

ARET – AUTOMATICKÉ ČTENÍ UČEBNÍCH TEXTŮ PRO ZRAKOVĚ POSTIŽENÉ STUDENTY

Jindřich MATOUŠEK, Zdeněk HANZLÍČEK, Zdeněk KRŇOUL, Michal CAMPR

Západočeská univerzita v Plzni, Fakulta aplikovaných věd, Katedra kybernetiky

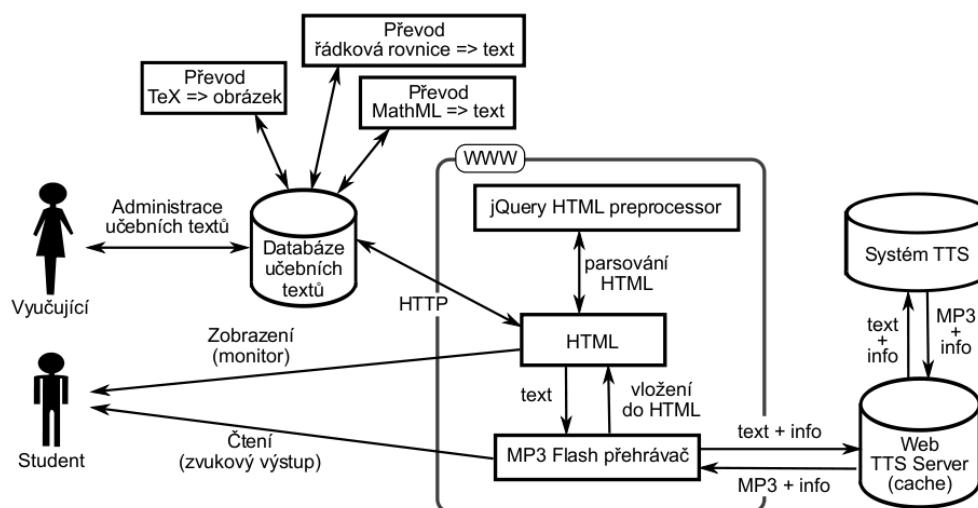
jmatouse@kky.zcu.cz, zhanzlic@kky.zcu.cz, zdkrnoul@kky.zcu.cz, mcamp@students.zcu.cz

Anotace: Příspěvek popisuje současný stav projektu „Automatické čtení učebních textů pro zrakově postižené studenty“. V rámci projektu jsou vytvářeny automaticky čtené učební pomůcky s cílem inovovat a rozvíjet výuku a rozšířit možnosti vlastní domácí přípravy zrakově postižených dětí. K automatickému čtení se využívá technologie syntézy řeči z textu (text-to-speech, TTS).

Úvod

Tento příspěvek popisuje současný stav projektu „Automatické čtení učebních textů pro zrakově postižené studenty“ (Automatic Reading of Educational Texts, ARET, <http://aret.zcu.cz>). Cílem projektu je inovace a rozvoj výuky, především pak rozšíření možností vlastní domácí přípravy zrakově postižených dětí a žáků. Projekt je řešen na katedře kybernetiky, Fakultě aplikovaných věd Západočeské univerzity v Plzni (ZČU) ve spolupráci se Základní školou a Mateřskou školou pro zrakově postižené v Plzni. V rámci projektu je vytvářen systém pro automatické čtení učebních pomůcek. Pedagogičtí pracovníci ze ZŠ systém využívají k přípravě a správě učebních textů. Učební texty jsou žákům přístupné pomocí webového rozhraní systému, v němž jsou zobrazovány a předčítány pomocí počítačové syntézy řeči. Je přitom možné využít „vestavěný“ hlas a systém syntézy řeči z textu (text-to-speech, TTS) vyvíjený na ZČU nebo standardní čtečku, screen reader, nainstalovanou na počítači uživatele. Přístup do systému je vysoce konfigurovatelný – je možné jednoduše nastavit různá práva pro editory (pedagogické pracovníky), uživatele (žáky) i ostatní (návštěvníky stránek s učebními texty). S ohledem na cíle projektu a potřeby hlavního partnera (ZŠ) byly vybrány učební texty z matematiky a fyziky na úrovni 2. stupně ZŠ. Systém automatického čtení se tak musí vypořádat s přítomností matematických a fyzikálních vzorečků, a to jak ve fázi vytváření učebních textů, tak i ve fázi jejich automatického čtení.

Čtením technických dokumentů nebo matematických vzorců se zabývaly i jiné projekty. Problematicke automatického čtení matematických vzorců se věnovaly například systémy AsTeR (Audio System for Technical Readings [1]) nebo AudioMath (vyvíjený na univerzitě v Portu [2]). Pro český jazyk byl na Masarykově univerzitě v Brně vyvinut Lambda editor [3], který kromě hlasového výstupu podporuje také Braillovo písmo. V rámci projektu ARET je vyvíjen nový systém pro čtení matematických vzorců.



Obr. 1: Blokové schéma systému automatického čtení učebních textů.

Popis systému

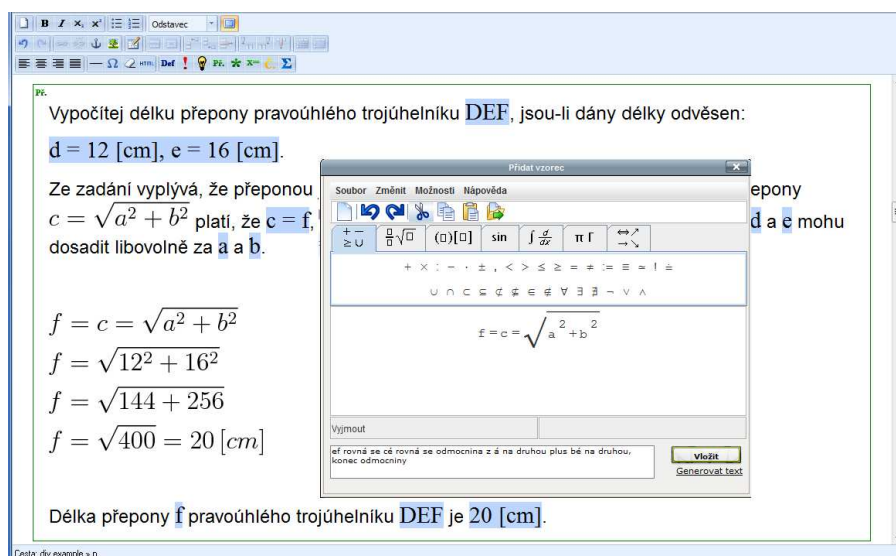
Vyvíjený systém je založen na architektuře klient-server a využívá moderní technologie a programovací jazyky (PHP, JavaScript, JQuery, Java, Python) používané při vývoji webovských aplikací. Výsledný systém běží na open-source web serveru *Apache* s databázovým systémem *MySQL*. Jádro systému je založeno na *Symfony*, open-source PHP „knihovně“ pro vytváření webovských aplikací. Ke správě databáze se využívá *Doctrine*, pomocí níž je možné zadávat dotazy do databáze v objektové orientovaném SQL dialektu DQL (Doctrine Query Language).

Systém je vyvíjen jako webová aplikace a je tedy logicky rozdělen na dvě části: veřejnou a administrační. Veřejná část slouží k prohlížení a ozvučení (pomocí počítačem generovaného hlasu) jednotlivých učebních textů (uspořádaných jako „témata“). Administrační část na druhou stranu umožňuje správu učebních textů, v níž jednotlivá témata mohou být vytvářena a upravována. Blokové schéma celého systému je znázorněno na obrázku 1.

Administrační část systému

Na obrázku 2 je zobrazena ukázka administrační části systému. V pozadí je textový editor, pomocí něhož mohou editoři (pedagogičtí pracovníci) zadávat vlastní náplň témat – učební texty. Editor byl pro potřeby projektu obohacen o možnost vkládat šablony a vzorce. Šablony se využívají k jednoznačnému zápisu určitých částí dokumentu, např. šablona *Pozor!* se využívá ke zvýraznění důležité informace, které by žáci měli věnovat zvýšenou pozornost, nebo šablona *Příklad* se využívá k zápisu různých příkladů probírané látky. Současná verze rozlišuje pět šablon: *Definici*, *Pozor!*, *Poznámku*, *Příklad* a *Řešení*.

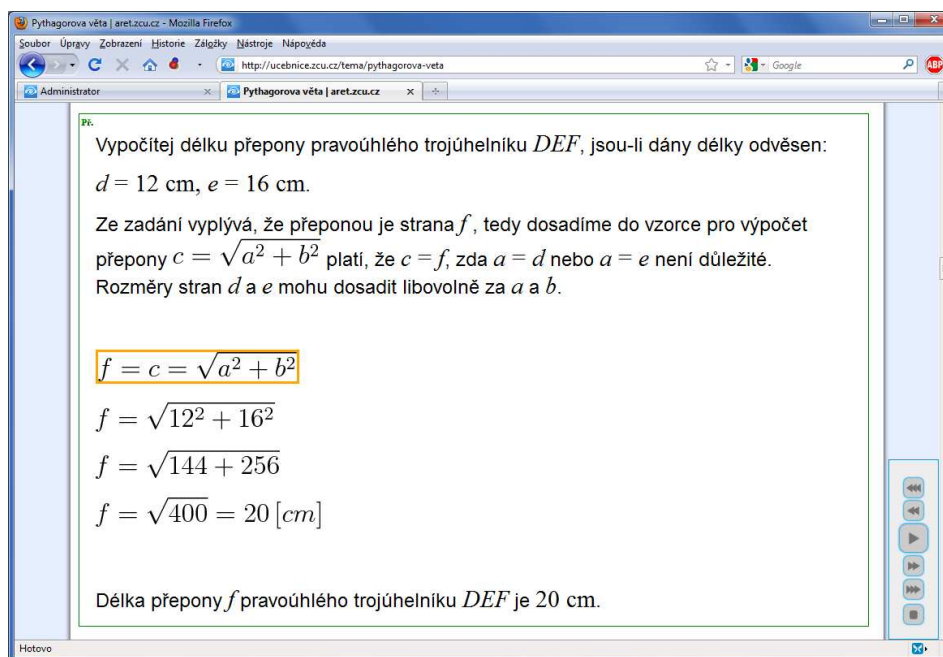
Matematické a fyzikální vzorce je možné do vytvářených textů vkládat dvěma různými způsoby. Jednoduché vzorce je možné zadávat jako *řádkové rovnice*, které mohou být ve speciálním tvaru zapsány v jednom řádku textu (např. $x+1$, na obr. 2 jim odpovídají vzorce-texty s namodralým pozadím). Ke vkládání komplexnějších vzorců se využívá editor *DragMath* (viz dialogové okno na popředí obrázku 2). Oba typy reprezentace vzorců jsou poté zpracovány, aby bylo vzorce možné zobrazit (vzorce jsou přitom zobrazeny jako obrázek) a ozvučit prostředky počítačové syntézy řeči při prohlížení učebních textů ve veřejné části systému (pro tyto účely jsou vzorce reprezentovány ve formátu MathML a poté pomocí speciálních pravidel převedeny na odpovídající slovní textový zápis).



Veřejná část systému

Veřejnou část systému tvoří webové rozhraní, pomocí něhož se jednotlivá témata zobrazují a jsou přihlášeným žákům automaticky předčítána. Témata pokrývají jednotlivé oblasti matematiky a fyziky 2. stupně ZŠ a jsou neustále doplňována a rozšiřována. Před vlastním zobrazením a čtením dochází k automatickému zpracování stránek, HTML dokumentů. Dokumenty jsou jednak upraveny a optimalizovány pro prohlížení a dále jsou z nich vyčleněny vlastní obsahové texty (bez formátovacích značek), včetně přepsaných řádkových rovnic a textově reprezentovaných vzorců z DragMathu.

V případě vestavěného systému TTS jsou „vyčištěné“ texty postupně po jednotlivých částech (odstavcích a textově reprezentovaných vzorcích) posílány na *Web TTS server*, který zajišťuje jejich automatické čtení. Vlastní převod textu na řeč provádí systém TTS, který pro daný vstupní text vygeneruje odpovídající zvukový soubor ve formátu MP3. Podporována je přitom i konfigurovatelná vnitřní paměť (*cache*) systému, která pro jednotlivé vstupní texty zvukové soubory uchovává a zabraňuje tak jejich opětovnému generování systémem TTS. K přehrávání zvukových souborů se využívá open-source flashový přehrávač *JPlayer* (viz obr. 3 vpravo dole). Přehrávač umožňuje jednoduchou navigaci v učebních textech a přehrávaných zvukových souborech – je možné se pohybovat po jednotlivých podkapitolách témat a po jednotlivých odstavcích (včetně vzorců). Konkrétní text, který je právě předčítán, je navíc zvýrazněn (viz obr. 3). V případě využití čtečky je čtení textů plně v režii konkrétní čtečky nainstalované na počítači uživatele.



Obr. 3: Ukázka veřejné části systému.

Využití TTS a problémy specifické pro projekt ARET

K automatickému čtení učebních textů se využívá technologie syntézy řeči z textu (TTS – bližší informace lze nalézt například v [4]). Úkolem systému TTS je převést libovolný vstupní text na odpovídající řeč, tj. „přečíst“ daný text. V projektu ARET se pro čtení textů primárně využívá český TTS systém ARTIC (Artificial Talker in Czech) [5] vyvíjený na katedře kybernetiky, Fakultě aplikovaných věd Západočeské univerzity v Plzni ve spolupráci s firmou SpeechTech.

Zjednodušeně se celý proces vytváření řeči z textu dá popsat následujícími kroky:

- předzpracování textu (přepis zkratk, číslic, akronymů apod. do plné slovní formy);
- fonetický přepis textu (včetně přepisu slov cizího původu);
- popis prozodických vlastností textu;
- vytváření řeči s požadovanými fonetickými a prozodickými vlastnostmi na signálové úrovni [6].

V rámci projektu ARET je navíc nutné věnovat zvláštní pozornost matematickým a fyzikálním vzorcům. Díky speciálnímu značení vzorců ve vytvářených učebních textech (ve formě řádkových rovnic nebo MathML) není nutné vzorce v textu detekovat. Pro převod takto zapsaných vzorců do odpovídajícího textového zápisu byla navržena speciální pravidla, která respektují typ vzorce (unární nebo binární operace, typ operace jako sčítání nebo násobení, typ operandů jako číslo nebo proměnná apod.) a zajišťují správné skloňování jednotlivých operandů. Současná verze systému obsahuje sadu pravidel přizpůsobených tématům řešeným v projektu ARET, tj. matematice a fyzice 2. stupně ZŠ. Sadu pravidel je však podle dané problémové oblasti možné jednoduše rozšířit o nová pravidla s novými matematickými operacemi. Proces převodu vzorců na text je symbolicky znázorněn na obrázku 1 v blocích „Převod řádková rovnice => text“ a „Převod MathML => text“.

Závěr

V tomto příspěvku byl prezentován současný stav projektu „Automatické čtení učebních textů pro zrakově postižené studenty“ (Automatic Reading of Educational Texts, ARET, <http://aret.zcu.cz>). Vzhledem k tomu, že je projekt řešen ve spolupráci se Základní školou a Mateřskou školou pro zrakově postižené v Plzni, soustředí se tematicky na předměty matematika a fyzika pro 2. stupeň ZŠ. Vyvíjený systém byl nicméně navržen obecně, aby jej bylo možné rozšířit i na jiné předměty a tematické oblasti. Přestože projekt ARET stále probíhá, první učební texty již jsou dostupné na <http://ucebnice.zcu.cz>. Některá z témat jsou dostupná všem návštěvníkům, jiná témata jsou po přihlášení přístupná žákům partnerské ZŠ.

Další práce na projektu se budou soustředit zejména na následující oblasti:

- **Témata.** Počet vytvořených učebních textů bude postupně narůstat s cílem pokrýt řešené tematické oblasti.
- **Funkcionalita systému.** Plánuje se zvyšovat funkcionalitu vyvíjeného systému (veřejnou i administrační část). Jako příklad uveďme opravu chyb způsobených systémem TTS, přidávání dalších pravidel pro čtení nových typů vzorců atd. Zvažujeme také možnosti personalizace systému pro každého uživatele, např. možnost změny designu webových stránek (např. změna barvy písma, šablon apod.) nebo možnost změny hlasu pro automatické čtení.
- **Kompatibilita s jinými nástroji pro zrakově postižené.** Naší snahou také bude, ve spolupráci s partnerskou ZŠ, zvýšit kompatibilitu vyvíjeného systému s jinými nástroji a systémy pro zrakově postižené. Jako příklad uveďme optimalizaci klávesových zkratk. Řešit hodláme i problémy s nežádoucím konkurenčním čtením učebních textů vestavěným TTS systémem a čtečkou, kterou je žák zvyklý používat pro čtení jiných druhů informace.

Literatura

- [1] Raman, T. V.: Audio System for Technical Readings. Disertační práce. Cornell University, New York, NY, 1994.
- [2] Ferreira, H., Freitas, D.: Enhancing the Accessibility of Mathematics for Blind People: The AudioMath Project. Lecture Notes in Computer Science, vol. 3118, 2004, pp. 678-685.
- [3] Návrh české osmibodové normy, matematický editor Lambda (informační portál). <http://www.teiresias.muni.cz/czbraille8>. (cit. 2011-02-23).
- [4] Psutka, J., Müller, L., Matoušek, J., Radová, V.: Mluvíme s počítačem česky. Academia, Praha, 2006.
- [5] Matoušek, J., Tihelka, D., Romportl, J.: Current State of Czech Text-to-Speech System ARTIC. Lecture Notes in Computer Science, vol. LNAI 4188, 2006, pp. 439-446.
- [6] Tihelka, D., Kala, J., Matoušek, J.: Enhancements of Viterbi Search for Fast Unit Selection Synthesis. Proceedings of Interspeech 2010, pp. 261-264, September 26-30, Makuhari, Japonsko.

Poděkování



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



ZÁPADOČESKÁ
UNIVERZITA
V PLZNI

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Projekt „Automatické čtení učebních textů pro zrakově postižené studenty“, CZ.1.07/1.2.00/08.0021, je spolufinancován z Evropského sociálního fondu a státního rozpočtu České republiky.