

Data Management and Artificial Intelligence Lab Class 9

Task 1 *Functional Dependencies* (50 minutes)

Please recall the notion of functional dependencies from the lecture. One way to formalize the entailment of a set of dependencies were the use of Armstrong's axioms:

1. Reflexivity: If $Y \subseteq X$, then $F| = X \rightarrow Y$.
2. Augmentation: If $F| = X \rightarrow Y$, then $F| = X, W \rightarrow Y, W$.
3. Transitivity: If $F| = X \rightarrow Y$ and $F| = Y \rightarrow Z$, then $F| = X \rightarrow Z$.

Note that the variables X, Y, Z above are possible sets of attributes, while W is a single attribute. For this task, please implement the above three axioms and finally compute the closure of an input set F_input , i.e., F_plus . In addition, print the overall procedure to obtain F_plus , which is essentially a sequence of axiom applications.

For an input set $F_input = \{A \rightarrow B, B \rightarrow C, C \rightarrow D, GD \rightarrow H\}$, the desired closure F_plus has a size of 1,815. A possible procedure to obtain F_plus can be *Ref - Aug - Tran - Aug - Tran - Aug - Aug*.

Hint: You may follow the steps below to solve this task:

1. Determine the datatypes used to represent FD and set of FD. You may consider a possible combination of list, tuple, set or dictionary.
2. Implement a recursive function *get_subsets* to generate all the subsets of an input set S . This function will be utilized in *Reflexivity*.
3. Implement the function *get_attribute_set* which receives the input set F_input and returns the universe set of attributes.
4. Implement the function *Reflexivity* which receives an FD set F and the attribute set obtained in subtask 3. Utilize the subset function in subtask 2 to perform reflexivity, and add the generated FDs to F to form an expanded set.
5. Implement the function *Augmentation* which receives an FD set F and the attribute set. Perform the augmentation operation to F and add the generated FDs to F . If there are new FDs added to F , update the boolean variable *change* to *True*. Finally, return F and *change*.
6. Implement the function *Transitivity* which receives an FD set F and the attribute set. Perform the transitivity operation to F and add the generated FDs to F . If there are new FDs added to F , update the boolean variable *change* to *True*. Finally, return F and *change*.

7. Finally, implement the main function *computeClosure* which receives the input FD set *F_input* and returns the desired closure *F_plus*. Think about how to organize the three axiom functions and how to implement the terminate condition of loop. In addition, print the procedure to obtain *F_plus*, which is essentially a sequence of axiom applications.

```
def get_subsets(S):
    #TODO
    return all_subsets

def get_attribute_set(F_input):
    #TODO
    return att_set

def Reflexivity(F, att_set):
    #TODO
    return F

def Augmentation(F, att_set):
    change=False
    #TODO
    return F, change

def Transitivity(F):
    change=False
    #TODO
    return F, change

def computeClosure(F_input):
    #TODO
    return F_plus

F_plus=computeClosure(F_input)
print(len(F_plus))
```

Task 2 *Translate XML files to csv files (25 minutes)*

1. Download the xml file with the name 'books collection.xml'. Some information of 12 books (including their ids, authors, titles ...) are saved in the file.
2. Use Python to translate the data in this XML file to a csv file, i.e., a table. The final results are shown as Table 3.

Task 3 *Seaborn on bike stations (20 minutes)*

Perform some statistical analysis on the bike stations in homework. For instance, check whether there is a correlation between the distance of a station's location from the city center of London and the number of bikes available. Visualize the results of your analysis using seaborn.

Table 3: Book collections

	author	title	genre	price	publish_date	description
bk101	Gambardella, Matthew	XML Developer's Guide	Computer	44.95	2000-10-01	An in-depth look ...
bk102	Ralls, Kim	Midnight Rain	Fantasy	5.95	2000-12-16	A former architect battles ...
bk103	Corets, Eva	Maeve Ascendant	Fantasy	5.95	2000-11-17	After the collapse of a ...
bk104	Corets, Eva	Oberon's Legacy	Fantasy	5.95	2001-03-10	In post-apocalypse ...
bk105	Corets, Eva	The Sundered Grail	Fantasy	5.95	2001-09-10	The two daughters ...
bk106	Randall, Cynthia	Lover Birds	Romance	4.95	2000-09-02	When Carla meets ...
bk107	Thurman, Paula	Splash Splash	Romance	4.95	2000-11-02	A deep sea ...
bk108	Knorr, Stefan	Creepy Crawlies	Horror	4.95	2000-12-06	An anthology of ...
bk109	Kress, Peter	Paradox Lost	Science Fiction	6.95	2000-11-02	After an inadvertant ...
bk110	O'Brien, Tim	Microsoft .NET: The Programming Bible	Computer	36.95	2000-12-09	Microsoft's .NET ...
bk111	O'Brien, Tim	MSXML3: A Comprehensive Guide	Computer	36.95	2000-12-01	The Microsoft MSXML3 ...
bk112	Galos, Mike	Visual Studio 7: A Comprehensive Guide	Computer	49.95	2001-04-16	Microsoft Visual Studio 7 ...