Finding localization in disordered system.

Jakub Wronowicz 245053

1 Functional relations:

To find functional relations we have to find what can we obtain from particular columns. Let's call particular columns as:

1. Country $\rightarrow$ A,
2. LANGUAGE $\rightarrow$ B,
3. area $\rightarrow$ C,
4. national_day $\rightarrow$ D,
5. country_code2 $\rightarrow$ E,
6. country_code3 $\rightarrow$ F,
7. Region $\rightarrow$ G,
8. Continent $\rightarrow$ H,
9. region_area $\rightarrow$ I,
10. percent_of_region $\rightarrow$ J,
11. YEAR $\rightarrow$ K,
12. population $\rightarrow$ L,
13. gdp $\rightarrow$ M

Now we can find our relations:

$$
\begin{aligned}
F = \{ & A \to ACDEFGHIJ, \\
& C \to ACDEFGHIJ, \\
& E \to ACDEFGHIJ, \\
& F \to ACDEFGHIJ, \\
& G \to GHI, \\
& I \to GI, \\
& AB \to ABCDEFGHIJ, \\
& BC \to ABCDEFGHIJ, \\
& BE \to ABCDEFGHIJ, \\
& BF \to ABCDEFGHIJ, \\
& AK \to ACDEFGHIJKLM, \\
& CK \to ACDEFGHIJKLM, \\
& EK \to ACDEFGHIJKLM, \\
& FK \to ACDEFGHIJKLM, \\
& AL \to ACDEFGHIJKLM, \\
& CL \to ACDEFGHIJKLM, \\
& EL \to ACDEFGHIJKLM, \\
& FL \to ACDEFGHIJKLM, \\
& AM \to ACDEFGHIJKLM, \\
& CM \to ACDEFGHIJKLM, \\
& EM \to ACDEFGHIJKLM, \\
& FM \to ACDEFGHIJKLM, \\
& ABK \to ABCDEFGHIJKLM, \\
& CBK \to ABCDEFGHIJKLM, \\
& EBK \to ABCDEFGHIJKLM, \\
& FBK \to ABCDEFGHIJKLM, \\
& ABL \to ABCDEFGHIJKLM, \\
& CBL \to ABCDEFGHIJKLM, \\
& EBL \to ABCDEFGHIJKLM, \\
& FBL \to ABCDEFGHIJKLM, \\
& ABM \to ABCDEFGHIJKLM, \\
& CBM \to ABCDEFGHIJKLM, \\
& EBM \to ABCDEFGHIJKLM, \\
& FBM \to ABCDEFGHIJKLM, \\
\}
\end{aligned}
\tag{1}
$$

Now we can find all candidate to be key. It is just these set of attributes which can give us information about all attributes. In our case it is anyone from following set: {ABK, CBK, EBK, FBK, ABL, CBL, EBL, FBL, ABM, CBM, EBM, FBM }. Now let us notice that we can divede attributes into two different category: key attributes, non-key attributes:

- key attributes: ABCEFKLM
- non-key attributes: DGHIJ

Before we will come to normalization we'll first minimize our functional relations F:

$$\begin{aligned} F = \{ & A \to CDEFHGIJ, \\ & C \to ADEFGHIJ, \\ & E \to ACDFGHIJ, \\ & F \to ACDEGHIJ, \\ & G \to HI, \\ & I \to G, \\ & AK \to LM, \\ & CK \to LM, \\ & EK \to LM, \\ & FK \to LM, \\ & AL \to KM, \\ & CL \to KM, \\ & EL \to KM, \\ & FL \to KM, \\ & AM \to KL, \\ & CM \to KL, \\ & EM \to KL, \\ & FM \to KL, \\ \} & \end{aligned} \tag{2}$$

Now from relation $A \to D$ we can notice that this relation break 2NF so this table is maximal 1NF normalized. Moreover we can notice that first four relations are breaking the 2NF, due to existing attribute D, G, I and J on the right hand side. So let's split it with them.

$$\begin{aligned} F_1 = \{ & A \to CDEFGHIJ, \\ & C \to ADEFGHIJ, \\ & E \to ACDFGHIJ, \\ & F \to ACDEGHIJ \} \end{aligned} \tag{3}$$

So $R_1 = ABCDEFGHIJ$ and candidate of key are A, C, E, F. In this case we can easly notice that it is 3NF so it's fully normalize. So let's back to second result of our split:

$$
\begin{aligned}
F_2 = \{ & A \to CEF, \\
& C \to AEF, \\
& E \to ACF, \\
& F \to ACE, \\
& G \to HI, \\
& I \to G, \\
& AK \to LM, \\
& CK \to LM, \\
& EK \to LM, \\
& FK \to LM, \\
& AL \to KM, \\
& CL \to KM, \\
& EL \to KM, \\
& FL \to KM, \\
& AM \to KL, \\
& CM \to KL, \\
& EM \to KL, \\
& FM \to KL, \\
\} &
\end{aligned}
\tag{4}
$$

The candidates for key are exactly the same as in case before split - {ABK, CBK, EBK, FBK, ABL, CBL, EBL, FBL, ABM, CBM, EBM, FBM }. R_2 = ABCEFKLM. So now F_2 are 2FN (because nor G or I are not subset of any of key). To obtain 3NF we have to split it with relations: $G \to HI$ and $I \to G$. So let's do it:

$$
\begin{aligned}
F_{2.1} = \{ & G \to HI, \\
& I \to G\}
\end{aligned}
\tag{5}
$$

In this case $R_{2.1}$ = GHI and candidate for key is equal G. Of course it's 3NF. Let's go to second case of this split:

$$
\begin{aligned}
F_{2.2} = \{ & A \to CEF, \\
& C \to AEF, \\
& E \to ACF, \\
& F \to ACE, \\
& AK \to LM, \\
& CK \to LM, \\
& EK \to LM, \\
& FK \to LM, \\
& AL \to KM, \\
& CL \to KM, \\
& EL \to KM, \\
& FL \to KM, \\
& AM \to KL, \\
& CM \to KL, \\
& EM \to KL, \\
& FM \to KL, \\
\} &
\end{aligned}
\tag{6}
$$

In this case $R_{2.2}$ = ABCEFJKLM and candidate for key is equal $\{$ABK, CBK, EBK, FBK, ABL, CBL, EBL, FBL, ABM, CBM, EBM, FBM $\}$. In this case we didn't obtain any member of non-key attribute so it have to be 3NF.