A Progress Report on Encoding an Old Map in Japan

OHYA Kazushi

Tsurumi University

София, Oct. 28, 2005

Target Map Data

『改増伊豆七島全図附無人島八十与島相武房総海岸図』

"Revised and enlarged map of Seven Izu Islands with a map of eighty uninhabited islands and a coast map of Sagami, Musashi, Awa, and Simousa counties."

published in 1842(壬寅之春、天保13年春)

Why Old Maps?

- Old maps contain useful information.
- Old maps are sometimes difficult to treat and read.
- Old maps bring us too many challenging problems.

  e.g.

A List of Problems

1. How should we describe Characters beyond Unicode?
2. How should we describe land routes and sea routes?
3. Where should we describe information on image data?
4. How should we describe printing direction for text areas in maps?
5. We would like to regard each actual data as a module.
6. We need a global attribute for transliteration.
7. How should we describe places of map features in image data?
8. How should we describe the extent of time and space for map features in old maps?

Characters beyond Unicode

Markup constructs for representation of characters and glyphs in P5.

```
<charDesc>
 <char id="aenl">
  <charName>LATIN LETTER ENLARGED SMALL A</charName>
   <charProp>
    <localName>entity</localName>
    <value>aenl</value>
   </charProp>
   <mapping type="standardized">a</mapping>
 </char>
</charDesc>
```

```
<p><g ref="aenl">a</g></p>
```

substitution

Characters beyond Unicode

CASE 1: A digraph (one word with two words/morphemes)

1 `<expan abbr="` `">ども/domo/</expan>`

   →The element <g> cannot be used in attribute values.

2 `<abbr expan="ども/domo/">`
   `<g ref="somewhere">` ? `</g></abbr>`

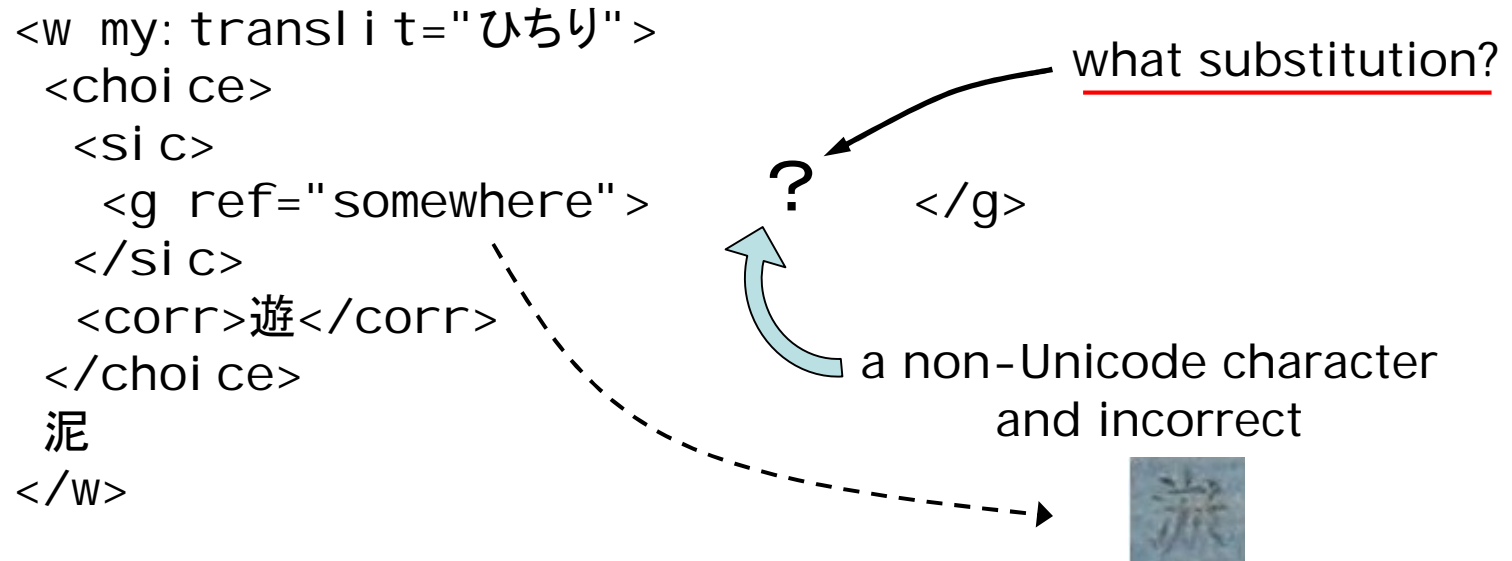   →What character should we use as a substitution to a Non-Unicode-character? "ども" or something else?

3 `<abbr expan="ども/domo/"><g ref="somewhere" /></abbr>`

Characters beyond Unicode 3/3

CASE 2: A word comprised of two characters is
       equivalent to a group of three characters.
       e.g. "遊泥" = "ひちり/hitiri/"

```
<w my:translit="ひちり">
  <choice>
    <sic>
      <g ref="somewhere">      ?      </g>
    </sic>
    <corr>遊</corr>
  </choice>
  泥
</w>
```

what substitution?

a non-Unicode character
and incorrect

→The character on the map is beyond Unicode, and incorrect.
What should we describe as contents in the element <g>?

An Overview of our Project

Objectives

– Encoding old maps in the Edo period(1603-1868).

– Examining how the present TEI schema can be applied to map-type contents.

– Making a guideline for Japanese bibliographers to encode their resources in TEI.

Plan

– Encoding a map in 2005.

– Defining a basic framework for old map encoding, and making the guideline in 2006.

– Encoding multiple maps, and starting the next phase in 2007.

Long-term Objectives

  We have a plan to make a DB system for historical map features. The DB system will use maps as interfaces for registering and viewing information.

  We would like to use a TEI specification as a schema for such kinds of systems, and make it more popular in a community of Japanese bibliographers.

Poster Presentation

11:00～12:30, Oct. 29.
Х О Т Е Л Р О Д И Н А

Please help us by your comments.

Б и х т е л и м и п о м о г н а л т
Д о у т р е!

OHYA Kazushi
Tsurumi University