

SMIEŠKOVÁ, E.: Augmentatíva a ich miesto v slovníku spisovného jazyka. In: Slovenská reč, 21, 1956, s. 79–81.

Sprachwissenschaftliches Lexikon. Rudi Conrad. Leipzig 1988.

Wörterbuch grammatischer Termini. Spiewok und Kol. Greifswald 1976.

Problémovo orientovaná analýza prirodzeného jazyka

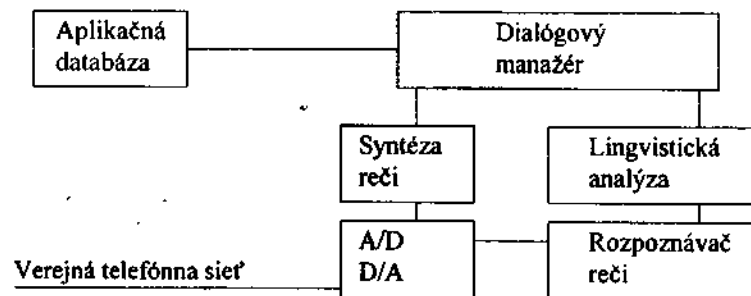
Jozef Ivanecký

Fakulta elektrotechniky a informatiky TU, Košice

Príspevok predstavuje časť vlakového informačného systému, ktorý prostredníctvom verejnej telefónnej linky poskytuje informácie o cestovnom poriadku. Vyvíja sa v rámci projektu Copernikus v spolupráci s univerzitami v Erlangene, Lubľane a Plzni. Výsledkom má byť viacjazyčný systém, ktorý poskytuje informácie o príchodoch a odchodoch vlakov v železničnej sieti zúčastnených krajín. Komunikácia by mala prebiehať automaticky v jazyku používateľa. V nasledujúcej časti v stručnosti predstavíme architektúru celého systému, ďalej sa podrobnejšie zameriame na lingvistickú analýzu, na problémy spojené s jej návrhom a na príklady s ukázkami riešení.

Architektúra systému

Štruktúru celého systému ukazuje obr. 1 (Nöth, 1995). Verejný telefón je pripojený na AD/DA prevodník, ktorý je súčasťou počítača. Aplikačná databáza sa môže nachádzať aj na inom, vzdialenejšom počítači pripojenom do siete a zodpovedá formalizovanému cestovnému poriadku železničnej siete.



Obr. 1. Bloková štruktúra systému

Rozpoznávač reči

Vstupný rečový signál z telefónu sa vzorkuje AD prevodníkom a po spracovaní sa privádza na vstup samotného rozpoznávača. Na rozpoznávanie bol použitý systém ISADORA (Schukat–Talamazzini, 1993, s. 19–36) vytvorený na univerzite v Erlangene, ktorý pracuje na princípe skrytých markovských modelov. Výstupom z rozpoznávača je veta rozpoznaná s najväčšou pravdepodobnosťou, ktorá tvorí vstup pre *lingvistický analyzátor*.

Lingvistický analyzátor

V tejto časti systému, ktorá je podrobnejšie opísaná nižšie, sa vykonáva analýza s cieľom získať zo vstupnej vety základné relevantné informácie (napr.: cieľová stanica, čas odchodu, čas príchodu a podobne) a vytvoriť z nich SIL (Semantic Interpretation Language) objekty, ktoré sú vstupom do *dialógového manažéra*.

Dialógový manažér

Úlohou dialógového manažéra je riadenie dialógu s pýtajúcim sa, s používateľom. Ak sa z prvej vety nezískajú všetky údaje potrebné na vyhľadanie požadovanej informácie, dialógový manažér generuje otázky vo forme SIL objektov. Tie sa *syntézou* prevádzajú na výstupné vety pre používateľa. Dialóg prebieha dovtedy, kým nie sú získané všetky potrebné údaje alebo dialóg nie je používateľom ukončený.

Po získaní potrebných informácií od používateľa sa v databáze vyhľadá príslušné vlakové spojenie. Dialógový manažér bol tiež vyvinutý na univerzite v Erlangene. Keďže pracuje nad SIL objektami, je jazykovo nezávislý.

Syntéza reči

Výstup z dialógového manažéra, ktorý je v tvare SIL objektov, je potrebné previesť do prirodzeného jazyka. To je úlohou bloku syntézy. Ku každému možnému SIL objektu, vystupujúcemu z dialógového manažéra, je priradená skupina možných výstupných viet. Z nich sa náhodným výberom vyberie jedna a po transformácii na číslcovú formu sa vyšle na vstup D/A prevodníka pripojeného na telefónny prístroj.

Lingvistická analýza rečového prejavu

Vo väčšine systémov zaoberajúcich sa sémantickou analýzou sa ako vstup do systému berie písaný text. Vtedy je možné použiť na analýzu všetky bežne známe algoritmy využívajúce skupinu prepisovacích pravidiel, pomocou ktorých sú popísané všetky gramaticky správne vety. Samotná analýza spočíva v prehľadávaní stromov, ktoré je možné na základe daných prepisovacích pravidiel generovať.