Normalization of the Database of Regional Division of the Relief

Zdena Dobešová

Department of Geoinformatics, Faculty of Science, Palacký University in Olomouc
Head of Department: Assoc. Prof. Dr. Vít Voženílek

KEY WORDS: Database, Normalization, Geomorphological units

Abstract

This article describes application of the normalization process on the database of the Regional Division of the Relief of the Czech Republic. An example of searching information in this geographical database is also provided.

1. Introduction

Relief of the Czech Republic is divided into a system of geomorphological units.
Jaromír Demek and his colleagues built up the Geographic Lexicon of the Czechoslovak Republic: Mountains and Lowlands (Zeměpisný lexikon ČSR: Hory a nížiny). The Lexicon provides hierarchical division of the geographic relief of the Czech Republic. Furthermore it unifies usage of geographical names for both professionals and the public. This work describes geomorphological units and a large amount of peaks. It contains simple maps with boundary lines and codes of the units and also the network of rivers and big settlements.
A disadvantage of the Lexicon is that the units are sorted alphabetically. Therefore information about certain area can only be found if you know the name of the particular unit. We can get better access to information by transferring the data to a database i.e. to Geographical Information System (GIS). There will be only area of the Czech Republic in the proposed information system. The main task is to transfer the data about regional division of the relief to the database. Process of normalization will be used when designing relational database model. The main advantages of normalization are elimination of redundancy, easy access to required information, unambiguous answer to a question and also data update on a single spot.

2. Normalization

First part of the process is design of the database model. This means that rules of the normal forms will be applied on the original data structure, which is not in a normal form. In most cases the first, second and the third normal forms are used. Other normal forms like the Boyce-Codd, fourth and the fifth are rarely used.

Normalizing the database means removing redundancies by the mechanism of the decomposition without loss of information.

First Normal Form

A relation is in the first normal form if all of its attributes are defined within scalar ranges of values (domains). This means further division of attributes is not possible.

One column cannot have more than one piece of information. The value of a column must not be a relation. [Riordan, 2000]

The criteria of scalar attributes are sometimes not met with values such as composed codes and flags. This is also the case of the Czech Republic geographical relief coding.

3. Geomorphological Division of the Czech Republic

In 1987 Jaromír Demek and his colleagues built up the Geographic Lexicon of the Czechoslovak Republic: Mountains and Lowlands. The lexicon provides hierarchical division of the geographic relief of the Czech Republic. Furthermore it unifies usage of geographical names for both professionals and the public. This work describes geomorphological units and a large amount of peaks.

Each geomorphological unit (hereafter the unit) has its numeric and alphabetic code (hereafter the code). The code is set up so that its structure determines belonging of the lower unit to the higher. This code instead of verbal description is used in the maps attached to the lexicon.

Geomorphological types are as follows: systém, subsystém, provincie, soustava, podsoustava, celek, podcelek a okrsek.

Example:

Hercynský	systém
Hercynská pohoří	subsystém

	Česká Vysočina	provincie
I	Šumavská soustava	soustava
IA	Českoleská podsoustava (oblast)	podsoustava
IA-1	Český les	celek
IA-1A	Čerchovský les	podcelek
IA-1A-a	Haltravská hornatina	okrsek
IA-1A-b	Nemanická vrchovina	
IA-1A-c	Ostrovská vrchovina	
IA-1B	Kateřinská kotlina	
IA-1C	Přimdský les	
IA-1C-a	Málkovská vrchovina	
IA-1C-b	Plešivecká vrchovina	
IA-1C-c	Havranská vrchovina	
IA-1C-d	Rozvadovská pahorkatina	

[Demek, 1987]

The value of the code is composed in the following way:

Types of geomorphological units	**Description**	**Range**
systém	without description	
subsystém	without description	
provincie	without description	
soustava (new subprovincie)	roman digit	I to X
podsoustava (new oblast)	capital letter	A to D
celek	arabic digit	1 to 8
podcelek	capital letter	A to H
okrsek	small letter	a to i

Tab. 1: Code of units

4. Database Model

The database is designed in the Microsoft Access 2000 CZ environment. Seven relations (tables) are used in the design of the first database model. Each relation corresponds to a geomorphological type of units. Therefore proposed relations will be **System**, **Subs_Prov**, **Soustavy**, **Podsoustavy**, **Celky**, **Podceleky** and **Okrsky**.

There will be attributes **kod**, **nazev** in the relations above. From the normalization point of view it can be seen that the relations do not correspond to the first normal form. The code attribute is not scalar. It is necessary to divide it into separate attributes. It is possible to leave the original code as a separate attribute, which can serve as a link to the map. It is necessary to

realize that the length of the code attribute is variable. Therefore it is divided in a various number of attributes in the particular relations. In the Okrsky relation it is divided into 5 separate attributes.

The resulting relations have the following contents:

System : Tabulka

	ID	nazev
▶	1	Hercynský
+	2	Alpsko-Himalájský
*	(é číslo)	

Fig. 1: Content of the System relation

The System relation consists of the attributes **ID** - primary key (automatic numbering data type), **Nazev**.

Subs-Prov : Tabulka

	ID-system	subsystem	provincie	nazev
▶ +	1	Hercynská pohoří	1	Česká vysočina
+	1	Epihercynské nížiny	2	Středoevropská nížina
+	2	Karpaty	3	Západní Karpaty
+	2	Panonská pánev	4	Západopanonská pánev

Fig. 2: Content of the Subs_Prov relation

There are both the subsystem and the province in the Subs_Prov relation. These two types of geomorphological units are together in one relation because the relation between them is 1:1 for the area of the Czech Republic. Numeric index for the province (1, 2, 3, 4) was added to the database model. The three highest levels do not have any code.

Soustavy : Tabulka

	provincie	soustava	nazev
▶ +	1	I	Šumavská soustava
+	1	II	Česko-moravská soustava
+	1	III	Krušnohorská soustava
+	1	IV	Krkonošsko-jesenická soustava
+	1	V	Poberounská soustava
+	1	VI	Česká tabule
+	2	VII	Středopolské nížiny
+	3	VIII	Vněkarpatské sníženiny
+	3	IX	Vnější západní Karpaty
+	4	X	Vídeňská pánev

Fig. 3: Content of the Soustavy relation

Podsoustavy : Tabulka

kod	soustava	podsoustava	nazev
IA	I	A	Českoleská podsoustava
IB	I	B	Šumavská hornatina
IIA	II	A	Středočeská pahorkatina
IIB	II	B	Jihočeské pánve
IIC	II	C	Českomoravská vrchovina
IID	II	D	Brněnská vrchovina
IIIA	III	A	Krušnohorská hornatina
IIIB	III	B	Podkrušnohorská podsoustava
IIIC	III	C	Karlovarská vrchovina
IVA	IV	A	Krkonošská podsoustava
IVB	IV	B	Orlická podsoustava
IVC	IV	C	Jesenická podsoustava

Fig. 4: Content of the Podsoustavy relation

Celky : Tabulka

kod	soustava	podsoustava	celek	nazev
IID-1	II	D	1	Boskovická brázda
IID-2	II	D	2	Bobravská vrchovina
IID-3	II	D	3	Drahanská vrchovina
IIIA-1	III	A	1	Smrčiny
IIIA-2	III	A	2	Krušné hory
IIIA-3	III	A	3	Děčínská vrchovina
IIIB-1	III	B	1	Chebská pánev
IIIB-2	III	B	2	Sokolská pánev
IIIB-3	III	B	3	Mostecká pánev
IIIB-4	III	B	4	Doupovské hory
IIIB-5	III	B	5	České středohoří

Fig. 5: Part of content of the Celky relation

Podcelky : Tabulka

kod	soustava	podsoustava	celek	podcelek	nazev
VIB-2A	VI	B	2	A	Středojizerská tabule
VIB-2B	VI	B	2	B	Dolnojizerská tabule
VIB-3A	VI	B	3	A	Nymburská kotlina
VIB-3B	VI	B	3	B	Čáslavská kotlina
VIB-3C	VI	B	3	C	Mělnická kotlina
VIB-3D	VI	B	3	D	Mrlinská tabule
VIB-3E	VI	B	3	E	Českobrodská tabule
VIC-1A	VI	C	1	A	Cidlinská tabule
VIC-1B	VI	C	1	B	Chlumecká tabule
VIC-1C	VI	C	1	C	Pardubická kotlina

Fig. 6: Part of content of the Podceleky relation

Okrsky : Tabulka

kod	soustava	podsoustava	celek	podcelek	okrsek	nazev
IA-1A-a	I	A	1	A	a	Haltravská hornatina
IA-1A-b	I	A	1	A	b	Nemanická vrchovina
IA-1A-c	I	A	1	A	c	Ostrovská vrchovina
IA-1C-a	I	A	1	C	a	Málkovská vrchovina
IA-1C-b	I	A	1	C	b	Plešivecká vrchovina
IA-1C-c	I	A	1	C	c	Havranská vrchovina
IA-1C-d	I	A	1	C	d	Rozvadovská pahorkatina
IA-1D-a	I	A	1	D	a	Štokovská vrchovina

Fig. 7: Part of content of the Okrsky relation

It is easier to define relations between tables in this model. Primary keys for the relations are simple or composite keys. The keys unambiguously define the geomorphological units. Cardinality (relationship) is 1:N.

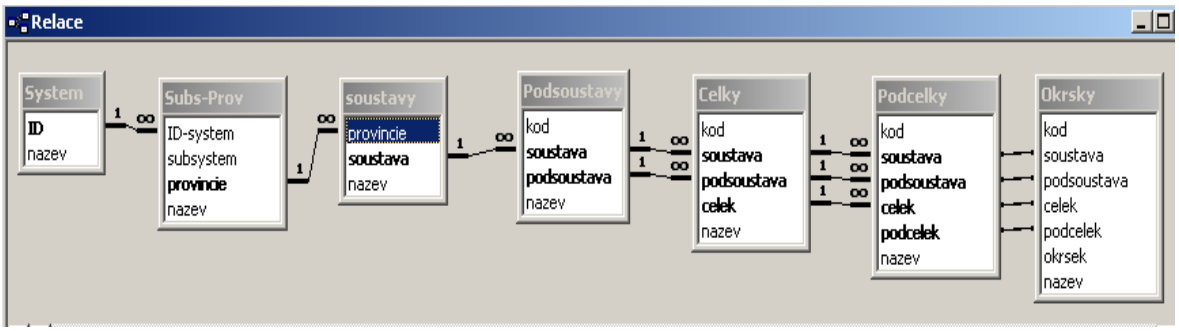

Fig. 8: Diagram of the relations between relations (tables)

5. Queries

By using the normalization process the tables have been divided into a number of the separate relations (tables). Now some method is needed to join the divided data from different tables together. For practical use it is necessary to prepare queries creating of not normalized relation, which includes verbal description of all of the superior hierarchy units.

In Microsoft Access we can create query that can find data in more than one table.

The example of a query follows. We are looking for all of the Celky type containing word „les“. There are 3 celky types as a result of the query:

Ždánický les, Slavkovský les a Český les.

In the report we can see the superior hierarchic structure and placing of the celky type in the structure of the geomorphological units. In this way we get quickly not only the required units

but also names of the superior units. In the original database only the code could be obtained. We would get the names of the superior units only by further analysis of the code.

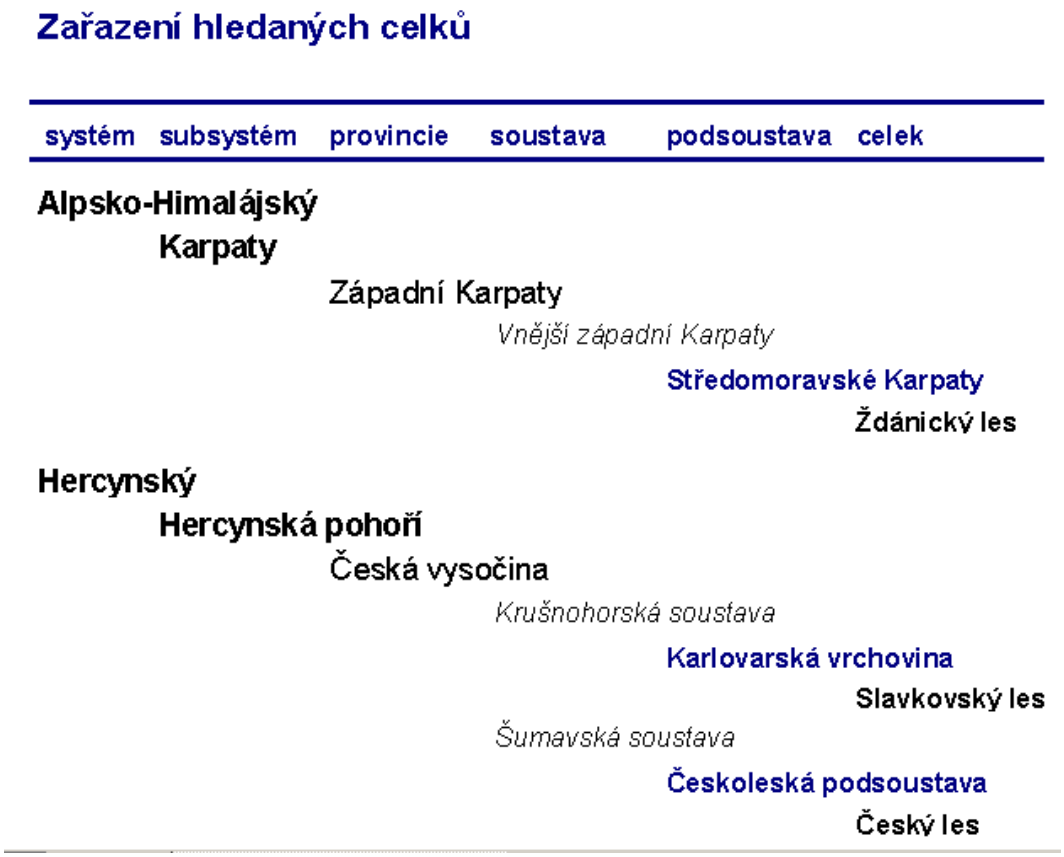

Fig.9: Output report

Various queries can be created to obtain the names of the units and their belonging to superior units. It is also possible to select a specific unit and to obtain all of the subordinate units to the specified level. The system (database) can answer all types of these queries.

Summary

In the presented example of the geographical database of the regional division of the relief of the Czech Republic we can see that by converting into a relational database it can be the basement of the information system.

By the process of normalization we have obtained the database composed of a number of separate relations (tables). The data can be joined again by queries.

It is easy to determine the placement of any unit in the hierarchic system of the units.

Relations of the data model could be easily extended by another attributes containing e.g. verbal description of the geomorphological unit or pictures.

Souhrn

Na uvedeném příkladu geografické databáze regionálního členění reliéfu České republiky vidíme, že jejím převodem do relační databáze získáváme informační systém, který může být základem geografického informačního systému.
Procesem normalizace jsme získali databázi, která je složena z několika samostatných relací. Sedm výsledných relací (tabulek) odpovídá geomorfologickým typům jednotek (systém, subsystém, provincie,… až okrsek). Mezi tabulkami jsou vytvořeny vztahy pomocí primárních a cizích klíčů. Pomocí dotazů jsme opětně schopni sloučit data dohromady. Lze sestavit různé dotazy. Například lze jednoduše určit zařazení kterékoliv jednotky v hierarchickém členění.
Relace datového modelu by bylo možné snadno rozšířit o další atributy, které by obsahovaly například textový popis geomorfologické jednotky nebo obrazovou dokumentaci.

References

Demek, J. a kol. : Zeměpisný lexikon ČSR, Hory a nížiny, Academia, Praha 1987
Riordan, R.M.: Vytváříme relační databázové aplikace,Computer Press, Praha 2000, ISBN 80-7226-360-9

Ing. Zdena Dobešová
Department of Geoinformatics
Faculty of Science
Palacký University in Olomouc
Tř. Svobody 26
771 46 Olomouc
Czech Republic
dobesova@prfnw.upol.cz

Reviewed: Dr. Ing. Jiří Horák