

Bonito verze 1.4 – Příručka uživatele

Pavel Rychlý, TextForge

Květen, 2003

1 Úvod

Bonito je grafické uživatelské rozhraní korpusového manažeru Manatee. Umožňuje tvořit dotazy a klást je na různé korpusy. Výsledek je potom přehledně zobrazen, je možné ho různě upravovat a lze na něm počítat rozličné statistiky.

1.1 Korpus

Korpus je zde chápán jako posloupnost tzv. *pozic*. Každou pozici tvoří jedno slovo, číslo, interpunkční znaménko apod. Vlastní rozdělení do pozic je u většiny korpusů provedeno automaticky externími nástroji, nijak nezávisí na systémech Manatee či Bonito a může se tedy korpus od korpusu lišit.

Každá pozice je tvořena množinou *pozičních atributů*. Každý atribut obsahuje nějakou jednoduchou textovou informaci (slovo, základní tvar, slovní druh atd.). Pozice libovolného korpusu obsahuje vždy minimálně atribut se jménem *word*, který obsahuje vlastní slovo na dané pozici. Různé korpusy obsahují různé soubory atributů. Některé obsahují pouze zmíněný atribut *word*, jiné obsahují u každého slova jeho základní tvar (atribut *lemma*) a gramatickou informaci (atribut *tag*). Některé korpusy mají gramatickou informaci rozdělenou do více specifických atributů.

V korpusu mohou být zaznamenány také různé *strukturní značky* jako například hranice vět, odstavců, dokumentů apod. Některé typy těchto značek mají navíc u každého výskytu značky v textu uloženu další informaci. Například mnoho korpusů má celý text rozdělený na dokumenty strukturní značkou `<doc>`, která většinou obsahuje identifikátor zdroje dokumentu.

U některých značkových korpusů je zvykem gramatickou informaci uloženou u každého slova nazývat značkou. Tento typ značek ovšem netvoří výše definované strukturní značky, naopak jsou značky (gramatické informace) uloženy v některém z pozičních atributů.

1.2 Korpusový manažer

Výsledkem dotazu na korpus je tzv. *konkordanční seznam*, který tvoří všechny pozice korpusu odpovídající zadanému dotazu. Konkordanční seznam se potom zobrazuje ve formátu *KWIC* (Key Word(s) In Context), kdy jsou hledaná slova se svými kontexty zobrazena přehledně pod sebou. Konkordanční seznam je někdy zkráceně nazýván *konkordancí*. Zkratka **KWIC** reprezentuje v programu hledané slovo či posloupnost slov.

2 Vzhled aplikace

Hlavní okno manažeru obsahuje v horní části menu a pod ním vstupní pole pro zadávání dotazů. Vpravo na konci vstupního řádku je tlačítko s aktuálním jménem korpusu, po stlačení tlačítka je možné vybrat ze zobrazeného seznamu jiný korpus.

Většinu plochy okna tvoří konkordanční seznam, ve kterém se zobrazují výsledky dotazu. Ve spodní části je potom oblast detailu, ve kterém je zobrazen zvolený KWIC s větším kontextem a kontext je možné interaktivně zvětšovat: **šipka nahoru** rozšíří kontext doleva, **šipka dolů** naopak doprava. Na úplně spodním okraji hlavního okna je stavový řádek, který informuje o právě probíhající operaci. Pokud žádná operace neprobíhá, informuje o zobrazených řádcích.

Informace o zobrazených řádcích je za slovem *Zobrazeno* a může mít dvě podoby:

- Pokud jsou zobrazeny všechny řádky výsledku, obsahuje stavový řádek slovo *vše* následované lomítkem (/) a počtem řádků. (např. *vše/13*).
- Pokud je zobrazena jenom část výsledku, informace o zobrazených řádcích mají na první pohled podobu složitějšího matematického výrazu (např. *141+50/363 (13%)*). První číslo (před znakem plus (+)) označuje, od kterého řádku jsou řádky zobrazeny, dále následuje počet zobrazených řádků, za lomítkem (/) je potom celkový počet řádků výsledku a v závorkách je počet procent zobrazených řádků vzhledem ke všem řádkům výsledku. Uvedený příklad tedy můžeme číst jako: *od řádku 141 je zobrazeno 50 řádků z celkových 363, což je 13%*

Stavový řádek dále za slovem *Řádek* obsahuje číslo aktuálního řádku a pokud je nějaký řádek vybrán, následuje za slovem *Vybráno* počet vybraných řádků.

Pro přesun mezi jednotlivými vstupními poli zadávání dotazu, konkordančním seznamem a detailem je možné použít **tabelátor**, klávesa **F12** přepíná mezi konkordančním seznamem a vlastním vstupním polem dotazu. Mezerník na nabídkovém tlačítku zobrazí nabídku (menu) se všemi dostupnými hodnotami.

3 Hlavní menu

Tato kapitola popisuje postupně všechny položky menu. V textu je používáno označení **Menu>Položka** pro položku *Položka* z menu *Menu*.

3.1 Manažer

Připojení

Provede připojení na server. Bez připojení nelze provádět žádné operace s korpusem. Je možné se připojit buď na vzdálený server pomocí Internetu/intranetu, nebo spustit server lokálně. Pokud je v konfiguraci nastaveno spuštění serveru lokálně, provede se po startu Bonita automaticky připojení. V opačném případě se automaticky vybere tato položka menu a v následně zobrazeném okně je nutné zadat jméno a heslo uživatele.

Připojit přes internet

Připojení se provede na vzdálený server, jehož adresa je ve vstupním řádku *Adresa serveru*, s použitím zadaného jména a hesla.

Spustit lokálně

Server se spustí lokálně, tedy na stejném počítači, na kterém běží Bonito. V tomto případě musí být celý manažer i se všemi použitými korpusem nainstalován na daném počítači. Jméno a heslo uživatele je ignorováno a přístup ke zvoleným korpusem je dán přístupovými právy k datovým souborům jednotlivých korpusem. Pro spuštění serveru se použije příkaz ve vstupním řádku *Příkaz serveru*.

Adresa serveru:

Adresa počítače, na kterém server běží, například `aurora.fi.muni.cz`.

Uživatel, Heslo:

Jméno (login) a heslo uživatele pro přístup ke Bonito (může se lišit od jména i/nebo hesla pro běžný přístup) na daném serveru. Text hesla se ve vstupním řádku nezobrazuje.

Příkaz serveru:

Příkaz, kterým se provede spuštění serveru (například: `manateesrv "desam susanne" 2>errlog`).

Změna hesla

Změna hesla pro přístup ke Bonito. Je nutné zadat současné (původní) heslo a dvakrát heslo nové. Text žádného z hesel se ve vstupních řádcích nezobrazuje.

Obnovení výsledku

Provede opětovné zobrazení řádků výsledku (podle zadání v **►Zobrazení►Rozsah**).

Nastavení

Umožní nastavit různé uživatelské volby aplikace, případně uložit volby pro příští spuštění. Implicitní hodnoty voleb jsou nastaveny tak, aby vyhovovaly většině uživatelů. Pokročilí uživatelé mohou měnit tyto volby:

- *Zobrazování vysvětlivek* – odpovídá položce menu **►Nápověda►Zobrazení vysvětlivek**
- *Příkaz pro spuštění prohlížeče (s nápovědou)* – příkaz systému, který zajistí spuštění WWW prohlížeče se zadanou stránkou. Používá se pro zobrazení on-line nápovědy (menu **►Nápověda►Dokumentace**).
- *Maximální počet kroků zpět* – viz menu **►Konkordance►Zpět k předchozí**
- *Počet dotazů v historii* – udává počet dotazů, které budou v historii uloženy. Při zadávání dalších dotazů se nejstarší z historie dotazů odstraňují.
- *Počet "nových" řádků při posunu o stránku* – viz menu **►Zobrazení►O stránku nahoru** a **►Zobrazení►O stránku dolů**.
- *Počet najednou načtených řádků při zobrazování*.
Zobrazování řádků konkordančního seznamu probíhá průběžně po zadaném počtu řádků, jak přicházejí data ze serveru. Pokud máte rychlejší (novější) počítač, může nastavení větší hodnoty urychlit celkové zobrazení velkého počtu (více než 1000) řádků. Zrychlení se projeví, pouze má-li Bonito dostatek dat ze serveru. Pokud tedy máte pomalé spojení se serverem (připojení po modemu) nebo jste zadali složitý dotaz, který server vyhodnocuje pomalu, Bonito zobrazí vždy pouze ty řádky, které má k dispozici, a změna nastavení nebude mít na rychlost zobrazení žádný vliv.
- *Připojit k internetovému serveru* – pokud je zatrženo, bude v přihlašovacím okně vybrána volba *Připojit přes internet*, v opačném případě bude vybráno *Spustit lokálně*.
- *Adresa serveru, Číslo portu na serveru* – internetová adresa počítače a číslo portu, na kterém běží serverová část manažeru (manateesrv).
- *Příkaz pro spuštění serveru lokálně* – implicitní hodnota pro vstupní pole *Příkaz serveru* v připojovacím okně (viz menu **►Manažer►Připojení**)
- *Počáteční korpus*. Po připojení na server je zvolen první z dostupných korpusů. Pokud většinou pracujete z jiným korpusem, napište jeho jméno do tohoto vstupního řádku a zmáčkněte **Uložit**. Při dalším spuštění Bonito bude zvolen zadaný korpus. Jméno musí přesně odpovídat jménu korpusu a to i ve velikosti písmen.
- *Počáteční velikost kontextu* – viz menu **►Zobrazení►Kontext**.
- *Počáteční rozsah zobrazení* – viz menu **►Zobrazení►Rozsah**.

Tlačítkem **Použít** se nastavené hodnoty uchovají pouze do ukončení aplikace. Tlačítkem **Uložit** se hodnoty navíc uloží do konfiguračního souboru a budou použity i při dalším spuštění aplikace. Některé hodnoty (*Počáteční ...*) mají smysl pouze pro uložení.

Změna jazyka

Volba jazyka uživatelského rozhraní. Zvolený jazyk bude použit i při příštím spuštění programu.

Konec

Ukončení aplikace.

3.2 Korpus

Souhrnné informace

Zobrazí souhrnné informace o vybraném korpusu: jeho jméno a případná dodatečná informace o korpusu z jeho konfiguračního souboru, velikost (celkový počet pozic), poziční atributy a strukturní značky. Pro každý poziční atribut je zobrazen počet různých prvků ve slovníku daného atributu (dle atributu tedy počet různých slov, základních tvarů, značek atd.). Pro každou strukturní značku je zobrazen její celkový počet výskytů v korpusu.

Statistiky

Pro zadané hodnoty atributů vypočte a zobrazí četnosti hodnot v celém korpusu, četnost bigramu v korpusu, množství vzájemné informace (MI-score) a T-score.

Bigramy se hledají s použitím zadaného okna. Implicitní hodnoty *od 1 do 1* znamenají, že má druhé slovo přímo následovat první slovo. *Od 1 do 5* znamená velikost okna 5, tedy mezi prvním a druhým slovem mohou být až 4 jiné pozice. Záporná znaménka znamenají opačné pořadí výskytu. *Od -1 do -1* tedy znamená, že druhé slovo má přímo předcházet první slovo. *Od -5 do 5* potom znamená, že slova mají být od sebe vzdálená maximálně 5 pozic a mohou být v libovolném pořadí.

Vytvoření subkorpusu

Ze zvoleného korpusu vytvoří subkorpus podle zadaných podmínek. Subkorpus je možné použít pouze pro hledání. Pro výpočet všech statistik se používá vždy celý korpus.

Základní korpus:

Jméno korpusu, ze kterého se bude subkorpus vytvářet. Jedná se o aktuální korpus z hlavního okna aplikace.

Jméno subkorpusu:

Po úspěšném vytvoření subkorpusu se v seznamu korpusů objeví nový korpus ve tvaru *základní korpus:jméno subkorpusu*. Pokud například vytvoříme subkorpus se jménem *press* z korpusu *susanne* v seznamu korpusů bude *susanne:press*.

Značka:

Jméno značky (strukturního atributu), podle které se má subkorpus vytvářet. V subkorpusu budou potom všechny pozice, které jsou součástí značky vyhovující zadané podmínce. Většinou se tvoří subkorpora podle značky nejvyšší úrovně, které odpovídají jednotlivým dokumentům či textům korpusu.

Podmínka:

Podmínka omezující některé atributy zadané značky. Například v korpusu *Susanne* jsou tiskové zprávy v dokumentech, jejichž jména začínají na *A*. Dokumentům odpovídá značka *<doc>*, jména dokumentů jsou v atributu *file*. Pro vytvoření subkorpusu tiskových zpráv z korpusu *Susanne* tedy zapíšeme podmínku:

```
file = "A.."
```

Smazání subkorpusu

Zobrazí aktuální seznam subkorpusů, ze kterého je možné zvolené subkorpora mazat.

Seznam slov

Zobrazí všechna slova ze slovníku zadaného atributu, která vyhovují danému vzoru.

Implicitní atribut

Umožňuje změnit implicitní atribut, podle kterého se provádí vyhodnocování dotazů bez zadaného jména atributu.

3.3 Dotaz

Vyhodnocení

Provede vyhodnocení dotazu (resp. P filtru, N filtru nebo kolokace). Vyhodnocení probíhá na serveru, Bonito pouze čeká na předání výsledku. Operaci lze zastavit pomocí tlačítka **Stop**, které se aktivací funkce zobrazí vpravo od jména korpusu.

Seznam pojmenovaných dotazů

Zobrazí seznam pojmenovaných dotazů, ze kterého je možné vybrat dotaz pro vyhodnocení, popř. smazat některý dotaz z tohoto seznamu.

Seznam předloh

Zobrazí seznam předloh (viz kapitola 4. *Předlohy*), ze kterého lze vybrat předlohu pro vyhodnocení, smazat některou předlohu z tohoto seznamu, přidat novou předlohu, popř. změnit text nebo popis vybrané předlohy.

Grafické vytváření

Zobrazí okno, ve kterém je možné vytvářet dotaz pomocí grafických prvků.

Vytvoření předlohy

Přidá do seznamu předloh novou předlohu.

Jméno:

jméno předlohy, které se potom bude používat v dotazech

Předloha:

vlastní text předlohy

Popis:

volitelný popis předlohy

Import předloh

Po zvolení souboru s předlohami (například od jiného uživatele) přidá nové předlohy do seznamu předloh. Soubory s předlohami mají koncovku `tpl`.

Import pojmenovaných dotazů

Po zvolení souboru s dotazy (například od jiného uživatele) přidá nové dotazy k pojmenovaným dotazům. Soubory s dotazy mají koncovku `qry`. Importovat je možné i soubor s historií (`history.qry`), ze kterého se vyberou pouze pojmenované dotazy.

3.4 Konkordance

Souhrn

Zobrazí souhrnné informace o konkordanci: jméno korpusu, velikost konkordance a seznam jednotlivých operací (počínaje dotazem), které vedly k vytvoření konkordance.

Uložení

Uloží zvolené řádky konkordančního seznamu do souboru.

Kódování:

výběr kódování znaků uložených řádků (hodnota ‘-’ znamená kódování, ve kterém je korpus uložen na serveru)

Hlavička:

výběr formátu hlavičky pro uložení informací o dotazu

zobrazené řádky:

uloží pouze aktuálně zobrazené řádky

všechny řádky:

uloží všechny řádky konkordančního seznamu

číslovat řádky:

při zaškrtnutí budou ve výstupním souboru jednotlivé řádky očíslovány

zarovnat KWIC:

při zaškrtnutí budou klíčová slova zarovnána pod sebou

Tlačítkem **Kontext** je možné změnit kontext pro uložení, který je implicitně stejný s kontextem pro zobrazování (viz ▶**Zobrazení**▶**Kontext**). Po stlačení tlačítka **Uložit** je nutné zadat nebo vybrat jméno souboru, do kterého se má konkordanční seznam uložit.

Tisk

Vytiskne zvolené řádky konkordančního seznamu na tiskárně, zadávají se stejné informace jako pro ▶**Konkordance**▶**Uložení**.

Smazání vybraných

Smaže označené řádky konkordančního seznamu (viz menu ▶**Výběr**). Počet vybraných řádků je zobrazen ve stavovém řádku u spodního okraje hlavního okna. V závislosti na aktuálním rozsahu zobrazených řádků některé vybrané řádky nemusí být právě zobrazeny.

Redukce

Redukce počtu řádků konkordančního seznamu. Zadává se, které řádky mají ve výsledku zůstat a kolik řádků, procent či setin procenta z původního počtu řádků má zůstat.

Jednoduché třídění

Setřídí konkordanční seznam podle zadaných kritérií.

Počet tříděných pozic:

počet pozic, podle kterých se bude třídit

Třídít:

Udává, které pozice z řádku se budou při třídění porovnávat. V následujícím schématu předpokládáme *Počet tříděných pozic* = 3. Číslo 1 znamená nejvýznamnější pozici (bude se podle ní třídit nejdříve), číslo 3 nejméně významnou pozici (bude se podle ní třídit až nakonec).

	<KWIC KWIC>
levý kontext	3 2 1	<	>
KWIC odleva		<1 2 3	>
KWIC odprava		<	3 2 1>
pravý kontext		<	> 1 2 3

Ignorovat velikost:

Při třídění nebude záležet na velikosti písmen.

Retrográdně:

Jednotlivá slova se budou třídit od posledního písmene k prvnímu.

Atribut:

Podle kterého pozičního atributu se má třídění provádět – výběr z možných pozičních atributů daného korpusu.

Obecné třídění

Seřadí konkordanční seznam podle seznamu třídících podmínek. Každá třídící podmínka určuje jednu pozici, podle které se budou jednotlivé řádky porovnávat.

Vynechat duplicitu:

Pokud je zatrženo, ze všech řádků, které se shodují ve všech třídících intervalech, se ve výsledku ponechá pouze jeden.

Přidat:

přidá další třídící podmínku

Smazat:

smaže vybraný interval (výběr se provádí kliknutím myši)

Budiž:

provede třídění

Zavřít:

zavře okno bez jakéhokoli třídění

Každá třídící podmínka obsahuje:

Třídít pozici:

číslo pozice, která se bude porovnávat. Opět záporná čísla znamenají pozice *před*, kladná znamenají za zvoleným počátkem.

od:

udává počátek počítání pozic stejně jako při zadávání filtrů nebo kolokací (viz kap. 3. *Tvorba dotazu*).

Jméno tříděného atributu

z výběru pozičních atributů korpusu

AaAa:

ignorovat velikost písmen

udazdO:

Odzadu – třídít retrográdně

Četnostní třídění

Provede třídění podle průměrné četnosti slov z zadaném kontextu. Na prvních řádcích konkordance jsou po třídění kontexty obsahující nejčastější slova.

Třídění podle skupin

Seskupí řádky podle přiřazených čísel skupin. Jednotlivé skupiny budou setříděny vzestupně (první bude skupina číslo 1), na konci konkordance budou řádky, kterým nebyla žádná skupina přiřazena.

Zpět k předchozí

Provede návrat k předchozí konkordanci a zobrazí ji. Touto operací lze zrušit provedené změny v konkordanci (redukce, mazání, použití filtrů) nebo provedené třídění. Funkci je možné použít opakovaně, maximální počet uchovaných konkordancí lze měnit v ►**Manažer**►**Nastavení**, implicitně je 5.

Vpřed k následující

Zopakuje operaci, která byla zrušena akcí *Zpět k předchozí*.

Pojmenovat

Pojmenuje zobrazenou konkordanci. Pojmenované konkordance jsou přímo dostupné v menu **►Konkordance►Pojmenované**.

Pokud potřebujeme pracovat opakovaně s výsledkem jediného dotazu, pojmenujeme si výslednou konkordanci a můžeme se k ní kdykoliv rychle vrátit, aniž by se musel znovu vyhodnocovat dotaz.

Smazání pojmenovaných

Slouží k zrušení některých pojmenovaných konkordancí. Zobrazí se seznam všech pojmenovaných konkordancí. Výběrem ze seznamu a stlačením tlačítka **Smazat** se zvolená konkordance zruší.

3.4.1 Konkordance ► Statistiky

Frekvenční distribuce

Spočítá četnosti slov nebo jiných atributů nebo jejich posloupností na zadaných pozicích.

Limit:

Do výsledku se budou počítat a zobrazovat pouze posloupnosti s četností vyšší než je zadaný limit. Implicitní limit 0 znamená, že se budou počítat všechny hodnoty. Pro konkordanční seznam s velkým počtem řádků to může znamenat velké množství předávaných dat ze serveru a následné náročné třídění výsledku. Pokud se má zobrazit více řádků (konkrétní číslo závisí na výkonu počítače, běžně více než několik tisíc) může jejich zobrazení trvat příliš dlouho.

Každá podmínka obsahuje:

Atribut:

jméno atributu (z výběru pozičních atributů korpusu)

Pozice:

číslo pozice, která se bude porovnávat.

Práce se seznamem je stejná jako v případě Obecného třídění. Po stisku tlačítka [Budiž] se provede výpočet a zobrazí se okno s výsledky.

V zobrazeném okně lze dodatečně měnit způsob zobrazení pomocí následujících ovládacích prvků:

Limit:

Analogicky jako u zadání se ve výsledku nezobrazí ty řádky, které mají spočítanou četnost menší nebo rovnu zadanému limitu. Počet zobrazovaných řádků se vždy spočítá a ukáže vedle.

Pro každou zadanou pozici je možné vybrat ze tří možností:

zobrazit

Slova na dané pozici se budou normálně zobrazovat.

součet

Slova na dané pozici se budou zobrazovat a navíc se pro ně bude počítat mezisoučet. Pro poslední pozici jsou volby *zobrazit* a *součet* totožné, protože se pro ni součet počítá vždy.

schovat

Daná pozice se v přehledu vůbec nebude počítat a zobrazovat.

Po libovolné změně limitu nebo zobrazovacích podmínek se provede nové zobrazení výsledku. Šířku jednotlivých sloupců lze upravit tažením nadpisu sloupce levým tlačítkem myši.

Kolokace

Vypočítá nejčtenější kolokace v zadaném kontextu podle následujících parametrů:

Atribut

Jméno atributu (z výběru pozičních atributů korpusu).

V rozsahu od, do

Počáteční resp. konečná pozice kontextu. Kladné hodnoty se počítají doprava od konce KWIC, záporné se počítají od začátku KWIC doleva.

Minimální četnost v korpusu

Statistiky se budou počítat pouze pro slova s celkovou četností v korpusu vyšší než zadaná četnost.

Minimální četnost v daném rozsahu

Statistiky se budou počítat pouze pro slova s četností v zadaném kontextu vyšší než zadaná četnost.

Maximální počet zobrazených řádků

Pokud je ve výsledku více řádků, zobrazí se pouze zadaný počet nejčtenějších.

Setřídít podle četnosti

Udává třídění, podle kterého se zobrazí výsledné řádky. Je důležité pouze pro výběr nejčtenějších řádků (viz předchozí parametr), zobrazené řádky lze potom třídit podle kterékoliv statistiky.

Třídění podle **absolutní** četnosti je blízké T-score, třídění podle **relativní** četnosti je shodné s MI-score. (Podrobnější popis níže.)

Výsledek je zobrazen ve formě tabulky, kterou lze uložit do souboru pomocí tlačítka **Uložit**. Tabulku je možné setřídít dle libovolného sloupce kliknutím pravým tlačítkem myši na nadpis požadovaného sloupce. Šířku jednotlivých sloupců lze upravit tažením nadpisu sloupce levým tlačítkem myši.

Hodnoty v jednotlivých sloupcích mají následující význam:

1. sloupec

je nadepsaný jménem počítaného atributu (např. *word*). Obsahuje hodnoty daného atributu (např. slova), pro které se statistiky počítaly.

MI-score

– MI-score (vzájemná informace) slova a konkordance

T-score

– T-score slova a konkordance

Rel. f

– relativní četnost slova (kolik procent ze všech výskytů slova v korpusu je ve zvoleném kontextu)

Abs. f

– absolutní četnost slova (kolikrát se slovo objevilo v zadaném kontextu)

Stiskem pravého tlačítka myši na zvoleném slově 1. sloupce se zobrazí lokální menu obsahující dvě položky: *P filtr* a *N filtr*. Výběrem jedné z položek se aplikuje příslušný filtr na aktuální konkordanci.

Rozložení

Zobrazí okno, které uvádí počet řádků výsledku (četnost) a tzv. *redukovanou četnost* konkordance. Dále je v něm graficky zobrazeno rozložení jednotlivých řádků výsledku v rámci celého korpusu. Na ose *x* jsou jednotlivé pozice korpusu, na ose *y* je počet výskytů v daném místě korpusu.

Pokud jsou jednotlivé řádky konkordančního seznamu v celém korpusu rovnoměrně rozloženy jsou jednotlivé čáry v grafu stejně dlouhé a jsou zobrazeny rovnoměrně po celé délce okna. Pokud naopak je většina řádků pouze z jednoho "místa" korpusu (například z jednoho dokumentu) je v jenom místě okna výrazně více delších čar.

Okno rozložení může sloužit i pro rychlé přemístění do "zajímavé" části konkordance. Kliknutím na libovolnou čáru v grafu se aktuální řádek konkordančního seznamu přemístí na řádek odpovídající zvolené čáře pod kurzorem myši.

3.5 Zobrazení

Reference

Určuje, jaké reference budou u jednotlivých řádků výsledku zobrazeny. Pokud je zatrženo *číslo pozice*, zobrazí se číslo pozice začátku KWIC. Zatržením jména značky (například *doc*) se zobrazí pořadové číslo dané značky v korpusu, ve které se KWIC nachází (například *doc#91* znamená, že se dané KWIC nachází v 91. dokumentu od začátku korpusu). Při zatržení jména atributu značky (například *doc.file*) budou zobrazeny hodnoty daného atributu (například *doc.file=A03* znamená, že dané KWIC se nachází v dokumentu, v němž má atribut *file* hodnotu *A03*). Reference se zobrazují zelenou barvou na začátku každého řádku.

Atributy

Určuje, které poziční atributy se mají zobrazovat.

Pouze v KWIC:

Vybrané atributy se budou zobrazovat pouze pro pozice v KWIC.

Pro všechny pozice:

Vybrané atributy se budou zobrazovat pro pozice v KWIC i pro všechna slova v zobrazeném kontextu.

Značky

Určuje, které značky (strukturní atributy) se budou zobrazovat. U značek, které obsahují atributy (například identifikátory vět *id* u značky *<s>*), je možné zatrhnout zvolený atribut. Potom se budou hodnoty tohoto atributu zobrazovat v příslušné značce (například *<s id=12/3>*).

Kontext

Určuje, v jakém kontextu se mají slova zobrazovat. Pro pravou i levou stranu může být jednotkou *znak*, *pozice* nebo libovolná ze strukturních značek. Pokud je kontext ve znacích (jednotka *znak*), zobrazí se celá slova tak, aby byl zobrazen minimálně zadaný počet znaků.

Rozsah

Udává, které řádky a v jakém počtu se mají zobrazovat. Pokud je počet zadaných řádků roven 0, zobrazují se všechny řádky.

O stránku nahoru/dolů

Načtou a zobrazí se další (nezobrazené) řádky (nová stránka) před resp. za zobrazenými řádky. Počet těchto řádků lze měnit v **Manažer>Nastavení**, implicitně je 20.

Skok na řádek

Zobrazí zvolený rozsah řádků (viz **Zobrazení>Rozsah**) od zadaného řádku. Číslo řádku, od kterého jsou řádky zobrazeny, je uveden ve stavovém řádku před znakem +.

3.6 Výběr

Vybrané řádky je možné přenést do schránky (viz níže), nebo smazat z konkordančního seznamu (menu **Managér>Smazání vybraných**). Vybrané řádky jsou v hlavním okně odlišeny modrým pozadím. Jejich počet je zobrazen ve stavové řádce na spodním okraji hlavního okna.

Výběr lze provádět myší nebo klávesami. **Levé tlačítko myši** nebo **mezerník** vybere nevybraný řádek, u vybraného řádku výběr zruší. **Shift+levé tlačítko myši** vybere všechny řádky mezi zvoleným řádkem a řádkem aktuálním.

Všechny

Označí všechny řádky.

Žádné

Zruší veškeré označení.

Inverze

Označí neoznačené řádky a zruší označení u označených řádků.

Vložit do schránky

Označené řádky vloží do schránky (clipboard) pro použití v jiných aplikacích.

Předat CEDu

Předá vybrané řádky korpusovému editoru CED. Tato funkce je dostupná pouze na platformě UNIX.

3.7 Nápořěda

O aplikaci

Zobrazí stručnou informaci o aplikaci a číslo verze.

Zobrazení vysvětlivěk

Při zapnutém stavu se automaticky zobrazují krátké vysvětlivky pro některé prvky uživatelského rozhraní, pokud se nad nimi na krátkou chvíli zastaví kurzor myši.

Dokumentace

Spustí se prohlížeč s dokumentací.

Licence

Zobrazí licenci, pod kterou je možné program používat.

4 Tvorba dotazu

Dotaz může být libovolný dotaz dotazovacího jazyka nebo šablona (viz níže).

Typ dotazu:**Nový dotaz**

(Ctrl-Q) Vytvoří základní konkordanční seznam, který lze potom upravovat (aplikovat filtry, mazat, redukovat nebo třídít řádky atd.).

P filtr

(Ctrl-P) Pozitivní filtr – v konk. seznamu zůstanou pouze řádky, které v zadaném intervalu obsahují nějaké hledané pozice

N filtr

(Ctrl-N) Negativní filtr – v konk. seznamu zůstanou pouze řádky, které v zadaném intervalu neobsahují žádné hledané pozice

kolokace

Pro každý řádek nalezne v zadaném intervalu jeden výskyt hledaných pozic. Tyto pozice jsou potom v konk. seznamu zvýrazněny.

První/Poslední

Pro kolokace udává, kolikrát výskyt v zadaném intervalu se má vzít jako kolokace.

Od, Do:

Začátek a konec intervalu pro P/N filtr nebo hledání kolokace.

Číslo udává počet jednotek. Kladná čísla znamenají počet jednotek doprava (před zvolený počátek), záporná doleva (za zvolený počátek). Číslo 0 znamená přesně danou pozici. Jednotka je podle výběru buď pozice (*Pozic*), nebo strukturní značka. Poslední výběrové tlačítko *od* udává počátek, odkud se má daný počet jednotek počítat. Hodnotu lze vybrat z následujících možností:

< **KWIC**

– od začátku nalezených slov

KWIC >

– od konce nalezených slov

< **n.kol.**

– od začátku n-té kolokace

n.kol. >

– od konce n-té kolokace

Každý zadaný dotaz se uloží do historie dotazů. Neukládá se pouze v případě, pokud je dotaz sám i jeho jméno zcela shodný s předešlým dotazem. Ve vstupním řádku dotazu lze používat kurzorové šipky nahoru a dolů za účelem procházení této historie.

Pokud u dotazu zadáme jméno (následující vstupní pole za dotazem), uloží se dotaz do seznamu pojmenovaných dotazů

Seznam pojmenovaných dotazů, seznam předloh a historie se automaticky ukládají na disk pro příští spuštění při ukončení aplikace (►**Manažer**►**Konec**).

5 Předlohy

Předloha je určitý typ dotazu, který umožňuje snadné zadávání opakujících se dotazů nebo opakovaných dotazů stejného typu. Zjednodušení spočívá v tom, že složitý opakující se dotaz vytvoříme pouze jednou jako předlohu a při vlastních voláních pouze měníme hodnoty parametrů pro tuto předlohu.

5.1 Vytvoření předlohy

Předloha se skládá ze jména, vlastního textu předlohy a volitelného popisu. Jméno je jednoznačný identifikátor předlohy a používá se při vyvolání dotazu. Všechny tři položky se zadávají v okně ►**Dotaz**►**Vytvoření předlohy**. Text předlohy a popis je možné dodatečně modifikovat v seznamu předloh (►**Dotaz**►**Seznam předloh**).

V textu předlohy lze použít *proměnných*, za které se při aplikaci předlohy dosadí skutečné parametry. Tyto proměnné jsou tvořeny znakem dolar (\$) následovaným číslicí. První parametr se tedy bude nahrazovat za všechny výskyty proměnné \$1, druhý za všechny výskyty proměnné \$2, třetí za \$3 atd.

Například předloha pro všechny slovní tvary pravidelných anglických sloves, by mohla mít text:

```
[word="$1" | word="$1s" | word="$1ing" | word="$1ed"]
```

5.2 Použití předlohy

Vyvolání předlohy se zapisuje do stejného řádku jako běžný dotaz. Od běžného dotazu se odlišuje tím, že prvním znakem zleva je vykřičník (!), následuje jméno předlohy, dvojtečka (:) a mezerou oddělené případné parametry. Pokud by jméno výše uvedené předlohy bylo *verb*, zadání dotazu by vypadalo

```
!verb: help
```

6 Konkordanční seznam

Konkordanční seznam zobrazuje (některé) řádky z výsledku dotazu. Jednotlivá KWIC jsou zobrazena přehledně pod sebou. Pro navigaci v konk. seznamu lze použít následující klávesy a akce myši:

Kurzorové šipky nahoru/dolů:

Aktuální řádek se posune o jeden řádek nahoru resp. dolů.

PageUp/PageDown:

Aktuální řádek se posune o jednu stránku (počet řádků na obrazovce) nahoru resp. dolů.

Mezerník:

Změní se výběr aktuálního řádku.

Enter (Return), dvojklik:

Pro aktuální řádek se zobrazí větší kontext v okně detailu.

Ctrl+Enter (Ctrl+Return), kliknutí pravým tlačítkem

Pro aktuální řádek se zobrazí plný odkaz v okně detailu.

Control-PageUp/PageDown (šipky v pravém dolním rohu seznamu):

Načtou a zobrazí se další (nezobrazené) řádky (novou stránku) před resp. za zobrazenými řádky.

Pokud je v menu ▸**Zobrazení**▸**Rozsah** zvolen **Náhodný** rozsah zobrazovaných řádků šipka nahoru resp. dolů změní rozsah zobrazení na **První** resp. **Poslední**.

F12:

přepíná kurzor mezi vstupním řádkem dotazu a konkordančním seznamem.

6.1 Skupiny

Každému řádku je možné přiřadit číslo skupiny. Číslo skupin mohou být od 1 do 99. Skupiny 1 až 9 lze přiřazovat stiskem odpovídající klávesy. Skupiny 10 až 99 se přiřazují postupným stiskem tří kláves: nejdříve *Control-E* a potom dva stisky čísel reprezentujících dvouciferné číslo. Klávesou 0 se přiřazení skupiny zruší.