

PPKE BTK / PPKE ITK

**A komplex PhD-vizsga nyelvtechnológiai tételsora**

2017/2018

1. Az n-gramok mibenléte. n-gram-alapú statisztikai megfigyelések. Az n-gramok használata a nyelvtechnológia különböző területein.
2. Véges automaták, véges fordítók és kiterjesztéseik. Reguláris kifejezések. Műveletek automatákkal. Az elemzés és a generálás hasonlósága és különbsége (a nyelvi jelenségek és a lexikon szerepe). A kétszintes leírás alapelvei.
3. Természetes nyelvek mondatszerkezetének ábrázolása formális nyelvtanokkal. Függőség és összetevős szerkezet: hasonlóságok és különbségek. Szintaktikai elemzési algoritmusok. A gépi és az emberi mondatelemzés összevetése.
4. Az unifikáció és az unifikálhatóság fogalma. Jegyszerkezetek és kapcsolatuk az irányított körmentes gráfokkal. Példák az unifikáció alkalmazására a morfológiában és a szintaxisban.
5. Lexikai és mondatjelentés-reprezentációk. Fogalmi hálók, ontológiák. A WordNet és továbbfejlesztései. Szóbeágyazási modellek.
6. A gépi fordítás alapvető módszerei. Szabály-alapú közelítések. Párhuzamos korpuszok és felhasználásuk. A statisztikai gépi fordítás alapjai. A neurális fordítás alapgondolata.