**Nazwa przedmiotu:**

Inteligentne systemy informacyjne

**Koordynator przedmiotu:**

prof. Mieczysław Muraszkiewicz

**Status przedmiotu:**

Fakultatywny dowolnego wyboru

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Informatyka

**Grupa przedmiotów:**

Przedmioty techniczne - zaawansowane

**Kod przedmiotu:**

ISI

**Semestr nominalny:**

4 / rok ak. 2012/2013

**Liczba punktów ECTS:**

4

**Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się:**

30 h - udział w wykładach,
30 h - przygotowanie do wykładów (korzystając m.in. z slajdów dołączonych do wykładu, wskazanej literatury, wykonanie fakultatywnej pracy domowej, ew. dyskusje z uczestnikami wykładu),
15 h - przygotowanie do egzaminów,
30 h - udział w ćwiczeniach,
60 h - projekt: przygotowanie eksperymentu/kodowanie.

Razem: 30 h + 30 h +15 h + 30 h + 60 h = 165 h

**Liczba punktów ECTS na zajęciach wymagających bezpośredniego udziału nauczycieli akademickich:**

**Język prowadzenia zajęć:**

polski

**Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym:**

**Formy zajęć i ich wymiar w semestrze:**

|  |  |
| --- | --- |
| Wykład:  | 30h |
| Ćwiczenia:  | 30h |
| Laboratorium:  | 0h |
| Projekt:  | 0h |
| Lekcje komputerowe:  | 0h |

**Wymagania wstępne:**

Podstawowe wiadomości z zakresu:
1. Systemów zarządzania bazami danych.
2. Logiki pierwszego rzędu.

**Limit liczby studentów:**

30

**Cel przedmiotu:**

Celem wykładu jest nabycie przez studenta głównych komponentów wiedzy z zakresu metod i technik informatycznych traktowane jako "inteligentne", które znajdują lub mogą znaleźć zastosowanie w projektowaniu i budowaniu systemów informacyjnych. Wykład pomyślany jest jako droga prowadząca od morfologii klasycznego systemu informacyjnego, czyli określenia jego składników z możliwie jak największą dokładnością, a następnie wiodąca przez poszukiwanie odpowiedzi na pytanie: które z poszczególnych składników można wyposażyć w „inteligencję" za pomocą znanych technik z zakresu sztucznej inteligencji i dziedzin pokrewnych, by - w rezultacie - doprowadzić do opracowania mapy architektury generycznego systemu informacyjnego z zaznaczeniem tych miejsc, gdzie, w jaki sposób i w jakim stopniu można wprowadzać do systemu „inteligentne” komponenty.

Celem wykładu jest także praktyczne "zetknięcie się" z "inteligencją" systemów informacyjnych. Dlatego wykładowi towarzyszy projekt, który obejmuje albo zaplanowanie i wykonanie prostego eksperymentu z wykorzystaniem wybranej techniki omawianej na wykładzie lub innej techniki z zakresu sztucznej inteligencji i dziedzin pokrewnych, albo stworzenie prostej aplikacji „inteligentnej”, która mogłaby zostać zintegrowana z klasycznym systemem informacyjnym. Przykładami pro-jektów są: przeprowadzenie eksperymentu eksploracji danych faktograficznych / tekstowych, opracowanie parsera dla pewnego języka atomów semantycznych, implementacja mechanizmu wyszukiwania w sieciach semantycznych pewnej klasy, opracowanie programu do identyfikacji preferencji użytkownika na podstawie analizy jego kwerend. Nie stawia się żadnych wstępnych warunków na narzędzia programistyczne, które zostaną zastosowane do realizacji prac projektowych.

Mottem wykładu jest powiedzenie R. Hamminga: „The purpose of computing is insight, not num-bers”.

**Treści kształcenia:**

Poniżej scharakteryzowano moduły tematyczne wykładu.

Terminy podstawowe. Omówiona zostanie podstawowa terminologia dotycząca inteligencji, systemów informacyjnych oraz sztucznej inteligencji. Podana zostanie szkieletowa definicja systemu informacyjnego jako n-tki uporządkowanej: funkcja celu, model danych, model procesów, baza danych, język interakcji (zwłaszcza wyszukiwanie), sprzęt, oprogramowanie, wydajność, administracja systemu i utrzymanie go w ruchu, mechanizmy ochrony danych i praw dostępu, ludzie (personel, użytkownicy), aspekty finansowe, otoczenie itd. Towarzyszyć temu będzie pierwsza, intuicyjna próba wskazania tych komponentów lub ich składników, gdzie już stosuje się lub możliwe wydaje się zastosowanie technik "inteligentnych".

Relacje. Za istotne w tym wykładzie uważa się możliwie dokładne określenie wzajemnych relacji i wzajemnych wpływów takich dziedzin jak: systemy informacyjne, systemy baz danych, sztuczna inteligencja oraz metody reprezentacji wiedzy. W badaniu tych kwestii pod uwagę będą brane m.in. takie czynniki, jak: dane, model danych; złożone struktury danych; język manipulacji; język wyszukiwania; wnioskowanie; transakcje; optymalizacja; badanie wiarygodności danych; lokalizacja / systemy rozproszone; ochrona danych i dostępu; wydajność. Zostaną wskazane dwa nurty w pracach nad sztuczną inteligencją i dziedzinami pokrewnymi, a mianowicie: (i) wykorzystanie metod "racjonalnych", "sztucznych" (logika, algebra, lingwistyka), oraz (ii) symulacja rozwiązań stworzonych przez "naturę": sieci neuronowe, ewolucja, Darwinowski dobór naturalny.

Wiedza. Zostaną pokazane i przedyskutowane definicje wiedzy używane na gruncie informatyki oraz będzie podjęta próba zdefiniowania na użytek tego wykładu terminu wiedza. Wyjaśnione zostaną powody reprezentowania wiedzy, role jakie spełniają reprezentacje wiedzy i problemy związane z reprezentacją. Punktem kulminacyjnym tej części wykładu będzie pokazanie jak zbudowane są metody reprezentacji wiedzy oraz wymienienie najważniejszych z nich tj. język naturalny, metody stosowane w obszarze baz danych, logika matematyczna (klasyczna, niestandardowa), reguły produkcji (production rules), sieci semantyczne (semantic networks), atomy semantyczne (ang. semantic atoms), ontologie, sieci neuronowe (neural nets), ramy, skrypty (frames, scripts), zbiory przybliżone (rough sets), zbiory rozmyte (ang. fuzzy sets), algorytmy genetyczne.

Logiki klasyczne i nieklasyczne. Zostaną przypomniane najważniejsze terminy logik klasycznych pierwszego rzędu oraz sposoby użycia narzędzi opartych na takich logikach do specyfikacji systemów informacyjnych oraz opisu języków kwerend (SQL) i budowania kodu (Prolog). Wskazane zostaną ograniczenia tego rodzaju narzędzi (twierdzenie Goedla, duża złożoność obliczeniowa, trudności w reprezentacji informacji niepewnej i niekompletnej). Mimo jej ograniczeń logice przypisuje się szczególnie duże, wręcz paradygmatyczne znaczenie ze względu na jej precyzję i walory edukacyjne - w wykładzie jest ona punktem odniesienia i porównania z innymi metodami. Wyjaśniony zostanie termin monotoniczność logiki i jego znaczenie dla systemów informacyjnych korzystających ze środków logicznych. W tej części wykładu wprowadzone zostaną również główne pojęcia dotyczące logik nieklasycznych oraz powody ich zastosowania w systemach informacyjnych. Studenci dowiedzą się także czym są definitywne i niedefinitywne systemy baz danych, poznają znaczenie założenia o zamkniętości / otwartości świata (ang. Close/Open World Assumption). Tę część wykładu zamkną informacje o Reitera logice defoltów, systemach weryfikacji przekonań (ang. Truth Maintenance System), Pawlaka zbiorach przybliżonych i Zadeha zbiorach rozmytych.

Wybrane metody reprezentacji wiedzy

Reguły produkcji, czyli napisy o konstrukcji jeśli - to są prostą metodą reprezentacji wiedzy o sporym jednak potencjale deskrypcyjnym i inferencyjnym; są one często stosowane w systemach doradczych. Wyraźnie zostanie podkreślona różnica między reprezentacją za pomocą lo-gik pierwszego rzędu i reguł produkcji. Reguły produkcji są przydatne przy wspieraniu wyszukiwania i budowaniu heurystyk wyszukiwawczych.

Sieci semantyczne zostaną przedstawione jako typowy produkt strukturalistycznego podejścia do zagadnień technicznych, gdzie semantykę obiektu określają wzajemne relacje istniejące po-między tym obiektem i innymi obiektami. Sieci są grafami (zbliżonymi do Quillianowskiego modelu pamięci), w których wnioskowanie jest szczególnie ułatwione (choć nie wolne od pułapek), gdyż ścieżka inferencyjna jest po prostu konkatenacją tych łuków (krawędzi) grafu, które wiodą od przesłanek do tezy. Zastosowania sieci semantycznych w kontekście systemów informacyj-nych są różnorodne, a w tym do: konceptualnego opisu projektowanych systemów informacyjnych, rozumienia języka naturalnego, rozpoznawania i rozumienia mowy, budowy interfejsów do systemów baz danych opartych na języku naturalnym, budowania systemów odpowiadających na pytania (np. SIR).

Atomy semantyczne oparte na Demokryta i Russela poglądach atomistycznych są próbą - i tak zostaną przedstawione - opisu złożonych semantycznie obiektów za pomocą niewielkiego zbioru (kilkunastu lub kilkudziesięciu) tzw. prymitywów semantycznych. Zastosowanie atomów se-mantycznych polega zatem na tłumaczeniu zdań z danego języka na język korzystający wyłącznie z atomów. W podejściu tym za wiedzę uznaje się wszystko to, co można wyrazić w jakimś języku (zgodnie z przekonaniem Wittgensteina: „Die Grenzen der Sprache sind die Grenzen der Welt”) - a po dokonaniu tłumaczenia - w języku opartym na atomach semantycznych. Na marginesie zostanie skomentowane ciekawa opinia Wittgenstaina, że: „"Einen Satz verste-hen, heisst, eine Sprache verstehen. Eine Sprache verstehen, heisst eine Technik be-herrschen". Tę część wykładu zamknie krótka prezentacja Schanka ontologii zdarzeń, Concep-tual Dependency, oraz przykłady zestawów atomów semantycznych za Wilksem i Wierzbicką.

Ramy (Minsky, 1975) i skrypty (Schank i Abelson, 1977) nawiązują do znanych z psychologii poznawczej obserwacji, że bieżące zachowania są oparte na stereotypowych sekwencjach zdarzeń, które utworzono na podstawie doświadczeń wyniesionych z przeszłości. Rama będzie przedstawiona jako zbiór tzw. slotów (ang. slots), które zawierają deklaratywną lub procedural-ną informację, przy czym ramy zwykle zorganizowane są w struktury hierarchiczne (z dziedziczeniem własności). Ramy i skrypty oprócz wiedzy zawierają w sobie mechanizmy inferencyj-ne. Zostanie zwrócona uwaga na to, że stosowane obecnie coraz częściej w różnych obszarach infromatyki podejście obiektowe pod wieloma względami nawiązuje do koncepcji ram i skryptów. Jako ciekawe zastosowania tej koncepcji będą wspomniane systemy KRL i Parmenides.

Sieci neuronowe i algorytmy genetyczne reprezentują grupę metod i technik, które w jakimś stopniu są symulacją rozwiązań stworzonych przez „naturę”. Zostanie podkreślona zdolność uczenia się, odporność na uszkodzenia oraz możliwość przetwarzania równoległego programów opartych na tych metodach. W odniesieniu do sieci neuronowych szczególnie dokładnie zosta-nie omówiony model ADALINE (Adaptative Linear Element). Algorytmy genetyczne będą zapre-zentowane w wersji z dwoma operatorami genetycznymi tzn. krzyżowaniem i mutacją.

Ontologie jako formalna reprezentacja pewnej dziedziny wiedzy, na którą składa się zapis zbiorów pojęć (ang. concept) i relacji między nimi. Zapis ten tworzy schemat pojęciowy, który będąc opisem danej dziedziny wiedzy, może służyć jednocześnie jako podstawa do wnioskowania o właściwości opisywanych ontologią pojęć. Ontologie oprócz „statycznej” wiedzy odziedzinie mo-gą zawierać aspekty dynamiczne (behawioralne).

Eksploracja danych, odkrywanie wiedzy. W tej część wykładu studenci zostaną wprowadzeni w problematykę eksploracji danych, rozumianą jako odkrywanie z dostępnych zasobów danych różnego rodzaju uogólnień, regularności, prawidłowości, reguł, a zatem czegoś, co stanowi pewną wiedzę zawartą implicite w tych zasobach. Eksploracja może być traktowana jako rozszerze-nie języka zapytań. Zostaną pokazane metody oparte na klasyfikacji, regresji, klasteringu, koja-rzeniu i regułach epizodycznych.

**Metody oceny:**

Na ocenę końcowa składają się trzy komponenty:
- ocena za pracę nad projektem,
- ocena z egzaminu pisemnego,
- ocena z egzaminu ustnego.
Przed przystąpieniem do egzaminu ustnego wyliczna jest ocena wstępna jako następująca średnia arytmetyczna:
 ((ocena z projektu -1) + ocena z egzaminu pisemnego))/2
Jeśli student zaakceptuję średnią arytmetyczną, to egzamin ustny jest pomijany i zaakceptowana ocena jest wpisywana do indeksu.

**Egzamin:**

tak

**Literatura:**

Literatura główna:
[1] Brachman R., Levesque H., Knowledge Representation and Reasoning, The Morgan Kaufmann Series in Artificial Intelligence, 2004.
Literatura dodatkowa:
[2] Barr A., Feigenbaum E. A., The Handbook of Artificial Intelligence, vol. I, II, III, William Kaufmann Inc, 1981.
[3] Gasevic D., Djuric D, Devedzic V, Model Driven Architecture and Ontology Development, Springer, 2006.
[4] van Harmelen F., Lifschitz V., Porter B. (editors), Handbook of Knowledge Representation (Foundations of Artificial Intelligence) (Foundations of Artificial Intelligence), Elsevier Science, 2007.
[5] Xuan F. Zha (editor), Artificial Intelligence and Integrated Intelligent Information Systems: Emerging Technologies and Applications, IGI Global, 2006.

**Witryna www przedmiotu:**

https://studia.elka.pw.edu.pl/pl

**Uwagi:**

Wykład prowadzony jest w formule konwersatorium tzn. studenci mogą w trakcie wykładu zadawać pytania i/lub wypowiadać własne komentarze.
Studentów zachęca się do składania propozycji co do ulepszenia formy i treści wykładu.

## Efekty przedmiotowe

### Profil ogólnoakademicki - wiedza

**Efekt W1:**

Student, który zdał egzamin ma wiedzę z zakresu podstaw inteligentnych systemów informacyjnych opartą na pogłębionej znajomości zakresu znaczeniowego terminów podstawowych (inteligencja, informacja, wiedza, systemy informacyjne) oraz poznaniu głównych metodami reprezentacji wiedzy wykorzystywanych w inteligentnych systemach informacyjnych.

Weryfikacja:

Ocena projektu i egzamin pisemny oraz ew. egzamin ustny

**Powiązane efekty kierunkowe:** K\_W01, K\_W11

**Powiązane efekty obszarowe:** T2A\_W01, T2A\_W03, T2A\_W04, T2A\_W07

### Profil ogólnoakademicki - umiejętności

**Efekt U1:**

Po zakończeniu kursu student umie wybrać i uzasadnić wybór metod(y) reprezentacji wiedzy dla projektowanego inteligentnego systemu informacyjnego

Weryfikacja:

Ocena projektu i egzamin pisemny oraz ew. egzamin ustny

**Powiązane efekty kierunkowe:** K\_U01, K\_U04, K\_U06, K\_U09, K\_U10

**Powiązane efekty obszarowe:** T2A\_U01, T2A\_U05, T2A\_U08, T2A\_U09, T2A\_U11, T2A\_U15