

KLÁRA HALASOVÁ

ALGORITMICKÝ POPIS ČESKÉ FORMÁLNÍ MORFOLOGIE SUBSTANTIV A ADJEKTIV

1. ÚVOD

Cílem tohoto článku je pokus o dynamický popis české formální morfologie substantiv a adjektiv. Východiskem byl klasický statický popis, jak jej podává Mluvnice češtiny (2) Tvarosloví (dále MČ (2)).

V článku je nastíněn problém analýzy jednotlivých tvarů ohebných slov. Hlavním cílem je podrobná subklasifikace českých substantivních a adjektivních vzorů. Díleční vzory umožňují formulaci pravidel syntézy a analýzy ohebných tvarů substantiv a adjektiv v jazyce PROLOG na základě seznamu tvarotvorných základů a tvarotvorných formantů uložených v prologovské databázi. Na základě takto formulovaných pravidel je možné pomocí prologovského interpretu generovat textová slova.

2. KLASICKÝ POPIS ČESKÉ FORMÁLNÍ MORFOLOGIE

V tomto oddíle stručně připomeneme některé pojmy týkající se našeho předmětu, jak je uvádí MČ (2).

Základní jazykovou jednotkou v rámci našeho zkoumání je slovo. Je známo, že slovo je v různých rovinách jazyka definováno různým způsobem. V rámci formální morfologie nás zajímá slovo jako jednotka morfemicko-morfologická.

MČ (2) rozlišuje dvě stránky pojmu slovo, a to (1) „slovo“ jako reálně vyčlenitelnou jednotku jazykového projevu — textu, jako sled, řetězec morfů, a (2) „slovo“ jako potenciální jednotku jazyka — jazykového systému —, která je prvním slovem reprezentována. Z této definice vyplývá též rozlišování textového slova a systémového slova, nebo také slovoformy a lexému. Textové slovo je realizací systémového slova.

Zatímco v klasických popisech české formální morfologie (MČ (2), Havránek — Jedlička: Česká mluvnice), které jsou ve své podstatě statické, nacházíme pouze popis systémových slov, je cílem algoritmického popisu nakonec vytvoření programu, který generuje, popřípadě analyzuje libovolná textová slova na základě popisu systému jazyka, tedy popisu dynamický.

V jazycích s rozvinutou flexí, jako je čeština, má zmíněné rozlišení textového a systémového slova velký dosah. Zamyslíme-li se nad skutečností, že jednomu systémovému slovu odpovídá 4 až 12 textových slov u substantiv a 5 až 13 textových slov u adjektiv, je nanejvýš pravděpodobné, že rodilý mluvčí má ve své paměti uloženy jisté informace o systému české flexe, z nichž vychází při realizaci promluvového aktu.

V našich úvahách o možnostech algoritmického popisu české substantivní a adjektivní flexe jsme vyšli ze segmentace textových slov.

2.1. Segmentace textových slov

„Slovní tvar, který je realizací abstraktního systému, je lineární segment promluvy, charakterizovaný celistvostí jak významově funkční, tak i zvukovou a grafickou, se zásadní samostatností, projevující se v jeho přemístitelnosti (omezené ovšem zákonitostmi pořádku slov ve větě).“ MČ (2).

Takto jsou definovány jednoduché (syntetické) slovní tvary, které stojí v popředí našeho zájmu. Slovními tvary jsou v češtině vyjadřovány gramatické významy morfologických kategorií příslušných slovních druhů. V případě substantiv a adjektiv jsou to významy rodu, čísla a pádu.

Slovní tvar lze dále segmentovat na jednotky nižšího řádu. MČ (2) pracuje s termíny tvarotvorný základ a tvarotvorný formant. Tvarotvorný základ je lexikální složka slovního tvaru, která je všem tvarům ohebného slova společná. Tvarotvorný formant je gramatická složka slovního tvaru, nese gramatické významy, mění se během ohýbání slova.

MČ (2) uvádí také termíny kmen (tvarotvorný kmen a koncovka). Koncovku definuje jako tvarotvornou příponu stojící v absolutním konci slova (pádovou, osobní, infinitivní). Kmen je ta část slovního tvaru, která zbývá po odpojení koncovky. S tímto členěním je spjato několik problémů. Termíny tvarotvorný základ, tvarotvorný formant, kmen, koncovka se totiž částečně překrývají a částečně liší. Například tvarotvorný základ rozšířený o kmenotvornou příponu (konektém, který spojuje základ s koncovkou,) je kmen. Z hlediska dichotomie základ — formant je kmenotvorná přípona součástí tvarotvorného formantu. Ze synchronního hlediska je pojem kmenotvorné přípony živý pouze u slo-

ves. U substantiv se můžeme spokojit s tím, že tvarotvorný základ je v naprosté většině případů totožný s kmenem a tvarotvorný formant s koncovkou. Při sestavování algoritmického popisu analýzy slovesných tvarů by bylo zřejmě nejvhodnější provést potrojnou segmentaci na tvarotvorný základ, kmenotvornou příponu a koncovku osobní nebo infinitivní, případně příponu participiální a koncovku jmenného rodu. Podobný postup lze zvolit i při popisu stupňování adjektiv.

2.2. Alternace kmene

Morfologické alternace jsou změny hláskové podoby základu. Podrobný popis alternací podává MČ (2). Důležité je rozdělení na alternace samohláskové a souhláskové.

2.2.1. SAMOHLÁSKOVÉ ALTERNACE

Výčet samohláskových alternací, které se vyskytují v české flexi, je uveden v MČ (2). Pravidla jejich výskytu u jednotlivých českých substantiv nelze formulovat jinak než ve formě výčtu. V případech samohláskových alternací tzv. vkladného „e“ se můžeme opřít o slovtvorbu. Alespoň do určité míry můžeme zachytit stav v češtině následujícími pravidly.

Pravidlo 1.: Má-li základ substantiva samohláskovou alternaci, pak platí, že jedna podoba základu se pojí s koncovkou $-\emptyset$ a druhá podoba s ostatními koncovkami.

Poznámka: K určitému kolísání dochází u samohláskových alternací feminin skloňovaných podle vzoru *žena*. Toto kolísání nelze popsat formálními pravidly.

Pravidlo 2.: K samohláskovým alternacím u adjektiv dochází pouze u adjektiv přivlastňovacích tvořených příponou $-\dot{u}v$, pojí-li se k základu koncovka $-\emptyset$. Ve spojení s jinými koncovkami dochází v příponě k alternaci $\dot{u}-o$.

2.2.2. SOUHLÁSKOVÉ ALTERNACE

Jednotlivé typy alternací základových finál jsou opět popsány v MČ (2). Souhláskovým alternacím věnujeme větší pozornost, neboť jsou lépe formálně popsatelné. Vzhledem k tomu, že algoritmus, který jsme připravili, je zaměřen na analýzu psané podoby jazyka, zajímají nás především ty případy alternací, které způsobují změnu grafické podoby základu. (Tedy např. alternace velár a ne alternace dentál.)

Příklad: *vlk* — *vlci* [*vlk*] — [*vlci*]
had — *hadí* [*hat*] — [*hadí*]

Všeobecně lze tvrdit, že alternace základu závisí na koncovce, s níž se základ pojí. Z toho vycházíme při formulaci pravidel alternací velár a alternace finály *r*.

Pravidlo 3.: Má-li kmen substantiva nebo adjektiva alternací velár *k—c*, *h—z*, *g—z*, *ch—š*, pak kmen s alternací se pojí s koncovkou *-i*, *-e*, *-ich* u substantiv a s koncovkou *-í* u adjektiv.

Pravidlo 4.: Má-li kmen substantiva nebo adjektiva souhláskovou alternací *r—ř*, pak se kmen s alternací pojí u substantiv s koncovkami *-e*, *-í* a u adjektiv s koncovkou *-í*.

Zájem o grafickou podobu analyzovaných slov nás přivedl k úvahám o závislostech změn grafické podoby základu a koncovky. Touto problematikou se MČ (2) nezabývá. Zmíněné závislosti lze popsat přesně formulovanými pravidly. Pro náš popis jsou tato pravidla stejně důležitá jako pravidla morfologických alternací.

2.2.3. ZMĚNY GRAFICKÉ PODOBY KONCOVEK A ZÁKLADŮ

Pravidlo 5.: Substantiva skloňovaná podle tvrdých vzorů *hrad* a *město* mají v lokálu singuláru vedle koncovky *-u* alternativní koncovku *-e/-ě*. Substantiva skloňovaná podle vzoru *žena* mají v dativu a lokálu singuláru koncovku *-e/-ě*.

Volba grafému *-e/-ě* závisí na základové finále. Je-li finála *-d*, *-t*, *-n*, *-b*, *-p*, *-v*, *-f*, *-m*, napíšeme *-ě*. Je-li finála *-l*, *-s*, *-z*, popřípadě *-c*, *-z*, *-š*, *-ř*, vzniklá morfologickou alternací z *-k*, *-h*, *-g*, *-ch*, *-r*, píšeme *-e*.

Pravidlo 6.: Substantiva skloňovaná podle měkkých vzorů *muž*, *stroj*, *soudce*, *růže*, *píseň*, *moře* a 1. a 4. pád singuláru substantiv skloňovaných podle vzoru *kuře* mají v různých pádech koncovky *-e/-em/-emi/-ě/-ěm/-ěmi*.

Volba grafické podoby koncovky závisí opět na finále základu. Je-li finála *-p*, *-b*, *-v*, *-f*, *-m*, píšeme *-ě/-ěm/-ěmi*. Je-li finála *-l*, *-s*, *-z*, *-č*, *-š*, *-ř*, *-ž*, *-j*, píšeme *-e/em/emi*. Je-li finála *-d*, *-t*, *-ň*, píšeme *-ě/-ěm/-ěmi* a finálu píšeme jako *-d*, *-t*, *-n*. Grafická podoba finály *-d*, *-t*, *-n*, zůstává dále také před koncovkami *-i*, *-ich*, *-im*, *-í*.

Poznámka: Pro vzor *píseň* platí pravidlo 6 s následujícím omezením. Skupina substantiv zakončených na *-ev/-v* (původních *u*-kmenů) přibírá koncovky *-e/-emi*, a ne *-ě/ěmi*, jak by plynulo z pravidla 6. Totéž omezení platí i pro skupinu substantiv skloňovaných částečně podle vzoru *hrad* a částečně podle vzoru *stroj* (bývalé mužské *n*-kmeny) a pro skupinu substantiv rodu středního typu *rameno*, skloňovaných částečně podle vzoru *město* a částečně podle vzoru *moře* (bývalé střední *n*-kmeny). I tato substantiva přijímají koncovky *-e/-em*, na místě podle pravidla 6. očekávaných koncovek *-ě/-ěm*.

2.3. SYSTÉM SLOVNÍCH TVARŮ

Základním termínem, se kterým se při popisu tvaroslovného systému pracuje, je paradigma. Morfologické (tvaroslovné) paradigma je soubor tvarů ohebného slova vyjadřující systém jeho mluvnických kategorií. MČ (2) poukazuje na možnost vyčlenit z obecného paradigmatu paradigma dílčí. S dílčími paradigmaty se běžně setkáváme u sloves (dílčí paradigma přítomný, imperativní atd.). U substantiv se někdy uvádí dílčí paradigma pro singulár a nebo plurál. Na definování dílčích paradigmat substantiv a adjektiv je založena klasifikace vzorů pro algoritmičtý popis české substantivní a adjektivní deklinace.

Vzor je v MČ (2) definován jako reprezentace tvaroslovného paradigmatu paradigmatem určitého slova. V tomto článku budeme termín vzor používat i ve smyslu tvaroslovné charakteristiky slova (tj. systému koncovek v úloze slovotvorného formantu).

2.3.1. VYDĚLENÍ DÍLČÍCH PARADIGMAT — SUBKLASIFIKACE VZORŮ

Za základ klasifikace vzorů bereme rozdělení na 14 vzorů (tj. pro maskulina životná vzory *pán, muž, předseda, soudce*, pro maskulina neživotná vzory *hrad a stroj*, pro femínina *žena, růže, píseň, kost*, pro neutra *město, moře, stavení, kuře*).

Tento systém vzorů byl pro úplný formální popis, byť jen pravidelných jevů, nedostatečný. Bylo proto třeba vytvořit systém podvzorů na základě definování dílčích paradigmat. Při definování dílčích paradigmat bylo použito následujících kritérií:

1. Část paradigmatu má v závislosti na koncovce vyjadřující určitý pád, číslo, rod, alternaci, a to
 - a) samohláskovou,
 - b) souhláskovou,
 - c) mění se grafická podoba kmene.
2. Část paradigmatu má v závislosti na kmenové finále různou grafickou podobu koncovky (koncevek).
3. Část paradigmatu má dubletní koncovky v závislosti na
 - a) významu slovního tvaru (*na hradě — o hradu*)
 - b) způsobu tvoření slova (tvarotvorné příponě)
 - muži — mužové — učitel-ě*
 - páni — pánové — obč-an-ě*
 - předsedové — hus-it-ě — social-ist-ě*
 - růži — obr-ýň-ø — ul-ic-ø*
 - moř-i — hř-išť-ø*)
 - c) kmenové finále
 - (o tangu — o městě)*

Pod zorným úhlem předcházejících kritérií se běžně uváděné vzory rozpadnou na jádro (základ) vzoru nebo tvaroslovné charakteristiky, které tvoří koncovky pádů, jež nezpůsobují alternace, nemění se v závislosti na finále a nemají dublety, a jednotlivá dílčí paradigmatata, která budou dále definována.

Tabulka základních paradigmat jednotlivých vzorů

	pán(v1)	hrad(v2)	muž(v3)	stroj(v4)	předseda(v5)	soudce(v6)
N. sg.					-a	-e
G. sg.	-a				-y	-e
D. sg.	-u/-ovi	-u	-i/-ovi	-i	-ovi	-i/ovi
A. sg.	-a				-u	-e
L. sg.	-u/-ovi		-i/-ovi	-i	-ovi	-i/ovi
I. sg.	-em	-em			-ou	-em
N. pl.		-y	-i			-i
G. pl.	-ů	-ů	-ů	-ů	-ů	-ů
D. pl.	-ům	-ům	-ům	-ům	-ům	-ům
A. pl.	-y	-y			-y	-e
L. pl.			-ich	-ich		-ich
I. pl.	-y	-y	-i	-i	-y	-i

	žena(v7)	růže(v8)	píseň(v9)	kost(v10)	
N. sg.	-a				
G. sg.	-y			-i	
D. sg.		-i	-i	-i	
A. sg.	-u				
L. sg.		-i	-i	-i	
I. sg.	-ou	-i	-i	-i	
N. pl.	-y	-i		-i	
G. pl.			-i	-i	
D. pl.	-ám	-ím	-ím	-em	
A. pl.	-y			-i	
L. pl.	-ách	-ich	-ich	-ech	
I. pl.	-ami			-mi	

	město(v11)	moře(v12)	stavení(v13)	kuře(v14)	-et/-at
N. sg.	-o		-í		
G. sg.	-a		-í		-e
D. sg.	-u	-i	-í		-i
A. sg.	-o		-í		
L. sg.		-i	-í		-i
I. sg.	-em		-ím		-em
N. pl.	-a		-í		-a
G. pl.			-í		-ø
D. pl.	-ům	-ím	-ím		-ům
A. pl.	-a		-í		-a
L. pl.		-ích	-ích		-ech
I. pl.	-y	-i	-ími		-y

Definice dílčích paradigmat

Definice 1: Všechna substantiva, do jejichž tvaroslovné charakteristiky patří koncovka $-\emptyset$, se skloňují podle vzoru $v-x$. (Srov. pravidlo 1.)

Tabulka dílčích vzorů $v-x$

	pán(v1x)	hrad(v2x)	muž(v3x)	stroj(v4x)	žena(v7x)	
N. sg.	$-\emptyset$	$-\emptyset$	$-\emptyset$	$-\emptyset$		
A. sg.		$-\emptyset$		$-\emptyset$		
G. pl.					$-\emptyset$	

	růže(v8x)	píseň(v9x)	kost(v10x)	moře(v12x)	
N. sg.		$-\emptyset$	$-\emptyset$		
A. sg.		$-\emptyset$	$-\emptyset$		
G. pl.	$-\emptyset$			$-\emptyset$	

Definice 2: Tvrdá substantiva mužského a středního rodu se v lokálu plurálu skloňují podle vzoru $v-q$, pojí-li se s koncovkou $-ích$, podle vzoru $v-z$, pojí-li se s koncovkou $-ech$, nebo podle vzoru $v-p$, pojí-li se s koncovkou $-ách$.

Tabulka vzorů v-z, v-q a v-p

	pán(v1z)	v1q)	hrad(v2z)	v2q)	v2p)
L. pl.	-ech	-fch	-ech	-fch	-ách

	předseda(v5z)	v5q)	město(v11z)	v11q)	v11p)
L. pl.	-ech	-fch	-ech	-fch	-ách

Definice 3: Tvrdá substantiva rodu mužského životného, do jejichž tvaroslovné charakteristiky patří koncovka *-i*, se v nominativu plurálu skloňují podle vzoru *v-y*.

	pán(v1y)	předseda(v5y)
N. pl.	-i	-i

Definice 4a: Tvrdá substantiva se podle zakončení základu rozpadají do dvou dílčích paradigmat. Je-li finála *-d, -t, -n, -p, -b, -v, -f, -m*, skloňují se podle vzoru *v-b*, je-li finála *-s, -l, -z*, popřípadě *-c, -z, -š, -ř*, vzniklé alternací z *-k, -h, -g, -ch, -r*, skloňují se podle vzoru *v-a*.

Definice 4b: Měkká substantiva se podle zakončení základu rozpadají do dvou dílčích paradigmat. Je-li finála *-č, -š, -ř, -ž, -j, -s, -c, -z, -l*, skloňují se podle vzoru *v-a*. Je-li finála *-b, -p, -v, -f, -m*, popřípadě *-d, -t, -ň*, psané jako *-d, -t, -n*, skloňují se podle vzoru *v-b*.

Poznámka: Podle vzoru *v-a* se skloňují bývalé ů-kmeny a n-kmeny.

Tabulka vzorů v-a a v-b

	hrad(v2a)	v2b)	žena(v7a)	v7b)	město(v11a)	v11b)
D. sg.			-e	-ě		
L. sg.	-e	-ě	-e	-ě	-e	-ě

	muž(v3a)	v3b)	stroj(v4a)	v4b)	růže(v8a)	v8b)	píseň(v9a)	v9b)
N. sg.					-e	-ě		
G. sg.	-e	-ě	-e	-ě	-e	-ě	-e	-ě
A. sg.	-e							
I. sg.	-em	-ěm	-em	-ěm				
N. pl.			-e	-ě	-e	-ě	-e	-ě
A. pl.	-e	-ě	-e	-ě	-e	-ě	-e	-ě
I. pl.					-eml	-ěml	-eml	-ěml

	moře(v12a)	v12b)
N. sg.	-e	-ě
G. sg.	-e	-ě
A. sg.	-e	-ě
I. sg.	-em	-ěm
N. pl.	-e	-ě
A. pl.	-e	-ě

Definice 5: Substantiva tradičně zahrnovaná pod vzor *hrad* mají v genitivu singuláru dubletní koncovku *-u/-a* a substantiva skloňovaná podle vzorů *město* a *hrad* mají v lokálu singuláru koncovky *-u/-e/-ě*. Podle toho se skloňují v genitivu buď podle vzoru *v-c*, nebo podle vzoru *v-d* a v lokálu podle vzoru *v-e*, nebo *v-a/v-b*.

Tabulka podvzorů *v-c*, *v-d* a *v-e*

	hrad(v2a)	v2b)	v2c)	v2d)	v2e)	město(v11a)	v11b)	v11e)
G. sg.			-u	-a				
L. sg.	-e	-ě			-u	-e	-ě	-u

Definice 6: Substantiva rodu mužského životného mají v nominativu plurálu vedle základní koncovky *-l* většinou ještě koncovku *-ově* nebo *-ě*. Nominativ plurálu tvoří u těchto substantiv dílčí paradigma *v-f* a *v-g*.

Tabulka podvzorů v-f a v-g

	pán(v1f)	v1g)	muž(v3f)	v3g)	předseda(v5f)	v5g)
N. pl.	-ové	-é	-ové	-é	-ové	-é

Definice 7: Substantiva skloňovaná podle vzorů *růže* a *moře* mají v genitivu plurálu základní koncovku *-í*. Vzhledem k tomu, že skupina substantiv typu *obryně*, *ulice*, *hřiště* má v genitivu plurálu koncovku *-ø*, a patří tedy do dílčích paradigmat *v8x* a *v12x*, tvoří genitiv plurálu těchto substantiv dílčí podvzor *v-e*.

Tabulka podvzorů v-e a v-x

	růže(v8e)	v8x)	moře(v12e)	v12x)
G. pl.	-í	-ø	-í	-ø

Na základě předložené klasifikace je každé substantivum reprezentováno ne jedním vzorem, ale celou řadou (systémem) podvzorů.

Podobným způsobem byly doplněny i vzory adjektivní. Základní vzory — tvrdý (*chudý*), měkký (*jarní*) a přivlastňovací (složený, jmenný (*otcův/ matčín*)) mají mnohem méně podvzorů než vzory podstatných jmen.

Při definování dílčích paradigmat u adjektiv jsme vycházeli pouze z kritéria 1. Ostatní se v adjektivní deklinaci neuplatňují.

Uveďme nejprve jádro vzorů adjektiv.

Tabulka základních vzorů adjektiv

	chudý(v1)	mz	mn	n	f	jarní(v2)	mz	mn	n	f
N. sg.		-ý	-ý	-é	-á		-í	-í	-í	-í
G. sg.			-ého		-é			-ího		-í
D. sg.			-ému		-é			-ímu		-í
A. sg.		-ého	-ý	-é	-ou		-ího	-í	-í	-í
L. sg.			-ém		-é			-ím		-í
I. sg.			-ým		-ou			-ím		-í
N. pl.			-é	-á	-é			-í		

	chudý(v1)	mz	mn	n	f	jarní(v2)	mz	mn		f
G. pl.			-ých					-ích		
D. pl.			-ým					-ím		
A. pl.		-é	-é	-á	-é			-í		
L. pl.			-ých					-ích		
I. pl.			-ými					-ími		

	otcův/otcov-(v3)	mz	mn	n	f
N. sg.				-o	-a
G. sg.			-a		-y
D. sg.			-u		-ě
A. sg.		-a		-o	-u
L. sg.			-ě		-ě
I. sg.			-ým		
N. pl.		-í	-y	-a	-y
G. pl.			-ých		
D. pl.			-ým		
A. pl.		-y	-y	-a	-y
L. pl.			-ých		
I. pl.			-ými		

Definice 8: Adjektiva tvrdá se v nominativu plurálu rodu mužského životného skloňují podle vzoru *vly*.

Tabulka *vly*

	chudý(v1)	mn	n	f	chudý(vly)	mz
N. pl.		-é	-á	-é		-í

Definice 9: Adjektiva skloňovaná podle vzoru smíšeného (*otcův*) se v nominativu singuláru maskulin životných i neživotných a v akuzativu singuláru maskulin neživotných skloňují podle vzoru *v3x*.

Tabulka v3x

	otcův {v3x}	mz	mn.	otcov- {v3}	mz	n	f
N. sg.		-ø	-ø			-o	-a
A. sg.			-ø		-a	-o	-u

3. ALGORITMICKÝ POPIS ČESKÉ FORMALNÍ MORFOLOGIE SUBSTANTIV A SLOVES

3.1. Využití programového jazyka PROLOG pro práci s přirozeným jazykem

Programovací jazyk PROLOG si v poslední době získal značnou pozornost, zejména díky své jednoduchosti a velkým vyjadřovacím schopnostem. Způsob zápisu programu je přístupný i uživateli neprogramátorovi.

Zhruba řečeno umožňuje PROLOG uživateli vytvoření báze poznatků z určité oblasti (tzv. databáze — db). Databáze obsahuje klausule, které definují vztahy mezi objekty reálného světa, v našem případě přirozeného jazyka. Klausule mohou být buď fakta definovaná výčtem, nebo pravidla, která definují vztahy obecněji, nad fakty a pravidly už dříve definovanými.

Programováním v PROLOGU rozumíme jednak vytváření db (tj. deklarování faktů a pravidel), jednak kladení otázek nad uloženými klausulemi. Otázka ověřuje prostřednictvím interpretu PROLOGU pravdivost údajů z otázky (odpovídá otázce zjišťovací v přirozené komunikaci), nebo přímo zjišťuje hodnotu údajů, na které se uživatel ptá, a to buď přímo, porovnáním s db, nebo nepřímou, dedukcí z pravidel uložených v db (odpovídá otázce doplňovací v přirozené komunikaci).

Jazyk PROLOG umožňuje vyjádření některých logických operací. Je to např. konjunkce označovaná symbolem „;“, disjunkce označovaná „;“, implikace označovaná „:-“. Důležitá je funkce otázky „?—“.

Operace, které s využitím PROLOGU provádí počítač, jsou založeny na dokazovacím mechanismu. Systém se snaží splnit úkoly, které mu předkládáme. K tomu účelu prohledává db. Jako splnění zadaného cíle vybírá z db tu klausuli, která byla úspěšně porovnána s cílem.

3.2. Využití jazyka PROLOG při popisu systému deklinace substantiv a adjektiv

Z možností programovacího jazyka PROLOG vyplývá, že cílem programu pro automatickou analýzu a syntézu ohebných slov je vytvoření báze

poznatků o systému deklinace v termínech tohoto jazyka. Do db můžeme ukládat informace o českém skloňování jednak ve formě faktů definovaných výčtem, jednak ve formě pravidel. Pravidlo se v jazyce PROLOG skládá ze dvou částí: levá strana (hlava) pravidla udává definovaný vztah, pravá strana (tělo) pravidla je tvořena logickou formulí nad již známými fakty (pravidly).

3.2.1. DATABÁZE

V teoretické části (kap. 2) jsme vyšli z předpokladu, že každé ohebné slovo lze dále segmentovat na dvě části — tvarotvorný základ a tvarotvorný formant. Zatímco základů je velký počet a navíc tvoří otevřený inventář, je počet formantů do značné míry omezen. U substantiv je počet možných formantů 26 (počítáme všechny grafické podoby jednotlivých formantů), u adjektiv 23.

Z hlediska formálního popisu můžeme jednotlivé tvary ohebných slov považovat za kombinace základů a formantů. Tyto kombinace nejsou ovšem nikterak libovolné. Kritérií by se našlo více. Nejdůležitějším a nejspolehlivějším je příslušnost základu k ohýbacímu typu, který je v naprosté většině případů typický pro jeden slovní druh a u substantiv také pro jmenný rod.

Tvarotvorné formanty vyjadřují, jak už bylo řečeno výše, gramatické významy mluvnických kategorií. V českém flektivním systému neplatí rovnice jeden formant = jeden význam (kombinace významů — číslo + pád, rod + číslo + pád). Existuje až 50% homonymie (srov. MČ(2)). Např. morfologická *-ø* u substantiv vyjadřuje u maskulin životných skloňovaných podle vzorů *pán* a *muž* gramatické významy nom. sg., u maskulin neživotných skloňovaných podle vzorů *hrad* a *stroj* významy nom. sg. a ak. sg., u feminin skloňovaných podle vzorů *píseň* a *kost* taktéž, u feminin skloňovaných podle vzoru *žena* a podle vzoru *růže* typu *ulice* a u neuter skloňovaných podle vzoru *město* a *moře* typu *hřiště* vyjadřuje *-ø* význam genitiv plurálu. Tých tvarotvorný formant je součástí tvaroslovné charakteristiky substantiv nebo adjektiv skloňovaných podle různých vzorů.

Shrňme tedy: možnost kombinace základu a formantu je dána příslušností základu k ohýbacímu typu a formantu do tvaroslovné charakteristiky téhož ohýbacího typu.

Do db musíme zahrnout následující skupiny faktů a pravidel:

1. Seznam tvarotvorných formantů českých substantiv a adjektiv s udáním gramatických významů pádu a čísla u substantiv a rodu, pádu a čísla u adjektiv a vzoru, u něhož dané gramatické významy vyjadřují.
2. Slovník základů s udáním rodu (u substantiv) a vzoru (podvzorů).
3. Pravidla kombinací základů a formantů, které jsou podmíněny příslušností obou složek k témuž vzoru.

Seznam českých substantivních a adjektivních formantů je v db uložen ve formě faktů definovaných výčtem.

Příklad: $t(-\emptyset, 1, S, /v1x, v2x, v3x, v4x, v9x, v10x/)$.

$t(-\emptyset, 4, S, /v2x, v4x, v9x, v10x/)$.

$t(-\emptyset, 2, P, /v7x, v8x, v11x, v12x/)$.

Slovně: Tvarotvorný formant $-\emptyset$ má význam 1. pád singuláru ve spojení se základy skloňovanými podle vzorů *pán, hrad, muž, stroj, píseň a kost*, význam 4. pád singuláru ve spojení se základy skloňovanými podle vzorů *hrad, stroj, píseň, kost* a význam 2. pád plurálu ve spojení se substantivy skloňovanými podle vzorů *žena, růže [ulice], města a moře [hřiště]*.

Jako fakta definovaná výčtem jsou do db zahrnuty slovníky základů substantiv a adjektiv.

Příklad: $ka(star, v1)$.

$ka(stař, v1y)$.

Slovně: Tvarotvorný základ *star-* je základ adjektiva skloňovaného podle vzoru *chudý* pro všechny pády obou čísel všech rodů kromě nom. pl. mz. Nominativ plurálu rodu mužského životného se tvoří od tvarotvorného základu *stař-* skloňovaného podle vzoru *v1y*.

Podstatou pravidel PROLOGU jsou implikace, v nichž antecedent tvoří konjunkce predikátů, z nichž první odpovídá tvarotvornému základu a druhý tvarotvornému formantu. Podmínkou úspěšné konjunkce je shoda vzoru u tvarotvorného základu a tvarotvorného formantu. Konsekvent je pak syntéza základu a formantu, vytvoření textového slova.

Příklad: $sb(ks, t) :- ks(Ks, Gen, v1), t(T, Cas, Uum, v1)$.

Slovně: Jestliže *ks* je tvarotvorný základ substantiva rodu *Gen* (mz = maskulina životného, mn = maskulina neživotného, f = feminina, n = neutra) skloňovaného podle vzoru *v1* (*pán*), a *t* tvarotvorný formant je pro vzor *v1* (*pán*) nositelem gramatických významů *Cas* (nominativ, genitiv, ...) a *Num* (singulár, plurál), pak *ks* a *t* tvoří dohromady tvar substantiva *sb*.

V konkrétním případě je v části db — slovníku kmenů — uložen tvarotvorný základ *vlc-* a v části db — tabulce koncovek — tvarotvorný formant *-i* nebo *-ich*. V termínech PROLOGU zapíšeme tyto údaje následujícím způsobem.

$ks(vlc, mz, /v1y, v1q/)$.

$t(-i, 1, P, /v1y, v3, v5y, v10/)$.

$t(-ich, 6, P, /v1q, v2q, v3, v4, v5q, v6, v8, v9, v11q, v12, v13/)$.

Pravidlo generuje na základě této části db tvary substantiva *vlk* následujícím způsobem.

$sb(vlci) :- ks(vlc, mz, v1y), t(i, 1, P, v1y)$.

$sb(vlcich) :- ks(vlc, mz, v1q), t(ich, 6, P, v1q)$.

Počet pravidel popisujících českou substantivní deklinaci je 65, počet pravidel popisujících adjektivní deklinaci je 5. Čísla jsou dána, jak je zřejmé, počtem podvzorů.

L I T E R A T U R A

- BABKA, O.—ROHANOVÁ, J.—JANČÍK, J.: Jazyk PROLOG (Výzkumná zpráva). Vysoká škola strojní a elektrotechnická v Plzni, Výpočetní středisko, Plzeň 1986.
- DOKULIL, M.: K otázce morfologických protikladů. SaS, 19, 1958, s. 81n.
- HAVRÁNEK, B.—JEDLIČKA, A.: Česká mluvnice. SPN, Praha 1981.
- CHOMSKY, N.: Syntaktické struktury. Academia, Praha 1966.
- LAMPRECHT, A.—ŠLOSAR, D.—BAUER, J.: Historická mluvnice češtiny. SPN, Praha 1986.
- Mluvnice češtiny (1), Mluvnice češtiny (2). Academia, Praha 1987.
- RUSÍNOVÁ, Z.: Současná česká morfologie. SPN, Praha 1984.
- SGALL, P.: Generativní popis jazyka a česká deklinace. ČSAV, Praha 1967.
- SLAVÍČKOVÁ, E.: Retrogradní morfematičtý slovník češtiny. Academia, Praha 1975.

DIE ALGORITHMISCHE BESCHREIBUNG DER TSCHECHISCHEN FLEXIONSMORPHOLOGIE DER SUBSTANTIVE UND ADJEKTIVE

In diesem Artikel wird die Problematik der tschechischen Flexionsmorphologie behandelt.

Man vergleicht die Beschreibungsweise des Deklinationssystems in den traditionellen Grammatiken, wie z. B. bei Havránek—Jedlička, oder in Mluvnice češtiny (2) — Tvarosloví, mit der Möglichkeit, diesen System dynamisch zu beschreiben.

Das Zentralproblem der formalen Deskription des Deklinationssystems ist die Klassifizierung der einzelnen Deklinationstypen. Diese Klassifizierung unterscheidet sich von der in den genannten Grammatiken üblichen. Zu jedem Deklinationstyp wird nach folgenden Kriterien eine Reihe von Teiltypen gebildet.

1. die Alternation des Vokals oder des Konsonantes im Stamm;
2. alternative Schreibweise der Endung;
3. Endungsdubletten.

