

# Uporaba glagolskih predlog pri strojnem prevajanju

**Peter Holozan**

Amebis d.o.o.  
Bakovnik 3, 1241 Kamnik, Slovenija  
peter.holozan@amebis.si

Prevajalni sistem Presis je namenjen strojnemu prevajanju besedil, za zdaj sta izdelana prevajalnika med slovenščino in angleščino in obratno. Po začetni fazi, ko je bilo prevajanje večinoma na nivoju besed z nekaj osnovne stavčne analize, je Presis zdaj prešel na uporabo glagolskih predlog pri stavčni analizi in generaciji prevodov.

Predloga opiše, kako se določen glagol uporablja. Pove, s katerimi predmeti se veže, katere omejitve so pri tem, katera prislovna določila so lahko ob glagolu, katere predložne zveze so tipično povezane z glagolom in podobno. Predloge so potem povezane v pomene, pri čemer so pripisane tudi potrebne stilne oznake, ki jih prevajalnik uporablja za boljše izbiranje prevodov. Da bi se izognili prevelikemu ponavljanju pri predlogah, je del informacij, ki so značilne za veliko število glagolov v posameznem jeziku, rešen programsko na nivoju prevajalnika (analizatorja oz. generatorja) in so v tem primerih v predlogah poudarjene le izjeme. Tak primer je npr. vezava s prislovnimi določili, kjer so tipično dovoljena dodatno k predlogu poljubna prislovna določila, razen prislovnega določila kraja, po katerem se vprašamo kam. Analizator in generator tudi skrbita za pravilen vrstni red prislovnih določil v posameznem jeziku.

S pomočjo predlog lahko razrešimo pri prevajanju dvoumnosti, ki se slovnično razlikujejo, recimo glede prehodnosti/neprehodnosti, povratnosti pri slovenskih glagolih, frazne glagole v angleščini... Prav tako je možno stalne fraze v celoti vpisati v predloge, s čimer lahko dosežemo pravilne prevode fraz. Ta zapis fraz je veliko bolj prilagodljiv, kot bi bil zapis z enostavnim zaporedjem besed, saj so elementi predloge lahko v poljubnem vrstnem redu (znotraj pravil posameznega jezika). Tako lahko Presis zdaj brez težav prevede tudi stavek »Bobu je včeraj Janez rekel bob.« v »Janez called a spade a spade yesterday.« Pri tem sicer najde tudi druge možne prevode (brez uporabe fraze), vendar da na prvo mesto rešitev, kjer predloga pokrije čim večji del stavka (postopek za razvrščanje najdenih analiz uporablja sicer več kriterijev za razvrstitev, vendar je ta med pomembnejšimi in ponavadi prevlada).

Predloge določijo tudi stavčne člene, tako da je zaradi tega mogoče povezovati pomene glagolov s pomeni samostalnikov glede na vlogo, v kateri se pojavljajo. Tako je mogoče določiti, da sta glagol (oz. glagolski pomen) »skuhati« in samostalniki »kosilo« povezana tako, da je kosilo tipično predmet v tožilniku (oz. rodilniku v zanikanih stavkih, za kar pa poskrbi sam analizator in tega v vzorcih ni treba posebej pisati) pri tem glagolu, ne pa npr. osebek ali predmet v dajalniku.

Pri predlogah je možno napisati tudi tipične predložne zveze, s katerimi se povezujejo. S tem lahko po eni strani izboljšamo prepoznavanje predlog, po drugi strani pa tudi dosežemo, da se predlog ob določenem glagolu prevaja drugače, kot bi se sicer običajno. »V našega Janeza se je zaljubila lepa Micka.« tako Presis prevede v »Beautiful Micka fell in love with our Janez.«

Presis zna predloge tudi prilagoditi različnim časom in trpniku. Morebitne izjeme (predloge, ki ne morejo uporabljati v trpniku ali se lahko uporabljajo le v trpniku, omejitve pri časih) so označene z uporabo zastavic pri predlogah. Če stavka, ki je bil v originalu v trpniku, v ciljnem jeziku ni mogoče zapisati na ta način (to se običajno zgodi pri prevajanju iz angleščine v slovenščino, kadar je znan tudi osebek, ki je bil v angleščini naveden z »by«), ga generator pretvori v tvorni način. Primer za to je prevod iz »John was made happy by George.« v »George je osrečil Johna.«

Posebnost slovenščine so tudi glagoli, ki dobijo v neosebni rabi drug pomen. Tak primer je »Gre za ta problem.«, kar Presis s pomočjo predloge prevede v »It is this problem.« oz. kot druga možnost »It is about this problem.« Najde pa seveda tudi kopico drugih možnosti, recimo »It goes behind this problem.«

Predloge so lahko tudi take, da kot del zahtevajo nov glagol. Tipičen primer so predloge za modalne glagole (oz. v slovenščini tudi modalne prislove, recimo »lahko« in »rad«). V takem primeru zna Presis sestaviti predlogi. Primer za tak prevod je »Why could John fall in love with Mary?«, kar se prevede v »Zakaj bi se John lahko zaljubil v Mary?«

Ocenjujemo, da bo za zadovoljivo delovanje prevajalnika potrebnih med 10.000 in 20.000 vnesenih predlog na jezik, jezik pa bo večinoma pokrit pri okoli 100.000 predlogah. Da bi Presis dovolj dobro deloval tudi v času, ko še nima vnesenih dovolj predlog, smo pri glagolih, ki nimajo vpisanih predlog, uvedli možnost, da poskuša Presis uganiti predlogo. Težava pri ugibanju predlog so v slovenščini povratni glagoli in glagoli v neosebni rabi, v angleščini pa frazni glagoli, zato imajo ti pri vnosu predlog prednost (skupaj z najbolj pogostimi glagoli).