



OSTIS-2013

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

КОМПОНЕНТНОЕ ПРОЕКТИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ ОБУЧАЮЩЕЙ СИСТЕМЫ ДЛЯ ПОДГОТОВКИ ШКОЛЬНИКОВ К ЦЕНТРАЛИЗОВАННОМУ ТЕСТИРОВАНИЮ ПО ИНОСТРАННЫМ ЯЗЫКАМ

Елисеева О.Е., Русецкий К.В.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

olae@open.by

rusetski.k@gmail.com

В работе представлено компонентное проектирование интеллектуальной обучающей системы для подготовки к централизованному тестированию по иностранному языку. Выделенные компоненты ИОС позволяют выполнять разработку отдельных ее частей почти автономно с участием специалистов различного профиля.

Ключевые слова: централизованное тестирование; интеллектуальная обучающая система; ИОС; база знаний.

ВВЕДЕНИЕ

Основной формой контроля знаний абитуриентов при поступлении в высшие и средние специальные учебные заведения Республики Беларусь в настоящее время является централизованное тестирование (ЦТ). Чтобы успешно сдать ЦТ, абитуриентам необходимо владеть комплексом знаний по соответствующей дисциплине в объеме средней школы.

Для поступления на популярные сегодня специальности экономического профиля, специальность «Международные отношения», а также ряд гуманитарных специальностей, требуется сдавать ЦТ по иностранному языку. По данным Республиканского института контроля знаний, число сдававших тесты по иностранному языку в 2012 году составило более 34 тысяч человек. Это свидетельствует о том, что общее количество школьников, готовящихся к централизованному тестированию по иностранному языку достаточно велико. Многие из них прибегают к помощи частных репетиторов, некоторые готовятся к испытаниям самостоятельно. Однако и тем, и другим необходимы постоянные занятия по выполнению тестов, подобных тем, которые они будут проходить в процессе централизованного тестирования. Очевидно, что процесс такой «тренировки» предполагает наличие достаточно большого количества разнообразных упражнений-тестов, а также быструю обратную связь, т.е. информацию о правильности/неправильности прохождения теста. Такая обратная связь в

большинстве случаев представлена в виде заранее заготовленного списка ответов к тестам, сверившись с которыми учащиеся узнают, правильно ли они выполнили тест. Однако простой констатации факта об успешности/неуспешности прохождения теста недостаточно. Для качественной подготовки к тестированию абитуриенту крайне важно сообщить, какая допущена ошибка и какие правила необходимо повторить, чтобы ее исправить при прохождении нового теста. Именно эту задачу в большинстве своем решают преподаватели и репетиторы в процессе подготовки учащихся к ЦТ.

1. Цель работы и постановка задачи проектирования

Резюмируя сказанное выше, сформулируем две основные задачи, которые встают в процессе подготовки к ЦТ по иностранному языку:

- подготовка большого количества тестов;
- анализ и исправление допущенных в процессе тренировочного тестирования ошибок.

Решение обеих указанных задач можно автоматизировать путем использованием технологий искусственного интеллекта. Будем при этом учитывать, что подобного рода задачи вполне успешно решаются в рамках интеллектуальных обучающих систем (ИОС) [Chad Lane, 2006], [Рыбина, 2008]. В связи с этим **целью** данной работы ставится проектирование ИОС для подготовки к ЦТ. Так как указанная система является достаточно сложной и требует привлечения специалистов из различных

предметных областей (см. ниже), то становится очевидным, что при ее проектировании необходимо выделить компоненты, разработка которых возможна автономным образом. Фактически речь идет о выделении вполне самостоятельных компонентов, создание которых возможно с привлечением специалистов различных направлений и квалификации (различных компетенций), а также с учетом особенностей обработки информации различного типа и требуемых для этого методов и технологий.

2. Компетенции

Создание любой компьютерной системы обучения предполагает привлечение специалистов различного профиля. Когда речь идет о создании ИОС, следует учитывать, что ее разработка носит так называемый междисциплинарный характер.

Учитывая сказанное, в контексте данной работы выделим основные необходимые для проектирования и реализации рассматриваемой ИОС компетенции:

- глубокие знания иностранного языка (на лингвистическом уровне);
- владение методикой составления тестов различного вида для централизованного тестирования по иностранным языкам;
- знание технологии организации централизованного тестирования;
- владение современными технологиями программирования;
- владение технологиями представления данных и знаний;
- владение технологиями представления, хранения и обработки сложно структурированной информации;
- владение технологиями проектирования и реализации интерфейсов обучающих систем;
- владение технологиями защиты информации и организации авторизованного доступа к данным.

Каждая из перечисленных выше компетенций требует учета особенностей организации информации и знаний соответствующего типа. В процессе компонентного проектирования ИОС для подготовки к ЦТ будем ориентироваться на различные виды подлежащих формализации и обработке знаний, с одной стороны, и различные целевые группы пользователей – с другой.

3. Целевая аудитория (пользователи) ИОС для подготовки к ЦТ по иностранному языку

Основной целевой аудиторией (конечными пользователями) проектируемой ИОС являются школьники старших классов и абитуриенты, готовящиеся к прохождению испытаний на централизованном тестировании. Будем исходить из предположения, что обучаемые положительно

мотивированы к обучению. К другой категории пользователей ИОС можно отнести лиц старшего возраста, которые занимаются управлением и подготовкой к сдаче тестов. В большинстве случаев они будут принимать решение о целесообразности использования предлагаемого программного продукта в качестве инструмента для такой подготовки. Речь идет о родителях, преподавателях подготовительных курсов, репетиторах. Очевидно, что каждая из обозначенных целевых групп (категорий) пользователей будет использовать ИОС для достижения различных целей (рис.1).



Рисунок 1 – Цели пользователей ИОС по подготовке к ЦТ по иностранному языку

Из рисунка1 видно, что:

- Школьники и абитуриенты могут использовать ИОС для повторения, углубления и отработки знаний иностранного языка, приобретенных ими в школе, а также получения практических навыков решения тестовых заданий и заполнения тестовых бланков. При этом особое значение приобретает формирование навыков самоконтроля, что в рамках школы, как правило, не происходит в силу специфики организации учебного процесса, предполагающего постоянный контроль со стороны учителя.
- Преподаватели подготовительных курсов и репетиторы могут воспользоваться ИОС для организации учебной деятельности, а также в качестве источника многообразных тестовых упражнений, которые в данной системе генерируются в автоматизированном режиме.
- Родители школьников и абитуриентов с помощью ИОС получают возможность наблюдать за успеваемостью и общим уровнем подготовки своих детей к предстоящему тестированию.

4. Концепция и компоненты ИОС

Далее для наглядности в качестве примера будем рассматривать ИОС, рассчитанную на подготовку школьников к тестированию по конкретному иностранному языку. В качестве такового избран английский язык, так как он является наиболее распространенным. Создание в качестве прототипа ИОС по подготовке к ЦТ по одному языку позволит

выявить ее достоинства и недостатки, отладить методику ее практического использования, а также создать технологию тиражирования ИОС на случай других иностранных языков.

Исходя из выше сказанного, база знаний ИОС в идеале должна содержать информацию, соответствующую курсу английского языка в объеме средней школы, который разбивается на лексические и грамматические темы. При проектировании и реализации рассматриваемой ИОС будем учитывать возможность постепенного наращивания базы знаний различными темами. Так, для первой версии системы можно ограничиться несколькими грамматическими темами: «Времена глаголов», «Употребление артиклей», «Образование формы множественного числа существительных» и несколькими лексическими: «О себе», «Моя семья», «Город Минск».

Следуя общепринятой технологии проектирования прикладных интеллектуальных систем и ИОС, в частности, выделим в качестве основных компонентов проектируемой системы следующие (рис.2):

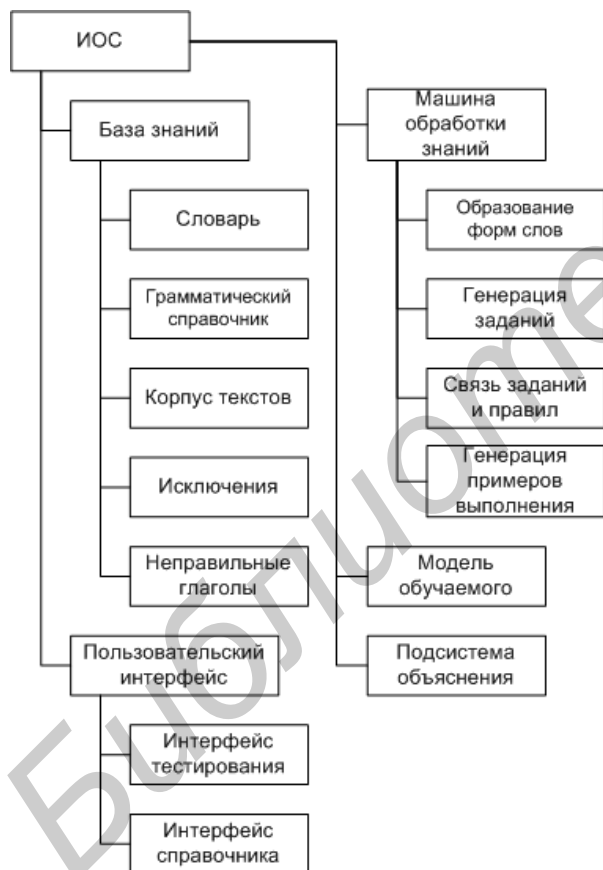


Рисунок 2 – Компоненты ИОС для подготовки к ЦТ по иностранному языку

- База знаний – информационно-содержательное ядро ИОС, основу которого составляет необходимый для подготовки к ЦТ учебный материал по иностранному языку;

- Машина обработки знаний – функциональное ядро ИОС, основной задачей которого является автоматизация создания

тренировочных тестов и реализация различных режимов тестирования знаний;

- Модель обучаемого – специально выделенный из базы знаний компонент, в рамках которого хранится и обрабатывается информация об учащихся и их достижениях;

- Подсистема объяснений – относительно автономный компонент машины обработки знаний, в задачи которого входит объяснение ошибок, допущенных учащимися в процессе выполнения заданий тестов;

Пользовательский интерфейс – стандартный компонент любой системы, но, с учетом функциональных особенностей рассматриваемой ИОС подразделяемый на два отличающихся друг от друга подвида интерфейса.

4.1. База знаний

Как указывалось выше, база знаний является информационно-содержательным ядром ИОС, включающем знания об изучаемой предметной области, а именно об иностранном языке. При проектировании базы знаний иностранного языка имеет смысл учитывать как общие лингвистические закономерности, которые берутся за основу изучения языка, так и частные особенности строения каждого конкретного языка (как лингвистического, так и методического характера). Так, например, вне зависимости от того, по какому языку будет вестись подготовка к ЦТ, в составе базы знаний ИОС выделены компоненты «Словарь», «Грамматический справочник», «Корпус текстов». С учетом особенностей, например, английского языка нами выделены также компоненты «Исключения» и «Неправильные глаголы», т.к. именно этим тематикам уделяется особое внимание в процессе подготовки к ЦТ по английскому языку.

Накопление достаточного словарного запаса является неотъемлемой частью изучения любого иностранного языка. Поэтому одним из важнейших компонентов ИОС является **Словарь**. В отличие от большинства существующих в настоящее время словарей иностранных языков, мы рассматриваем словарь не в «традиционном» смысле (как набор упорядоченных по алфавиту слов с транскрипцией и переводом), а семантически структурированный. Это означает, что в рассматриваемом словаре:

- каждому слову приписывается грамматическая информация (часть речи, ее признаки);

- каждое слово связано с другими словами различными семантическими отношениями (например, быть синонимом, антонимом, гипонимом, гиперонимом и пр.);

- словам ставятся в соответствие примеры их употребления в текстах, которые хранятся и систематизируются в компоненте «Корпус текстов».

Поясним вышесказанное на конкретном примере.

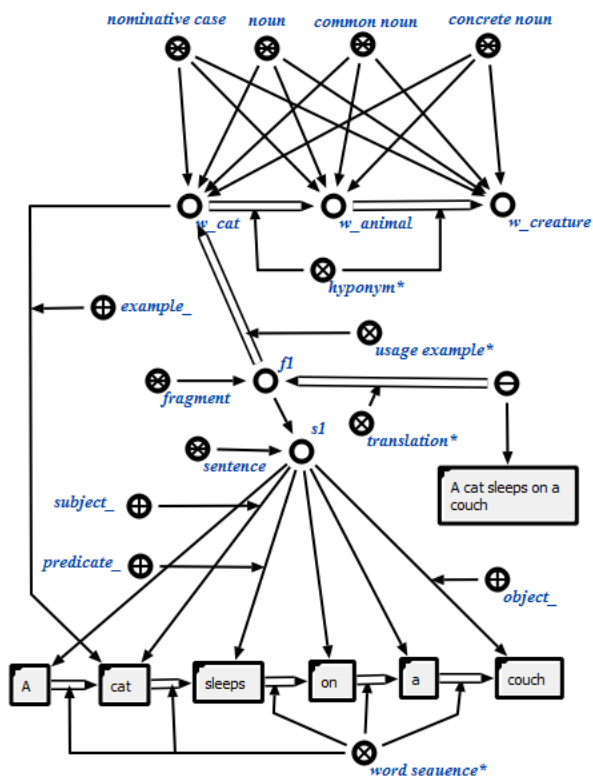


Рисунок 3 – Пример представления словарной статьи в памяти интеллектуальной обучающей системы на языке SCg.

На рисунке 3 предложено описание слова в семантически структурированном словаре, выполненное на языке SCg и отвечающее описанным выше требованиям. Подробно этот язык описан в работе [Голенков и др, 2001]. Язык SCg был выбран в силу того, что он позволяет одновременно наглядно и при этом формально корректно изображать лингвистические структуры в графической форме.

Произведем содержательную интерпретацию приведенной SCg-конструкции. Под словом будем понимать множество реализаций слова в различных предложениях. Заметим также, что к идентификатору знака слова добавляется префикс w_* , чтобы отличать знак слова в семантической сети от знака понятия, которое это слово обозначает. Слово «cat» (кошка, кот) является конкретным нарицательным существительным в именительном падеже. Это слово является гипонимом по отношению к слову «animal» (животное), которое, в свою очередь является гипонимом по отношению к слову «creature» (живое существо), что отражено с помощью бинарных ориентированных связей отношения *гипоним** (гипоним*). Слова «animal» и «creature» также являются конкретными нарицательными существительными в именительном падеже. Примером употребления слова «cat» является фрагмент текста $f1$, который состоит из предложения $s1$, которое записывается на естественном языке как «Cat sleeps on a couch» (Кот спит на диване). В предложении $s1$ входят слова «cat», «sleeps», «on», «a», «couch», причем слово «A» предшествует слову «cat», которое

предшествует слову «sleeps», и так далее. Слово «cat» в предложении $s1$ является подлежащим; слово «sleeps» в предложении $s1$ является сказуемым; а слово «couch» в этом предложении является дополнением, что обозначается соответствующими атрибутами.

Особенностью представления **Грамматического справочника** ИОС для подготовки к ЦТ по иностранному языку является то, что информация в нем рассчитана как на использование (изучение) учащимися, так и на реализацию операций автоматического создания заданий тестов. Основу справочника составляет набор грамматических правил соответствующего языка, которые представлены в двух видах: 1) формализованных правил образования словоформ, словосочетаний, предложений; 2) учебных естественно-языковых текстов объяснений для учащихся указанных правил.

Для создания разнообразных заданий и тестов необходимы источники иноязычных текстов [Централизованное, 2012]. Практически все тесты создаются на основе оригинальных текстов. Именно поэтому в рассматриваемой ИОС специально выделен компонент «**Корпус текстов**», который планируется использовать с двумя целями: а) для автоматизированного создания заданий тестов; б) для обеспечения связей слов из Словаря с их словоупотреблениями в текстах (пример такой связи приведен выше на рисунке 3). В связи с этим рассматриваемый корпус размечается с учетом того, какие тестовые задания могут в дальнейшем создаваться на их основе.

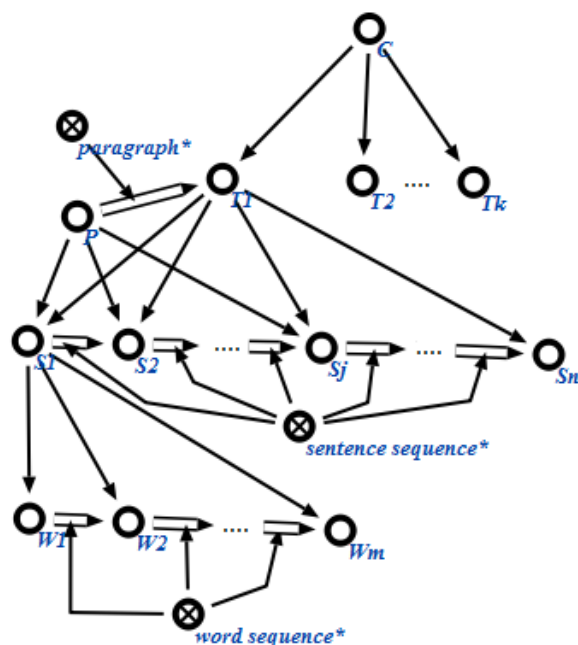


Рисунок 4 – Схематичное представление корпуса текстов в памяти интеллектуальной обучающей системы

Дадим интерпретацию структуре корпуса, схематично изображенной на рисунке 4. Корпус текстов C представляет собой неориентированное множество текстов:

$$C = \{T_1, T_2, \dots, T_k\} \quad (1)$$

Текст T_1 представляет собой ориентированное множество предложений:

$$T_1 = \langle S_1, S_2, \dots, S_j, \dots, S_n \rangle \quad (2)$$

Предложение S_1 представляет собой последовательность слов

$$S_1 = \langle W_1, W_2, \dots, W_m \rangle \quad (3)$$

Отметим, что каждое из слов имеет описание, подобное приведенному на рисунке 3.

Подмножество предложений текста, выделяемое на письме красной строкой, будем называть абзацем (на рисунке 4 – множество P)

$$P = \langle S_1, S_2, \dots, S_j \rangle \quad (4)$$

Принадлежность знаков текстов, предложений и слов соответствующим множествам не приводится, чтобы не загромождать иллюстрацию. Разделители (пробелы, знаки препинания) формализовывать не требуется, так как те части текста, которые они отделяют, можно формализовать как подмножества соответствующих структур. Например, простая часть сложносочиненного предложения будет являться подмножеством предложения. Проиллюстрируем данный факт на примере предложения «It rained, so I took an umbrella» (Шел дождь, поэтому я взял зонтик).

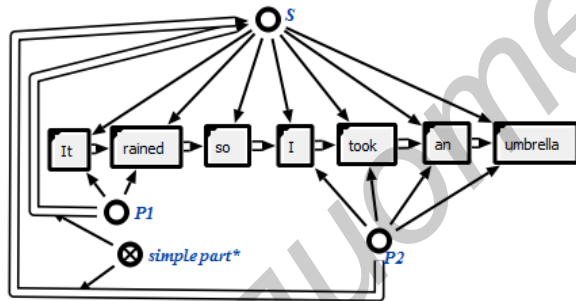


Рисунок 5 – Пример выделения простых частей в сложносочиненном предложении.

Приведенное на рисунке 5 предложение состоит из двух простых частей, границы которых определяются подмножествами P1 и P2 соответственно. Значит, формализовывать запятую, которая разделяет части предложения на письме, нет необходимости.

Учитывая особенности английского языка, в базе знаний выделены компоненты «Исключения» и «Неправильные глаголы», в рамках которых предусматривается хранение и формальное описание всевозможных исключений образования, например, форм множественного числа некоторых слов, а также временных форм неправильных глаголов. Все формы слова, которое является исключением из правил формообразования, представляются в базе знаний в явном виде, как на

рисунке 6.

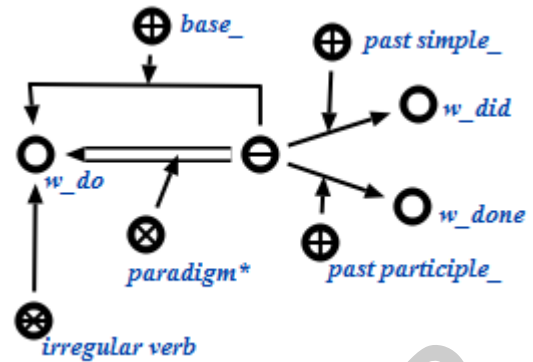


Рисунок 6 – Представление информации о неправильном глаголе в памяти системы.

На приведенном выше рисунке неправильный глагол «do» связан со множеством своих форм отношением *paradigm** (парадигма*). Вхождения форм в множество словоформ помечаются соответствующими атрибутами *base_*, *past simple_*, *past participle_* (инфинитив_, прошедшее время_, причастие прошедшего времени_).

Как указывалось выше, относительно автономным элементом базы знаний является **модель обучаемого**, в рамках которой сохраняется информация об индивидуальных особенностях учащихся. Частично модель обучаемого представляет собой профиль пользователя, подобный тем, которые создаются в процессе регистрации учетных записей пользователей в любой информационной системе. В отличие от «традиционных» систем, модель обучаемого рассматриваемой ИОС предусматривает хранение протокола работы с системой и фиксацию допускаемых при выполнении заданий тестов ошибок, с целью дальнейшего их исправления.

Кроме того, база знаний ИОС для подготовки к ЦТ должна содержать знания о тестировании – структуру тестов, типы и структуру тестовых заданий, и т.п.

4.2. Машина обработки знаний

Главной задачей машины обработки знаний ИОС для подготовки к ЦТ по иностранному языку является автоматизация рутинных функций по созданию заданий тестов на основе грамматических правил, словаря и корпуса текстов из базы знаний. Среди механизмов (операций), обеспечивающих указанную автоматизацию, выделим следующие наборы:

- образование словоформ и их поиск в текстах по заданным признакам;
- генерация тестовых заданий различного типа на основе оригинальных текстов;
- обеспечение связей сгенерированных заданий и использованных для этого грамматических правил;

- генерация примеров выполнения заданий тестов для проверки правильности выполнения тестов учащимися и обеспечения режима обучения.

4.3. Подсистема объяснения

Как указывалось выше, подсистема объяснения является частью машины обработки знаний. В рамках подсистемы объяснения обеспечиваются следующие режимы работы с ИОС:

- переход к различным разделам грамматического справочника из заданий тестов, а также свободная навигация по нему;
- поиск и навигация по словарю;
- тестирования в режиме обучения (тренировки) с пояснениями и отсылками к правилам, без ограничений по времени и количеству ошибок;
- реализация экзаменационного тестирования, без объяснений, ограниченный по времени (120 минут).

4.4. Пользовательский интерфейс ИОС

Пользовательский интерфейс предназначен для организации эффективного диалога пользователя и обучающей системы в процессе достижения целей обучения. Исходя из этого, разработка интерфейса обучающей системы должна начинаться с разработки примеров диалога. Образцы диалогов являются отправной точкой для выделения основных понятий и отношений, а также являются основой этапа тестирования и отладки системы.

Важнейшим элементом диалога является цель общения. Цель пользователя - изучение некоторого иностранного языка в объеме, достаточном, чтобы успешно сдать ЦТ. Цель системы – осуществить подготовку пользователя, обеспечивающую это.

Наилучшим вариантом интерфейса для обучающей системы является естественно-языковой интерфейс (ЕЯИ), так как такой интерфейс наиболее адекватен для неопытного пользователя. Он позволяет ему эффективно взаимодействовать с системой, несмотря на невысокий уровень подготовки и с максимальной быстротой формулировать произвольные запросы к системе. Однако трудоемкость разработки ЕЯИ довольно велика, и для первой версии разрабатывать такой интерфейс нецелесообразно. Вместо этого для начала предлагается использовать в качестве языка общения уже упомянутый ранее язык SCg. Однако не стоит упускать из виду и ЕЯИ, поэтому приведем фрагмент диалога между пользователем (П) и системой (С) параллельно на естественном языке и SCg.

П: Как подразделяются предложения по сложности?

С: По сложности предложения подразделяются на простые, сложные и составные.

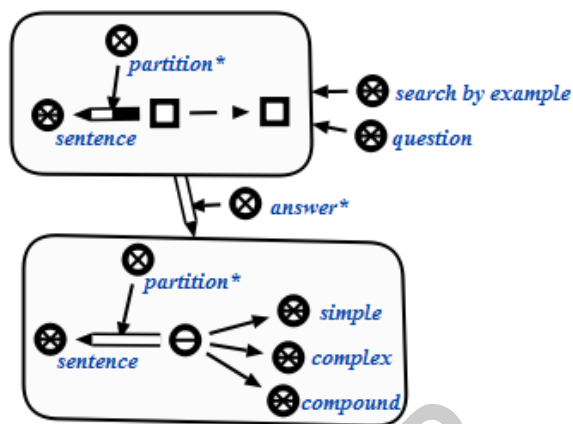


Рисунок 7 – Фрагмент диалога на SCg

Заключение

В завершение данной работы отметим, что выделенные в процессе проектирования компоненты ИОС для подготовки к ЦТ по иностранному языку требуют дальнейшего детального рассмотрения с привлечением различных компетенций. Хотя реализация ЕЯИ и не предполагается в первой версии системы, но лингвистические компоненты, не касающиеся непосредственно ЦТ, лягут в основу разработки ЕЯИ в будущих версиях.

Библиографический список

- [Chad Lane, 2006] Chad Lane H. Intelligent Tutoring Systems: Prospects for Guided Practice and Efficient Learning. 2006.
- [Голенков и др, 2001] Голенков, В.В. Представление и обработка знаний в графодинамических ассоциативных машинах /В. В. Голенков [и др.] – Мн. : БГУИР, 2001.
- [Рыбина, 2008] Рыбина Г.В. Обучающие интегрированные экспертные системы: некоторые итоги и перспективы // Искусственный интеллект и принятие решений 1/2008. – С.22-46.
- [Централизованное, 2012] Централизованное тестирование. Английский язык : сборник тестов / Респ. ин-т контроля знания М-ва образования Респ. Беларусь. – Минск: Аверсэв, 2012.

COMPONENT DESIGN OF INTELLIGENT TUTORING SYSTEM TO PREPARE STUDENTS FOR CENTRALIZED TESTING IN A FOREIGN LANGUAGE

Yeliseyeva O.E., Rusetski K.V. *

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

olae@open.by
rusetski.k@gmail.com

In the process of designing an intelligent tutoring system to prepare for centralized testing in a foreign language, select the component, further design and development of which is possible in offline mode with specialists in various fields.