Студијски програм: Докторске студије информатике

Назив предмета: Р465 - Обрада природних језика

Наставник: Душко Витас и други наставници Катедре за рачунарство и информатику

Статус предмета: Изборни

Број ЕСПБ: 15

Услов: Р415

Циљ предмета: Циљ курса је преглед основних метода у обради природних језика, заснованих на моделирању знања о језику и на статистичким моделима природних језика. Курс уводи студента у основне проблеме анализе природних језика и методе њиховог решавања. Илуструје се на материјалу српског језика.

Исход предмета: Студент познаје основне проблеме анализе природних језика и методе њиховог решавања, са акценотом на обради српског језика.

Садржај предмета: У оквиру предмета се обрађују следеће теме: Сегментација текста; препознавање речи и препознавање реченица. Морфологија– препознавање и генерисање; коришћење коначних аутомата и лексикона, комбиноване методе. Статистички модели. Етикетирање врсте речи; избор скупа етикета; методе засноване на правилима, методе засноване на трансформацијама; отклањање вишезначости – статистичке методе и методе засноване на правилима. Синтаксичка анализа; анализа контекстно слободним граматикама – анализа надоле и нагоре; плитко парсирање и коначни аутомати; формализам заснован на ограничењима - структуре својстава (“features”) и унификација. Специфични проблеми анализе; одређивање антецедента, координација, конгруенција. Све теме се илуструју примерима примене на српски језик.

Литература:

1. Ruslan Mitkov, The Oxford Handbook of Computational Linguistics, Oxford University Press, 2003,
2. Daniel Jurafsky & James H. Martin, Speech and Language Processing, Prentic Hall, New Jersey, 2000
3. S. Nirenburg et al: Ontological Semantics, MIT Press, 2004
4. Kenneth R. Beesley and Lauri Karttunen: Finite State Morphology, CSLI Publications, Stanford Univ. 2003
(наставник може изабрати другу одговарајућу актуелну литературу)

Бр. час. акт. наставе: 10	**Теоријска настава: 4**	**Прак. настава: -**	**Лаб.вежбе: -**	**СИР: 6**

Методе извођења наставе: Фронтални, групни, индивидуални и практични.

Оцена знања (максималан број поена је 100)

Предиспитне обавезе	**поена**	**Завршни испит**	**поена**
активност у току предавања	-	писмени испит	50
практична настава	-	усмени испит	-
колоквијум-и	-	писмено-усмени испит	-
семинар-и	50		