

# Kapitola 2

## Publikace materiálů na WWW

V první části této kapitoly se naučíte nezbytné minimum nutné k tomu, abyste mohli vyvážet vlastní WWW stránky. Druhá nepovinná část je věnována současným možnostem publikování matematického textu na Webu a výhledu do budoucnosti.

### 2.1 Udělejte si vlastní WWW stránku

Vlastní tvorba dokumentů pro WWW se provádí v jazyce HTML (Hypertext Markup Language). HTML je jedním z jazyků, definovaných pomocí ISO standardu SGML (Standard Generalized Markup Language). Jedná se o jazyk, určený k definici strukturálních jazyků (metajazyků). Umožňuje popsat pravidla, kterými se musí řídit text v daném jazyce. Definice HTML je tedy tvořena sadou SGML pravidel, určujících co se smí a co nikoli - čili syntax jazyka. Tato sada pravidel se nazývá Document Type Definition (DTD).

Jazyk HTML patří do kategorie „markup language“. Základní myšlenkou všech jazyků kategorie „markup language“ je, že vycházejí z čistého textového souboru. Ten obsahuje jak vlastní text, který chcete sdělit čtenáři, tak příkazy ovlivňující jeho vzhled a strukturu. Jsou zapisovány prostřednictvím určitých speciálních znakových řetězců, kterým byl v daném jazyce přiřazen jistý význam. Markup language jazyky nedbají na tvar zdrojového textu. Konec řádku v nich má zpravidla stejný význam jako mezera.

Zmíněné vlastnosti má i HTML. Příkazům se zde říká značky (tagy) a zapisují se mezi znaky ”menší než” a ”větší než” - například `<EM>` zahájí zvýrazňovaný text. Ve jménech značek se nerozlišují malá písmena od velkých. Pokud se rozhodnete vytvářet své stránky ručně, doporučujeme stanovit si jednotný způsob zápisu a ten dodržovat.

Řada značek je párových. To znamená, že se vyskytují ve dvojicích - jedna vždy zahajuje a druhá ukončuje danou konstrukci. Jejich účinek se vztahuje na text, který je uzavřen mezi zahajující a ukončující značkou. Obě značky v páru mají shodné jméno, v ukončující značce mu však předchází znak ”lomítko”. Například

```
raz <EM>dva tři</EM> čtyři
```

signalizuje, že prostřední dvě slova mají být zvýrazněna (zpravidla pro ně bude použita kurzíva).

A teď přijde velmi důležitá informace: HTML soubor interpretuje klient WWW. Tento program musí vykonávat dvě základní funkce. Musí se spojit se serverem a získat od něj zdrojový text stránky (tato úloha klienta je celkem obvyklá a setkáte se s ní i v jiných službách). Když se tohle podaří, musí získaný text analyzovat, zformátovat a předložit výslednou podobu uživateli. Praktický důsledek: vzhled stránky závisí na klientovi. Stejná stránka může vypadat velmi odlišně, bude-li prohlížena dvěma různými klienty. Naší snahou bude ale pokud možno vytvářet WWW stránky tak, aby vzhled byl pokud možno na klientovy nezávislý.

### 2.1.1 Co budete potřebovat, než začnete tvořit vlastní stránky?

HTML dokumenty jsou obvyčejné textové soubory. K jejich editaci proto postačí libovolný editor, který umí pracovat s čistým ASCII textem – např. edit z MS-DOSu nebo notepad z Windows. Většina prohlížečů umožňuje zobrazit stránku i v její zdrojové podobě zapsané pomocí jazyka HTML (obr. 2.1).



Obrázek 2.1: Zápis stránky v HTML zobrazený pomocí programu Netscape

U každé stránky se tedy můžete podívat, jakým způsobem byla vytvořena a získat tím inspiraci pro vlastní tvorbu.

Druhou nezbytnou podmínkou pro tvorbu vlastních WWW stránek je znalost jazyka HTML, s jehož základy se seznámíte v této kapitole. WWW stránky je možno vytvářet také pomocí integrovaných editorů jako je např. FrontPage, HotMetal, Netscape Composer atd. nebo exportem z Wordu a Excelu. Domníváme se, že efektivnější je nejdříve zvládnout základy jazyka HTML a teprve poté experimentovat s podpůrnými nástroji.

Pokud si budete chtít výsledky své práce zkontrolovat, je nezbytné opatřit si nějaký prohlížeč. V současné době mezi dva nejpoužívanější patří Netscape Navigátor a Microsoft Internet Explorer a naleznete je na CD, které jste obdrželi. Na CD najdete i experimentální prohlížeč W3C konsorcia Amaya.

Při samotné tvorbě dokumentů nepotřebujete být připojeni k síti Internet. Většina prohlížečů umožňuje práci v tzv. off-line režimu, kdy jsou WWW stránky čteny přímo z lokálního disku počítače a ne přes síť pomocí protokolu HTTP.

Zpřístupnění stránky okolnímu světu se liší podle přístupu k Internetu (přímé spojení nebo přístup pomocí modemu a telefonní sítě — komutované spojení). Konkrétní informace je vždy nutné zjistit od správce sítě či poskytovatele připojení.

## 2.1.2 Struktura stránky

WWW stránka má předepsanou pevnou formu, kterou by měla dodržovat. Vypadá následovně:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<HTML>
<HEAD>
<TITLE>Nadpis stránky </TITLE>
<META HTTP-EQUIV="content-type" CONTENT="text/html;
      charset=ISO-8859-2">
</HEAD>
<BODY>
tělo stránky
</BODY>
</HTML>
```

První řádek obsahující <!DOCTYPE . . . > (prolog) je zde kvůli kompatibilitě s SGML a identifikuje verzi HTML – v našem případě verzi 4.01. Dále je celá stránka uzavřena do párové značky <HTML>. Uvnitř se pak dělí na dvě části. Záhlaví (obklopené značkou <HEAD>) obsahuje především informace popisující některé vlastnosti stránky - jak se jmenuje, kde je umístěna, informace o kódování češtiny a podobně. Značka <TITLE> zde určuje titulek (čili nadpis okna) dokumentu. Párová značka <BODY> pak obklopuje tělo stránky. To bude hlavním cílem vašeho snažení, protože tvoří viditelný obsah, který bude zobrazen uživateli.

## Odstavce

HTML ignoruje tvar zdrojového textu stránky. Konce řádků, konce odstavců - to vše mu musí být řečeno vhodnými značkami. Ty nejzákladnější uvádí tabulka [2.1](#)

Každý odstavec je třeba zahájit značkou <P>. Lze ji používat také jako párovou, kdy na začátku odstavce stojí <P> a na jeho konci </P>.

Chcete-li ukončit stávající řádek a další text umístit na začátek následujícího, poslouží vám značka <BR>. U většiny klientů se její chování podobá <P>. Na rozdíl od něj však nevynechává svislou mezeru, která bývá vkládána před nový odstavec.

<P>	nový odstavec
 	nový řádek
<HR>	vodorovná čára

Tabulka 2.1: Členění textu

Vraťte se ještě ke značce <P>, na které si předvedete další významný prvek HTML. Je jím atribut, který umožňuje ovlivnit chování značky. Zapisuje se za její jméno, oddělen mezerou. Zpravidla ve tvaru

atribut="hodnota"

Je-li hodnotou číslo nebo jeden z předem pevně daných řetězců, lze uvozovky vynechat. Pro každou značku je definováno, jaké má atributy, jaké jsou jejich významy a implicitní hodnoty. Pro značku <P> je zajímavý atribut ALIGN, určující způsob zarovnání odstavce. Přípustné hodnoty jsou LEFT (na prapor vlevo, implicitní hodnota), RIGHT (na prapor vpravo) a CENTER (centrováno na střed).

### 2.1.3 Písmo

Možnost měnit vzhled písma je u moderních systémů pro práci s textem neodmyslitelná. Jak je na tom WWW? Kolik a jaká písma umí? Odpověď zní: těžko říct. Pracuje-li WWW klient v textovém režimu, bude nabídka písem chudičká. Ačkoli jazyk HTML nabízí řadu příkazů pro změny použitého písma, výsledek je bez záruky. Stojí a padá se schopnostmi klienta.

Právě z tohoto důvodu se doporučuje dávat při práci s písmem přednost tak zvanému logickému vyznačování (viz tabulka 2.2). Jeho prostřednictvím říkáte: tento text má být zvýrazněn, tento text je ukázkou klávesnicového vstupu. Nechávejte přitom na klientovi, aby posoudil, jak nejlépe dosáhnout kýženého efektu.

<EM>	zvýraznění
<STRONG>	silné zvýraznění
<CITE>	citace
<SAMP>	ukázka zdrojového kódu
<KBD>	vstup z klávesnice
<VAR>	jméno proměnné

Tabulka 2.2: Logické značení

Všechny značky jsou párové a ovlivňují text, který je uvnitř páru. Kromě nich existuje ještě druhá skupina značek, ovlivňujících písmo. Umožňují fyzické vyznačení, při kterém výslovně požadujete „pro tento text použij kurzívu“. Souhrn značek pro fyzické vyznačení najdete v tabulce 2.3.

<B>	tučné
<I>	kurzíva
<TT>	neproporcionální
<U>	podtržené, nepřliř podporováno
<BIG>	velké
<SMALL>	malé
<SUB>	index
<SUP>	exponent

Tabulka 2.3: Fyzické značení

## 2.1.4 Formátování

Po poměrně podrobném úvodu, který byl nezbytně nutný pro pochopení principu jazyka HTML, uvádíme další problematiku již jen zkráceně. Kompletní kurzy jazyka HTML (které vyžadují mnohem více času, než máte k dispozici) jsou k dispozici např. na

<http://www.kosek.cz/clanky/html/index.html>

Vytváříme si domovskou stránku od Jirky Koska a na

<http://www.kit.vslib.cz/~satrapa/docs/wwwprik1/index.html>

HTML v příkladech od Pavla Satrapy. Z těchto kurzů jsme také čerpali při přípravě tohoto materiálu.

Jelikož byl jazyk HTML původně určen k vyjádření struktury textu, odehrávají se formátovací kouzla převážně v logické rovině. Část textu například označíme jako nadpis, část jako seznam a podobně.

### Nadpisy

Pro vytvoření nadpisu poskytuje HTML značku <Hn>. Na místě n stojí číslice od jedné do šesti, která vyjadřuje významnost (a tudíž zpravidla i velikost) nadpisu. <H1> obaluje nadpis celé stránky - tedy největší. Měl by být právě jeden. Jednotlivé části stránky bývají zahájeny nadpisem, uzavřeným do značky <H2>, jejich části používají pro své nadpisy <H3> a tak dále. Všechny značky jsou párové. Za nadpis je považován text uzavřený mezi oběma členy páru.

### Číslovaný seznam a seznam s odrážkami

HTML nabízí celkem tři druhy seznamů, které lze použít pro zápis nejrůznějších výčtů, pracovních postupů a řadu dalších strukturovaných textů. Velmi jednoduše vytvoříte seznam číslovaný a seznam s odrážkami. Celý seznam je vždy uzavřen do patřičné párové značky. Každá jeho položka je zahájena značkou <LI>:

```
<OL>
<LI> text první položky
<LI> text druhé položky
</OL>
```

<OL>	číslovaný seznam
<UL>	seznam s odrážkami
<LI>	položka v <OL> či <UL>
<DL>	seznam definic
<DT>	pojem v <DL>
<DD>	vysvětlení v <DL>

Tabulka 2.4: Značky pro seznamy

## Předformátovaný text

Pomocí značky <PRE> lze potlačit obvyklé formátovací mechanismy HTML. Text, který je uzavřen mezi <PRE> a </PRE>, má být klientem zobrazen přesně v té podobě, ve které se nachází. To znamená, že se budou dodržovat mezery i konce řádků a bude použito písmo s pevnou šířkou znaku. Značka <PRE> nachází uplatnění zejména při ukázkách programů či konfiguračních souborů.

## 2.1.5 Obrázky

### Formáty obrázků

V současném Webu jsou podporovány dva základní formáty grafických souborů: GIF a JPEG. První z nich je považován za vhodnější zejména pro kresby s ostrými hranami mezi jednotlivými barevnými plochami. Formát JPEG je doporučován především pro obrázky charakteru fotografií, oplývající plynulými přechody barev, světel a stínů. Tento formát nezachovává kompletní obrazovou informaci. Při ukládání do JPEG souboru dává většina programů možnost zvolit, jak vysokou kvalitu výstupu požadujete. Vyšší hodnota znamená věrnější obrázek, ale zároveň větší objem dat. Je třeba najít vhodný kompromis.

### Vložení obrázku

Vložení obrázku provedete pomocí značky <IMG>. Jedná se o nepárovou značku, jejíž chování určují hodnoty atributů (viz tabulka 2.5). Nejdůležitějším je atribut SRC. Udává, kde má klient

SRC=" . . . "	URL vkládaného obrázku
ALT=" . . . "	alternativa pro textové klienty
ALIGN=" . . . "	zarovnání vůči okolnímu textu
WIDTH= . . .	šířka
HEIGHT= . . .	výška
ISMAMP	obrázek s klikou

Tabulka 2.5: Atributy značky <IMG>

vzít dotyčný obrázek. Pro obrázky, uložené ve stejném adresáři jako stránka, stačí uvést jméno jejich souboru. Druhým povinným atributem je ALT (od verze HTML 4.01), který nahrazuje obrázek negrafickým klientům.

### Několik rad závěrem

- Pro obrázky charakteru kresby používejte formát GIF, pro fotografie formát JPEG. GIF umožňuje i animace a průhledné pozadí.
- Opatrně s používáním WIDTH a HEIGHT. Lepší je upravit velikost obrázku vhodným grafický programem.
- Ikony a symboly by neměly přesáhnout velikost několika málo KB. Ostatní obrázky by neměly přesahovat velikost 100 KB. V opačném případě může čekání na načtení vaší stránky odradit potenciální zájemce.
- Pokud používáte obrázek jako odkaz lze atributem BORDER=0 odstranit rámeček.

### 2.1.6 Odkazy

Síla Webu spočívá ve vzájemném propojování dokumentů. Schopnost vkládat do stránek odkazy na stránky jiné odlišuje WWW od běžných systémů pro formátování textu a činí z něj službu počítačové sítě.

Pro vložení odkazu do WWW stránky slouží značka <A> (Anchor - kotva). Přináší dvě informace - kam vede a jak má na stránce vypadat. Cíl odkazu určuje atribut HREF značky. O její vzhled se stará text (a případné značky), uzavřený mezi <A> a párové </A>. Dostane-li klient zdrojový text

```
Server  
<A HREF="http://www.math.muni.cz/">Sekce matematiky</A>  
Přírodovědecké fakulty Masarykovy Univerzity ...
```

zobrazí jej asi takto (obrázek 2.2).

Server [Sekce matematiky](http://www.math.muni.cz/) Přírodovědecké fakulty Masarykovy Univerzity ...

Obrázek 2.2: Odkaz na jiný dokument

Slova „Sekce matematiky“ budou aktivní a pokud je zvolíte, klient se přesune na stránku s URL <http://www.math.muni.cz/>. Klient obvykle odlišuje aktivní text barevně a navíc jej podtrhne.



## 2.1.7 Validace

Předtím než vystavíte svou stránku světu je užitečné vědět, zdali vytvářená stránka skutečně půjde načíst všemi zamýšlenými browsery anebo jenom, jestli ve zdrojovém textu někde nechybí zakončovací lomítko. K tomu slouží HTML validace. Stačí otevřít stránku

<http://validator.w3.org/>,

zadat do formuláře adresu stránky, požadovaný rozsah kontroly a počkat na výsledek. Aby mohla služba zkontrolovat vaši stránku, nesmíte zapomenout na prolog (viz strana 22).



Obrázek 2.3: Výsledek validace

### ■ Příklad 2.1

Prohlédněte si zdrojový kód stránky

[http://www.math.muni.cz/~plch/ukazka\\_validace.html](http://www.math.muni.cz/~plch/ukazka_validace.html).

Zkuste najít chyby. Poté pokusnou stránku zvalidujte s porovnejte výsledek analýzy s vaším předpokladem a s výsledkem na obrázku 2.3. Vidíte, že validátor odhalil všechny chyby. Oznamil