

**პროგრამული აკლიკაციების მონაცემთა ბაზების დავრომქტება
პატემორიალური მიდგომით**

გია სურგულაძე, ნინო თოფურია, მარინე კაშიბაძე, დავით სალინაძე
საქართველოს ტექნიკური უნივერსიტეტი

რეზიუმე

განიხილება დავროგრამების UML ტექნოლოგიის და ობიექტ-როლური მოდელირების ORM ინსტრუმენტის გამოყენებით არსთა-დამოკიდებულების ERM მოდელის ავტომატიზებულ რეჟიმში აგების ამოცანა. კატეგორიალური მიდგომისა და ლოგიკურ-ალგებრული მეთოდების გამოყენებით განხორციელებულია საპრობლემო სფეროს მონაცემთა სტრუქტურებისა და ცოდნის ობიექტ-ორიენტირებული მოდელირება.

საკვანძო სიტყვები: მონაცემთა ბაზა.. UML. ORM. ERM. კატეგორიალური მიდგომა. ლოგიკურ-ალგებრული შეზღუდვები.

1. შესავალი

ინფორმაციული ტექნოლოგიების განვითარების თანამედროვე ეტაპზე განსაკუთრებით აქტუალურია სამეცნიერო და საინჟინრო სამუშაოების შესრულება განაწილებული მართვის ავტომატიზებული სისტემების დასაპროექტებლად და სარეალიზაციოდ ობიექტ-ორიენტირებული მიდგომის საფუძველზე. უნიფიცირებული მოდელირების ენის (UML) ინსტრუმენტების გამოყენებით მიიღწევა დავროექტებისა და დავროგრამების ურთულეს პროცესთა ავტომატიზაცია, რაც საგრძობლად ამცირებს საპროექტო-საინჟინრო ეტაპების შესრულების დროს და ამალღებს საბოლოო პროდუქტის, მართვის სისტემის ხარისხს [1,2].

საუნივერსიტეტო განათლების სისტემა კომპლექსური და მეტად მნიშვნელოვანი ობიექტია მართვის პროცესების შემდგომი სრულყოფის თვალსაზრისით ახალი საინფორმაციო ტექნოლოგიების ბაზაზე [4]. აგრეთვე, ჩვენი ქვეყნის მრავალი დარგი, კერძოდ ენერგეტიკა, მრეწველობა, სოფლის მეურნეობა, ჯანმრთელობის დაცვა და ა.შ. ინფორმაციკისა და მართვის ამოცანების გადასაჭრელად ფართოდ იყენებს რელაციური ბაზების მართვის სისტემებს, რომელთა შემდგომი სრულყოფა მათემატიკური მოდელირების სემანტიკური პრობლემების აღმოსაფხვრელად და ბაზების სტრუქტურების ავტომატიზებულ რეჟიმში დასაპროექტებლად – მეტად მნიშვნელოვანი საკითხია.

მონაცემთა რელაციური მოდელი ფორმალური მათემატიკური ობიექტია და მისი საშუალებით საგნობრივი სფეროს არაფორმალური თვისებების ასახვას მივყავართ რთულ სემანტიკურ პრობლემამდე. იგულისხმება სემანტიკის (შინარსის) მათემატიკური მოდელირების სპეციფიკური პრობლემები. მათი სირთულე ძირითადად განისაზღვრება მონაცემთა ბაზის რეორგანიზაციის (განახლების) პროცედურების სირთულით, სტატიკური და დინამიკური შეზღუდვების (პრედიკატების) სისწორის შემოწმებით. შეზღუდვები, გამოსატავს მონაცემთა ბაზის მდგომარეობის (რელაციები)

ზოგად, აბსტრაქტულ თვისებებს ანუ მონაცემთა ბაზის სემანტიკას. შეზღუდვების გამოხატვის ყველაზე ბუნებრივი ხერხია გამონათქვამები I-რიგის პრედიკატების ენაზე, რომელიც განსაზღვრავს მონაცემთა ბაზის დასაშვებ მდგომარეობათა სიმრავლეს. ამ გამონათქვამების ერთობლიობა მოდელიდან ამოსარჩევი ელემენტის სახელებთან ერთად (ატრიბუტების დასახელება) ქმნის მონაცემთა ბაზის ე.წ. სტატიკურ სქემას. მონაცემთა ბაზა, გადადის რა ერთი მდგომარეობიდან მეორეში (რელაციური ცვალებადობა), აღწერს მონაცემთა მოდელში გარკვეულ ტრაექტორიას. ყოველი მომდევნო მდგომარეობა შეიძლება დამოკიდებული იყოს მის წინა მდგომარეობაზე. ამ მდგომარეობებს შორის კავშირები აღიწერება მთლიანობის დინამიკური შეზღუდვებით (პრედიკატებით), რომლებიც ქმნის მონაცემთა ბაზის ე.წ. დინამიკურ სქემას.

მთლიანობის პრედიკატების ცოდნის საფუძველზე, რომლის ერთ-ერთ მნიშვნელოვან ნაწილს ფუნქციონალური დამოკიდებულებები შეადგენს, ხდება სინამდვილის (საგნობრივი სფეროს) მოდელირება ნორმალურ ფორმათა თეორიის საშუალებით. ფორმირდება მონაცემთა ბაზის სქემა, რომელიც ამ დამოკიდებულებათა სასრულ ერთობლიობას შეადგენს. მაგრამ ფუნქციონალური დამოკიდებულებები ასახავს სინამდვილის შინაარსობრივი მხარის მხოლოდ ნაწილს; მონაცემთა ბაზის მოდიფიკაცია კი სემანტიკური ბუნების მატარებელია. აქედან გამომდინარე, არაა ცხადად წარმოდგენილი, თუ როგორ შეიძლება დამოკიდებულებათა ნორმალიზაციის სინტაქსური პროცედურებით გადაწყდეს სემანტიკური პრობლემები.

საგნობრივი არის სტრუქტურისა და კანონზომიერებების მათემატიკური მოდელირებისას რელაციურ მოდელში გამოიყენება სიმრავლეთა თეორიის ელემენტები. ამიტომ ისმის კითხვა, რამდენად შეესაბამება მონაცემთა ბაზაში ფორმალურად ასახული ცნებები სიმრავლეთა თეორიის ცნებებს (ტერმინოლოგიის მათემატიკური სიზუსტის თვალსაზრისით).

რელაციური ბაზა გამოიხატება დამოკიდებულებათა ქვესქემების ერთობლიობის საშუალებით, ე.ი. გვაქვს რელაციების (ცხრილების) ერთობლიობა, რომლებშიც ასახულია ობიექტი (R_i), მათი თვისებები (A_j), კონკრეტული ურთიერთკავშირებით (Z_k). კონკრეტული საგნობრივი სფეროდან გამომდინარე, რელაციური მოდელის აგების დროს ობიექტები, თვისებები და მნიშვნელობები შინაარსობრივად ურთიერთშენაცვლებადია, ე.ი. არ არსებობს მოდელირების საბოლოო ცალსახა ალგორითმი ამ საკითხის გადასაწყვეტად. ამასთანავე, რელაციური ბაზის მოდელში არაა ასახული ობიექტებს შორის კავშირების სემანტიკა. ხშირად კი, შინაარსობრივად ერთი და იგივე თვისება შეიძლება წარმოდგენილ იქნეს სინტაქსურად განსხვავებული ფორმებით (სინონიმებით). სისტემას კი სჭირდება დამატებითი ინფორმაციული საშუალებები, რათა ეს მომენტები გაითვალისწინოს.

ჩვენ შემოვიფარგლეთ ზოგიერთი ძირითადი პრობლემური საკითხების ანალიზით, რომლებიც დამახასიათებელია რელაციული მოდელისათვის. შეიძლება აღინიშნოს, რომ ასეთი პრობლემების

გადაჭრა შესაძლებელია როგორც რელაციური მოდელის შედარებით გართულებით (დამატებითი ინფორმაციის შემოტანით), ასევე პროგრამული და აპრატული საშუალებების სრულყოფით მომავალში.

კატეგორიები, რომლებსაც აქ ვიხილავთ, არის ლოგიკური გრამატიკის შემადგენელი ნაწილი, რომელიც, როგორც ზოგადი ენის საძირკველი, ბევრად მდიდარია, ვიდრე ტრადიციული პრედიკატების ლოგიკა და რომელიც ცხადია დამოუკიდებელი უნდა იყოს ბუნებრივი ენის ემპირიული გრამატიკისგან. პრედიკატების ლოგიკა არის ლოგიკური გრამატიკა დანართების შინაარსის ასაგებად, მაგრამ ძალზე ღარიბია და ნაკლებად განვითარებული - იგი არ ფლობს სინტაქსურ კატეგორიებს, რათა განსხვავდეს ერთმანეთისგან მთავარი და არამთავარი პრედიკატები. პრედიკატული ლოგიკა უნდა განვიხილოთ, როგორც ლოგიკური გრამატიკის გაფართოება.

2. ძირითადი ნაწილი

ORM არის მოდელირება ფაქტების საფუძველზე, სადაც საპრობლემო არე განიხილება, როგორც ობიექტების ერთობლიობა, რომლებიც თამაშობს განსაზღვრულ როლებს [3]. კონცეპტუალური მოდელირება მიიღწევა არსთა დამოკიდებულების (ER) მოდელირებას, თუმცა იგი შეიძლება გამოვიყენოთ მას შემდეგ, რაც დაპროექტების პროცესი დამთავრებულია. ER-დიაგრამა შორსაა ბუნებრივი ენისაგან, ვერ ხერხდება ამა თუ იმ მოვლენის ფაქტით შევსება, დამალულია ინფორმაცია იმ სემანტიკური დომენების შესახებ, რომლებიც ქმნის მოდელს.

ამრიგად, კონცეპტუალური მოდელირების განვითარებულ ტექნიკას წარმოადგენს ობიექტ-როლური მოდელირება. სწორედ ORM-ს შეუძლია უზრუნველყოს სხვადასხვა პროფესიის ადამიანთა შეთანხმებული მუშაობა, რომელთა მომზადების დონე ინფორმაციული სისტემების დაპროექტების სფეროში შეიძლება მნიშვნელოვად განსხვავდებოდეს.

დაპროექტების აღნიშნული პროცესი რამდენიმე ეტაპს მოიცავს: 1. ORM -ის კონცეპტუალური მოდელირების სქემის პროცედურა ანუ CSDP (*conceptual schema design procedure*) ყურადღებას ამახვილებს მონაცემების ანალიზზე და დაპროექტებაზე. ელემენტარული ფაქტების ფორმირება და მათი აღქვასტურობის შემოწმება; 2. ფაქტების ტიპებისათვის დიაგრამის აგება და სისრულის შემოწმება; 3. იმ ობიექტთა ტიპების შემოწმება, რომლებიც უნდა გაერთანდეს და მათი მათემატიკური წარმომავლობის დაფიქსირება; 4. დაემატოს უნიკალურობის შეზღუდვა და შემოწმდეს ფაქტების ტიპების ოპერანდების რაოდენობა; 5. დაემატოს როლების იძულებითი შეზღუდვები და შემოწმდეს მათი ლოგიკური წარმომავლობა; 6. დაემატოს ელემენტები, სიმრავლეთა შედარება და ქვეტიპის შეზღუდვები; 7. დაემატოს სხვა შეზღუდვები და მოხდეს საბოლოო შემოწმება.

პირველი ეტაპი ყველაზე მნიშვნელოვანია, ვინაიდან აქ ხდება სხვადასხვა სახის ინფორმაციის შეგროვება ბუნებრივ სალაპარაკო ენაზე. ასეთი ინფორმაცია ხშირად არის შემაჯავლი და გამომავალი

ფორმების ან ხელნაწერის სახით. შესაძლებელია მოდელის დამპროექტებელს მოუხდეს უშუალო კონსულტაციის მიღება კლიენტთან, რათა ზუსტად ჩამოაყალიბოს სისტემის მოთხოვნები.

ფაქტი ჩაიწერება შემდეგი ფორმით:

- *f1*: თანამშრომელს ნომრით 25 აქვს გვარი ‘ლოლიძე’

- *f2*: თანამშრომელი ნომრით 17 მუშაობს კონტრაქტით თარიღამდე ‘12.31.08’

თითოეული ფაქტი არის ბინარული დამოკიდებულება ორ ობიექტს შორის. პრედიკატს შეიძლება ჰქონდეს 1,2,3,.. ოპერანდი, თუმცა რადგან პრედიკატი ელემენტარულია 3-4 ოპერანდზე მეტი იშვიათად გვხვდება. უმრავლეს შემთხვევაში პრედიკატი არის ორობითი. ასეთი პრედიკატებისათვის არსებობს ინვერსული პრედიკატი. ისე, რომ ფაქტი შეიძლება წავიკითხოთ ორივე მიმართულებით.

მეორე ეტაპზე ხდება ფაქტების ტიპებისათვის დიაგრამის აგება. ობიექტები გამოისახება ელიფსებით, პრედიკატები მართკუთხედებით, მნიშვნელობის ტიპი წყვეტილი ელიფსით. პრედიკატი იკითხება მარცხნიდან-მარჯვნივ და ზემოდან-ქვემოთ მანამ, სანამ არ შეხვდება ნიშანი “<<”, რომელიც ცვლის წაკითხვის მიმართულებას საწინააღმდეგო მიმართულებით.

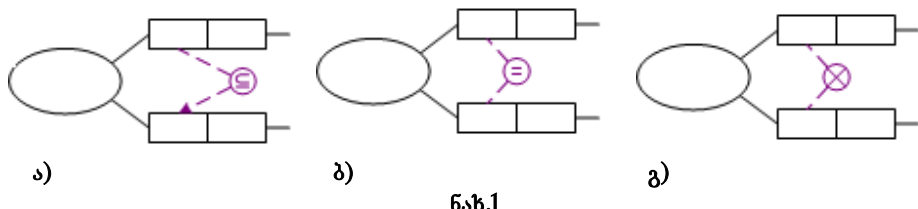
შემდეგ ბიჯებზე ხდება შეზღუდვების დაწესება. ობიექტ-როლური მოდელირება გამოიყენებს: უნიკალურობის, იძულების, წრიული, ქვესიმრავლეთა შედარების, სისშირის, ინდექსურ და მნიშვნელობის სახის შეზღუდვებს.

ახლა განვიხილოთ ობიექტ-როლური მოდელის აგების ტიპური ამოცანები ალგებრულ-ლოგიკური შეზღუდვების გამოყენებით.

2.1. სიმრავლეთა შედარება: ქვესიმრავლის და ტოლობის შეზღუდვები

თუ ორი სხვადასხვა როლი თამაშდება ერთი ობიექტის მიერ, მნიშვნელოვანი ხდება მათი შედარება. მონაცემთა ბაზებში არსებობს შედარების სამი ოპერატორი, რომლებიც ORM-დიაგრამაზე გამოისახება ასეთი სახით:

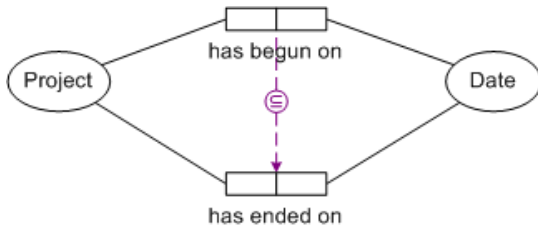
- ქვესიმრავლის შეზღუდვა – Subset constraint (ნახ.1-ა);
- ტოლობის შეზღუდვა – Equality constraint (ნახ.1-ბ);
- წყვილის გამორიცხვის შეზღუდვა – pair-exclusion constraint (ნახ.1-გ)



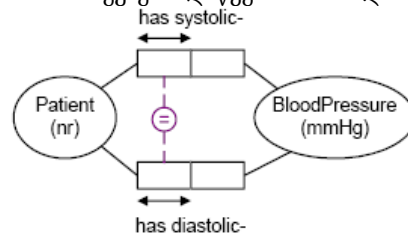
ქვესიმრავლის შეზღუდვა – პირველი როლის ნაკრები არის მეორე როლის ნაკრების ქვესიმრავლე. მაგალითად,
- *f1* : პროექტი (*nr*) დაიწყო განსაზღვრულ დღეს (*mdy*)

- f2 : პროექტი (nr) დამთავრდა განსაზღვრულ დღეს (mdy)

პროექტს ვერ ექნება დამთავრების თარიღი მანამ, სანამ არ ექნება დაწყების თარიღი (ნახ.2).



ნახ.2



ნახ.3

ტოლობის შეზღუდვა გამოიყენება მხოლოდ მაშინ, როდესაც როლები თავსებადია. იგი გვიჩვენებს, რომ ორი როლი პირობითად ტოლია.

- f1 : თუ პაციენტს აქვს არტერიული წნევა, მაშინ მას აუცილებლად ექნება გულის წნევა.

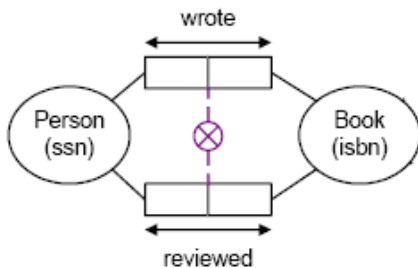
თუ ცნობილია ავადმყოფის არტერიული წნევა, მაშინ ცნობილია გულის წნევაც და პირიქით (ნახ.3).

წყვილის გამორიცხვის შეზღუდვა აღწერს ორი როლის ურთიერთგამორიცხულობას. მაგალითად,

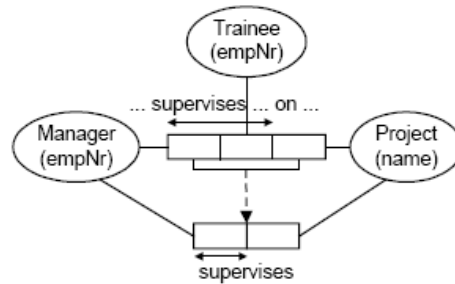
- f1 : პერსონამ (ssn) დაწერა წიგნი (isbn)

- f2 : - პერსონამ (ssn) დაწერა რეცენზია წიგნზე (isbn)

პერსონა საკუთარი წიგნის შესახებ ვერ დაწერს რეცენზიას (ნახ.4).



ნახ.4



ნახ.5

აღვწეროთ ფაქტები:

- f1 : მენეჯერი ხელმძღვანელობს პროექტს (ნახ.5).

- f2 მენეჯერი ასევე აკონტროლებს ტრენინგს ამ პროექტისათვის.

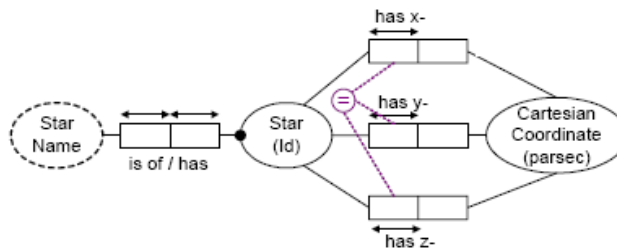
- f3 თითოეული მენეჯერი რომელიც აკონტროლებს ტრენინგს ასევე ხელმძღვანელობს პროექტს.

ა) - f1: ვარსკვლავს აქვს სახელი (ნახ.6).

- f2: ვარსკვლავს აქვს x-დეკარტული კოორდინატი.

- f3 : ვარსკვლავს აქვს y-დეკარტული კოორდინატი.

- f4 : ვარსკვლავს აქვს z-დეკარტული კოორდინატი.



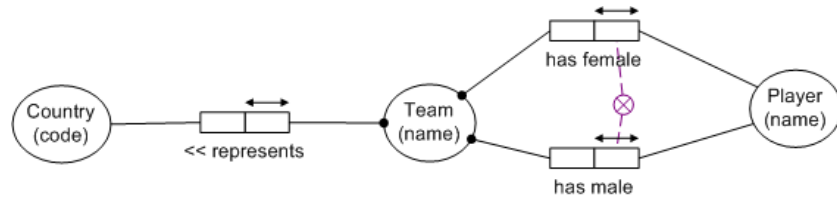
ნახ.6

ბ) - f1: გუნდი წარმოადგენს ქვეყანას (ნახ.7).

- f2: გუნდი შედგება ქალი მოთამაშეებისაგან.

- f3: გუნდი შედგება კაცი მოთამაშეებისაგან.

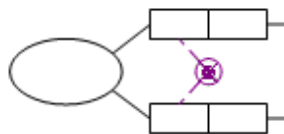
- f4: არ არეობს ისეთი გუნდი, სადაც იქნებინ როგორც ქალი, ისე კაცი მოთამაშეები.



ნახ.7

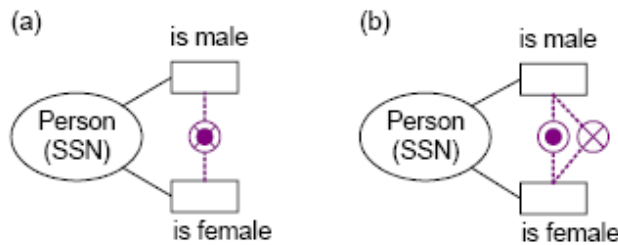
2.2. სიმრავლეთა შედარების შეზღუდვა: იძულების დიზუნქცია

იძულების დიზუნქციის შეზღუდვა – inclusive-or constraint (ნახ.8)



ნახ.8

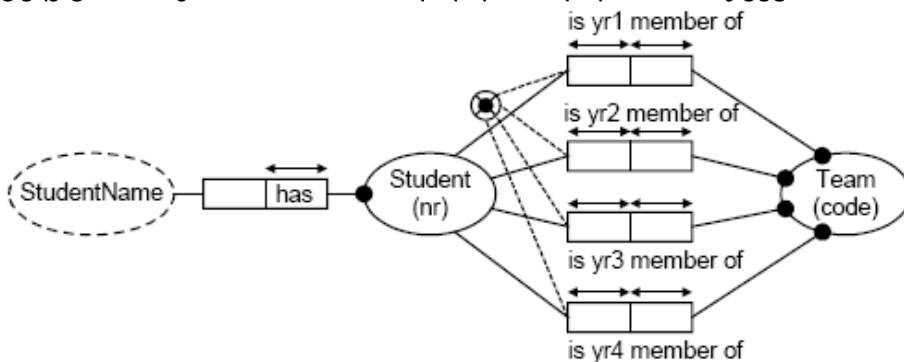
იგი გვჩვენებს, რომ ობიექტი ვალდებულია შეზღუდვაში ასრულებდეს მხოლოდ ერთ როლს და არა ორივეს ერთად. მაგალითად, პერსონა შეიძლება იყოს ქალი ან კაცი და არა ორივე ერთად ORM-დიაგრამაზე გამოისახება ასეთი სახით (ნახ.9).



ნახ.9

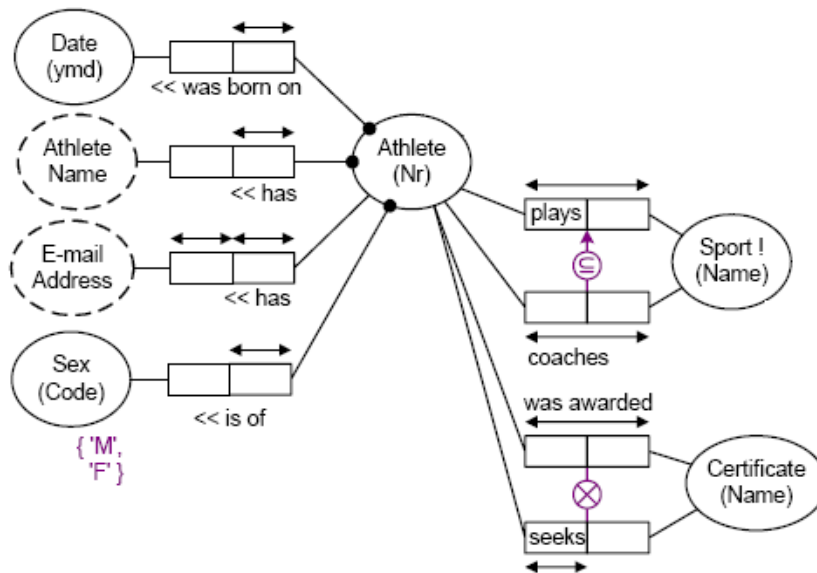
აღვწეროთ შემდეგი ფაქტები:

- ა) - f1 : სტუდენტს აქვს გვარი (ნახ.9).
- f2 : სტუდენტი 'აბაშიძე რ. №1' არის მხოლოდ და მხოლოდ '608836' ჯგუფში.
- f3 : სტუდენტი 'ბაქრაძე დ. №5' არის მხოლოდ და მხოლოდ '608835' ჯგუფში.
- f4 : სტუდენტი 'ღვალა ს. №9' არის მხოლოდ და მხოლოდ '608837' ჯგუფში.
- f5 : სტუდენტი 'იობაძე რ. №35' არის მხოლოდ და მხოლოდ '608839' ჯგუფში.



ნახ.10

- ბ) - f1: ათლეტს აქვს სახელი (ნახ.11).
 - f2: ათლეტი დაიბადა 1967/01/18.
 - f3: ათლეტს აქვს იმეილი.
 - f4: ათლეტს აქვს სქესი მამრ. ან მდედრ. (M.F)
 - f5: ათლეტი მონაწილეობს .. სახეობაში.
 - f6: ათლეტი ვარჯიშობს .. სახეობაში.
 - f7: თითოეული ათლეტი მონაწილეობს მხოლოდ იმ სახეობის შეჯიბრში, რომელშიც ვარჯიშობს.
 - f8: ათლეტს მინიჭებული აქვს სერთიფიკატი.
 - f9: ათლეტი იღწვის სერთიფიკატის მოსაპოვებლად.
 - f10: ათლეტი, რომელმაც უკვე მოიპოვა სერთიფიკატი აღარ მოიპოვებს მას მეორედ.



ნახ.11

3. დასკვნა

მონაცემთა ბაზის სტრუქტურის დაპროექტების პროცესების ავტომატიზაცია ეფექტურად ხორციელდება ობიექტ როლური მოდელირების ინსტრუმენტით, რომლის თეორიულ საფუძველსაც კატეგორიული მიდგომა წარმოადგენს. ლოგიკურ-ალგებრული მეთოდების გამოყენებით ხორციელდება საპრობლემო სფეროს მონაცემთა სტრუქტურებისა და ცოდნის ობიექტ-ორიენტირებული მოდელირება.

ლიტერატურა

1. Booch G., Rumbaugh J., Jacobson I., The Unified Modeling Language User Guide. Addison Wesley Longman, 2004.
2. სურგულაძე გ., ვედეკინდი ჰ., თოფურია ნ. განაწილებული ოფის-სისტემების მონაცემთა ბაზების დაპროექტება და რეალიზაცია UML-ტექნოლოგიით. მონოგრაფია. სტუ, თბილისი. 2006.
3. Halpin T.A., Information Modeling and relational Databases, Morgan Kaufmann Publishers, San Francisco, 2001. www.mkp.com/books_catalog/catalog.asp/ISBN=1-55860-672-6.

4. გოგინაშვილი გ., ფრანგიშვილი ა., სურგულაძე გ. ინფორმატიკა, პროგრამული ტექნოლოგიები და მათი განვითარების და სწავლების თანამედროვე მიმართულებანი. სტუ შრ.კრ.,მას“ №1(2), 2007.

5. Николаишвили В., Сургуладзе Г. Топурия Н., Кашибадзе М. Категориальный подход разработки абстрактных моделей данных для объектно-ориентированных, реляционных баз данных. Тез.Докл.Интерн.Конф. Киев, 2006.

DESIGNING DATABASES OF APPLICATIONS WITH THE CATEGORIAL APPROACH

Surguladze Gia, Topuria Nino, Kashibadze Marine,
Saginadze David
Georgian Technical University

Summary

There is considered the problem of automated construction of ER-model on the basis of UML and ORM. By means of the categorial approach and logical-algebraic methods object-oriented modelling for structures of data and knowledge of Application is carried out.

**ПРОЕКТИРОВАНИЕ БАЗ ДАННЫХ ПРОГРАММНЫХ АППЛИКАЦИЙ
КАТЕГОРИАЛЬНЫМ ПОДХОДОМ**

Сургуладзе Г., Топурия Н., Кашибадзе М., Сагинадзе Д.
Грузинский Технический Университет

Резюме

Рассматривается задача автоматизированного построения ER модели на основе применения UML технологии и ORM инструмента. С помощью категориального подхода и логико-алгебраических методов осуществляется объектно-ориентированное моделирование структур данных и знания проблемной области.