UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte STEVEN EDWARD ATKIN

Appeal 2008-4352
Application 09/891,341
Technology Center 2600

Decided: January 30, 2009

Before JOHN C. MARTIN, JOSEPH F. RUGGIERO, and
JOHN A. JEFFERY, *Administrative Patent Judges.*

MARTIN, *Administrative Patent Judge.*

DECISION ON APPEAL

STATEMENT OF THE CASE

This is an appeal under 35 U.S.C. § 134(a) from the Examiner's rejection of claims 1-15, all of the pending claims, under 35 U.S.C. § 103(a).

We have jurisdiction under 35 U.S.C. § 6(b).

We reverse the Examiner's § 103(a) rejection of claims 1-15 and enter (a) a new ground of rejection of claims 1-15 under § 112, second paragraph, (b) a new ground of rejection of claims 1-4, 9-13, and 15 under § 101, and (c) a new ground of rejection of claims 9-12 and 15 under § 112, first paragraph (enablement requirement).

A. Appellant's invention

Appellant's invention relates to human interfaces (input, display, output) concerning computer network and World Wide Web addresses in languages which require bidirectional display and presentation. Specification [0001].

By way of background, Appellant notes that the current technologies supporting the World Wide Web are "English-centric" due to the roots of the beginning of the Internet being an American and European effort. *Id*. at [0003]. As such, many of the conventions and standards employed in servers, routers, e-mail protocols, etc., employ an English alphabet with English-like syntax. *Id*. This ASCII-based domain name system (*id*. at [0007]) encodes only scripts written and displayed in a left-to-right order. *Id*. at [0010].

According to Appellant, there is a need for a system and a method that allow domain names to be handled and displayed with different (non-English) reading orders (*id*. at [0013]), such as Arabic, which is read right-to-left. *Id*. at [0057]. Furthermore, such a system and method should be

readily usable within the currently deployed technologies of the World Wide Web. *Id*. at [0013].

Appellant's invention preserves the characteristic of domain names that a full stop (i.e., a period) always serves to separate a domain name into its individual parts or "labels." *Id*. at [0054]. In addition, the invention preserves the strict hierarchy regarding the order of the labels, which is that the most general part of the domain name is always the rightmost label, while the most specific part of the name appears as the leftmost label. *Id*.

We note that the Brief,[1] in explaining how Appellant's invention works, appears to equate a URL, which is not recited in claim 1, with a domain name, which is. *See* Br. 7 ("[A] domain name or Universal Resource Locator ('URL') is defined by those in the industry as having a protocol identifier (e.g. http or https, etc.), a top-level identifier (e.g..com, .org, .net, etc.), a registered domain server name or second-level identifier, an optional third-level identifier (e.g. www, www2, etc.), zero or more subdomains, zero or more subdirectories, and zero or more resource names."). The Specification, on the other hand appears not to consider "http://" to be part of the domain name. *See* Specification at [0033] ("The determination as to whether a stream contains a domain name is rather straightforward if the domain name is preceded by some special identifier[,] [s]pecifically, "http://', 'ftp://', or 'telnet://'."). We note that considering

[1] References herein to the Brief are to the "Appeal Brief (First
(Continued on next page.)

"http://" to be separate from the domain name is also consistent with *Wikipedia*, which provides:

> The following example illustrates the difference between a URL (Uniform Resource Locator) and a domain name:
>
> URL: http://www.example.net/index.html
> Domain name: www.example.net
> Registered domain name: example.net

Wikipedia, http://en.wikipedia.org/wiki/Domain_name (last visited Jan. 28, 2009) (copy enclosed).

The Brief (at 5-6) describes several examples of URLs ("Uniform Resource Locators"[2]), which are referred to in the Specification (at 18:17) as "Universal Resource Locators." The first URL, "http://www.anycompany.com," is described as having three labels ("http://www," "anycompany," and "com"). The second URL, "http://www.help.ibm.com," is described as having four labels ("http://www," "help," "ibm," and "com").

Appellant explains that a likely starting point for choosing an allowable set of characters from which domain names may be constructed in other languages is the character repertoire available in the well-known Unicode/ISO10646 standard. Specification [0009]. The range of characters available in Unicode accommodates most modern written scripts, including Arabic and Hebrew scripts (*id*.), in which at least some characters are written

---

Reinstatement)" filed May 29, 2007.

[2] *Wikipedia*, http://en.wikipedia.org/wiki/Domain_name, *supra*.

right-to-left. *Id*. at [0077]. Thus, in contrast to ASCII, Unicode encodes scripts that are written right-to-left as well as those written left-to-right (*id*. at [0010]). Also, in Unicode it is perfectly "legal" to intermix these scripts. *Id*.

The Feinberg reference,[3] which does not concern domain names, explains that

> [s]ome spoken languages, such as Hebrew, are rendered *bi-directionally*. That is, certain portions of such languages are rendered left-to-right while other portions of those languages are rendered right-to-left. For example, according to the rules of the Hebrew language, text is rendered on a computer display or printout in a right-to-left configuration, but numeric formulas are rendered from left-to-right.

Feinberg, col. 1, ll. 22-29 (emphasis added).

Appellant's invention is described as "provid[ing] a one-to-one mapping between names in logical order and names in display order." Specification [0023]. In the Unicode system, the "logical order" is "[t]he order in which text is typed on a keyboard." *Glossary of Unicode Terms*, http://www.unicode.org/glossary (last visited Jan. 28, 2009), pages 1 and 23 (copy enclosed) of 41 printout pages. The "display order," on the other hand, is "[t]he order of glyphs presented in text rendering." *Id*. at page 13 (copy enclosed).

[3] Feinberg U.S. Patent 6,944,820 B2.

Unicode characters are either strong or indeterminate. The meaning of "strong" in the Unicode context is explained in *Wikipedia* as follows:

> In Unicode encoding, all non-punctuation characters are stored in writing order. This means that the writing direction of characters is stored within the characters. If this is the case, the character is called "strong". Punctuation characters however, can appear in both LTR [left-to-right] and RTL [right-to-left] languages. They are called "weak" characters because they do not contain any directional information.

Wikipedia, http://en.wikipedia.org/wiki/Bi-directional_text (last visited Jan. 28, 2009) (hypertext formatting omitted) (copy enclosed).

Appellant's invention "resolves the direction of indeterminate characters, such as the full stop or 'dot', hyphen-minus, Arabic numeral, and European numeral." Specification [0066]. Feinberg gives the following example of the problems presented by the indeterminate hyphen character in the Hebrew language:

> In the Hebrew language, certain neutral characters such as the hyphen character ("—") create special difficulties for displaying Hebrew text. As discussed above, according to the rules of the Hebrew language, text is rendered in a right-to-left configuration, while numeric formulas are rendered in a left-to-right configuration. For example, the text "I live in the house on the left" rendered according to the rules of the Hebrew language would be rendered in a right-to-left configuration. However, the numeric formula "3-2=1" would be rendered in a left-to-right configuration according to the rules of the Hebrew language.

Feinberg, col. 1, ll. 55-65.

Generally speaking, Appellant's invention involves breaking each domain name into a plurality of individual labels separated by full stop characters (i.e., periods) and independently evaluating each label for proper bidirectional display order. Specification [0022]. The resulting mapping of logical order to display order is described as providing unambiguous resolution of multilingual domain names. *Id*.

The Specification describes an example of a multi-language domain name consisting of an Arabic-language, right-to-left letter group "ABC" (which would be displayed as "CBA") and two English-language, left-to-right letter groups "ibm" and "com."[4] This example is depicted in Figure 3, which is reproduced below:[5]

---

[4] The Specification employs the following conventions to represent Arabic, Hebrew, and European letter and numerals:

> (a) uppercase letters "A" through "M" are used to represent Arabic letters,
>
> (b) uppercase letters "N" through "Z" are used to represent Hebrew letters,
>
> (c) digits "0" (zero) through "4" are used to indicate European numerals, and
>
> (d) digits "5" through "9" are used to indicate Arabic numerals. Specification [0034].

The lowercase letters (a-z) presumably represent English letters.

[5] In this example, the domain name is not preceded by a special identifier, such as "http://." *Id*. at [0033].

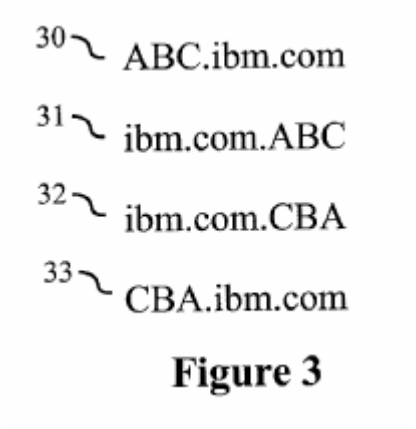


Figure 3

Figure 3 is described as illustrating issues involved with interpreting and displaying typical domain names using the well-known Unicode BiDi[6] algorithm, and the ambiguity which can result. *Id*. at [0006], [0017]. To simplify the following discussion, we will limit our discussion of Figure 3 to lines 30 and 33. Line 30 is described as "a 'normal' domain name in presented logical order" (*id*. at [0057]), which as explained above is the order (reading left to right) in which the characters would be typed on a keyboard.

The Specification explains that

> [t]he method of the invention provides a more desirable multilingual output (4) [*sic*; line 33] as illustrated in FIG. 3, wherein the "ABC" label is a right-to-left language component of the domain name, and the "ibm" and "com" labels are left-to-right components of the multilingual domain name. This output

[6] "**Bidirectional script support** is the capability of a computer system to correctly display bi-directional text. The term is often shortened to the jargon term **BiDi** or **bidi**." *Wikipedia*, http://en.wikipedia.org/wiki/Bi-directional_text, *supra* (hypertext formatting omitted).

> is consistent with the current structure of domain names. In this case the full stop characters are ignored, and the bidirectional algorithm is applied to each of the individual labels of the domain name.

Id. at [0061]. Thus, in the display order depicted in line 33, the Arabic characters ABC appear in their correct "CBA" display order while the remaining characters also appear in their correct display order.

As already noted, Appellant's invention also resolves the direction of indeterminate characters, such as the hyphen-minus character. Figure 4 is reproduced below.

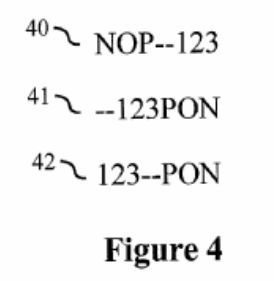


Figure 4

Figure 4 shows a logical order character string 40 that consists of Arabic letters NOP and European numerals 123 separated by a hyphen-minus character. Specification [0062]. Character string 41 is the incorrect display order obtained when using the Unicode BiDi algorithm, which inappropriately treats the hyphen-minus character as a European Terminator character. *Id*. Character string 42 is the correct display order obtained by treating the hyphen-minus character as a white space. *Id*. Also, the display order character string 42 is bidirectional in that it includes left-to-right European numerals 123 and right-to-left Arabic letters NOP.

Appellant's method includes inferencing and reordering phases, as explained in the following paragraphs of the Specification:

> [0066] The method of the invention is divided into two phases: inferencing and reordering. Inferencing resolves the direction of indeterminate characters, such as the fall stop or "dot", hyphen-minus, Arabic numeral, and European numeral. During this phase each character is assigned a strong direction, either left or right. The reordering phase takes the fully resolved characters and generates a display ordering for them.
>
> [0067] The inferencing phase is accomplished in several passes. In the first pass Arabic and Hebrew letters are assigned the right-to-left direction, while full stops and other alphabetic characters are assigned the left-to-right direction. The next set of passes resolves the directions of digits.

The manner in which the directions of Arabic and European numerals and hyphen-minus characters are resolved is explained in paragraphs [0068] and [0069]:

> [0068] There are two rules for resolving the direction of Arabic and European numerals. All Arabic numerals are assigned the right-to-left direction. European numerals are assigned the left-to-right direction, unless the European numeral is surrounded by right-to-left characters (Arabic or Hebrew letters), in which case it takes the right-to-left direction. This is accomplished in two passes -- a forward pass and a reverse pass.
>
> [0069] The final set of passes resolves the directions of hyphen-minus characters. There are two rules for the resolution of hyphen-minus characters. All hyphen-minus characters become left-to-right, unless the hyphen-minus is surrounded by characters whose direction is right-to-left in which case the hyphen-minus becomes right-to-left. This is the same

> resolution as digits, but occurs after digit resolution. At this point each character in the domain name has a strong direction.

Appellant's algorithm accommodates two different groups of domain name creators. *Id*. at [0078]. One group knows what it wants to register (i.e., the logical order) but is unsure how it will be displayed. *Id*.[7] The other group knows what it wants to see displayed (i.e., the display order) but is unsure which logical sequence of characters should be registered. *Id*. Appellant's algorithm eliminates the need for specialized individual algorithms. *Id*.

B. The claims

The independent claims before us are claims 1, 5, and 9, of which claim 1 reads:

> 1. A method for converting a unidirectional domain name to a bidirectional domain name, said method comprising the steps of:
>
> establishing a plurality of labels within a unidirectional domain name by using a pre-determined full stop punctuation mark as a delimiter between said labels, said labels having an original label display order as encountered from left to right;
>
> within each said label, performing inferencing through resolving the direction of indeterminate characters by assigning

[7] In accordance with Appellant's invention, "multilingual domain name registration is made in logical order. . . . [C]onsistent with how bidirectional data is generally stored in files today." Specification [0065].

> a strong direction left or right to each indeterminate character; and
>
> reordering said characters within each said label of said unidirectional domain name into character display order using the fully resolved characters previously inferenced, thereby converting said uni-directional domain name to a bidirectional domain name in which said original label display order is preserved, and bidirectionality of characters within each label is produced.

Comparing claim 1 to Appellant's Figure 3, the recited "unidirectional domain name" appears to read on line 30 ("ABC.ibm.com"), which shows the "logical order" of the characters in each label, i.e., the order in which the characters would be written, regardless of whether those characters should be displayed in a left-to-right or a right-to-left order. The recited "bidirectional domain name" reads on line 33 (CBA.ibm.com"), which shows the display order of the characters in each label of the domain name.

The claim limitations concerning indeterminate characters do not read on Figure 3, which does not show any indeterminate characters. Instead, those limitations appear to be directed to Figure 4, wherein character string 40 represents the original order of a character string containing a hyphen-minus symbol.

The recited "reordering of said characters . . . of said unidirectional domain name into character display order using the fully resolved characters previously inferenced" therefore appears to be directed to converting a logical order character string like 40 in Figure 4 into a display order character string like 42 in Figure 4. The language "bidirectionality of

characters *within each label* is produced" (emphasis added) in the last two lines of the claim appears to refer to a single label that contains left-to-right *and* right-to-left characters, such as character string 42 in Figure 4.

The recited "original label display order as encountered left to right" in the unidirectional domain name refers to the left-to-right order of the labels as they appear in the logical order of the domain name, which label order the claim recites is preserved in the bidirectional (i.e., display order) domain name.

Appellant does not separately argue independent claim 5, which recites a computer readable medium encoded with computer software for accomplishing method steps like those recited in claim 1, or independent claim 9, which recites a "system" that comprises elements for performing the functions represented by those steps. Claims App., Br. 14, 16. Nor does Appellant separately argue any of the dependent claims.

C. The references and rejection

The Examiner relies on the following references:

Abir	US 6,738,827 B1	May 18, 2004
Feinberg	US 6,944,820 B2	Sep. 13, 2005

Claims 1-15 stand rejected under 35 U.S.C. § 103(a) for obviousness over Abir in view of Feinberg.

NEW GROUNDS OF REJECTION

Before addressing the merits of the Examiner's § 103(a) rejection, we are hereby, pursuant to our authority under 37 C.F.R. § 41.50(b), entering the following new grounds of rejection:

1. Claims 1-15 are rejected under 35 U.S.C. § 112, second paragraph, for being indefinite.

2. Claims 1-4, 9-13, and 15 are rejected under 35 U.S.C. § 101 as not directed to patent-eligible subject matter as defined in *In re Bilski*, 545 F.3d 943 (Fed. Cir. 2008) (en banc).

3. Claims 9-12 and 15 are also rejected under the enablement requirement of § 112, first paragraph.

A. The new § 112, second paragraph, rejection of claims 1-15

Independent claims 1, 5, and 9 are indefinite in a way that is common to all of those claims. Therefore, only claim 1 will be discussed.

It is not clear whether the phrase "a plurality of labels" in the first, "establishing" step refers to: (1) all of the labels within a domain name; or (2) to only those labels that contain indeterminate characters (thereby requiring at least two labels containing indeterminate characters). The language of the first step suggests the former interpretation. However, the latter interpretation is suggested by the second and third steps, of which the second step provides that "within *each* said label [of said plurality of labels], performing inferencing through resolving the direction of indeterminate

characters by assigning a strong direction left or right to each indeterminate character" (emphasis added), and the third step calls for "reordering said characters within *each* said label of said unidirectional domain name into character display order using the fully resolved characters previously inferenced . . . " (emphasis added). However, the claim language thus construed would limit the claim to domain names having at least two labels (the recited "plurality of labels") each containing bidirectional characters, which would not appear to be Appellant's intention.

B. The new § 101 and § 112, first paragraph (nonenablement) rejections

1. Claims 1-4 and 13

As noted above, Appellant describes the invention as a "single universal algorithm" (Specification [0078]) that provides a standardized way to convert between the logical order and the display order of domain names. *See also id*. at [0023] (describing an object of the invention as "provid[ing] a one-to-one mapping between names in logical order and names in display order.").

The exclusive test for patent-eligibility of a process under § 101 is whether the process: (1) is tied to a particular machine or apparatus, or (2) transforms a particular article into a different state or thing. *See Bilski*, 545 F.3d at 964 ("[T]he machine-or-transformation test is the only applicable test and must be applied, in light of the guidance provided by the Supreme Court and this court, when evaluating the patent-eligibility of

process claims."); *Gottschalk v. Benson*, 409 U.S. 63, 70 (1972) ("Transformation and reduction of an article 'to a different state or thing' is the clue to the patentability of a process claim that does not include particular machines."). Method claim 1 does not recite any machine or apparatus or call for transforming an article into a different state or thing. A domain name is simply a series of characters representing the address of a resource, such as a server, on the World Wide Web. Specification [0004], [0008], [0013]. All of the steps are data manipulation steps.

Nor do any of claims 2-4 and 13, which depend on claim 1, satisfy either test for patent eligible subject matter.

We are therefore rejecting claims 1-4 and 13 under § 101 for reciting patent ineligible subject matter.

2. Claims 9-12 and 15

Claim 9 recites essentially the same limitations as claim 1 but in "system" form:

> 9. A system for converting a unidirectional domain name to a bidirectional domain name comprising:
>
> a label definer adapted to establish a plurality of labels within a unidirectional domain name by using a pre-determined full stop punctuation mark as a delimiter between said labels, said labels having an original label display order as encountered from left to right;
>
> an inferencer adapted to, within each said label, resolve the direction of indeterminate characters by assigning a strong direction left or right to each indeterminate character; and

> a character reorderer adapted to reorder said characters within each said label of said unidirectional domain name into character display order using the fully resolved characters previously inferenced, thereby converting said uni-directional domain name to a bidirectional domain name in which said original label display order is preserved, and bidirectionality of characters within each label is produced.

Claims App. Br. 16.

The term "system" in the preamble is broad enough to read on a method and thus does not imply the presence of any apparatus. Although the body of the claim recites a "label definer," an "inferencer," and a "character reorderer," those recitations fail to serve as structural limitations because (1) they are not "means" recitations subject to interpretation under 35 U.S.C. § 112, sixth paragraph, and (2) they would not have been understood in the art as implying any particular structure. Therefore, in accordance with *Ex parte Miyazaki*, 89 USPQ2d 1207 (BPAI 2008) (precedential), we are treating claim 9 as encompassing any and all means for performing the recited functions and are accordingly rejecting that claim rejected under § 112, first paragraph, as based on a nonenabling disclosure in accordance with *Miyazaki*. Specifically, *Miyazaki* (a) held that the claim phrase "sheet feeding area operable to feed . . . " is not subject to interpretation under § 112, sixth paragraph (*id*. at 1216), (b) found that the term "sheet feeding area" has not been shown to have a definite structural meaning in the art (*id*.), and (c) concluded that the claim phrase in question therefore "encompasses any and all structures or acts for performing a

recited function, including those which were not what the applicant had invented" (*id*, at 1217), with the result that "the disclosure fails to provide a scope of enablement commensurate with the scope of the claim and the claim would violate the prohibition of *Halliburton* [*Oil Well Cementing Co. v. Walker*, 329 U.S. 1 (1946)]." *Id*.

Dependent claims 10-12 and 15 fail to recite any structure or include "means plus function" language and thus are also rejected under the enablement requirement of § 112, first paragraph.

Although *Miyazaki's* above-noted treatment of claim scope did not arise in the context of a ground of rejection for patent eligibility under § 101, that claim construction principle would appear to be applicable to determinations of patent eligibility under § 101. We are therefore additionally rejecting claims 9-12 and 15 under § 101 for reciting patent ineligible subject matter because, when construed in accordance with *Miyazaki* in the manner explained above in the discussion of the new rejection of those claims under § 112, first paragraph, for nonenablement, those claims encompass any and all structures for performing the recited functions. As a result, claims 9-12 and 15 are at least as broad as method claims 1-4 and 13, which we have determined recite patent ineligible subject matter under *Bilski*.[8]

[8] We leave it to the Examiner to determine in the first instance whether claims 5-8 and 14, which recite a "computer readable medium encoded with computer executable software," recite patent eligible subject
(Continued on next page.)

THE EXAMINER'S § 103(a) REJECTION

In view of the new ground of rejection entered above against claims 1-15 under 35 U.S.C. § 112, second paragraph, for indefiniteness, we are hereby reversing the Examiner's prior-art rejection of those claims. *In re Steele*, 305 F.2d 859, 862-63 (CCPA 1962) (reversing § 103 rejection because based on considerable speculation as to meaning of terms of claims and assumptions as to their scope). However, in the interest of completeness, we offer the following observations regarding the merits of the § 103(a) rejection.

A. Principles of law

"[T]he examiner bears the initial burden, on review of the prior art or on any other ground, of presenting a *prima facie* case of unpatentability." *In re Oetiker*, 977 F.2d 1443, 1445 (Fed. Cir. 1992). A rejection under 35 U.S.C. § 103(a) must be based on the following factual determinations: (1) the scope and content of the prior art; (2) the level of ordinary skill in the art; (3) the differences between the claimed invention and the prior art; and (4) any objective indicia of non-obviousness. *DyStar Textilfarben GmbH & Co. Deutschland KG v. C.H. Patrick Co.*, 464 F.3d 1356, 1360 (Fed. Cir. 2006) (citing *Graham v. John Deere Co.*, 383 U.S. 1, 17 (1966)).

---

matter.

"The combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results." *Leapfrog Enters., Inc. v. Fisher-Price, Inc.*, 485 F.3d 1157, 1161 (Fed. Cir. 2007) (quoting *KSR Int'l Co. v. Teleflex Inc.*, 127 S. Ct. 1727, 1739 (2007)). "[W]hen a patent claims a structure already known in the prior art that is altered by the mere substitution of one element for another known in the field, the combination must do more than yield a predictable result." *KSR*, 127 S. Ct. at 1740.

Discussing the obviousness of claimed combinations of elements of prior art, *KSR* explains:

> When a work is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or a different one. If a person of ordinary skill can implement a predictable variation, §103 likely bars its patentability. For the same reason, if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill. *Sakraida* [*v. AG Pro, Inc.*, 425 U.S. 273 (1976)] and *Anderson's-Black Rock*[, *Inc. v. Pavement Salvage Co*., 396 U.S. 57 (1969)] are illustrative—a court must ask whether the improvement is more than the predictable use of prior art elements according to their established functions.

Id. If the claimed subject matter "involve[s] more than the simple substitution of one known element for another or the mere application of a known technique to a piece of prior art ready for the improvement," *id*.,

> it will be necessary . . . to look to interrelated teachings of multiple patents; the effects of demands known to the design community or present in the marketplace; and the background knowledge possessed by a person having ordinary skill in the art, all in order to determine whether there was an apparent reason to combine the known elements in the fashion claimed by the patent at issue.

Id. at 1740-41. "To facilitate review, this analysis should be made explicit." *Id*. at 1741. That is, "there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *Id*. (quoting *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006)).

B. The level of skill in the art

Appellant faults the Examiner for failing to ascertain the level of skill in the art, citing *Graham*. Reply Br. 4. In the absence of other evidence addressing the level of skill, it is presumed to be represented by the references themselves. *See In re Oelrich*, 579 F.2d 86, 91 (CCPA 1978) ("the PTO usually must evaluate both the scope and content of the prior art and the level of ordinary skill solely on the cold words of the literature"); *In re GPAC Inc*., 57 F.3d 1573, 1579 (Fed. Cir. 1995) (Board did not err in adopting the approach that the level of skill in the art was best determined by the references of record).

C. The Abir reference

> Abir's invention
>
> > contemplates transformation of a conventional URL to an identifier that is familiar to the user. (This familiar identifier will be referred to as a "friendly" identifier.) This transformed identifier would be displayed to the user. Internally, the apparatus would continue to use the conventional URL and would use this URL in accessing resources on the Internet.

Abir, col. 2, l. 62 to col. 3, l. 1.

Abir's invention also can be used to convert a friendly identifier that has been input by the user into a conventional identifier to be used by the apparatus. *See id.*, col. 2, ll. 58-61 ("It is an object of the invention to provide a system and method for a user of an apparatus to access a HTML page by inputting to the apparatus a resource identifier that is simpler or otherwise more desirable than the conventional URL.").

Abir discloses a specific example of converting a conventional (i.e., English-language) URL to a "friendly identifier" in Hebrew. *Id.*, col. 4, ll. 29-41.

Figure 1 is reproduced below.

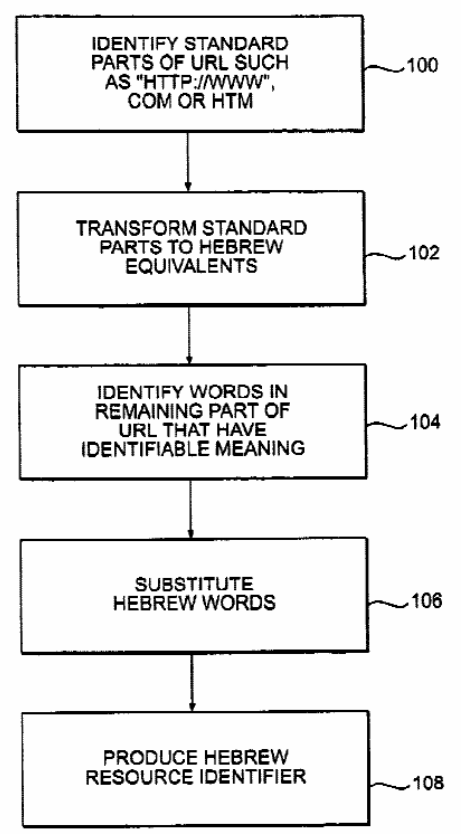


FIG. 1

Figure 1 is a block diagram of a "conversion" algorithm for transforming a conventional resource identifier into a friendly resource identifier, using conversion to Hebrew as an illustrative example. *Id*., col. 4, ll. 23-29. The operation of the algorithm is described as follows:

> In step 100, standard parts of conventional resource identifiers such as "http://www" "com" and "htm" are identified. In step 102, the standard parts are converted to well-known Hebrew equivalents such as {character pullout} for "http://www" and {character pullout} for "com". In step 104, the remaining parts of the conventional resource identifier is analyzed for words

> that have identifiable meanings. For example, the words "health" and "insurance" would be recognized in the word "healthinsurance". In step 106, the Hebrew word {character pullout} would be substituted for "health" and the Hebrew word {character pullout} would be substituted for insurance. In step 108, the complete Hebrew resource identifier would be produced.

Id., col. 4, ll. 29-42.

Abir's Figure 3 is reproduced below.

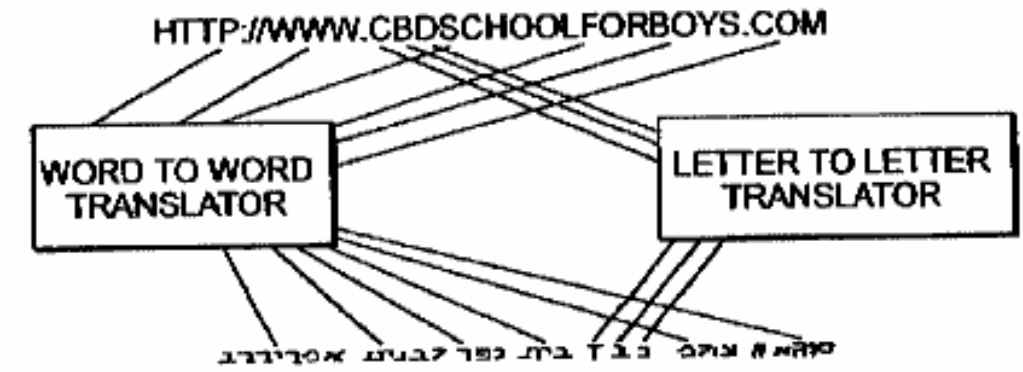


FIG. 3

Figure 3 is block diagram showing the application of Abir's conversion algorithm to the conventional resource locator "http://www.cdbschoolfor boys.com." *Id*., col. 3, ll. 13-14. A letter to letter translator is used to convert letters that are not parts of words recognized by the word to word translator. *Id*., col. 4, ll. 42-55.

Abir also describes reversing the order of words when converting words into Hebrew words:

> [T]he address in the case of the address http://www.healthinsurance.com, the system will isolate the part of the address that comes after "http://www" and before the ".com" (also after the ".com" for sub sites.) The system will then convert the words into the Web surfer's native language using the system's simple English Hebrew dictionary. In this case health insurance. Then the system may reverse the order of the Hebrew words[,] add the Hebrew variation of the "http://www and the ".com", and display the address in the surfer's native language.

Id., col. 6, ll. 22-31. Presumably, reversal of the words also involves reversal of the letters of each word.

Abir does not address domain names that contain left-to-right and right-to-left characters. Nor does Abir address indeterminate (i.e., "weak") characters, such as hyphens. Furthermore, Abir fails to address a domain name that includes a subdomain, such as represented by "help" in Appellant's above-noted URL example, "http://www.help.ibm.com." Br. 5. As a result, Appellant's characterization of Abir as "teach[ing] treating the entire set of characters which are not 'standard parts' as a string to be converted to the alternate language" (Br. 8) is overly broad. Abir only discusses URLs that do not include subdomains.

D. The Feinberg reference

Feinberg's invention relates to detecting and correcting the reading order of text rendered in a bi-directionally rendered language environment (Feinberg, col. 1, ll. 6-10). As correctly noted by Appellant, "Feinberg is

silent regarding applying any part or form of their invention to URL's . . . ." Br. 12.

Feinberg's invention is specifically directed to resolving ambiguities when converting bidirectional Unicode text from the logical order to the correct display order. Feinberg, col. 1, l. 41 to col. 2, l. 14. Such ambiguities arise, for example, when the directional text includes a combination of letters, hyphens and numbers, such as the text "I live in house--12." *Id*., col. 2, ll. 10-14.

In Feinberg's invention, a text selection is scanned for a portion of the text selection that must be rendered differently than other portions of the text selection according to text rendering rules of the language to which the text selection belongs. *Id*., col. 2, ll. 55-58. The beginning and end of the portion of the text selection is marked, and the portion of the text selection is rendered differently than other portions of the text selection according to the rules of the language to which the text selection belongs. *Id*., col. 2, ll. 59-63.

More particularly, a text selection is scanned to locate a hyphen character indicating a beginning of a portion of text that may need to be rendered in left-to-right reading order. *Id*., col. 2, l. 66 to col. 3, l. 1. The text selection is further scanned to locate a second character indicating an end of the portion of text that may need to be rendered in left-to-right reading order. *Id*., col. 3, ll. 1-4. A determination is made as to whether the portion of text must be rendered in left-to-right reading order. *Id*., col. 3, ll.

4-6. If so, that portion of the text selection is rendered in a left-to-right reading order. *Id*., col. 3, ll. 6-7.

During operation of Feinberg's invention on text rendered according to the rules of the Hebrew language, the module 205 (Fig. 2) scans every character looking for a dash "--." *Id*., col. 6, ll. 30-33. When the module 205 locates a dash, the text including the dash is highlighted and the user is given a prompt to ask the user whether the text surrounding the dash should be reversed or flipped into a left-to-right configuration. *Id*., col. 6, ll. 33-37. For example, say the text includes the equation "3-2=1", and upon scanning the text the string "1=2-3" is highlighted. *Id*., col. 6, ll. 37-39. The user is asked via a prompt whether the highlighted text should be corrected. *Id*., col. 6, ll. 39-40. If the user accepts, the string is flipped to correctly read "3-2=1," while the remaining text continues to be configured in a right-to-left configuration according to the rules of Hebrew text rendering. *Id*., col. 6, ll. 40-43.

Alternatively, the module 205 may be set to automatic correction in which case the highlighted text is flipped to the opposite reading order automatically. *Id*., col. 9, ll. 2-4.

As noted by the Examiner (Answer 4), Feinberg's invention also involves recognizing separators, such as periods:

> Referring back to FIG. 4c, if at step 442 a determination is made that the character obtained in step 438 is not a dash, the method then proceeds along the "No" branch to step 448, where a determination is made whether the character is a separator,

> such as a colon, a period, or a comma. It should be understood that the colon, the period and the comma are not defined as separators according to the Unicode standard, but are defined as separators for operation of an exemplary embodiment of the present invention.

Id., col. 9, ll. 54-62. The Examiner characterizes this passage as "specifically teach[ing] the period can be used as a separator to mark or label sections of text for processing." Answer 8. Appellant does not disagree with this characterization but argues that Feinberg treats periods as delimiting *sentences* of text rather than delimiting labels in a domain name. *See* Br. 11 ("Feinberg is most certainly addressed to natural language processing, but URL's are not 'natural language' in the sense that Feinberg addresses natural language. The period characters, or full stop characters, in a URL do not delimit full sentences of 'spoken language' (col. 1, line 18).").

Appellant more particularly explains, without contradiction by the Examiner, how Feinberg will translate text containing two periods:

> When applying conventional natural language translation techniques, a Latin period "." character is typically interpreted as signaling the end of a sentence construct within a paragraph, unless it is immediately followed by a paragraph termination character, such as a hard line feed ("LF") or carriage return ("CR") character. So, for example, if the *words of* the phrase:
>
> "I own a dog. It is a good dog. <CR>"
>
> were re-ordered for right-to-left languages and interpreted using conventional natural language translation techniques, it would appear in the following order:

> ".dog good a is It. dog a own I."
>
> Notice that the *sentences* reversed order, as well as the words within the sentences. This is a fundamental problem of the Unicode Bidirectional Algorithm ("BIDI") as applied to domain names, which arises due to the fact that the algorithm was designed to process natural language text (e.g. sentences and paragraphs), not URLs.

Br. 9.

E. The merits of the rejection

The Examiner found that "[t]he system of Abir determines and detects the standard parts of a URL (http://www, '.com', etc) without specifically disclosing the implementation of parsing the domain name into 'labels' based on detected delimiters." Answer 3. More particularly, the Examiner found that "Abir processes the URL to separate the URL into parts but failed to specifically teach what processing was implemented so as to achieve the parts." *Id*. at 9. The Examiner then explained that

> Feinberg was cited for teaching the processing of text to detect for various delimiters or separator characters (Feinberg specifically suggests various characters can be used as separators - colon, period, comma, hyphen, dash, slash) so as to mark or label the beginning and/or end of text that needs to be corrected or processed. Therefore, the Examiner maintains that the combination of Abir and Feinberg provide adequate support for the claim language.

Id.

It seems to us the Examiner's reliance on Feinberg for a teaching of using a period as a delimiter is unnecessary because that feature is inherent in Abir. Although Abir does not use the terms "delimiter" or "labels" to describe separating the URL "http://www.schoolforboys.com" into the standard parts "http://www" and "com" and the nonstandard part "schoolforboys," the claim term "delimiter" reads on the periods and the claim term "labels" reads on the parts separated by those periods. Appellant does not contend otherwise.

Regarding the recited bidirectionality of characters, which is not disclosed in Abir, the Examiner concluded:

> It would have been obvious to one of ordinary skill at the time of the invention to modify the system of Abir to implement the bi-directional text processing of Feinberg, for the purpose of ensuring that the alternate Internet and resource locators of Abir are rendered in the proper reading order for bi-directional or regular text, as suggested by Feinberg.

Final Action 3-4; Answer 4. In view of this explanation, Appellant is incorrect to assert that at page 11 of the Brief that "[t]here is no statement . . . why Feinberg was employed in a § 103 combination."

At page 2 of the Reply Brief, Appellant acknowledges the Examiner's above-quoted "proper reading order" rationale but argues that combining Abir and Feinberg will not result in preserving the "original label display order," as required by the claims. Specifically, Appellant contends that the "'[p]roper reading order' [in Feinberg] refers to natural language syntax for spoken languages, and not only specifies a left-to-right or right-to-left

order of reading, but also specifies orders of sentences relative to each other." *Id*. This argument apparently relates back to the explanation in the Brief of how Abir would convert Appellant's example URL, "http://www.applyforaloan.bigbank.com." Br. 9. Appellant argues that in contrast to Appellant's invention, which would divide this URL into four labels ("http://," "applyforaloan," "bigbank," and "com") and preserve their display order even if converted to the Hebrew language, Abir's invention would divide this URL into two standard parts ("http://www" and "com") and one nonstandard part of ("applyforaloan.bigbank") and would lose the original label display order of "applyforaloan" and "bigbank" because the nonstandard character string "applyforaloan.bigbank" would be treated as two sentences, whose sentence order and letter order would be reversed as follows during conversion of the URL to the Hebrew language:

<A>knabgib.naolarofylppa<B>,

where <A> is Abir's substitution for "http://www" and <B> is Abir's substitution for "com." Br. 9.

In our view, Abir's Figure 14, reproduced below, suggests that Appellant has the positions of <A> and <B> reversed.

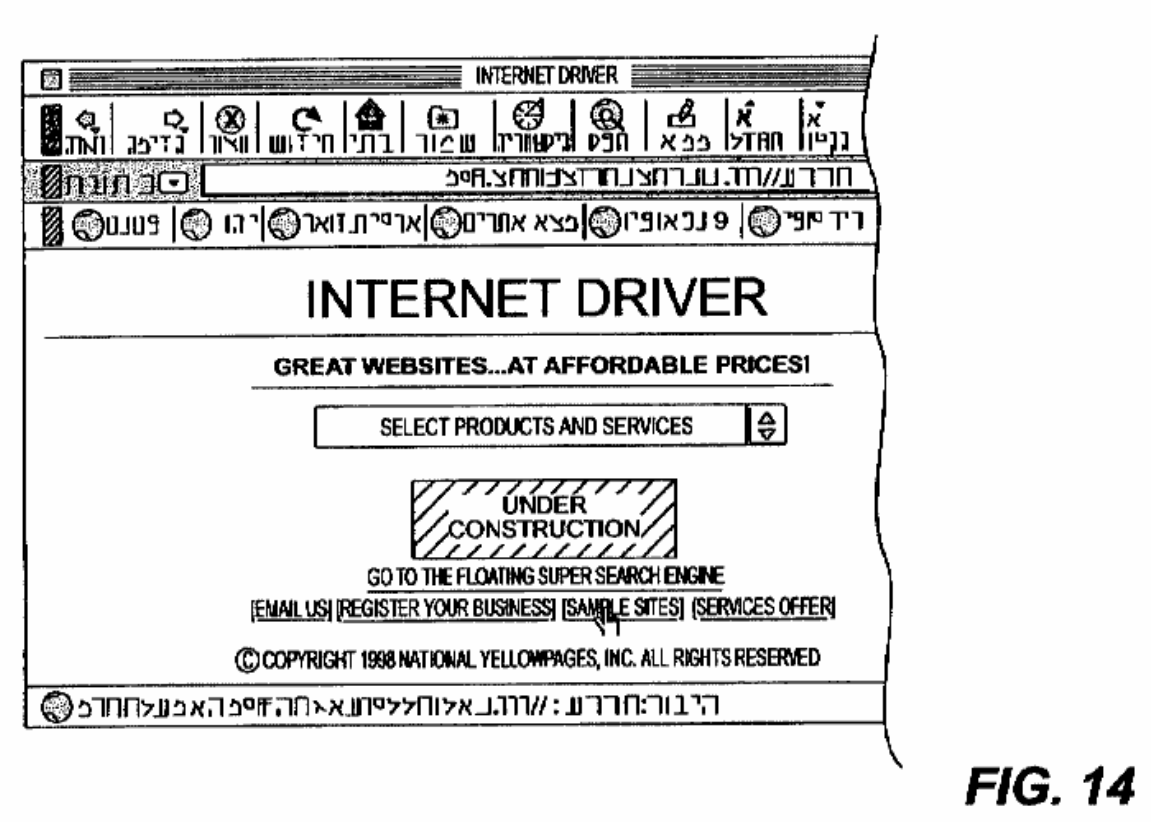


FIG. 14

Figure 14 represents a screen shot of a Hebrew-language browser employing Abir's invention. Abir, col. 3, ll. 32-37. The Hebrew-language equivalent for "http://www" given at column 4, lines 31-33 appears to be located at the right end of the address line (albeit minus the colon, which we assume is an oversight), in which case the three Hebrew letters at the left end of the address line presumably represent "com."[9] It would therefore appear that the result of applying Abir as modified in view of Feinberg to Appellant's example should be represented as follows:

[9] We say "presumably" because the Hebrew letters do not strongly
(Continued on next page.)

<B>knabgib.naolarofylppa<A>

where <A> is Abir's substitution for "http://www" and <B> is Abir's substitution for "com." This represents complete reversal of the original label display order.

In any case, the Examiner, in response to Appellant's argument that Abir fails to preserve the label display order, stated that "[t]he processing and translation of subsites and the parts after the '.com' of the URL to ensure a complete and proper translation requires that the order of domains and subdomains or subsites are maintained to ensure the proper cites [*sic*; sites] are accessed after transformation." Answer 10. This explanation is not understood, because it fails to explain why changing the order of the Hebrew labels displayed in Abir's Figure 14 URL address line would have ensured accessing of the proper sites. The Hebrew version of the URL in Abir is used for display purposes only; the computer would continue to use the conventional English-language URL to access resources on the Internet. Abir, col. 2, l. 62 to col. 3, l. 1.

Appellant also argued that Feinberg's processing method is not "fully automatic," because the user is prompted by highlighting to indicate how questionable text is to be displayed. Br. 12. The Examiner correctly held (Answer 12) that this argument improperly reads a limitation from Appellant's Specification into the claims, which do not require "fully

---

resemble the Hebrew equivalent letters for "com" given in column 4, line 33.

automatic" operation. Appellant's argument also fails to take into account the automatic operation described by Feinberg at column 9, lines 2-5.

SUMMARY

We have entered the following new grounds of rejection pursuant to our authority under 37 C.F.R. § 41.50(b):

1. Claims 1-15 are rejected under § 112, second paragraph, for indefiniteness.

2. Claims 1-4, 9-13, and 15 are rejected under § 101 for reciting patent ineligible subject matter.

3. Claims 9-12 and 15 are rejected under § 112, first paragraph, as based on a nonenabling disclosure.

In view of the new ground of rejection of claims 1-15 under 35 U.S.C. § 112, second paragraph, for indefiniteness, we have reversed the Examiner's rejection of claims 1-15 under § 103(a) for obviousness over Abir in view of Feinberg. *Steele*, 305 F.2d at 862-63.

APPELLANT'S OPTIONS FOR RESPONDING TO THE NEW GROUNDS OF REJECTION

Regarding the new grounds of rejection entered pursuant to 37 C.F.R. § 41.50(b), that paragraph explains that "[a] new ground of rejection pursuant to this paragraph shall not be considered final for judicial review." Appellant, <u>within TWO MONTHS from the date of this decision</u>, must exercise one of the following two options with respect to the new grounds of

rejection to avoid termination of the appeal as to the rejected claims:

> (1) *Reopen prosecution*. Submit an appropriate amendment of the claims so rejected or new evidence relating to the claims so rejected, or both, and have the matter reconsidered by the Examiner, in which event the proceeding will be remanded to the Examiner. . . .
>
> (2) *Request rehearing*. Request that the proceeding be reheard under § 41.52 by the Board upon the same record. . . .

37 C.F.R. § 41.50(b) (2008).

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a).

REVERSED; 37 C.F.R. § 41.50(b)

msc

IBM CORPORATION (RHF)
C/O ROBERT H. FRANTZ
P. O. BOX 23324
OKLAHOMA CITY, OK 73123

Enclosures:

Glossary of Unicode Terms, http://www.unicode.org/glossary (last visited Jan. 28, 2009), pages 1, 13, and 23 of 41 printout pages.

Wikipedia, http://en.wikipedia.org/wiki/Domain_name (last visited Jan. 28, 2009).

Wikipedia, http://en.wikipedia.org/wiki/Bi-directional_text (last visited Jan. 28, 2009).

Glossary

A-Z index

To jump to a particular section of the glossary, click on a letter below.

A B C D E F G H I J K L M
N O P-Q R S T U V W X-Y Z

Finding Terms

Use the "Find (on this page)..." command in your browser to locate specific terms in this glossary.

Related Links

The Unicode Standard, Version 5.1
The Unicode Standard, Version 5.0
FAQ for more general information
Technical Reports for specific topics in depth
Terminology for translations of terms

Glossary of Unicode Terms

This glossary is originally based on *The Unicode Standard, Version 5.0*, and has been updated for Unicode Version 5.1. It will continue to be updated over time. Some definitions from the 5.0 book were slightly modified for the Web. In particular, some glyph examples used in definitions in the book are not included here. References to sections, tables, etc., refer to the relevant location in the 5.0 book, except where otherwise noted. For the glossary as published in *The Unicode Standard, Version 5.0*, see the 5.0 Glossary in the electronic edition of the standard.

Translations of Unicode and ISO/IEC 10646 terminology are also available.

A

Abjad. A writing system in which only consonants are indicated. The term "abjad" is derived from the first four letters of the traditional order of the Arabic script: *alef, beh, jeem, dal.* (See Section 6.1, Writing Systems.)

Abstract Character. A unit of information used for the organization, control, or representation of textual data. (See definition D7 in Section 3.4, Characters and Encoding.)

Abstract Character Sequence. An ordered sequence of one or more abstract characters. (See definition D8 in Section Section 3.4, Characters and Encoding.)

Abugida. A writing system in which consonants are indicated by the base letters that have an inherent vowel, and in which other vowels are indicated by additional distinguishing marks of some kind modifying the base letter. The term "abugida" is derived from the first four letters of the Ethiopic script in the Semitic order: *alf, bet, gaml, dant.* (See Section 6.1, Writing Systems.)

Accent Mark. A mark placed above, below, or to the side of a character to alter its phonetic value. (See also *diacritic*.)

Acrophonic. Denoting letters or numbers by the first letter of their name. For example, the Greek acrophonic numerals are variant forms of such initial letters.

Aksara. (1) In Sanskrit grammar, the term for "letter" in

linguistic information). Also called *diacritical mark* or *diacritical*. (See also *combining character* and *nonspacing mark*.)

Diaeresis. Two horizontal dots over a letter, as in *naïve*. The diaeresis is not distinguished from the *umlaut* in the Unicode character encoding. (See *umlaut*.)

Dialytika. Greek term for *diaeresis* or *trema*, used in Greek character names.

Digits. (See *Arabic digits*, *European digits*, and *Indic digits*.)

Digraph. A pair of signs or symbols (two graphs), which together represent a single sound or a single linguistic unit. The English writing system employs many digraphs (for example, *th, ch, sh, qu,* and so on). The same two symbols may not always be interpreted as a digraph (for example, *cathode* versus *cathouse*). When three signs are so combined, they are called a *trigraph*. More than three are usually called an *n-graph*.

Dingbats. Typographical symbols and ornaments.

Diphthong. A pair of vowels that are considered a single vowel for the purpose of phonemic distinction. One of the two vowels is more prominent than the other. In writing systems, diphthongs are sometimes written with one symbol and sometimes with more than one symbol (for example, with a *digraph*).

Direction. (See *paragraph direction*.)

Directionality Property. A property of every graphic character that determines its horizontal ordering as specified in Unicode Standard Annex #9, "The Bidirectional Algorithm." (See Section 4.4, Directionality—Normative.)

Display Cell. A rectangular region on a display device within which one or more glyphs are imaged.

Display Order. The order of glyphs presented in text rendering.

Double-Byte Character Set. One of a number of character sets defined for representing Chinese, Japanese, or Korean text (for example, JIS X 0208-1990). These

Informative property of characters that are used to write words.

Ligature. A glyph representing a combination of two or more characters. In the Latin script, there are only a few in modern use, such as the ligatures between "f" and "i" or "f" and "l". Other scripts make use of many ligatures, depending on the font and style.

Little-endian. A computer architecture that stores multiple-byte numerical values with the least significant byte (LSB) values first.

Locale Data Markup Language. The XML specification for the exchange of locale data, defined by Unicode Technical Standard #35, "Locale Data Markup Language (LDML)." (See also *Common Locale Data Repository*.)

Logical Order. The order in which text is typed on a keyboard. For the most part, logical order corresponds to phonetic order. (See Section 2.2, Unicode Design Principles.)

Logical Store. Memory representation.

Logosyllabary. A writing system in which the units are used primarily to write words and/or morphemes of words, with some subsidiary usage to represent just syllabic sounds. The best example is the Han script.

Lowercase. (See *case*.)

Low-Surrogate Code Point.. A Unicode code point in the range U+DC00 to U+DFFF. (See definition D73 in Section 3.8, Surrogates.)

Low-Surrogate Code Unit. A 16-bit code unit in the range $DC00_{16}$ to $DFFF_{16}$, used in UTF-16 as the trailing code unit of a surrogate pair. Also known as a *trailing surrogate*. (See definition D74 in Section 3.8, Surrogates.)

LSB. Acronym for *least significant byte*.

LZW. Acronym for *Lempel-Ziv-Welch*, a standard algorithm widely used for compression of data.

M

Domain name

From Wikipedia, the free encyclopedia

> *This article primarily discusses registered Internet domain names. See the Domain Name System article for technical discussions about the Domain Name System and the hostname article for general discussion of naming aspects of Internet hosts.*

The main purpose of a **domain name** is to provide symbolic representations, i.e., recognizable names, to mostly numerically addressed Internet resources. This abstraction allows any resource (e.g., website) to be moved to a different physical location in the address topology of the network, globally or locally in an intranet, in effect changing the IP address. This translation from domain names to IP addresses (and vice versa) is accomplished with the global facilities of Domain Name System (DNS).

By allowing the use of unique alphabetical addresses instead of numeric ones, domain names allow Internet users to more easily find and communicate with web sites and any other IP-based communications services. The flexibility of the domain name system allows multiple IP addresses to be assigned to a single domain name, or multiple domain names to be services from a single IP address. This means that one server may have multiple roles (such as hosting multiple independent websites), or that one role can be spread among many servers. One IP address can also be assigned to several servers, as used in anycast networking.

Contents

- 1 Defined
- 2 Examples
- 3 Top-level domains
- 4 Second-level and lower level domains
- 5 Official assignment
- 6 Abuses
- 7 Generic domain names—problems arising from unregulated name selection
- 8 Unconventional domain names
- 9 Premium domain names
- 10 Resale of domain names
- 11 Domain aftermarket prices and trends
- 12 Popular domain prefixes - "E" and "I"
- 13 Branding with a domain name
- 14 Domain name confusion
- 15 References
- 16 See also
- 17 External links

Defined

By definition (RFC 1034 (http://tools.ietf.org/html/rfc1034)), domain names are restricted to the ASCII letters *a* through *z* (case-insensitive), the digits *0* through *9*, and the hyphen, with some other restrictions in terms of name length and position of hyphens. Since this does not allow the use of many characters

commonly found in non-English languages, and no multi-byte characters necessary for most Asian languages, the Internationalized domain name (IDN) system has been developed and is now in testing stage with a set of top-level domains established for this purpose.

The underscore character is frequently used to ensure that a domain name is not recognized as a hostname, as with the use of SRV records, for example, although some older systems such as NetBIOS did allow it. To avoid confusion and for other reasons, domain names with underscores in them are sometimes used where hostnames are required.

Domain names are often referred to simply as *domains* and domain name registrants are frequently referred to as *domain owners*, although domain name registration with a registrar does not confer any legal ownership of the name, only an exclusive right of use.

Examples

The following example illustrates the difference between a URL (Uniform Resource Locator) and a domain name:

URL: `http://www.example.net/index.html`
Domain name: `www.example.net`
Registered domain name: `example.net`

As a general rule, the IP address and the server name are interchangeable. For most Internet services, the server will not have any way to know which was used. However, the explosion of interest in the Web means that there are far more Web sites than servers. To accommodate this, the hypertext transfer protocol (HTTP) specifies that the client tells the server which name is being used. This way, one server with one IP address can provide different sites for different domain names. This feature goes under the name *virtual hosting* and is commonly used by web hosts.

For example, as referenced in RFC 2606 (http://tools.ietf.org/html/rfc2606) (Reserved Top Level DNS Names), the server at IP address 208.77.188.166 handles all of the following sites:

example.com
www.example.com
example.net
www.example.net
example.org
www.example.org

When a request is made, the data corresponding to the hostname requested is provided to the user.

Top-level domains

Every domain name ends in a top-level domain (TLD) name, which is always either one of a small list of generic names (three or more characters), or a two-character territory code based on ISO-3166 (there are few exceptions and new codes are integrated case by case). Top-level domains are sometimes also called first-level domains.

The generic top-level domain (gTLD) extensions are:

The country code top-level domain (ccTLD) extensions are:

Second-level and lower level domains

Below the top-level domains in the domain name hierarchy are the second-level domain (SLD) names. These are the names directly to the left of .com, .net, and the other top-level domains. As an example, in the domain *en.wikipedia.org*, *wikipedia* is the second-level domain.

Next are third-level domains, which are written immediately to the left of a second-level domain. There can be fourth- and fifth-level domains, and so on, with virtually no limitation. An example of a working domain with four domain levels is *www.sos.state.oh.us*. The www preceding the domains is a host name of the World-Wide Web server. Each level is separated by a dot, or period symbol. 'sos' is said to be a sub-domain of 'state.oh.us', and 'state' a sub-domain of 'oh.us', etc. In general, Sub-domains are domains subordinate to their parent domain. An example of very deep levels of subdomain ordering are the IPv6 reverse resolution DNS zones, e.g., 1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.ip6.arpa, which is the reverse DNS resolution domain for the IP address of a loopback interface, or the localhost name.

Second-level (or lower-level, depending on the established parent hierarchy) domain names are often created based on the name of a company (e.g., *microsoft*.com), product or service (e.g., *gmail*.com). Below these levels, the next domain name component has been used to designate a particular host server. Therefore, *ftp.wikipedia.org* might be an FTP server, *www.wikipedia.org* would be a World Wide Web server, and *mail.wikipedia.org* could be an email server, each intended to perform only the implied function. Modern technology allows multiple physical servers with either different (cf. load balancing) or even identical addresses (cf. anycast) to serve a single hostname or domain name, or multiple domain names to be served by a single computer. The latter is very popular in Web hosting service centers, where service providers host the websites of many organizations on just a few servers.

Official assignment

The Internet Corporation for Assigned Names and Numbers (ICANN) has overall responsibility for managing the DNS. It administers the root domain, delegating control over each TLD to a domain name registry. For ccTLDs, the domain registry is typically installed by the government of that country. ICANN has a consultation role in these domain registries but cannot regulate the terms and conditions of how domain names are delegated in each of the country-level domain registries. On the other hand, the generic top-level domains (gTLDs) are governed directly under ICANN, which means all terms and conditions are defined by ICANN with the cooperation of each gTLD registry.

Domain names are often seen in analogy to real estate in that (1) domain names are foundations on which a website (like a house or commercial building) can be built and (2) the highest "quality" domain names, like sought-after real estate, tend to carry significant value, usually due to their online brand-building potential, use in advertising, search engine optimization, and many other criteria.

A few companies have offered low-cost, below-cost or even cost-free domain registrations with a variety of models adopted to recoup the costs to the provider. These usually require that domains be

hosted on their website within a framework or portal that includes advertising wrapped around the domain holder's content, revenue from which allows the provider to recoup the costs. Domain registrations were free of charge when the DNS was new. A domain holder (often referred to as a domain *owner*) can give away or sell infinite number of subdomains under their domain name. For example, the owner of *example.edu* could provide subdomains such as *foo.example.edu* and *foo.bar.example.edu* to interested parties.

Abuses

As domain names became interesting to marketers because of their advertising and marketing potential, rather than just being used to label Internet resources in a technical fashion, they began to be used in manners that in many cases did not reflect the intended purpose of the label of their top-level domain. As originally planned, the structure of domain names followed a hierarchy in which the TLD indicated the type of organization (commercial, governmental, etc.), and addresses would be nested down to third, fourth, or further levels to express complex structures, where, for instance, branches, departments and subsidiaries of a parent organization would have addresses in subdomains of the parent domain. Also, hostnames were originally intended to correspond to actual physical machines on the network, generally with only one name per machine.

As the World Wide Web became popular, site operators frequently wished to have memorable addresses, regardless of whether they fit properly into the structure; thus, because the .com domain was the most popular and therefore most prestigious, even noncommercial sites began to obtain domains directly within that gTLD, and many sites desired second-level domain names in .com, even if they were already part of a larger entity where a subdomain would have been logical (e.g., *abcnews.com* instead of *news.abc.com*).

Shorter, and therefore more memorable, domain names are thought to have more appeal. As a convenience methods were implemented to reduce the amount of typing required when entering a web site address into the location field of a web browser. A website found at "http://www.example.org" will often be advertised without the *http://*, since the HTTP protocol is implicitly assumed when referring to web sites. In many cases, web sites can be also be reached by omitting the www prefix, as in this given example. This feature is usually implemented in DNS by the website administrator. In the case of a .com, the website can sometimes be reached by just entering *example* (depending on browser versions and configuration settings, which vary in how they interpret incomplete addresses).

The popularity of domain names also led to uses which were regarded as abusive by established companies with trademark rights; this has become known as *cybersquatting*, in which a person registers a domain name that resembles a trademark in order to profit from visitors looking for that address. To combat this, various laws and policies were enacted to allow abusive registrations to be forcibly transferred, but these were sometimes themselves abused by overzealous companies committing reverse domain hijacking against domain users who had legitimate grounds to hold their names. Such legitimate uses could include the use of generic words that are contained within a trademark, but used in a particular context within the trademark, or their use in the context of fan or protest sites with free speech rights of their own.

As of 2008, the four major Registrars have all sub-contracted their expiring domain lists to certain reseller and auctioneer partnerships, for the purpose of keeping the domain name at the original registrar and continuing to extract revenue off the renewal of premium registered names. Since this policy is not

explicitly banned at ICANN, the practice has become more commonplace and as a result, complaints from individual registrants about losing their domains has tracked higher over the past two years [1] (http://www.phraseologist.com/2008/07/how-to-keep-your-existing-domain-name.html).

Laws that specifically address domain name conflicts include the Anticybersquatting Consumer Protection Act in the United States and the Trademarks Act of 1999 in India. Alternatively, domain registrants are bound by contract under the UDRP to comply with mandatory arbitration proceedings should someone challenge their ownership of a domain name.

Often email phishing scams will abuse subdomain names to appear to be a legitimate site. For instance, an email might purport to be from Bank of America, and include a link to a fake login screen hosted on http://www.bankofamerica.com.abc.def.ghi.jkl/ In this case, the actual domain is ghi.jkl, but appears at first glance to be bankofamerica.com.

Generic domain names—problems arising from unregulated name selection

Within a particular TLD, parties are generally free to register an undelegated domain name on a first come, first served basis, resulting in Harris's lament, *all the good ones are taken*. For generic or commonly used names, this may sometimes lead to the use of a domain name which is inaccurate or misleading. This problem can be seen with regard to the ownership or control of domain names for a generic product or service. By way of illustration, there has been tremendous growth in the number and size of literary festivals around the world in recent years. In the current context, a generic domain name such as *literary.org* is available to the first literary festival organization that is able to obtain the registration, even if the festival in question is very young or obscure. Some critics argue that there is greater amenity in reserving such domain names for the use of, for example, a regional or umbrella grouping of festivals. Related issues may also arise in relation to noncommercial domain names.

Unconventional domain names

Due to the rarity of one-word dot-com domain names, many unconventional domain names, domain hacks, have been created. They make use of the top-level domain as an integral part of the Web site's title. Two popular domain hack Web sites are `del.icio.us` and `blo.gs`, which spell out "delicious" and "blogs", respectively. Delicious.com later reverted to a normal domain name, as the unconventional one was difficult to remember.[2] (http://blog.delicious.com/blog/2008/07/oh-happy-day.html)

Unconventional domain names are also used to create unconventional email addresses. Non-working examples that spell 'James' are `j@m.es` and `j@mes.com`, which use the domain names `m.es` (of Spain's .es) and `mes.com`, respectively.

Premium domain names

In the business of marketing domain names, "premium" domain names are often valuable, and have particular characteristics. For example, the names are short and memorable, or may contain words that are regularly searched on search engines, or keywords that help the name gain a higher ranking on search engines. They may contain generic words, so the word has more than one meaning, and they may

contain common typos.

Resale of domain names

The business of resale of previously registered domain names is known as the "domain aftermarket".

Various factors influence the perceived value or market value of a domain name. They include 1) the natural or "organic" traffic that can be attributed to web surfers typing in a domain name in their web browser as opposed to doing a search for the site through a search engine. 2) Branding Opportunity. The ability to have a term recognized and easily recalled as a brand for a company or entity. 3) Re-sale value. The ability to spot trends and predict the value of a name based on its length (short is preferred), clarity, and commercial use. The word loan is far more valuable than the word sunshine.

Generic domain names have sprung up in the last decade. Certain domains, especially those related to business, gambling, pornography, and other commercially lucrative fields of digital world trade have become very much in demand to corporations and entrepreneurs due to their importance in attracting clients.

The most expensive public sale of an Internet domain name to date, according to DNJournal, is porn.com which was sold in 2007 for $9.5 million cash.

There are disputes about the high values of domain names claimed and the actual cash prices of many sales such as Business.com. Another high-priced domain name, sex.com, was stolen from its rightful owner by means of a forged transfer instruction via fax. During the height of the dot-com era, the domain was earning millions of dollars per month in advertising revenue from the large influx of visitors that arrived daily. The sex.com sale may have never been final as the domain is still with the previous owner. Also, that sale was not just a domain but an income stream, a web site, a domain name with customers and advertisers, etc. Two long-running U.S. lawsuits resulted, one against the thief and one against the domain registrar VeriSign [1]. In one of the cases, *Kremen v. Network Solutions*, the court found in favor of the plaintiff, leading to an unprecedented ruling that classified domain names as property, granting them the same legal protections. In 1999, Microsoft traded the name Bob.com with internet entrepreneur Bob Kerstein for the name Windows2000.com which was the name of their new operating system. [2]

One of the reasons for the value of domain names is that even without advertising or marketing, they attract clients seeking services and products who simply type in the generic name. This is known as Direct Navigation or Type-in Traffic. Furthermore, generic domain names such as movies.com (now owned by Disney) or Books.com (now owned by Barnes & Noble) are extremely easy for potential customers to remember, increasing the probability that they become repeat customers or regular clients. In the case of Movies.com, Disney has built a stand-alone portal featuring branded content. More and more large brands are beginning to employ a more comprehensive domain strategy featuring a portfolio of thousands of domains, rather than just one or two.

Although the current domain market is nowhere as strong as it was during the dot-com heyday, it remains strong and is currently experiencing solid growth again. [3] Annually tens of millions of dollars change hands in connection with the resale of domains. Large numbers of registered domain names lapse and are deleted each year. On average, more than 25,000 domain names drop (are deleted) every day.

It is important to remember that a domain (name, address) must be valued separately from the website (content, revenue) that it is used for. The high prices have usually been paid for the revenue that was generated from the website at the domain's address (URL.). The intrinsic value of a domain is the registration fee. It is difficult to appraise a current market value for a domain. The Fair Market Value of a domain can be anything from nearly nothing to millions of dollars. Factors involved may include previous sales data of similar domains, however a single letter difference can completely alter the value. The value of the domain (or any sum resp. division etc.) are usually added to the current or expected revenue from the web content (advertising, sales, etc.). The price of a domain (name + ext.) should not be confused with that of a website (content + revenue).

An estimate by an appraiser is always the addition of what they would like a domain to be worth together with the effective/expected/desired revenue from the web content. Some people put value on the length of the SLD (name) and other people prefer description capability, but the shorter an SLD is, the less descriptive it can be. Also, if short is crucial, then the TLD (extension) should be short too. It is less realistic to get a domain like LL.travel or LL.mobi than a domain travel. LL or mobi. LL. This illustrates the relativity of domain value estimation. It is safe to say that the revenue of web (content) can be easily stated, but that the value of a domain (SLD.TLD aka name.ext) is a matter of opinion and preference. In the end, however, any sale depends on the expectations of the domain seller and the domain buyer.

A webmaster creating a new web site either buys the domain name directly from a domain name registrar, or indirectly from a domain name registrar through a domainer. People who buy and sell domain names are known as domainers. People who sell value estimation services are known as appraisers.

Domain aftermarket prices and trends

Domain name sales occurring in the aftermarket are frequently submitted to the DN journal. The sales are listed weekly and include the top aftermarket resellers which include but are not limited to Sedo, Traffic (auctions), Afternic, NameJet, Moniker and private sales.

To date, and according to Guinness World Records and MSNBC, the most expensive domain name sales on record as of 2004 were[3]:

- Business.com for $7.5 million in December 1999
- AsSeenOnTv.com for $5.1 million in January 2000
- Altavista.com for $3.3 million in August 1998
- Wine.com for $2.9 million in September 1999
- CreditCards.com for $2.75 million in July 2004
- Autos.com for $2.2 million in December 1999

The week ending January 27, 2008, DNJournal reported that CNN, a cable news channel purchased iReport.com for $750,000. This signifies another turning point in domain name sales. This name has neither organic traffic, nor does it have a dictionary term alone. Instead it is a highly brandable domain name utilizing the second most popular prefix for a "dictionary" and commercial word.

Popular domain prefixes - "E" and "I"

In addition to a domain placing value on the shortness of the word, ease in spelling, commercial appeal, and organic capacity to generate natural traffic, today's domain names are being valued for the branding potential. The domain name sale iReport although not an organic or dictionary term alone, is actually preferred as a highly brandable term, in that it is has a popular pre-fix "i" which indicates the "report" to be online.

The prefixes and dashes between words were once considered second, but now due to brandability, if the term is a commercial term, a prefix is often preferred. Example eLoans markets with an e to indicate to its potential customers that a loan may be obtained online.

The two primary prefixes are "E", for electronic, and "I", for Internet. Both indicate the word or phrase to be accessible online. Because of that, in terms of branding, an i or e combined with a commercial term are highly desirable. In domain sales typically an **e** has been preferred, and **i** slightly less in terms of demand. eBrooklyn sold for approximately $2500 whereas once it would have been available to register at the price of a domain name (which ranges from $8 to $30 us dollars depending on the registrar). The rapidly increasing use of prefixes in conjunction with main dictionary and or commercial terms is here and for some predominantly internet based companies, or high technology, high profile companies, the prefix is now preferred.

One of the details that make a domain with a prefix more valuable for a brand, is the ability to simply promote the name without the use of ".com" in the promotion. If a domain owner had report.com he would be forced to use the .com to indicate it was on the net at that address, however a domain name with a one letter prefix does not need to use the ".com".

Someone could promote "iReport" as a brand, and assuming it was a world class brand, visitors would know they could find it at "iReport.com without seeing the .com. However if it was a .net, it would be wise to state iReport.net. This option to simply state the name of the company or entity is particularly valuable in that it is brief and clear in indicating that a report can be either made or found on the "i"nternet.

eLoans similarly does not have to state "eLoans.com". eLoans, in the minds of most is clearly an online entity offering electronic loan applications.

Some alternative domains that avoid the use of ".com" in their promotion are "WebMD" as the word web as a prefix suffice to indicate the information is online and likely at a .com extension.

Branding with a domain name

Brands are greatly affected by the ability of the company to obtain the matching domain name. If a company builds a brand around a name to which it does not own the domain name, it can end up directing traffic to another domain owner's site. If it is a competitor, this would be a problem.

Today's advertising development of a great brand is strictly confined to the availability to synchronize the brand with a domain name. Any confusion might result in a competitor gaining valuable internet traffic and possible customers.

Domain name confusion

Intercapping is often used to clarify the meaning of a domain name. However, DNS is case-insensitive, and some names may be misinterpreted when converted to lowercase. For example: Who Represents, a database of artists and agents, chose `whorepresents.com`; a therapists' network thought `therapistfinder.com` looked good; and another website operating as of August 2007, `cummingfirst.com`, website of the Cumming First United Church in Cumming, GA and `powergenitalia.com`, a website for an Italian Power Generator company. In such situations, the proper wording can be clarified by use of hyphens. For instance, Experts Exchange, the programmers' site, for a long time used `expertsexchange.com`, but ultimately changed the name to `experts-exchange.com`.

Leo Stoller threatened to sue the owners of StealThisEmail.com on the basis that, when read as `stealthisemail.com`, it infringed on claimed trademark rights to the word "stealth". [4] There is no word mark for "stealth" in the USPTO trademark database and Leo Stoller's trademarks on this term were canceled.

References

1. ^ Sex.com Settles With VeriSign (http://www.wired.com/news/business/0,1367,63142,00.html)
2. ^ Windows2000.com owner sells domain to Microsoft (http://www.theregister.com/1999/11/11/windows2000_com_owner_sells_domain/)
3. ^ *a b* Domain name sells for $2.75 million (http://www.msnbc.msn.com/id/5467584)
4. ^ Steal This v. Stealth Is: Community Technology Collective Bullied Over Misreading of URL (http://www.interactivist.net/stealth/pressrelease)

See also

- Domaining
- Domain hack
- Domain hijacking
- Domain name warehousing
- Domain tasting
- Fully qualified domain name
- Geodomain
- ICANN
- Internationalized domain name
- Name generator
- Public domain issues of domain names
- Uniform Resource Locator
- Web page
- Web site
- World Wide Web

External links

- IANA generic TLD (http://www.iana.org/gtld/gtld.htm)
- IANA Two letter Country Code TLD (http://www.iana.org/root-whois/)
- ICANN (http://www.icann.org/) - Internet Corporation for Assigned Names and Numbers.
- Internic.net (http://www.internic.net/), public information regarding Internet domain name registration services.

- RFC 1034 (http://tools.ietf.org/html/rfc1034), Domain Names — Concepts and Facilities, an Internet Protocol Standard.
- RFC 1035 (http://tools.ietf.org/html/rfc1035), Domain Names — Implementation and Specification, an Internet Protocol Standard.
- UDRP (http://www.icann.org/udrp/udrp.htm), Uniform Domain-Name Dispute-Resolution Policy.

Bi-directional text

From Wikipedia, the free encyclopedia

Bi-directional text is used as some writing systems of the world, notably the Arabic (including variants such as Nasta'liq), Persian and Hebrew scripts, are written in a form known as right-to-left (RTL), in which writing begins at the right-hand side of a page and concludes at the left-hand side. This is different from the left-to-right (LTR) direction used by most languages in the world. When LTR text is mixed with RTL in the same paragraph, each type of text should be written in its own direction, which is known as *bi-directional text*. This can get rather complex when multiple levels of quotation are used.

Many computer programs fail to display bi-directional text correctly. For example, the Hebrew name Sarah (שרה) should be spelled shin (ש) resh (ר) heh (ה) from right to left. Some Web browsers may display the Hebrew text in this article in the opposite direction.

Unicode
Character encodings
- Comparison
- UTF-7, UTF-1
- UTF-8, CESU-8
- UTF-16/UCS-2
- UTF-32/UCS-4
- UTF-EBCDIC
- SCSU, BOCU-1
- Punycode (IDN)
- GB 18030

UCS
Mapping
Bi-directional text
BOM
Han unification
Unicode and HTML
Unicode and E-mail
Unicode typefaces

Contents

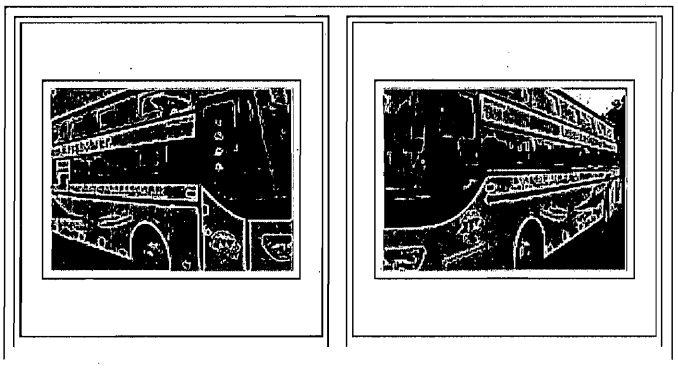
- 1 Languages using bi-directional text
- 2 Unicode support
- 3 See also
- 4 References
- 5 External links

Languages using bi-directional text

There are very few scripts that can be written in either direction.

Such was the case with Egyptian hieroglyphics, where the signs had a distinct "head" that faced the beginning of a line and "tail" that faced the end.

Chinese characters can also be written in either direction, especially in signs (but the orientation of the individual characters is never changed). This can often be seen on tour buses in China, where the company name customarily runs from the front of the vehicle to its rear - that is, from right to left on the right side of the bus, and from left to right on the left side of the bus.



The right side (text runs from right to left)	The left side (text runs from left to right)

Another variety of writing style, called *boustrophedon,* was used in some ancient Greek inscriptions, Tuareg, and Hungarian runes. This method of writing alternates direction, and usually reverses the individual characters, on each successive line.

Unicode support

Bidirectional script support is the capability of a computer system to correctly display bi-directional text. The term is often shortened to the jargon term **BiDi** or **bidi**.

Early computer installations were designed only to support a single writing system, typically for left-to-right scripts based on the Latin alphabet only. Adding new character sets and character encodings enabled a number of other left-to-right scripts to be supported, but did not easily support right-to-left scripts such as Arabic or Hebrew, and mixing the two was not practical. It is possible to simply flip the left-to-right display order to a right-to-left display order, but doing this sacrifices the ability to correctly display left-to-right scripts. With bidirectional script support, it is possible to mix scripts from different scripts on the same page, regardless of writing direction.

In particular, the Unicode standard provides foundations for complete BiDi support, with detailed rules as to how mixtures of left-to-right and right-to-left scripts are to be encoded and displayed.

In Unicode encoding, all non-punctuation characters are stored in writing order. This means that the writing direction of characters is stored within the characters. If this is the case, the character is called "strong". Punctuation characters however, can appear in both LTR and RTL languages. They are called "weak" characters because they do not contain any directional information. So it is up to the software to decide in which direction these "weak" characters will be placed. Sometimes (in mixed-directions text) this leads to display errors, caused by the bidi-algorithm that runs through the text and identifies LTR and RTL strong characters and assigns a direction to weak characters, according to the algorithm's rules.

In the algorithm, each sequence of concatenated strong characters is called a "run". A weak character that is located between two strong characters with the same orientation will inherit their orientation. A weak character that is located between two strong characters with a different writing direction, will inherit the main context's writing direction (in an LTR document the character will become LTR, in an RTL document, it will become RTL). If a "weak" character is followed by another "weak" character, the algorithm will look at the first neighbouring "strong" character. Sometimes this leads to unintentional display errors. To correct or prevent these errors, you can use "pseudo-strong" characters. These Unicode control characters are called "marks". The mark (U+200E LTR or U+200F RTL) is to be inserted into a location to make an enclosed weak character inherit its writing direction.

For example, to have the trademark symbol ™ (™; U+8482) for an English name brand (LTR) in an Arabic (RTL) passage display correctly, you need to add an LTR mark after the trademark symbol if the symbol is not followed by LTR text. This is because if you do not add the LTR mark, the weak character ™ will be neighboured by a strong LTR character and a strong RTL character. Hence, in an RTL context, it will be considered to be RTL, and displayed in an incorrect order.

See also

- Internationalization and localization
- Horizontal and vertical writing in East Asian scripts
- Writing system (section on directionality)

References

External links

- Unicode Standards Annex #9 (http://www.unicode.org/reports/tr9/) The Bidirectional Algorithm
- W3C guidelines on authoring techniques for bi-directional text (http://www.w3.org/TR/i18n-html-tech-bidi/) - includes examples and good explanations
- GNU FriBidi (http://fribidi.org/) A free implementation of the Unicode bidirectional algorithm
- ICU (http://www.icu-project.org/) International Components for Unicode contains an implementation of the bidirectional algorithm — along with other internationalization services
- UCData: "Pretty Good Bidi Algorithm Library" (http://crl.nmsu.edu/~mleisher/ucdata.html) A small and fast bidirectional reordering algorithm that works pretty good, but not necessarily compliant to the Unicode algorithm
- Bidirectional Scripts in Desktop Software (http://bidi.info/) Working group for supporting BiDi in Free Software. Contains several links to readings and implementation regarding BiDi in computer systems.
- Another Wiki about BiDi (http://mac.plonter.co.il/plonwiki/BidiWiki)
- Bidirectional text - Examples and practical advice (http://freenet-homepage.de/prilop/bidirectional-text.html)
- .Net BiDi Implementation (http://nbidi.sf.net/)

Retrieved from "http://en.wikipedia.org/wiki/Bi-directional_text"
Categories: Character encoding | Unicode algorithms | Unicode
Hidden categories: Articles lacking sources from September 2008 | All articles lacking sources