

Разработка систем адаптивного тестирования для MOOC

МАЗОРЧУК МАРИЯ, К.Т.Н., ДОЦЕНТ

MOOC

Massive open online course — массовые (широкодоступные, публичные) открытые дистанционные онлайн курсы

Признаки MOOC:

- привлечение преподавателей лучших университетов;
- наличие графика, расписания, дедлайнов;
- наличие многочисленных каналов обратной связи слушатель-преподаватель, слушатель-слушатель;
- бесплатность;
- как следствие, массовость и глобальность: тысячи, десятки и сотни тысяч пользователей со всего мира

Coursera

The screenshot shows the Coursera website interface in Russian. At the top, there is a navigation bar with the Coursera logo, links for "Учреждения" (Institutions), "Войти" (Log In), and a "Зарегистрироваться" (Sign Up) button. The main banner features the headline "Учитесь онлайн на лучших курсах мира" (Learn online on the best courses in the world) and a search bar with the placeholder text "Чему бы вы хотели научиться?" (What would you like to learn?). Below the search bar is a link to "или посмотрите каталог >" (or view the catalog >). The banner also displays statistics: "15 981 343 learners", "1 470 курсы" (courses), and "135 partners".

The "Популярные Специализации" (Popular Specializations) section contains the following items:

- Big Data** (California University)
- Python for Everybody** (Michigan University)
- Excel to MySQL: Analytic Techniques for Business** (Duke University)
- Data Science** (Johns Hopkins University)
- Machine Learning** (University of Washington)
- Business Analytics** (University of Pennsylvania)
- Введение в принципы и практику управления...** (California University)
- TESOL Certificate, Part I: Teach English Now!** (Arizona State University)

The browser's taskbar on the right shows the Windows logo, several application icons (Chrome, VLC, File Explorer, PowerPoint), and system tray information including the language "RU", a notification for "M52 Доступ к И...", the time "20:38", and the date "02.11.2015".

edX

The screenshot shows the edX website homepage in a browser window. The browser's address bar displays `https://www.edx.org`. The page features a navigation menu with links for Courses, How It Works, Schools & Partners, and About. A search bar contains the text "I want to learn about...". A prominent banner image shows two students looking at a laptop, with the text "Learn from the best. Anytime. Anywhere." and "Join our growing global community of over 5 million learners". A "Find Courses" button is visible below the banner. The footer of the banner lists partner institutions: MIT Massachusetts Institute of Technology, HARVARD UNIVERSITY, Berkeley UNIVERSITY OF CALIFORNIA, THE UNIVERSITY OF TEXAS SYSTEM, TU Delft, and Australian National University. Below the banner, a section titled "Popular Courses Starting Soon" displays four course cards: MongoDB M101x, UC3Mx IT.1.1x, Microsoft DEV204x, and MITx 6.00.2x. The browser's taskbar on the right shows icons for Windows, Chrome, YouTube, File Explorer, and PowerPoint. The system tray at the bottom right indicates the language is RU, the time is 20:37, and the date is 02.11.2015.

MOOC - Пошук G... x Доверие к серти... x Coursera - лидер... x edX - Пошук Goo... x edX | Free online c... x EdX - Вікіпедія x

`https://www.edx.org`

edX Courses ▾ How It Works ▾ Schools & Partners ▾ About ▾ I want to learn about... Q Sign In Register

Learn from the best. Anytime. Anywhere.

Join our growing global community of over 5 million learners

Find Courses

MIT Massachusetts Institute of Technology HARVARD UNIVERSITY Berkeley UNIVERSITY OF CALIFORNIA THE UNIVERSITY OF TEXAS SYSTEM TU Delft Australian National University >

Popular Courses Starting Soon

Course	Institution	Verified
MongoDBx M101x	MIT	Yes
UC3Mx IT.1.1x	UC3M	Yes
Microsoft DEV204x	Microsoft	Yes
MITx 6.00.2x	MIT	Yes

M52 Доступ к И...
20:37
02.11.2015

Udacity

The screenshot shows a web browser window displaying the Udacity website. The browser's address bar shows the URL <https://www.udacity.com>. The website's navigation bar includes links for 'Nanodegree', 'Catalog', 'Sign In', and 'Sign Up'. The main content area features a large banner with the text 'Become a Full Stack Developer' and 'Earn a Nanodegree Credential', accompanied by an orange 'Learn more' button. Below the banner, there is a 'FEATURED' badge and a mention of 'Udacity featured in the New York Times!'. At the bottom of the banner, logos for 'Courses built by' include Google, facebook, mongoDB, cloudera, and at&t. A system tray notification for 'M52' is visible on the right side of the desktop.

MOOC - | x | Доверие x | Udacity - x | Online Co x | Coursera x | coursera - x | Coursera x | edX | Free x | EdX — Bl x

<https://www.udacity.com>

UDACITY

Nanodegree Catalog Sign In Sign Up

Become a
Full Stack Developer
Earn a Nanodegree Credential

Learn more

FEATURED Udacity featured in the [New York Times!](#)

Courses built by Google facebook mongoDB cloudera at&t

M52
Доступ к И...

RU

20:42
02.11.2015

Khan academy

MOOC - Поиск Google x Доверие к сертификатам x Khan academy - Поиск x About Khan Academy x

https://www.khanacademy.org/about

KHANACADEMY Subjects

Мазорчук Мария

Our mission is to provide a free, world-class education for anyone, anywhere.

Centrifugation 55% Fibrinogen (+clotting factors) electrolytes, nutrients white blood cells / platelets

About Team Interns Content Specialists Board Supporters Volunteer Blog Donate

A personalized learning resource for all ages

Khan Academy offers practice exercises, instructional videos, and a personalized learning dashboard that empower learners to study at their own pace in and outside of the classroom. We tackle math, science, computer programming, history, art history, economics, and more. Our math missions guide learners from kindergarten to calculus using state-of-the-art, adaptive technology that identifies strengths and learning gaps. We've also partnered with institutions like NASA, The Museum of Modern Art, The California Academy of Sciences, and MIT to offer specialized content.

3D cube, The Starry Night, 20% progress, Line graph, 3D cube, The Starry Night

M52 Доступ к Интернету

RU

20:47 02.11.2015

Массивные данные

>72
million
users to date

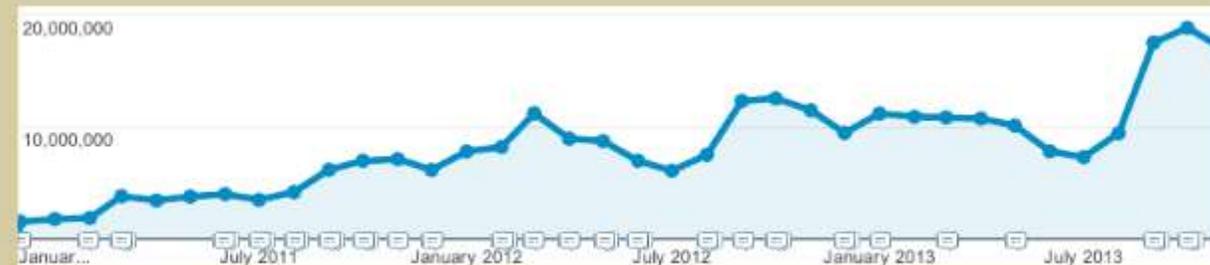
>1.6 billion
problems answered

>800,000
hours of video watched per month

>216
countries

>250,000
self-identified teachers

Visits per month



Moocology

MOOCology ua@moocology.org +380 67 4647 001

**WE WILL EXPAND
MOOC MOVEMENT
IN UKRAINE**

New eco-system of
professional education

REINVENTING EDUCATION

We will gather community of best ukrainian experts to make professional education effective.
Are you one of them?

MOOCOLGY DREAM-MISSIONS MOOCOLGY ORGANISATION

RU
M52
Доступ к И...
20:58
02.11.2015

PROMETHEUS

The screenshot shows a web browser window with the URL `courses.prometheus.org.ua/courses/KPI/Algorithms101/2015_Spring/about`. The website header features the Prometheus logo and navigation links. The main content area displays the course title "Розробка та аналіз алгоритмів. Частина 1" and a "REGISTER FOR ALGORITHMS101" button. A chalkboard graphic with the word "ALGORITHMS" and various terms like "discrete", "array", "search", "effective", "complexity", "logic", and "complexity" is visible. The page includes a table with course details and a sidebar with social media icons and a language selector.

Номер Курсу	Algorithms101
Старт	26 Січ. 2015
Час	8-9 годин на тиждень

огляд

Курс "Розробка та аналіз алгоритмів" присвячений вивченню та побудові програмних алгоритмів. У рамках цього курсу ви зможете познайомитись із різними підходами до розробки алгоритмів, дослідження та порівняння їх ефективності.

Курс побудований від простого до складного. Кожного тижня розглядається новий підхід до створення алгоритмів та кілька прикладів алгоритмів в рамках цього підходу.

RU

M52
Доступ к И

20:56
02.11.2015

Недостатки MOOC

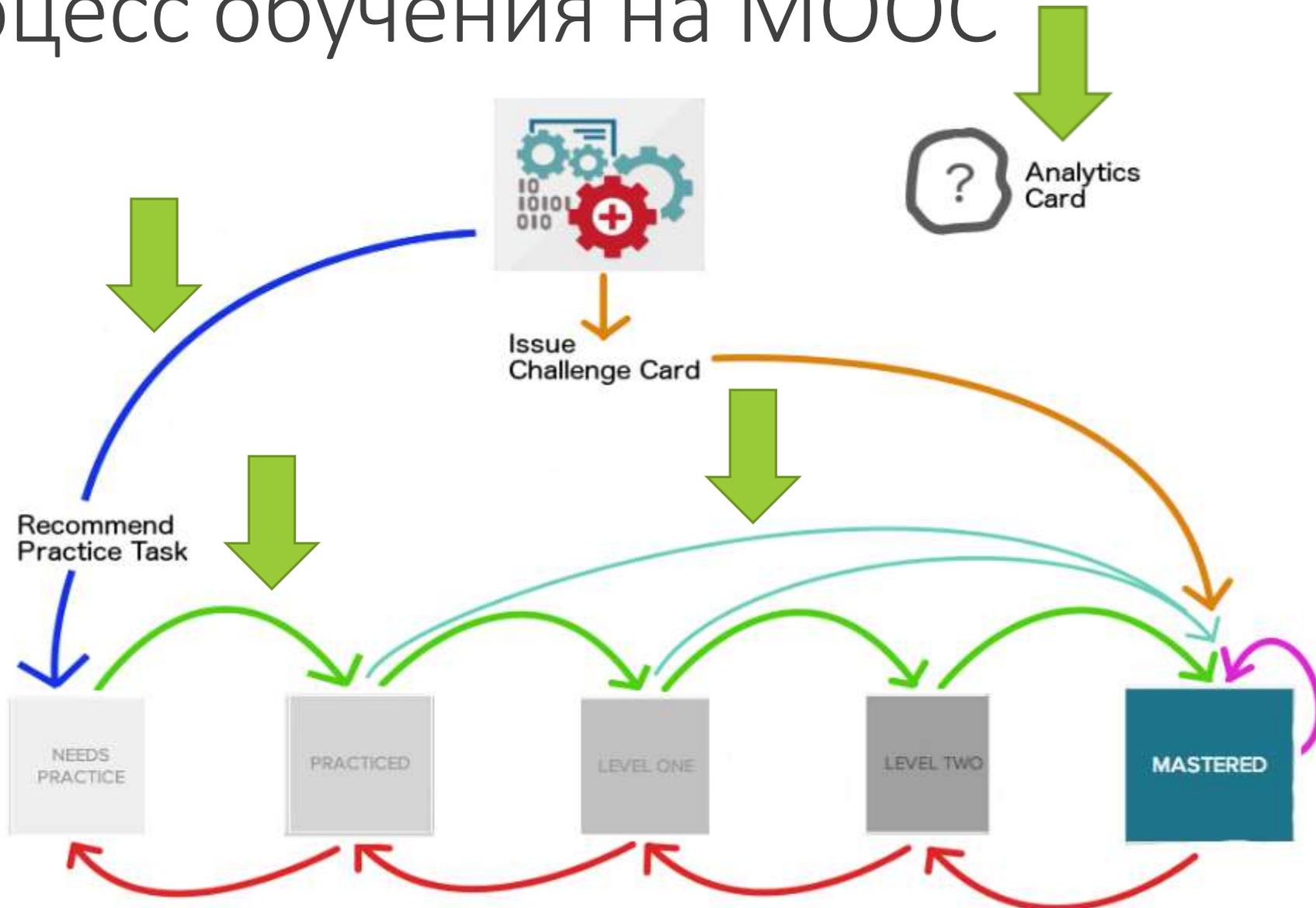
С точки зрения пользователя:

- Необходимость самоорганизации для самостоятельной работы
- Не всегда возможно обеспечить обратную связь
- Доверие к результатам обучения

С точки зрения разработчиков:

- Время
- Деньги
- Качество

Процесс обучения на MOOC



Как проверить полученные знания?

- Тестирование
- Адаптивное тестирование
- Компьютерное адаптивное тестирование (САТ)

Тестирование

- Классическая теория тестов (СТТ) - экспериментальный контроль обеспечивается посредством согласования и случайного отбора
- Теория тестовых заданий (IRT) решает проблему контроля с помощью статистического согласования (van der Linden & Humbleton, 1997)

Классическая теория тестов

Три основные понятия:

- тестовый балл (который часто называют еще видимым баллом)
- истинный балл (балл)
- балл погрешности.

Истинный балл - математическое ожидание баллов участника тестирования при прохождении параллельных вариантов теста.

Балл погрешности - разница между истинным баллом участника и тестовым баллом.

Предположения СТТ

- Наблюдательный балл (X) является линейной связью истинного балла (T) и ложного балла (E)

$$X = T + E$$

- Истинные и ложные баллы не коррелируют между собой
- Средний ложный балл по множеству равен нулю
- Ложные баллы в параллельных тестах некоррелированы
- Стандартная ошибка измерения одинакова по всей измеренной шкале

Критика СТТ

Экзамен и характеристики теста не являются обособленными:

- Характеристики ТЗ зависят от группы
- Характеристики экзаменованного зависят от теста
- Определение стандартной ошибки измерения и надежности
- Направлена на тест, а не на ТЗ

Современная теория тестирования (Item Response Theory)

Допущение в современной теории тестирования

Результат ответа на тест есть величина которая зависит от двух латентных (скрытых) характеристик :

- трудность задания β
(измеряется в логитах)
- уровень подготовки тестируемого θ
(измеряется в логитах)



Заявленные обещания IRT

- ❑ Статистика теста не зависит от групп, для которых она рассчитывается
- ❑ Баллы, описывающие успешность тестируемого, не зависят от сложности теста
- ❑ Модели теста обеспечивают основу для согласования тестовых заданий по уровням умений
- ❑ Модели теста, не требуют точного параллельного теста для оценки надежности

Сравнительный анализ характеристик методов тестирования

Характеристика	Классическая теория тестов (СТТ)	Теория ответов на тестовые задания (IRT)
Тип модели	Линейная	Нелинейная
Уровень моделирования	Тест в целом	Задания теста
Допущения (условия использования модели)	Слабые: данные тестирования легко удовлетворят условия СТТ	Сложные: данные тестирования не всегда удовлетворяют условиям IRT
Отношения между характеристиками тестового задания и подготовленностью участников тестирования	Не специфические	Характеристическая кривая заданий и способностей обучаемого
Инвариантность характеристик заданий и подготовленности участников (независимость выборок заданий и участников тестирования)	Нет: параметры заданий и участников тестирования зависят от выборок заданий и участников тестирования	Да: параметры заданий и участников не зависят от выборок заданий и участников тестирования (при условии, что данные удовлетворяют условиям IRT)
Статистика задания	p – сложность задания d – дифференцирующая способность r – корреляция	b, a, c – параметры модели, отражающие информационную функцию задания
Размер выборки (для определения величины параметров задания и участников тестирования)	Относительно небольшие – 200-500 испытуемых	Зависит от модели IRT, но обычно более 500 испытуемых

Однопараметрическая модель Г.Раша

Например

$$P_j(\theta) = \frac{e^{1,7(\theta - \beta_j)}}{1 + e^{1,7(\theta - \beta_j)'}}$$

$$P_i(\beta) = \frac{e^{1,7(\theta_i - \beta)}}{1 + e^{1,7(\theta_i - \beta)'}}$$

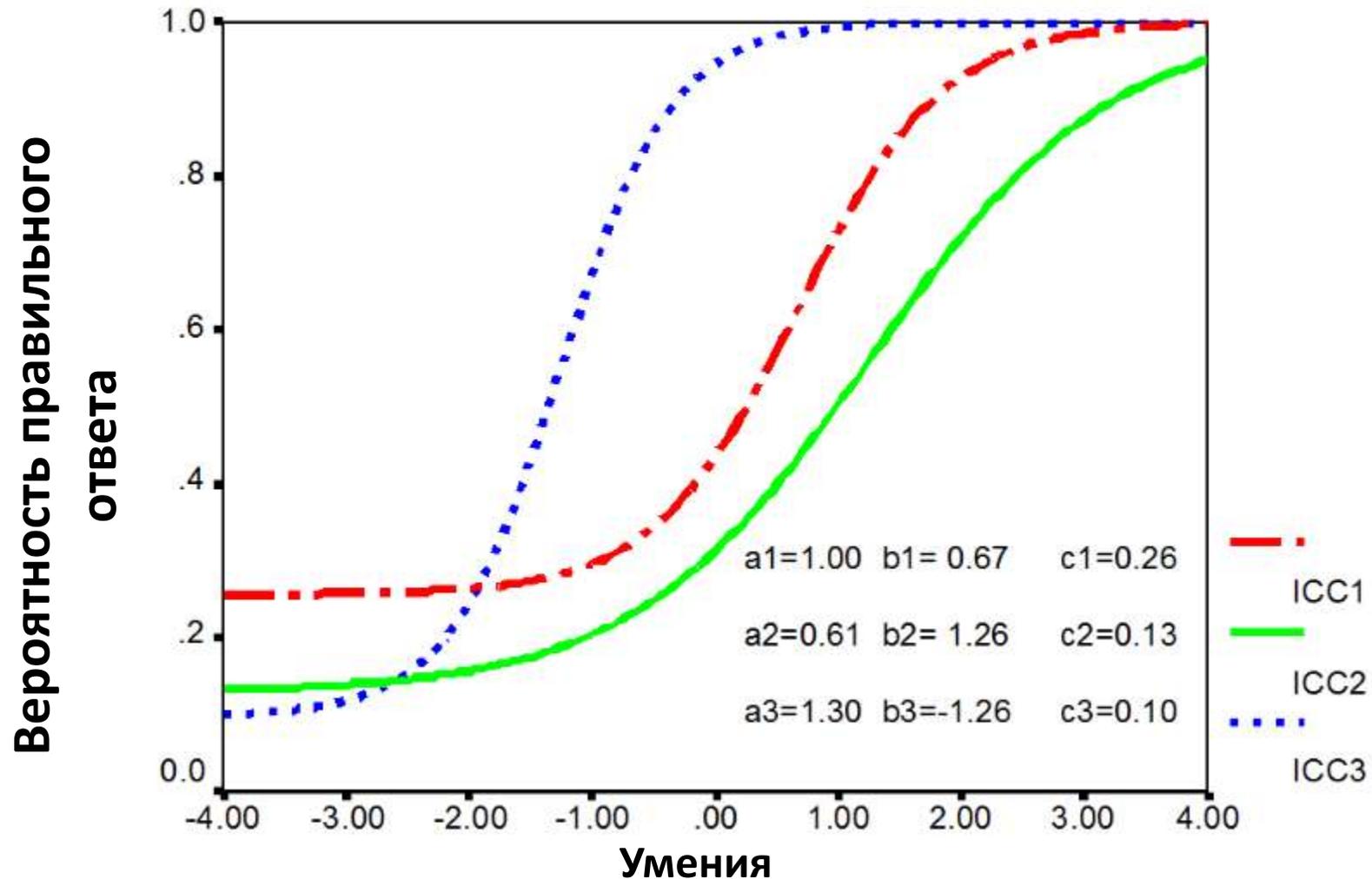
Пусть в тесте

Задание 2 сложнее Задания 1.

Тогда при той же подготовленности испытуемого (или при тех же самых знаниях) вероятность правильного ответа на Задание 2 будет меньше, чем вероятность правильного ответа на Задание 1



Характеристические кривые заданий



Характеристические кривые

Кривую теста (ХКГ) можно построить как сумму характеристических кривых заданий (ХКЗ), включенных в теста.

Ее можно использовать для прогнозирования ожидаемых баллов экзаменуемых на данных уровнях подготовки.

При разработке двух тестов для измерения одинакового уровня подготовки их баллы можно легко уравнять благодаря характеристическим кривым теста.

Двух- и однопараметрические модели

В двухпараметрической модели предполагается, что на вероятность правильного ответа влияет трудность (b) и дискриминативность (a) задания, и не влияет угадывание ($c = 0$)

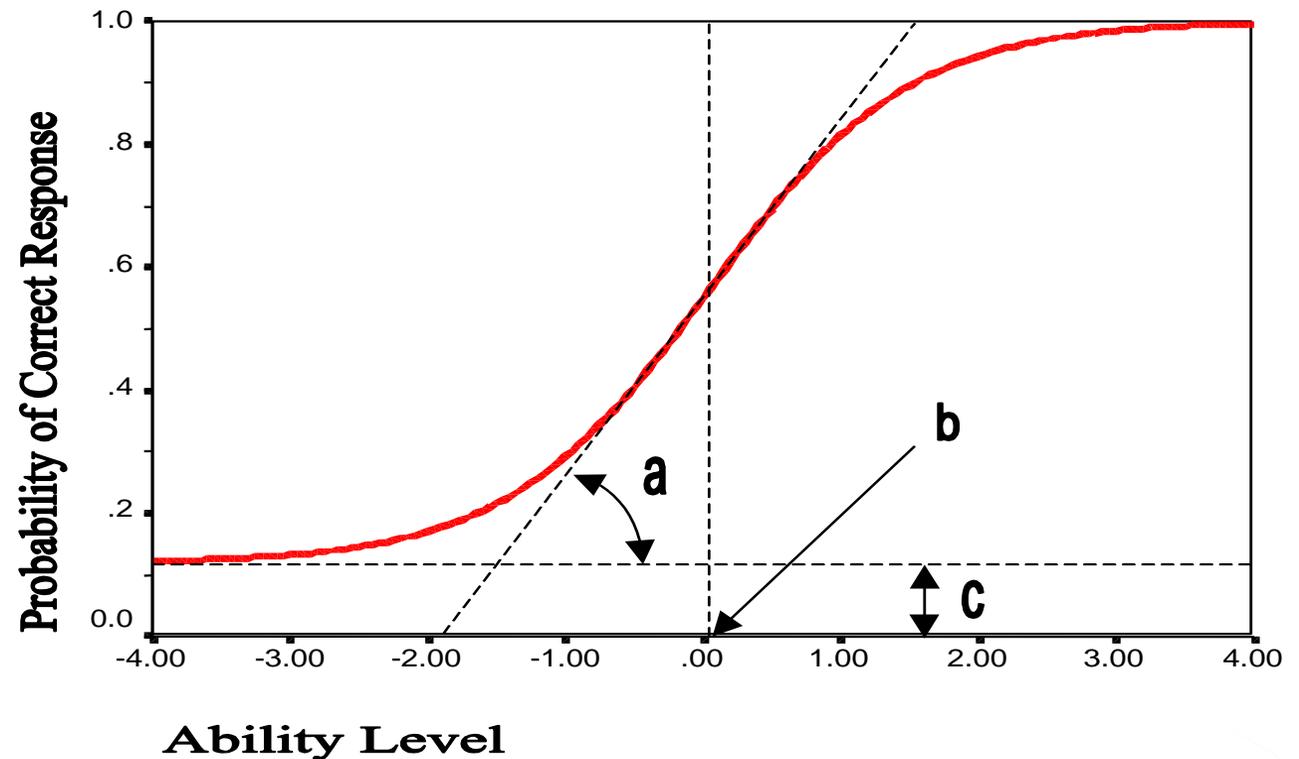
$$P(\Theta) = \frac{1}{1 + e^{a_j(\Theta - b_j)}}$$

В однопараметрической модели предполагается, что достаточно лишь параметра сложности ($c = 0, a = -1.7$)

Трех-параметрическая модель

$$P(\Theta) = c_j + (1 - c_j) \frac{1}{1 + e^{-a_j(\Theta - b_j)}}$$

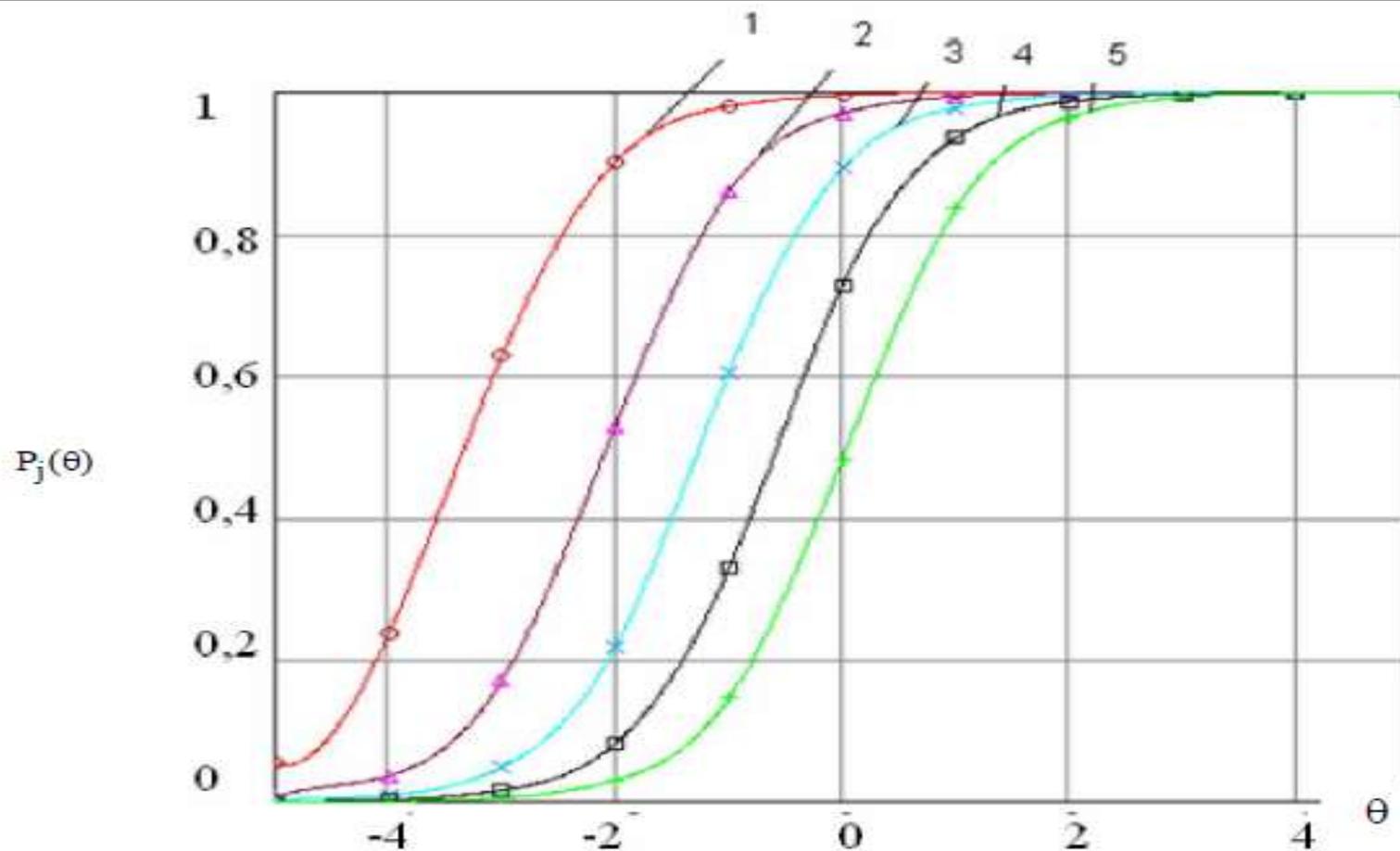
3PL Item Characteristic Curve



Политомические модели IRT

- ❑ Модель градуированных ответов (Graded Response Model (GRM; Samejima, 1969))
- ❑ Политомическое ТЗ трактуется как ряд дихотомий вокруг предельных точек $k-1$, расположенных между категориями k
- ❑ Есть обобщением двухпараметрической модели

Характеристические кривые для пяти категорий



Информационные функции

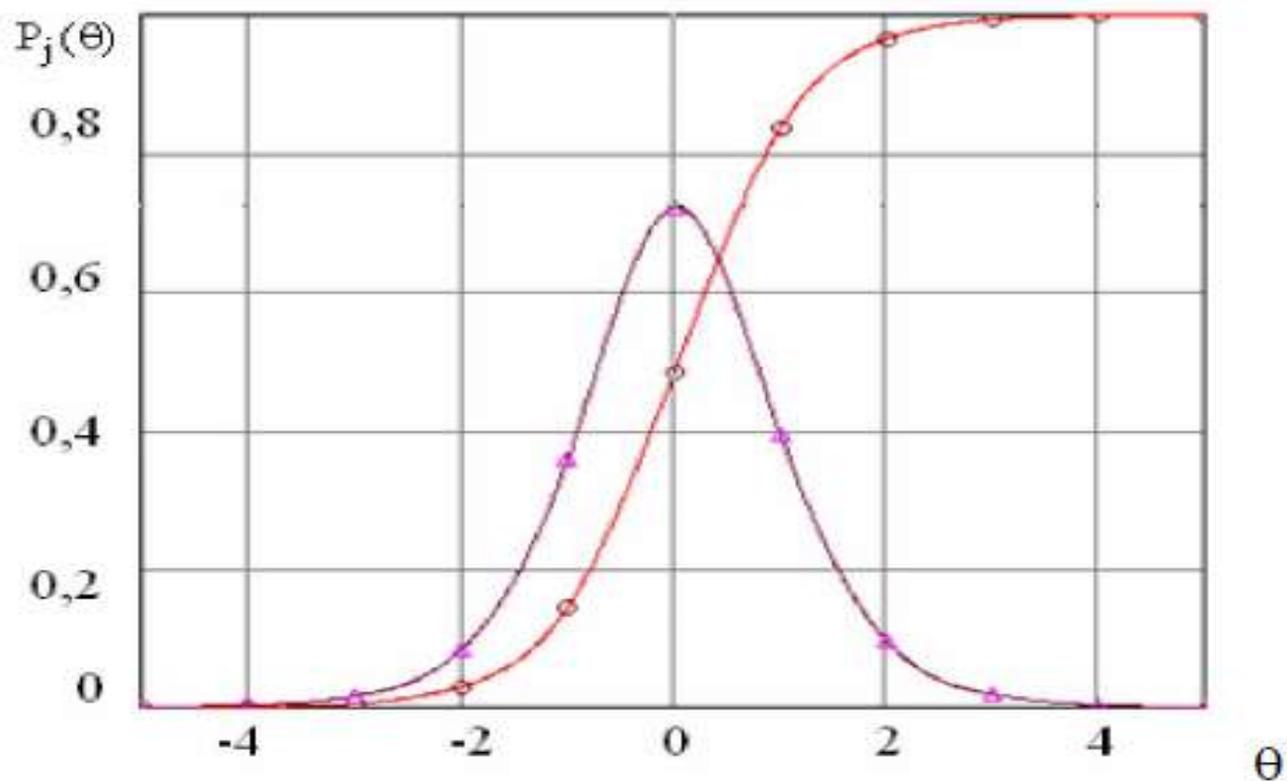
Одна из самых полезных особенностей моделей IRT - это существование информационной функции задачи (IIF)

- эта функция показывает вклад отдельных тестовых заданий в оценивания умений на определенных уровнях.
- IIF является обратной по отношению к дисперсии ошибки измерения на данных уровнях умений.

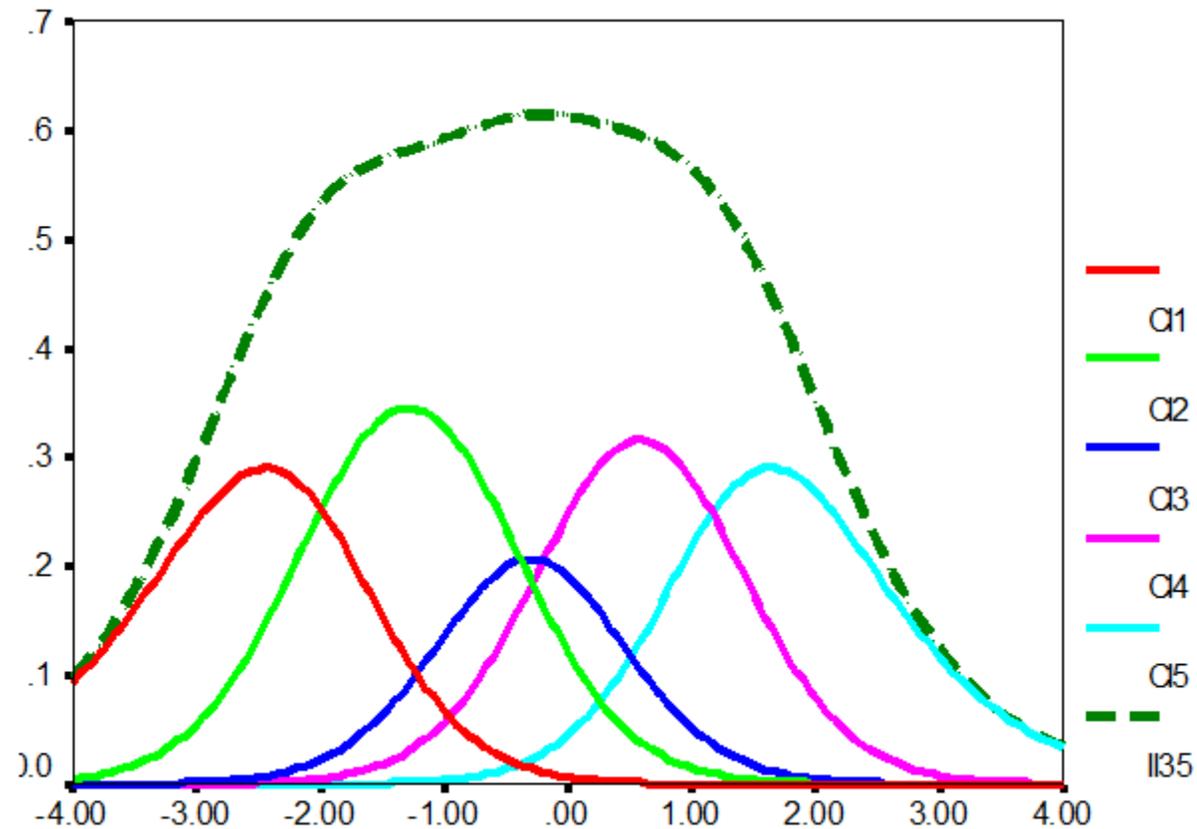
Информационные функции задач суммируются в информационную функцию теста (TIF), что демонстрирует надежность, по которой измеряются умения на определенных уровнях.

- TIF очень полезна в построении теста.
- Разработчики тестов могут определить целевую TIF, а затем выбрать тестовые задания по IIF, что будет приближать желаемый TIF к удовлетворительной

Информационная функция задания с нулевой трудностью



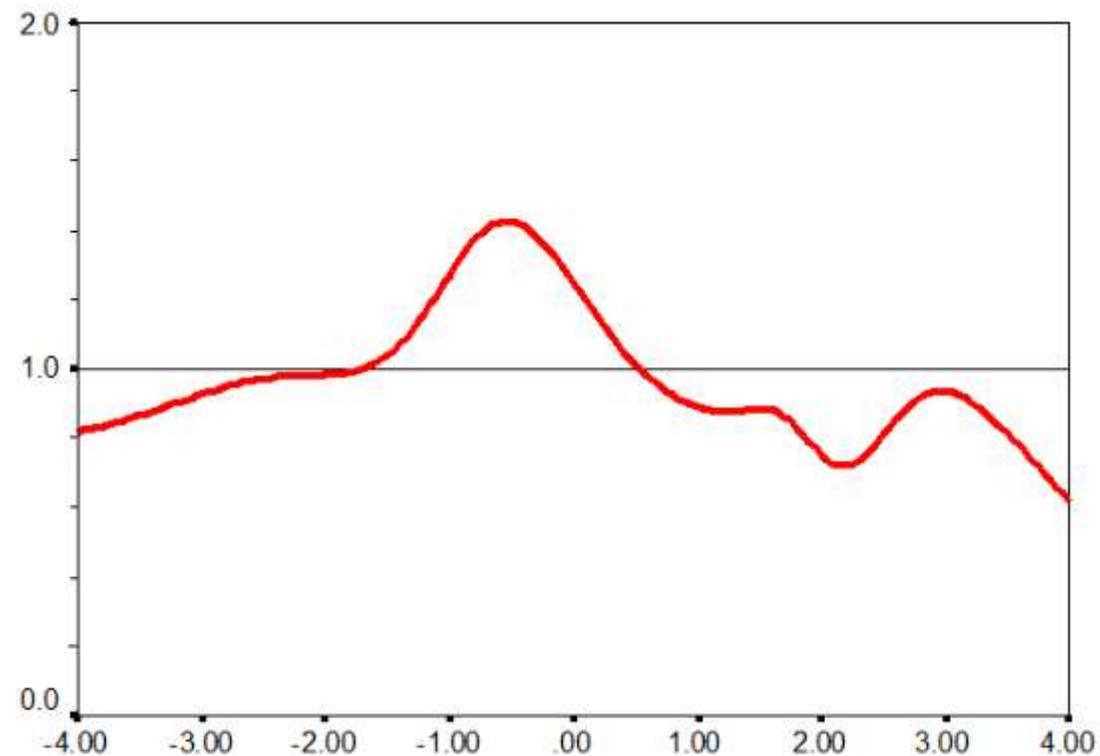
Информационная функция политомического задания



Относительная эффективность

Кривая относительной эффективности строится в качестве соотношения между двумя кривыми информативности теста

Это демонстрирует связь между надежностью измерения, осуществляется путем двух сопоставимых тестов в каждой точке континуума умений



Использование теории IRT

- Подготовка тестовых заданий и конструирования тестов с заданными характеристиками (например, тест SAT в США (ETS))
- Построение систем компьютерного адаптивного тестирования (система CAT для проведения массовых онлайн тестов при поступлении, система CAT **Khan academy**)
- Мониторинговые исследования качества образования (например, системы международных сравнительных исследований в области математики и естественных дисциплин TIMSS, PISA).
- Использование IRT для шкалирования результатов (шкалирование ЕГЭ в России)

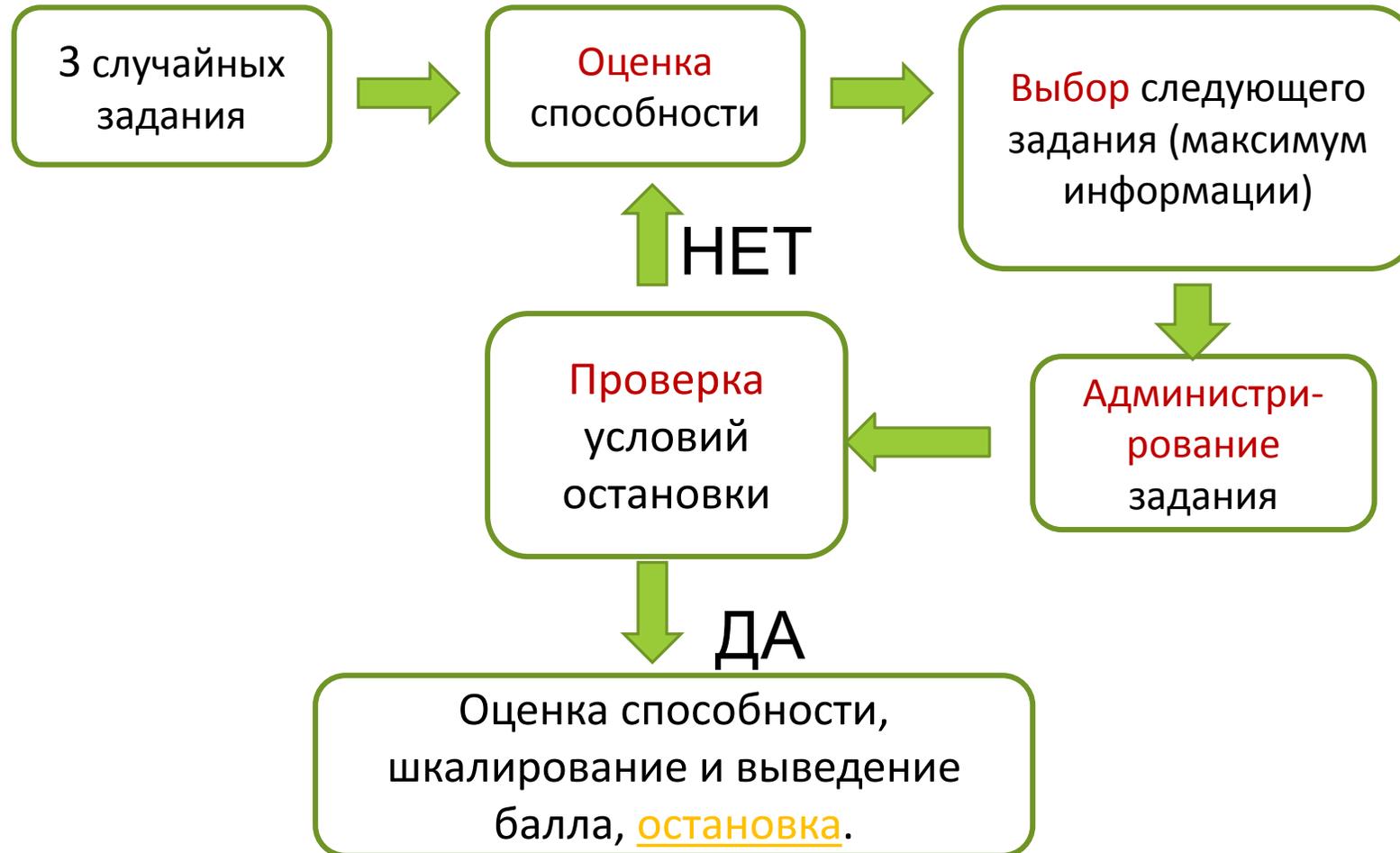
Что такое CAT?

CAT – это не только милый пушистый друг...



Это формат тестирования на основе моделей теории IRT (теории ответов на тестовые задания)

Алгоритм CAT



Типичная сессия CAT

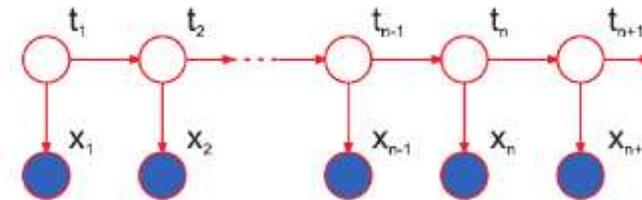
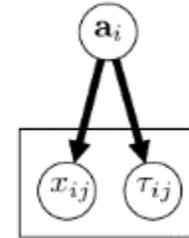


Проблемы использования IRT

- Тесты должны быть гомогенны
- Вопросы теста не связаны с компетенциями (моделью компетенций курса)
- Нет возможности оценить уровень подготовки в соответствии с курсом (как учил, что использовал, результаты онлайн подготовки)
- Большие ошибки прогноза на начальном этапе
- Большие банки тестовых заданий
- Проблемы в калибровке тестов и т.д.

Решение

- ❑ Модели ВКТ – Байесовская трассировка знаний
- ❑ Скрытые Марковские сети



Student Modeling

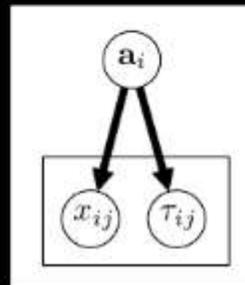
Currently on site:

- Adaptive pretest driven by multi-dimensional item response theory

$$p(\mathbf{a}_i) = \mathcal{N}(0, \mathbf{I})$$

$$p(x_{ij} = 1 | \mathbf{a}_i) = \sigma(\mathbf{W}_j \mathbf{a}_i + b_j)$$

$$p(\log \tau_{ij} | \mathbf{a}_i) = \mathcal{N}(\mathbf{V}_j \mathbf{a}_i + c_j, \alpha \mathbf{I})$$



$\mathbf{a}_i \equiv$ abilities vector for student i

$x_{ij} \in \{0, 1\} \equiv$ correctness for student i on exercise j

$\tau_{ij} \equiv$ response time for student i on exercise j

$\mathbf{W}_j, \mathbf{V}_j \equiv$ weights vectors for exercise j

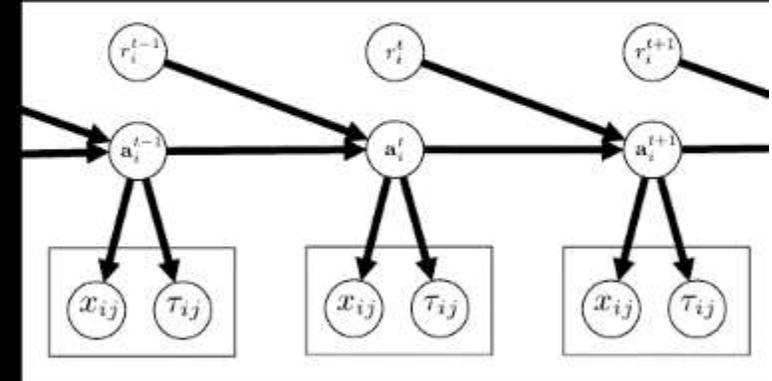
$b_j, c_j \equiv$ bias terms for exercise j

Student Modeling

Under development:

- Temporal Multidimensional Item Response Theory (TMIRT)

$$p(\mathbf{a}_i^{t+1} | \mathbf{a}_i^t, r_i^t) = \mathcal{N}(\mathbf{a}_i^t + \Phi_{r_i^t} \mathbf{a}_i^t + \theta_{r_i^t}, \Sigma_{r_i^t})$$



$r_i^t \equiv$ index of resource used by student i at time t

$\Phi_r \equiv$ linear update weights for resource r

$\theta_r \equiv$ linear update bias for resource r

$\Sigma_r \equiv$ update covariance for resource r



Спасибо за внимание!

