Localizing to right-to-left languages: Main Issues and Best Practices

Murhaf Hossari
University College Dublin
Murhaf.hossari@gmail.com
Outline

- Right-to-Left Languages
- Localizing to right-to-left languages
  - Right-to-left layout
  - Text justification
  - Directionality support
    - Unicode Bidirectional Algorithm
    - Issues and areas of improvements
- Numbers
- Best Practices
- Summary
Outline

- Right-to-Left Languages
  - Localizing to right-to-left languages
    - Right-to-left layout
    - Text justification
    - Directionality support
      - Unicode Bidirectional Algorithm
      - Issues and areas of improvements
  - Numbers
- Best Practices
- Summary
Right-to-left languages

Story

"Welcome aboard, Peterson."
Outline

- Right-to-Left Languages
- Localizing to right-to-left languages
  - Right-to-left layout
  - Text justification
  - Directionality support
    - Unicode Bidirectional Algorithm
    - Issues and areas of improvements
  - Numbers
- Best Practices
- Summary
Localizing to RTL Languages

RTL Layout

[Image of a Mac OS X Sharing settings window with a check for Bluetooth Sharing]

- Computer Name: Murha's MacBook Pro
- Computers on your local network can access your computer at: Murha's-MacBook-Pro.local
- Bluetooth Sharing: On
- Use Bluetooth Sharing preferences and set up your computer to share files with other Bluetooth enabled computers and devices.
  - When receiving items: Ask What to Do
  - Folder for accepted items: Downloads
  - When other devices browse: Ask What to Do
  - Folder others can browse: Public

Click the lock to prevent further changes.
Localizing to RTL Languages

RTL Layout
Outline

- Right-to-Left Languages
  - Localizing to right-to-left languages
    - Right-to-left layout
    - Text justification
    - Directionality support
      - Unicode Bidirectional Algorithm
      - Issues and areas of improvements
  - Numbers
- Best Practices
- Summary
Localizing to RTL Languages

Text Justification

You can help Apple improve its products and user support by having your Mac automatically send Apple information from time to time about how it's working and how you use it.

- This information is collected anonymously in a form that does not personally identify you.
- Click on the question mark for more information on the types of data sent.
- To learn more about Apple's Privacy Policy, see: http://www.apple.com/privacy

Enable Location Services
Allow the apps below to determine your location.

Indicates an app that has requested your location within the last 24 hours.

Click the lock to make changes.
Localizing to RTL Languages

Text Justification

Apple

You can help Apple improve its products and user support by having your Mac automatically send Apple information from time to time about how it’s working and how you use it.

- This information is collected anonymously in a form that does not personally identify you.
- Click on the question mark for more information on the types of data sent.
- To learn more about Apple’s Privacy Policy, see: http://www.apple.com/privacy

Enable Location Services

Indicates an app that has requested your location within the last 24 hours.

Click the lock to make changes.
Outline

- Right-to-Left Languages
- Localizing to right-to-left languages
  - Right-to-left layout
  - Text justification
  - Directionality support
    - Unicode Bidirectional Algorithm
    - Issues and areas of improvements
- Numbers
- Best Practices
- Summary
Left-to-right text:
Doubt is a pain too lonely to know that faith is his twin brother.

Right-to-left text:
.rehtorb niwt sih si htiap taht wonk ot ylenol oot niap a si tbuoD

Bidirectional text:
Doubt is a pain too lonely to know rehtorb niwt sih si htiap taht.
Directionality Support
Unicode Bidirectional Algorithm (UBA)

• Text is always entered in the same way
• Different languages have different character flow
• Reorder the characters and define their directionality properties (R, L, N...etc)
• Set of rules to produce the final correct visual representation
**Directionality Support**
Unicode Bidirectional Algorithm

**Example**

```
<table>
<thead>
<tr>
<th>RRRRRR</th>
<th>WS</th>
<th>RRRRRRR</th>
<th>WS</th>
<th>RR</th>
<th>WS</th>
<th>RRRRRRR</th>
<th>WS</th>
<th>N</th>
<th>LLLL</th>
<th>WS</th>
<th>LLLLLL</th>
<th>WS</th>
<th>LLLL</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRRRRR</td>
<td>R</td>
<td>RRRRRRR</td>
<td>R</td>
<td>RR</td>
<td>R</td>
<td>RRRRRRR</td>
<td>R</td>
<td>R</td>
<td>LLLL</td>
<td>L</td>
<td>LLLLLL</td>
<td>L</td>
<td>LLLL</td>
<td>R</td>
</tr>
<tr>
<td>111111</td>
<td>1</td>
<td>1111111</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>1111111</td>
<td>1</td>
<td>1</td>
<td>2222</td>
<td>2</td>
<td>222222</td>
<td>2</td>
<td>2222</td>
<td>1</td>
</tr>
</tbody>
</table>
```

CANNOT CONNECT TO SERVER "mail server name"
Directionality Support
Unicode Bidirectional Algorithm - Issues

• Correct layout for most of the cases
• Issues with some cases due to:
  o *Paragraph direction isn’t detected correctly*
  o *Complicated nesting of strings of different types*
  o *Strings with special nature such as part numbers, names ... etc*
  o *Ambiguous strings even for human eye*
Directionality Support

Unicode Bidirectional Algorithm - Issues

**Examples**

Incorrect Paragraph Direction (Base Direction)

شَرِكَةٌ أَمْريَكَيَّةٌ مَتَعَدَّدَةٌ الجِنْسِيَّاتِ Apple
Apple شركة أمريكية متعددة الجنسيات
Directionality Support
Unicode Bidirectional Algorithm - Issues

Examples

Weak Characters

Yahoo شركة

تطبيق doc

تطبيق doc

✔
Directionality Support
Unicode Bidirectional Algorithm - Issues

Examples

Weak Characters 2

VPN (PPTP)
VPN (PPTP) قطع الاتصال
VPN (PPTP) قطع الاتصال
Directionality Support

Unicode Bidirectional Algorithm – Directional Codes

- Invisible characters to help UBA showing the correct layout
- Explicit directional codes manually added to enforce the correct layout
Directionality Support
Unicode Bidirectional Algorithm – Directional Codes - Pitfalls

- Manually Added
- Not trivial to use
- Invisible!
- Need to check them at runtime
Directionality Support
Areas of improvements

• Finding patterns in the cases that usually need help
• Detect those patterns in the early stages by parsing the strings
• Deal with those cases in a special way by UBA
Outline

- Right-to-Left Languages
  - Localizing to right-to-left languages
    - Right-to-left layout
    - Text justification
    - Directionality support
      - Unicode Bidirectional Algorithm
      - Issues and areas of improvements
  - Numbers
- Best Practices
- Summary
Localizing to RTL Languages

Numbers

• Some scripts use different numbers
• Localizing numbers can be tricky
• Support for localized numerals in web is very little

(١، ٢، ٣، ٤، ٥، ٦، ٧، ٨، ٩، ٠)
(1, 2, 3, 4, 5, 6, 7, 8, 9, 0)
Outline

- Right-to-Left Languages
- Localizing to right-to-left languages
  - Right-to-left layout
  - Text justification
  - Directionality support
    - Unicode Bidirectional Algorithm
    - Issues and areas of improvements
- Numbers
- Best Practices
- Summary
Best Practices

- Support of RTL Layout
- Avoid composed strings
- Avoid weak and neutral characters
- Don’t enforce text direction
- Support of localized numbers
- Support of different locales
Outline

- Right-to-Left Languages
- Localizing to right-to-left languages
  - Right-to-left layout
  - Text justification
  - Directionality support
    - Unicode Bidirectional Algorithm
    - Issues and areas of improvements
- Numbers
- Best Practices
- Summary
• Need to pay extra attention when localizing to RTL languages
• Current support for RTL languages is very good
• Some improvements can be added to overcome persistent issues