

Počítači, přečti mi pohádku

Stále častěji se začínají objevovat volně šířené utility pro rozpoznávání a syntézu řeči, které dotahují dostupné komerční produkty. Je bohužel fakt, že tato hlasová řešení jsou většinou nadstavby pro rozhodně ne dokonalé řečové enginy a na profesionální systémy, které se domácím uživatelům do ruky nedostanou, jednoduše nemají.

Bez ohledu na to, že doposud nedošlo ke kvalitativnímu průlomům v takové zajímavé a perspektivní oblasti, jakou je syntéza řeči či rozpoznávání hlasu a hlasové řízení, nové utility pro stávající technologie se objevují se záviděníhodnou pravidelností. Mezi nimi je možné narazit na produkty nejrůznějšího původu, bohužel českých programů zatím mnoho není. V přehledu, který je pro vás připraven, jsme si nevytyčili úkol ukázat jednoho vítěze - účastníci jsou určeni pro řešení různých úkolů, tudíž pracují různě. Každá z představených utilit má určité nedostatky a řešit jakékoli závažné úkoly pomocí hlasového řízení je nemožné. Ale ty funkce, které jsou k dispozici, jasně hovoří o tom, jakým směrem půjdou vývojáři metod hlasového řízení počítače. Myslím si, že možnost velet svému PC z pohodlného křesla nikdo neodmítne. Zvláště zmíníme TTS engine (Text To Speech) s podporou českého jazyka od firmy GB-SOFT. Po jeho instalaci máte možnost většinu následujících programů naučit mluvit česky.

A. Jevdokimov, P. Zákostelný

Užitečné odkazy

Jestliže některý z uvedených programů nebude fungovat, stáhněte si (či použijte náš lednový Chip CD) a nainstalujte si komplet hlasových knihoven Microsoft:

Pokud žádný zvuk neuslyšíte, přeinstalujte SAPI:

<http://activex.microsoft.com/activex/controls/sapi/spchapi.exe> (824 KB).

Engine pro rozpoznání řeči: <http://activex.microsoft.com/activex/controls/agent2/actcnc.exe> (6 MB).

Control panel přidá ikonu do Nastavení: <http://agent.microsoft.com/agent2/spchcpl.exe> (927 KB, ne pro XP).

Podpora řeči:

angličtina: <http://activex.microsoft.com/activex/controls/agent2/lhhttpseng.exe> (2,6 MB);

němčina: <http://activex.microsoft.com/activex/controls/agent2/lhhttpsged.exe> (2,2 MB);

čeština: <http://www.gb-soft.cz> (896 KB).

Co se týče programů třetích stran, pro ně je nutná zvláštní podpora:

<http://activex.microsoft.com/activex/controls/agent2/msagent.exe> (400 KB, v 2K/XP/Me již obsaženo); česká podpora pro dialogy (ne řeč): <http://activex.microsoft.com/activex/controls/agent2/AgtX0405.exe> (128 KB).

Zajímavé odkazy

CS-VOICE 97: český hlasový výstup, cca 5000 Kč, www.frog.cz

Ukázka profesionálních projektů - webové generování řeči (čte zadaný text)

www.speechtech.cz

www.bell-labs.com/project/tts/voices.html

www.research.att.com/projects/tts

Zpívající sólista

Nevím, jestli je to špatně, nebo dobře, ale programy založené na technologii MS Agent se vytvářejí hlavně pro zábavu dětí nejrůznějšího věku. Samozřejmě že jednotliví vývojáři se snaží přidat svým programům alespoň špetku respektu a závažnosti, ale kvůli tomu jsou často jejich výtvoři o to komičtější. Jiní se naopak vůbec nesnaží skrývat zábavný charakter svých výtvořů, jako například tvůrce utility MUSH. MUSH je zkratka z Music Scripting Helper. Tato utilita vytváří skripty pro hudební MIDI a KAR (text Karaoke) soubory. Používat MUSH není složité. Je zapotřebí pouze vybrat agenta (nejlépe shodného se základním), označit všechny body kromě La[180] the words, rozhodnout, kam uložit hotový klip, klepnout na Convert a najít vhodný soubor MIDI nebo KAR (Karaoke - hudební MIDI s textem) pro zpracování. Hodí se libovolná skladba, ale pouze s anglickými slovy - české texty MUSH nepodporuje. Mimochodem, když se klepne na tlačítko Find same Karaoke Files, v prohlížeči se otevře stránka s množstvím KAR souborů. Pokud se rozhodnete, kreslená postavička vám zazpívá vybranou písničku. Samozřejmě že to nebude takové jako od Luciana Pavarottiho, ale bude to svým způsobem originální. Pokud vám

interpretace vyhoví, MUSH uloží výsledek. Ze tří možných variant jsem si osobně vybral tu nejjednodušší a nejdostupnější - export do Internet Exploreru (Basic web output). Získaný HTML soubor bude obsahovat pouze jedno tlačítko Play. Když na toto tlačítko klepnete, objeví se zpívající pomocník. Program MUSH hledejte na stránkách výrobce: www.pepedog.com/mush.zip (132 KB).

Software na Chip CD 1/03

Následující programy a utility bohužel vzhledem k naplněnosti prosincového CD naleznete až na Chip CD 1/03.

řečové knihovny a enginy Microsoft, freeware
český řečový engine GB-SOFT, freeware
CyberBuddy 2.12, freeware
Web Talkster 3.9, freeware
Speakonia 1.03, freeware
GUI Reader 2.0, freeware
Voice Navigator 1.01, demo
MUSH 2.1.64, freeware

Web Talkster 3.9

Snad žádný program pracující ve Windows nemá tolik variací a následovníků jako MS Internet Explorer. Těžko říci, co je důvodem tohoto stavu - přes mnohé nedostatky samotného prohlížeče se může těšit ohromné popularitě mezi uživateli. Fakt zůstává, že se snad každý měsíc objevuje nějaká jeho další varianta. Podle mého názoru by bylo překvapivé, kdyby vývojáři hlasových programů zůstali stranou od všeobecného nadšení programující mládeže a nevytvořili by "čtecí zázrak" a la Internet Explorer. Příkladem může být Web Talkster - bezplatný prohlížeč s podporou technologií text-to-speech a MS Agent. Prohlížeč je samozřejmě vybudován na základu Internet Exploreru, ale zabírá přitom 1,5krát větší objem operační paměti než microsoftský prapředek. Okno tohoto prohlížeče, bez ohledu na velká tlačítka a možnost změny barev, neomylně také Explorer připomíná. Je pravda, že řada důležitých funkcí (kopírování, tisk, hledání na stránce atd.) je schována v menu Options - File, ale všechno ostatní je udělané jako v originálu, plus tři nová tlačítka: Speak (čtení veškerého textu na celé webové stránce), Selected (čtení pouze označených odstavců) a QuitSpk (zrušení ozvučení). Web Talkster používá sympatické animované postavičky - pomocníky. Podpora této technologie společně s podobou čaroděje Merlina se nastavuje během procesu instalace prohlížeče. Mimochodem, knihovny MS Speech API a anglická varianta řečového enginu Lernout & Hauspie budou také nabídnuty. Vlastní hlas, přesněji řečeno jeho zabarvení a rychlost čtení, je možné nastavit v menu Options - Speech. Bohužel nastavení je možné pouze přibližné. Proto na úrovni rychlosti Medium bude prohlížeč Web Talkster číst texty příliš rychle, ve variantě Slow přespříliš pomalu. Pomocí menu Options - Character Options můžete také optimalizovat svého pomocníka, vybrat jeho podobu (Select character) nebo ho úplně vypnout (Move - Off screen). Je zajímavé, že i bez zobrazení pomocníka bude Web Talkster pokračovat ve čtení stránek. Z doplňkových možností mluvícího prohlížeče bych chtěl upozornit na vestavěnou funkci hledání na Google.com (jako v prohlížeči Opera) a také na maličký zápisník pro poznámky - My Surf'n Notes.

Web Talkster 3.9s

Výrobce Code-it Software
Internet www.code-it.com
Velikost distribučního souboru 9 MB
Podmínky šíření freeware
Operační systém Windows 9x/ME/2000/XP

- + přidává hlasové funkce do Internet Exploreru
- + používá animované postavy
- nelze přesně nastavit parametry čtení
- vyžaduje mnoho operační paměti

CyberBuddy 2.12

Můj oblíbený Agent Reader (www.aldoatools.com) se v tomto roce stal sharewarem, bez boje odevzdal vedoucí pozici v třídě neplacených hlasových programů virtuálnímu sekretáři CyberBuddy, jehož výrobcem je John DeFino. A právem. Tato utilita byla v posledních verzích podstatně přepracována. Za

prvé se v jejich nastaveních objevila možnost výběru klíčových frází pro různé jazyky, čeština ale chybí (General Advanced Options - Select Language Files), a za druhé byla opravena nepříjemná chyba spojená s chybějícím korektním ukončením (bez stlačení Ctrl+Alt+Del) programu po skončení práce. Nyní, potom, co se zbavil hříšků mládí, nabízí CyberBuddy výkonný univerzální program, který plní množství užitečných funkcí pro široký okruh uživatelů: funkci mluvčího poštáka (E-mail), organizéru (Remind), internet pageru (Buddies - Send), zprostředkovatele příspěvků (News), také hlídače změn na webových stránkách (WWW) a pohybu kurzů na burzách (Stocks). Vše je ozvučeno, ozdobeno animovanými postavičkami, a navíc zcela zadarmo. Zcela jistě má smysl to vyzkoušet, pouze hlasový engine a knihovny se musí stáhnout odděleně, neboť nejsou součástí distribuční verze CyberBuddy. V tomto programu je možné vše nastavit podle vlastního vkusu. Na záložce General Character and Speech Options si nastavíte podle sebe různé parametry pomocníků a hlasových enginů. Na záložce General Advanced Options vyberte skin pro základní okno programu a napište, pokud chcete, vlastní fráze pro animované postavičky (Custom Message). Na záložce E-mail budete muset zadat adresu své poštovní schránky a také uživatelské jméno a heslo pro přístup k e-mailu. Nyní, po příkazu Get E-mail Now, budete dostávat poštu a po vyznačení příslušného dopisu uslyšíte jeho obsah (Read E-mail). Pokud program instalujete z CD, je třeba jako první spustit soubor "tv_enua.exe" pro instalaci hlasového enginu. Jinak se program bude tento soubor snažit nahrát z internetu.

CyberBuddy 2.12

Výrobce John M. DeFino

Internet www.thecyberbuddy.com

Velikost distribučního souboru 1,9 MB

Podmínky šíření freeware

Operační systém Windows 9x/ME/2000/XP

- + mnohofunkčnost
- + vestavěný poštovní klient a organizér
- + pružný systém nastavení

Speakonia 1.03

Jednoduchým a dostupným způsobem, jak ozvučit text v počítači, je použít program Speakonia. Jediné, co budete muset udělat, je nainstalovat vybraný hlasový engine (Lermout & Hauspie, Digalo nebo Sakrament). Co se týče knihoven MS Speech API a podpory anglického jazyka (American English), ty jsou již výrobcem přiloženy v instalačním souboru. Rozhraní Speakonia, to je vzor minimalismu. Nic zbytečného, pouze základní hlasové funkce. Ovládání je úplně stejné jako v multimediálních přehrávačích: tlačítko spuštění čtení, pauzy a úplného zastavení. Parametry přednášení se v zásadě dají nastavovat (Speech - Properties). Můžete si vybrat hlasový engine, rychlost, zbarvení a zvučnost hlasu - nutno přiznat, že pouze pro engine Lernout & Hauspie. Ale v každém případě hodnoty, které jsou nastaveny implicitně, vám určitě budou vyhovovat. Pokud se zamotáte v nastaveních, můžete klidně kliknout na tlačítko Reset a nespletete se. Tím spíše, že kvalita čtení nezávisí na vlastním programu Speakonia. V Editoru uživatelských slovníků, který se objeví po klepnutí myší na tlačítko Lexicon na záložce Opravy, je možné upřesnit nesprávně ozvučená slova nebo zkratky. Mimochodem, není to až tak nutné, protože Speakonia vcelku dobře ozvučuje text ze schránky, textové soubory, stránky na internetu (File - Read Webpage) a také aktuální čas a datum. Po registraci na internetových stránkách programu (Tools - Options Registration) se objeví možnost uložení strojového čtení do WAV souboru (File - Save as Wav File).

Speakonia 1.03

Výrobce CFS Technologies

Internet: www.cfs-technologies.com

Velikost distribučního souboru: 2,6 MB

Podmínky šíření: freeware

Operační systém: Windows 9x/ME/2000/XP

- + jednoduché rozhraní
- + možnost uložení čteného textu do WAV souboru
- nastavení je možné pouze pro engine Lernout & Hauspie

GUI Reader 2.0

Ani ruští vývojáři nezapomínají na hlasové programové vybavení. Příkladem může být GUI Reader Andreje Jefimova, který podporuje všechny potřebné funkce pro programy obdobného druhu, ale přitom zabírá pětikrát (!) méně operační paměti než Web Talkster a CyberBuddy: 800 KB ve srovnání se 4 MB - to je zřejmá přednost. V klidovém stavu je GUI Reader zastoupen ikonkou v pravé části hlavního panelu a nijak nepřekáží uživateli při práci. Pro jeho aktivaci je nutné klepnout pravým tlačítkem myši na tuto ikonku a vybrat v rozbalovacím menu to, co je zapotřebí. Je možné požádat GUI Reader, aby ozvučil textový soubor (Read File) nebo obsah schránky (Read Clipboard). Program bohužel neobsahuje prakticky žádná nastavení. K dispozici jsou pouze standardní systémová menu vlastností a doplňkových parametrů. Pouze tak je zde možné nastavit pomocníka a jeho hlas. Zato je v GUI Readeru povoleno vybírat jazyk ozvučeného textu, hlavní však je, aby ve vašem systému byl nainstalován příslušný hlasový engine s podporou vámi požadovaného jazyka. V instalačním balíku chybí pouze hlasové engine, které je nutné instalovat zvlášť.

GUI Reader 2.0

Výrobce CFS Technologies

Internet Andrej Jefimov

Internet www.mediapg.ru

Velikost distribučního souboru 3,7 MB

Podmínky šíření freeware

Podporované operační systémy Windows 98/NT/2000/XP

- + nenáročnost na zdroje
- + možnost výběru jazyka textu
- má pouze všeobecná systémová nastavení

Voice Navigator 1.01

Voice Navigator - to je jediný komerční program v tomto přehledu a jediná utilita, která je určena pro hlasové řízení počítače. Jako zajímavost lze uvést, že Navigator vytvořilo petrohradské Centrum řečových technologií, to samé, které se zabývalo rozšířením nahrávek ponorky Kursk, která se před dvěma lety potopila v Barentsově moři. Oblíbeným dítkem tohoto centra je Truffaldino, komplex hlasového řízení domácí a kancelářské techniky. Vývojáři slibují, že v něm bude obsažena úplná podpora televizí, telefonů, a dokonce mikrovlnných trub. Je pravda, že jak to vypadá v praxi, je možné se pouze dohadovat - demoverze Truffaldina nic takového nedemonstruje. Ale demoverze Voice Navigatoru se ukázala z té nejlepší strany. Sice ne ihned, ale až po menším tréninku. Pro procvičování je nutné vyznačit myší jeden z pěti bodů, které jsou přístupné v demoverzi, klepnout myší na tlačítko Improve Recognition a zřetelně vyslovit příslušný povel. Může to být libovolné slovo, a dokonce i zvuk. Například já si spouštím Kalkulačku na zapísknutí. Ihned po učení vyzkoušejte vyslovit do mikrofonu příkaz pro spuštění nějakého programu, např. Poznámkového bloku, a ověřte reakci. Když skončíte práci, vyslovte podmíněný příkaz pro Zavřít a aktivní okno se zavře. Příkazy se provádí poměrně rychle a přesně, ale čas od času je zapotřebí je opakovaně učit. Kromě všeho ostatního je pro příkazy uživatele možné zadat určitou hlasovou odpověď v tzv. Sound File Manageru. Pokud vám nevyhovuje implicitní varianta, klepněte na Replace a uložte nebo nadiktujte svou vlastní.

Voice Navigator 1.01

Výrobce Centrum řečových technologií, Petrohrad, Rusko

Internet www.stc.rus.net

Velikost distribučního souboru 1,2 MB

Podmínky šíření komerční demo, 20 USD

Operační systém Windows 9x/ME/2000/XP

- + vysoká přesnost rozeznávání
- + možno zadat hlasovou odpověď počítače
- požaduje opakované učení příkazů