

Czesław JĘDRZEJEK\*, \*\*

Rafał RENK\*\*

Lilianna ANIOŁA-JĘDRZEJEK\*\*\*

\* Instytut Technik Telekomunikacyjnych i Informatycznych w Poznaniu (ITTI)

\*\* Akademia Techniczno-Rolnicza w Bydgoszczy (ATR)

\*\*\* Studium Języków Obcych, Politechnika Poznańska

## NOWE MOŻLIWOŚCI TECHNICZNE W KOMPUTEROWEJ OCENIE WYNIKÓW NAUCZANIA

### 1. WSTĘP

#### 1.1. WPŁYW INTERNETU NA EDUKACJĘ I POMIAR WYNIKÓW KSZTAŁCENIA

Obecnie na świecie następuje gwałtowny rozwój edukacji za pośrednictwem internetu. Powstają wirtualne uniwersytety, liczące kilkadziesiąt tysięcy studentów. Rynek usług szkoleń dla firm przez internet na świecie wyniesie w r. 2003 wg Merrill Lynch 11 mld USD.

Jako przykład aktywności biznesowej można podać fakt, że Pearson, firma działająca na polu edukacji i mediów, nawiązała współpracę z IBM w celu stworzenia oferty interaktywnych usług edukacyjnych. Firmie Pearson przyświeca idea edukacji bez ograniczeń geograficznych i czasowych. Pierwsze testowe kursy będą organizowane wkrótce w Stanach Zjednoczonych oraz w kilku krajach europejskich i azjatyckich. Usługa „nauki na żądanie” może być realizowana w standardowych sieciach internetowych/intranetowych, a jako terminal może służyć komputer PC, urządzenie typu *set-top box* lub zaawansowany telefon komórkowy.

Wydarzeniem o ogromnym znaczeniu jest deklaracja Massachusetts Institute of Technology bezpłatnego udostępnienia w internecie wszystkich materiałów edukacyjnych (wykładów) przez kilka lat w ramach projektu OpenCourseware<sup>1</sup>. Jest to wyzwanie dla prywatyzacji wiedzy, ale też przejaw jeszcze większej jej globalizacji.

Sprzężenie efektywności i dostępności internetu z postępami w oprogramowaniu i interfejsach graficznych wymusza zmiany rynkowe w kierunku stosowania komputerowych metod oceny. W r. 2000 CAT, Inc., Thomson Prometric i Vue przeprowadziły razem 1,9 mln testów rocznie głównie w dziedzinie kompetencji, dotyczących systemów firm Microsoft i Novell.

W USA kolejne rodzaje egzaminów o charakterze ponaduczelnianym (*high-stake*), np. SAT, GRE, przeprowadzane są w sposób komputerowy<sup>2</sup>. Europa jest w tym względzie bardziej konserwatywna i zmiany mogą rozciągnąć się na okres 10 lat.

Ocena jest integralną częścią procesu nauczania, służy także do diagnozowania procesów edukacyjnych.

<sup>1</sup> <http://web.mit.edu/newsoffice/nr/2001/ocw.html>.

<sup>2</sup> R. E. Bennett, *How the internet will help large-scale assessment reinvent itself*, „Education Policy Analysis Archives” 9, 5, Feb. 2001.

Niniejsza praca przedstawia nowe możliwości technologiczne, które mogą być wykorzystane do oceny. W pewnych przypadkach dają one możliwość wyjścia poza pytania zamknięte (zadania wyboru wielokrotnego — choć autorzy użyliby raczej wyrażenia „wyboru z wielu możliwości”) i mogą wzbogacić sprawdzanie na bieżąco rozumienia materiału przez studenta oraz formalny pomiar wyników kształcenia.

Autorzy zdają sobie sprawę, że oprócz uwarunkowań techniczno-finansowych zastosowanie tych metod wymaga rozległych stosowanych badań pedagogicznych i psychologicznych.

W szczególności otwarte jest zagadnienie, na ile nowe możliwości informatyczne wpływają na rzetelność, trafność pomiaru oraz poprawiają obiektywizm w ocenianiu<sup>3</sup>. Jeśli chodzi o głębsze aspekty testowania, a w szczególności — konsekwencje etyczne, odsyłamy czytelnika do omówienia idei S. Messicka<sup>4</sup>.

## 1.2. FUNKcjONALNOŚCI PODSTAWOWEGO NARZĘDZIA TESTOWEGO I SYSTEM INFORMATYCZNY

Wykonanie pracy zostało oparte na wykorzystaniu systemu zarządzania zdalnego nauczania, rozwijanego przez ITTI w ramach projektu „Leonardo da Vinci”, pt. „Multimedia Distance English Courses for Polish Users in Legal, Banking and Finance, Science and Technology, and Safety Training Sectors with Elements of European Union Regulations and Standards” (Pilot Project N: PL/98/1/86519/PI/III.1.a/CONT)<sup>5</sup>. Ryc. 1.1. przedstawia architekturę systemu. Ryc. 1.2. przedstawia rozszerzoną wersję warstwowej programowej architektury systemu.

W systemie tym zaimplementowane są typy pytań automatycznie oceniane przez system oraz typy pytań, których ocena należy do prowadzącego dany kurs (tutora). System dostarcza możliwości implementacji pytań jako testowych lub egzaminacyjnych. Wszystkie typy pytań zostały zaimplementowane w skryptach, które odpowiadają poszczególnym typom pytań, przy tym zachowano jednolity graficzny interfejs użytkownika.

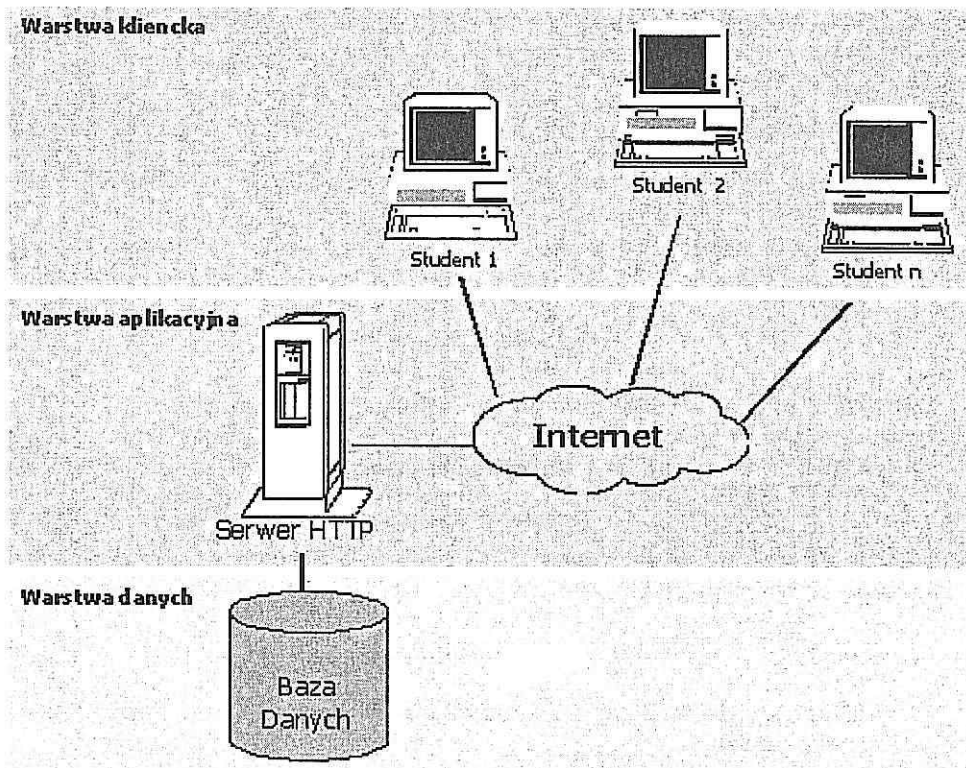
Ćwiczenia egzaminacyjne różnią się od ćwiczeń testowych tym, że nie występują na poszczególnych stronach materiału kursowego. Zestaw ćwiczeń egzaminacyjnych może być zdefiniowany dla każdego kursu. Z tego zestawu pytań system automatycznie losuje zadaną liczbę ćwiczeń dla każdego studenta, tworząc jednocześnie formularz egzaminacyjny. Formularz egzaminacyjny jest gotowy do wypełnienia przez studenta. Po wypełnieniu wyniki wpisywane są do bazy, a ocena ćwiczeń, które mogą być ocenione w sposób automatyczny, jest wykonywana zaraz po zakończeniu egzaminu. Formularz egzaminacyjny można także wydrukować i rozdać studentom w formie arkuszy papierowych.

System umożliwia umieszczanie w egzaminach i testach następujących rodzajów ćwiczeń:

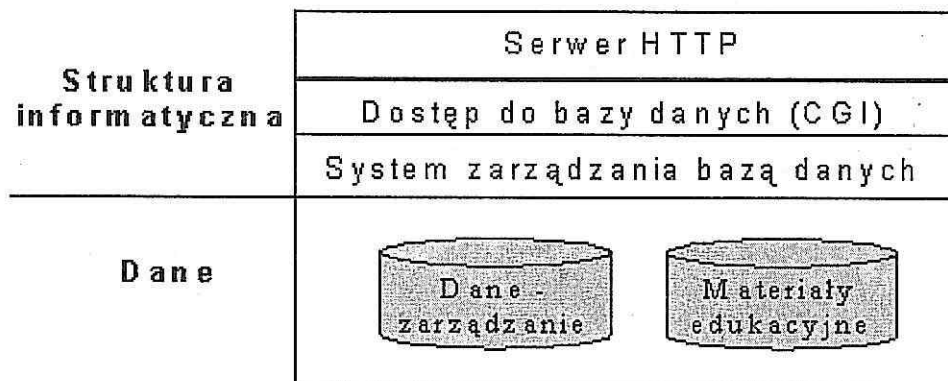
<sup>3</sup> D. B. Niemierko, *Czy zadanie wyboru wielokrotnego nadaje się do diagnozowania procesów edukacyjnych?*, materiały sympozjum „Zadania wyboru wielokrotnego, Wałbrzych 7–8 grudnia 2000.

<sup>4</sup> R. E. Orton, *Samuel Messick's consequential validity*, „Philosophy of Education”, [www.cd.uiuc.edu/EPS/PES-Yearbook/1998/orton.html](http://www.cd.uiuc.edu/EPS/PES-Yearbook/1998/orton.html).

<sup>5</sup> L. Anioła-Jędrzejek, A. Adamczyk, *Distance education system and thematic English learning courses in the framework of Leonardo da Vinci Programme*, „ODL Networking For Quality Learning”, Lisbon 2000, s. 200–207.



Ryc. 1.1. Trójwarstwowa architektura systemu



Ryc. 1.2. Rozszerzona architektura systemu

1. Wybierz jedną możliwą odpowiedź — przyciski typu „radio”

Przykład:

*Which word is connected with speed?*

- green  
 soft  
 quick

2. Wybierz jedną lub kilka poprawnych odpowiedzi — przyciski typu „checkbox”

Przykład:

*Which word is connected with speed?*

- vivid  
 quick  
 fast

3. Zaklasyfikuj elementy — lista rozwijana z jedną poprawną odpowiedzią

Przykład:

*Classify the elements according to the tense of the word.*

- go   
 hear   
 woke   
 gone

4. Uzupełnij tekst, wybierając poprawną odpowiedź tekstową — lista rozwijana z jedną poprawną odpowiedzią

Przykład:

When Angela was younger, she   
 school. She  schools three times  
 between the ages of 11 and 16. She never  
 except before exams,  
 and she  studying altogether when  
 she was fourteen.

5. Uzupełnij, tekst wpisując poprawną odpowiedź tekstową — pola tekstowe  
Przykład:

When Angela was younger, she   
school. She  schools three times  
between the ages of 11 and 16. She never  
 except before exams,  
and she  studying altogether when  
she was fourteen.

6. Uzupełnij tekst, zmieniając zawartość w polach tekstowych na poprawną — pola tekstowe z predefiniowaną zawartością  
Przykład:

When Angela was younger, she   
school. She  schools three  
times between the ages of 11 and 16.  
  
except before exams,

7. Udziel tekstowej odpowiedzi na pytanie — puste pole edycji tekstu wysyłane do prowadzącego

8. Wyślij odpowiedź tekstową do kolegi lub koleżanki — puste pole edycji tekstu wysyłane do innego studenta z grupy

9. Dopasuj elementy do siebie, przeciągając myszką — oprogramowanie w JavaScript/DHTML (dynamiczny HTML) do obsługi opcji *przeciągnij i upuść*  
Przykład:

*Match the present and past forms of these irregular verbs.*

go  
tell  
get  
come  
hear  
say  
know


*drag & drop*

told	came	went	heard	said
knew	got			

W ramach poleceń i zamiast tekstu w treści ćwiczenia można używać obiektów multimedialnych dostępnych w HTML (obrazków, komiksów, zdjęć, krótkich animacji, muzyki, odsłuchiwanie muzyki lub tekstu, sekwencji wideo itp.)

Z syntetycznego punktu widzenia testy sprowadzają się (choć są różnie technicznie implementowane) do wyboru z wielu możliwości lub z pola tekstowego, lecz odpowiedzi nie są skorelowane. Jednak nawet wersja podstawowa wykracza poza możliwości „papierowej” wersji egzaminu.

### 1.3. OGÓLNE FUNKCJONALNOŚCI ROZSZERZONEGO NARZĘDZIA TESTOWEGO I SPOSÓB TESTOWANIA

W systemach zdalnego nauczania istnieje możliwość zaimplementowania pewnych funkcjonalności, które dają dodatkowe możliwości analizy i oceny postępów studentów. Dzięki w pełni skomputeryzowanej implementacji ćwiczeń testowych i pytań egzaminacyjnych możliwe jest otrzymanie informacji, które w przypadku tradycyjnego egzaminu byłyby niemożliwe do uzyskania. Te informacje mogą być gromadzone w celu ich dalszej obróbki, np. dane dotyczące trafności i czasu odpowiedzi na poszczególne typy pytań dla każdego studenta i dla każdego pytania.

Wykorzystanie dodatkowych danych daje możliwość np.:

- analizy statystycznej lub wykorzystania z góry ustalonych korelacji pomiędzy pytaniami;
- oceny efektywności różnych technik odpowiedzi (różnych typów pytań), np. wybór odpowiedzi z listy rozwijanej, wybór odpowiedzi z zestawu, w którym żadna lub wiele odpowiedzi jest poprawna itp.;
- możliwość oceny trudności czy też poprawności sformułowania pytań przy wykorzystaniu korelacji pomiędzy zarejestrowanym czasem odpowiedzi i efektywnością odpowiedzi na każde z pytań;
- kontroli i odpowiedniej automatycznej reakcji na postępy w nauce dokonywane przez studenta podczas przyswajania materiału poprzez badanie odpowiedzi na pytania testowe umieszczane po każdej większej części kursu.

Wszystkie wyżej wymienione analizy są rozważane w dalszej części artykułu.

Moduł egzaminacyjny w zmodyfikowanej wersji został wykorzystany przy egzaminowaniu studentów Akademii Techniczno-Rolniczej w Bydgoszczy na kierunku studiów Systemy Telekomunikacyjne i Multimedia, z przedmiotu „Cyfrowe audio i wideo”. Cały egzamin składał się z dwóch części: egzaminu pisemnego i zdalnego egzaminu przeprowadzanego z serwera znajdującego się w ITTI w Poznaniu.

Ta część pracy została w skrócie przedstawiona na konferencji Online-Educa<sup>6</sup>.

Część pisemna przeprowadzona była w sposób tradycyjny. Egzaminator określił pytania i studenci odpowiadali na nie pisemnie. Po egzaminie egzaminator sprawdzał wszystkie odpowiedzi i oceniał je osobiście.

<sup>6</sup> L. Anioła-Jędrzejek, C. Jędrzejek, A. Adamczyk, R. Renk, *Database enhanced assessment in distance learning*, Conference Online Educa 2000, Berlin, s. 40–45; L. Anioła-Jędrzejek, A. Grabowska, A. Adamczyk, T. Shannon-Little, *Multimedia distance English courses and learning management system in the framework of Leonardo Da Vinci programme*, Konferencja EDEN 2001, Sztokholm, maj 2001.

W części egzaminu zdalnego, składającego się z 17 pytań, egzaminowano cztery grupy studentów. Studenci mieli określony czas na udzielenie odpowiedzi na wszystkie pytania. Ocena poprawności odpowiedzi i zliczanie punktów były realizowane przez system na podstawie udzielanych odpowiedzi.

Dodatkowo wprowadzono ujemne punkty za udzielenie niepoprawnej odpowiedzi. Za każde pytanie typu „wybór w przypadku jednej poprawnej odpowiedzi” (przycisk „radio” lub lista rozwijana z jedną poprawną odpowiedzią) można było uzyskać następującą liczbę punktów:

- +1 za poprawną odpowiedź na pytanie,
- -1 za niepoprawną odpowiedź na pytanie,
- 0 za wybranie odpowiedzi „brak odpowiedzi” lub niezatwierdzenie odpowiedzi (wysłanie odpowiedzi do serwera poprzez kliknięcie klawisza „Zatwierdź odpowiedź”)

Zasada ta jest podobna do używanej w konkursach „Kangur”. Jej celem jest zniechęcenie studentów do zgadywania odpowiedzi oraz do odpisywania, co jest plagą polskiej edukacji. Skala punktacji była różna w zależności od tego, czy dane pytanie było skorelowane z innym pytaniem, czy też nie.

## 2. KORELACJE W PYTANIACH

### 2.1. TYPY KORELACJI

W pytaniach egzaminacyjnych wyróżniono dwa typy korelacji: korelację zewnętrzną (hierarchiczną) i korelację wewnętrzną. Korelacja zewnętrzna zachodzi pomiędzy dwoma różnymi pytaniami, natomiast korelacja wewnętrzna odnosi się do jednego pytania, w którym występuje kilka pytań szczegółowych lub możliwość wyboru więcej niż jednej poprawnej odpowiedzi.

Przypadek korelacji zewnętrznej zachodzi dla pary pytań, z których jedno, nazywane pytaniem nadrzędnym, „master”, jest pytaniem ogólnym (np. pytaniem o definicję), a drugie, nazywane pytaniem podrzędnym, „slave”, jest pytaniem o konkretne zastosowanie definicji z pytania „master”. Jest mało prawdopodobne, aby student mógł udzielić poprawnej odpowiedzi na pytanie „slave” nie odpowiadając poprawnie na pytanie „master”. Wystąpienie takiej sytuacji (tzn. poprawnej odpowiedzi na pytanie typu „slave” i niepoprawnej odpowiedzi na pytanie typu „master”) wskazuje na przypadkowość odpowiedzi studenta i brak zrozumienia całości zagadnienia, jakie obejmuje dane pytanie. Takie sytuacje przy ocenie pytań były analizowane i przydzielano odpowiednio mniej punktów, niż by to wynikało ze zliczenia poprawnych odpowiedzi.

Podczas egzaminu wystąpiły trzy pary pytań zewnętrznie skorelowanych. Jedna z takich par pytań przedstawiona jest poniżej. Poprawne odpowiedzi na podpytania zaznaczono czcionką półgrubą.



Treść pytania „master” — dwie listy rozwijane:

Następujący wzór  $G = \frac{\sigma_{ff}^2}{\sigma_n^2}$  opisuje:

- a) transmitancję odwrotną do  $B(z)$ ,
- b) współczynnik kompresji danych,
- c) szum gaussowski,
- d) brak odpowiedzi.

Kompresję danych uzyskujemy kiedy współczynnik  $G$  jest:

- a) = 1,
- b) < 1,
- c) > 1,
- d) brak odpowiedzi.

Treść pytania „slave” — wybór jednej odpowiedzi z wielu możliwych (test wyboru 1 z N — *radiobutton*):

Które z równań opisuje stosunek sygnału do szumu:

a)  $\eta_Q = \frac{\sigma_x^2}{\sigma_Q^2}$

b)  $\eta_Q = \frac{\sigma_Q^2}{\sigma_x^2}$

c)  $\eta_Q = \frac{\sigma_x^2 \cdot \sigma_Q^2}{12\Delta}$

d)  $\eta_Q = \frac{\sigma_Q^2 + \sigma_x^2}{\sigma_x^2}$

e) brak odpowiedzi

Podczas egzaminu wystąpiły cztery przypadki korelacji zewnętrznej dla wyżej przedstawionych pytań: dwa w pierwszej egzaminowanej grupie i dwa w grupie czwartej. Dla pozostałych dwóch par skorelowanych pytań korelacje w odpowiedziach nie wystąpiły.

Korelacja wewnętrzna sprawdza powiązania pomiędzy podpytaniami (np. wiele list rozwijanych w jednym pytaniu), umieszczonymi w jednym pytaniu. W tym typie korelacji nie można określić pytania nadrzędnego i pytań podrzędnych — wszystkie podpytania są jednakowo ważne i korelowane na jednakowym poziomie. Na przeprowadzo-



nym egzaminie było jedno pytanie z korelacją wewnętrzną. Treść pytania przedstawiona jest poniżej.

Treść pytania z korelacją wewnętrzną — cztery listy rozwijane:

Wypisz cechy cyfrowych systemów liniowych opisane poniżej:

1.  $y = h(x) = Ax, x = c_1x_1 + c_2x_2$   
 $y = h(x) = A [c_1x_1 + c_2x_2] = c_1 [Ax_1] + c_2 [Ax_2] = c_1y_1 + c_2y_2$ 
  - a) przyczynowość,
  - b) liniowość,
  - c) odwracalność,
  - d) brak odpowiedzi.
2. Jeżeli odpowiedź systemu na deltę Kroneckera  $\delta(n)$  jest  $h(n)$ , to odpowiedź  $\delta(n - k)$  jest  $h(n - k)$ .
  - a) stabilność,
  - b) liniowość,
  - c) stacjonarność,
  - d) brak odpowiedzi.
3.  $h_n = 0$  dla  $n < 0$ 
  - a) przyczynowość,
  - b) odwracalność,
  - c) stacjonarność,
  - d) brak odpowiedzi.
4. Suma jego odpowiedzi impulsowych jest skończona
  - a) odwracalność,
  - b) stacjonarność,
  - c) stabilność,
  - d) brak odpowiedzi.

## 2.2. Zasady poprawności odpowiedzi w pytaniach skorelowanych i typu „checkbox”

Całkowita ocena studenta składa się z sumy punktów całej odpowiedzi. Niestety z powodu występowania pytań skorelowanych i typu „checkbox”, aby porównywać pytania między sobą, zachodzi konieczność zdefiniowania poprawności odpowiedzi na dane pytanie. Reguły te są złożone i subiektywne. Ale podobnie subiektywna jest reguła przyznawania tej samej liczby punktów za pytania o różnym stopniu

trudności. Zagadnienia przez nas poruszone powinny stać się przedmiotem systematycznych badań, ponieważ nie jest oczywiste, że testy wychodzące poza wybór z wielu możliwości są rzetelne, choć usuwają pewne słabości testów wyboru z wielu możliwości.

Z powodu korelacji typu i pytań typu „checkbox” suma punktów za indywidualne odpowiedzi, także podpytań, nie równa się sumie punktów za odpowiedzi przy pewnej definicji poprawności odpowiedzi.

Przyjęto następujące reguły punktacji:

A) Jeśli pytanie składało się z kilku podpytań (np. kilku list rozwijanych), za każde podpytanie, przy braku korelacji wewnętrznej pomiędzy podpytaniami, można było uzyskać następującą liczbę punktów:

- +1 za poprawną odpowiedź na podpytanie,
- -1 za niepoprawną odpowiedź na podpytanie,
- 0 za wybranie odpowiedzi „brak odpowiedzi” lub niezatwierdzenie odpowiedzi (wysłanie odpowiedzi do serwera poprzez kliknięcie klawisza „zatwierdź odpowiedź”).

B) W przypadku korelacji wewnętrznej punkty przyznawano wg następującego schematu: sumowano punkty za poszczególne podpytania zgodnie ze schematem przedstawionym w punkcie 1.3., z tą różnicą, że za odpowiedzi typu „brak odpowiedzi” przydzielano -0,5 punktu.

C) W pytaniach typu: wybór w przypadku wielu możliwych poprawnych odpowiedzi (przyciski „checkbox”) można było uzyskać następującą liczbę punktów:

- poprawna odpowiedź +1 (dobra),
- „brak odpowiedzi” 0 (brak odp.),
- niepoprawna odpowiedź -1 (zła),
- niezaznaczona poprawna odpowiedź -0,5 (zła),
- niezaznaczona niepoprawna odpowiedź 0,5 (dobra).

D) Opcja „wybór w przypadku wielu możliwych poprawnych odpowiedzi” nie może być częścią korelacji wewnętrznej.

E) Suma punktów za poszczególne podpytania składała się na liczbę punktów, jakie można było uzyskać za całe pytanie. Przyjęto zasadę, że poprawną odpowiedź na pytanie uzyskujemy wówczas, gdy odpowiadający uzyskał dodatnią liczbę punktów za całe pytanie (suma punktów zdobytych za poszczególne podpytania) oraz dał poprawną odpowiedź na co najmniej 50% podpytań. Należy tak ustalić punktację, aby student nie mógł uzyskać poprawnej odpowiedzi, losowo zaznaczając odpowiedzi na podpytania.

F) W przypadku korelacji zewnętrznej punkty przyznawano wg następującego schematu:

- przy poprawnej odpowiedzi na pytanie „master” sumowano punkty dla pytania „master” i „slave” wg schematu przyznawania punktów podanego w punkcie A, niezależnie od odpowiedzi na pytanie typu „slave”;
- przy „braku odpowiedzi” lub niepoprawnej odpowiedzi na pytanie typu „master”, a poprawnej odpowiedzi na pytanie typu „slave”, punkty przyznawano wg następującej reguły: suma punktów za pytanie typu „master” (punktowane wg reguł opisanych w punkcie 1.3.) i pytanie typu „slave”, pomnożone przez współczynnik ważności korelacji. Współczynnik ważności korelacji określa siłę korelacji pomiędzy skorelowa-

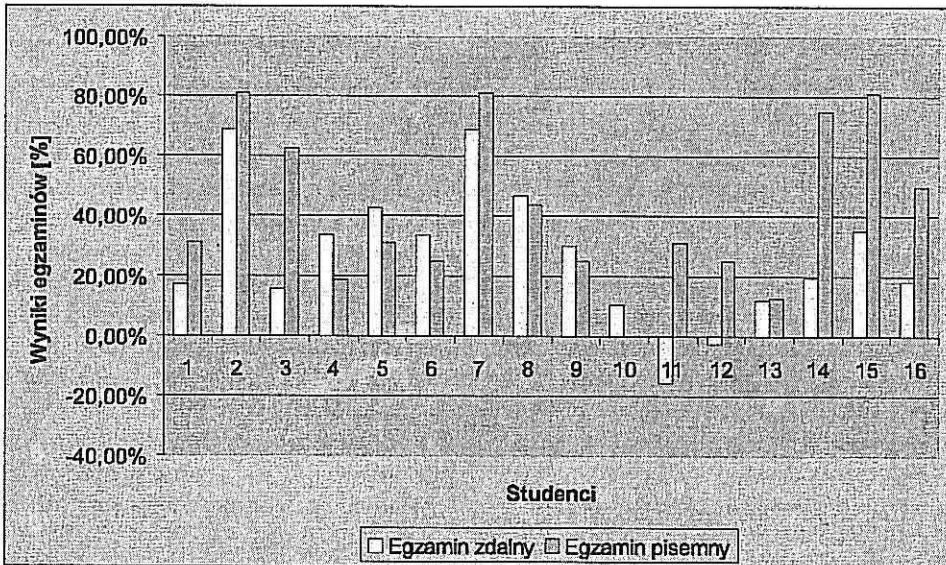
nymi pytaniami, tzn. wskazując, jak dalece odpowiedź na jedno pytanie jest powiązana z odpowiedzią na drugie ze skorelowanych pytań. Dostępne były trzy poziomy ważności korelacji i odpowiadające im współczynniki: korelacja duża (współczynnik 0,15), korelacja „średnia” (współczynnik 0,30) i korelacja mała (współczynnik 0,60). Podczas egzaminu wykorzystywana była korelacja normalna;

— przy „braku odpowiedzi” lub złej odpowiedzi na pytanie typu „master” i braku odpowiedzi lub złej odpowiedzi na pytanie typu „slave” sumowano punkty dla pytania „master” i „slave” wg schematu przyznawania punktów podanego w punkcie 1.3.

### 3. EFEKTYWNOŚĆ ODPOWIEDZI ZE WZGLĘDU NA RÓŻNE TECHNIKI JEJ UDZIELANIA

Poniższy rysunek przedstawia porównanie efektywności odpowiedzi studentów z egzaminu pisemnego i zdalnego.

Ponad połowa studentów (56,25%) ma zbliżoną (różnica pomiędzy wynikami z obydwu typów egzaminu nie przekraczała co do wartości bezwzględnej 15%) efektywność w obu typach egzaminu. Wskazywałoby to na fakt, że efektywność odpowiedzi jest w dużej mierze zależna od wiedzy studenta, a nie od formy egzaminu. Rezultat lepszy z egzaminu pisemnego niż z egzaminu zdalnego uzyskało 37,5% studentów (różnica pomiędzy egzaminem pisemnym a zdalnym przekraczała 15%). Studenci ci mogli być zaskoczeni formą egzaminu zdalnego i podejść do niego z pewnym dystansem lub preferowali formę egzaminu pisemnego. Reszta studentów (6,25%) uzyskała lepszy wynik z egzaminu zdalnego (różnica pomiędzy egzaminem zdalnym a pisemnym przekraczała 15%), prawdopodobnie



Ryc. 3.1. Porównanie wyników egzaminu pisemnego i zdalnego

wykorzystując szansę przypadkowego trafienia dobrej odpowiedzi, zgodnie z rachunkiem prawdopodobieństwa.

Przy ocenie efektywności odpowiedzi ze względu na różne techniki jej udzielania rozważone zostaną trzy typy pytań (możliwych odpowiedzi), na które studenci musieli odpowiadać podczas zdalnego egzaminu. Pierwszy typ pytania to pytanie polegające na teście wyboru jednej poprawnej odpowiedzi z kilku możliwych (test wyboru 1 z N — przycisk typu „radio”, *radiobutton*), drugi typ pytania to pytanie typu: test wyboru z możliwymi kilkoma poprawnymi odpowiedziami (test wyboru M z N, *checkbox*), i trzeci typ pytania to listy rozwijane z jedną poprawną odpowiedzią (*pull-down lists*). Liczba pytań z poszczególnymi typami była zbliżona do siebie dla każdego typu. We wszystkich typach pytań student, jeśli nie chciał odpowiedzieć na zadane pytanie, mógł wybrać możliwość „brak odpowiedzi”. Każda odpowiedź na pytanie musiała być potwierdzona przez studenta przed wysłaniem. Student mógł również dowolną liczbę razy zmieniać odpowiedź na konkretne pytanie. Niepotwierdzenie odpowiedzi na pytanie traktowane było przez system jako „brak odpowiedzi”.

Tab. 3.1. przedstawia procentowy udział odpowiedzi poprawnych, niepoprawnych i „brak odpowiedzi” w poszczególnych typach pytań.

TABELA 3.1

Typy pytań i ich wpływ na efektywność odpowiedzi studentów

Typy pytań	listy rozwijane	test wyboru (1 z N)	test wyboru (M z N)
średnio odpowiedzi poprawnych	56%	40%	53%
średnio odpowiedzi niepoprawnych	29%	46%	31%
średnio „brak odpowiedzi”	15%	14%	16%

W pytaniach typu lista rozwijana większość odpowiedzi była poprawna. Przyczyną wysokiego współczynnika poprawnych odpowiedzi może być to, że ten typ pytania w większości przypadków występował jako podpytanie, na które było łatwiej odpowiedzieć niż na inny typ pytania. Dla pytania typu test wyboru 1 z N prawdopodobieństwo poprawnych (40%) i błędnych (46%) odpowiedzi było zbliżone. Dla obydwu powyższych typów odpowiedzi prawdopodobieństwo braku odpowiedzi było niskie (odpowiednio 15% i 14%). Taki wynik może brać się z faktu, że studenci są pewni, że jedna z odpowiedzi jest poprawna, i jeśli jej nie znają, próbują trafić na nią przypadkowo. Zupełnie inna sytuacja wystąpiła w przypadku pytań typu: test wyboru M z N, w których zbiór poprawnych odpowiedzi jest znacznie większy niż w przypadku poprzednich pytań. Pomimo tej różnicy w stosunku do poprzednich typów pytań prawdopodobieństwo poprawnych odpowiedzi (53%), prawdopodobieństwo niepoprawnych odpowiedzi (31%), a także prawdopodobieństwo odpowiedzi „brak odpowiedzi” (16%) pozostały na zbliżonych poziomach.

#### 4. OCENA STOPNIA TRUDNOŚCI I SPOSOBU PRZEDSTAWIENIA PYTANIA W RELACJI DO WIEDZY STUDENTÓW

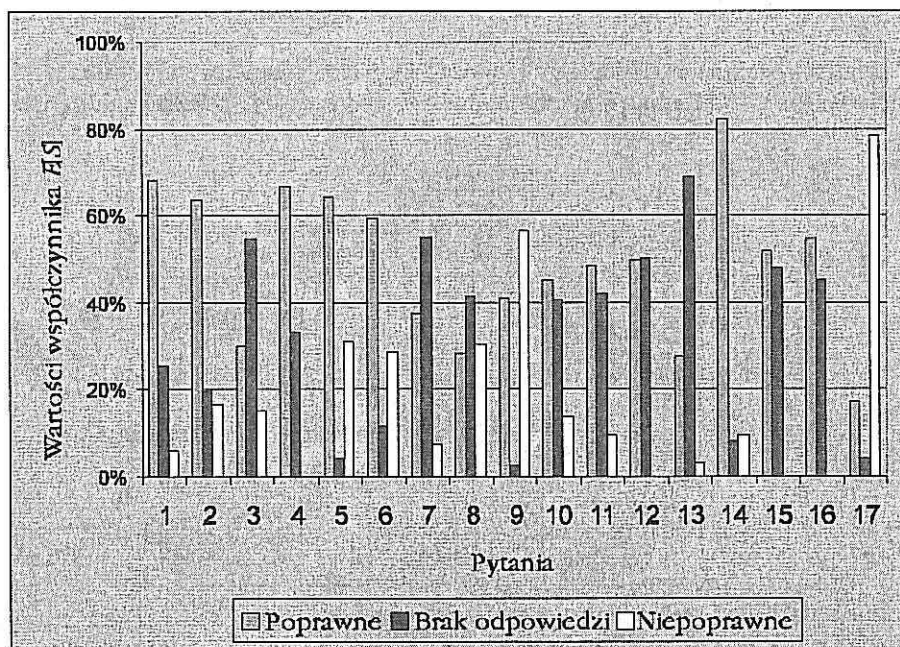
Ocena stopnia trudności i sposobu przedstawienia pytania w relacji do wiedzy studentów bazuje na efektywności odpowiedzi na poszczególne pytania i związanego z tym czasem odpowiedzi (rejestracji czasu wyświetlenia strony i czasu przesłania odpowiedzi na serwer). Przed przystąpieniem do analizy pytań określiliśmy współczynnik  $E[S]$  (zależność pomiędzy możliwymi odpowiedziami wyrażaną średnim czasem udzielania odpowiedzi a całkowitym czasem odpowiedzi dla wszystkich rodzajów pytań) dla: odpowiedzi poprawnych  $E[S]_{correct}$ , niepoprawnych  $E[S]_{incorrect}$  i „braku odpowiedzi”  $E[S]_{no\_answer}$  wyrażony odpowiednio dla poszczególnych typów pytań jako:

$$E[S]_{correct} = \frac{\text{średni czas odpowiedzi dla pytań z odpowiedziami poprawnymi}}{\text{średni czas odpowiedzi dla wszystkich typów odpowiedzi}}$$

$$E[S]_{no\_answer} = \frac{\text{średni czas odpowiedzi dla pytań z odpowiedziami „brak odpowiedzi”}}{\text{średni czas odpowiedzi dla wszystkich typów odpowiedzi}}$$

$$E[S]_{incorrect} = \frac{\text{średni czas odpowiedzi dla pytań z odpowiedziami niepoprawnymi}}{\text{średni czas odpowiedzi dla wszystkich typów odpowiedzi}}$$

Poniższy rysunek przedstawia wartości współczynnika  $E[S]$  dla poszczególnych pytań w egzaminie zdalnym.



Ryc. 4.1. Wartości współczynnika  $E[S]$  dla poszczególnych pytań

Na podstawie współczynnika  $E[S]$  można wyróżnić trzy schematy (wzory) odpowiedzi na pytania:

- liczba poprawnych, „brak odpowiedzi” i niepoprawnych spada monotonicznie;
- duża liczba odpowiedzi „brak odpowiedzi” przy małej liczbie odpowiedzi poprawnych i niepoprawnych;
- duża liczba poprawnych i niepoprawnych odpowiedzi przy małej liczbie odpowiedzi „brak odpowiedzi”.

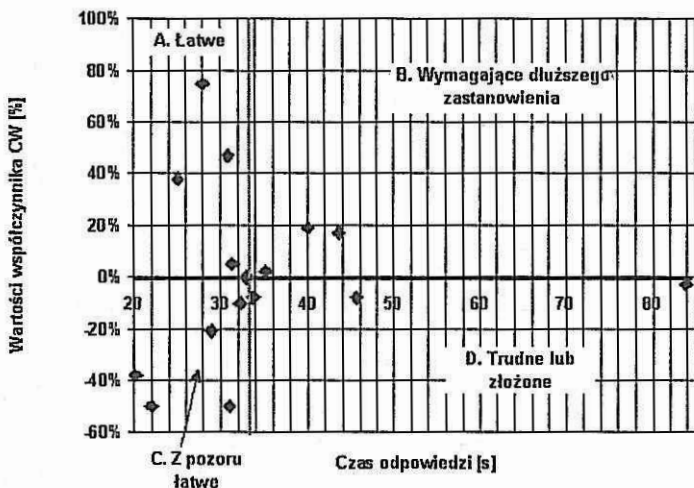
Pierwszy schemat odpowiedzi jest najczęściej spotykany. Ponad połowa pytań może zostać zaklasyfikowana do tego typu, co oznacza, że pytania są poprawnie zdefiniowane. Drugi schemat odpowiedzi reprezentuje pytania, które wyglądają na skomplikowane lub zbyt trudne dla studentów. Z tego też względu studenci w znacznej mierze rezygnują z udzielenia odpowiedzi na te pytania. Ostatni schemat odpowiedzi wskazuje pytania, które wydają się łatwe, albo studenci oceniają je jako pytania, na które jest szansa przypadkowego trafienia poprawniej odpowiedzi.

Innym współczynnikiem, jaki zdefiniowaliśmy w celu oceny stopnia trudności i sposobu przedstawienia pytania w relacji do wiedzy studentów, jest współczynnik  $CW$ . Stanowi on różnicę pomiędzy procentowo wyrażoną liczbą poprawnych odpowiedzi oraz sumą procentowo wyrażonych odpowiedzi „brak odpowiedzi” i niepoprawnych odpowiedzi dla poszczególnych pytań (patrz wzór poniżej).

Przy analizie indywidualnych pytań traktujemy je jako nieskorelowane, inaczej niż dla sumy punktów całej odpowiedzi.

$$CW(\%) = \text{odpowiedzi poprawne} - (\text{odpowiedzi niepoprawne} + \text{odpowiedzi „brak odpowiedzi”})$$

Poniższy rysunek przedstawia umiejscowienie pytania na wykresie w zależności od współczynnika  $CW$  i średniego czasu odpowiedzi studentów dla poszczególnych pytań.



Ryc. 4.2. Charakterystyka pytania w zależności od współczynnika  $CW$  i średniego czasu odpowiedzi



Pionowa linia odpowiada średniemu czasowi odpowiedzi na wszystkie pytania. Pozioma linia — średniej współczynnika *CW* dla wszystkich pytań. Współczynnik *CW* określa poziom trudności pytania. Wyżej wymienione linie dzielą cały obszar wykresu na cztery podobszary: A, B, C i D. Pytania z obszaru A zaliczyć można do pytań łatwych, pytania z obszaru B — do pytań wymagających dłuższego zastanowienia, pytania z obszaru C — to pytania, które z pozoru wydają się łatwe, i pytania z obszaru D — trudne lub zbyt złożone.

Rozkład pytań na wykresie jest niemal równomierny dla poszczególnych obszarów. Jedno pytanie, dla którego średni czas odpowiedzi przekracza 80 s, stanowi wyjątek. Jest to pytanie z wewnętrzną korelacją, prezentowane wcześniej. Wskazywałoby to na konieczność podzielenia tego pytania na kilka pytań cząstkowych.

## 5. NARZĘDZIE AUTORSKIE DO TWORZENIA ZAWARTOŚCI KURSÓW

System zdalnego nauczania składa się z trzech podstawowych komponentów: systemu zarządzania nauczaniem LMS (*learning management system*), służącego do administrowania kursami i studentami, wirtualnego środowiska nauczania VLE (*virtual learning environment*), wyposażonego w elastyczny i łatwy w użyciu interfejs oraz narzędzi autorskich do tworzenia materiałów kursowych.

Dwa pierwsze komponenty systemu już istnieją i działają i mogą być wykorzystane do przeprowadzonego egzaminu. Trwają prace nad trzecim komponentem.

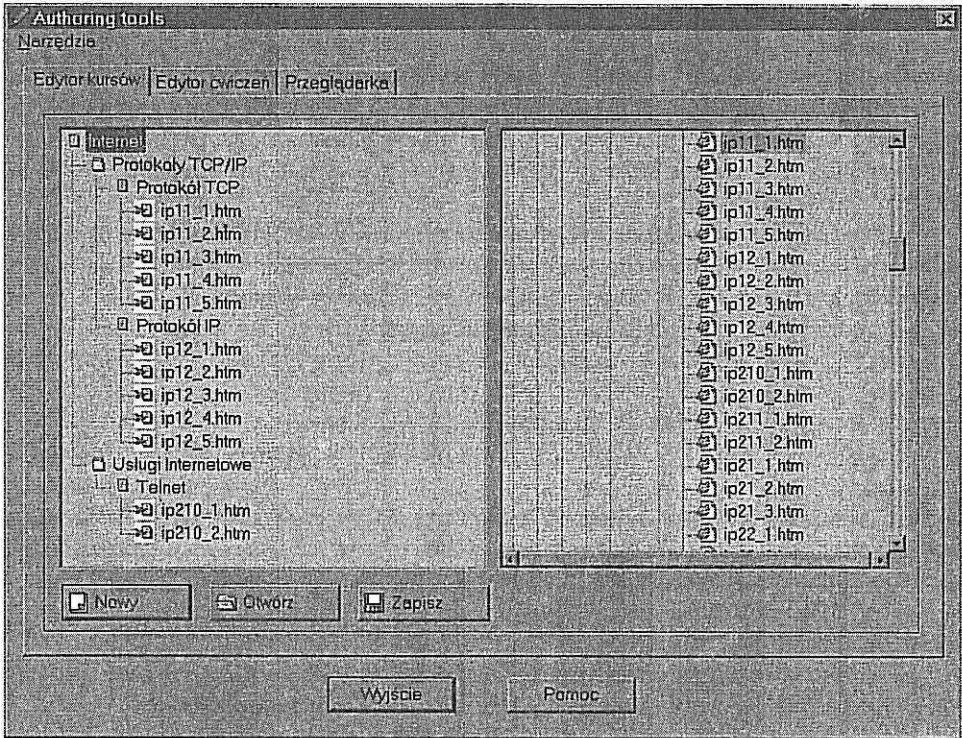
Narzędzie autorskie składa się z dwóch modułów: modułu edycji zawartości kursów i modułu edycji (tworzenia) pytań. Pierwszy moduł jest odpowiedzialny za umieszczanie materiałów do kursu w odpowiedniej strukturze danych. Tak przygotowane materiały mogą być wykorzystane w dowolnym VLE, zgodnym ze standardem opakowywania materiałów IMS Content Packaging<sup>7</sup>. Druga część narzędzia autorskiego pozwala na konstruowanie różnych typów pytań egzaminacyjnych i testowych. Po utworzeniu są one zapisywane w skrypcie instalacyjnym, zgodnym ze standardem IMS Question and Testing Interoperability. Informacje zapisywane są w plikach w języku XML.

Poniższy rysunek przedstawia przykładowe okno modułu, umożliwiającego tworzenie struktury kursu. Tworzenie struktury odbywa się w lewej części okna poprzez dodawanie nowych poziomów w hierarchii kursu. Łączenie zawartości kursu, czyli stron HTML, ze strukturą kursu odbywa się poprzez przeciągnięcie odpowiedniej strony HTML z okna prawego do odpowiedniej sekcji w oknie lewym. Istnieje również możliwość manipulowania stronami w samej hierarchii kursu poprzez ich kasowanie czy też przeniesienie w inne miejsce, do innej sekcji. W dowolne miejsce materiałów można wstawić pytania testowe, a na koniec — pytania egzaminacyjne utworzone w drugim module narzędzia autorskiego.

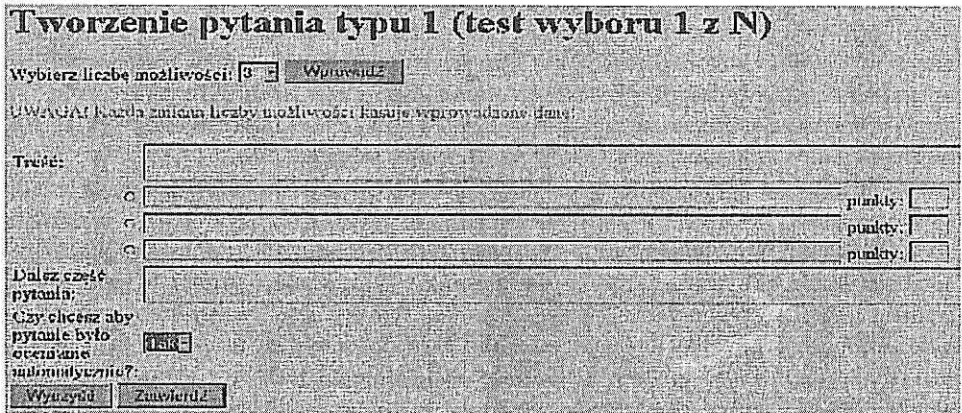
Trwają także prace nad umożliwieniem autorowi kursu, korzystającemu z narzędzia autorskiego, przeglądania i tworzenia zawartości kursu w trybie WYSIWYG (na powyż-

<sup>7</sup> <http://www.imsproject.org/specifications.html>.





Ryc. 5.1. Przykładowe okno programu autorskiego



Ryc. 5.2. Przykład tworzenia pytania typu „radiobutton”

szym rysunku zakładka o nazwie „Przeglądarka”). Do tej pory każdy z autorów musiał samodzielnie, przy wykorzystaniu dowolnego edytora HTML, tworzyć strony z materiałami kursowymi.

Drugi moduł umożliwia tworzenie różnego typu pytań, zarówno testowych, jak i egzaminacyjnych. W systemie przewidziano 8 różnych typów pytań: testy wyboru, listy rozwijane, listy wyboru wielu wartości, tekstowe pola edycyjne, tekstowe pola edycyjne z predefiniowaną wartością, pole opisowe (typu *text area*) i pytania typu przeciągnij i upuść (ang. *drag & drop*). Poniższy rysunek przedstawia przykład tworzenia pytania typu: test wyboru 1 z N.

Pierwszym krokiem przy tworzeniu pytania typu: test wyboru 1 z N jest wybór liczby możliwych odpowiedzi. Następnie należy wpisać treść pytania (pole „Treść”), treści poszczególnych opcji do wyboru (pole edycyjne poniżej pola „Treść”) i można dodać dalszą treść pytania, która powinna ukazać się na ekranie po opcjach wyboru (pole „Dalsza część pytania”). Przy każdej opcji wyboru należy wprowadzić odpowiednią liczbę punktów możliwych do uzyskania w przypadku wybrania danej opcji podczas odpowiedzi (pola „Punkty”). Ostatnią czynnością jest określenie, czy dane pytanie ma być ocenianie automatycznie przez system czy odpowiedź ma być przesłana do oceny przez prowadzącego kurs (ang. *tutor*) (pole „Czy chcesz, aby pytanie było oceniane automatycznie”). Po wprowadzeniu pytania i możliwych odpowiedzi wraz z odpowiadającą im liczbą punktów należy zatwierdzić pytanie, klikając przycisk „zatwierdź”.

## 6. SYSTEM KONTROLI I REAKCJI NA POSTĘPY W NAUCE DOKONYWANE PRZEZ STUDENTA

System kontroli i reakcji na postępy w nauce, dokonywane przez studenta podczas przyswajania materiału, realizowany jest poprzez badanie odpowiedzi na pytania testowe umieszczane po każdej większej części kursu. Każdemu pytaniu testowemu zostaje przypisany poziom trudności na podstawie liczby punktów, przydzielonych przez autora pytania. Dalsza ewaluacja poziomu trudności pytania następuje na podstawie odpowiedzi studentów po każdym zakończonym kursie. Zmiana poziomu pytania zależna jest od wyliczonego współczynnika *CW* i średniego czasu odpowiedzi studentów na pytanie (patrz ryc. 4.2.). Liczbę obszarów, do których może trafić pytanie, można zmieniać dowolnie, dzieląc każdy z obszarów: A, B, C i D na cztery podobszary, a te z kolei na następne cztery itd. Poziom trudności pytania (pytanie najłatwiejsze to poziom 1, trudniejsze to poziom 2 itd.) zależy od obszaru, do którego zakwalifikowano pytanie, i jest określany zgodnie ze schematem przedstawionym na poniższym rysunku.

Początkowy podział zakładał 4 poziomy (na ryc. cyfry bez nawiasów) ponumerowane odpowiednio: 1 (lewy górny obszar), 2 (prawy górny obszar), 3 (lewy dolny obszar), 4 (prawy dolny obszar). Po podzieleniu obszaru 1 na dodatkowe poziomy trudności pytań zmieniła się liczba i numeracja poziomów. Obszar 1 został podzielony, zgodnie ze schematem początkowego podziału, na cztery dodatkowe poziomy, które zastąpiły cały ten obszar. W systemie możliwy jest trzykrotny podział każdego z pierwotnych czterech obszarów.

Po przypisaniu pytaniom poszczególnych poziomów trudności wybieramy te ze środkowymi wartościami (w naszym przypadku pytania z poziomem 4) jako pytania testowe dla studenta na początku kursu. Następnie system rejestruje odpowiedzi studenta

(1)	(2)	2 (5)
1		
(3)	(4)	
3 (6)		4 (7)

Ryc. 6.1. Przykład klasyfikacji pytań do odpowiednich poziomów trudności

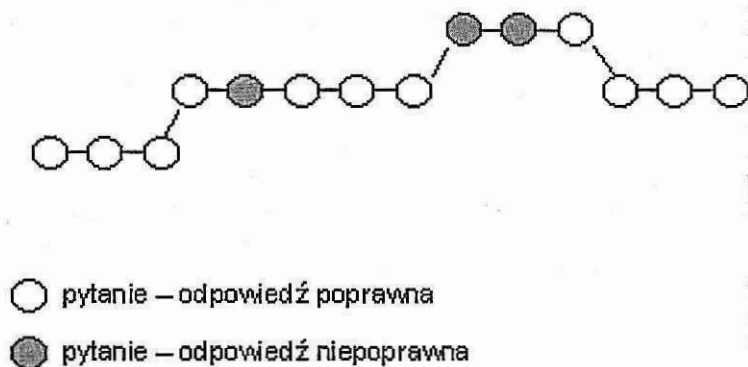
na poszczególne pytania. W przypadku stwierdzenia na przykład trzech niepoprawnych odpowiedzi następujących po sobie system obniża poziom trudności zadawanych pytań o 1. W przypadku trzech poprawnych odpowiedzi następujących po sobie system podnosi poziom trudności zadawanych pytań. Liczbę odpowiedzi następujących po sobie czy też kombinacje odpowiedzi, po których ma nastąpić zmiana poziomu pytania, ustala prowadzący zajęcia przed rozpoczęciem kursu, korzystając z funkcji systemu. Taki system pozwala na dostosowanie poziomu trudności pytań do indywidualnych możliwości studenta określonych na podstawie wiedzy przyswojonej przez studenta po przeczytaniu materiałów

poziom pytań 6

poziom pytań 5

poziom pytań 4

poziom pytań 3



Ryc. 6.2. Przykład kontroli systemu nad przyswajaniem wiedzy przez studenta podczas zdalnego kursu

kursowych. Ponadto system ten pozwala informować studenta w sposób ciągły o postępach w nauce lub ich braku w zależności od poziomu pytań, na jakie aktualnie odpowiada i chwalić go za te postępy bądź też mobilizować do dalszej intensywniejszej pracy w przypadku, gdy poziom pytań, na które student odpowiada, jest niski. Ryc. 6.2. przedstawia w formie graficznej przykładowy sposób działania takiego systemu.

Co pewien czas (po losowej liczbie udzielonych odpowiedzi bądź też w ściśle określonych miejscach kursu) wysyłane są do studenta komunikaty odpowiedniej treści, informujące go o postępach w nauce. Treści komunikatów zależne są od poziomu pytań, na jakie odpowiada student, i mogą być tworzone przez prowadzącego zajęcia.

## 7. WZORY MATEMATYCZNE I OCENA POPRAWNOŚCI ROZWIĄZYWANIA ZADAŃ MATEMATYCZNYCH

Jako przykład możliwości technologii przedstawiamy wykorzystanie do tworzenia pytań symboli lub grafiki, a nie tylko tekstu. Wzory matematyczne do tej pory przedstawiane były na stronach WWW w sposób statyczny w postaci obrazów graficznych. W naszym systemie udostępniamy narzędzie pozwalające studentowi na manipulowanie wzorami zapisanymi w formacie MathML<sup>8</sup> z poziomu przeglądarki WWW. Całe narzędzie podzielone jest na kilka modułów, realizujących różne funkcjonalności. Do modułów tych zaliczamy:

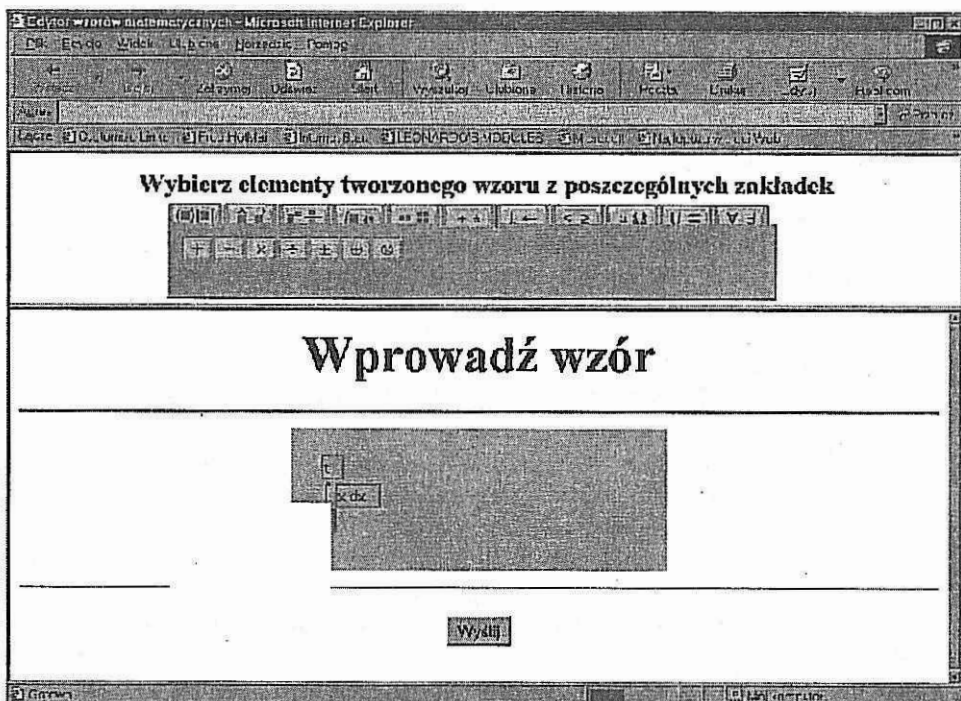
- moduł odpowiedzialny za wyświetlanie wzorów;
- moduł realizujący funkcje edytora równań (wykorzystanie zbioru predefiniowanych symboli matematycznych);
- moduł realizujący funkcje interakcji z użytkownikiem (albo przez wysłanie i odebranie danych z serwera, albo poprzez wykorzystanie języka JavaScript).

Ryc. 7.1. przedstawia przykładowe okno aplikacji realizującej funkcje edytora równań. W prawej ramce mamy do dyspozycji zbiór predefiniowanych symboli matematycznych w postaci obrazków bitowych. Po naprowadzeniu kursora na wybrany element zostaje on podświetlony. W lewej ramce przedstawione jest pole edycyjne, w którym umieszczane są wybrane poprzez kliknięcie na nie symbole matematyczne z lewej ramki. Istnieje również możliwość dokonywania zmian w już wprowadzonym wzorze w ramce lewej w polu edycyjnym wzoru. Aby dokonać zmiany, np. usunięcia elementu, należy najpierw zaznaczyć interesujący nas symbol matematyczny, a następnie, klikając na niego, dokonać odpowiedniej zmiany (wybór akcji, jaką należy wykonać, dokonywany jest z menu ukazującego się po kliknięciu wybranego symbolu matematycznego).

Za pomocą tego narzędzia możliwe jest m.in.:

- wyświetlanie wzorów matematycznych na stronie WWW;
- tworzenie pytań typu: jaka potęga powinna znaleźć się we wzorze przy symbolu np.  $c$ ;
- wyświetlenie np. układu równań liniowych do rozwiązania i prześledzenie poszczególnych kroków rozwiązania zaproponowanego przez studenta wraz z wynikiem;

<sup>8</sup> Mathematical Markup Language (MathML). Specyfikacja pod adresem internetowym: <http://www.w3.org/math/>.



Ryc. 7.1. Przykład modułu realizującego funkcje edytora równań

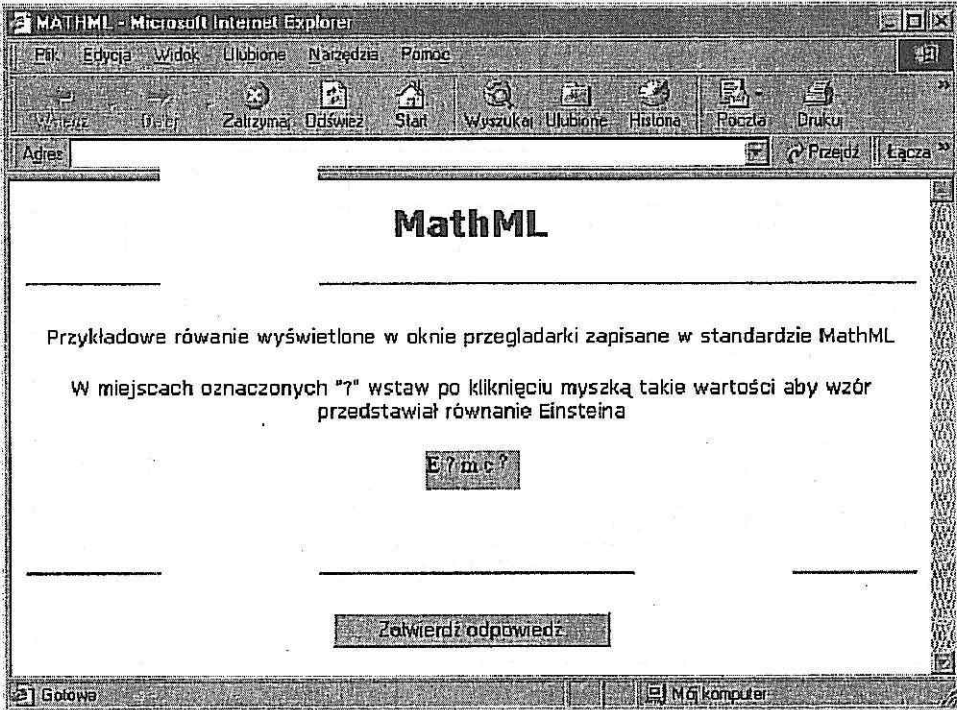
- wyświetlenie równania matematycznego z możliwością zmiany pewnych jego parametrów i umożliwienie studentowi obserwowania, jak zmiany tych parametrów wpływają na wynik lub wykres;
- umożliwienie prześledzenia poszczególnych kroków rozwiązania zadania matematycznego poprzez np. klikanie klawisza „następny krok”;
- zadawanie pytań testowych w postaci zadań matematycznych, w których po kliknięciu na odpowiedni element pojawi się na ekranie poprawna odpowiedź.

Innym przykładem możliwości narzędzia jest ćwiczenie wykorzystujące język MathML, przedstawione na ryc. 7.2.

Zadaniem studenta w tym ćwiczeniu jest zastąpienie znaków „?” we wzorze odpowiednimi symbolami, tak aby cały wzór przedstawiał równanie Einsteina. Po kliknięciu myszką na znak „?” wyświetlone zostaje okienko, w którym student może wpisać swoją odpowiedź. Po uzupełnieniu całego równania i kliknięciu przycisku „zatwierdź odpowiedź” odpowiedź studenta zostaje przesłana w języku MathML do serwera, gdzie jest porównywana z odpowiedzią poprawną.

Wymiana danych pomiędzy klientem i serwerem, a także sprawdzanie odpowiedzi z odpowiedzią, zapisane w bazie danych, następuje przy wykorzystaniu standardu MathML. Jest to standard konsorcjum W3C, używany do zapisu wzorów matematycznych pod kątem zarówno ich wizualizacji (np.  $\sqrt{2}$ ), jak i matematycznej zawartości wykorzystywanej do





Ryc. 7.2. Edycja wzoru matematycznego połączona z oceną

obliczeń (zawartość określa czy zapis  $V^2$  oznacza wektor, macierz czy liczbę). MathML jest standardem opartym na DOM-API<sup>9</sup> i bazuje na standardzie XML<sup>10</sup>. DOM API jest to interfejs niezależny od platformy sprzętowej i języka programowania, który pozwala programom i skryptom na dynamiczny dostęp i zmianę zawartości, struktury i wyglądu dokumentu.

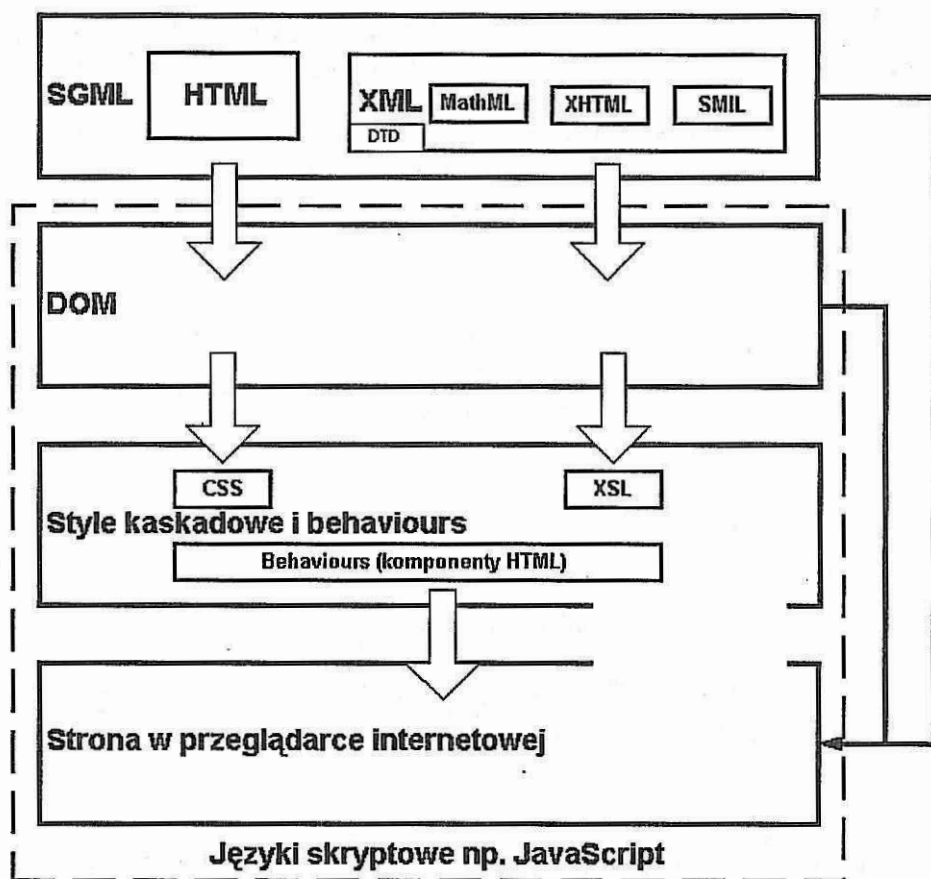
## 8. STANDARDY ORGANIZACJI W3C

Standardy wymienione w poprzednim rozdziale mogą być wykorzystane do budowy powiązanych struktur materiałów.

Ryc. 8.1. przedstawia w sposób uproszczony podstawowe relacje pomiędzy poszczególnymi standardami organizacji W3C. Języki HTML i XML są oparte na języku SGML, stanowią one jednak osobne języki programowania. Podstawową różnicą pomiędzy HTML

<sup>9</sup> Document Object Model (DOM); Specyfikacja pod adresem internetowym: <http://www.w3.org/DOM>.

<sup>10</sup> Extensible Markup Language (XML). Specyfikacja pod adresem internetowym: <http://www.w3.org/XML>.



Ryc. 8.1. Miejsce standardów W3C w architekturze informatycznej prezentacji i tworzenia

i XML jest odseparowanie w języku XML treści od ich formy prezentacji, np. w przeglądarce. Ponadto XML w przeciwieństwie do języka HTML daje możliwość tworzenia swoich własnych znaczników. Dokument napisany w języku XML może być tzw. *well-formed* lub *valid*. Dokument „well-formed” może zawierać dowolnie znaczniki, ale musi być napisany zgodnie z podstawowymi regułami zapisu w języku XML, np. każdy znacznik otwierający musi mieć odpowiadający mu znacznik zamykający. Dokument „valid” natomiast zawiera dowolnie zdefiniowaną strukturę znaczników w pliku DTD (typ definicji danych), z którą tworzony dokument XML musi być zgodny.

Na bazie języka XML tworzonych jest wiele nowych języków, jak np. MathML (do opisu równań), XHTML (odpowiednik języka HTML bazujący na XML) czy SMIL (język prezentacji multimedialnych) i wiele innych.

DOM jest to niezależny od języka i systemu operacyjnego interfejs umożliwiający manipulowanie (interakcję) obiektami stron napisanych w języku HTML lub XML. Inter-



fejs ten może być zaimplementowany w dowolnym języku programowania. Wyróżnia się różne poziomy (ang. *level*) interfejsu DOM. Poziom 1 odpowiedzialny jest za nawigację i manipulowanie dokumentami HTML i XML. Poziom 2 dostarcza model obiektowy dla arkuszy stylów, pozwalający na manipulowanie właściwościami stylów przypisanych do dokumentu. Ponadto określa m.in. takie funkcjonalności, jak możliwość przechodzenia po obiektach strony HTML lub XML czy model zdarzeń. Poziom 3, nad którym aktualnie trwają prace, określać ma możliwości otwierania i zapisywania dokumentów z możliwością sprawdzania ich składni, grupowania zdarzeń, formatowania sposobu przedstawiania dokumentu. Dalsze poziomy DOM miałyby specyfikować m.in. zagadnienia związane z bezpieczeństwem, interfejsem do języka zapytań czy sposobami powiadamiania użytkownika o różnego typu zdarzeniach.

Następnym etapem na drodze dokumentu do wyświetlenia go w oknie przeglądarki są style kaskadowe i *behaviours*<sup>11</sup>. Mechanizmy te dostarczają funkcjonalności umożliwiające formatowanie i określanie sposobu wyświetlania dokumentów HTML i XML.

Istnieje możliwość wyświetlania dokumentów HTML i XML bezpośrednio w przeglądarce bez pośrednictwa modelu DOM, a także z pominięciem stylów kaskadowych.

## 9. WNIOSKI

W artykule przedstawiono technologię umożliwiającą ocenę wyników testów przeprowadzonych za pośrednictwem internetu, w których uzyskuje się informacje wychodzące poza możliwości wyboru z wielu odpowiedzi, oraz przeanalizowano dokładniej wyniki takiego testu, przeprowadzonego w ATR w Bydgoszczy z wykorzystaniem systemu ProEDUS.

Z punktu widzenia technicznego technologia ta jest dojrzała, natomiast aby oddać ją w ręce dydaktyków, należy stworzyć narzędzie autorskie do konstrukcji testów, aby mogła z niego korzystać osoba nie będąca zawodowym informatykiem.

W drugiej części pracy przedstawiamy projekt narzędzia autorskiego do tworzenia testów oraz wstępną demonstrację, wychodzącą poza formę tekstową zadawania pytań, np. wymagającą manipulacji na równaniach w przypadku wzorów matematycznych.

Poważnym problemem jest kwestia skali punktacji pytań skorelowanych i pytań typu „checkbox”. Wynika to z faktu, że suma punktów za całą odpowiedź nie jest równa sumie punktów uzyskanych z indywidualnych odpowiedzi, a pytania wnoszą różne wartości do całkowitej odpowiedzi.

Niestety, z powodu występowania pytań skorelowanych i pytań typu „checkbox” zachodzi konieczność zdefiniowania poprawności odpowiedzi na dane pytanie, aby porównywać pytania między sobą.

Powstaje problem takiej konstrukcji oceny, aby przyjęte skale nie naruszały jej rzetelności i trafności.

<sup>11</sup> Behavioral Extension to CSS (HTML Components). Specyfikacja pod adresem internetowym: <http://www.w3.org/TR/becss>.

Technologia tutaj prezentowana wychodzi poza najlepsze produkty rynkowe<sup>12</sup> oraz standard IMS. Można sobie wyobrazić jeszcze bardziej złożone narzędzia oceny, np. umożliwiające przybliżoną ocenę odpowiedzi otwartej na podstawie słów kluczowych z opcją kontekstu skorelowanych z możliwościami zbliżonymi do używanych w wyszukiwarkach. Choć powszechne wykorzystanie takich narzędzi w Polsce nie nastąpi wcześniej niż za 10 lat, upowszechnienie internetu i spadek ceny oraz pojawienie się nowych terminali dostępowych przybliży ten moment.

---

<sup>12</sup> Narzędzie firmy QuestionMark, Perception for Web, [www.Questionmark.com](http://www.Questionmark.com); a w Polsce: [www.premiere.com.pl](http://www.premiere.com.pl).