

Od terminológií ku klasifikáciám – redukcia informácií

Redakčná rada

Predseda: prof. MUDr. Milan Pavlovič, CSc.

Podpredseda: doc. MUDr. Anna Holomáňová, PhD

Tajomník: MUDr. Oskár Kadlec, CSc.

Členovia redakčnej rady

PhDr. Mária Bujalková, CSc., prof. MUDr. Ivan Ďuriš, DrSc., Mgr. Jana Levická, PhD, PhDr. Katarína Martinková, prof. MUDr. Dušan Meško, CSc., doc. PhDr. Ľudmila Ozábalová, PhD, doc. PhDr. František Šimon, CSc.

Členovia terminologickej komisie

doc. MUDr. Kamil Belej, CSc., prof. MUDr. Jozef Bilický, CSc., doc. Daniel Böhmer, PhD, doc. MUDr. Indrig Brucknerová, PhD, RNDr. Luboš Danišovič, PhD, prof. MUDr. Štefan Galbavý, PhD, prof. MUDr. Peter Gavorník, CSc., doc. MUDr. Jozef Glasa, CSc., prof. MUDr. Karol Holomáň, CSc., doc. MUDr. Marián Holomáň, CSc., doc. MUDr. Ján Chandoga, PhD, MUDr. Michaela Kostičová, PhD, prof. MUDr. Štefan Krajčík, CSc., prof. MUDr. Elena Elena Kukurová, CSc., prof. MUDr. Tibor Marček, CSc., doc. PhDr. Alojz Nociar, CSc., doc. MUDr. Daniela Ostatníková, CSc., doc. Vojtech Ozorovský, CSc., prof. Ľudmila Podracká, CSc., prof. MUDr. RNDr. Rudolf Pullmann, PhD, prof. MUDr. Ivan Schréter, PhD, prof. Ing. Vít Šajter, CSc., prof. MUDr. Ladislav Turecký, CSc.

Obsah

1 Úvod	3
2 Klasifikácia	4
3 Redukcia informácií	6
3.1 Od reality k pozorovaniu	6
3.2 Od pozorovania k lekárskeým záznamom	7
3.3. Od lekárskeým záznamom k diagnóze	7
3.4 Od diagnóz ku kódom a diagnostickým skupinám	7
4 Stanovenie informačného obsahu na každej vrstve granularity	8
5 Proces kódovania	9
5.1 Je kódovanie prirodzene odvoditeľné?	9
5.1.1 Odvodzovanie v hierarchickom strome	9
5.1.2 Rozdiel medzi živočíšnym systémom a systémom chorôb	10
6 Kódovanie artériovej hypertenzie v MKCH-10	11
6.1 Sémantické dimenzie	11
6.2 Problém súčet–sčítance	12
6.2.1 Aké následky má pridanie sčítance	12
6.2.2 Sú klasifikácie vôbec potrebné?	12
7. Závery	13

1 Úvod

V praxi sa obyčajne zamieňa pojem **chorý jedinec a choroba**. Riešenie tohto problému oživuje **opozíciu realizmu a nominalizmu**¹. Namiesto logických kategórií založených na klasifikačných pravidlách sa začali analyzovať „prirodzené“ kategórie, ktoré inšpirovali bádateľov k **prototypovej teórii** založenej na idei, že kategorizácia prebieha na základe zistenej podobnosti s prototypom. Ukázalo sa však, že **kategorizácia založená na príbuznosti** exemplára (pacienta) s prototypom (chorobou) **je neadekvátne**, pretože táto príbuznosť zahŕňa asociačné vzťahy medzi exemplármi, ktoré nevyhnutne nepredpokladajú prototyp. Prototypová organizovanosť kategórie je totiž len jedna z možností.

Princípom klasifikácie je identifikácia objektov a javov, na základe ktorej tieto dostávajú svoje miesto v systéme. Klasifikácia nie je zameraná primárne na riešenie úloh, t. j. získavanie nových poznatkov, ale na utvorenie celkového obrazu o tom, čo už poznáme. Dôležitou vlastnosťou jazyka je však z tohto hľadiska jeho klasifikačný potenciál.

Klasifikačný potenciál jazyka je schopnosť vniesť určitý systém do univerza objektov a javov teórie. Objekty, ktorých význam sa určoval pôvodne na základe kontextu ich objavu, dostávajú nový význam na základe svojho miesta v systéme. Často sa to, čo sa pokladalo za jedinečné stáva typickým. Keď „**prehovoria nemé parametre**“, vynoria sa okrem už známych objektov nové objekty zodpovedajúce novým hodnotám parametrov (Kvasz, 2010).

S klasifikačným potenciálom jazyka súvisí **zmena kategoriálneho systému**. Kategórie²

sú fundamentálne pojmy, ktoré v dôsledku ich fundamentálnosti jazyk neumožňuje explicitne vymedziť. Ilustráciou je posun v kategoriálnom systéme geometrie pri zmene, ktorú použil Alfred Tarski, keď namiesto troch základných objektov (bod, priamka, rovina) zaviedol jediný typ – body. Vďaka zmene kategoriálneho systému geometrie mohol Tarski geometrické pojmy klasifikovať podľa toho, aký vzorec (formulu) potrebujeme na ich definíciu. Zmena kategoriálneho systému teda úzko súvisí s klasifikačným potenciálom jej jazyka.

Informácie dostupné o pacientovi sú vždy chudobnejšie, ako sú dáta získané pozorovaním reálneho pacienta. Jazyk používaný na opis pacienta sa môže líšiť v závislosti od rozličných charakteristík: v ostatnom čase sa diskutuje najmä o rozdieloch medzi **ontologickým** (týkajúcim sa podstaty) a **epistemologickým hľadiskom** (týkajúcim sa poznania) (Bodenreider a spol., 2004 a Burgun a spol., 2005, cit. Straub a spol. 2006).

Jazyky sa však líšia aj svojou **granularitou**, t. j. informačným obsahom, resp. úrovňou podrobnosti identifikácie jednotlivých složek informačného objektu. Z tohto hľadiska možno rozlišovať voľne prístupné prirodzené jazyky, štandardizované a štruktúrované terminológie (s jemnou granularitou a veľkým informačným obsahom) a klasifikácie (ako je SNOMED CT) a klasifikácie s hrubou granularitou a malým informačným obsahom (ako je MKCH-10). Z informácií nachádzajúcich sa v termínoch a kódach klasifikácie sa nedá extrahovať informácia, ktorú obsahuje jemne granulovaný jazyk (terminológia). Možno však postupovať opačne

vlastnosť a vzťah. Podľa **Kanta** sa kategórie nevzťahujú na poznávaný objekt, ale patri k rozumu ako schopnosti poznania. Odvodzujú sa od štyroch základných hľadísk, podľa ktorých sa delia všetky sudy (kvalita, kvantita, vzťah a modalita). V **biologických taxonómiiach** sa pod pojmom kategória (angl. *rank*) rozumie klasifikačná úroveň v hierarchickej stupnici (čiže druh, rod, čeľaď, rad ap.). Označuje postavenie taxónu v hierarchickej klasifikácii.

¹**Nominalizmus** je smer vo filozofii, ktorý tvrdil, že reálne jestvujú iba jednotlivé veci s ich individuálnymi vlastnosťami a že všeobecné pojmy sú iba mená jednotlivých predmetov.

²**Kategória** je podľa **Aristotela** základný rod bytia, trieda predikátov, ktoré sa dajú použiť na každý predmet. Existujú tri základné kategórie: vec,

a priradiť nejaký prípad ku klasifikácii pomocou termínov v samotnej terminológii? Na prvý pohľad sa to zdá samozrejmé, ale pri podrobnejšom štúdiu s tým vznikajú problémy.

2 Klasifikácia

Klasifikácia (angl. *classification* z l. *classis* trieda + l. *facere* robiť) je triedenie, zaraďovanie jednotlivých faktov, údajov a meraní do vopred určených tried na základe zistenia spoločnej charakteristiky. Ide o postup, ktorý sa zakladá na rozdelení množiny, resp. súboru prvkov (javov, objektov, prípadov ap.) do určitých podmnožín podľa zvolených variantov tzv. klasifikačného (triediaceho) znaku; **1.** metóda určenia vzťahov (rodovo-druhových a i.) medzi sémantickými jednotkami, nezávisle od stupňa hierarchie systému a od toho, či sa použije pri tradičnom alebo mechanizovanom informačnom prieskume; **2.** usporiadanie pojmov do tried a podtried s cieľom vyjadriť sémantické vzťahy medzi nimi. Triedy sa vyjadrujú pomocou **notácie**³.

Rozlišuje sa klasifikácia **analytická** (postup od základnej množiny k triedam a prvkom) a **syntetická** (postup v opačnom smere). Základná množina je spravidla neusporiadaná, preto klasifikácia znamená aj jej usporiadanie priradením prvkov do príslušných tried. Základnú množinu možno určiť **enumeráciou** (vymenovaním všetkých prvkov) alebo **definične** (na základe znaku alebo sústavy znakov), ktoré rozhodujú o príslušnosti prvku do danej základnej množiny (napr. všetky gravidné ženy s hypertenziou). Vstupným prepokladom klasifikácie je možnosť jednoznačne rozhodnúť o tom, či daný prvok patrí alebo nepatrí do základnej množiny.

Jednotlivé podmnožiny zodpovedajú jednotlivým **klasifikačným triedam** alebo **kategóriám**⁴. Kategorizácia sa vykonáva v

³ **Notácia** je sústava triednikov a pomocných znakov prijatých pre danú klasifikačnú schému.

⁴ **Kategória** je najvšeobecnejší významový útvar, ktorý je výsledkom poznania všetkých entít reality z istého (napr. filozofického, špeciálnocedného

rámci priameho pozorovania, exploračných metód, pri analýze dokumentov (najmä pri obsahovej analýze). Je pojmovým ekvivalentom klasifikácie. Pri **jednostupňovej klasifikácii** (v prípade používania len jedného klasifikačného znaku) zodpovedá každá trieda jedinému vyčlenenému variantu klasifikačného znaku (napr. pohlavie muži—ženy). Pri **viacstupňových**, jemnejších **klasifikáciách**, kde sa používajú 2 a viaceré triediace znaky, zodpovedá každá trieda jedinečnej kombinácii variantov všetkých triediacich znakov (napr. muži—ženy, 18 — 45- roční, s hypertenziou). Počet možných tried závisí od počtu zvolených klasifikačných znakov a počtu ich vymedzených variantov. Pri dvoch dichotomických znakoch sa počet tried rovná 4, pri troch 8, pri troch znakoch s troma variantmi 17 atď.

Korektnosť klasifikácie závisí od rešpektovania elementárnych pravidiel, ktoré majú logický a heuristický charakter a vyjadrujú určité **podmienky** (Schenk, in: Veľký sociologický slovník 1996, 486 — 487): **1. funkčnosť** (klasifikácia sa odvodzuje z konkrétneho výskumného zámeru, ktorý

atď.) hľadiska. Je výsledkom poznania najvšeobecnejších stránok reality. Treba rozlišovať kategóriu a pojem kategórie. **Kategória** podľa Aristotela je základný rod bytia, trieda predikátov, ktoré sa dajú použiť na každý predmet. Existuje 10 kategórií: vec, vlastnosť, vzťah, kvantitat, kvalita, miesto, čas, poloha, manie, činnosť a trpenie. Hlavnými kategóriami sú prvé tri z nich. **Pojem kategórie** je kognitívny významový útvar, ktorý je výsledkom poznania kategórie (napr. pojem, súd, úsudok, hypotéza, teória, pravidlo, vedecký zákon...). **Špeciálnovedné kategórie** je kľúčový pojem určitej špeciálnej vedy (napr. fyziky). Sú nosnými piliermi, určovateľmi a usmerňovateľmi vedeckého hľadania, ukazovateľmi ciest formovania vedeckých hypotéz a výstavby teórií.. Všetky ostatné pojmy špeciálnej vedy sa musia vyjadriť (definovať) prostredníctvom špeciálnovedných kategórií. Sú základnou podmienkou možnosti špeciálnej vedy.

determinuje výber klasifikačných znakov i vymedzenie ich variantov); **2. jednodimenzionalnosť** (klasifikačný znak musí byť presne vymedziteľný a kvalitatívne homogénny z hľadiska vecného i logického — je rozdiel napr. medzi triedením osôb so zvýšenou celkovou hmotnosťou tuku v tele a hmotnosťou brušného tuku — žiadny z variantov klasifikačného znaku sa nesmie vzťahovať k inému kontínuu); **3. nezávislosť** (požiadavka, aby sa klasifikačné triedy neprekrývali, nemali ani jeden spoločný prvok — z logického hľadiska to znamená, že prienik ľubovoľných tried, ktoré vznikli pri klasifikácii, musí byť prázdny); **4. úplnosť** (klasifikácia musí byť vyčerpávajúca, každý prvok základnej množiny musí byť principiálne zaraditeľný do jednej z utvorených klasifikačných tried; pri správnej klasifikácii je zjednotenie všetkých tried, resp. podmnožín ekvivalentné základnej množine); **5. indikácia** (o každom, prvku základnej množiny sa musí dať jednoznačne rozhodnúť, do ktorej triedy patrí). Z logického hľadiska je klasifikácia korektná vtedy, keď sú splnené podmienky 2 až 5., t. j. keď je každý prvok základnej množiny jednoznačne zaradený (zaraditeľný) výlučne do jedinej z možných tried.

Klasifikácia je špeciálny prípad logickej operácie delenia rozsahu pojmov, ktorý predstavuje určitý súhrn delenia (delenia triedy na druhy, delenia týchto druhov atď.). Je to logické usporiadanie množiny klasifikovaných pojmov na základe vzťahov a závislostí medzi charakteristickými znakmi týchto pojmov. Ide o metódu určenia vzťahov (rodovo-druhových a i.) medzi sémantickými jednotkami, nezávisle od stupňa hierarchie systému a od toho, či sa použije pri tradičnom alebo mechanizovanom informačnom prieskume (FID/CR, 1973). Vo všeobecnosti ide o usporiadanie pojmov do tried a podtried s cieľom vyjadriť sémantické vzťahy medzi nimi. Triedy sú vyjadrené pomocou **notácie** (sústavy triednikov a pomocných znakov prijatých pre danú klasifikačnú schému). Triediaci znak je reprezentácia určitej triedy pomocou notácie klasifikačnej schémy (ISO 5127).

Klasifikácia je najstaršia a najjednoduchšia vedecká metóda. Je to východisko formulovania teórie, bez ktorého nemožno odkrývať kauzálne vzťahy medzi vecami a javmi. Keď chceme potvrdiť jestvovanie triedy, musíme rozoznať jej spoločnú a najtypickejšiu zložku, napr. stavec pre stavovce, myslenie pre ľudí. Klasifikácia predpokladá rozlišovanie menších celkov, ktoré sú charakteristické pre väčší celok, t. j. triedu. **Cieľom klasifikácie je zistiť v systéme pravidelnosť, poriadok, spolupatričnosť, protikladnosť** (Filkorn, 1956).

Klasifikácia dominuje najmä v začiatkových štádiách rozvoja vednej disciplíny, resp. v jej predteoretickom štádiu. Klasickým príkladom je biológia (botanika a zoológia), chémia a i., v ktorých **sama klasifikácia znamenala dôležitý poznatok** alebo stimulovala významný objav alebo teoretický pokrok. Klasifikácia je aj prostriedkom prezentácie a systematizácie poznatkov. Je elementárnym štádiom a nástrojom poznávania, predpokladom a východiskom analýzy a predstupňom ďalších, nadväzujúcich úloh, postupov a metód.

Osobitný význam má klasifikácia v deskriptívnom výskume, pri aplikácii metód merania, štatistickom triedení, v teórii a pri tvorbe **typológií**.

Klasifikáciou pojmov vedného odboru sa zisťujú vzťahy medzi pojmi (termínmi). Z týchto vzťahov je najdôležitejší **vzťah rod – druh**. Ten sa výrazne manifestuje pri väčšine termínov so zhodnými alebo nezhdnými prívlastkami (Masár, 1991). Napr. encefalitída je rod, japonská encefalitída druh. Kritériom na zaradenie do triedy je vzťah medzi určujúcim a určovaným prvkom termínu.

Klasifikácia na základe podstatných znakov sa nazýva **typológia**; vychádza z pojmu **typ** ako jednotky rozčlenenia skúmanej reality, z pojmu konkrétneho ideálneho modelu historicky sa vyvíjajúcich objektov (biologické, jazykové, kulturologické ap. typológie).

Každá klasifikácia je výsledkom určitého zjednodušenia skutočných hraníc medzi druhmi, lebo tieto hranice sú vždy podmienené a relatívne. S rozvojom poznania sa klasifikácia spresňuje a mení.

Pri klasifikácii je dôležitým krokom vydelenie všeobecných pojmov, najmä tzv. fundamentálnych pojmov, v medicíne napr. pojmov chorôb.

Pri klasifikácii sa berú za základ delenia znaky, ktoré sú pre dané predmety podstatné. V tomto prípade **prirodzená klasifikácia**) odhaľuje podstatné zhody a rozdiely medzi predmetmi a má poznávací význam.

V iných prípadoch, keď cieľom klasifikácie je len systematizácia predmetov, berú sa za základ delenia znaky vhodné na tento cieľ, ale nie podstatné pre samy predmety (napr. abecedné katalógy). Takéto klasifikácie sa nazývajú **umelými klasifikáciami**.

Najväčšiu hodnotu majú klasifikácie založené na poznaní zákonov súvislostí medzi druhmi, na poznaní prechodov jedného druhu do iného vo vývojovom procese. Taká je napr. Mendelejevova klasifikácia chemických prvkov, klasifikácia organických zlúčenín, taxonómia rastlín a živočíchov).

Utvorenie hierarchického stromu si vyžaduje splnenie dvoch podmienok: podmienku výlučnosti (disjunktivity) a (unidimenzionality) jednosmernosti.

Disjunktivita znamená, že súrodenci na každej úrovni sa navzájom vylučujú. Keď je cicavec psom, nemôže byť súčasne mačkou.

Unidimenzionalita v hierarchii znamená, že vetvenie postupuje len jedným smerom: cicavce sa môžu diferencovať na psy, mačky, kravy, slony atď. Táto diferenciácia však nemôže postupovať opačným smerom: slony sú cicavce a nemôžu byť nikdy rybami. Medzi prirodzenou a umelou klasifikáciou je

3 Redukcia informácií

3.1 Od reality k pozorovaniu

Vyšetrenie pacienta je diagnostický proces založený na analýze a syntéze. Nevyhnutný je

podstatný rozdiel. Napr. systém chorôb sa nedá usporiadať do hierarchického stromu prirodzene, pretože tu absentuje disjunktivita a unidimenzionalnosť. Naproti tomu v zoológii je disjunktivita zaručená prirodzene skutočnosťou, že dva druhy sa nikdy nemiešajú. Jestvuje tu druhová bariéra, pretože napr. mačky a psi nemôžu mať spolu potomkov, a preto sú tieto dva druhy disjunktívne.

Podmienka unidimenzionality sa zakladá na histórii evolúcie druhov. Jednosmernosť evolúcie nemožno zvrátiť.

Takáto unidimenzionalnosť absentuje pri vývoji diagnóz. Choroby nevznikajú z iných chorôb ako živočíšne druhy, ktoré sa vyvíjajú v priebehu času z prechádzajúcich dávnych druhov.

Keďže pri klasifikácii nastáva dramatická redukcia informačného obsahu, vzniká otázka, či sú vôbec klasifikácie v medicíne potrebné. Odpoveď je kladná. Klasifikácie sú síce zjednodušením reálneho sveta, ale naše myslenie je schopné pojať všetky informácie reálneho sveta. Redukcia informácií sa preto pokladá za jednu z hlavných úloh ľudskej inteligencie (Straub a spol. 2006).

Keď máme porovnať dva systémy, musíme najprv zoskupiť ich príslušné jednotlivé prvky, čo znamená zjednodušenie a klasifikáciu. Štatistické porovnanie napr. rozličných terapeutických režimov v dvoch odlišných skupinách pacientov nie je možné bez ich zoskupenia.

Keď postupujeme od jemne granulovanej k hrubo granulovanej informačnej úrovni, teda od reality k terminológii alebo od terminológie ku klasifikáciám, stretávame sa vždy s problémom redukcie informácií. Klasifikovaná informácia znamená prídavnú hodnotu (Straub a spol. 2006).

prítom individuálny prístup, pretože ide o vždy o jedinečné objekty a neopakovateľné deje.

Pri pozorovaní reality (napr. vyšetrovaní pacienta) nastáva ohromná redukcia obsahu informácií, pretože pozorovateľ má len obmedzené možnosti vyšetrenia (úzke zorné

pole) a navyše svoju pozornosť aktívne zameriava určitým smerom (zámerne selektuje informácie relevantné pre konkrétny prípad). Príliš veľa informácií by totiž mohlo pozorovateľa zmiasť.

3.2 Od pozorovania k lekárskej záznamom

Nie každé pozorovanie je natoľko významné, aby sa zaznamenalo a len vlastne malá časť informácií nachádzajúcich sa v mysli lekára sa zaznamenáva. Záznamy môžu mať aj podobu kvantitatívnych dát, ako sú biometrické údaje, výsledky fyzikálneho, biochemického, imunologického a i. laboratórneho vyšetrenia, podobu grafov (elektroencefalogram, elektrokardiogram, elektromyogram...), obrazov (snímky rtg, CT, MRI ap.) alebo slov (kvalitatívne dáta). Hlavný obsah terminológií a klasifikácií v medicíne tvoria kvalitatívne informácie, získané napr. anamnézou.

3.3 Od lekárskeho záznamu k diagnóze

Diagnóza sa stanovuje priamym vyšetrením pacienta alebo vylúčením iných diagnóz prichádzajúcich do úvahy (*diagnosis per exclusionem*). Diagnózu možno určiť aj podľa úspechu alebo neúspechu terapie (*diagnosis ex iuvantibus*). Uvádza sa, že vyše 60 % diagnózy sa stanovuje na základe **anamnézy**. Jednotlivé diagnózy sa formulujú verbálne a tradične v latinčine.

Diagnostický proces je procesom abstrakcie. Je to empiricko-intuitívny proces utvárania všeobecného pojmu našim myslením vychádzajúci z konkrétneho a smerujúci k všeobecnému. Diagnóza predstavuje obyčajne len malú časť z celkových informácií v zdravotnom zázname.

3.4 Od diagnóz ku kódom a diagnostickým skupinám

Na základe neobmedzeného počtu diagnostických údajov – subjektívnych a objektívnych príznakov – sa pacienti zaraďujú do obmedzeného počtu

diagnostických skupín. Je pritom dôležité, aby bola štruktúra klasifikačného systému čo najprehľadnejšia a počet skupín nebol príliš veľký. Hlavným kritériom pre zaradenie do skupín je diagnóza pacienta alebo zdravotný výkon. V závislosti od toho, či sa za kritérium klasifikácie zvolí diagnóza, funkčné schopnosti človeka a ich obmedzenia alebo výkon sa rozlišujú klasifikácie chorôb, postihnutia a výkonov.

V **Medzinárodnej klasifikácii chorôb** (MKCH-10), ktorá vstúpila do platnosti 1. 1. 1993 (v Slovenskej republike 1. 1. 1994) sa nahradil tradičný číselný kódovací systém **alfanumerickou schémou**. MKCH.10 obsahuje klasifikáciu na trojmiestnej a (nepovinné) štvormiestnej úrovni, klasifikáciu morfológie nádorov, špeciálny tabuľkový zoznam definícií príčin chorôb a smrti a pravidla nomenklatúry. Tabuľkový zoznam chorôb obsahuje 298 príčin smrti.

V USA sa pre účely sociálnej poisťovne Medicare vypracoval r. 1983 **systém diagnostických skupín** (Diagnosis-Related Group, DRG) na klasifikáciu nemocničných prípadov do jednej z asi 500 skupín, založený na diagnózach Medzinárodnej klasifikácie chorôb, výkonoch, pohlaví, veku, stavu pri prepustení a prítomnosti komplikácií či komorbidít. Vychádzal z toho, že pacienti tej istej kategórie sú klinicky podobní a ich liečenie si vyžaduje tie isté náklady. Diagnostické skupiny sa ďalej klasifikujú do hlavných diagnostických kategórií (Major Diagnostic Categories, MDC)

Nie všetky diagnózy a výkony majú svoju skupinu a naopak každá skupina zahŕňa aj niekoľko desiatok diagnóz. Zaradenie pacienta do jednej skupiny má dve základné kritéria: **1.** majú podobný spôsob liečby; **2.** majú podobné náklady na liečbu. Používaný systém diagnostických skupín má 25 hlavných diagnostických kategórií (nadskupín), ktoré sa ďalej členia do 941 podskupín. Pri začleňovaní pacientov do jednotlivých skupín sa vykazujú najmä tieto údaje: základná diagnóza, vedľajšie diagnózy, výkony, dĺžka hospitalizácie a základné údaje o pacientovi (vek, pohlavie,

pôrodná hmotnosť...). Toto sú však dáta, ktorá každé zariadenie už vykazuje zdravotným poisťovňam a tie na ich základe vykonávajú vyúčtovanie. Na zaradovanie jednotlivých prípadov do skupín slúžia manuály alebo špeciálne počítačové programy (groupery). Grouper je väčšinou prepojený s informačným systémom nemocnice, ktorý pripravuje dáta pre zdravotné poisťovne.

Rozčlenenie prípadov do diagnostických skupín umožňuje vyrátať náklady na ich liečbu, hodnotiť nemocničnú starostlivosť a priame úhrady posteľovým zdravotníckym zariadeniam, objektívnejšie porovnávanie kvality zdravotnej starostlivosti, jej poskytovaní a nákladovosť, porovnávanie jednotlivých zariadení a hodnotiť využiteľnosť financií, oplánovanie, tvorba rozpočtu atp.

Výstupy systému diagnostických skupín sú využiteľné na všetkých stupňoch riadenia zdravotníctva – manažment nemocníc, zriaďovatelia nemocníc, orgány štátnej správy, inštitúcie s celoštátnou pôsobnosťou, odborné spoločnosti, ale aj samostatní platcovia zdravotnej starostlivosti. Systém DRG sa využíva aj v informačných databázach.

V tejto etape nastáva opäť redukcia objemu informácií a aj táto redukcia je zámerná. Čím je menej kódov a diagnostických skupín, tým ľahšie je štatistické porovnávanie skupín.

V ostatných 30 rokoch nastal posun v nazeraní na **hodnotenie zdravotného postihnutia**⁵, najmä v súvislosti s predĺžením života. Človek s postihnutím (*disability*) sa nepokladá za izolovaného jedinca s určitou diagnózou, ale za jedinca, ktorého problémy sú dané dynamickou interakciou medzi ním

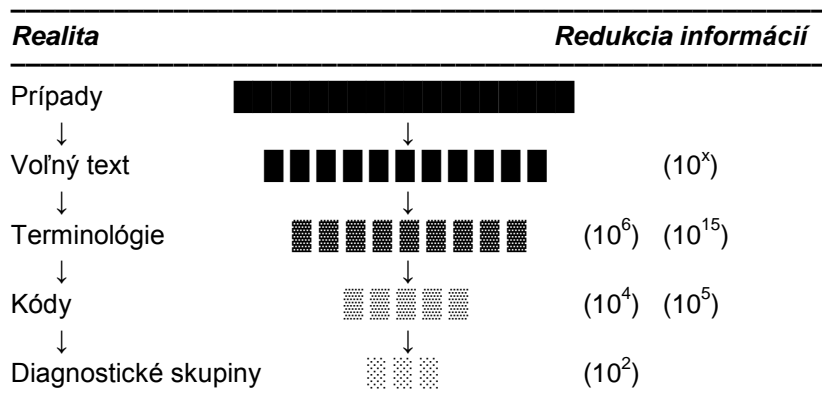
a prostredím. V súčasnosti sa hovorí aj o ďalšom posune, a to k občianskemu modelu. Ten však neznamená nahradenie biopsychosociálneho modelu, ale jeho rozšírenie o dôraz na aktívnu participáciu občanov s postihnutím. Ide o zvýšenie kvality života. Preto SZO vypracovala **Medzinárodnú klasifikáciu funkčných schopností, postihnutia a zdravia** (MKF). Kým MKCH-10 poskytuje „diagnózy“ chorôb, chýb a iných zdravotných problémov, MKF dopĺňa túto informáciu o funkčné schopnosti. Dvaja jedinci s rovnakou chorobou môžu mať rozdielny stupeň funkčnej schopnosti a dve osoby s rovnakým stupňom výkonnosti nemusia mať nevyhnutne rovnaké zdravotné problémy. Prídavné hodnotenie rozširuje informácie o zdravotnom stave jedinca.

4 Stanovenie informačného obsahu na každej vrstve granularity

Na obr. 1 je uvedený približný informačný obsah jednotlivých vrstiev granularity. Určiť prípustný počet prípadov vo vrstve možno na základe odhadu informačného obsahu vybraného prípadu (selektívny obsah informácie podľa Shannona, 1963 a MacKeya, 1969, cit. Straub a spol. 2006).

Diagnostických skupín je niekoľko stoviek a zahŕňajú obyčajne **menej ako 1000 skupín**. **MKCH-10** má v závislosti od jej príslušnej verzie asi **15 000 kódov**. **SNOMED CT** obsahuje **više 1 milióna** termínov. Pri porovnaní s týmto zatiaľ malým počtom nemožno kvantifikovať informačný obsah prípadu v realite. Na obr. 1 si treba predstaviť ohromné množstvo informácií ako mračno informácií s amorfnou konfiguráciou.

⁵ **Zdravotné postihnutie** je zníženie funkčných schopností na úrovni tela, jedinca alebo spoločnosti, ktoré vzniká, keď sa občan so svojím zdravotným stavom stretáva s bariérami prostredia. Neklasifikujú sa osoby, ale opisuje a klasifikuje sa situácia každého jedinca v rade okolností, vzťahujúcich sa na zdravie (Medzinárodná klasifikácia funkčných schopností, postihnutia a zdravia, WHO, 2001).



Obr. 1. Od reality k terminológiám a klasifikáciám

Čísla v zátvorke (a hustota obdĺžnikov v interpretačných vrstvách) odráža skutočnosť, že aj keď jestvujú rozličné kódy MKCH-10 pre každý prípad, je pre daný prípad len jedna diagnostická skupina. Informačný obsah jedného kódu MKCH-10 je násobkom a informačný obsah celku je súčin obsahov jednotlivých kódov. Predpokladá sa, že každá diagnostická skupina má dva kódy. Je to však len hrubý odhad. Nie každá kombinácia kódov je možná, každý prípad má však obyčajne viac ako dva diagnostické a terapeutické kódy.

Čo platí pre kódy, platí aj pre termíny. Mnoho termínov sa kombinuje a kóduje jedným kódom. Nie každý termín sa používa na kódovanie MKCH-10. Preto sa redukuje nielen informačný obsah jedného termínu SNOMED CT na jeden kód MKCH-10, ale viaceré termíny v lekárskom zázname majú len jeden kód.

Ako vyplýva z obr. 1, množstvo informácie zdola (od diagnostickej skupiny) nahor (k voľnému textu lekárskeho záznamu) natrastá. Informácia v reálnom prípade (mračno informácií) je opäť bohatšia ako informácií v lekárskom zázname. Opačným smerom, od reálneho prípadu ku kódom diagnostických skupín informačný obsah výrazne klesá.

5 Proces kódovania

Kódovanie je proces, pri ktorom sa každému znaku alebo postupnosti znakov daného súboru znakov jednoznačne priradí znak alebo postupnosť znakov z iného súboru

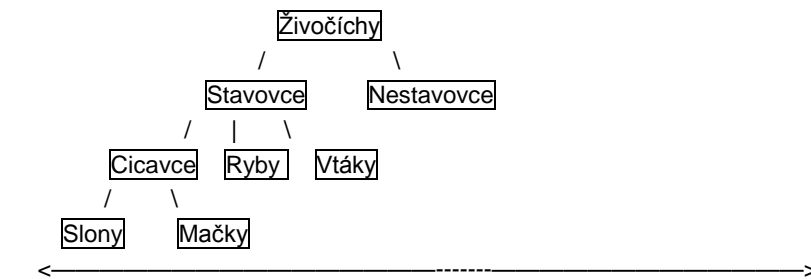
znakov. Kódovanie je teda transformácia určitej informácie z jednej formy na druhú pomocou určitého postupu – algoritmu, ktorý je väčšinou verejne známy. Na rozdiel od šifrovania nie je účelom kódovania utajenie informácie, len jej iná forma zápisu vybraná tak, aby sa informácia dala čo najlepšie alebo najúspornejšie uchovať alebo preniesť. Vďaka počítačom sa najčastejšie sa používa kódovanie údajov a informácií do číselnej podoby (tzv. digitalizácia). Používa sa však aj nečíselné kódovanie.

Straub a spol. (2006) vyvinuli počítačový program na **automatické kódovanie MKCH**, ktorý je schopný čítať voľný text (spojenia podstatných mien) a produkovať kódy MKCH-10. Keď vstup nebol dostatočne presný, program žiadal chýbajúcu informáciu vo forme kontextovo špecifickej otázky s viacerými možnosťami. Ako vnútorný reprezentačný jazyk sa použili pojmové molekuly, ktoré umožňujú presné a štruktúrované modelovanie opisného informačného obsahu slova v prirodzenom lekárskom jazyku, ako aj informácie obsiahnutej v kódoch MKCH-10.

5.1 Je výsledok kódovania prirodzene odvoditeľný?

5.1.1 Odvodzovanie v hierarchickom strome

Na obr. 2 je znázornená hierarchia s dvoma podmienkami: **disjunkcia vo vylučujúcom zmysle** (*disjunktivity*) a **jednosmernosť** (*unidirectionality*).



Obr. 2. Pravá hierarchia – Zoologický systém (podľa Strauba a spol. 2006)

Disjunkcia znamená, že súrodenci na každej úrovni sa navzájom vylučujú. Keď je pes cicavcom, nemôže byť pes súčasne mačkou.

Jednosmernosť v hierarchii znamená, že vetvenie prebieha len jedným smerom: cicavce sa môžu diferencovať na psov, mačky, kravy, slony atď. Táto diferencácia nie je však možná opačným smerom: slony sú cicavce a nemôžu byť nikdy rybami. Keby hierarchia nebola prísne jednosmerná, obsahovala by štruktúry reťazcov a nebola by to hierarchia, ale sieť.

Keď sa rešpektujú tieto dve podmienky, máme do činenia s pravou hierarchiou a môžeme spätne pomocou hierarchického stromu vyvodzovať z jeho listov závery o vetvách: keď vieme že objekt je slon, môžeme vyvodiť záver, že je cicavcom a že je stavovcom. Môžeme tiež prechádzať vlastnosťami prvkov v hornej vrstve k vlastnostiam nižších vrstiev. Slon má všetky vlastnosti cicavcov, ako aj stavovcov.

To je výhoda z hľadiska reprezentácie poznatkov: netreba opätovne uvádzať všetky informácie o každom druhu, o slonoch, psoch, mačkách atď., pretože stačí uviesť spoločnú informáciu raz na vyššej úrovni. Šetrí to priestor v reprezentácii, uľahčuje údržbu a zvyšuje transparentnosť.

Hierarchické stromy sú preto ideálne na reprezentáciu poznatkov.

Vlastnosti prechádzajú od koreňa k listom, od hrubej úrovne granularity k jemnej. Informácia o triedach je však odvoditeľná opačným procesom, od úrovne s jemnou

granularitou k úrovni s hrubou granularitou (od slona k cicavcom). Táto dedukcia je v hierarchickom strome samozrejmalá, ale závislá od uvedených dvoch podmienok.

Prirodzená odvoditeľnosť tohto druhu od úrovne s jemnou granularitou k úrovni s hrubou granularitou vyplýva z procesu kódovania. Keď sa dá informačný redukčný „lievik“ znázornený na obr. 1 označiť ako hierarchia, mal by sa prípad z medicíny identifikovať s úrovňou s hrubou granularitou a z nej opísať úroveň s jemnou granularitou. Mal by sa dať odvodiť kód MKCH-10 z opisu prípadu lekáskymi termínmi bez vonkajšej asistencie.

5.1.2 Rozdiel medzi živočíšnym systémom a systémom chorôb

Systémy chorôb sa nedajú prirodzene usporiadať do hierarchického stromu. Príčinou je absencia dvoch podmienok potrebných na utvorenie hierarchie, disjunktivity a jednosmernosti, ktoré sa uplatňujú v živočíšnej a rastlinnej ríši, ale nie v prípade chorôb.

V zoológii je podmienka vylúčenosti prirodzene zaručená skutočnosťou, že dva druhy sa nemôžu zamieňať (*druhovú bariéru*). Pretože mačky a psi nemôžu mať spolu potomkov, sú tieto dva druhy nezlučiteľné, disjunktívne.

Podmienka jednosmernosti sa zakladá na histórii vývoja druhov (fylogenéze). Pretože druhy sa vznikli v priebehu jednosmerného vývoja, evolúcia nemôže byť vratná. Zo slonov sa nikdy v budúcnosti nemôže vyvinúť ryba.

Evolúcia živočíšnych druhov je však v prírode osobitným prípadom. Skutočnosť, že tento systém má podobu perfektného hierarchického stromu je následkom fylogenetického vývoja.

Takáto história absentuje pri vývoji diagnóz. Choroby nevznikajú z iných chorôb ako živočíšne druhy zo svojich dávnejších predkov.

Choroby samozrejme navzájom súvisia, jedna choroba môže vyvolať inú chorobu. Ale tieto vzťahy sú oveľa zložitejšie ako v zoológii.

Pretože tieto dve podmienky – disjunktivita a jednosmernosť – sa neuplatňujú v systéme chorôb, tento systém sa prirodzene nevyskytuje v podobe hierarchie. Keď z neho chceme z praktických dôvodov utvoriť hierarchiu, musíme ju utvoriť umelo. Akonáhle je však štruktúra arteficiálna, jej tvar sa dá meniť a stáva sa ľubovoľná, arbitrárna.

Na zdokonalenie systému možno použiť štatistické metódy, napr. varianciu redukcie s ohľadom na cieľovú premennú, ako ju uskutočnil v prípade diagnostických skupín Fetter a spol (1991).

MKCH-10 sa označuje ako hierarchia. To má mnoho výhod, ale treba si uvedomiť, že jeho štruktúra je arbitrárna, akokoľvek dokonalý je jej dizajn. Keď chceme priradiť kódy MKCH-10 k diagnózam, musíme redukovať zložitú informáciu reálneho prípadu diagnózy, aby zapadla do arteficiálneho hierarchického stromu. Na určenie toho, aké informácie sa pritom stratili, treba aplikovať prídavné pravidlá – pravidlo zahrnutia a vylúčenia.

6 Kódovanie artériovej hypertenzie v MKCH-10

6.1 Sémantické dimenzie

Keď chceme kódovať diagnózu, musíme najprv analyzovať charakteristiky použité v cieľovom kódovacom systéme. Termíny použité na opis kódov sa najlepšie usporadúvajú do skupín s tým istým sémantickým „vzorom“.

Termíny toho istého vzoru predstavujú **tokeny**⁶ (bloky textu) toho istého typu. Každému vzoru možno zvyčajne priradiť jeden token nezávisle od diagnózy, takže diagnóza má toľko priradených tokenov, koľko je vzorov. Vzorky možno pokladať za sémantické dimenzie, ako **osi v sémantickom priestore** alebo **stupne voľnosti** (aby sa vyjadrila nezávislosť každej dimenzie). Sú príbuzné, nie však totožné s hranami (partitions) a charakteristikami (features, qualities) sémantickej siete.

Sémantické dimenzie alebo osi tvoria multifokálnu sieť, ktorú autori používajú pri metóde pojmových molekúl ako základ reprezentujúci význam termínu a sémantického vyvodzovania (semantic reasoning). Zo sémantických dimenzií nevznikajú ontologické výroky. Sú len spôsobom reprezentácie významu lekárskeho slov a tried v kódovacích systémoch a slúžia ako technický prostriedok na preklad z jazyka do kódov.

V MKCH-10 sa nachádza týchto 5 tried alebo dimenzií na segmentáciu (partitioning) diagnózy hypertenzie:

- Príčina: primárna, renovaskulárna, renoparenchymatózna, endokrinná
- Malignita: malígna, benígna
- Postihnutie orgánov: s kardiopatiou, s nefropatiou, bez orgánového postihnutia
- Príznaky orgánového postihnutia: s nedostatočnosťou srdca, s nedostatočnosťou obličiek, bez takýchto príznakov
- Hypertenzná kríza: s krízou, bez krízy

Týchto 5 dimenzií sa spojí do jednej dimenzie, ktorá je dostupná v MKCH-10. Po spojení 5 dimenzií do jednej, zisťujeme, že sa znázornia niektoré kombinácie, ktoré sú teoreticky možné, nie však všetky:

- s krízou/bez krízy možno pridať ako 5. číslicu ku každému 4-číslicovému kódu

⁶ **token** – v počítačovej vede základná textová jednotka použitá v analýze alebo spracovaní textu

- orgánová nedostatočnosť sa dá priradiť k len hypertenzii s orgánovým postihnutím
- nemožno znázorniť malignitu sekundárnej hypertenzie
- organické postihnutie a príčina hypertenzie sa nedajú znázorniť súčasne.

Pri postupe od úrovne s jemnou granularitou k úrovni s hrubou granularitou nastáva strata informácií. Táto strata informácií sa dá očakávať. Napr. pri prechode z 5-miestneho kódu na 3-miestny sa môžeme ešte domnievať, že máme prístup k predpovedateľným informáciám, preto prechod z terminológie ku klasifikácii nie je problémom. To však neplatí z dvoch dôvodov:

1. Sémantické stupne voľnosti sú v kódujúcom hierarchickom systéme určené arbitrárne. Nie je totiž dané prírodou, že jeden stupeň voľnosti bude taký efektívny na úrovni 3-miestneho kódu ako na úrovni 4-miestneho kódu. Sekvencia sémantických tried je arbitrárnou charakteristikou kódujúceho systému. Kódujúci systém má vlastné charakteristiky. Keď poznáme jeho štruktúru, môžeme bez problémov vyvodzovať závery o klasifikácii z terminológie.

2. Reálna prekážka spočíva však v kódoch, ktoré reprezentujú kombináciu dvoch alebo viacerých diagnóz. Keď napr. porovnáme benígnu hypertenziu (I10) a hypertenznú kardiopatiu (I11), prídavná informácia k I10 je „benígna“, kým k I11.0 je kardiálna insuficiencia. Na rozdiel od „benígnej“ je „kardiálna insuficiencia“ pre hypertenziu ako významného vraha civilizovanej spoločnosti pravdivejšia a dôležitejšia. Kardiálna insuficiencia má preto v ICD-10 vlastný kód, a to I59.-, ktorý sa dá odlíšiť pridaním dvoch ďalších číslic v ICD-10.

I11.0 je v skutočnosti kombináciou dvoch diagnóz, ktoré môžu vystupovať aj samostatne. I11.0 nie je navyše izolovaný prípad. Všetky prídavné diagnózy hypertenzie majú svoje vlastné kódy, keď sa vyskytujú s hypertenziou. Situácia tohto druhu, kde dve alebo viaceré jednotlivé diagnózy (sčítance) tvoria jeden kód (súčet), keď sa vyskytujú

spolu, je veľmi častá, a netýka sa len hypertenzie.

6.2 Problém súčet–sčítance

6.2.1 Aké následky má prídanie sčítanca?

Následkom problému súčet–sčítanec strácame nielen informácie periférneho významu. Postihnutý je sám koreňom hierarchie:

Keď sa neuvedie hypertenzná nefropatia s renálnou insuficienciou ako súčet N19 (renálna insuficiencia) + I12.0 (hypertenzná nefropatia) s N19 ako vedúcim kódom a postupujeme v hierarchii nahor, nájdeme „N“ ako prvú číslicu prvého kódu. To znamená že sa nachádzame v kapitole urogenitálne choroby a nie v kapitole kardiovaskulárne choroby, kde očakávame, že sa stretneme s artériovou hypertenziou. Postihnutý je teda koreňom hierarchie. (Renálna insuficiencia sa pokladá za klinicky dôležitejší stav ako hypertenzia, preto má mať prednosť).

6.2.2 Sú klasifikácie vôbec potrebné?

Klasifikácie ICD-10 sú zjednodušením reálneho sveta. Možno sa preto vzdať klasifikácií úplne? Nemáme sa vyjadrovať v pojmoch bohatšej terminológie s jemnejšou granularitou namiesto hrubej granularity jazyka klasifikácií? Je potrebné vôbec kódovanie?

Myslíme si, že kódovanie je potrebné. Keď chceme nájsť knihu v knižnici, musí byť kniha najprv kategorizovaná. Keď chceme študovať novú oblasť, musíme získať najprv prehľad o tejto oblasti, čo v podstate znamená zjednodušenie. Naše mysle nie sú schopné pojať všetky informácie reálneho sveta. Redukcia informácií sa teda môže pokladať za hlavnú úlohu ľudskej inteligencie.

Keď chceme porovnávať dva systémy, musíme zoskupiť ich príslušné prípady a to znamená zjednodušenie a klasifikáciu. Štatistické porovnanie, Napr. rozličných terapeutických režimov v rozdielnych skupinách pacientov v medicíne nie je možné bez utvorenia skupín.

Vo všeobecnosti sa s problémom redukcie informácií stretávame vždy, keď sa pohybujeme od úrovne s jemnou granularitou k úrovni s hrubou granularitou alebo od reality k terminológii. Redukcia informácií sa vyskytuje vo všetkých týchto prechodoch a uvedené problémy sú následkom každej redukcie informácií vôbec, nielen v medicíne.

Keby nebolo ťažké získať klasifikované informácie z terminologických informácií, klasifikované informácie by boli samozrejme a neprinášali by pridanú hodnotu. Klasifikácia znamená teda informačný prínos.

Uvedený problém súčtu–sčítancov sú **vzťahy ČASŤ...** (*part-of*), ktoré nielen pripájajú atribúty k diagnózam, ale aj diagnózy samé. V takých prípadoch môže byť diagnóza sčítancom iného súčtu diagnóz alebo nespojená diagnóza ako taká. To má za následok klasifikačný zmätok.

Schulz a spol. (2005) uvádzajú ten istý príklad s **výkonmi**. Analyzovali súčtový výkon „incízia žalúdka + odstránenie cudzieho telesa“ a jeho znázornenie v SNOMED CT. Aj tu sa kritické vzťahy spájajú s vrcholovými rámcovými entitami (výkonmi) vo forme komplexov súčet/sčítance, ako je to v prípade vrcholových rámcových entít (diagnóz) vo vyššie uvedenom príklade. Takáto situácia je najzraniteľnejšia, keď ide o počet a identitu vrcholových entít.

Schulz diskutuje o znázornení takýchto prípadov v systéme SNOMED CT. Vzťahy nestačia na modelovanie zložitej situácie, preto sa na štruktúrovanie súboru vzťahov použila množina vzťahov. Stačí poskytnúť vzťahy ako súbor a ako zložito môže byť súprava štruktúr organizovaná? Štruktúrovanie vzťahov je kľúčovým krokom pri riešení problému súčtu/sčítancov v praxi. S cieľom usporiadať komplexy rozličných vzťahov v dobre definovanej štruktúre sa utvorili molekuly pojmov, a tak sa dali použiť na riešenie problému súčtu/sčítancov.

7 Závěry

- Opis hrubšej granularity v zložitej tematickej oblasti sa nedá získať z opisu oblasti s jemnejšou granularitou.
- Kódy ICD-10 sa nedajú získať z informácií o konkrétnom pacientovi vo forme štandardizovaných termínov bez explicitných prídavných pravidiel.
- V každom kódujúcom systéme treba venovať pozornosť problému súčet–sčítance.
- Pri strojovom kódovacom systéme platia tie isté pravidlá.

Klasifikácie sú potrebné, aj keď štandardizované terminológie poskytujú presnejšie a podrobnejšie informácie.