



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Моделирование процесса адаптации контента MOOC

Дмитрий Алдунин

Аспирант 2-ого года Аспирантской школы по компьютерным наукам

Научный руководитель: Мальцева С.В., д.т.н.



В рамках подготовки диссертации на соискание учёной степени кандидата технических наук по теме

«Математические модели, методы и программное обеспечение формирования адаптивного контента образовательных интернет-ресурсов»

Цель

Разработать модели ключевых элементов и подпроцессов процесса адаптации контента массовых открытых онлайн-курсов к индивидуальным особенностям обучающегося для дальнейшего использования в стохастическом имитационном моделировании процесса обучения значительного числа обучающихся для последующей разработки программной системы поддержки образовательного процесса.

- 1) Определить основные сущности процесса электронного образования и их связи;
- 2) Применить совместные процессный и онтологический подходы к моделированию образовательного процесса;
- 3) Использовать преимущества цифровой трансформации образовательного процесса;

Внешние данные

- **Семантические карты областей знания** в форме ориентированных графов, отражающих иерархию компетенций (т.н. семантические сети).

Задача выявления компетенций и определения их иерархии не рассматривается, так как существует ряд решений, таких как принятый Европейской Комиссией “The Framework for Qualifications of the European Higher Education Area” [1].

- **Набор курсов и их компетентностные характеристики.**

[1] European Commission. Learning Opportunities and Qualifications in Europe.
<http://ec.europa.eu/ploteus/en> Retrieved 01.04.17

Входные параметры

- **Семантическая карта знаний обучающегося**, основанная на данных о его или её прошлом опыте обучения и/или вступительных испытаниях. Такая карта является подграфом семантической сети соответствующей области знания. Вершины также содержат процентное значение уровня владения компетенцией.
- **Набор компетенций, определённых обучающимся как желаемые.**
- **Данные об индивидуальных характеристиках обучающегося** (образовательный стиль по бинарной классификации Фельдера-Сильверман, культурные измерения Хофстеде и т.п.), определённые при вступительном тестировании.



Высокоуровневое описание модели

Выходные параметры

- **Семантическая карта знаний обучающегося**, дополненная приобретёнными компетенциями. Степени владения компетенциями обновлены.
- **Оценка эффективности** процесса получения желаемых компетенций.

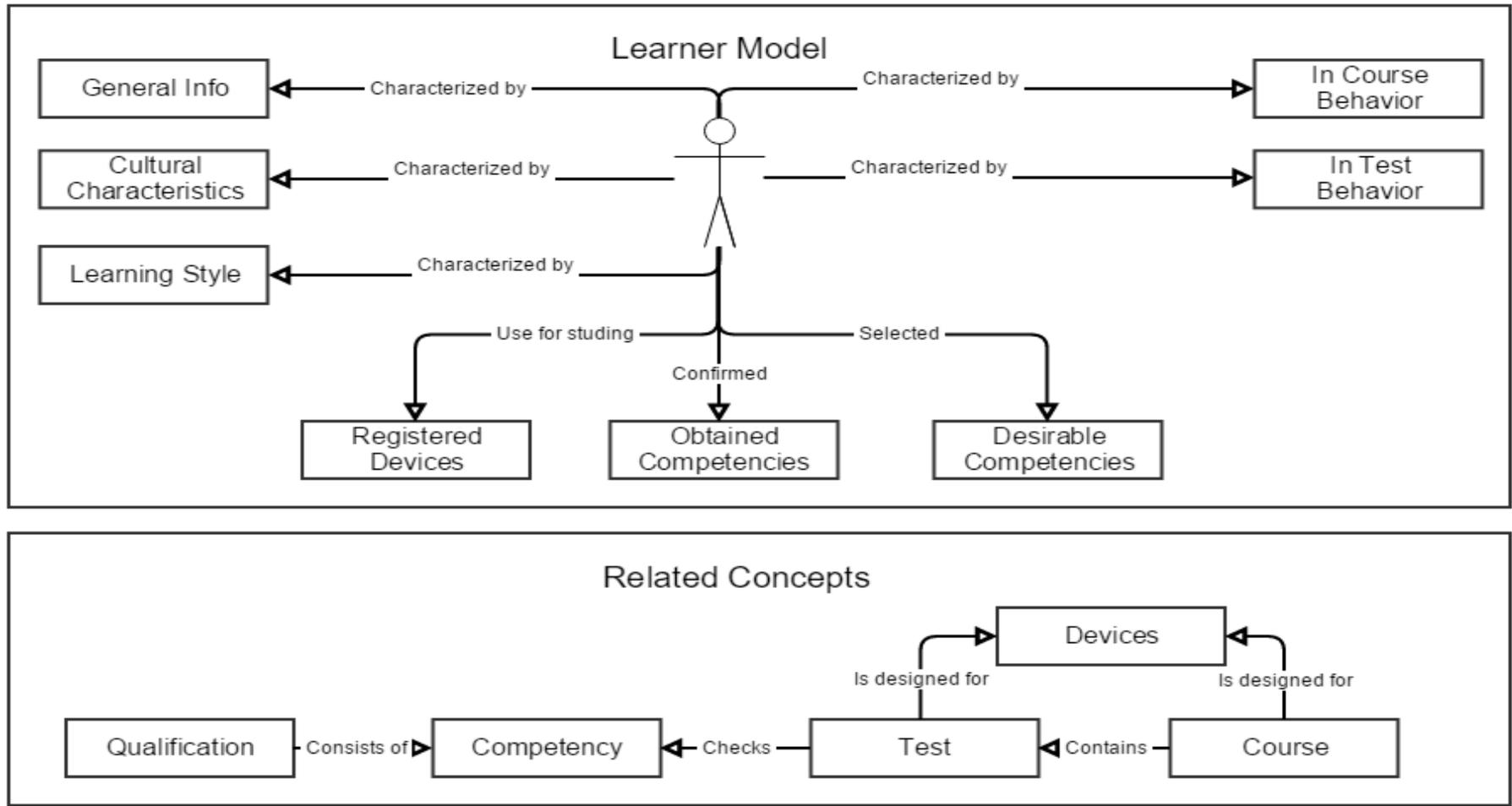


Рис. 1. Онтологическая модель обучающегося

Model Parts. Model of Learner

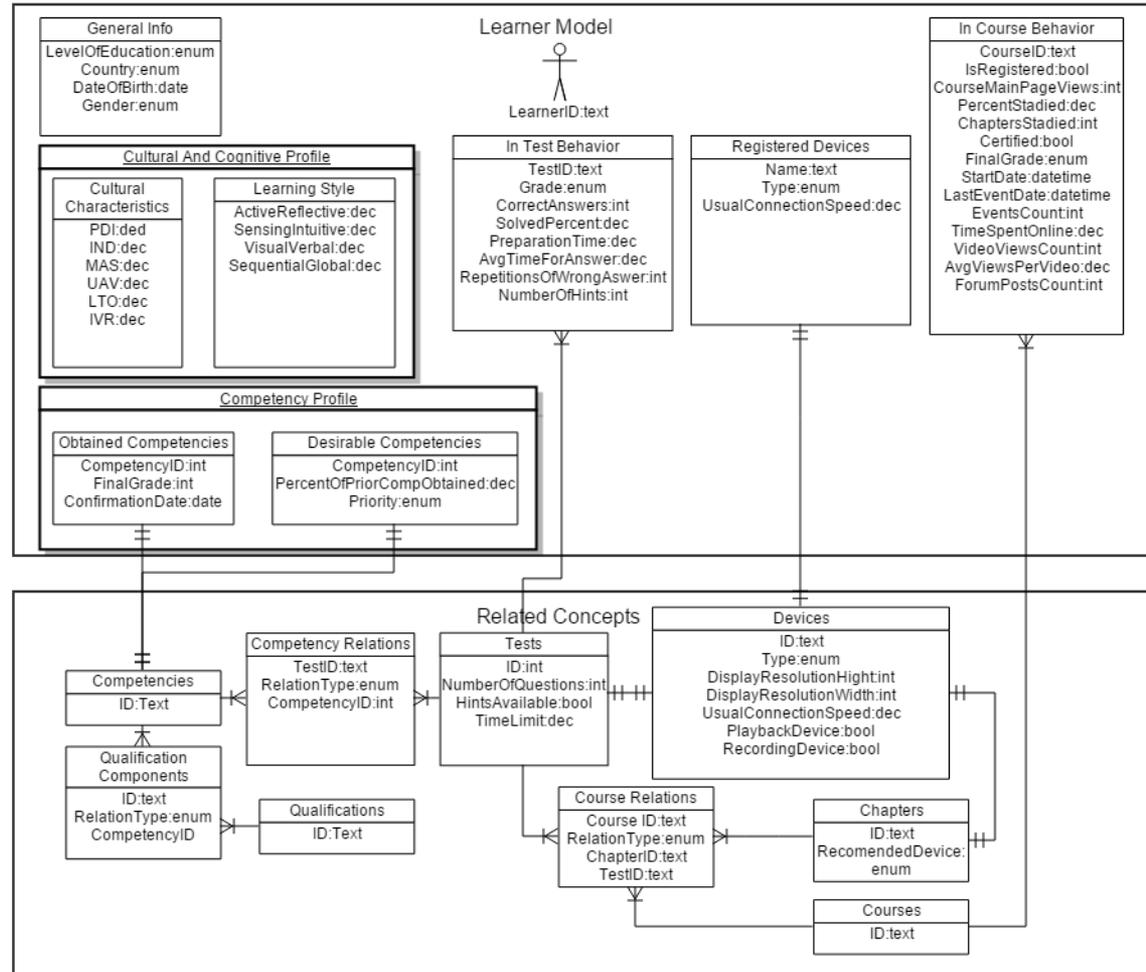


Рис. 2. Детализированная модель с атрибутами и связями

Части модели. Семантическая сеть области знания

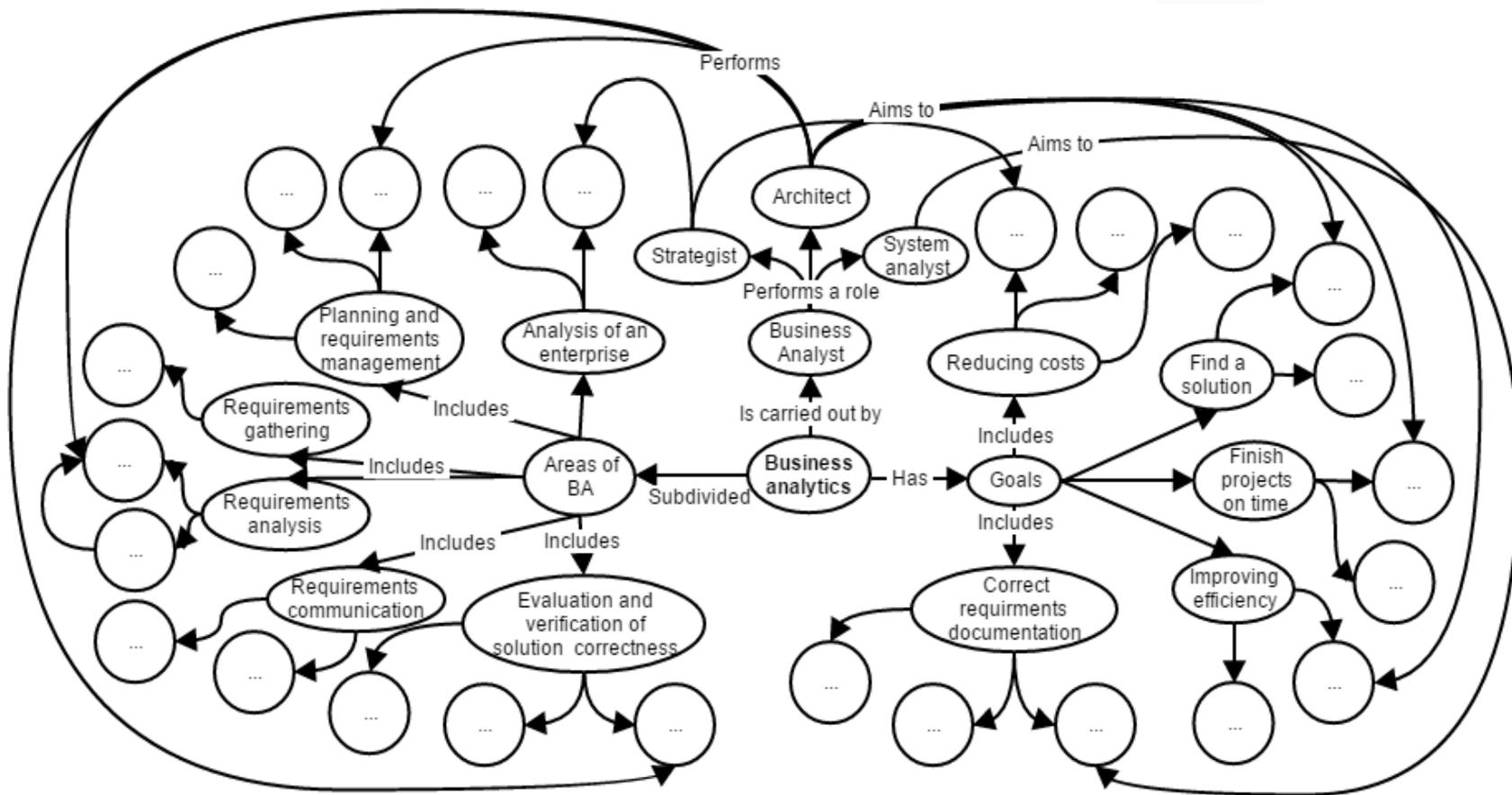


Рис. 3. Примерная семантическая сеть области знания «Бизнес-аналитика»

Параметры:

Область знаний D ;

Семантическая сеть области знаний $G^D = \{V^D, E^D\}$, где
 V^D – множество вершин (компетенций),
 E^D – множество связей графа;

Семантическая сеть – связный в сильном смысле ациклический орграф.

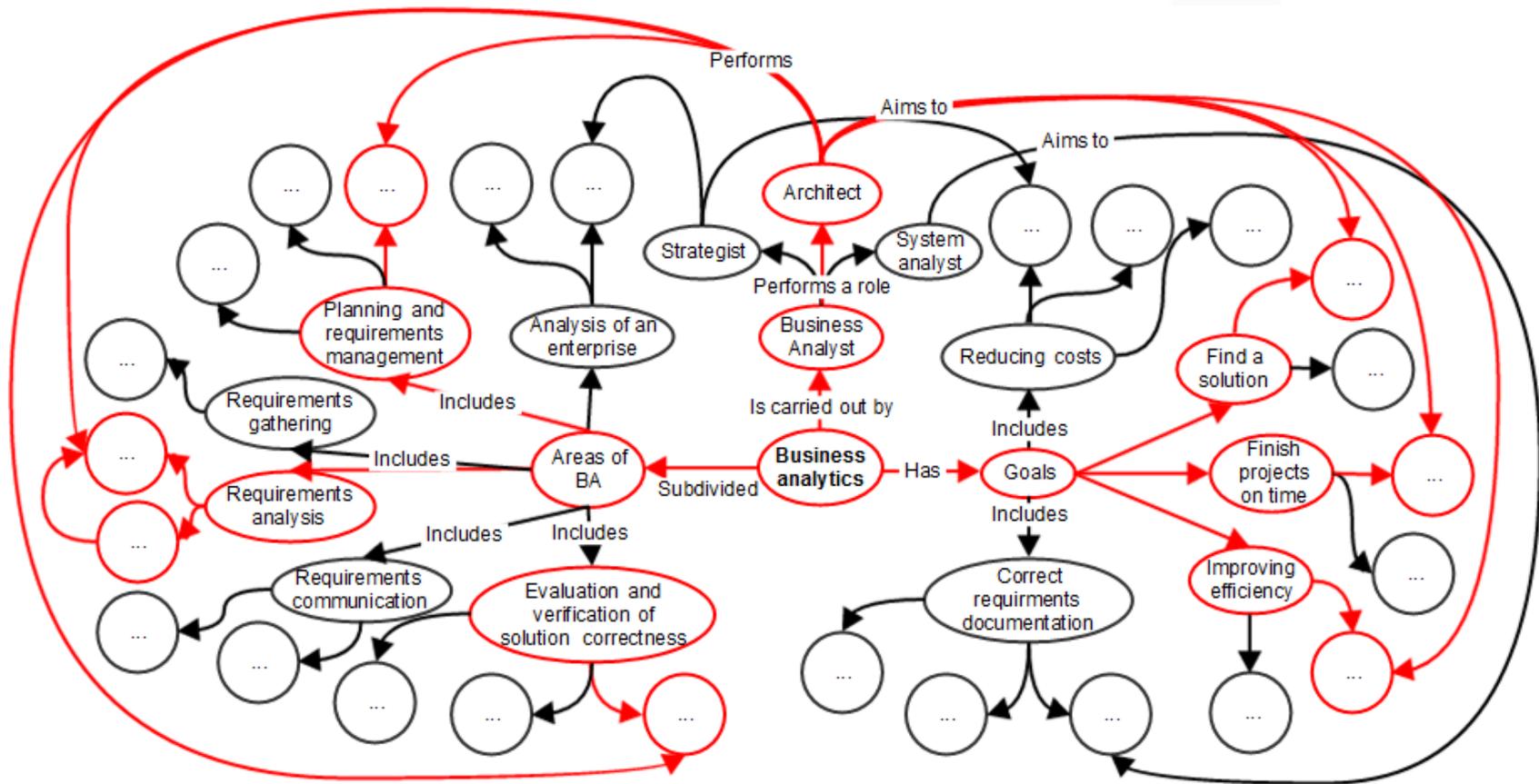


Рис. 4. Примерная семантическая сеть области знания «Бизнес-аналитика»

Параметры:

Обучающийся L ;

Семантическая сеть знаний обучающегося $G^L = \{V^L, E^L\}$, где

V^L – множество вершин (компетенций),

E^L – множество связей графа;

$V^L \subseteq V^D, E^L \subseteq E^D$;

Семантическая сеть знаний обучающегося – связный в сильном смысле ациклический орграф.

V_D^L – множество желаемых компетенций,

$V_D^L \subseteq V^D$;

Ограничения:

$|V_D^L \cap V^L| = 0$;

Допущение:

Граф G^L является связным подграфом графа G^D , т.е. в карте знаний обучающегося нет “пропусков”.

Части модели. Компетентностная модель курса

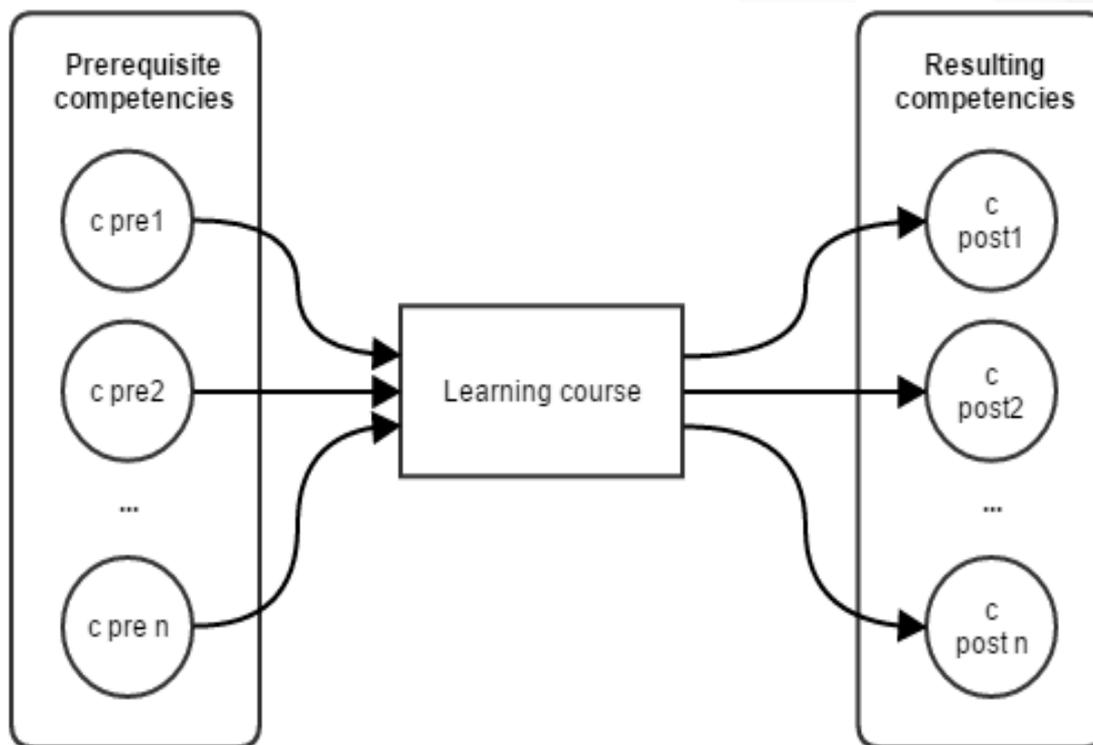


Рис. 5. Компетентностная модель курса

Параметры:

Множество курсов $C = \{C_i\}, i = 1..K;$

Множество компетенций-прerequisites $\{c_{pre\ i\ k}\}, k = 1, \dots, M;$

Множество компетенций-постреquisites $\{c_{post\ i\ k}\}, k = 1, \dots, N;$

Ограничения:

Мощность множества prerequisites курсов $|c_{pre\ i}| \geq 0;$

Мощность множества постerequisites курса $|c_{post\ i}| > 0;$

$|c_{pre\ i} \cap c_{post\ i}| = 0;$



Критерий подбора курсов

Контент курсов (образовательные материалы и проверочные задания), которые могут быть включены в индивидуальную образовательную траекторию, должен подходить для устройства, которое использует обучающийся.

Части модели. Модель индивидуальной образовательной траектории

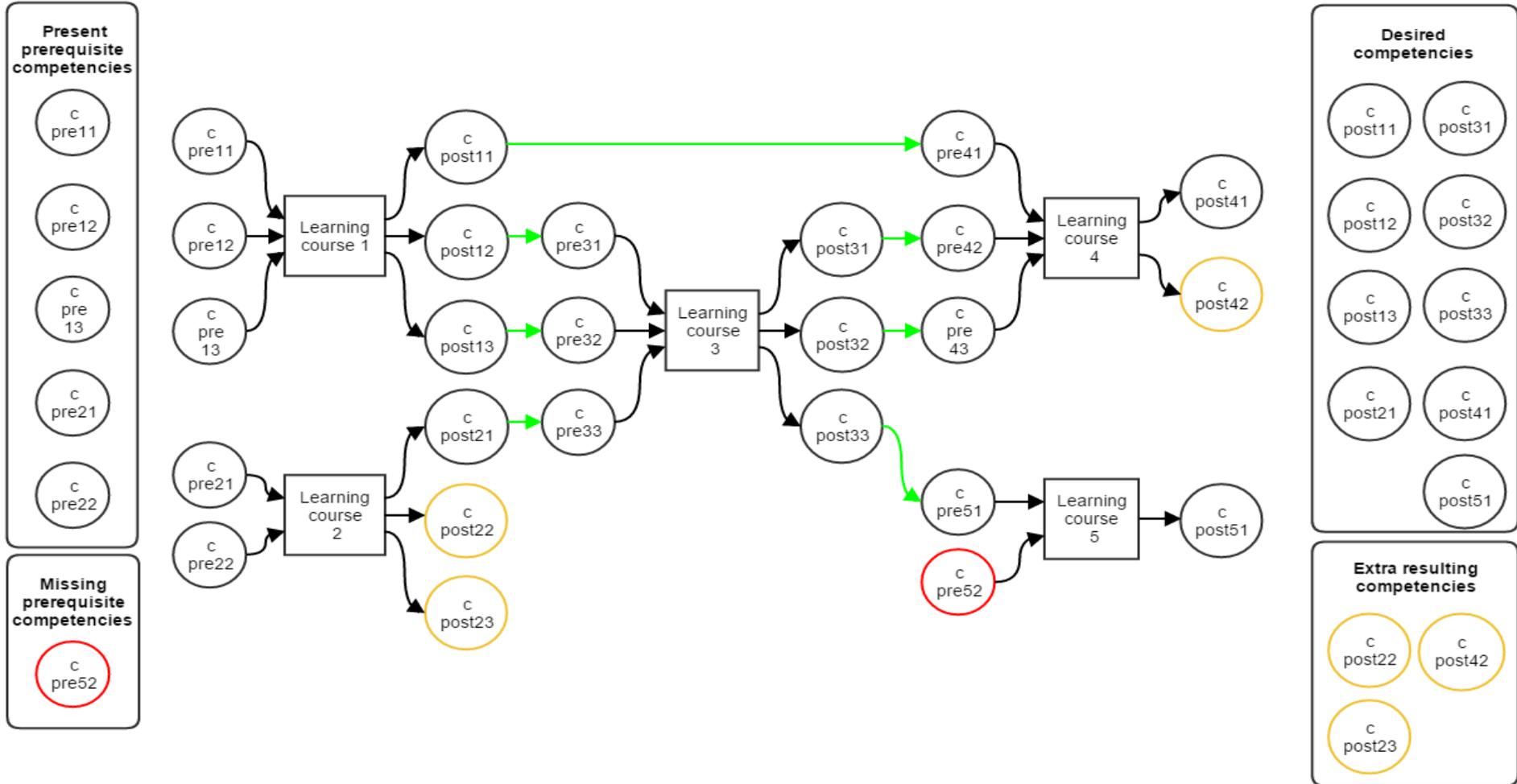


Рис. 6. Модель индивидуальной образовательной траектории

Задача многокритериальной оптимизации

Результат:

Связный орграф, представляющий траекторию $G^T = (C^T, E^T)$;

Ограничение:

Получение всех желаемых компетенций

$$\cup_{i \in C^T} c_{post\ i} \supseteq V_D^L;$$

Целевые функции:

➤ Максимизация взаимного использования компетенций курсами

$$\sum_{(C_i, C_j) \in E^T} |c_{pre\ i} \cap c_{post\ j}| \rightarrow \max;$$

➤ Минимизация использования недостающих компетенций

$$|\cup_{i \in C^T} c_{pre\ i} \setminus V^L| \rightarrow \min;$$

➤ Минимизация получения излишних компетенций

$$\begin{cases} c_{extra} \subset \cup_{i \in C^T} c_{post\ i}, \\ c_{extra} \cap V_D^L = \emptyset; \end{cases} \quad \max |c_{extra}| \rightarrow \min;$$

Для выбора из нескольких Парето-недоминируемых вариантов возможно учитывать исторические данные об успехах обучающихся со схожими характеристиками.

Части модели. Модель индивидуальной образовательной траектории

Задачи для дальнейшей работы:

- Определение приемлемого времени решения задачи
- Разработка алгоритма построения индивидуальной образовательной траектории

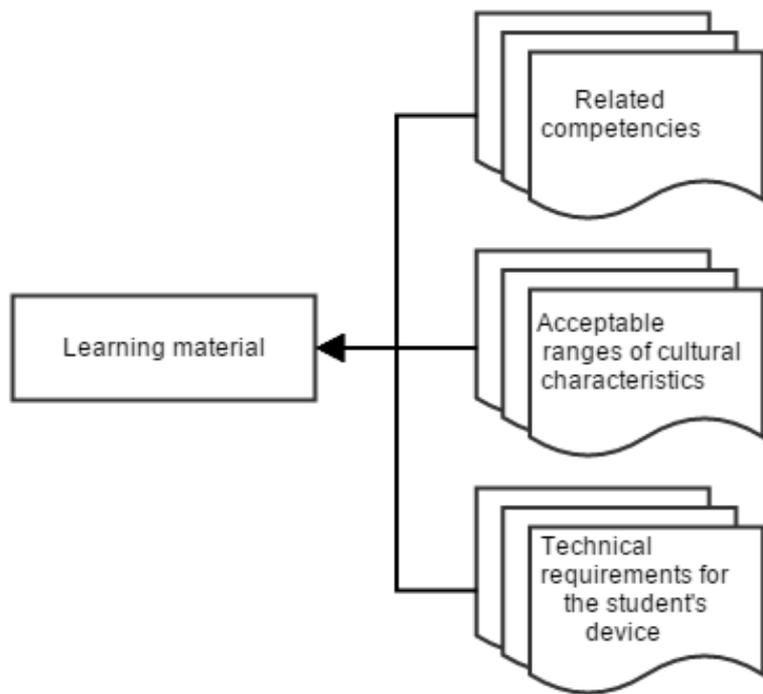


Рис. 7. Модель обучающего материала

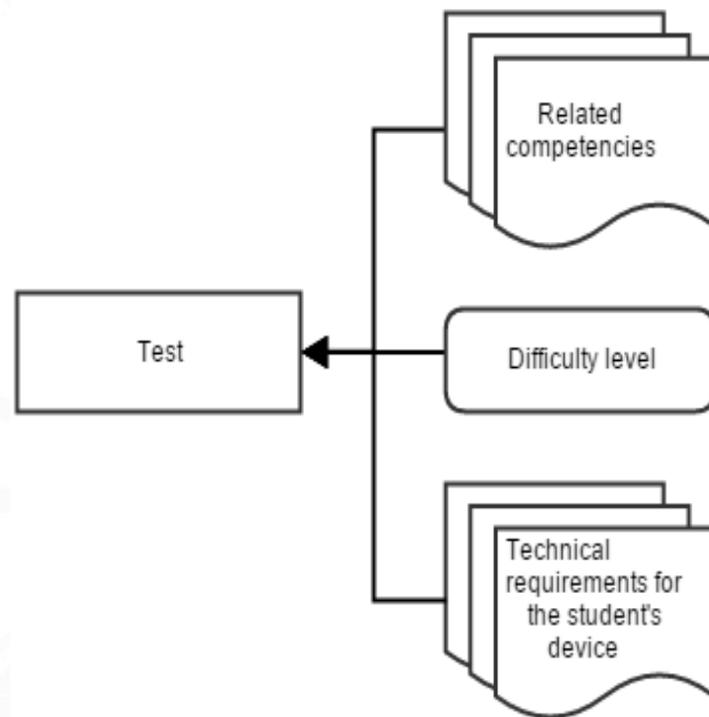


Рис. 8. Модель теста

Части модели. Модель приобретения компетенций

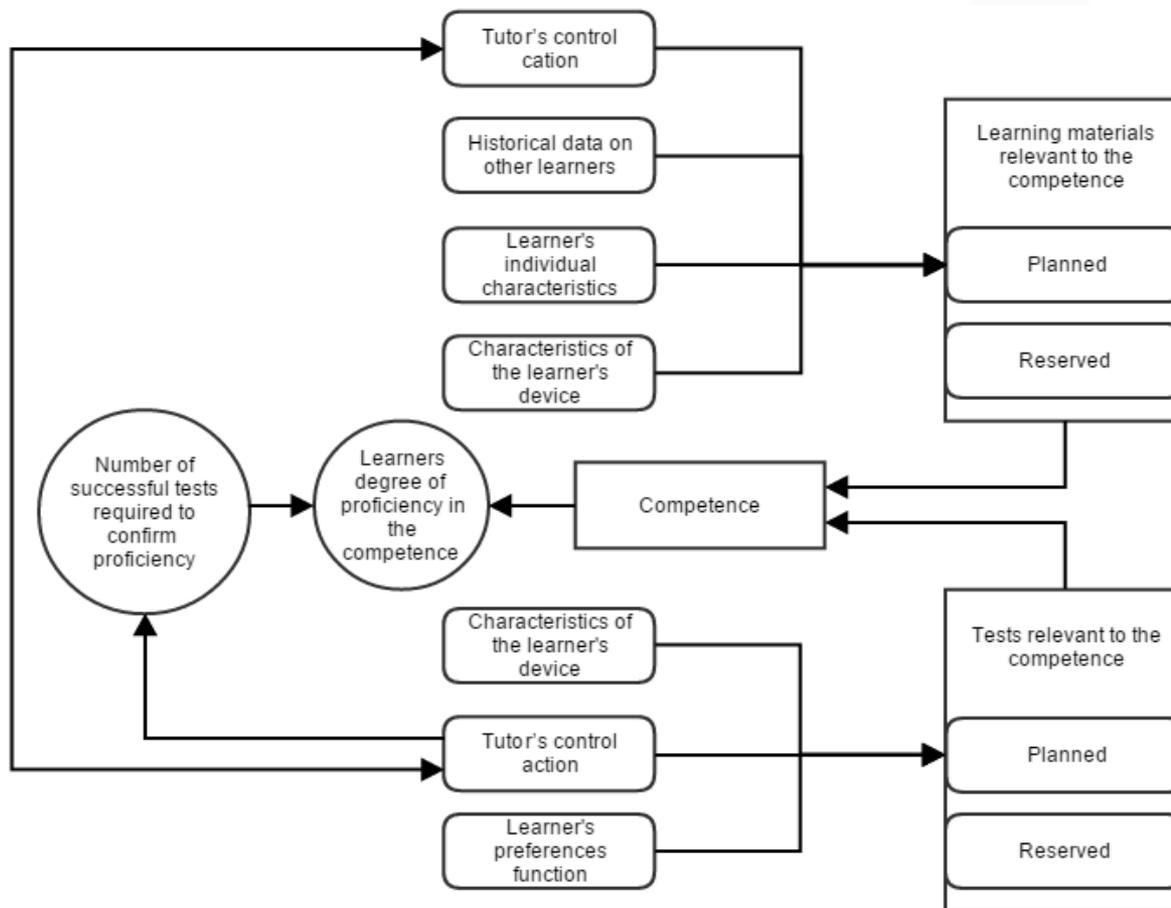


Рис. 9. Модель приобретения компетенций

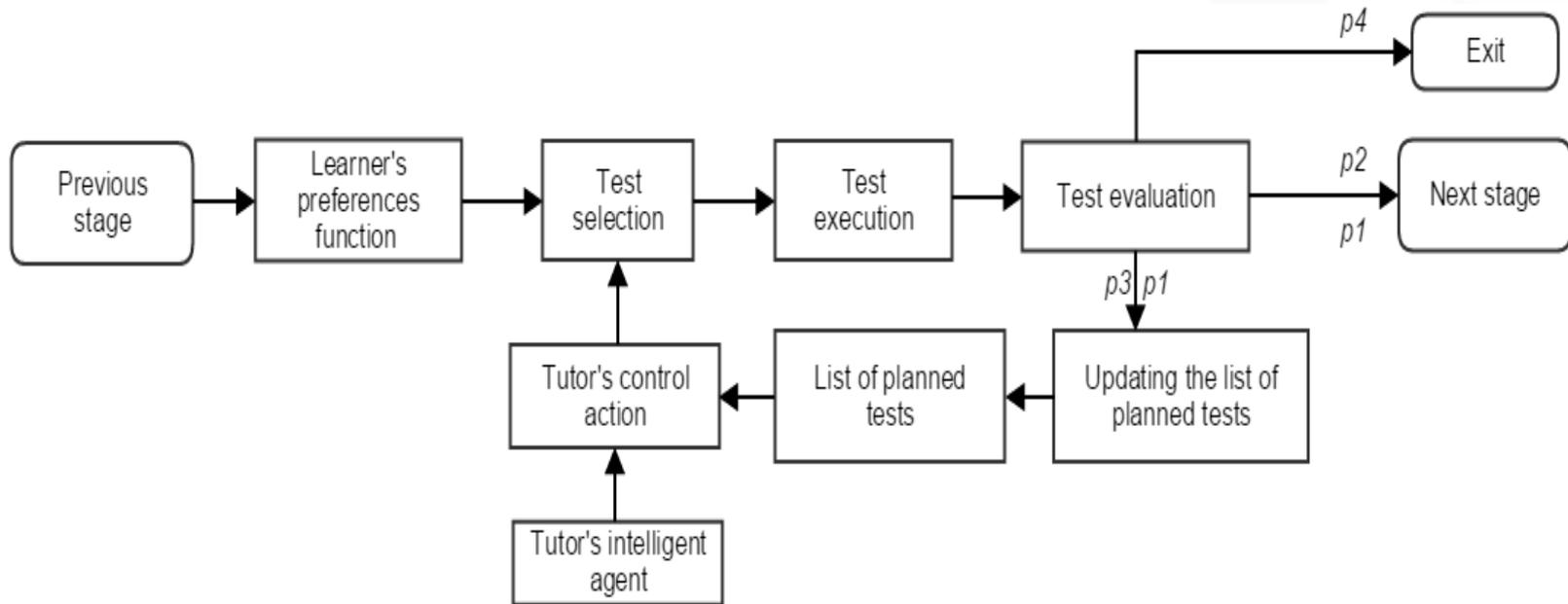


Рис. 10. Модель процесса тестирования владения компетенцией

p_1 – вероятность того, что обучающийся успешно справился с заданием и счёл его слишком простым;

p_2 – вероятность того, что обучающийся успешно справился заданием и счёл его сложность адекватной;

p_3 – вероятность того, что обучающийся был вынужден вернуться к выбору задания из-за сложности;

p_4 – вероятность того, что обучающийся покинул процесс;

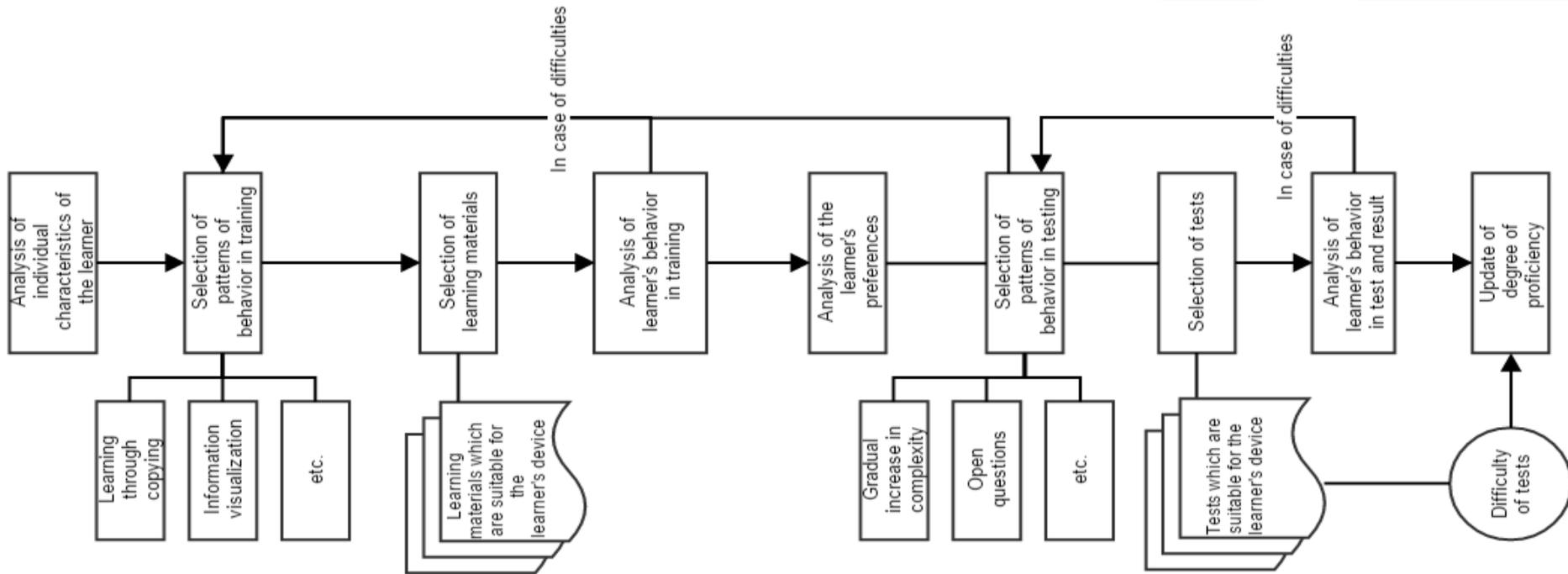


Рис. 11. Модель управляющего воздействия агента тьютора

Паттерны – XML-файлы с разметкой соответствия индивидуальных характеристик и особенностей контента



Дальнейшие задачи

- Объединение в единую модель в среде динамического моделирования (предположительно, IBM WebSphere Business Modeler)
- Извлечение из открытых данных сведений о распределении событий во времени
- Проведение имитационного моделирования и анализ результатов
- Применение полученных результатов для разработки системы поддержки электронного образовательного процесса



Вспомогательные задачи. Использование исторических данных для анализа активности обучающегося

	R ²	Number of interactions with course	Number of days of activity	Number of video views	Number of course chapters studied	Number of forum posts
Introduction to Solid State Chemistry	0.81	0.479	0.182	-0.237	0.462	-0.007
Introduction to Computer Science and Programming	0.78	0.370	0.406	-0.162	0.240	-0.005
Introduction to Biology	0.75	0.814	0.173	-0.328	0.164	-0.021
Electricity and Magnetism	0.82	0.676	0.341	-0.243	0.104	-0.006
Mechanics Review	0.84	0.467	0.050	-0.043	0.454	0.017
Health in Numbers	0.74	0.624	0.070	-0.295	0.424	-0.017
Human Health and Global Environmental Change	0.67	0.793	0.031	-0.286	0.236	0.015

Таблица 1. Регрессионный анализ исторических данных по нескольким курсам из открытого набора данных о курсах HarvardX и MITx за 2013 учебный год

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/26147>

Работа в группах на STEM специальностях повышает вероятность успеха в курсе на треть. [Scott Freeman]

Различные алгоритмы кластеризации в зависимости от параметров:

- Число групп;
- Размер групп;
- По одной или нескольким характеристикам.

Задача для дальнейшей работы:

- Подбор алгоритмов для конкретных задач разбиения на группы и обоснование их использования



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Спасибо
за внимание!

daldunin@hse.ru