

Semi-structured Data

6 - XPath in XML Schema

XPath in XML Schema

- XML Schema uses XPath expressions in:
 - **key** elements
 - **kefref** elements
 - **unique** elements
- but only a subset of XPath is supported

Keys and References

```
<?xml version="1.0"?>
<company>
  <employees>
    <employee emp_id="e1">
      ...
    </employee>
    ...
  </employees>
  <managers>
    <manager mgr_id="e1">
      ...
    </manager>
    ...
  </managers>
</company>
```

The diagram illustrates an XML structure with two main sections: `<employees>` and `<managers>`. In the `<employees>` section, the `emp_id="e1"` attribute of the `<employee>` element is circled in blue. An arrow points from this circle to the text "key attribute". In the `<managers>` section, the `mgr_id="e1"` attribute of the `<manager>` element is also circled in blue. An arrow points from this circle to the text "foreign key (refers to emp_id)".

key attribute

foreign key
(refers to emp_id)

Keys and References

```
<xsd:element name="company" type="companyType">
```

```
  <xsd:key name="empKey">
```

```
    <xsd:selector xpath="employees/employee"/>
```

```
    <xsd:field xpath="@emp_id"/>
```

```
  </xsd:key>
```

```
  <xsd:keyref name="empRef" refer="empKey">
```

```
    <xsd:selector xpath="managers/manager"/>
```

```
    <xsd:field xpath="@mgr_id"/>
```

```
  </xsd:keyref>
```

```
</xsd:element>
```

```
<xsd:complexType name="companyType">
```

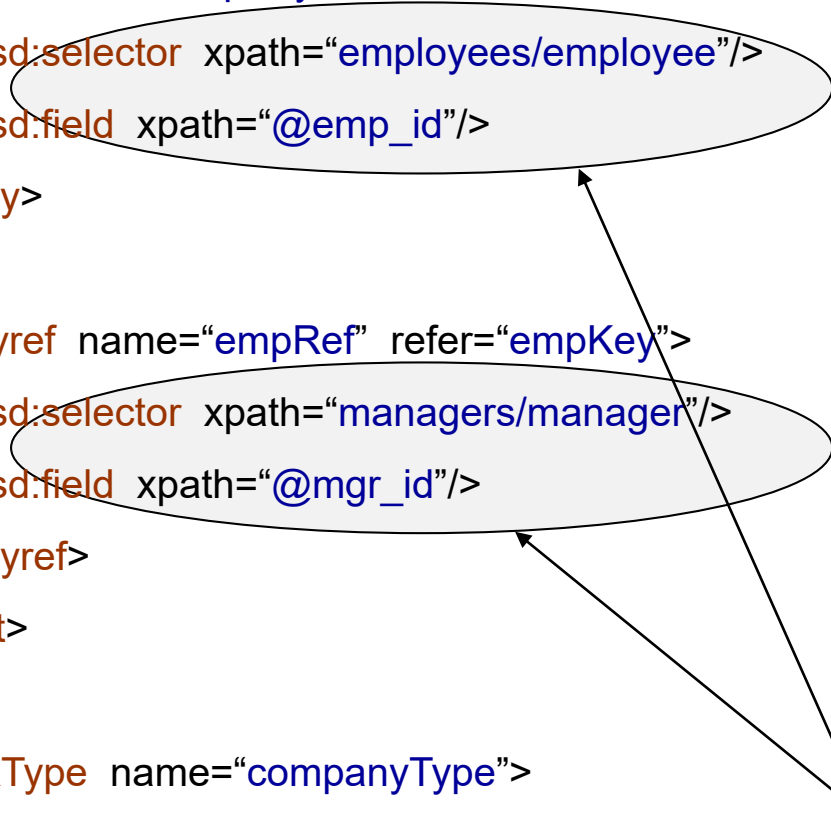
```
  ...
```

```
</xsd:complexType>
```

select emp_id

select mgr_id

XPath expressions



xsd:key vs. xsd:unique

- **xsd:key**: The field exists in all selected elements and its value is unique
- **xsd:unique**: If the field exist for a selected element the its value is unique

Assume not all employees have ids, but all managers have.

```
<xsd:element name="company" type="companyType">
  <xsd:unique name="empKey">
    <xsd:selector xpath="employees/employee"/>
    <xsd:field xpath="@emp_id"/>
  </xsd:unique>
  <xsd:keyref name="empRef" refer="empKey">
    <xsd:selector xpath="managers/manager"/>
    <xsd:field xpath="@mgr_id"/>
  </xsd:keyref>
</xsd:element>
<xsd:complexType name="companyType">
  ...
</xsd:complexType>
```

Restricted XPath in `xsd:selector`

- Only the child axis and at the begin of a path the descendant-or-self axis are allowed (with special syntax)
- Selects Elements (not attributes)
- The Good:

```
<xsd:selector xpath="employees/employee"/>
```



```
<xsd:selector xpath="employees/*"/>
```



```
<xsd:selector xpath="//employee"/>
```



```
<xsd:selector xpath="employees/employee | //manager"/>
```



Restricted XPath in `xsd:selector`

- Only the child axis and at the begin of a path the descendant-or-self axis are allowed (with special syntax)
- Selects Elements (not attributes)
- The Bad:

`<xsd:selector xpath="descendant::employee"/>` ✗

`<xsd:selector xpath="employees//employee"/>` ✗

`<xsd:selector xpath="parent::employee"/>` ✗

`<xsd:selector xpath="employees/child::nodes()"/>` ✗

`<xsd:selector xpath="employees/@employee"/>` ✗

Restricted XPath in `xsd:selector`

- Only the child axis and at the begin of a path the descendant-or-self axis are allowed (with special syntax)
- Selects Elements (not attributes)
- The Ugly:

```
<xsd:selector xpath="child::employees/child::employee"/>
```

Details in the W3C Recommendation :

<https://www.w3.org/TR/xmlschema11-1/#sec-c-selector-xpath>

Restricted XPath in `xsd:fields`

- Only the child axis and at the begin of a path the descendant-or-self axis are allowed (with special syntax)
- Allows attributes

`<xsd:field xpath="@emp_id"/>`



`<xsd:field xpath="employee"/>`



`<xsd:field xpath="//employee"/>`



`<xsd:field xpath="@employeeid | @managerid"/>`



Details in the W3C Recommendation:

<https://www.w3.org/TR/xmlschema11-1/#sec-c-fields-xpaths>

Multiple `xsd:field` elements in `xsd:key` and `xsd:unique`

We can use more than one attribute to define a key or uniqueness constraint.

- Two (or more) elements/attributes, but in an element at most one of them is present:

```
<xsd:field xpath="@employeeid | @managerid"/>
```

- Two (or more) elements/attributes being part of the key:

```
<xsd:key name="empKey">
```

```
  <xsd:selector xpath="employees/employee"/>
```

```
  <xsd:field xpath="group"/>
```

```
  <xsd:field xpath="@emp_id"/>
```

```
</xsd:key>
```

Multiple `xsd:field` elements in `xsd:keyref`

```
<xsd:key name="empKey">
```

```
  <xsd:selector xpath="employees/employee"/>
```

```
  <xsd:field xpath="group"/>
```

```
  <xsd:field xpath="@emp_id"/>
```

```
</xsd:key>
```

key

foreign key
(refers to empKey)

```
<xsd:keyref name="empRef" refer="empKey">
```

```
  <xsd:selector xpath="managers/manager"/>
```

```
  <xsd:field xpath="@group_id"/>
```

```
  <xsd:field xpath="@emp_id"/>
```

```
</xsd:key>
```