60	2	2	2
	2	25	3	3	3
	3	10	20	15	10
	4	3	25	20	15
	5	2	50	60	70

Таблица 3

Условные вероятности для задания a3

Условная вероятность (%)	Уровень сформированности компетенции	0	1	2	3
Оценка за компетентностно-ориентированное задание	1	70	2	2	2
	2	20	3	3	3
	3	5	15	15	10
	4	3	20	20	15
	5	2	60	60	70

(Таблицы 1-3 составлены автором)

Для создания воспользуемся программой, написанной на языке R с помощью библиотеки gRain.

Листинг 1

Код для создания байесовской сети

```
#устанавливаем и подключаем необходимые библиотеки
source("http://bioconductor.org/biocLite.R")
biocLite("RBGL")
library(gRbase)
library(gRain)
library(Rgraphviz)
# задаём уровень компетенций
lvl <- c("3","2","1","0")
# задаём шкалу оценок
marks <- c("5","4","3","2","1")
# задаём априорные вероятности для сформированности компетенций
# уровни равновероятны
A <- cptable(~A, values=c(25,25,25,25), levels=lvl)
# задаём условные вероятности для заданий a1, a2, a3
a1.A <- cptable(~a1 | A, values=c(80, 10, 5, 3, 2,
                                70, 15, 10, 3, 2,
                                60, 20, 15, 3, 2,
                                2, 3, 15, 30, 50),levels=marks)
a2.A <- cptable(~a2 | A, values=c(70, 15, 10, 3, 2,
                                60, 20, 15, 3, 2,
                                50, 25, 20, 3, 2,
                                2, 3, 10, 25, 60),levels=marks)
a3.A <- cptable(~a3 | A, values=c(60, 20, 15, 3, 2,
                                50, 25, 20, 3, 2,
                                45, 35, 25, 3, 2,
                                2, 3, 5, 20, 70), levels=marks)
# строим структуру графа с параметрами
cpt.list <- compileCPT(list(A, a1.A, a2.A, a3.A))
bnet <- grain(cpt.list)
bnet <- compile(bnet)
# нарисовать байесовскую сеть
plot (bnet$dag)
```

Результат выполнения команды plot представлен на рисунке 8.

Теперь мы можем посчитать по теореме Байеса вероятности для узлов $a1$, $a2$ и $a3$.
Выполняем следующий запрос:

```
# запрос вероятностей выполнения заданий  
querygrain(bnet, nodes=c("a1", "a2", "a3"))
```

Вывод по запросу:

```
a1  
      5      4      3      2      1  
0.5300 0.1200 0.1125 0.0975 0.1400  
a2  
      5      4      3      2      1  
0.4550 0.1575 0.1375 0.0850 0.1650  
a3  
      5      4      3      2      1  
0.3823 0.1995 0.1568 0.07181 0.1895
```

Обратим внимание на то, что априорная вероятность выполнения более лёгкого задания выше, что соответствует предпосылкам.

$$p(a1 = \text{«выполнено»}) = p(a1 > 2) = 0,53 + 0,12 + 0,1125 = 0,7625$$

$$p(a2 = \text{«выполнено»}) = p(a1 > 2) = 0,455 + 0,1575 + 0,1375 = 0,75$$

$$p(a3 = \text{«выполнено»}) = p(a1 > 2) = 0,3823 + 0,1995 + 0,1568 = 0,7386$$

Рассмотрим несколько ситуаций, связанных с получением разных свидетельств. Свидетельство может быть по каким-то причинам вообще не получено, т.е. не имеется данных относительно результатов выполнения задания. Такая ситуация возникает, например, в случаях, если задание ещё не было выдано или студент по уважительной причине пропустил срок сдачи задания. Учитывая это, в любой из вершин со свидетельствами возможны шесть разных значений. Следовательно, всего существует $6! = 720$ возможных вариантов. Не будем рассматривать их все, остановимся на наиболее показательных случаях.

1. Предположим, студент выполнил все три задания на оценку 5.

Установим свидетельство:

```
bnet.ev<-setEvidence(bnet, nodes = c("a1", "a2", "a3"),  
                    states = c("5", "5", "5"))
```

И выполним следующий запрос:

```
querygrain(bnet.ev, nodes=c("A"))
```

Вывод по запросу:

A				
	3	2	1	0
	0.5024	0.3140	0.1835	1.196e-05

Из вывода следует, что наиболее вероятно (50,24%), у студента компетенция сформирована на уровне 3. Также можно говорить о том, что с вероятностью 81,64% уровень сформированности компетенции у студента не ниже чем 2. Для того чтобы делать более уверенные суждения, т.е. для приближения к 1 или к 0, можно увеличить количество сложных заданий, изменить априорную вероятность, ввести в байесовскую сеть дополнительные факторы (мотивация, личностные качества и т.д.)

2. Предположим, студент выполнил задание a1 с оценкой 5, a2 – 4, a3 – 3. Аналогично предыдущему случаю установим свидетельства и выполним запрос. Вывод по запросу:

	3	2	1	0
	0.2247	0.3495	0.4255	0.0004

С уверенностью около 78% можно сказать, что уровень сформированности компетенции у студента 1 или 2.

3. Отметим также, что байесовская сеть способна рассчитать вероятность сформированности компетенции даже в случае отсутствующих значений. Предположим, что студент выполнил задание a1 с оценкой 5, a2 – 4, а задание a3 пока не выполнил.

В этом случае результат вычисления вероятностей следующий:

	3	2	1	0
	0.2923	0.3410	0.3653	0.0015

Эта особенность позволяет делать прогнозы относительно того, сможет ли студент выполнить задание, т.е. решается одна из основных задач интеллектуального анализа образовательных данных [12]. Например, сделав следующий запрос можно получить оценку для вероятности выполнения студентом задания a3, при условии, что задание a1 и a2 им уже выполнены на оценки 5 и 4 соответственно

```
querygrain(bnet.ev, nodes=c("a3"))
```

Вывод по запросу:

a3					
	5	4	3	2	1
	0.4953	0.2600	0.1952	0.0293	0.0203

То есть студент с вероятностью около 95% выполнит задание на положительную оценку 3, 4 или 5.

4. Предположим, студент не справился ни с одним заданием, получив по 2 балла за каждое из трёх заданий.

Вероятности для каждого уровня сформированности в таком случае следующие:

A				
	3	2	1	0
	0.0018	0.0018	0.00168	0.9948

Про такого студента с вероятностью более 99% можно сказать, что у него не сформирована компетенция.

Сформулируем *правило интерпретации* уровня сформированности компетенции (для случая с 4 уровнями в дескрипторе)

Для интерпретации уровня сформированности компетенции необходимо знать требуемый уровень и посчитать вероятность того, что у студента уровень сформированности не ниже, чем требуемый.

Обозначим $Level(A)$ как требуемый уровень сформированности компетенции A . $Level(A)$ может принимать значения от 1 до 3.

Обозначим (p_0, p_1, p_2, p_3) – как вероятность сформированности компетенции в соответствии с возможными уровнями.

Обозначим, $Suc(A)$ – как вероятность того, что требуемый уровень сформированности компетенции достигнут. $Suc(A)$ вычисляется по формуле:

$$Suc(A) = \sum_{i=Level(A)}^3 p_i \quad (3)$$

Обозначим t – как уровень достоверности. t задаёт требуемый уровень достоверности и может изменяться от 0 до 1.

Правило интерпретации уровня сформированности компетенции заключается в следующем, цель обучения достигнута, если $Suc(A) > t$ и не достигнута в обратном случае. ■

Алгоритм 2. Построение байесовской сети для оценки сформированности компетенций

Примечание 1: алгоритм не учитывает возможные связи между заданиями.

Примечание 2: алгоритм работает с заданиями, которые оцениваются по двум разным шкалам: пятибалльная или бинарная.

Примечание 3: в сети не используются факторы.

1. В байесовскую сеть, построенную по алгоритму 1, добавляется новый тип узлов: компетентностно-ориентированные задания. Узлы с заданиями связываются родительскими узлами: компетенциями.

2. Назначается уровень сложности заданий. Уровень соответствует одному из уровней сформированности компетенции в дескрипторе. Повторяется для всех узлов заданий.

4. Определяется шкала для оценки выполнения заданий. Выбор между пятибалльной и бинарной.

3. В соответствии со сложностью заданий по таблицам 1-3 рассчитываются параметры (распределения условных вероятностей).

4. Повторяется шаг 3 для всех узлов заданий. ■

Алгоритм 3. Оценка сформированности уровня компетенции

1. При наличии результатов выполнения заданий в байесовскую сеть, построенную по алгоритму 1, добавляются свидетельства. При отсутствии свидетельств для расчётов будут использованы априорные условные распределения вероятностей.

2. Рассчитывается апостериорная вероятность для уровня сформированности интересующей компетенции.

3. Вероятности интерпретируются по правилу интерпретации уровня сформированности компетенции (см. пример 2). ■

Выводы

В статье рассмотрены теоретические основы моделирования обучающихся, в том числе на основе байесовских сетей. Описаны алгоритмы построения компетентностной модели с учётом свойств компетенции и построения байесовской сети для оценки компетенции. Предложено правило интерпретации вероятностей для оценки достижения цели обучения. Рассмотрены примеры оценки уровня сформированности компетенций на основе построенной модели.

ЛИТЕРАТУРА

1. Атанов Г.А., Пустынникова И.Н. Обучение и искусственный интеллект, или Основы современной дидактики высшей школы. – Донецк : Изд-во ДОУ, 2002. – 504 с.
2. Бидюк П.И., Терентьев А.Н. Построение и методы обучения байесовских сетей // Таврійський вісник інформатики і математики, №2/2004 – с. 139-154.
3. Брусиловский П.Л. Интеллектуальные обучающие системы//Информатика. Информационные технологии. Средства и системы. 1990. №2. С.3-22.
4. Буль Е.Е. Обзор моделей студента для компьютерных систем обучения // Образовательные технологии и общество (Educational Technology & Society). 2003. Том. 6, № 4. С. 245-250. Режим доступа: http://ifets.ieee.org/russian/depository/v6_i4/html/G.html (дата обращения 01.07.2014)
5. Ефремова Н.Ф. Компетенции в образовании. Формирование и оценивание. - М.: Национальное образование, 2012. - 416 с.
6. Коляда М.Г. Виды моделей, обучаемых в автоматизированных обучающих системах // Искусственный интеллект. 2008. № 2. Режим доступа: http://www.nbu.gov.ua/portal/natural/ii/2008_2/JournalAI_2008_2/Razdel2/00_Kolyada.pdf
7. Лисицына Л.С. Методология проектирования модульных компетентностно-ориентированных образовательных программ. – СПб: СПбГУ ИТМО, 2009. – 50 с.
8. Петрушин В.А. Экспертно-обучающие системы. – К.: Наукова думка, 1992. – 196 с.
9. Рыбина Г.В. Обучающие интегрированные экспертные системы: некоторые итоги и перспективы // Искусственный интеллект и принятие решений, 1/2008. – стр. 22-56
10. Тулупьев, А. Л. Байесовские сети: логико-вероятностный подход / А. Л. Тулупьев, С. И. Николенко, А. В. Сироткин. — СПб.: Наука, 2006. — 607 с. 95
11. Тулупьев, А. Л. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах / А. Л. Тулупьев, А. В. Сироткин, С. И. Николенко. — СПб.: Изд-во С.-Петербур. ун-та, 2009. — 400 с.
12. Хлопотов М.В., Коцюба И.Ю. Методы интеллектуального анализа данных для мониторинга и диагностики качества образования // Дистанционное и виртуальное образование. - Москва, 2014. - № 5. - С. 18-25.
13. Хлопотов М.В., Шишкин А.Р. Система диагностики сформированности профессиональных компетенций студентов специальности "Информационные технологии в образовании" // Дистанционное и виртуальное обучение. – Москва, 2010. - № 2. - С. 59-69.
14. Almond, R. G. Modeling diagnostic assessments with Bayesian networks / R. G. Almond, L. V. DiBello, B. Moulder, J. D. Zapata-Rivera // Journal of Educational Measurement. — 2007. — N. 44. — P. 341–359.
15. Conati, C., Gertner, A., & Vanlehn, K. (2002). Using Bayesian networks to manage uncertainty in student modeling. User Modeling and User-Adapted Interaction, 12(4), 371–417.
16. Desmarais, M. C., & Baker, R. S. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. User Modeling and User-Adapted Interaction, 22(1–2), 9–38.

17. Desmarais, M. C. Bayesian student models based on item to item knowledge structures / M. C. Desmarais, M. Gagnon // Innovative approaches for learning and knowledge sharing. — Springer, 2006. — P. 111–124.
18. Jensen, F. V. Bayesian Networks and Decision Graphs / F. V. Jensen. — NY: Springer-Verlag, 2001. — 268 p.
19. Liu, C. A simulation-based experience in learning structures of Bayesian networks to represent how students learn composite concepts / C. Liu // International Journal of Artificial Intelligence in Education. — 2008. — V. 8, N. 3. — P. 237–285.
20. Mayo, M., & Mitrovic, A. (2001). Optimising ITS behavior with Bayesian networks and decision theory. International Journal of Artificial Intelligence in Education, 12, 124–153.
21. Millán E., Loboda, T., & Pérez-de-la-Cruz J. L. Bayesian networks for student model engineering. // Computers & Education, 55(4) – 2010. – pp. 1663–1683.
22. Millán E. A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation / E. Millán, J. L. Pérez-de-la-Cruz // User Modeling and User-Adapted Interaction. — 2002. — N. 12. — P. 281–330.
23. Murray W.R. Control for Intelligent Tutoring Systems: A Blackboard-based Dynamic Instructional Planner // Proceedings of the 4th International Conference on Artificial Intelligence and Education. Amsterdam, Netherlands. 1989. P. 150-168.
24. Neapolitan RE. Learning Bayesian Networks. - NJ: Prentice Hall, 2004. – 704 p.
25. Pearl, J. Probabilistic Reasoning in Expert Systems: Networks of Plausible Inference / J. Pearl. — San Francisco: Morgan Kaufmann, 1988. — 552 p.
26. Reye, J. Student Modelling based on Belief Networks / J. Reye // International Journal of Artificial Intelligence in Education. — 2004. — N. 14. — P. 1–33.
27. VanLehn, K. Student modeling from conventional test data: a Bayesian approach without priors / K. VanLehn, Z. Niu, S. Siler, A. Gertner // Proc. of 4th Int. Conf. ITS'96. — 1996. — P. 29–47.
28. Villano, M. (1992) Probabilistic student models: Bayesian belief networks and knowledge space theory. In C.Frasson, C. Gauthier, & G. McCalla (Eds.) Intelligent Tutoring Systems, Proceeding of the Second International Conference, ITS'92, Montreal, Canada (pp. 491-498). Berlin: Springer-Verlag.

Рецензент: Горлушкина Наталия Николаевна, к.т.н., доцент, заведующая кафедрой интеллектуальных технологий в гуманитарной сфере, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики».

Maksim Khlopotov

Saint Petersburg National Research University of Information Technologies, Mechanics and Optics
Russia, Saint Petersburg
khlopotov@niuitmo.ru

Bayesian network in student model engineering for competence level evaluation

Student model is a basic component of the intelligent tutoring systems. In general such model reflects the knowledge of learner or learning system about students. This knowledge includes facts about of subject knowledge and students' skills, competences levels. There are three types of student models: starting, reference and current. Starting model captures the characteristics of the student to the beginning of the learning process. The learning objective specifies a reference model obtained from the regulation documents. Such a reference model called normative. Thus the model received as a result of the analysis of employers needs can act as reference model too. The current model describes the properties of the learner at different stages of learning process at a predetermined frequency. The current model is designed to track the educational progress of a particular student. The mechanisms used to build the current model are diagnostics and testing.

Bayesian networks are a useful tool to describe complex processes and events with uncertainties. It is necessary to define structure of the graph and parameters of each node for building the Bayesian network. This information can be obtained directly from the data or from expert. The structure of the Bayesian network represents the structure of competences of the students. Such network is a tool for making judgments about the level of students' readiness for professional activity. This article describes the process of building a Bayesian network for modeling students. The main steps of the process: identifying variables, defining structure, specifying parameters. Once a Bayesian network has been constructed, it is ready for inference. After some evidence are introduced, posterior probabilities are computed. Those probabilities can then be used for diagnostic reasoning (to determine the most likely causes of a certain event) or predictive reasoning (to predict the results of a test).

The main purpose of the competency modeling is to build a Bayesian network in which all composite competence are the parents to competencies, and every elementary competence is a parent for several competence-oriented tasks. In this regard, we note the properties of the relationship between competencies that are important for the building competence model. These properties include: hierarchy, equivalence, aggregation, community and causality. The article describes the general scheme of working with a list of competencies, which are in the educational standards. Some examples of the evaluation of the competence level are introduced.

Keywords: student modeling; educational data mining; Bayesian network; competence evaluation; competence modeling; probabilistic reasoning; R programming language.

REFERENCES

1. Atanov G.A., Pustynnikova I.N. Obuchenie i iskusstvennyj intellekt, ili Osnovy sovremennoj didaktiki vysshej shkoly. – Doneck : Izd-vo DOU, 2002. – 504 s.
2. Bidjuk P.I., Terent'ev A.N. Postroenie i metody obucheniya bajesovskih setej // Tavrijs'kij visnik informatiki i matematiki, №2/2004 – s. 139-154.
3. Brusilovskij P.L. Intellektual'nye obuchajushhie sistemy//Informatika. Informacionnye tehnologii. Sredstva i sistemy. 1990. №2. S.3-22.
4. Bul' E.E. Obzor modelej studenta dlja komp'juternyh sistem obucheniya // Obrazovatel'nye tehnologii i obshhestvo (Educational Technology & Society). 2003. Tom. 6, № 4. S. 245-250. Rezhim dostupa: http://ifets.ieee.org/russian/depository/v6_i4/html/G.html (data obrashhenija 01.07.2014)
5. Efremova N.F. Kompetencii v obrazovanii. Formirovanie i ocenivanie. - M.: Nacional'noe obrazovanie, 2012. - 416 s.
6. Koljada M.G. Vidy modelej, obuchaemyh v avtomatizirovannyh obuchajushhih sistemah // Iskusstvennyj intellekt. 2008. № 2. Rezhim dostupa: http://www.nbu.gov.ua/portal/natural/ii/2008_2/JournalAI_2008_2/Razdel2/00_Kolyada.pdf
7. Lisicyna L.S. Metodologija proektirovanija modul'nyh kompetentnostno-orientirovannyh obrazovatel'nyh programm. – SPb: SPbGU ITMO, 2009. – 50 s.
8. Petrushin V.A. Jekspertno-obuchajushhie sistemy. – K.: Naukova dumka, 1992. – 196 s.
9. Rybina G.V. Obuchajushhie integrirovannye jekspertnye sistemy: nekotorye itogi i perspektivy // Iskusstvennyj intellekt i prinjatие reshenij, 1/2008. – str. 22-56
10. Tulup'ev, A. L. Bajesovskie seti: logiko-verojatnostnyj podhod / A. L. Tulup'ev, S. I. Nikolenko, A. V. Sirotkin. — SPb.: Nauka, 2006. — 607 s. 95
11. Tulup'ev, A. L. Bajesovskie seti doverija: logiko-verojatnostnyj vyvod v aciklicheskih napravlennyh grafah / A. L. Tulup'ev, A. V. Sirotkin, S. I. Nikolenko. — SPb.: Izd-vo S.-Peterb. un-ta, 2009. — 400 s.
12. Hlopotov M.V., Kocjuba I.Ju. Metody intellektual'nogo analiza dannyh dlja monitoringa i diagnostiki kachestva obrazovanija // Distancionnoe i virtual'noe obrazovanie. - Moskva, 2014. - № 5. - S. 18-25.
13. Hlopotov M.V., Shishkin A.R. Sistema diagnostiki sformirovannosti professional'nyh kompetencij studentov special'nosti "Informacionnye tehnologii v obrazovanii" // Distancionnoe i virtual'noe obuchenie. – Moskva, 2010. - № 2. - S. 59-69.
14. Almond, R. G. Modeling diagnostic assessments with Bayesian networks / R. G. Almond, L. V. DiBello, B. Moulder, J. D. Zapata-Rivera // Journal of Educational Measurement. — 2007. — N. 44. — P. 341–359.
15. Conati, C., Gertner, A., & Vanlehn, K. (2002). Using Bayesian networks to manage uncertainty in student modeling. User Modeling and User-Adapted Interaction, 12(4), 371–417.
16. Desmarais, M. C., & Baker, R. S. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. User Modeling and User-Adapted Interaction, 22(1–2), 9–38.
17. Desmarais, M. C. Bayesian student models based on item to item knowledge structures / M. C. Desmarais, M. Gagnon // Innovative approaches for learning and knowledge sharing. — Springer, 2006. — P. 111–124.
18. Jensen, F. V. Bayesian Networks and Decision Graphs / F. V. Jensen. — NY: Springer-Verlag, 2001. — 268 p.

19. Liu, C. A simulation-based experience in learning structures of Bayesian networks to represent how students learn composite concepts / C. Liu // *International Journal of Artificial Intelligence in Education*. — 2008. — V. 8, N. 3. — P. 237–285.
20. Mayo, M., & Mitrovic, A. (2001). Optimising ITS behavior with Bayesian networks and decision theory. *International Journal of Artificial Intelligence in Education*, 12, 124–153.
21. Millán E., Loboda, T., & Pérez-de-la-Cruz J. L. Bayesian networks for student model engineering. // *Computers & Education*, 55(4) – 2010. – pp. 1663–1683.
22. Millán E. A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation / E. Millán, J. L. Pérez-de-la-Cruz // *User Modeling and User-Adapted Interaction*. — 2002. — N. 12. — P. 281–330.
23. Murray W.R. Control for Intelligent Tutoring Systems: A Blackboard-based Dynamic Instructional Planner // *Proceedings of the 4th International Conference on Artificial Intelligence and Education*. Amsterdam, Netherlands. 1989. P. 150-168.
24. Neapolitan RE. *Learning Bayesian Networks*. - NJ: Prentice Hall, 2004. – 704 p.
25. Pearl, J. *Probabilistic Reasoning in Expert Systems: Networks of Plausible Inference* / J. Pearl. — San Francisco: Morgan Kaufmann, 1988. — 552 p.
26. Reye, J. Student Modelling based on Belief Networks / J. Reye // *International Journal of Artificial Intelligence in Education*. — 2004. — N. 14. — P. 1–33.
27. VanLehn, K. Student modeling from conventional test data: a Bayesian approach without priors / K. VanLehn, Z. Niu, S. Siler, A. Gertner // *Proc. of 4th Int. Conf. ITS'96*. — 1996. — P. 29–47.
28. Villano, M. (1992) Probabilistic student models: Bayesian belief networks and knowledge space theory. In C.Frasson, C. Gauthier, & G. McCalla (Eds.) *Intelligent Tutoring Systems, Proceeding of the Second International Conference, ITS'92*, Montreal, Canada (pp. 491-498). Berlin: Springer-Verlag.