Problem 267: Batch It Up

Difficulty: Hard

Author: Shelly Adamie, Fort Worth, Texas, United States

Originally Published: Code Quest 2025

Problem Background

At Lockheed Martin, many programs rely on being able to access data obtained from other systems. Some of these programs connect directly to the source databases, whereas others rely on manual export/import processes. Ensuring that your program can effectively "ingest" this external data is a key step towards developing a functional system. However, sometimes the amount of data you're trying to import can be a bit excessive. When that happens, it's important to break it down into smaller batches.

Problem Description

You're working with Lockheed Martin Aeronautics to develop a system to manage an inventory tracking system for parts for fighter aircraft. Your team has been provided with a series of XML files that contain information about the parts held at various Air Force bases, and you need to load this into your database. Unfortunately, there are some difficulties with this: your database can't read XML files directly, the format of the data is somewhat inconsistent, and the files are far larger than your database can handle at one time anyway. Your team needs to read the data from the XML files, process it into a legible format, and divide the data into smaller batches your database can handle.

Each part listed in the XML files has three main data points associated with it: its name, its serial number, and its part number. None of these data points are unique on their own, however the combination of part number and serial number will be unique for each part.

Within the XML files, parts may be listed in one of three formats. In each of the examples below, the part name is "Name", the part number is "PartNum", and the serial number is "SerialNum":

- An empty/self-closing element. Attributes may be presented in any order.
  ```
  <part name="Name" number="PartNum" serial="SerialNum" />
  ```
- A single element. The content of the element will be the part's serial number; the other attributes may be presented in any order.
  ```
  <part number="PartNum" name="Name">SerialNum</part>
  ```
- A set of nested elements, each listing one property of the part, and appearing in any order.
  ```
  <part>
  <name>Name</name>
  <serial>SerialNum</serial>
  ```

© 2025 Lockheed Martin Corporation. All Rights Reserved.

```
<number>PartNum</number>
</part>
```

Elements following the first two formats will be presented on a single line; the third format will be spread across five lines, as shown above.

Your solution will need to be able to process and store the information for these parts and report them in batches for addition to your database. Make sure that all records are reported.

Sample Input

The first line of your program's input, **received from the standard input channel**, will contain a positive integer representing the number of test cases. Each test case will include:

- A line containing a positive integer, representing the number of parts to include in each batch.
- A line containing the opening parent <parts> tag, representing the start of an XML file containing parts data.
- One or more lines containing the content of the XML file. These lines will contain valid XML elements matching the formats described above. Lines will not be indented in any way. Part numbers, serial numbers, and part names may contain any alphanumeric character and/or spaces.
- A line containing the closing parent </parts> tag, representing the end of the XML file.

```
1
3
<parts>
<part name="Thing" serial="123" number="ABC" />
<part serial="456" number="DEF" name="Other Thing" />
<part number="ABC" name="Thing">456</part>
<part number="123" name="Doohickey">ABC</part>
<part>
<number>JKL1</number>
<serial>42</serial>
<name>Wrench</name>
</part>
<part>
<name>Spanner</name>
<number>JKL1</number>
<serial>999</serial>
</part>
</parts>
```

Sample Output

For each test case, your program must print the part data, in the order it was listed in the original XML file, in batches not exceeding the batch size provided in the input. Use this format for each batch containing one or more parts:

© 2025 Lockheed Martin Corporation. All Rights Reserved.

- A line containing an integer with the batch number (the first batch in each test case should be batch 1; this resets to 1 with every new test case)
- A line for each part within the batch, containing:
    - The part number
    - A comma (,)
    - The serial number
    - A comma (,)
    - The part name

```
1
ABC,123,Thing
DEF,456,Other Thing
ABC,456,Thing
2
123,ABC,Doohickey
JKL1,42,Wrench
JKL1,999,Spanner
```

© 2025 Lockheed Martin Corporation. All Rights Reserved.