YAX PIP: Yet another XML Plug-in for Protégé

John H. Gennari & Ping Lin

Biomedical & Health Informatics

University of Washington

Motivation

- Existing text format is *ad hoc*
- XML is hot!
- The semantic web is **really hot!**
- Existing XML plugins were broken or incomplete

Intuitive design choice

- XML tags based on class definitions
- Backend has two outputs:
  - XSD file with concept definitions
  - XML file with instances
- For example…

Simple ontology





SlotOver Protégé-2000 (D:\gennari\talks\ProtegeWorkshop03\Examples\SlotOver.pprj)

Project Window Help

Classes & Instances | Slots | Forms | Queries

Relationship

- :THING
  - :SYSTEM-CLASS
  - people
    - adults (2)
    - children (3)
  - Toys

Class: adults

Direct Instances

- Frederick
- Thomas

Frederick (type=adults, name=t...

Name: Frederick

Age: 30

Gender: Male

Simple instances in XML

```
<adult>
   <name>Frederick<\name>
   <age>30<\age>
   <gender>Male<\gender>
<\adult>
<adult>
   <name>Thomas<\name>
   . . .
```

Corresponding classes in XSD

```
<xs:element name="adult">
    <xs:complexType>
        <xs:element name="name" minOccurs=1 maxOccurs=1 />
        <xs:element name="age" minOccurs=0 maxOccurs=1 />
        <xs:element name="gender" type=  gendertype
                    minOccurs=0 maxOccurs=1 />
    <xs:complexType/>
```

```
<xs:simpletype name="gendertype">
    <xs:restriction base=xs:string>
        <xs:enumeration value="Male" />
        <xs:enumeration value="Female" />
</xs:restriction>
```

John goes to Stanford

- John: “Here’s our work so far”
- Ray: “But this is all wrong. Completely WRONG. REALLY WRONG!”
- John: “but…”
- Ray: “Throw it away and start over!”
- John: “Okay, boss.”

Problems w/ first design

- Differences between knowledge models:
  - Inheritance?
  - Slot type overrides?
  - xs:simpleType versus slot types

- What about Metaclasses?
  - Storage of own slot values?
  - Hard-wiring ‘documentation’, ‘constraints’, etc.

Designing for the Protégé KM

- Regardless of user concepts, a **single**, uniform xsd file!
- Schema based on the Protégé knowledge model
- Backend has only **one** output – an XML file w/ both classes & instances

  (plus the .pprj file, of course)

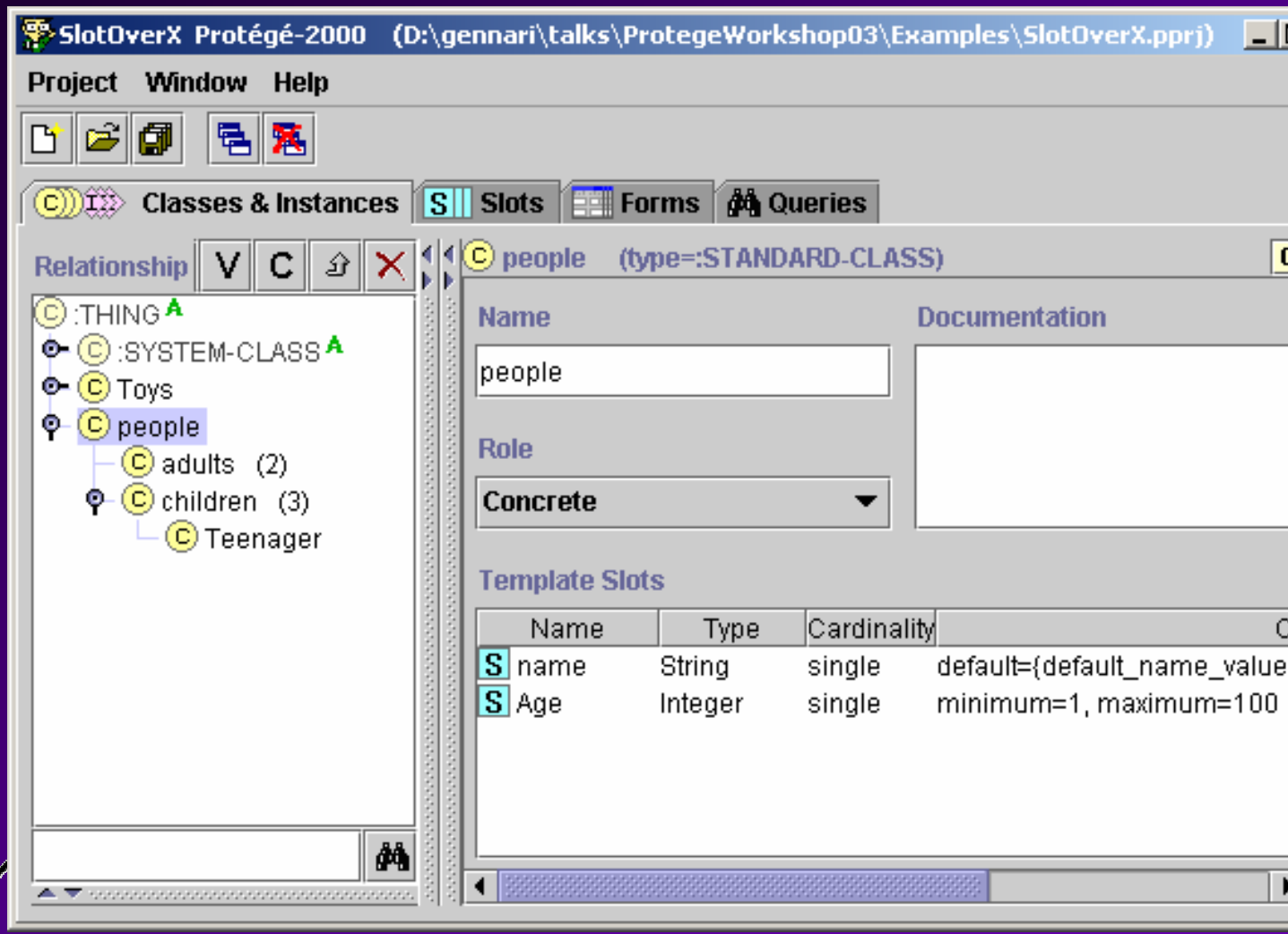
Examples, details

Demo of simple ontology,
corresponding XML file, and XSD file

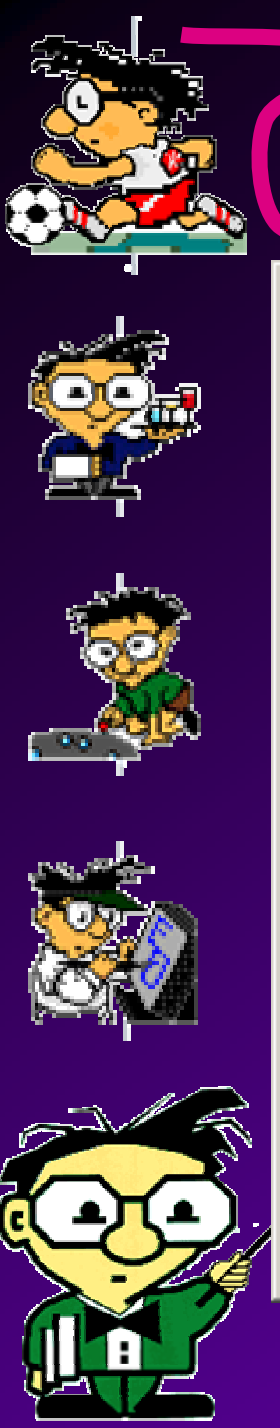
Metaclasses

- Old format: metaclasses are in both .pont and .pins file
- XML output:
  - Metaclasses are “just” another class, children of :Standard-Class
  - Instances of metaclasses are classes, with a non-standard :Direct-Type
  - (“instances” are really individuals)

Slot over-rides



Slot over-rides



SlotOverX Protégé-2000 (D:\gennari\talks\ProtegeWorkshop03\Examples\SlotOverX.pprj)

Project Window Help

Classes & Instances | Slots | Forms | Queries

Relationship

:THING
:SYSTEM-CLASS
Toys
people
adults (2)
children (3)
Teenager

adults (type=my-meta)

Name: adults

Documentation:

Role: Concrete

Template Slots

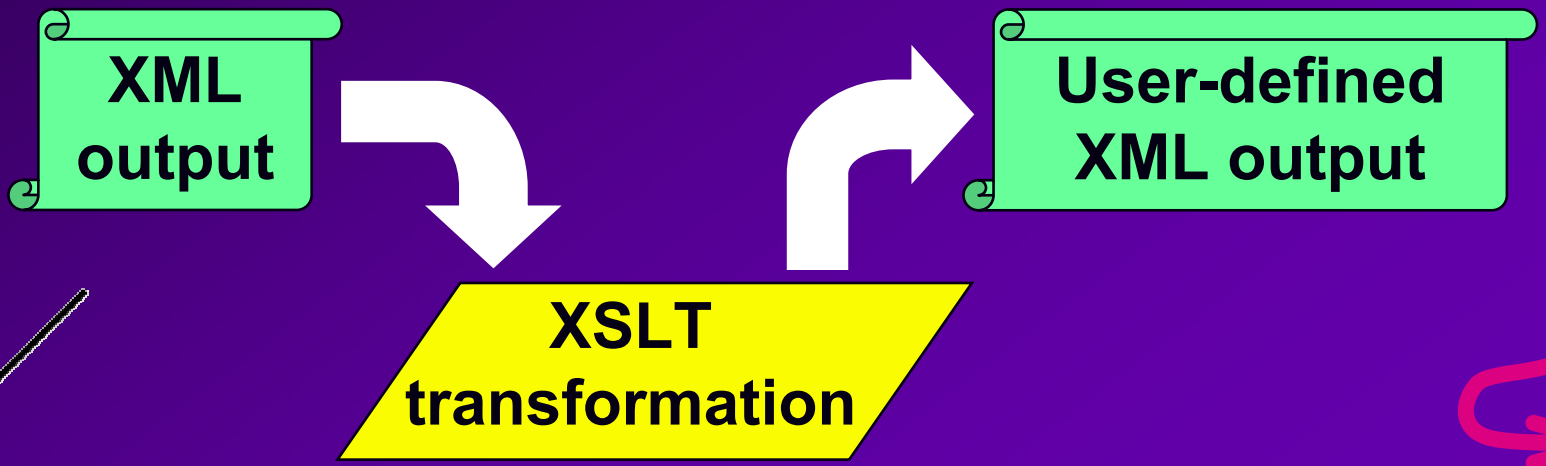
Name	Type	Cardinality	Oth
gender	Symbol	single	allowed-values={Male,Female}
name	String	single	default={default_name_value}
Age	Integer	single	minimum=18, maximum=100

Status

- Beta release in April ‘03
- Source code available on request
- Hand-off to Ray and Stanford team
  - Recent code updates to remove Xerces parser dependency (for Protégé 1.9)
  - Additional clean-up & implementation required for some details

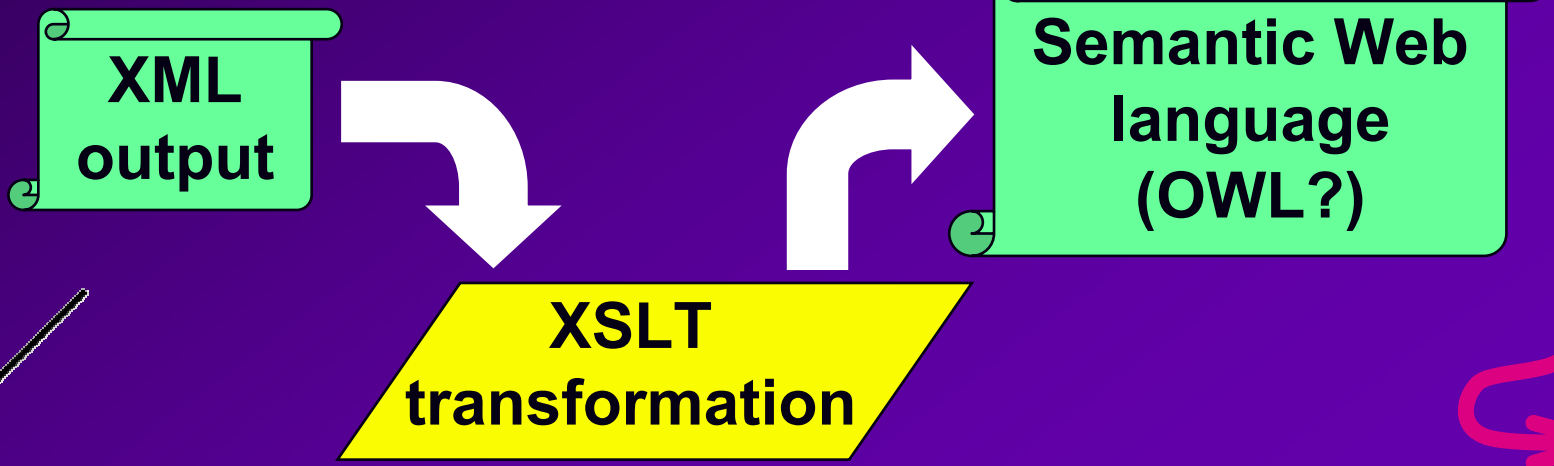
Future work

- XSLT for transforming XML
  - Use an easier-to-parse schema
  - Use an user-defined schema
  - Use a semantic web language schema



Future work

- XSLT for transforming XML
  - Use an easier-to-parse schema
  - Use an user-defined schema
  - Use a semantic web language schema



The big questions:

- Should XML be the default flat file storage format for Protégé?
  - For release 2.0?
  - Later?
- Is the XML plug-in just a stop-gap until Holger's OWL plug-in?

Thank You!