

Přednášky pronesené na 12. valném shromáždění České republiky v Karolinu
dne 15. května 2006

Eva Hajičová, MFF UK v Praze
„Jazyky a počítač: obtíže a možnosti“

Myšlenka na využití počítačů pro automatické zpracování textů se vynořila především ve spojení s aplikací na překlad z jednoho jazyka do druhého; již v druhé polovině čtyřicátých let se Warren Weaver v dopise zakladateli kybernetiky Norbertu Wienerovi zmiňuje, že na překlad je možné se dívat jako na dekodování a odkazuje na úspěšné válečné zkušenosti při dešifrování. První pokusy, přeceňované žurnalisty a opatrně přijímané vědci, však nevedly dost rychle k žádanému cíli. Hledaly se cesty, jak přirozený jazyk s jeho bohatým slovníkem i frazeologií a se složitou gramatikou a slovosledem uchopit v podobě, s níž by počítač mohl pracovat. To dalo podnět ke vzniku nového, po výtce interdisciplinárního oboru, počítačové (komputační) lingvistiky: komputační lingvistika se snaží dělat to, co dělá lingvistika, ovšem „komputačním“ způsobem. Uplatnění počítačů vyžaduje přesný a úplný, explicitní (formální) popis jazyka nebo jeho příslušné části. To se týká jak základního výzkumu, tak i aplikací. Strojový překlad je ovšem jen jednou z nich, a textový vstup či výstup jen jednou jazykovou formou, vedle jazyka mluveného. V přednášce budeme ilustrovat jednu z hlavních překážek pro to, aby počítač „porozuměl“ přirozenému jazyku, totiž víceznačnost jazyka, a zmíníme se i o možnosti, jak lze ke složitosti jazyka přistupovat z pohledu (relativně jednoduchého) jádra a (složitě, vrstevnatě) periferie. Ukážeme rovněž, jaké nové přístupy k počítačovému zpracování jazyka se rýsují. V závěru doložíme na příkladu mezinárodního projektu MALACH, jaké možnosti se nabízejí, přistupuje-li se k problémům počítačového zpracování multilingválních promluv (jde o videonahrávky svědectví přeživších, zachránců a svědků holocaustu) komplexně a s realistickým očekáváním, a jak teorie a aplikace na sebe mohou vzájemně plodně působit.