

# Adatbázisok 1.

## XML lekérdezőnyelvek

### – 1. rész

XPath

XQuery

# Az XPath/XQuery adatmodell

- A relációk megfelelője ebben a környezetben a *tételek listája (sequence of items)*.
- Ez azt jelenti, hogy a bemenetet, a köztes lépések eredményeit és a végeredményt is tételek listájaként kezeljük.
- Egy *tétel (item)* lehet:
  1. Egyszerű érték, pl.: egész vagy sztring.
  2. Csomópont.

# A csomópontok fő típusai

1. A *dokumentum-csomópontok* (*document nodes*) a teljes dokumentumot reprezentálják.
2. *Elemcsomópontok*: a tagek és a közöttük lévő dokumentumrészlet.
3. *Attribútumok*: a nyitó tagekben szerepelnek és ott kapnak értéket.

# Öt tétel szekvenciája

```
10
"tíz"
10.0
<Számrendszer bázis = "8">
    <Digit>1</Digit>
    <Digit>2</Digit>
</Számrendszer>
@val="10"
```

- Az első három tétel egyszerű típusú (egész, szöveg, valós sz.)
- A negyedik elemcsomópont típus
- Az utolsó egy attribútum-csomópont típus

# Dokumentum-csomópont

- A `doc(URL)` vagy `document(URL)` parancs hatására jön létre.
- **Példa:** `doc(http://abc.com/sales.xml)`
- Minden XPath (és XQuery) lekérdezés hivatkozik egy dokumentum-csomópontra.
  - **Példa:** az XML Sémában szereplő XPath kifejezések arra a dokumentumra hivatkoznak, amelyre a séma éppen vonatkozik.

# A használt példa DTD-je

```
<!DOCTYPE teázók [  
  <!ELEMENT teázók (teázó*, tea*)>  
  <!ELEMENT teázó (ár+)>  
    <!ATTLIST teázó név ID #REQUIRED>  
  <!ELEMENT ár (#PCDATA)>  
    <!ATTLIST ár melyikTea IDREF #REQUIRED>  
  <!ELEMENT tea EMPTY>  
    <!ATTLIST tea név ID #REQUIRED>  
    <!ATTLIST tea árulja IDREFS #IMPLIED>  
>
```

# Példa dokumentum

<teázók>

Elem csomópont.

<teázó név = "JoeTeázója">  
    <ár melyikTea = "Brisk">2.50</ár>  
    <ár melyikTea = "Pyramid">3.00</ár>  
</teázó> ...

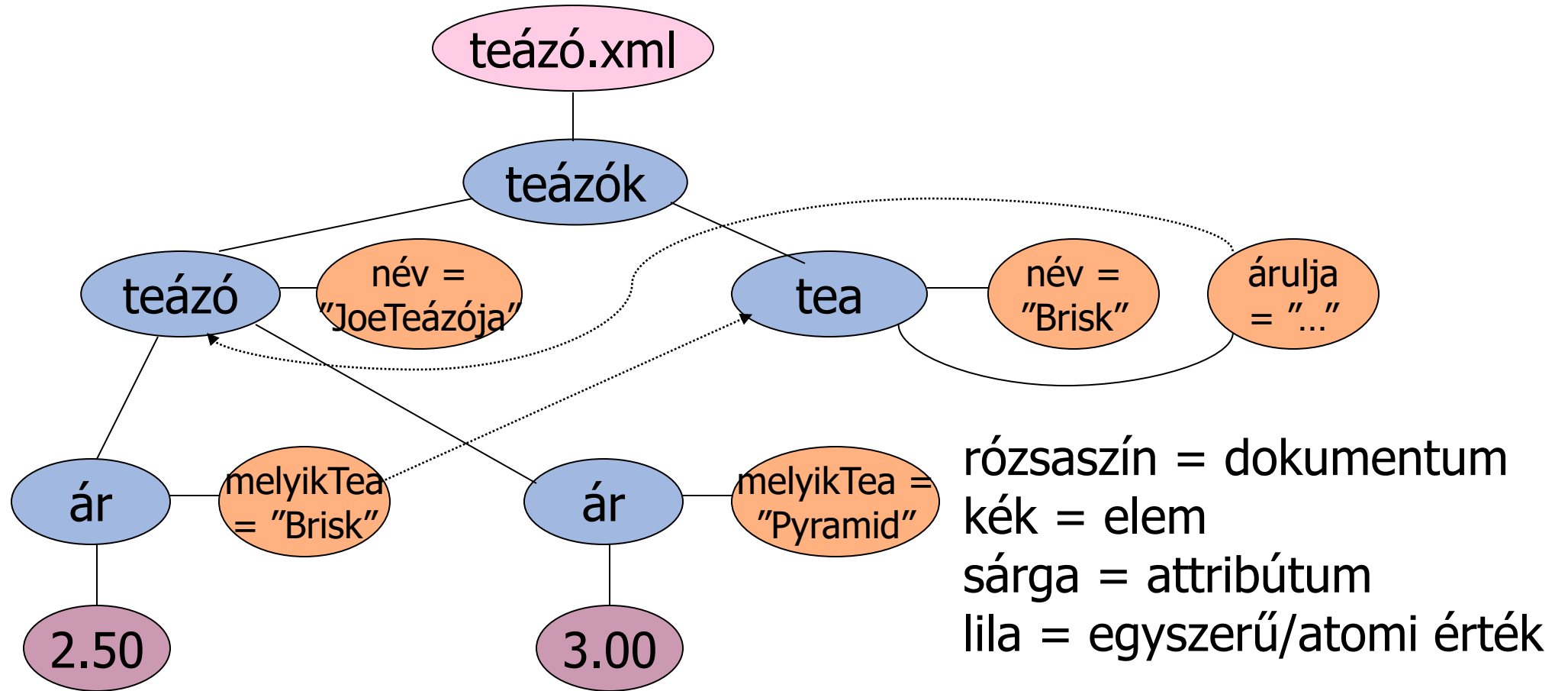
<tea név = "Brisk" árulja = "JoeTeázója  
SueTeázója ... "> ...

Attribútum csomópont

</teázók>

A dokumentum csomópont mindez, plusz az <? xml version... rész.

# A csomópontok típusa





# Utak (*paths*) az XML dokumentumban

- Az XPath segítségével az XML dokumentumokat járhatjuk be. Más szóval utakat adhatunk meg.
- Az utak mindig tételek egy listáját választják ki.

# Útkifejezések (*path expressions*)

- Egyszerű formájában az utak perjel (*slash*) és tagek sorozatából állnak. A kifejezés perjellel kezdődik.
  - Példa: /teázók/teázó/ár
- **Informális jelentés:** a dokumentum csomópontból kiindulva balról-jobbra haladva kövessük a tageket sorra.

# Útkifejezések kiértékelése

- Tegyük fel, hogy az első tag a gyökér.
  - Ennek a tagnek a feldolgozása a gyökér pontot tartalmazó, azaz egyelemű, sorozatot eredményezi.
- Tegyük fel, hogy rendelkezésünkre áll tételek egy sorozata és a következő tag  $X$ .
  - Minden elem pontot helyettesítsünk az  $X$  tagú alelemeivel a listában.

## Példa: /teázók

```
<teázók>
  <teázó név = "JoeTeázója">
    <ár melyikTea = "Brisk">2.50</ár>
    <ár melyikTea = "Pyramid">3.00</ár>
  </teázó> ...
  <tea név = "Brisk" árulja = "JoeTeázója
    SueTeázója ... "> ...
</teázók>
```

A teázók elem

## Példa: /teázók/teázó

<teázók>

<teázó név = "JoeTeázója">

<ár melyikTea = "Brisk">2.50</ár>

<ár melyikTea = "Pyramid">3.00</ár>

</teázó> ...

<tea név = "Brisk" árulja = "JoeTeázója

SueTeázója ... "> ...

</teázók>

A teázó elem, amit más teázó  
elemek is követhetnek.

## Példa: /teázók/teázó/ár

<teázók>

<teázó név = "JoeTeázója">

<ár melyikTea = "Brisk">2.50</ár>

<ár melyikTea = "Pyramid">3.00</ár>

</teázó> ...

<tea név = "Brisk" árulja = "JoeTeázója

SueTeázója ... "> ...

</teázók>

ár elemek, melyeket más  
teázók ár elemei követhetnek.

# Attribútumok az utakban

- Az attribútumokat a @ jel jelöli, ezután következik az attribútum neve.

Példa: /teázók/teázó/ár/@melyikTea

<teázók>

<teázó név = "JoeTeázója">

<ár melyikTea = "Brisk">2.50</ár>

<ár melyikTea = "Pyramid">3.00</ár>

</teázó> ...

<tea név = "Brisk" árulja = "JoeTeázója

SueTeázója ... "> ... A "Brisk" "Pyramid" értékek

</teázók>

belekerülnek az eredménybe.



## Ne felejtsük: tételek listája

- Eddig a tételek mindig elemek voltak.
- Ha egy útkifejezés attribútummal végződik, akkor a lista atomi típusú elemekből áll (az előző példában sztringekből).