



Pràctica 8.1: XPath

Grup: Lishi i Karen

Objectius

Els objectius d'aquesta pràctica d'XPath se centren a desenvolupar una comprensió fonamental de la sintaxi i les funcionalitats bàsiques d'XPath. En primer lloc, caldrà familiaritzar-se amb la sintaxi bàsica de l'XPath i adquirir habilitats per seleccionar elements específics d'un document XML mitjançant camins absoluts i relatius. Un segon objectiu important és la pràctica en la manipulació de text i atributs, que inclourà l'extracció de contingut de text d'elements i l'accés als valors dels atributs.

Lliuraments

Aquesta pràctica està formada per la part 8.1 i la part 8.2. El lliurament es farà en una sola entrega una vegada ambdues parts estiguin realitzades. Per aquesta raó, de moment no cal lliurar aquesta pràctica.

Els resultats d'aquesta part de la pràctica s'hauran d'entregar en format PDF i l'entrega pot ser a través de GIT* o el moodle.

*** S'ha d'entregar l'enllaç del GIT al moodle.**

Guió

Exercici 1

Completa la taula explicant els resultats esperats per a cada una de les expressions XPath donades. En casos on la resposta sigui **un objecte sigui substancialment gran**, només cal que indiquis de quin element/s es tracta.

- Si dona **error** o no dona **cap resultat**, explica'n la raó

Exercici 2

Una vegada tinguis la taula de l'exercici 1 plena, avalua les expressions XPath fent servir el codi Python 3 que es troba a l'arxiu *xpath_evaluator.py*.

- Assegura't tenir instal·lada la llibreria 'lxml'. Per instal·lar-la, pots executar al terminal: `pip install lxml`

	Ruta XPath	Explica el resultat
1	<code>/llenguatges/llenguatge/nom</code>	De tots els nodes nom que pengen del node llenguatge i aquest penja del seu node pare llenguatge, surt l'etiqueta nom, el contingut text i/o altres etiquetes si en té <code><nom>Kotlin</nom></code> <code><nom>C++</nom></code> <code><nom>Java</nom></code> <code><nom>Python</nom></code> <code><nom>JavaScript</nom></code>
2	<code>/llenguatges/llenguatge/nom/node()</code>	De tots els nodes nom que pengen del node llenguatge i aquest penja del seu node pare llenguatge, surt el contingut de l'etiqueta nom, text i/o etiquetes si en té, però sense incloure l'etiqueta nom. Kotlin C++ Java Python JavaScript
3	<code>/llenguatges/llenguatge/paradigmes[@tipat]</code>	De tots els nodes paradigmes que tingui com a mínim un atribut de nom "tipat", que penja de llenguatge i aquest de llenguatges, surt tots els nodes fills d'aquests paradigmes, les etiquetes paradigma i el seu contingut. <code><paradigmes tipat="..."></code> <code><paradigma>...</paradigma></code> <code>...</code> <code></paradigmes></code> <code>...</code>
4	<code>/llenguatges/llenguatge/paradigmes[@tipat="false"]</code>	De tots els nodes paradigmes que tingui com a mínim un atribut de nom "tipat" amb el valor "false", que penja de llenguatge i aquest de llenguatges, surt tots els nodes fills d'aquests paradigmes, les etiquetes paradigma i el seu contingut. <code><paradigmes tipat="false"></code> <code><paradigma>...</paradigma></code> <code>...</code> <code></paradigmes></code> <code>...</code>
5	<code>nom</code>	Torna només l'etiqueta nom amb el seu contingut, perquè no té cap barra que indiqui que penja d'algun node pare i llavors retorna l'etiqueta més pròxim de

		l'arrel i en aquest cas <nom>Això no és un llenguatge...</nom>
6	/nom	No retorna res, perquè no es selecciona res i el node nom, no és l'arrel, per tant el xpath no existeix.
7	/nom/	Error, perquè no existeix el xpath ni cap node nom té subnodes fills com indica la barra.
8	//nom	Mostra tots els nodes noms i el seu contingut que penja de qualsevol lloc, incloen el que penja del node arrel. <nom>Kotlin</nom> <nom>C++</nom> <nom>Java</nom> <nom>Python</nom> <nom>JavaScript</nom> <nom>Això no és un llenguatge..</nom>
9	//nom/node()	Mostra de tots els nodes noms només seu contingut que penja de qualsevol lloc, incloen el que penja del node arrel. Kotlin C++ Java Python JavaScript Això no és un llenguatge..
10	//llenguatge/nom	Mostra el contingut de tots els nodes nom i l'etiqueta nom que penja dels nodes llenguatge que penja desde qualsevol lloc. <nom>Kotlin</nom> <nom>C++</nom> <nom>Java</nom> <nom>Python</nom> <nom>JavaScript</nom>
11	//llenguatge/nom/node()	Mostra el contingut de tots els nodes nom que penja dels nodes llenguatge que penja desde qualsevol lloc. Kotlin C++ Java Python JavaScript
12	//nom/node() //popularitat/node()	Mostra el contingut de tots els nodes noms que penja de qualsevol lloc i el contingut de tots els nodes popularitat que penja de qualsevol lloc, mentre es compleixi qualsevol de les dues condicions. Kotlin, Creixent, C++, Alta, Java, Alta, Python, Molt

		alta, JavaScript, Alta, Això no és un llenguatge...
13	//mode_execucio/*	No retorna res, perquè no penja res d'aquest node, no té nodes fills.
14	//llenguatge[2]	Donaria tot el contingut de llenguatge i l'etiqueta llenguatge amb l'atribut inclus i el seu valor, del segon llenguatge que conté llenguatges[2], en aquest cas tot del C++, la seva etiqueta nom, els sibilings nodes mode_execucio, popularitat, paradigmes, continguts i subnodes fills. <llenguatge fundacio="1983"> <nom>C++</nom> <mode_execucio>Compilació</mode_execucio> <popularitat>Alta</popularitat> <paradigmes tipat="true"> <paradigma>Procedimental</paradigma> ... </paradigmes> </llenguatge>
15	//llenguatge[last()-1]	El mateix que l'anterior, tot el llenguatge i el seu contingut que correspon al penúltim , que és python, a xml comença a comptar per 1 i last() -1 no és l'últim. <llenguatge fundacio="1991"> <nom>Python</nom> <mode_execucio>Interpretat</mode_execucio> <popularitat>Molt alta</popularitat> <paradigmes tipat="false"> <paradigma>Imperatiu</paradigma> ... </paradigmes> </llenguatge>
16	//llenguatge[@fundacio]	Mostra tots els nodes llenguatge que penja de qualsevol lloc amb el seu contingut i que tenen un atribut de valor fundació, i mostra tant els nodes fills, com el propi atribut i el seu valor.
17	//@fundacio	Selecciona i mostra els valors de tots atributs que té de nom "fundacio". 2011, 1983, 1995, 1991, 1995
18	//@fundacio[.>2000]	Selecciona i mostra els valors tots els atributs que pengen d'on sigui amb el nom "fundacio" en que el valor és superior a 2000. 2011
19	//@fundacio[.>2000]/nom	No retorna res perquè no penja directament de l'atribut fundació cap node nom, ni te subnodes.

20	<code>//@fundacio[.>2000]/../nom</code>	Selecciona i mostra l'etiqueta nom i el seu contingut que corresponen amb un atribut de "fundació" que penja desde qualsevol lloc amb valor superior a 2000. <nom>Kotlin</nom>
21	<code>//llenguatge[mode_execucio='Java Virtual Machine']/nom/text()</code>	El text que conté totes les etiquetes nom que pegen de llenguatge que pegen de qualsevol lloc i que tenen una etiqueta mode_execucio de valor Java Virtual Machine. Kotlin, Java
22	<code>//llenguatge[nom='Kotlin']/popularitat/node()</code>	Selecciona el contingut de tots els elements popularitat que pegen de llenguatge que penja desde qualsevol lloc i que tenen una etiqueta nom de valor Kotlin. Creixent
23	<code>//mode_execucio[.='Interpretat']/..</code>	Tots els elements llenguatge, el propi node amb el seu atribut i valor, el seu contingut subnodes fills i en que els hi pegen el node mode_execució amb valor de Interpretat que corresponen a Python i a JavaScript .
24	<code>//paradigmes[node()='Imperatiu']/../nom</code>	Dona de resultat el node nom i el seu contingut, que és sibling de tots els paradigmes que pegen de qualsevol lloc i els hi penja un node paradigma de valor Imperatiu. <nom>Kotlin</nom> <nom>C++</nom> <nom>Python</nom>
25	<code>//*</code>	Selecciona tots els elements i el seu contingut.

```
<?xml version="1.0" encoding="UTF-8" ?>
<llenguatges>
  <llenguatge fundacio="2011">
    <nom>Kotlin</nom>
    <mode_execucio>Java Virtual Machine</mode_execucio>
    <popularitat>Creixent</popularitat>
    <paradigmes tipat="true">
      <paradigma>Orientat a objectes</paradigma>
      <paradigma>Imperatiu</paradigma>
      <paradigma>Funcional</paradigma>
    </paradigmes>
  </llenguatge>
  <llenguatge fundacio="1983">
    <nom>C++</nom>
    <mode_execucio>Compilació</mode_execucio>
    <popularitat>Alta</popularitat>
    <paradigmes tipat="true">
      <paradigma>Procedimental</paradigma>
      <paradigma>Imperatiu</paradigma>
      <paradigma>Orientat a objectes</paradigma>
    </paradigmes>
  </llenguatge>
</llenguatges>
```

```

        <paradigma>Programació genèrica</paradigma>
    </paradigmes>
</llenguatge>
<llenguatge fundacio="1995">
    <nom>Java</nom>
    <mode_execucio>Java Virtual Machine</mode_execucio>
    <popularitat>Alta</popularitat>
    <paradigmes tipat="true">
        <paradigma>Orientat a objectes</paradigma>
        <paradigma>Basat en classes</paradigma>
    </paradigmes>
</llenguatge>
<llenguatge fundacio="1991">
    <nom>Python</nom>
    <mode_execucio>Interpretat</mode_execucio>
    <popularitat>Molt alta</popularitat>
    <paradigmes tipat="false">
        <paradigma>Imperatiu</paradigma>
        <paradigma>Funcional</paradigma>
    </paradigmes>
</llenguatge>
<llenguatge fundacio="1995">
    <nom>JavaScript</nom>
    <mode_execucio>Interpretat</mode_execucio>
    <popularitat>Alta</popularitat>
    <paradigmes tipat="false">
        <paradigma>Orientat a objectes</paradigma>
        <paradigma>Funcional</paradigma>
    </paradigmes>
</llenguatge>
<nom>Això no és un llenguatge de programació :(</nom>
</llenguatges>

```

Exercici 3

El fitxer **ods.xml** conté els **Objectius de Desenvolupament Sostenible (ODS)** en català. Els ODS són un conjunt de 17 objectius interconnectats adoptats per les Nacions Unides per abordar els reptes mundials, com la pobresa, la desigualtat, el canvi climàtic i la justícia social, amb l'objectiu de millorar la vida de les persones i protegir el planeta.

<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

Dissenyeu l'expressió XPath per accedir a la informació sol·licitada al fitxer *ods.xml*. Recordeu que heu de satisfer les condicions de l'enunciat i, tot i conèixer l'XML, heu d'extreure la informació sense afegir cap informació addicional. Per exemple, si se us demana seleccionar l'ODS titulat "Igualtat de Gènere", **heu de seleccionar-lo exclusivament fent servir aquesta informació**. L'exercici seria incorrecte si seleccioneu la informació fent servir el fet que és l'ods=5.

1. El text de dins de l'etiqueta <titol> de **tots els ODS** (Objectius de Desenvolupament Sostenible).

Fi de la Pobresa
Fam Zero
Salut i Benestar
Educació de Qualitat
Igualtat de Gènere
[...]

/ods/grup/objectiu/titol/text()

2. La descripció (text) de l'ODS titulat "**Igualtat de Gènere**". XXX

Promou la igualtat de gènere i empoderar totes les dones i nenes. Inclou objectius com eliminar la violència de gènere i garantir la participació igualitària en la presa de decisions.

//objectiu[titol="Igualtat de Gènere"]/descripcio/text()

3. Les **accions** (text) de l'ODS que té el títol de "**Fi de la Pobresa**". XXX

Implementar polítiques socials inclusives
Garantir protecció social per a tots

/ods/grup/objectiu[titol="Fi de la Pobresa"]/accions/accio/text()

4. El **títol** (text) de les ODS **13, 14, 15 i 16**. XXX

Pau, Justícia i Institucions Sòlides
Acció pel Clima
Vida Submarina
Vida d'Ecosistemes Terrestres

//objectiu[@ods > 13 and @ods<=16]/titol/node()

//objectiu[@ods="13" or @ods="14" or @ods="15" or @ods="16"]/titol/node()

5. El **títol** i la **descripció** (text) de l'ODS amb **ods="10"**.

Reducció de les Desigualtats
Busca reduir les bretxes entre països i dins d'ells. Inclou objectius com empoderar les persones marginades i promoure polítiques inclusives.

//objectiu[@ods=10]/titol/node() | //objectiu[@ods=10]/descripcio/node()

6. El llistat d'**accions**, en XML, dels ODS que pertanyen al grup de **tipus econòmic**.

<accio>Fomentar l'emprenedoria i la innovació</accio>
<accio>Reduir la bretxa salarial de gènere</accio>
[...]

//grup[@tipus="economic"]/objectiu/accions/accio

7. El text de dins de l'etiqueta **<titol>** del **5é objectiu** dins del grup d'ODS de tipus **"ambiental"**.

Vida Submarina

//grup[@tipus="ambiental"]/objectiu[5]/titol/text()

8. Dins el grup de tipus **social**, selecciona el **7è objectiu** i retorna la **2a acció**.

<accio>Garantir igualtat d'oportunitats per a tothom</accio>

- com que penja directament del node arrel, no ha calgut posar res davant de grup (//).
grup[@tipus="social"]/objectiu[7]/accions/accio[2]

9. El **tipus del grup** al qual pertany l'ODS que té l'acció **"Promoure l'ús d'energies renovables"**.

ambiental

//accions[accio="Promoure l'ús d'energies renovables"]/../../@tipus

10. **Busca les etiquetes amagades <start> i <end>** (fes ctrl+F) i **troba el camí** per començar a l'etiqueta **<start>** i arribar a imprimir el text que hi ha dins l'etiqueta **<end>**. La teva ruta ha de començar per **//start** i has d'intentar fer l'expressió XPath més breu possible.

<end>Busca promoure l'ocupació digna, el creixement econòmic inclusiu i la protecció social. Inclou objectius com reduir la informalitat laboral i garantir salaris justos.</end>

//start/../../../../../grup/objectiu/descripcio/end/text()

//start/../../../../../end/text()

11. Les **5 primeres accions** per les ODS de tipus **ambiental**. [opcional]

<accio>Desenvolupar infraestructures per a energies renovables</accio>
<accio>Promoure l'eficiència energètica a la indústria</accio>
<accio>Invertir en investigació i desenvolupament</accio>
<accio>Desenvolupar infraestructures de transport sostenible</accio>
<accio>Fomentar el reciclatge i la reutilització de productes</accio>

(//grup[@tipus="ambiental"]//accio)[position()<=5]

//grup[@tipus='ambiental']/objectiu[position()<=2]//accio[position()<=2] |
//grup[@tipus='ambiental']/objectiu[3]//accio[1]