

ვიდეოთამაშებში მულტიაგენტური დასჯავლის მეთოდების შესახებ

გიორგი აბელაშვილი, ზურაბ ბოსიკაშვილი
საქართველოს ტექნიკური უნივერსიტეტი

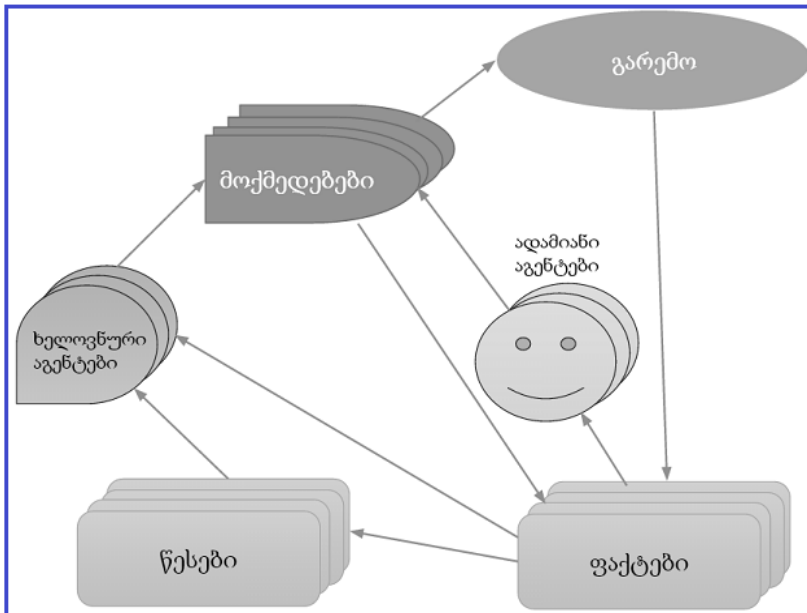
რეზიუმე

მულტიაგენტური ტექნოლოგია ყოველთვის ფართოდ გამოიყენებოდა კომპიუტერულ ვიდეოთამაშებში. იგი აერთიანებს თამაშების თეორიის, რთული სისტემების, გამოთვლითი სოციოლოგიისა და ხელოვნური ინტელექტის საკითხებს. არსებობს ისეთი სისტემები, რომელთა აგება ძალზე რთული ამოცანაა. ერთერთი მთავარი საკითხი ასეთ ამოცანებში არის აგენტის მიერ გარკვეული ცოდნის დასწავლა, ასევე ამ დასწავლილის სხვა აგენტისათვის მიწოდება - ცოდნის გაზიარება. ამ სტატიაში განვიხილავთ ცოდნის დასწავლის ერთერთ მეთოდს, ცოდნის მარკირების ამოცანას. ცოდნის მარკირების ამოცანა მოცემული იქნება კონკრეტული მაგალითით - ჭიანჭველების ამოცანით.

საკვანძო ტერმინები: ხელოვნური ინტელექტი. ინტელექტუალური აგენტი. მულტიაგენტური ტექნოლოგია. ცოდნის მარკირების ამოცანა. პროგრამული აგენტები.

1. შესავალი: ცოდნის ტიპები და ინტელექტუალური აგენტის ზოგადი მოქმედება

ვიდეოთამაშებში ადამია-მოთამაშებთან ერთად მონაწილეობას ღებულობს ხელოვნური (პროგრამული) აგენტები, რომლებიც მოთამაშესთან ერთად მოქმედებს დინამიკურ გარემოში, ფაქტებზე და წესებზე დაყრდნობით ღებულობს ინდივიდუალურ ან კოლექტიურ გადაწყვეტილებებს, ასევე აგენტები ასრულებს გარკვეულ ქმედებებს ინდივიდუალურად ან კოლექტიურად (ნახ.1).

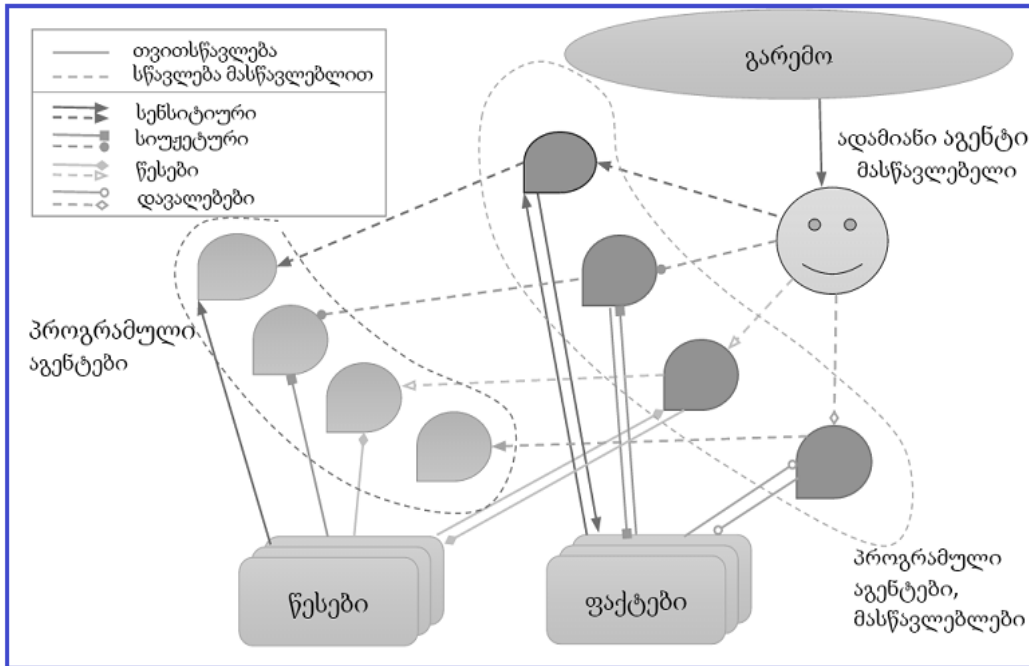


ნახ.1. ადამიანისა და ხელოვნური აგენტის მოქმედება

მოქმედებისთვის პროგრამულ აგენტებს უნდა გააჩნდეთ შესაბამისი ცოდნა. ეს ცოდნა აგენტებში შეიძლება ჩადებული იყოს სტატიურად კოდის სახით, ან მიეწოდებოდეს დინამიურად წესებისა და ფაქტების საშუალებით. რადგანაც აგენტები შეიძლება მოქმედებდნენ შემთხვევითად, ცვლად გარემოში, ასევე ადამიანი აგენტები (მოთამაშეები) შეიძლება იქცეოდნენ არადეტერ-

მინირებულად, შესაბამისად ეს ხელოვნური აგენტები აღჭურვილი უნდა იქნას სწავლების მექანიზმებით. სწავლება შეიძლება იყოს მასწავლებლის ან მასწავლებლის გარეშე (თვითსწავლება). აგენტები შეიძლება მოიპოვებდეს/უზიარებდეს სხვა აგენტებს შემდეგი სახის ცოდნას (ნახ.2):

- სენსიტიურ ცოდნას;
- ეპიზოდურ ცოდნას;
- ცოდნას წესების სახით;
- ცოდნას დავალებათა ბმულების სახით.



ნახ.2. სწავლების ტიპები და ნაკადები.

აგენტები შეიძლება ახორციელებდეს როგორც თვითსწავლებას, ასევე სხვების სწავლებას. სენსიტიური ცოდნა მოიცავს ინფორმაციას ობიექტების ადგილმდებარეობის, ზომის (რამდენი სიმბოლოსაგან შედგება) შესახებ. ეპიზოდური ცოდნა მოიცავს კონკრეტული ეპიზოდის აღწერას, ამ ეპიზოდში გამოყენებულ მოქმედებას და შედეგს. წესები განსაზღვრავს მოქმედების პირობებს ხოლო დავალებათა ბმულები აგენტების მიერ სხვა აგენტებზე გაცემულ დავალებებს (მიზანი, რესურსი).

სწავლების ზოგადი მიზანია კოოპერაციაში მყოფი აგენტების ისეთი მოქმედებების შერჩევის მოდელის შედგენა, რომელიც განაპირობებს აგენტების რაციონალურ ქმედებებს (სასარგებლოა როგორც თვითონ აგენტისთვის, ასევე კოოპერაციაში მყოფი სხვა აგენტებისთვისაც).

2. ცოდნის მარკირებისა და დასწავლის ამოცანა

ცოდნის მარკირება - არის გარკვეული სწავლების რაღაც ნიშნით გარკვეულ ადგილზე შენახვა, მარკირებული ნიშნის მიხედვით აგენტი პირდაპირი გზით სწავლობს შესაბამის მარკირებულ ფაქტს - იქნეს შესაბამის ცოდნას.

ვთქვათ, აგენტი x_i ღებულობს გარკვეულ ინფორმაციის X სიმრავლეს (ინფორმაცია შეიძლება იყოს ნებისმიერი სახის, სენსიტიური, ეპიზოდური და ა.შ.)

$$X = \{X_1, X_2, X_3, \dots, X_n\}$$

შესაბამისი X_i ინფორმაციას დისკრეტულ ორგანზომილებიან კოორდინატთა სისტემაში შეგვიძლია შევუსაბამოთ შესაბამისი $\{x_i, y_i\}$ (0) წერტილი, ანუ ეს არის აგენტის არსებული მდებარეობა X ინფორმაციის მიღების დროს.

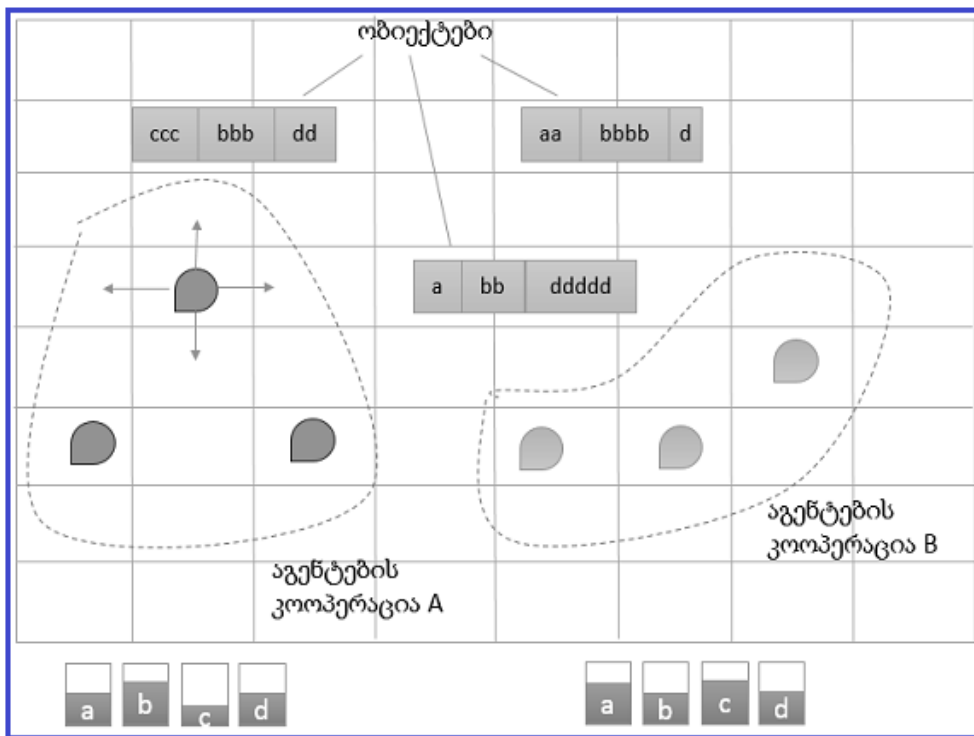
ნებისმიერი X_i აგენტი ნებისმიერი Y_k (0 ინფორმაციას გადაამუშავებს და ცოდნის ბაზაში შეინახავს, დაისწავლის, ხოლო ამ დასწავლილ შედეგს Y_k , იგი იმავე კოორდინატზე $\{x_k, y_k\}$ დატოვებს, გაუკეთებს მარკირებას. ყოველი შემდეგი აგენტი, რომელიც მივა სიბრტყის $\{x_k, y_k\}$ კოორდინატზე, იქ დახვდება x_k აგენტის მიერ დატოვებული y_k ცოდნა და ის ამ ცოდნას პირდაპირი გზით შეინახავს ცოდნის ბაზაში.

თამაშებში მნიშვნელოვანია ცოდნის გადამუშავებისა და შენახვის სისწრაფე ცოდნის ბაზაში, ნებისმიერი აგენტისთვის ცოდნის მოპოვება დამოუკიდებელი ამოცანაა, იგი დაკავშირებულია კომპიუტერულ რესურსებთან და, შესაბამისად, დროსთან, ცოდნის პირდაპირი გზით გადაცემით აგენტს დაწავლის სისწრაფე იზრდება.

3. ჭიანჭველების ამოცანა

შემდგომში ყველგან ვიგულისხმებთ, რომ პროგრამული აგენტები დელიბერაციულია და მხოლოდ სიმბოლების სამყაროში მოქმედებს.

ჩამოვყალიბოთ ჭიანჭველა-აგენტების ამოცანა. ჭიანჭველა-აგენტები და ობიექტები ორ განზომილებიან, უჯრებად დაყოფილ, შემოფარგულ სიბრტყეზე განთავსებული და შეუძლია ამ სიბრტყეში მოძრაობა მხოლოდ ვერტიკალური და ჰორიზონტალური მიმართულებებით (ნახ.3).



ნახ.3. ჭიანჭველების ამოცანის გრაფიკული პროტოტიპი

აგენტები ასევე შეიძლება გარკვეულ კოოპერაციაში იყოს გაერთიანებული. სიბრტყეზე შემთხვევითად ჩნდება ობიექტები (ფოთლები), რომლებიც წარმოადგენს შეზღუდული ანბანის შედგენილ მიმდევრობათა მიმდევრობას. აგენტს შეუძლია შემდეგი მოქმედებების განხორციელება:

- შეზღუდული ზომის ფოთლის გადაადგილება;
- ფოთლის დაჭრა იმ სიმბოლოზე რომელსაც იცნობს;
- გამოყოფილი სიმბოლოს შენახვა საცავში;
- თამაშის მიზანია გარკვეულ დროში კოოპერაციაში მყოფმა ჭიანჭველმა - აგენტებმა, რაც

შეიძლება მეტი სიმბოლოები დაავროვოს. აგენტები თავდაპირველად შემთხვევით პოზიციებში თავსდება და შეუძლია ერთდროულად მოძრაობა. ასევე კოოპერაციაში მყოფმა აგენტებმა შეიძლება მოქმედებები ერთდროულად შეასრულოთ, მაგალითად, სამმა აგენტმა გადაიტანოს დიდი ზომის ფოთლოვანი.

კონკრეტულ ამოცანაში მოცემულია რთული დავალება, რომელიც ინდივიდუალური აგენტის მიდგომით რთული ამოსახსნელია, საჭიროა აგენტთა კოოპერაცია. აგენტთა კოოპერაციაში უმნიშვნელოვანესი როლი უჭირავს ცოდნის გაზიარებას აგენტებს შორის, რათა მიზნის მიღწევა მოხდეს უფრო მარტივად. ცოდნის გაზიარების ერთერთი ეფექტური მეთოდია ცოდნის მარკირების ამოცანა.

4. ცოდნის მარკირება ჭიანჭველების ამოცანის მაგალითზე

კონკრეტულ მაგალითში, ვთქვათ, ნებისმიერ ჭიანჭველას კოლონიაში გააჩნია თავისი როლი:

მზვერავი – ჭიანჭველა, რომელიც ეძებს შესაბამისი ზომის ფოთლებს, შესაბამისი ფოთლის პოვნის შემთხვევაში იგი ამავე ფოთლის შესაბამის კოორდინატზე - $\{x_i, y_i\}$ ტოვებს ცოდნას, ფაქტს ფოთლის ვარგისიანობაზე, ახდენს ძებნის ინფორმაციის მარკირებას.

ფოთლების მჭრელი – ჭიანჭველები, რომლებსაც აქვს საწყისი ცოდნა სიმბოლოებზე, ანუ ისინი ცნობს მხოლოდ მარტივ სიმბოლოებს, თუ ფოთოლი შეიცავს ამ სიმბოლოს ეს ჭიანჭველები შესაბამისი ფოთლის ნაწილს მოაჭრიან.

გადამზიდავები - ჭიანჭველები, რომლებიც ეძებს უკვე ფოთლებისგან ჩამოჭრილ სიმბოლოებს და მიაქვთ საცავში.

მზვერავები მიდის სიბრტყეზე და ეძებს შესაბამის ვარგისიან ფოთლებს. ფოთლის ამოსაცნობად აგენტები იყენებს საწყის პარამეტრებს ცოდნის ბაზიდან. ფოთლის პოვნის შემთხვევაში მზვერავი ტოვებს კვალს, შესაბამისი ფოთლის კოორდინატზე, კვალი არის მარკირებული ინფორმაცია ფოთლის ვარგისიანობაზე, მზვერავი აგრძელებს გზას შემდეგი ფოთლის საძებნელად.

ფოთლის მჭრელი ჭიანჭველებიც მზვერავების პარალელურად მოძრაობს იმავე სიბრტყეზე, მათი მიზანია იპოვონ მზვერავის მიერ დატოვებული მარკირებული კვალი, რომელიც ფოთლის ვარგისიანობაზე მიუთითებს. მათ ამის შესახებ აქვთ წინასწარ გაწერილი ცოდნა ცოდნის ბაზაში. თუ მჭრელმა იპოვა მზვერავის ნიშანი, ესე იგი მას ვარგისიანი ფოთოლი უპოვია. კვალის პოვნის შემთხვევაში მჭრელი ამოწმებს მის მოსვლამდე იყო თუ არა მოსული სხვა, მსგავსი მჭრელი, რომელიც მის მსგავსად, იმავე სიმბოლოს ცნობს. თუ წინამორბედმა ვერ იპოვა საერთო სიმბოლო, იგი ამის შესახებ ნიშანს დატოვებდა, სწორედ ამ ნიშნის მიხედვით ყოველი შემდეგი მოსული იქცევა შემდეგნაირად, თუ ფოთოლში არ არის სიმბოლო და ეს მან ნიშნით გაიგო, გზას გააგრძელებს, თუ არა, მოძებნის თავის შესაბამის ამოცნობად სიმბოლოს, თუ იპოვა მოჭრის მას, თუ ვერა დატოვებს ნიშანს რომ შესაბამისი მისი სიმბოლო ფოთოლში ვერ მოიძებნა, ანუ ის

აგენტი იქნება პირველი ვინც შესაბამის ფოთოლთან მივიდა. ეს უკანასკნელი ცოდნა უკვე ყოველ შემდეგ მომსვლელ მჭრელს გამოადგება შესაბამისი მარკირების ნიშნით.

გადამზიდავების მოქმედება არის ბოლო ეტაპი, ისინი ეძებენ თითო სიმბოლოიან ფოთლებს და გადააქვთ საცავში.

ზემოთაღწერილი მაგალითი დამყარებულია აგენტთა მოქმედებაზე, რომლებიც ინდივიდუალურად მოქმედებს, თუმცა ერთმანეთს ცოდნას უზიარებენ და ამით ისინი ერთმანეთს აძლიერებენ.

კონკრეტული ამოცანა შეიძლება განვიხილოთ, როგორც ჯგუფურად მოქმედ აგენტთა ამოცანა, როდესაც გვაქვს ჭიანჭველათა ჯგუფი (რამდენიმე მძებნელი, რამდენიმე მჭრელი, და ა.შ.). შესაძლებელია ქცევის ცალსახა ცვლილებები ვიხილოთ, მაგალითად, მჭრელებმა შეიძლება უფრო რთული ფოთლები დაჭრან ერთად, გადამზიდებმა უფრო დიდი ფოთლები წაიღონ საცავში და მჭრელებმა მერე იქ დაჭრან. ეს გადაწყვეტილებები დამოკიდებულია უკვე გადაწყვეტილებათა სირთულეზე და დასკვნის გამოტანაზე, შესაბამისად განისაზღვრება ამოცანის ამოხსნის წარმალობა და დროში შესრულებული დავალებები.

[1] ნაშრომში აღწერილია აგენტთა გაძლიერებითი სწავლება და სწავლითი ალგორითმი, სადაც თვითოეულ ჯერზე გადაწყვეტილების მიღება დაფუძნებულია ჯილდოზე. ჯილდო პირობითად არის მიზნის მიღწევის შედეგად მიღებული სარგებელი, ეს სარგებელი არის კონკრეტულ თამაშზე დამოკიდებული (პირობითად, ჯილდო შეიძლება იყოს თამაშის მოგება).

თუ ჩვენ გვაქვს არსებული მდგომარეობა და ხელმისაწვდომი მოქმედება, მაშინ დასწავლითი ალგორითმი აირჩევს ყოველ -ს, რომლის ალბათობითი მნიშვნელობა მოიცემა ბოლცმანის ტოლობით [2]:

$$p(a_i|x) = \frac{e^{Q(x,a_i)/T}}{\sum_{\text{actions}} e^{Q(x,a_i)/T}} \quad (1)$$

სადაც T არის ტემპერატურის პარამეტრი და ის განსაზღვრავს გადაწყვეტილებათა მიღების შემთხვევითობას. აგენტი იღებს გადაწყვეტილებას და იღებს ჯილდო r შემდეგ კი გადადის ახალ მდგომარეობა y . ყოველ ჯერზე აგენტი ანახლებს $Q(x, a)$ - რეკურსიულად, უმატებს პოზიტიურობის წონას β და აკლებს სამომავლო რესურსებს:

$$Q(x, a) \leftarrow Q(x, a) + \beta(r + \gamma V(y) - Q(x, a)) \quad (2)$$

აქ $\gamma (0 \leq \gamma < 1)$ არის კლებადობის პარამეტრი და $V(x)$ მიიღება შემდეგი გამოსახულებით:

$$V(x) = \max_{b \in \text{actions}} Q(x, b) \quad (3)$$

5. დასკვნა

სტატიში არსებულმა კვლევამ დაგვანახა, რომ თამაშებში არსებული ხელოვნური ინტელექტის ამოცანების ერთ-ერთი უმნიშვნელოვანესი საკითხია ინტელექტუალური აგენტის ცოდნის დასწავლის მეთოდები, ეს შეიძლება იყოს ინდივიდუალურად ან ჯგუფურად. ინდივიდუალური აგენტის მთავარი პრობლემა არის ის, რომ იგი სხვა აგენტის ცოდნას აიგნორებს, ამ შემთხვევაში კი ამოცანის წარმალობა კლებულობს და შესრულებადობის დროც იზრდება.

ცოდნის მარკირების ამოცანით დავინახეთ, რომ ინდივიდუალურ აგენტებს შორის დასწავლა ვითარდება და საერთო ცოდნაც იზრდება, საერთო ცოდნის გაზრდის დროს კი თვითოეული

აგენტის შესაძლებლობაც იზრდება. მარკირების ამოცანის მოდელი შეგვიძლია სხვა კონკრეტულ ამოცანებსაც მოვარგოთ, ამისათვის საჭიროა შევქმნათ ხელოვნური ინტელექტის პროგრამული ძრავი, რომელიც უკვე სამომავლო კვლევის საგანია.

ლიტერატურა:

1. Ming Tan. Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents. GTE Laboratories Incorporated 40 Sylvan Road Waltham, MA 02254 tan@gte.com
2. Boltzmann Distribution - https://en.wikipedia.org/wiki/Boltzmann_distribution

METHOD FOR MULTI-AGENT LEARNING IN COMPUTER GAMES

Abelashvili Giorgi, Bosikashvili Zurab
Georgian Technical University

Summary

Multiple-agent technology has always been widely used in computer games. It combines game theories, compound systems, applied sociologies and artificial intelligence. There are systems which are hard to create. One of the case in such system is to make agent learn something, and afterwards to give this kind of knowledge to another agent - making agents share the knowledge. In this article, we will review the method of making agents share their knowledge - problem of knowledge marking. Problem of knowledge marking will be shown as an example of Ants Colony.

МЕТОД МНОГОАГЕНТНЫХ ИССЛЕДОВАНИЙ В КОМПЬЮТЕРНЫХ ИГРАХ

Абелашвили Г., Босикашвили З.

Грузинский Технический Университет

Резюме

Мультиагентная технология всегда широко используется в компьютерных играх. Она сочетает в себе игровые теории, составные системы, применяемые социологии и искусственный интеллект. Существуют такие системы, построение которых является очень сложной задачей. В таких случаях одним из важнейших вопросом является обучение агентом определенных знаний, а также предоставление этих знаний другим агентам. В статье рассматривается один из методов обучения-задача маркировки знаний. Задача маркировки знаний дана на конкретном примере-задача муравьев.