

# **Uživatelská dokumentace ke slovníkovému editoru MorfoEd**

David Kolovratník a Leoš Přikryl

2. června 2008

# Obsah

<b>1 Editor</b>	<b>3</b>
1.1 Spouštění . . . . .	3
1.2 Konfigurační soubory . . . . .	3
1.2.1 Soubor .Xdefaults . . . . .	3
1.3 Potřebné datové soubory . . . . .	4
1.4 Prohlížení slovníku . . . . .	5
1.5 Editace slovníkového hesla . . . . .	7
1.5.1 Vytvoření nového slovníkového hesla . . . . .	9
1.5.2 Editace existujícího slovníkového hesla . . . . .	9
1.5.3 Smazání existujícího slovníkového hesla . . . . .	10
1.6 Režim revize . . . . .	10
1.7 Typický scénář používání editoru . . . . .	10
<b>A DTD používaných dokumentů</b>	<b>12</b>
<b>Doporučená literatura</b>	<b>14</b>

# Kapitola 1

## Editor

### 1.1 Spouštění

Editor se spouští souborem `morfoEd`. První spuštění trvá trochu déle, protože se musí vytvořit index primárního souboru. Index zůstává uložen na disku, takže při příštím spuštění se už nemusí vytvářet (jedině, když se změní primární soubor).

Před spuštěním editoru musíte ve svém domovském adresáři vytvořit soubor `.morforc` a vyplnit v něm položku `author` (viz následující kapitola).

### 1.2 Konfigurační soubory

Vzhled a chování editoru ovlivňují tři konfigurační soubory: `.morforc`, `.morfoinfo` a `.Xdefaults`. Všechny soubory hledá editor v domovském adresáři aktuálně přihlášeného uživatele.

Soubor `.morforc` musí existovat, bez něj nelze editor spustit. Tento soubor si musí každý uživatel před prvním spuštěním editoru vytvořit ve svém domovském adresáři. Musí obsahovat alespoň položku `author`. Popis souboru `.morforc` najdete v sekci `XrefId[?Konfigurační soubor?]`.

V souboru `.morfoinfo` jsou uloženy interní informace editoru a není určen k ruční editaci. Pokud tento soubor smažete, bude znovu vytvořen při příštím spuštění editoru. Budou však zapomenuta některá nastavení (velikost a pozice oken, naposledy použitý soubor změn, apod.). Toho lze někdy s výhodou využít (např. když snížíte rozlišení obrazovky a po spuštění editoru je okno tak velké, že nevidíte jeho ovládací prvky).

Pomocí souboru `.Xdefaults` lze nastavit barvy, fonty a vzhled editoru.

#### 1.2.1 Soubor `.Xdefaults`

Editor využívá standardní nastavování vzhledu pomocí X Resource Database. Ta je řízena souborem `.Xdefaults` (na některých linuxových systémech to může být soubor `.Xresources`). Je velmi pravděpodobné, že tento soubor již ve svém domovském adresáři máte. V takovém případě do něj níže uvedené parametry pouze přidejte.

Pokud nenastavíte vzhled pomocí tohoto souboru, použijí se defaultní hodnoty. V takovém případě se může stát, že nebudou správně zobrazeny některé české znaky, protože se použije font, který nepodporuje UTF-8.

PARAMETRY NASTAVITELNÉ V SOUBORU `.XDEFAULTS`

**\*UKMorfoEditor\*background** barva pozadí oken editoru,

- \***UKMorfoEditor\*font** font používaný všude tam, kde není definovaný jiný,
- \***UKMorfoEditor\*Button\*font** font používaný na tlačítkách,
- \***UKMorfoEditor\*Entry\*font** font používaný v editačních polích (kvůli zobrazení značek doporučuji font s konstantní šířkou znaku),
- \***UKMorfoEditor\*Listbox\*conflictingAttr** barva, kterou jsou zobrazeny tvary slova, které nevznikly podle vzoru a jejichž atributy se neshodují s modusem,
- \***UKMorfoEditor\*Listbox\*derivedFrom** barva, kterou jsou zobrazeny tvary slova, které vznikly podle vzoru a jejichž atributy se shodují s modusem,
- \***UKMorfoEditor\*Listbox\*derivedAndConflict** barva, kterou jsou zobrazeny tvary slova, které vznikly podle vzoru a jejichž tvary se neshodují s modusem,
- \***UKMorfoEditor\*Listbox\*isSearched** barva, kterou jsou zobrazeny tvary, které mají stejné lemma, jako je lemma aktuálně vybrané v listboxu,
- \***UKMorfoEditor\*Listbox\*notSearched** barva, kterou jsou zobrazeny tvary, které nemají stejné lemma, jako je lemma aktuálně vybrané v listboxu,
- \***UKMorfoEditor\*Listbox\*newLemma** barva, kterou jsou v režimu revize zobrazena nová lemmata v seznamu lemmat,
- \***UKMorfoEditor\*Listbox\*changedLemma** barva, kterou jsou v režimu revize zobrazena změněná lemmata v seznamu lemmat,
- \***UKMorfoEditor\*Listbox\*deletedLemma** barva, kterou jsou v režimu revize zobrazena smazaná lemmata v seznamu lemmat,
- \***UKMorfoEditor\*Button\*revison** barva tlačítka na přepínání mezi režimem revize a editace,
- \***UKMorfoEditor\*Listbox\*activeRevision** barva tlačítka na přepínání mezi režimem revize a editace, pokud je nad ním kurzor.

**DŮLEŽITÉ**

Barvy `newLemma`, `changedLemma` a `deletedLemma` musejí být navzájem různé!

Vzorový `.Xdefaults` si můžete prohlédnout v souboru `slovník/Xdefaults`.

## 1.3 Potřebné datové soubory

Editor ke své práci potřebuje celou řadu datových souborů.

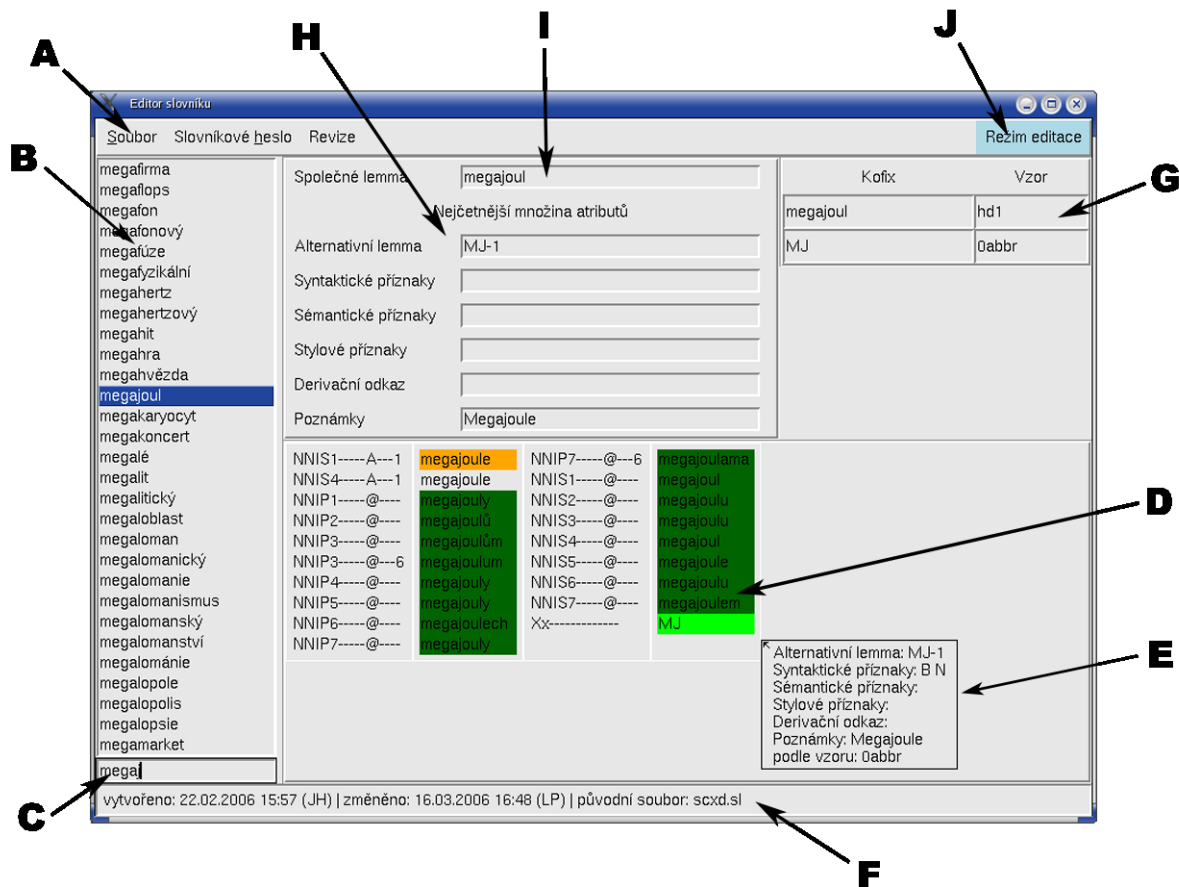
- Primární soubor - popis viz XrefId[?datové zdroje?].
- Seznam tagů - popis viz XrefId[?datové zdroje?].
- Seznam ohýbacích vzorů - popis viz XrefId[?datové zdroje?].
- Seznam derivačních vzorů - popis viz XrefId[?datové zdroje?].
- Seznam syntaktických příznaků - popis viz XrefId[?Programátorská dokumentace?].
- Seznam sémantických příznaků - popis viz XrefId[?Programátorská dokumentace?].
- Seznam stylových příznaků - popis viz XrefId[?Programátorská dokumentace?].
- Soubor s popisem jednotlivých pozic v tagu - popis viz XrefId[?Programátorská dokumentace?].
- Soubor s výjimkami při hádání vzorů - popis viz XrefId[?Programátorská dokumentace?].
- Soubor s tabulkou pro hádání vzorů - popis viz XrefId[?Programátorská dokumentace?].

Umístění všech těchto souborů lze nastavit v konfiguračním souboru XrefId[?.morforc?]. Pokud zde není nalezeno umístění, použije se zakompilovaná hodnota, která vede na příslušný soubor v adresáři, kam byla nainstalována data. Takže pokud nemáte speciální požadavky, nemusíte umístění těchto souborů nastavovat.

Pokud chcete slovník editovat, budete potřebovat soubor změn. Ten vytvoříte (případně otevřete) v editoru pomocí menu Soubor->Vybrat soubor změn.

## 1.4 Prohlížení slovníku

Obrázek 1.1 Hlavní okno editoru



## POPIS JEDNOTLIVÝCH ČÁSTÍ HLAVNÍHO OKNA

**A** Menu (popis je níže)

**B** Seznam lemmat - po seznamu lemmat se lze pohybovat pomocí šipek a kláves PageUp a PageDown.

**C** Pole pro inkrementální vyhledávání. Po zadání libovolného prefixu do tohoto pole se v listboxu vybere první slovo s tímto prefixem.

**D** Rozgenerované tvary slova. Všechny tvary vybraného slova včetně tagů. O barvách viz dále.

**E** Bublínková nápověda. Zobrazuje různé doplňkové informace o položce pod šipkou. Podrobnosti jsou popsány dále.

**F** Stavový řádek. Pokud je kurzor myši na nějakým tvarem slova, zobrazuje informace o vytvoření a poslední změně tohoto tvaru.

**H** Nejčtenější množina atributů mezi zobrazenými tvary. Většina tvarů v poli D má tyto atributy. Ostatní jsou barevně odlišeny - viz dále.

**I** Společné lemma pro toto slovníkové heslo. Je to lemma společné všem tvarům zobrazeným v poli D.

**J** Tlačítko pro přepínání mezi režimem revize a režimem editace. Pokud není vybrán soubor změn je toto tlačítko neaktivní a nelze použít (editor je v režimu prohlížení).

## MENU SOUBOR

**Vybrat soubor změn** Umožňuje vybrat soubor změn. Pokud je vybrán existující soubor, změny v něm uložené se ihned promítnou do editoru. Je také možné do spodního řádku zadat nový neexistující soubor. V tom případě bude vytvořen a změny se budou ukládat do něj. Otevřením souboru změn se zavírá a automaticky ukládá právě otevřený soubor změn.

**Export do starého formátu** Vyexportuje všechna slovníková hesla tak, jak jsou zobrazena v editačním režimu (tj. s aplikovaným souborem změn) do starého formátu slovníku. Při exportu je třeba vybrat adresář, ve kterém budou vytvořeny soubory starého formátu. Tento adresář musí existovat. Každé slovníkové heslo se uloží do souboru, jehož název odpovídá hodnotě „původní soubor“, která se zobrazuje ve stavovém řádku hlavního okna. Existující soubory v cílovém adresáři budou přepsány. Slovníková hesla, která nemají definovaný původní soubor (například nově vložená slova) se uloží do souboru, jehož jméno lze nastavit v Xrefld[?konfiguračním souboru .morforc?] parametrem old\_format\_default\_filename.

**Export do XML formátu** Vyexportuje všechna slovníková hesla tak, jak jsou zobrazena v editačním režimu (tj. s aplikovaným souborem změn) do XML formátu slovníku. Při exportu je potřeba zadat název XML souboru, od kterého bude slovník exportován. Pokud tento soubor již existuje, bude přepsán.

**Konec** Ukončí práci s editorem.

## MENU SLOVNÍKOVÉ HESLO

**Vložit** Zobrazí okno pro zadání nového slovníkového hesla. Klávesová zkratka <F3>.

**Editovat** Zobrazí okno pro editaci slovníkového hesla, které je právě vybráno v seznamu lemmat. Klávesová zkratka <F4>.

**Smazat** Smaže slovníkové heslo, které je právě vybráno v seznamu lemmat. Klávesová zkratka <F8>.

MENU REVIZE

**Zrušit změnu** Zruší veškeré změny provedené na právě vybraném slovníkovém hesle (tj. uvede ho do původního podoby z primárního souboru). Klávesová zkratka <F5>.

#### POZNÁMKA



Položky menu Slovníkové heslo nelze použít v režimu revize. Naopak položky menu Revize nelze použít v režimu editace.

Po vybrání lemmatu ze seznamu lemmat se v pravé části hlavního okna zobrazí informace o příslušném slovníkovém hesle. Podržíte-li kurzor myši nad většinou údajů, zobrazí se bublinková nápověda vztahující se k danému údaji. U tagu se zobrazí význam jednotlivých pozic, u příznaků význam všech příznaků.

V horní části okna je společné lemma a nejčtenější množina atributů. Atributy se mohou pro jednotlivé tvary lišit, ale u většiny tvarů jsou stejné. Právě tyto převládající atributy jsou zobrazeny v horní části okna. V dalším textu budu tuto nejčtenější množinu nazývat *modus*.

Pozadí pod slovními tvary ve spodní části okna může mít různé barvy.

- Šedá barva znamená, že slovní tvar je ručně zadaný (nevznikl podle žádného vzoru) a že atributy tohoto tvaru se shodují s modusem.
- Oranžová barva znamená, že slovní tvar je ručně zadaný a že se jeho atributy liší od modusu. Podržíte-li kurzor myši nad takovýmto slovním tvarem, zobrazí se bublinková nápověda s jeho atributy.
- Tmavě zelená barva znamená, že slovní tvar byl vytvořen podle nějakého vzoru a že se atributy tohoto tvaru neliší od modusu. Podržíte-li kurzor myši nad takovýmto slovním tvarem, zobrazí se bublinková nápověda se vzorem, podle kterého tvar vznikl.
- Světle zelená barva znamená, že slovní tvar byl vytvořen podle nějakého vzoru a že se atributy tohoto tvaru liší od modusu. Podržíte-li kurzor myši nad takovýmto slovním tvarem, zobrazí se bublinková nápověda s atributy tohoto tvaru a se vzorem, podle kterého tvar vznikl.

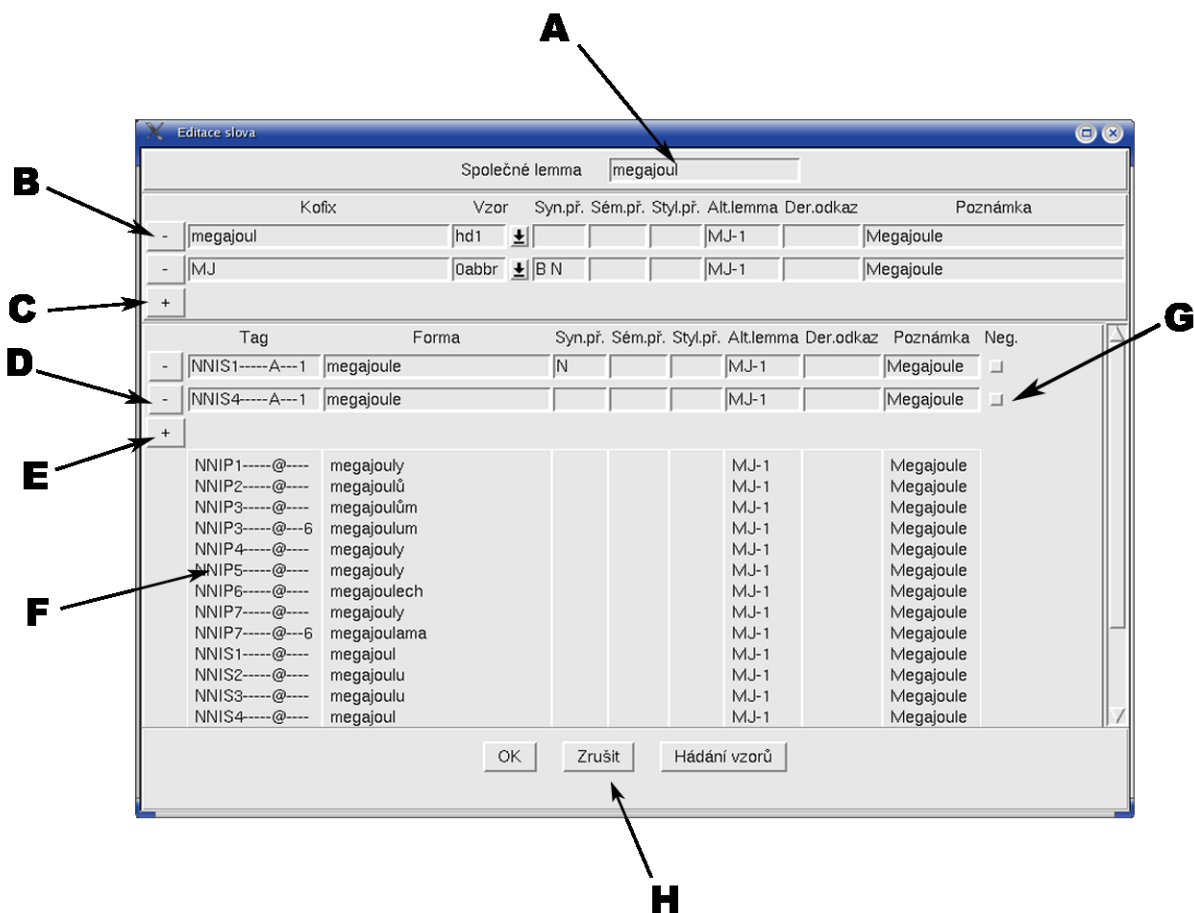
Také písmo, kterým jsou jednotlivé tvary napsány, může mít různé barvy. Slovníkové heslo totiž může obsahovat i takové slovní tvary, které mezi svými lemmaty nemají lemma, které je právě vybráno v seznamu lemmat. Do stejného slovníkového hesla ale patří proto, že mají stejné společné lemma. Tvary, které mezi svými lemmaty nemají lemma, které je právě vybráno, jsou napsány bílou barvou. Ty, které ho mezi svými lemmaty mají, jsou napsány černě.

Barvy popisované výše jsou pouze výchozím nastavením. Všechny barvy je možné změnit pomocí souboru [.Xdefaults](#).

## 1.5 Editace slovníkového hesla

POPIS JEDNOTLIVÝCH ČÁSTÍ EDITAČNÍHO OKNA

Obrázek 1.2 Editáčnı́ okno



A Lemma, společně pro všechny tvary.

B Tlačítko, sloužící k odebrání derivačního vzoru.

C Tlačítko, sloužící k přidání derivačního vzoru.

D Tlačítko, sloužící k odebrání jednotlivého slovního tvaru.

E Tlačítko, sloužící k přidání jednotlivého slovního tvaru.

F Slovní tvary vzniklé rozgenerováním vzorů uvedených v horní části okna. Po najetí nad slovní tvar se zobrazí bublinková nápověda s odpovídajícím vzorem.

G Zaškrtnutí pole, určující, zda jde slovní tvar negovat nebo ne.

H Tlačítkem OK potvrdíte změny. Změny jsou okamžitě ukládány do souboru změn. Tlačítkem Zrušit změny zrušíte. Tlačítkem Hádání vzorů spustíte hádání vzorů.

Při prvním spuštění je editor v prohlížečím režimu. V tomto režimu nelze provádět žádné změny, ani přepnout do jiného režimu. Pokud chcete provádět změny, musíte vybrat soubor změn, tj. soubor, kam se budou ukládat veškeré změny. Do primárního souboru se nic neukládá - zůstává celou dobu



nezměněn. Tento přístup umožňuje práci více uživatelů nad společným primárním souborem. Každý má svůj soubor změn, do kterého mu nikdo další nezasahuje. Soubor změn je také mnohem menší než celý primární soubor, takže ho lze snadno přenést na jiný počítač.

Soubor změn vyberete pomocí menu Soubor->Vybrat soubor změn. Můžete vybrat již existující soubor změn (v tom případě se změny již obsažené v tomto souboru promítnou do slovníku) nebo do spodního řádku dialogového okna zadat název nového souboru. Pokud nezadáte příponu souboru, bude automaticky doplněna (.xml). Při každém dalším spuštění editor nabízí naposledy použitý soubor změn k okamžitému otevření (pokud tento soubor stále existuje).

### 1.5.1 Vytvoření nového slovníkového hesla

Nové slovníkové heslo lze vytvořit pomocí položky menu Slovníkové heslo->Vložit nebo klávesové zkratky <F3>. Program se nejdříve zeptá na lemma. Zadané lemma se stane společným lemmatem nového slovníkového hesla a bude použito pro hádání vzorů. Pokud zadáte již existující lemma, bude nabídnuto automatické přidání sufixu ve tvaru -číslo.

Po potvrzení lemmatu klávesou Enter se zobrazí okno hádání vzorů. Okno se skládá z několika sloupců. V záhlaví každého sloupce je název jednoho z možných vzorů. Pod každým vzorem je seznam všech tvarů, vzniklých podle tohoto vzoru. Podržíte-li kurzor myši nad některým z tvarů, zobrazí se všechny tagy, příslušející k danému tvaru. Správný vzor (vzory) vyberete kliknutím na název vzoru. Vybrané vzory jsou barevně označeny. Výběr správných vzorů potvrdíte tlačítkem OK. Tlačítko OK jde stisknout pouze když je vybrán nějaký vzor. Může se také stát, že žádný z nabízených vzorů nebude ten pravý. V tom případě buď nevyberte žádný, stiskněte Zrušit a všechny tvary zadáte ručně, nebo vyberte nejpodobnější, který pak převedete na ručně zadávané tvary a ty upravíte (viz dále). Více o hádání vzorů najdete v XrefId[?programátorské dokumentaci?].

Po výběrů vzorů se dostanete do editačního okna a dále již postupujete jako při editaci existujícího slovníkového hesla.

### 1.5.2 Editace existujícího slovníkového hesla

Stisknutím klávesy <F4> nebo pomocí menu Slovníkové heslo->Editovat zahájíte editaci slovníkového hesla, které je právě vybráno v seznamu lemmat. Otevře se editační okno, které umožňuje změnu všech atributů slovníkového hesla.

V horní části editačního okna je možné editovat vzory, ke kterým dané slovo přísluší. Vzorek lze odebrat stiskem tlačítka „-“ vlevo od příslušného vzoru. Přidat vzor umožňuje tlačítko „+“ vlevo pod vzorem. Jednotlivé tvary vzniklé podle vzoru se zobrazují v části F editačního okna. Každá změna vzoru se zde okamžitě projeví.

Vzory lze vybírat buď ze seznamu, který se zobrazí po stisku šipky vlevo od vzoru, nebo je možné je zadat přímo. V takovém případě lze zadat pouze prefixy existujících vzorů nebo celé vzory. Jiné řetězce není možno zadat. Ručně zadaný vzor potvrdíte klávesou Enter. Pokud jde jen o prefix, automaticky se doplní první vzor s tímto prefixem.

Ve části F editačního okna jsou, kromě tvarů vzniklých podle vzorů, také ručně zadané tvary. Ruční zadávání tvarů slouží k podchycení nepravidelností, které se vymykají dostupným vzorům. Tvary lze přidávat a odebírat pomocí tlačítek + a - podobně jako vzory.

Často se stává, že vzor postihuje tvary slova takřka přesně a jen několik málo tvarů je špatně. V takovém případě je možné dvojklikem na některou z forem příslušejících ke vzoru tento vzor odebrat a všechny jeho tvary převést na ručně zadávané tvary, které už lze editovat jednotlivě.

Při ručním zadávání tagu se zobrazuje nápověda k jednotlivým pozicím. Hodnoty na jednotlivých pozicích se dají vybírat buď ze seznamu, nebo zadávat přímo.

Do polí pro příznaky (syntaktické, sématické a stylové) lze zadat pouze přípustné hodnoty. Jejich seznam se zobrazí v bublinkové nápovědě, podržíte-li kurzor myši nad příslušným polem.

Mezi jednotlivými poli v části F editačního okna (pole s tvary) se dá pohybovat pomocí Ctrl+kurzorových šipek.

Změny ve slovníkovém hesle potvrdíte tlačítkem OK. Po jeho stisknutí se ještě provádějí některé kontroly (například, jestli jsou zadány všechny povinné hodnoty, jestli všechny zadané tagy existují, apod.). Pokud tyto kontroly naleznou nějakou chybu, zobrazí se její popis a kurzor se nastaví na první chybné pole. Po opravení chyb můžete změny opět potvrdit tlačítkem OK.

Tlačítko hádání vzorů v dolní části editačního okna umožňuje spustit hádání vzorů stejné jako při zadávání nového slovníkového hesla. Toto dodatečné hádání navíc kontroluje shodu s již zadanými tvary a nabídne jen takové vzory, které odpovídají již zadaným formám a tagům. Pokud se nenajde žádný takový vzor, je spuštěno normální hádání. Pokud vyberete vzor, podle kterého by vznikl nějaký, již ručně zadaný, tvar, je tento tvar automaticky odstraněn.

### 1.5.3 Smazání existujícího slovníkového hesla

Stisknutím klávesy <F8> nebo pomocí menu Slovníkové heslo->Smazat smažete slovníkové heslo, které je právě vybráno v seznamu lemmat. Smazání jednoho slovníkového hesla může ovlivnit více lemmat ze seznamu. Smazáním slovníkového hesla se totiž smaže společné lemma i všechna alternativní lemmata.

## 1.6 Režim revize

Režim revize umožňuje zkontrolovat všechny změny uložené v právě otevřeném souboru změn a případně je zrušit. Do režimu revize a zpět se přepíná pomocí tlačítka v pravé horní části hlavního okna.

V režimu revize může mít pozadí pod lemmaty v seznamu lemmat různé barvy.

- Červenou barvou jsou zobrazena lemmata smazaných slovníkových hesel. Po vybrání takového lemmatu se v pravé části okna zobrazí slovníkové heslo tak, jak vypadalo před smazáním.
- Zelenou barvou jsou zobrazena lemmata nově přidaných slovníkových hesel. Po vybrání takového lemmatu se zobrazí toto nové slovníkové heslo.
- Oranžová a šedá lemmata jsou vždy v páru. Jsou to lemmata změněných slovníkových hesel. Po vybrání šedého lemmatu se zobrazí slovníkové heslo před změnami, po vybrání oranžového se zobrazí slovníkové heslo po provedených změnách. Tak je možné snadno zjistit, jaké změny byly provedeny.

Stiskem klávesy <F5> nebo pomocí menu Revize->Zrušit změnu zrušíte všechny změny provedené na aktuálně zobrazeném slovníkovém hesle. Heslo se vrátí do původního stavu, v jakém je uloženo v primárním souboru. U změněných to znamená zrušení změn, u smazaných obnovení a u nových jejich smazání.

## 1.7 Typický scénář používání editoru

Editor je navržen pro práci více autorů nad jedním slovníkem. Autoři mohou používat jak sdílenou instalaci, tak i svoji osobní. Při použití sdílené instalace má každý autor pouze svoje konfigurační soubory a soubor změn. Veškeré změny jsou ukládány pouze do souboru změn konkrétního autora, takže se u nikoho jiného neprojeví. Primární soubor může být společný a zůstává nezměněn.

Primární soubor může měnit pouze pověřená osoba - správce slovníku (typicky jeden, aby nevznikaly konflikty). U něj se shromažďují soubory změn od jednotlivých autorů. Správce slovníku projde každý soubor změn v režimu revize a odstraní nevhodné změny. Potom soubor vyexportuje do XML formátu. Tím získá nový primární soubor s aplikovanými změnami z daného souboru změn. Tímto souborem nahradí aktuální primární soubor (změna nelze provádět za běhu editoru, protože editor má primární soubor stále otevřený). Při dalším spuštění editoru je přestavěn index a je zobrazen nový slovník s provedenými změnami. Tento postup opakuje správce slovníku pro každý soubor změn.

Po změně primárního souboru by měla pověřená osoba zajistit jeho distribuci mezi všechny autory, aby všichni pracovali nad aktuální verzí. V případě použití sdíleného primárního souboru tento problém

odpadá. Správce slovníku však v takovém případě musí zajistit, aby v době aktualizace nebyl primární soubor nikým používán.

Při aplikaci souboru změn na jiný primární soubor, než na kterém byl soubor změn vytvořen, mohou nastat konflikty, které způsobí, že se změny v režimu revize zobrazí jinak, než na původním primárním souboru. Jejich význam a vliv na primární soubor však zůstane zachován. Pokud jsem například přidal slovo do primárního souboru a někdo jiný donese soubor změn, kde toto slovo také přidal, zobrazí se mi jeho přidání jako změna toho mého již vloženého slova. Když se totéž stane se smazaným slovem, slovo se v seznamu lemmat zobrazí jako smazané, ale po jeho výběru se nezobrazí původní slovníkové heslo, ale prázdné slovníkové heslo (protože informace o smazaném lemmatu již nikde neexistuje). V režimu editace se tyto případné konflikty nijak neprojeví - všechny změny se správně promítnou do slovníku.

V případě potřeby si kdokoliv může vyexportovat svoji kopii slovníku s provedeným libovolným souborem změn a to jak do XML tak do starého formátu.

Popisovaný scénář je pouze doporučený způsob práce s editorem. Editor lze samozřejmě využívat i jiným způsobem.

# Dodatek A

## DTD používaných dokumentů

### DTD seznamu značek.

```
<!ELEMENT tagcatalog      (tagentry)+>
<!ELEMENT tagentry       (tag, shorthand)>
<!ELEMENT tag            (#PCDATA)>
<!ELEMENT shorthand      (#PCDATA)>
```

### DTD seznamu ohýbacích vzorů.

```
<!ELEMENT patterncatalog (patternentry)+>
<!ELEMENT patternentry   (form)+>
<!ELEMENT form           (ending, tag+)>
<!ELEMENT ending         (#PCDATA)>
<!ELEMENT tag            (#PCDATA)>

<!ATTLIST patternentry   name      CDATA      #REQUIRED
                        neg        (yes|no) #REQUIRED>
```

### DTD seznamu derivačních vzorů.

```
<!ELEMENT derivationcatalog (derivationentry)+>
<!ELEMENT derivationentry   (rule)+>
<!ELEMENT rule              (kofixending, pattern, derivation, lemmaending,
                             refererderivation, refererlemmaending)>
<!ELEMENT kofixending       (#PCDATA)>
<!ELEMENT pattern           (#PCDATA)>
<!ELEMENT derivation        (#PCDATA)>
<!ELEMENT lemmaending       (#PCDATA)>
<!ELEMENT refererderivation (#PCDATA)>
<!ELEMENT refererlemmaending (#PCDATA)>

<!ATTLIST derivationentry name      CDATA      #REQUIRED>
<!ATTLIST rule            createreferer (yes|no) #REQUIRED>
```

**DTD morfologického slovníku.**

```
<!ENTITY % list      "#PCDATA | S">
<!ENTITY % list_attr  "cnt      CDATA      #IMPLIED">

<!ELEMENT dictroot   (entry)*>
<!ELEMENT entry      (stat, rt, pat, lstag?, lslem,
                      lsder?, lssyn?, lssem?, lssty?, lscom?,
                      origfile, created, modified?)+>
<!ELEMENT stat       EMPTY>
<!ELEMENT rt         (#PCDATA)>
<!ELEMENT pat        (#PCDATA)>
<!ELEMENT origfile   (#PCDATA)>
<!ELEMENT created    (#PCDATA)>
<!ELEMENT modified   (#PCDATA)>
<!ELEMENT lstag      (%list;)*>
<!ELEMENT lslem      (%list;)*>
<!ELEMENT lssty      (%list;)*>
<!ELEMENT lssyn      (%list;)*>
<!ELEMENT lssem      (%list;)*>
<!ELEMENT lscom      (%list;)*>
<!ELEMENT lsder      (%list;)*>

<!ELEMENT S          EMPTY>
<!ATTLIST entry      ver      CDATA      #IMPLIED
                  operation (delete|undo) #IMPLIED>
<!ATTLIST stat       val      (A)        #REQUIRED>
<!ATTLIST created    uid      CDATA      #REQUIRED>
<!ATTLIST modified   uid      CDATA      #REQUIRED>
<!ATTLIST lstag      %list_attr;>
<!ATTLIST lslem      %list_attr;>
<!ATTLIST lssty      %list_attr;>
<!ATTLIST lssyn      %list_attr;>
<!ATTLIST lssem      %list_attr;>
<!ATTLIST lscom      %list_attr;>
<!ATTLIST lsder      %list_attr;>
```

# Doporučená literatura

- [1] *Manual for Morphological Annotation*, Jiří Hana a Hana Hanová, CKL Technical Report TR-2002-14, Charles University, Czech Republic , 2002, dostupné v HTML <<http://ufal.mff.cuni.cz/pdt/Corpora/PDT.1.0/References/mman.html>>.
- [2] *Disambiguation of Rich Inflection*, Jan Hajič, Karolinum, Prague, Czech Republic , 2001.
- [3] *Manuál ke starší verzi morfologického editoru (MCLASS)*, Jan Hajič, dostupné na instalačním cd, v souboru doc/mclass.il2.
- [4] *Finite State Morphology*, Kenneth R. Beesley a Lauri Karttunen, CSLI Publications, 2003, dostupné v HTML <<http://www.fsmbbook.com>>.
- [5] *Algorithms and Data Structures Research and Reference Material*, L. Allison, Monash University, Australia , dostupné v HTML <<http://www.csse.monash.edu.au/~lloyd/tildeAlgDS/>>.