

საქართველოს ტექნიკური უნივერსიტეტი

რევი ჩოგოვაძე, რამაზ ხუროძე

ნეიროინფორმატიკის საფუძვლები

I ნაწილი



დამტკიცებულია სტუ-ს
სასწავლო-მეთოდური
საბჭოს მიერ

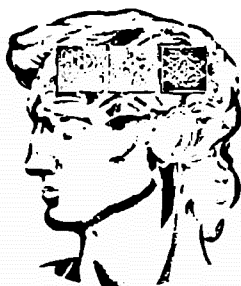
თბილისი
2005

GEORGIAN TECHNICAL UNIVERSITY

Revi Chogovadze, Ramaz Khurodze

The Principles of the Neuroinformatics

Part I



*Approved by the Educational
Methodical Council of the GTU*

**Tbilisi
2005**

განხილულია ხელოვნური ნეირონული ქსელების თეორიის ძირითადი მიმართულებები. ავტორები ეყრდნობიან ქართულ და საზღვარგარეთის შესაბამის ლიტერატურულ წყაროებში ასახულ სამეცნიერო და სასწავლო-მეთოდოლოგიურ მასალებსა და საკუთარ ნაშრომებს ნეირონულ ქსელებში, გამოთვლით სისტემებსა და სახეთა ამოცნობაში. აღწერილი ხელოვნური ნეირონული ქსელების უმეტესობის რეალიზება შესაძლებელია პერსონალურ კომპიუტერზე.

გათვალისწინებულია ასპირანტების, მაგისტრანტებისა და ბაკალავრიატის სტუდენტებისათვის, გამოადგებათ სხვადასხვა დარგის სპეციალისტებსაც, რომელთაც აქვთ სათანადო მომზადების დონე და ნეიროკომპიუტერული ტექნოლოგიების გამოყენების დაინტერესება.

The presented work outlines basic principles of the artificial neural network theory. The authors' findings are based upon their own research in neural networks, high parallelism systems and pattern recognition, as well as upon scientific, educational and methodological materials available in relevant foreign resources.

Knowledge of the presented neural informatics theory will enable IT professionals to adapt most artificial neural networks described in the work to personal computers.

This work is recommended to Bachelor, Master and Graduate students. Additionally, other professionals, interested in introducing this promising technology in their own field of expertise, will find the work useful.

რეკენზენტები: აკად. მინდია საღუქვაძე,
პროფ. ოთარ ვერულავა

წინასიტყვაობა

ნეიროინფორმატიკა ხელოვნური ინტელექტის დისციპლინის შემადგენელი ნაწილია. ხელოვნური ნეირონული ქსელების თეორია და პრაქტიკა ინტენსიური განვითარების სტადიაში იმყოფება, ნეიროქსელები მნიშვნელოვნად აფართოებენ გამოთვლების ცნებას, რაც განაპირობებს მათდამი ინტერესის გაზრდას ბოლო წლებში.

ნაშრომში სისტემატიზებულია ხელოვნური ნეიროქსელების სამეცნიერო და სასწავლო-მეთოდოლოგიური მასალები. ავტორები ეყრდნობოდნენ როგორც საკუთარ გამოკვლევებს ნეიროინფორმატიკაში, სახეთა ამოცნობასა და მაღალი პარალელურობის გამოთვლით სისტემებში, ასევე საზღვარგარეთის მრავალფეროვან პუბლიკაციას. ამ მასალებზე დაყრდნობით მომზადდა და პირველად საქართველოში, ტექნიკური უნივერსიტეტის ხელოვნური ინტელექტის დეპარტამენტში, იკითხება ნეიროინფორმატიკის ლექციების კურსი.

ხელოვნური ნეირონული ქსელების თეორიაში გამოყენებული ცნებების შესაბამისი ტერმინოლოგია როგორც ქართულ, ასევე სხვა ენებში საბოლოოდ არ არის ჩამოყალიბებული, ამიტომ სათანადო ყურადღება ექცევა რიგი განსაზღვრებების დაზუსტებას.

ნაშრომის ზოგიერთი ასპექტის დახვეწას ხელი შეუწყო მისმა განხილვამ პროფესორ ოთარ ვერულავას ხელმძღვანელობით ხელოვნური ინტელექტის დეპარტამენტში ჩატარებულ სემინარებზე, რომელშიც აქტიურად მონაწილეობდნენ დოცენტი თეა თოდუა, ინჟინერ-პროგრამისტები: გიორგი ვერულავა, დავით მიქაძე.

ყველა საკმიან შენიშვნასა და წინადადებას ავტორები კმაყოფილებით მიიღებენ და გაითვალისწინებენ “ნეიროინფორმატიკის საფუძვლების” მეორე ნაწილის გამოცემაში.

თავი I. შესავალი

1.1. ნეიროინფორმატიკის იდეოლოგიის წყაროები

ხელოვნური ნეირონული ქსელებისადმი ინტერესი განსაკუთრებით გაიზარდა ბოლო ათწლეულში, მიუხედავად იმისა, რომ პირველი ნაშრომები ამ მიმართულებით მეოცე საუკუნის შუა წლებში გამოჩნდა. ეს ინტერესი განაპირობა როგორც თეორიულმა მიდევნებმა, ასევე პრაქტიკულმა შედეგებმა ნეირონული ქსელების დარგში. სპეციალისტები ისეთი განსხვავებული სფეროებიდან, როგორიც არის მედიცინა, ქიმია, ბიოლოგია, ფიზიკა, ფსიქოლოგია, ფიზიოლოგია და სხვა, აქტიურად მუშაობენ თავიანთ დისციპლინებში ამ პერსპექტიული ტექნოლოგიის თეორიული ასპექტების და გამოყენების შესაძლებლობების გასაფართოებლად.

მეცნიერებისა და ტექნიკის ნებისმიერი სფერო აყენებს მრავალ ამოცანას, რომელთა ნაწილი წარმატებით წყდება თანამედროვე – ფონ ნეიმანის არქიტექტურის [1] კომპიუტერების მეშვეობით. კომპიუტერით ამოცანის გადაწყვეტის პროცესი შეიძლება დაეყოს სამ ძირითად ეტაპად:

1. კომპიუტერში ამოცანის ამოხსნის პროგრამა და საწყისი მონაცემები შეჰყავს მომხმარებელს;
2. კომპიუტერი შეყვანილი მონაცემების და პროგრამის გამოყენებით აწარმოებს გამოთვლებს;
3. კომპიუტერს გამოჰყავს გამოთვლების შედეგები გარე მოწყობილობებზე.

ამავე დროს, ამოცანის გადაწყვეტის მთავარი პრობლემა – ამოცანის ამოხსნის იმ ალგორითმის პოვნა, რომლის შესაბამისადაც მომხმარებელი ადგენს პროგრამას, ამ სქემაში არ არის. დგება საკითხი – შესაძლებელია თუ არა ნებისმიერი ამოცანის ამოხსნის ალგორითმის შექმნა. არსებობს ამოცანათა მთელი კლასი, რომელთა გადასაწყვეტად ადამიანი იღებს სწორ

გადაწყვეტილებას იმ შემთხვევაშიც კი, თუ მოცემული სიტუაცია მას ადრე არ შეხვედრია. ამგვარი ამოცანების გადასაწყვეტი ალგორითმების აგებას აძნელებს, ზოგჯერ კი შეუძლებელს ხდის ის გარემოება, რომ ადამიანს არ შესწევს უნარი გაითვალისწინოს ყველა მოსალოდნელი ფაქტორი და დაადგინოს ყველა ის კანონზომიერება, რომელიც ერთმანეთთან აკავშირებს ამოცანის პირობებს შედეგებთან.

ასეთი ტიპის რთული, ძნელად ფორმალიზებადი ამოცანა მრავლად გვხვდება ბიოლოგიაში, მედიცინაში, კრიმინალისტიკაში, სოციოლოგიაში, ფიზიკისა და ტექნიკის იმ სფეროებში, რომლებიც დაკავშირებულია სახეთა ამოცნობასა და საუკეთესო გადაწყვეტილების არჩევასთან. სწორედ ამ სფეროებშია ყველაზე უფრო შეზღუდული ტრადიციული არქიტექტურის კომპიუტერების გამოყენება. ამავე დროს, აღნიშნულ სფეროებში უკვე იყენებენ ხელოვნური ნეირონული ქსელების ტექნოლოგიებს (ნეიროკომპიუტერები, ნეიროპროგრამები). ინტენსიურად ყალიბდება მეცნიერების ახალი დარგი – ნეიროინფორმატიკა (ნეიროკიბერნეტიკა, ნეირომათემატიკა, ნეიროკომპიუტინგი). თავსართი „ნეირო“ მიუთითებს, რომ ნეიროკომპიუტერებისა და ნეიროპროგრამების მუშაობის პრინციპები (გამარტივებული სახით) ერთი მხრივ, მსგავსია ნერვული სისტემის უჯრედების – ბიოლოგიური ნეირონების ურთიერთქმედებისა მათი დენტრიტების, სინაფსებისა და აქსონების მეშვეობით, და მეორე მხრივ, განსხვავებით ჩვეულებრივი კომპიუტერებისა და პროგრამებისგან, ნეიროკომპიუტერები და ნეიროპროგრამები თვითსწავლებადია (თუ სისტემა აკმაყოფილებს სწავლადობის გარკვეულ პირობებს).

ხელოვნური ნეირონული ქსელები აფართოებს გამოთვლების ცნებას, მათ საფუძველზე მოსალოდნელია ისეთი ავტომატების სინთეზი, რომლებიც შეასრულებს იმ ფუნქციებს, რომლებიც დღეისათვის მხოლოდ ადამიანის ტვინის პრეროგატივაა.

ხელოვნური ნეირონული ქსელების თეორია სწრაფად ვითარდება, რაც იმის საწინდარია, რომ უახლოეს მომავალში შეიქმნება ნეიროქსელების ახალი, უფრო დახვეწილი პარადიგმები, ხოლო მათ საფუძველზე აგებული ახალი თაობის გამოთვლითი ტექნიკის - ნეიროკომპიუტერების შესაძლებლობები მნიშვნელოვნად გაფართოვდება. ა. გორბანის [2] თანახმად, გამოიკვეთა ხელოვნური ნეირონული ქსელების იდეოლოგიის სამი წყარო; ეს არის წარმოდგენები ადამიანის თავის ტვინის აგებულებისა და ფუნქციონირების, აგრეთვე, სწავლების პროცესების შესახებ და კიდევ ერთი წყარო, რომელიც დაკავშირებულია მონაცემთა დამუშავების პარალელურ-მიმდევრობითი ალგორითმების განსაკუთრებულ წარმოდგენასთან, რაც ორიენტირებულია სპეციალიზებულ თანაპროცესორზე, რომელიც პარალელურად ასრულებს მონაცემების მასივებზე ერთგვაროვან ოპერაციებს (მაღალი პარალელიზმის გამოთვლითი სისტემები).

თანამედროვე ნეირობიოლოგია, რომელიც აწყდება ბუნებრივი ნეირონების ურთიერთქმედების საბოლოო შედეგების განმსაზღვრელი ახლად აღმოჩენილი ფაქტორების რაოდენობის განუწყვეტელ ზრდას, ჯერ ვერ იძლევა ტვინის ფუნქციების ერთმნიშვნელოვან თეორიულ კონცეფციებს; ამიტომ უფრო სწორი იქნება, თუ ვიტყვით, რომ ხელოვნური ნეირონული ქსელების იდეაში ჩადებულია რეალობასთან მეტ-ნაკლებად მიახლოებული ფორმალიზებული წარმოდგენები ტვინის სტრუქტურისა და ფუნქციონირების შესახებ. იდეის ძირითადი ბირთვია ის, რომ ბუნებრივი ნეირონების მოდელირება შეიძლება მარტივი ავტომატებით, ხოლო ტვინის ფუნქციონირების მთელი სირთულე, მოქნილობა და სხვა თვისებები განისაზღვრება ნეირონთაშორისი კავშირებით, ანუ განსაკუთრებული მნიშვნელობა ენიჭება კავშირების სტრუქტურას [2].

სამეცნიერო-ტექნიკური მიმართულება, რომელიც განისაზღვრება ნეირონული ქსელების აღწერილი

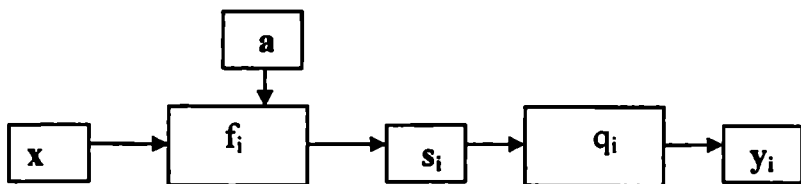
წარმოდგენით, ცნობილია როგორც კონექციონიზმი (connection კავშირი). კონექციონიზმს მჭიდროდ უკავშირდება ერთგვაროვნების, სიჭარბისა და პოლო-გრაფიულობის იდეები, ანუ მარტივი ელემენტებისაგან რთული ერთგვაროვანი პარალელური სისტემების, ხოლო შედარებით არასაიმედო ელემენტებისაგან საიმედო სისტემების შექმნის შესაძლებლობა; ამავე დროს, ასეთი სისტემები რომელიმე ნაწილის დაზიანების შემთხვევაში ინარჩუნებს სასარგებლო თვისებებს.

ნეიროინფორმატიკის იდეის ყველა ფორმალიზაცია და, მითუმეტეს, რეალიზება გულისხმობს განსაკუთრებული მოწყობილობის არსებობას (ფორმალურად ან რეალურად), რომელიც მიიღებს მონაცემთა ვექტორს და გამოითვლის მის გარკვეულ ფუნქციებს, რომლებიც შეიძლება დამოკიდებული იყოს შეთანაწყობადი პარამეტრების ნაკრებზე. ამ სპეციალიზებული მოწყობილობის ფუნქციონირების აღწერისას იგულისხმება, რომ პარამეტრების მნიშვნელობები წინასწარ არის განსაზღვრული (შეთანაწყობა ჩატარებულია) და ეს მნიშვნელობები ინახება მეხსიერების განსაკუთრებულ მოწყობილობაში.

პარამეტრების შეთანაწყობის ალგორითმები ეფექტურად უნდა იყენებდეს იმავე მოწყობილობას (სპეციალიზებულ თანაპროცესორს) და გადასცემდეს მას შესასრულებლად იმ ერთგვაროვან გამოთვლებს, რომელთა მოცულობაც იზრდება ვექტორის განზომილების პროპორციულად ან უფრო სწრაფად.

სპეციალიზებული თანაპროცესორის (სურ.1.1) მუშაობა შეიძლება წარმოვადგინოთ ფორმალური სახით.

ვთქვათ, მოცემულია: x - მონაცემთა ვექტორი; a - პარამეტრების ვექტორი; F - ორი ვექტორული ცვლადის ფუნქციების სასრული სიმრავლე; Q - ერთი ცვლადის ფუნქციების სასრული სიმრავლე. მაშინ თანაპროცესორი შეასრულებს შემდეგ ოპერაციებს:



სურ.1.1. სპეციალიზებული თანაპროცესორის მუშაობის ფორმალიზაცია

მოცემული $f_i \in F$ ფუნქციით და x_i და a_i ვექტორების მიხედვით s_i ვექტორის გამოთვლა:

$$s_i = f_i(x_i, a_i), \quad (1)$$

მოცემული $q_i \in Q$ და s_i ვექტორის მიხედვით y_i ვექტორის გამოთვლა:

$$y_i = q_i(s_i). \quad (2)$$

როგორც წესი $f(x, a)$ ფუნქცია წარმოადგენს სკალარულ ნამრავლს:

$$f(x, a) = \sum x_k a_k. \quad (3)$$

ზოგიერთ სპეციალურ შემთხვევაში $f(x, a)$ ფუნქცია შეიძლება იყოს არაწრფივით. Q სიმრავლეში შემავალი ფუნქციების მაგალითებია:

$q_1(x_i) = |x_i|$ — აბსოლუტური მნიშვნელობა;

$q_2(x_i) = \tanh(x_i)$ — ჰიპერბოლური ტანგენსი;

$q_3(x_i) = x_i^2$ — კვადრატული ფუნქცია.

Q ფუნქციებიდან ზოგს შეიძლება პქონდეს შეთანაწყობის პარამეტრები. ვექტორების განზომილებების ზოგიერთი შეზღუდვის შემთხვევაში მიიჩნევა, რომ ყველა ეს ოპერაცია პარალელურად სრულდება.

ზემოთ აღნიშნული იყო, რომ ნაშრომში [2] ნეირონფორმატიკის იდეოლოგიის ერთ-ერთ წყაროდ მითითებულია „წარმოდგენები სწავლების პროცესებზე“.

რაც ეჭვს არ იწვევს. თუმცა, ეს დებულება, ჩვენი აზრით, უნდა გაფართოვდეს – ნეიროინფორმატიკის იდეოლოგიის ერთ-ერთ წყაროდ უნდა მივიჩნიოთ არა მხოლოდ წარმოდგენები სწავლების პროცესებზე, არამედ მთლიანობაში სახეთა ამოცნობის თეორია, რომელიც მნიშვნელოვნად ასაზრდოებს მას იდეურად, აგრეთვე შესაძლებელს ხდის პრაქტიკული ამოცანების გადასაწყვეტად ხელოვნური ნეირონული ქსელების სინთეზის პროცესებში გამოყენებულ იქნეს სახეთა ამოცნობაში შემუშავებული მრავალი მეთოდი, სახელდობრ:

- ნეიროქსელების მუშაობა პირველ მიახლოებაში შეიძლება წარმოვიდგინოთ შემდეგი პროცედურით – ნეიროქსელში შემავალ სიგნალებს, რომლებიც ეკუთვნის N განზომილების სივრცეს, ნეიროქსელი განაზოგადებს და მიაკუთვნებს გარკვეული რაოდენობის კლასებს. მათემატიკური თვალსაზრისით, ეს ხდება ჰიპერსიბრტყეებით სივრცის დაყოფის გზით. ჩანაწერი წრფივი ფორმალური ნეირონებისათვის იქნება

$$f(x, a) = \sum_i x_i a_{ik}, \quad k=1 \dots m, \quad i=1 \dots N. \quad (4)$$

სახეთა ამოცნობაში (4) გამოსახულებით მოცემულ ფუნქციას ეწოდება მსგავსების ზომა სკალარული ნამრავლით; მივიღებთ კლასების რაოდენობას არა უმეტესს 2^m , სადაც m ქსელის გამოსასვლელების რაოდენობაა. ეს პროცედურა ანალოგიურია სახეთა ამოცნობაში წრფივი დისკრიმინანტული ფუნქციისა, ანუ მისი მოდელირებით მივიღებთ კლასიკურ მაკ-კალოკისა და პიტსის ნეირონს, ხოლო თუ ამ მიდგომას განვაზოგადებთ არაწრფივი დისკრიმინანტული ფუნქციების მიმართ, მაშინ შესაძლებელია სხვადასხვა სახის მრავალშრიანი ნეირონული ქსელების მიღება (იგულისხმება, რომ სქემას ყველგან ემატება აქტივაციის ელემენტი);

- ნეირონული ქსელების სინთეზის და ფუნქციონირების პროცესებში გამოყენებულია სახეთა ამოცნობის ისეთი მეთოდები და პროცედურები, როგორიცაა: გადასაწყვეტი ამოცანის მაგალითებისაგან მონაცემთა

ბაზის შედგენა, ამ ბაზის დაყოფა სასწავლო და საკონტროლო სიმრავლეებად, ამოცანის შესაბამისი ნიშანთა სივრცის შერჩევა, მონაცემთა წარმოდგენა რიცხვითი ფორმით და მათი ნორმირება, ნეიროქსელის სწავლებისა და ფუნქციონირების კრიტერიუმების შერჩევა, სწავლებაში ოპტიმიზაციის მრავალფეროვანი მეთოდების გამოყენება და სხვ.;

- ისევე როგორც სახეთა ამომცნობი სისტემების მიერ მიღებული გადაწყვეტილებები, ნეიროქსელების მიერ მიღებული გადაწყვეტილებებიც ყოველთვის არ არის კატეგორიული, ანუ ქსელი გადაწყვეტილებასთან ერთად იძლევა მისი საიმედოობის მაჩვენებელსაც, რაც კრიტიკულად უნდა შეაფასოს მომხმარებელმა.

სხვა სიტყვებით რომ ვთქვათ, ნეიროქსელებშიც მსგავსი პრობლემაა, რაც სახეთა ამოცნობაში, ანუ ზემოთ ჩამოთვლილი პროცედურების ეფექტურობა (სასწავლო ნაკრები, მისი კორექტულობა, ქსელის პარამეტრების შერჩევა და სხვ.) მნიშვნელოვნად არის დამოკიდებული მომხმარებლის სუბიექტურ მოსაზრებებზე.

ამოცანის ზუსტი დასმა, იმის მკაფიოდ წარმოდგენა, თუ რა უნდა „დაისწავლოს“ ნეიროქსელმა, საკმაოდ რთული საქმეა, ანუ ხაზი უნდა გაესვას იმ გარემოებას, რომ ხელოვნური ნეირონული ქსელი არ წარმოადგენს ყველა პრობლემის გადაწყვეტის პანაცეას, არამედ ის არის მოქნილი ინსტრუმენტი სათანადოდ მომზადებული მომხმარებლის ხელში.

ამრიგად, მოყვანილი მოსაზრებების საფუძველზე ნათლად იკვეთება ნეიროინფორმატიკის იდეოლოგიის სამი წყარო:

1. ადამიანის ცენტრალური ნერვული სისტემის ფიზიოლოგია (ნეირობიოლოგია, როგორც მრავალდონიანი მეცნიერება);
2. მაღალი პარალელიზმის გამოთვლითი სისტემები (პარალელური ერთგვაროვანი პროცესორები);
3. სახეთა ამოცნობის თეორია.

ნეიროინფორმატიკის იდეოლოგიის წყაროების – ნეირობიოლოგიის, მაღალი პარალელიზმის გამოთვლითი სისტემებისა და სახეთა ამოცნობის თეორიის ის მნიშვნელოვანი განსაზღვრებები, დებულებები და მეთოდები, რომლებიც ხშირად გამოიყენება ხელოვნური ნეირონული ქსელების თეორიასა და პრაქტიკაში, მოკლედ არის განხილული მომდევნო თავებში.

12. ბუნებრივი ნეირონისა და ნეირონული ქსელების ბიომორფოლოგიური აღწერა

12.1. ბუნებრივი ნეირონი და მისი სტრუქტურული ელემენტები

ადამიანის ცენტრალური ნერვული სისტემა (ცნს) შედგება თავის ტვინისა და ზურგის ტვინისაგან. ცენტრალური ნერვული სისტემის ძირითადი ფუნქციებია:

1. ორგანიზმის ყველა ქსოვილისა და ორგანოს მოქმედების რეგულირება და მათი ერთ მთლიანობაში გაერთიანება;
2. ორგანიზმის გარემოს პირობებთან შეგუება, ანუ ორგანიზმის მოთხოვნილებების ადეკვატური ქცევის უზრუნველყოფა.

ცენტრალური ნერვული სისტემის წამყვანი ნაწილია დიდი ნახევარსფეროების ქერქი, ანუ ნეოკორტექსი, რომელიც მართავს ადამიანის ცხოველქმედების ყველაზე უფრო რთულ ფუნქციებს – ფსიქიკურ პროცესებს. მისი სისქე ≈ 3 მმ, ხოლო საერთო ფართობი $\approx 2,500$ სმ² უდრის.

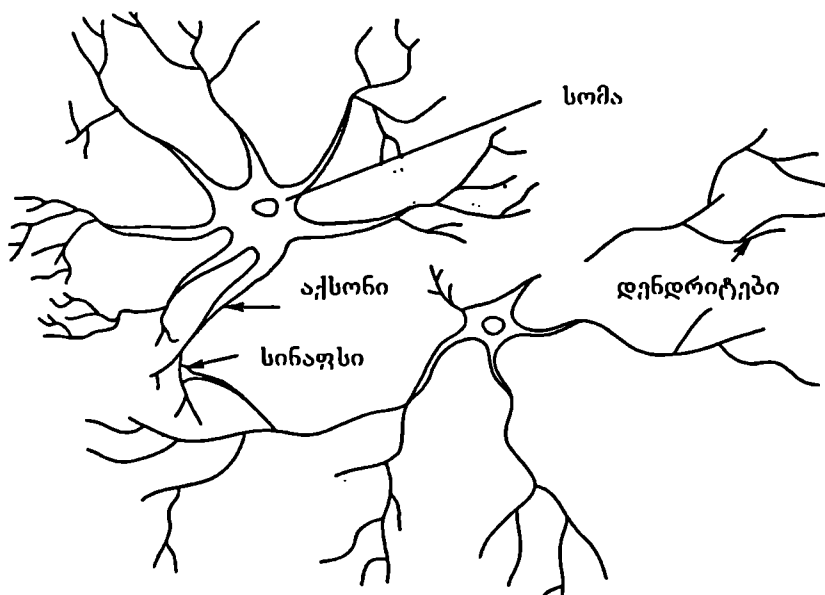
ცენტრალური ნერვული სისტემის ძირითადი სტრუქტურული ელემენტია ნერვული უჯრედი, ანუ ნეირონი. ნეირონების მეშვეობით ხდება ინფორმაციის გადაცემა ცნს-ის სხვადასხვა უბნებს შორის, აგრეთვე ინფორმაციის გაცემა ნერვულ სისტემასა და სხეულის

სხვადასხვა ორგანოებს შორის. ნეირონებში მიმდინარეობს ინფორმაციის დამუშავების ძირითადი პროცესები, მათი საშუალებით ფორმირდება ორგანიზმის საპასუხო რეაქციები (რეფლექსები) გარე და შიგა გაღიზიანებებზე.

ნეირონების სამ ძირითად ტიპს განარჩევენ: აფერენტულ (მგძნობიარე), ეფერენტულ (მამოძრავებელ) და საშუალო (ქართულ) ნეირონებს. აფერენტული ნეირონები (ცენტრისკენული) ინფორმაციას გადასცემენ რეცეპტორებიდან ცენტრალურ ნერვულ სისტემას; ეფერენტული ნეირონები (ცენტრიდანული) უზრუნველყოფენ ინფორმაციის გადაცემას ცენტრალური ნერვული სისტემიდან მუშა ორგანოებისაკენ; საშუალო ნეირონები (ინტერნეირონები) შედარებით მცირე უჯრედებია, რომლებიც აკავშირებენ ერთმანეთთან აფერენტულ და ეფერენტულ ნეირონებს და ქმნიან გამტარ გზებს.

ნეირონის სხვადასხვა სტრუქტურულ ელემენტებს ახასიათებს ფუნქციური თავისებურება და სხვადასხვა ფიზიოლოგიური მნიშვნელობა. ნერვული უჯრედის ძირითადი შემადგენელი ელემენტებია: ნერვული უჯრედის სხეული ცენტრალური ბირთვითურთ ანუ სომა (Soma), დენდრიტები (dendrites) და აქსონი (axon). დენდრიტები წარმოადგენს ნერვული უჯრედის განტოტვილ მრავალრიცხოვან წანაზარდებს (მორჩებს). ისინი შეიძლება განვიხილოთ, როგორც ნეირონის შესასვლელები, რომელთა გავლითაც სიგნალები მიეწოდება სომას. სომადან გამომავალი ერთადერთი წანაზარდი — აქსონი წარმოადგენს ნეირონის გამოსასვლელს; იგი გადასცემს ნერვულ იმპულსებს სხვა ნეირონებს ან მუშა ორგანოებს (კუნთი, ჯირკვავი). ნერვული უჯრედების ფორმები, წანაზარდების სიგრძე და განლაგება ძალზე მრავალფეროვანია და დამოკიდებულია მათ ფუნქციურ დანიშნულებაზე (სურ. 12). ნეირონების განივკვეთის ზომები მერყეობს 5-დან 90 მიკრონამდე. ერთ ნეირონს

შეიძლება ჰქონდეს რამდენიმე ათასი დენდრიტი, რომელთა დიამეტრი დაახლოებით 1 მიკრონის ტოლია.



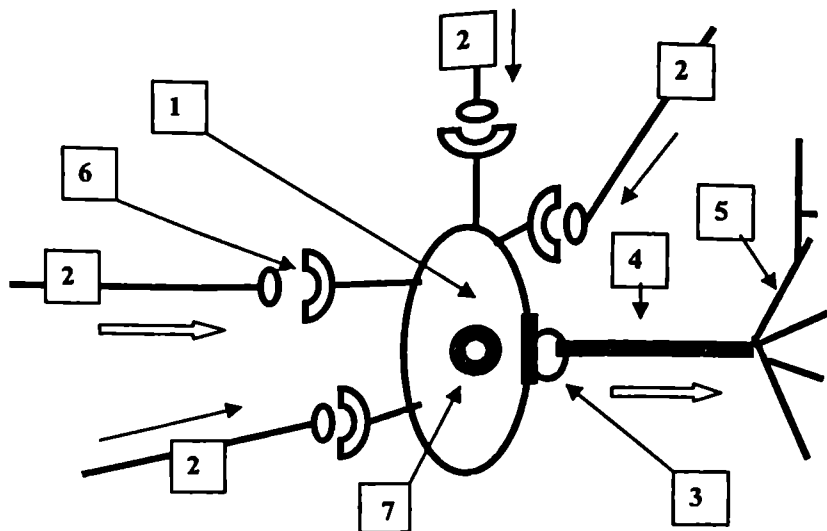
სურ. 12. ბუნებრივი ნეირონი და მისი სტრუქტურული ელემენტები.
(ილუსტრაცია აღებულია [21] ნაშრომიდან)

აქსონები გამოირჩევა სიგრძით – რამდენიმე მილიმეტრიდან ერთ მეტრამდე. დენდრიტების და სომაზე მემბრანაზე განლაგებულია სინაფსები (Sinaps) – სხვა ნეირონების აქსონების (იშვიათად საკუთარი აქსონის) დაბოლოებები (სურ. 13). სინაფსების განლაგების სიმჭიდროვე შედარებით მეტია დენდრიტების განშტოებათა კვანძებსა და ლოკალური გაფართოების უბნებში. სინაფსების ზომები სხვადასხვაა – სომაზე არის $0,2 \div 0,4$ მკმ, ხოლო დენდრიტებზე $\approx 1,0$ მკმ. დენდრიტების სოკოსებრ გამონაზარდებზე განლა-

გებულია აგრეთვე რამდენიმე სპეციალური „კბილისებრი სინაფსი“. ზურგის ტვინის თითოეულ მსხვილ მოტონეირონს აქვს 15000 ÷ 20000 სინაფსი.

ნეირონების სიმჭიდროვე ყველაზე მაღალია ადამიანის თავის ტვინის ქერქში და ერთ კუბურ მილიმეტრში 40 ათას ნეირონს აღწევს. ზოგიერთი შეფასებით ცენტრალურ ნერვულ სისტემაში არის დაახლოებით ასი მილიარდი ნეირონი. ნეირონების გრძელი მორჩების კონები შეადგენს ნერვებს, რომლებიც ცენტრალური ნერვული სისტემისა და სხვადასხვა ორგანოების დამაკავშირებელი ტრაქტებია. ტრაქტებს აქვს თეთრი ფერი, ხოლო ნეირონის სომასა და დენდრიტებს რუხი ფერი, ამიტომ ტვინში განასხვავებენ თეთრ ნივთიერებას (გამ-ტარი გზები) და რუხ ნივთიერებას (ქერქული სტრუქტურები და ქერქქვეშა ბირთვები).

ბიოლოგიური ნეირონის სქემა და სინაფსური კავშირები ნაჩვენებია 1.3 სურათზე, სადაც: 1 არის სომა; 2 – დენდრიტები; 3 – აქსონის ბორცვი; 4 - აქსონი; 5 – აქსონის განშტოებები; 6 – სინაფსური კავშირი; 7 – ცენტრალური ბირთვი. ისრებით ნაჩვენებია სიგნალების გავრცელების მიმართულებები ნეირონში.



სურ. 13. ბიოლოგიური ნეირონის გამარტივებული სქემა და მისი სინაფსური კავშირები

ნეირონებში აღინიშნება ნივთიერებათა ცვლის მაღალი სიჩქარე და აერობული პროცესების სიჭარბე. ტვინი მოითხოვს ჟანგბადის დიდ რაოდენობას: მისი წონა დაახლოებით ორგანიზმის წონის 2 პროცენტია, ხოლო ჟანგბადის მოხმარება სისხლის მეშვეობით აღწევს 45 პროცენტს (≈ 46 მლ/წთ). ჟანგბადის მიწოდების შეფერხება იწვევს შეუქცევადი პროცესების განვითარებას: ზურგის ტვინში ≈ 30 წთ-ის შემდეგ, თავის ტვინში ≈ 20 წთ-ის შემდეგ, ხოლო თავის ტვინის ქერქში ≈ 6 წთ-ის შემდეგ.

ტვინისათვის ენერგიის ძირითადი წყარო არის გლუკოზა, რომელსაც იგი დღე-ღამეში ≈ 15 გრამს მოიხმარს. ადამიანის ორგანიზმის მთელი სადღეღამისო ენერგია-დანახარჯის $12 + 17$ პროცენტი მოდის ტვინის წილად. ნეირონების მიერ ჟანგბადის და გლუკოზის მაღალ მოთხოვნილებას უზრუნველყოფს სისხლის ინტენსიური მიწოდება. სისხლი ტვინში 5-7-ჯერ უფრო მეტი სიჩქარით

მოძრაობს, ვიდრე უმოქმედო კუნთებში. ტვინი უხვად არის უზრუნველყოფილი სისხლძარღვებით, თვითოეულ მსხვილ ნეირონს აქვს რამდენიმე საკუთარი კაპილარი, რომლებიც ტვინის ქერქის მოცულობის 10 პროცენტს შეადგენს.

ნერვული უჯრედების კეებასა და ნივთიერებათა ცვლაში მონაწილეობენ აგრეთვე გლიური უჯრედები (ნეიროგლია), რომლებიც ავსებენ ტვინის ნეირონთა შორის სივრცეებს. ნეოკორტექსში მათი რაოდენობა 5-ჯერ მეტია, ვიდრე ნეირონებისა. მთლიანად გლიური უჯრედების რაოდენობა დაახლოებით 130 მილიარდია. ცენტრალური ნერვული სისტემის კაპილარები თითქმის მთლიანად (80 %-მდე) დაფარულია გლიური უჯრედებით, ამ უჯრედებს შეუძლიათ გადაადგილდნენ სივრცეში ყველაზე უფრო აქტიური ნეირონებისაკენ. შესაძლებელია, ისინი მონაწილეობას იღებდნენ აგრეთვე ტვინის პირობით-რეფლექსურ მოქმედებასა და მეხსიერების პროცესებში. ზოგიერთი მკვლევარის აზრით ნეიროგლია ინფორმაციის შენახვასთან ერთად აგრეთვე „აპროგრამებს“ ნეირონებს.

12.2. ბიოლოგიური ნეირონის ძირითადი ფუნქციები

ნერვული უჯრედის ანუ ბიოლოგიური ნეირონის ძირითადი ფუნქციებია:

გარე გაღიზიანებების აღქმა (რეცეპტორული ფუნქცია);

- გაღიზიანებების სიგნალების გადამუშავება (ინტეგრაციული ფუნქცია);

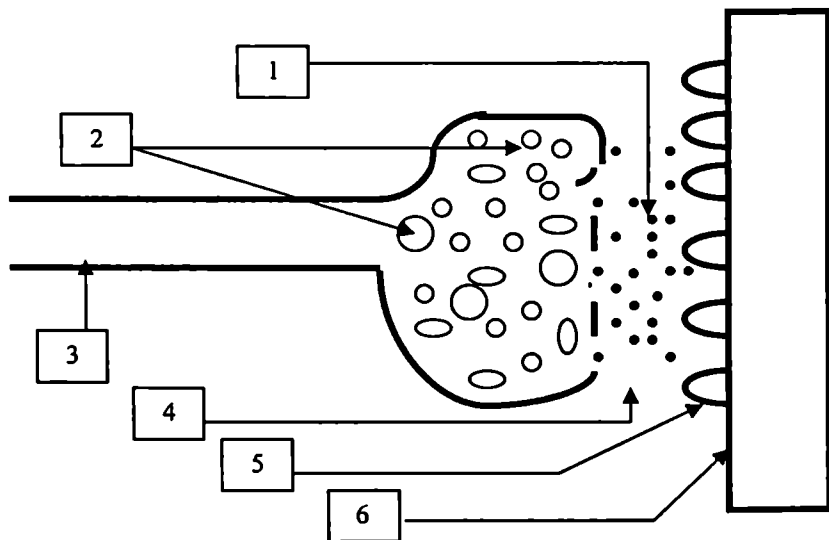
- ნერვული უმოქმედების გადაცემა სხვა ნეირონებისათვის ან სხვადასხვა მუშა ორგანოებისათვის (ეფექტორული ფუნქცია).

ნერვულ სისტემაში შემოსული ყველა აგზნება ნეირონს გადაეცემა სინაფსური კონტაქტების მიმდებარე მემბრანის გარკვეული უბნების მეშვეობით. უმეტესად ეს გადაცემა ხორციელდება ქიმიური გზით სპეციალური

ნივთიერებების მედიატორების (ნეიროტრანსმიტერების) საშუალებით. ნეირონის პასუხი არის მემბრანული პოტენციალის სიდიდის ცვლილება. აქსონებს დაბო-ლო-ებაზე აქვს განშტოებები, რომელთაც შეუძლიათ სინაფსების წარმოქმნა დენდრიტებზე (აქსოდენდრიტული სინაფსები), სომაზე (აქსოსომატური სინაფსები) და ზოგჯერ აქსონზე (აქსოაქსონური სინაფსები). ყველაზე მეტი (50 პროცენტამდე) სინაფსი განლაგებულია დენდრიტებზე.

სინაფსური კონტაქტის პრესინაფსურ ნაწილში მოსულ ნერვულ იმპულსს შეჰყავს მედიატორი (აცეტილქოლი, ნორადრენალინი და სხვ.) სინაფსურ ნაპრაღში.

1.4 სურათზე ნაჩვენებია სინაფსური კონტაქტის ნაპრაღ-ში მედიატორის გამოყოფა, სადაც: 1 არის ნეიროტრანსმიტერი; 2 – სინაფსური ბუშტულები; 3 – პრესინაფსური აქსონი; 4 – სინაფსური ნაპრაღი;



სურ. 1.4. სინაფსური კონტაქტის ნაპრაღში მედიატორის გამოყოფა

5 - ნეიროტრანსმიტერული რეცეპტორი; 6 - პოსტსინაფსური დენდრიტი.

სინაფსის აქტივაციის დროს წარმოქმნილი ეფექტები შეიძლება იყოს ამგზნები ან მამუხრუჭებელი. ეს დამოკიდებულია მედიატორის და პოსტსინაფსური მემბრანის თვისებებზე. მოსვენებულ მდგომარეობაში მემბრანა პოლარიზებულია, ანუ მემბრანის გარეთ დადებითი მუხტი რეგისტრირდება, ხოლო შიგნით - უარყოფითი. ნეირონის მოსვენების მემბრანული პოტენციალი ≈ 70 მვოლტია. ამგზნები ზეგავლენისას იზრდება მემბრანის შეღწევადობა, რაც განაპირობებს ნატრიუმის დადებითად დამუხტული იონების შესვლას უჯრედში და შედეგად, მემბრანის სხვადასხვა მხარეზე პოტენციალთა სხვაობის შემცირებას, ანუ მის დეპოლარიზაციას. უჯრედის მოცემული უბნის პოსტსინაფსურ მემბრანაში ამ დროს რეგისტრირდება მემბრანული პოტენციალის უარყოფითი რხევა ამპლიტუდით ≈ 10 მვოლტი, ანუ ამგზნები პოსტსინაფსური პოტენციალი (აპსკ), რომელიც 1,2 მილიწამის განმავლობაში აღწევს თავის მაქსიმუმს, ხოლო შემდეგ კლებულობს. მამუხრუჭებელი ზეგავლენისას მემბრანის შეღწევადობა უმნიშვნელოდ იზრდება, ძირითადად კალიუმის იონებისათვის (კალიუმის იონის დიამეტრი ნაკლებია, ვიდრე ნატრიუმის). იმის გამო, რომ კალიუმის იონების რაოდენობა მეტია უჯრედის შიგნით, ისინი გამოდიან გარეთ, იზრდება მემბრანის პოლარიზაცია, ანუ ხდება მისი ჰიპერპოლარიზაცია. ამ დროს რეგისტრირდება დადებითი რხევა ≈ 5 მვოლტის ამპლიტუდით, ანუ მამუხრუჭებელი პოსტსინაფსური პოტენციალი (მპსკ). ეს პოტენციალები, როგორც წესი, გრძელდება რამდენიმე მლწამის განმავლობაში, თუმცა თავის ტვინის ზოგიერთ ნეირონში აპსკ-ს ხანგრძლივობა 80 მლწამამდე, ხოლო მპსკ - 100 მწამამდე აღწევს.

4. ნეირონის მემბრანული პოტენციალის საერთო ცვლილება არის დენდრიტებსა და სომაზე განლაგებული ყველა მრავალრიცხოვანი აქტივირებული სინაფსის ადგილობრივი აპსკ და მპს პოტენციალების

რთული ურთიერთქმედების (ინტეგრაციის) შედეგი. ნეირონის მემბრანაზე ხდება პოტენციალის დადებითი და უარყოფითი რხევების ალგებრული აჯამება. მამუხრუჭებელი სინაფსური ზემოქმედების გადაჭარბების დროს მოხდება მემბრანის პიპეროლარიზაცია და ნეირონის მოქმედების დამუხრუჭება, ხოლო თუ მემბრანული პოტენციალი გადაიხარა დეპოლარიზაციისაკენ – გაიზრდება ნეირონის აგზნებადობა. ნეირონის საპასუხო განმუხტვა მოხდება მხოლოდ მაშინ, როდესაც მემბრანული პოტენციალის სიდიდე მიაღწევს დეპოლარიზაციის კრიტიკულ დონეს – ზღურბლურ მნიშვნელობას. ამისათვის ამგზნები პოსტსინაფსური პოტენციალის სიდიდე უნდა გახდეს დაახლოებით 10 მეოლტი. დეპოლარიზაციის კრიტიკული დონის მიღწევისთანავე იწყება ნატრიუმის იონების უჯრედში შესვლის ზეავისებრი პროცესი და ამ დროს რეგისტრირდება ე.წ. მოქმედების პოტენციალი (მპ). მოქმედების პოტენციალში განასხვავებენ ხანმოკლე, შედარებით მაღალვოლტიან პიკურ ნაწილს, ანუ სპაიკს და ხანგრძლივ დაბალი ამპლიტუდის რხევებს. მოტონეირონების მოქმედების პიკის ამპლიტუდა არის 80 – 100 მეოლტი, ხოლო მისი ხანგრძლივობა ≈ 1.5 მწამი. ნერვული უჯრედების დეპოლარიზაციის სიდიდე წრფივ დამოკიდებულებაშია გამღი-ზიანებელი იმპულსების სიხშირესთან. ტვინის უმაღლესი დონის განყოფილებები სხვადასხვა სიხშირის იმპულსების გაგზავნით ცნს-ის ქვედა დონის ნეირონებისათვის არეგულირებენ მათ აგზნებადობას, რითაც ახორციელებენ კონტროლს ორგანიზმის საპასუხო რეპციებზე.

5. მოქმედების პოტენციალის გაჩენისთანავე, ნერვული იმპულსი იწყებს გავრცელებას სომადან აქსონის გაყოლებით სხვა ნეირონისაკენ ან მუშა ორგანოსაკენ, ანუ ახორციელებს ნეირონის ეფექტორული ფუნქცია. ამრიგად, აქტიურ ნეირონში მიმდინარე პროცესები შეიძლება წარმოდგენილი იქნას შემდეგი თანამიმდევრობით:

1. წინამდებარე ნეირონის პრესინაფსურ დაბოლოებაში ნერვული იმპულსის მოსვლა;
2. სინაფსურ ნაპარაღში მედიატორის გამოყოფა;
3. პოსტსინაფსური მემბრანის შედწევალობის გაზრდა;
4. მემბრანის დეპოლარიზაცია (აქსკ) ან პიპერპოლარიზაცია (მქსკ);
5. ნეირონის დენდრიტებსა და სომაზე აქსკ და მქსკ პოტენციალების ურთიერთქმედება (აღგებრული აჯამევა);
6. ამგზნები ზემოქმედების გადაჭარბების შემთხვევაში, მემბრანული პოტენციალის წანაცვლება დეპოლარიზაციისაკენ;
7. დეპოლარიზაციის კრიტიკული დონის მიღწევა;
8. მოქმედების პოტენციალის წარმოქმნა;
9. ნერვული იმპულსის გაერცვლება აქსონში;
10. მედიატორის გამოყოფა აქსონის დაბოლოებებში, ანუ ნერვული პროცესის გადაცემა სამიზნე უჯრედზე – მომდევნო ნეირონზე ან მუშა ორგანოზე.

ამრიგად, ინფორმაციული პროცესები ნეირონებში ხორციელდება ორი გზით: ქიმიური მექანიზმით (მედიატორები) და ელექტრული მექანიზმით (ამგზნები პოსტსინაფსური პოტენციალი, მამუხრუჭებელი პოსტსინაფსური პოტენციალი, მოქმედების პოტენციალი).

1.2.3. ნერეული უჯრედების ურთიერთქმედების მექანიზმები და სისტემები

ნერეული უჯრედები ფუნქციონირებენ ერთმანეთთან მჭიდრო ურთიერთქმედებით, რაც ძირითადად ხორციელდება ორი მექანიზმით:

1. ნერეული უჯრედების ელექტრული ველების ზემოქმედებით (ელექტროტონული ზემოქმედებით);

2. ნერეული იმპულსების ზემოქმედებით.

ელექტრული ველების ზემოქმედების მექანიზმი შემდეგია: ნერეული უჯრედის ელექტრული მუხტი მის გარშემო ქმნის ელექტრულ ველს, რომლის რხევებიც იწვევს მეზობელი ნეირონების ელექტრული ველების ცვლილებებს, რაც ცვლის ამ ნეირონების აგზნებადობას, ლაბილურობასა და გამტარობას. ამ ველის განფენილობა არ არის დიდი, ≈ 100 მიკრონია.

ნერეული იმპულსების ზემოქმედების მექანიზმი უზრუნველყოფს ნერეული ზემოქმედების ზუსტ და სწრაფ ($100 \div 140$ მ/წმ) გადაცემას როგორც ახლო, ასევე შორ მანძილებზე. ამრიგად, კავშირის ძირითადი საშუალება ნეირონებს შორის არის ნერეული იმპულსი. ნეირონთა შორის ურთიერთქმედებაში გამოიყენება სიხშირული კოდი, ანუ ერთი ნეირონის საპასუხო რეაქციის ხასიათისა და ფუნქციური მდგომარეობის ცვლილების კოდირება ხდება იმ იმპულსების სიხშირის ცვლილებით, რომლებსაც ეს ნეირონი უგზავნის სხვა ნეირონებს. ნეირონის მნიშვნელოვანი ფიზიოლოგიური მანევრებული არის დროის ერთეულში ნეირონის მიერ გაგზავნილი იმპულსების საერთო რაოდენობა, ანუ მისი ჯამური იმპულსური აქტივობა. ნერეული იმპულსების ნეირონთა შორის გადაცემის პროცესებში უმნიშვნელოვანეს როლს ასრულებენ სინაფსები. სინაფსები მიერ ნერეული იმპულსების გადაცემის უნარი ცვალებადია – იზრდება სინაფსების აქტიური მოქმედების შედეგად და მცირდება აქტივობის

არარსებობის დროს. სინაფსები ნერვული სისტემის ყველაზე მგრძნობიარე ელემენტებია. უკვე რამდენიმე დღის უმოქმედობის შედეგად. სინაფსების ზომები მცირდება, კლებულობს მედიატორის მარაგი, იზრდება სინაფსური შეყოვნება, ხოლო სინაფსების სრული დაზიანების შემთხვევაში, ნეირონებს შორის კავშირი საერთოდ წყდება.

ნერვული სისტემის ორგანიზაციის უმარტივესი ფორმა არის დიფუზიური ნერვული ქსელი, რომელიც აქვთ უდაბლეს უხერხემლო ცხოველებს (ღრუბელი, მედუზა, აქტინია). ასეთი ნერვული ქსელის ნეირონებს აქვთ მრავალრცხოვანი ურთიერთკავშირები, რომლებითაც აგზნება დიფუზურად ვრცელდება ყველა მირთულეებით.

ნერვული სისტემის ორგანიზაციის შედარებით რთული ფორმაა სეგმენტური ნერვული სისტემა, რომელიც აქვთ უმაღლეს უხერხემლო ცხოველებს (მწერები, ჭიაყელები). მათი სხეულები შედგება ერთნაირი აგებულების სეგმენტებისაგან, რომელთა ინერვაცია ხდება ამავე სეგმენტებში განლაგებული ნეირონებით.

ევოლუციის შემდეგი საფეხური ხერხემლიან ცხოველებს უეითარებს ნერვული სისტემის სათავე დაბოლოებას - ენცეფალონს, სადაც თავს იყრიან ნერვული უჯრედების ჯგუფები - ნერვული ცენტრები, რომლებიც მართავენ უმნიშვნელოვანეს ფუნქციებს. ამ დროს ცალკეული სეგმენტები კარგავს ავტონომიურობას და ფუნქციების მეტი ნაწილი გადაეცემა ნერვული სისტემის ზემდგომ განყოფილებებს, ანუ მიმდინარეობს ფუნქციების ენცეფალიზაცია (ცენტრალიზაცია). თავის ტვინის დიდი ნახევარსფეროების ქერქის - ნეოკორტექსის ფორმირებასთან ერთად, მიმდინარეობს მისდამი ნერვული სისტემის ყველა სხვა განყოფილების დამორჩილება, ანუ კორტიკალიზაციის პროცესი. ადამიანის ცნს-ის აგებულება და ფუნქციები ასახავს ამ ევოლოციურ პროცესს. მაგალითად,

სეგმენტური ორგანიზაციის ნიშნები აქვს ზურგის ტვინს, რომლის ნერვული ცენტრები უშუალოდ მართავს სხეულის იმავე დონის უბნებს, ხოლო ნეოკორტექსს არა აქვს უშუალო კავშირები სხეულის ორგანოებთან, არამედ იგი მათ მოქმედებებს მართავს სეგმენტური ცენტრების მეშვეობით.

1.2.4. ნეირონული დოქტრინა და პარადიგმების ცვლა

ნაშრომში [3] მოცემულია ე.წ. თანამედროვე ნეირონული დოქტრინა, რომლის თანახმად ნერვული სისტემის ანატომიურ, გენეტიკურ, ფუნქციურ, პოლარიზებულ, დამაზიანებელ შემოქმედებათა მიმართ მგრძნობიარე და რეგენერაციის უნარის მქონე ერთეულს წარმოადგენს ნეირონი. დოქტრინა ეფუძნება შემდეგ ექვს დებულებას:

1. ნეირონი ნერვული სისტემის მთავარი შემადგენელი ერთეულია, რომლის ძირითადი ელემენტებია უჯრედის სხეული (სომა), დენდრიტები, სინაფსები, აქსონი;
2. ნეირონი გენეტიკური ერთეულია, ვითარდება დამოუკიდებელი ნეიროპლასტიდან, აღჭურვილია გენეტიკური კოდით, რაც განსაზღვრავს მის სტრუქტურას, მეტაბოლიზმს, კავშირებს, რომლებიც დასწავლის პროცესისას ვარიანბელურია;
3. ნეირონი ფუნქციური ერთეულია, მას აქვს სტიმულების მიმღებლობის, იმპულსების გენერაციისა და გატარების უნარი, ამასთან, იგი ჩართულია რომელიმე ნერვული წრედის შემადგენლობაში.
4. ნეირონი არის პოლარიზებული ერთეული, ანუ იმპულსებს მხოლოდ ერთი მიმართულებით ატარებს, ამასთან სხვა უჯრედზე აგზნება მხოლოდ სინაფსური გზით გადაიცემა (ეს უკანასკნელი დებულება, როგორც ქვემოთ იქნება

ნაწევნები, ბოლო ნეირობიოლოგიური მონაცემების საფუძველზე უარყოფილია);

5. ნეირონი წარმოადგენს დამაზიანებელ ზემოქმედებათა მიმართ მგრძნობიარე ერთეულს;
6. ნეირონს აქვს რეგენერაციის უნარი.

რუსეთის მეცნიერებათა აკადემიის თეორიული და ექსპერიმენტული ბიოფიზიკის ინსტიტუტის მეოცე საუკუნის ეკლევების შემაჯამებელ მოხსენებაში „ნეირომეცნიერება მეორე ათასწლეულის ბოლოს: პარადიგმების ცვლა“ [4], მოცემულია ბოლო წლებში დაგროვილი ახალი ნეირობიოლოგიური მონაცემების ანალიზი, რასაც, როგორც აღნიშნულია მოხსენებაში, მიყავართ დიდი ხნის წინ ჩამოყალიბებული და საყოველთაოდ მიღებული პოსტულატების რადიკალურ გადახედვასთან, ანუ საჭირო ხდება ნერვული უჯრედების სიცოცხლისა და მუშაობის შესახებ ახალი წარმოდგენების შემუშავება.

ბუნებრივი ნეირონის ფუნქციონირების შესახებ შეხედულებები მნიშვნელოვნად იცვლება, სახელდობრ: ცალკეულ ბუნებრივ ნეირონს შეუძლია იმოქმედოს არა ერთი, არამედ რამდენიმე ტრანსმიტერული ნივთიერების გამოყოფის მეშვეობით; დენდრიტებზე ზემოქმედებამ შეიძლება გამოიწვიოს აგზნების აქტიური გავრცელება; აგზნების პარამეტრების ზეგავლენით აქსონებს შეუძლიათ დიფერენცირებულად გამოათავისუფლონ ნეიროტრანსმიტერები. ასტროგლია მონაწილეობს ყველა ნეირონული ფუნქციის ორგანიზებასა და რეგულირებაში; იგი მნიშვნელოვან როლს ასრულებს აგრეთვე ნეირონების ექსიტოტოქსიკური დაღუპვისაგან დამცავი ნეიროტრანსმიტერებისა და სხვა ნივთიერებების სინთეზში. ნეირონებს შორის სინაფსური ურთიერთქმედების გარდა ბოლო წლებში იკვლევენ განსხვავებული ტიპის კომუნიკაციების სხვადასხვა სისტემებს. მაგალითად, სინაფსების უმრავლესობა წარმოადგენს ღია კონტაქტს, საიდანაც ტრანსმიტერი შეიძლება გავრცელდეს როგორც სამგანზომილებიანი დიფუზიური სიგნალი

იზოტროპიულ ფოროვან გარემოში, ან ტრანსმიტერი შეიძლება გამოიყოს არასინაფსური ვარიაციისგან; ნაპრალოვანი კონტაქტები აგრეთვე წარმოადგენს ნეირონთაშორისი მეტაბოლური და ელექტრული კომუნიკაციების დამატებით სისტემას, სადაც ხდება უშუალოდ იონებისა და მცირე მოლეკულების გაცვლა – კავშირის ეს ფორმა ნეირონების, აგრეთვე ნეირონებისა და ასტროციტების გაერთიანებაში მონაწილეობს ფუნქციურ ანსამბლებად.

ნეირონთაშორისი კომუნიკაციების რთული სურათი, რომელიც შეიცავს მრავლობით სინაფსურ, პრესინაფსურ და პარასინაფსურ ურთიერთქმედებას, კიდევ უფრო რთულდება ამ პროცესებში ნეიროტროფული ნივთიერებებისა და ციტოკინების მონაწილეობის შედეგად. იმ ფაქტორების რაოდენობის ზრდა, რომლებიც განსაზღვრავს ნეირონების ურთიერთქმედების საბოლოო შედეგს, ქმნის ახალ სიძნელეებს ტვინის ფუნქციების თეორიული კონცეფციებისა და მოდელირების განვითარებაში.

ამრიგად, ბოლო წლებში მოლეკულური ნეირობიოლოგიისა და ნეიროგენეტიკის მიერ გამოვლენილი სხვადასხვა ნეირობიექტებს შორის რთული ურთიერთქმედებების დეტალების სიმრავლე (და ეს პროცესი ყოველდღიურად მიმდინარეობს) თეორიული ნეირობიოლოგიის მიერ არ არის ჯერჯერობით ინტეგრირებული ტვინის ფუნქციების ორგანიზაციის კოჰერენტულ წარმოდგენაში, ეს მონაცემები, რა თქმა უნდა, აფართოებს ცოდნას, მაგრამ, ამავე დროს, ართულებს ტვინისა და მისი განყოფილებებისა და ცალკეული სისტემების ინტეგრალური მუშაობის პრინციპების კონცეპტუალური წარმოდგენების ფორმირებას.

ნეირონიდან გამომავალი საბოლოო სიგნალი დამოკიდებულია პირობების იმდენად დიდ სიმრავლეზე, რომ ჰიპოთეზური თეორიული კონსტრუქციების საიმედოობა სულ უფრო მცირდება, ეს ეხება ნერვული სისტემის მოდელირების პრობლემასაც. ადრე, როდესაც

ნეირონებს წარმოიდგენდნენ როგორც ამგზნები ან მამუხრუჭებელი მუდმივი თვისებების მქონე აბსტრაქტულ ობიექტებს, მათი მოდელირება შედარებით მარტივი იყო. ბოლო წლების მონაცემების გათვალისწინებით კი ნეირონთაშორის ურთიერთქმედებებში შესაძლებელი ხდება მრავალი ვარიანტის მონახვა, რაც პრაქტიკულად გზას უხსნის განსხვავებული თეორიებისა და მოდელების შემუშავებას.

ნეირობიოლოგია არის მრავალდონიანი მეცნიერება, რომელიც კვლევებს აწარმოებს მოლეკულურ – სუბუჯრედულ – უჯრედულ – სტრუქტურულ – სტრუქტურათშორის (ტვინის) – კოგნიტიურ (ქცევით) დონეებზე. თითოეული დონე იყენებს წინამდებარე დონის სინთეზურ მონაცემებს და საანალიზო მონაცემებით უზრუნველყოფს მომდევნო დონეს. მდენად, ლოგიკური და მოდელური კონსტრუირების პროცესებში გათვალისწინებული უნდა იყოს თითოეულ დონეზე მიღებული შედეგები.

13. მაღალი პარალელიზმის გამოთვლითი სისტემები

13.1. კომპიუტერების არქიტექტურის ევოლუციის ძირითადი ეტაპები

გამოთვლითი ტექნიკის სისტემების ყველაზე უფრო გავრცელებული არქიტექტურა ეფუძნება ჯონ ფონ ნეიმანის [1] პარადიგმას, ანუ ეს სისტემები სტრუქტურულად შედგება მართვის მოდულისაგან, აგრეთვე მესხიერების, ინფორმაციის შეყვანა-გამოყვანის, ლოგიკური და არითმეტიკული დამუშავების ფუნქციური მოდულებისაგან, რომლებიც წინასწარ შედგენილი ალგორითმის მიხედვით ასრულებენ ინფორმაციის

დამუშავების ოპერაციებს. მათემატიკური პროგრამირების საფუძველზე მომუშავე კომპიუტერების სტრუქტურა ხისტი და არ იცვლება (hardware), მაგრამ შესაძლებელია იცვლებოდეს პროგრამული უზრუნველყოფა (software).

მწარმოებლობის, საიმედოობის, ტექნოლოგიურობის, გამოყენების არეალის გაზრდისა და, ამავე დროს, მართვისა და დაპროგრამების, კონტროლისა და დიაგნოსტიკის პროცესების გამარტივების მიზნით, განუწყვეტლივ მიმდინარეობს კომპიუტერების სტრუქტურისა და ფუნქციების განმსაზღვრელი კონცეფციების განვითარება.

ანალიზი გვინვენებს, რომ ბოლო ათწლეულების განმავლობაში გამოთვლითი სისტემების არქიტექტურის განვითარების ძირითადი ხაზი პარალელიზმისა და ერთგვაროვნების ხარისხის ამაღლებას უკავშირდება. ცნობილია, რომ გამოთვლითი ტექნიკის ორმოცდაათწლიანი ისტორიის განმავლობაში გამოთვლითი მანქანების მწარმოებლურობა გაიზარდა დაახლოებით 10^7 -ჯერ, აქედან $5 \cdot 10^2$ ჯერ ელემენტური ბაზის სწრაფქმედების ამაღლების ხარჯზე და $2 \cdot 10^4$ - ჯერ ახალი არქიტექტურული კონცეფციების წყალობით, რასაც ძირითადად ინფორმაციის დამუშავების პარალელიზმის ამაღლება (მათ შორის, კონვეიერული დამუშავება) უზრუნველყოფს. ასე მაგალითად, გამოთვლითი მანქანა EDSAC (1949წ.) ასრულებდა 1 წამში 100 არითმეტიკულ ოპერაციას (ტაქტის დრო იყო 2 მიკროწამი), ხოლო თანამედროვე სუპერკომპიუტერ GRAY C90-ის მწარმოებლობა 1 წამში აღწევს 1 მილიარდ არითმეტიკულ ოპერაციას (ტაქტის დრო - 4 ნანოწამი). მრიგად, კომპიუტერის მწარმოებლობის ზრდას ძირითადად მისი არქიტექტურის სრულყოფა განაპირობებდა. სავარაუდოა, რომ ეს ტენდენცია გადრმავედება, ვინაიდან ელემენტური ბაზის სწრაფქმედება ზღვარს უახლოვდება.

ხელოვნური ნეირონული ქსელების იდეოლოგიის ერთ-ერთი წყარო, როგორც ხაზგასმულია ნაშრომში [2], უკავშირდება ინფორმაციის დამუშავების პარალელური

ალგორითმების განსაკუთრებულ წარმოდგენას, რაც ორიენტირებულია იმ სპეციალიზებული მაღალი პარალელუზიმის პროცესორების ფუნქციონირებაზე, რომლებიც ერთგვაროვან ოპერაციებს ასრულებენ მონაცემთა დროში სწრაფად ცვალებად დიდ მასივებზე. ამასთან დაკავშირებით საინტერესოა კომპიუტერების პარალელუზიმისა და ერთგვაროვნების ხარისხის ამაღლების ძირითადი ეტაპების განხილვა.

გამოთვლითი სისტემების მაღალსიჩქარიანი ელექტრონული მოწყობილობებისა და დაბალსიჩქარიანი ელექტრომექანიკური მოწყობილობების მუშაობის შეთანაწყობის მიზნით, კ. სტრეინიმ [5] წამოაყენა პროცესორის დროის დაყოფის იდეა სხვადასხვა პროგრამებს (ამოცანებს) შორის, რამაც უზრუნველყო პროცესორისა და სხვა მოდულების მუშაობის უმარტივესი განპარალელება. ამ ნაბიჯით პროცესორი უფრო თანაბრად და ერთგვაროვნად დაიტვირთა დროში, ხოლო პროგრამების ნაკადმა უფრო ერთგვაროვანი სტრუქტურა მიიღო მსხვილი პროგრამების დანაწევრების გამო.

მიკროპროგრამული მართვის მეთოდი შეიმუშავა მ. უილკსმა [6], რითაც პროგრამული მართვის პრინციპი განავრცო კომპიუტერის მართვის სქემებზე. ამით მართვის მოდულმა მიიღო მეტი ერთგვაროვნება და ე.წ. მმართველი მექანიზმები, რომლის გამოსასვლელებიც უშუალოდ უკავშირდება შემსრულებელი სქემების შესასვლელებს.

დ. სლოტნიკმა [7] წამოაყენა მატრიცული პროცესორის იდეა, ანუ გაზარდა ინფორმაციის დამმუშავებელი სქემების სიერცული ერთგვაროვნება და პარალელიზმი. ამასთან, თითოეული სქემა, რომელიც მრავალჯერ მეორდებოდა, წარმოდგენდა მრავალფუნქციურ მოდულს. მატრიცულ პროცესებში მნიშვნელოვნად ადვილდება ორგანიზომილებიანი მასივების დამუშავება, მარტივდება დაპროგრამება და პროცესორის მართვა, იზრდება მწარმოებლობა, განსაკუთრებით, ორგანიზომილებიანი

ინფორმაციული ნაკადების, მატრიცული ალგებრისა და სხვა მსგავსი ამოცანების გადაწყვეტისას.

შემდეგი ნაბიჯი იყო გამოთვლითი სისტემების მართვის ფუნქციების (ცალკეული მოწყობილობებისა და ინფორმაციის ნაკადების მართვა, სხვა ოპერატორული ფუნქციები) აპარატურული ან პროგრამული ავტომატიზაცია. სისტემის მართვის ამ ნაწილს წარმოადგენს თანამედროვე ოპერაციული სისტემა.

უკუკავშირის სიგნალების გამოყენების ფონ ნეიმანის კონცეფციის განვითარებად უნდა ჩაითვალოს კომპიუტერის ზოგიერთი ტიპური მდგომარეობის ამოცნობის ფუნქციის განხორციელება (მიმდინარე დროის ანალიზი და კონტროლი, თანრიგების გადავსების გამოვლენა, მეხსიერების გამოყოფილი უბნიდან გამოსვლის კონტროლი და სხვ.), რაც კომპიუტერის სტრუქტურის მნიშვნელოვან გაუმჯობესებას წარმოადგენდა.

„ვირტუალური“ მეხსიერების კონცეფციამ, ანუ მეხსიერების სხვადასხვა დონეებს შორის ინფორმაციის ოპერატიული გადასროლის მექანიზმმა შედარებით გააადვილა ფიზიკურად მრავალდონიან მეხსიერებასთან ურთიერთობა. ეს შეიძლება განვიხილოთ იმის ცდად, რომ მომხმარებლისათვის მეხსიერება ყოფილიყო ერთდონიანი, ანუ უფრო ერთგვაროვანი.

ამრიგად, კომპიუტერის სტრუქტურისა და ფუნქციების გასაუმჯობესებლად შემუშავებული კონცეფციები ძირითადად მიმართულია, ერთი მხრივ, მისი პარალელულობის და, მეორე მხრივ, ერთგვაროვნების გაზრდისაკენ. ფონ ნეიმანის პარადიგმის მნიშვნელოვანმა განვითარებამ ხელი შეუწყო პარალელური არქიტექტურის მქონე მრავალი სხვადასხვა სახის კომპიუტერული სისტემების შექმნას.

13.2. კომპიუტერების არქიტექტურის კლასიფიკაცია

ლიტერატურაში აღწერილია კომპიუტერების არქიტექტურის კლასიფიკაციის სხვადასხვა სქემები. მათ შორის ყველაზე ხშირად მიმართავენ ფლინის ტაქსონომიას, რომელიც ეფუძნება კომპიუტერების მუშაობის აღწერას მონაცემთა ნაკადებისა და ბრძანებების ნაკადების ცნებების გამოყენებით. ნაკადის ცნება აქ გულისხმობს პროცესორების მიერ დამუშავებული ბრძანებების ან მონაცემების ელემენტების თანამიმდევრობას.

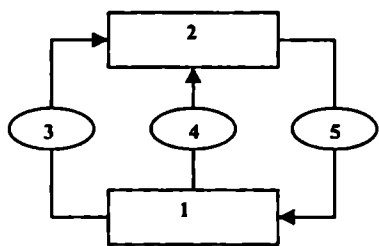
ფლინის კლასიფიკაცია (სურ. 1.5) ყველა შესაძლო სახის აქიტექტურის კომპიუტერებს ყოფს ოთხ კატეგორიად:

ა. **SISD** – კომპიუტერები (Single Instruction Stream – Single Date Stream) – ბრძანებების ერთი ნაკადი და მონაცემთა ერთი ნაკადი;

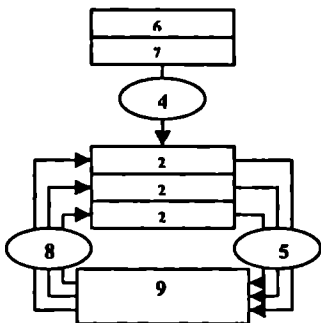
ბ. **SIMD** – კომპიუტერები (Single Instruction Stream – Multiple Date Stream) ბრძანებების ერთი ნაკადი და მონაცემთა ნაკადების სიმრავლე;

გ. **MISD** – კომპიუტერები (Multiple Instruction Stream – Single Date Stream) – ბრძანებების ნაკადების სიმრავლე და მონაცემთა ერთი ნაკადი;

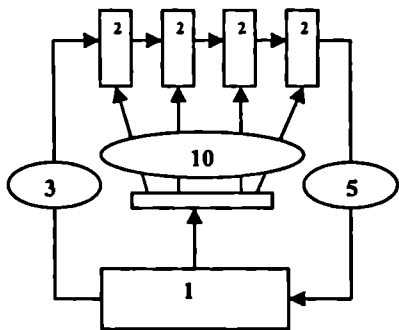
დ. **MIMD** – კომპიუტერები (Multiple Instruction Stream – Multiple Date Stream) – ბრძანებების ნაკადების სიმრავლე და მონაცემთა ნაკადების სიმრავლე.



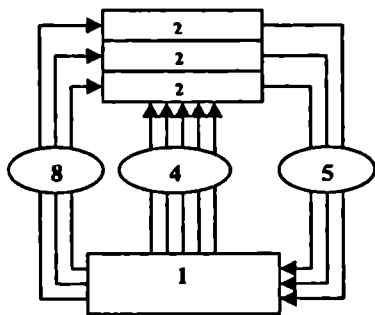
ა.



ბ



გ



დ

სურ. 15. კომპიუტერების არქიტექტურის კლასიფიკაცია

1.5 სურათზე მიღებულია შემდეგი აღნიშვნები:

1. მეხსიერების მოდული;
2. პროცესორი;
3. მონაცემთა ერთი ნაკადი;
4. ბრძანებების ერთი ნაკადი;
5. შედეგები;
6. პროგრამების მეხსიერება;
7. მართვის მოდული;
8. მონაცემთა ნაკადების სიმრავლე;
9. მონაცემთა მეხსიერება;
10. ბრძანებების ნაკადების სიმრავლე.

SISD კლასის კომპიუტერებს (სურ. 1.5 ა) მიეუთვნება კლასიკური კომპიუტერები ანუ ფონ ნეიმანის მანქანები, სადაც ბრძანებების მხოლოდ ერთი ნაკადია, თითოეული ბრძანება ინიცირებს ერთ ოპერაციას მონაცემთა ერთ ნაკადზე. ამ კლასს ზოგჯერ მიაკუთვნებენ ისეთ მანქანებსაც, სადაც არითმეტიკული ოპერაციების დასაჩქარებლად გამოიყენება კონვეიერული დამუშავება (მაგალითად, კომპიუტერები CDC 6600 და CDC 7600).

SIMD კლასის მანქანებში (სურ. 1.5 ბ) ბრძანებების ერთი ნაკადია, მაგრამ ეს ნაკადი, განსხვავებით წინა კლასისაგან, შეიცავს ვექტორულ ბრძანებებს, რაც შესაძლებლობას იძლევა ერთი ჯგუფური ოპერაცია ჩაუტარდეს ერთდროულად მონაცემთა სიმრავლეს – ვექტორების ელემენტებს. ვექტორული ოპერაციები შეიძლება ჩატარდეს სხვადასხვა მეთოდით, მაგალითად, პროცესორების მატრიცით (ILIAC-IV) ან კონვეიერის საშუალებით (CRAY-1).

MISD კლასში (სურ. 1.5 გ) მოიაზრება ისეთი კომპიუტერული სისტემები, სადაც პროცესორთა სიმრავლე მონაცემთა ერთ ნაკადს ამუშავებს. მკვლევართა ნაწილი ამ კლასში აერთიანებს კონვეიერულ კომპიუტერებს, ნაწილი კი თვლის, რომ ეს კლასი ჯერჯერობით ცარიელია.

MIMD კლასი (სურ. 1.5 დ) არის ყველაზე წარმომადგენლობითი, ვინაიდან მოიცავს სხვადასხვა სახის არქიტექტურის მქონე მასობრივი პარალელიზმის სისტემებს, დაწყებული ერთთანრიგიანი ლოკალურ მეხსიერებიანი პროცესორების მატრიცით და დამთავრებული

ზემაღალი მწარმოებლურობის მულტიპროცესორული სისტემებით (Intel Paragon, Denelcor HEP, BBN Butterfly, CRAY T3D და მრავალი სხე.).

13.3. ნეიროკომპიუტერები და პარალელური პროცესორები.

ნეიროკომპიუტერები, რომელთა საფუძველია ხელოვნური ნეირონული ქსელები, შეიძლება მივაკუთვნოთ მაღალი პარალელიზმის კომპიუტერების MIMD კლასს [19]. სიგნალების პარალელური დამუშავების პრინციპის განხორციელება მიიღწევა დიდი რაოდენობით ფორმალური ნეირონების სხვადასხვა კონფიგურაციის მქონე ჯგუფებად ან შრეებად გაერთიანებით, ხოლო შრეების ერთმანეთთან კასკადური ან სხვა ტიპის შეერთებით.

ხელოვნური ნეირონული ქსელების თემატიკა დისციპლინათაშორისი ხასიათისაა, იგი ინტენსიური ფორმირების სტადიაში იმყოფება; ამიტომ ნეიროკომპიუტერის ერთმნიშვნელოვანი განსაზღვრება, რომელიც მისაღები იქნებოდა სხვადასხვა სამეცნიერო მიმართულები-სათვის, ჯერ არ არსებობს. მაგალითად, ნაშრომში [19] ნეიროკომპიუტერი ზოგადად განსაზღვრულია, როგორც აპარატურული და პროგრამული უზრუნველყოფის ისეთი არქიტექტურის მქონე გამოთვლითი სისტემა, რომელიც ნეიროქსელურ ლოგიკურ ბაზისში წარმოდგენილი ალგორითმების შესრულების ადეკვატურია.

ჩვენი აზრით, გამოთვლითი ტექნიკისა და ინფორმატიკის კუთხით შესაძლებელია ნეიროკომპიუტერის განსაზღვრება ჩამოყალიბდეს შემდეგი სახით:

განსაზღვრება 1. ნეიროკომპიუტერი არის MIMD არქიტექტურის გამოთვლითი სისტემა, სადაც ერთგვაროვანი სტრუქტურის პროცესორული ელემენტი გამარტივებულია ფორმალური ნეირონის დონემდე; ამასთან, გადასაწყვეტი ამოცანების ადეკვატურად არის გართულებული ელემენტთაშორისი კავშირები, ხოლო გამოთვლითი სისტემის დაპროგრამება გამოიხატება ფორმალური ნეირონებისა და ნეირონთაშორისი კავშირების პარამეტრების ცვლილებებით.

შემოთავაზებული განსაზღვრება მკაცრი არ არის და შესაძლებელია მისი შემდგომი დახვეწა.

ტერმინი „ნეიროკომპიუტერი“ ზოგჯერ იხმარება აგრეთვე ხელოვნური ინტელექტის სისტემათა აგების სამუშაოთა მთელი სპექტრის აღსანიშნავად იმ მიდგომის ჩარჩოებში, რომელიც ეფუძნება ცენტრალური ნერეული სისტემის სხვადასხვა ელემენტებისა და სტრუქტურების ფუნქციების მოდელირებას. იმის გამო, რომ ამჟამად გამოკვლევები ძირითადად ნეირონული ქსელების მოდელების დონეზე მიმდინარეობს, ხშირად ტერმინი „ნეიროკომპიუტერი“ ნეირონული ქსელების სინონიმადაც მოისაზრება.

რეალიზების ხერხის შესაბამისად ნეიროკომპიუტერები შეიძლება დაიყოს ოთხ დონედ:

- I. თეორიული დონე. ნეირონული ქსელების მოდელები წარმოდგენილია მათემატიკური, ალგორითმული, სიტყვიერი ან რაიმე სხვა ფორმის აღწერების სახით;
- II. პროგრამული დონე. ნეირონული ქსელების მოდელები რეალიზებულია პროგრამულად მიმდევრობით (ფონ ნეიმანის) კომპიუტერებზე;
- III. პროგრამულ-აპარატურული დონე. ნეირონული ქსელების მოდელირების პროცესის დასაჩქარებლად რეალიზებულია სპეციალური თანაპროცესორი;
- IV. აპარატურული დონე. ნეირონული ქსელები მთლიანად რეალიზებულია ფიზიკურად.

ნეირონული ქსელების მოდელების სტრუქტურისა და ფუნქციონირების მაღალი პარალელიზმი, აგრეთვე ნეირონული ოპერაციების სპეციფიკურობა მნიშვნელოვნად აძნელებს მათ რეალიზებას მიმდევრობით კომპიუტერებზე, ამიტომ ბოლო წლებში გამოჩნდა ნეირონული ქსელების მოდელების სპეციალიზებული გამოთვლითი მოწყობილობები – ნეიროკომპიუტერები ტერმინის ვიწრო გაგებით, რომლებიც შეესაბამება ზემოთმოყვანილი კლასიფიკაციის III და IV დონეებს, თუმცა, ნეიროქსელური ტექნოლოგიები პარალელიზმის გამოყენების გარეშეც ფუნქციური აპარატია რიგი ამოცანების გადაწყვეტისათვის.

მრიგად, ნეიროკომპიუტინგის იდეა უკავშირდება პარალელურ ერთგვაროვან პროცესორებს. განვიხილოთ ამ კლასის ზოგიერთი წარმომადგენლის ძირითადი არქიტექტურული მახასიათებლები.

1.3.4. პარალელური მატრიცული პროცესორები

პარალელური კომპიუტერების მნიშვნელოვანი სეგმენტი უკავია მატრიცულ პარალელურ კომპიუტერებს. მატრიცული კომპიუტერების დამახასიათებელი ნიშანია მათი მოდულების მატრიცული განლაგება, როდესაც ყველა მოდული ერთნაირია და ერთნაირი წესით არის შეერთებული მეზობელი მოდულების გარკვეულ რაოდენობასთან. ასეთი სისტემების მაგალითებია სისტემა SOLOMON და მისი მომდევნო სისტემა ILLIAC-IV, რომელიც წარმოადგენს 64 პროცესორის (8X8) მატრიცას, სადაც ცენტრალური მართვის ფუნქციას ასრულებს კომპიუტერი B6500. მატრიცული ამოცანების გადაწყვეტისას სისტემის მწარმოებლობა 200 მილიონ ოპ/წმ-ს აღწევს.

მატრიცული პარალელური პროცესორების ერთ-ერთი სახეობა არის ასოციაციური პარალელური პროცესორი [8], რომელიც დიდი რაოდენობის (1000-4000) ოპერანდების წყვილებზე ერთდროულად აწარმოებს იდენტურ ლოგიკურ და არითმეტიკულ ოპერაციებს ისეთი ამოცანების გადაწყვეტისას, სადაც შესაძლებელია ჯგუფური ოპერაციების შესრულება. ინფორმაციის მასივების დამუშავების დროს მისი მწარმოებლობა არის 100 მილიონი ოპ/წმ, რაც მიიღწევა არა მაღალი სწრაფქმედების ელემენტური ბაზის გამოყენებით, არამედ სტრუქტურული მეთოდების ხარჯზე. ასეთ პროცესორებში ინფორმაციის ჩაწერა, წაკითხვა და ძიება ხორციელდება არა მისამართებით (როგორც ტრადიციულ კომპიუტერში), არამედ შინაარსობრივი (ასოციაციური) ნიშნების მიხედვით. პროცესორის ძირითადი ნაწილია მეხსიერების მატრიცა, რომელიც შედგება $m \times n$ უჯრედისაგან, სადაც m არის მანქანური სიტყვების რაოდენობა, ხოლო n – სიტყვის თანრიგების რაოდენობა. თითოეული უჯრედი შეიცავს მეხსიერებას (ტრიგერი) და ინფორმაციის ჩაწერის, შენახვის, ძიების, შედარებისა და წაკითხვის მართვის სქემას. ასეთ ასოციაციურ მატრიცაში არითმეტიკული, ლოგიკური და საძიებო ოპერაციის შესრულება ხორციელდება მიკროპროგრამების მეშვეობით, რაც აღწერილია [8] ნაშრომში.

მართალია, მატრიცულ პარალელურ პროცესორებს ორი-სამი რიგით უფრო მაღალი მწარმოებლობა აქვთ ვიდრე ტრადიციულ კომპიუტერებს, მაგრამ ასეთი პროცესორები ხისტი კავშირებით ეფექტურია შედარებით ვიწრო კლასის კარგად განპარალელებადი ამოცანებისათვის. მათი ეფექტურობის გაზრდა კავშირების ცვლილებების პროგრამულ მართვას მოითხოვს.

აღწერილი ასოციაციური პროცესორი პარალელურად ამუშავებს ოპერანდების სიმრავლეს სიტყვების მიხედვით, მაგრამ მიმდევრობით – თანრიგების მიხედვით, რაც არითმეტიკული ოპერაციების შესასრულებლად

ნეირონული ქსელების მოდელების სტრუქტურისა და ფუნქციონირების მაღალი პარალელიზმი, აგრეთვე ნეირონული ოპერაციების სპეციფიკურობა მნიშვნელოვნად აძნელებს მათ რეალიზებას მიმდევრობით კომპიუტერებზე, ამიტომ ბოლო წლებში გამოჩნდა ნეირონული ქსელების მოდელების სპეციალიზებული გამოთვლითი მოწყობილობები – ნეიროკომპიუტერები ტერმინის ვიწრო გაგებით, რომლებიც შეესაბამება ზემოთმოყვანილი კლასიფიკაციის III და IV დონეებს, თუმცა, ნეიროქსელური ტექნოლოგიები პარალელიზმის გამოყენების გარეშეც ეფექტური აპარატია რიგი ამოცანების გადაწყვეტისათვის.

მრიგად, ნეიროკომპიუტინგის იდეა უკავშირდება პარალელურ ერთგვაროვან პროცესორებს. განვიხილოთ ამ კლასის ზოგიერთი წარმომადგენლის ძირითადი არქიტექტურული მახასიათებლები.

13.4. პარალელური მატრიცული პროცესორები

პარალელური კომპიუტერების მნიშვნელოვანი სეგმენტი უკავია მატრიცულ პარალელურ კომპიუტერებს. მატრიცული კომპიუტერების დამახასიათებელი ნიშანია მათი მოდულების მატრიცული განლაგება, როდესაც ყველა მოდული ერთნაირია და ერთნაირი წესით არის შეერთებული მეზობელი მოდულების გარკვეულ რაოდენობასთან. ასეთი სისტემების მაგალითებია სისტემა SOLOMON და მისი მომდევნო სისტემა ILLIAC-IV, რომელიც წარმოადგენს 64 პროცესორის (8X8) მატრიცას, სადაც ცენტრალური მართვის ფუნქციას ასრულებს კომპიუტერი B6500. მატრიცული ამოცანების გადაწყვეტისას სისტემის მწარმოებლობა 200 მილიონ ოპ/წმ-ს აღწევს.

მატრიცული პარალელური პროცესორების ერთ-ერთი სახეობა არის ასოციაციური პარალელური პროცესორი [8], რომელიც დიდი რაოდენობის (1000-4000) ოპერანდების წყვილებზე ერთდროულად აწარმოებს იდენტურ ლოგიკურ და არითმეტიკულ ოპერაციებს ისეთი ამოცანების გადაწყვეტისას, სადაც შესაძლებელია ჯგუფური ოპერაციების შესრულება. ინფორმაციის მასივების დამუშავების დროს მისი მწარმოებლობა არის 100 მილიონი ოპ/წმ, რაც მიიღწევა არა მაღალი სწრაფქმედების ელემენტური ბაზის გამოყენებით, არამედ სტრუქტურული მეთოდების ხარჯზე. ასეთ პროცესორებში ინფორმაციის ჩაწერა, წაკითხვა და ძიება ხორციელდება არა მისამართებით (როგორც ტრადიციულ კომპიუტერში), არამედ შინაარსობრივი (ასოციაციური) ნიშნების მიხედვით. პროცესორის ძირითადი ნაწილია მეხსიერების მატრიცა, რომელიც შედგება $m \times n$ უჯრედისაგან, სადაც m არის მანქანური სიტყვების რაოდენობა, ხოლო n – სიტყვის თანრიგების რაოდენობა. თითოეული უჯრედი შეიცავს მეხსიერებას (ტრიგერი) და ინფორმაციის ჩაწერის, შენახვის, ძიების, შედარებისა და წაკითხვის მართვის სქემას. ასეთ ასოციაციურ მატრიცაში არითმეტიკული, ლოგიკური და საძიებო ოპერაციის შესრულება ხორციელდება მიკროპროგრამების მეშვეობით, რაც აღწერილია [8] ნაშრომში.

მართალია, მატრიცულ პარალელურ პროცესორებს ორი-სამი რიგით უფრო მაღალი მწარმოებლობა აქვთ ვიდრე ტრადიციულ კომპიუტერებს, მაგრამ ასეთი პროცესორები ხისტი კავშირებით ეფექტურია შედარებით ვიწრო კლასის კარგად განპარალელებადი ამოცანებისათვის. მათი ეფექტურობის გაზრდა კავშირების ცვლილებების პროგრამულ მართვას მოითხოვს.

აღწერილი ასოციაციური პროცესორი პარალელურად ამუშავებს ოპერანდების სიმრავლეს სიტყვების მიხედვით, მაგრამ მიმდევრობით – თანრიგების მიხედვით, რაც არითმეტიკული ოპერაციების შესასრულებლად

გრძელი მიკროპროგრამების აუცილებლობას იწვევს და ზრდის ოპერაციების შესრულების დროს.

ნაშრომებში [9], [10] აღწერილია ასოციაციური პარალელური პროცესორი მოდულის გაფართოებული ფუნქციებით, რომლის ლოგიკური სტრუქტურა უზრუნველყოფს 10-15-ჯერ უფრო მაღალ მწარმოებლობას [8] ნაშრომში აღწერილ პროცესორთან შედარებით. მწარმოებლობის ამაღლება მიიღწევა მოდულში – პროცესორის ძირითად შემადგენელ ელემენტში (მოდულის რეალიზაციით დიდი ინტეგრალური სქემის ერთ კრისტალზე) არითმეტიკული, მართვისა და კომუტაციის სქემების შეყვანით, რაც არითმეტიკული და სხვა ფუნქციების აპარატურული შესრულების შესაძლებლობას იძლევა. ეს კი, შესაბამისად, მნიშვნელოვნად ამოკლებს დამუშავების მიკროპროგრამებს, ამცირებს ოპერაციების დროსა და მიკროპროგრამული მეხსიერებისა და მმართველი მოწყობილობების მოცულობას, ანუ ამცირებს პროცესორის არარეგულარული ნაწილის ხვედრით წონას და ზრდის მის ერთგვაროვნებას. ასეთი ასოციაციური პარალელური პროცესორი მოდულის გაფართოებული ფუნქციებით ასრულებს ჯგუფურ არითმეტიკულ და ლოგიკურ ოპერაციებს პრაქტიკულად ნებისმიერი მოცულობის მონაცემთა მასივებზე. გამრავლების, გაყოფის და სხვა ჯგუფური ოპერაციების ალგორითმები და ლოგიკური ორგანიზაცია შემუშავებულია ნაშრომებში [11], [12], იქვე მოცემულია ჯგუფური ოპერაციების შესასრულებლად საჭირო დროის დანახარჯების შეფასებები.

აღწერილ პარალელურ ასოციაციურ პროცესორთან შედარებით მეტი ფუნქციური შესაძლებლობები და მაღალი მწარმოებლობა ახასიათებს [13] ნაშრომში წარმოდგენილ მატრიცულ ასოციაციურ პროცესორს. იგი შედგება იდენტური ელემენტარული პროცესორების სიმრავლისაგან, რომლებიც განლაგებულია მატრიცულად, თითოეული მათგანი უკავშირდება ოთხივე მეზობელს. ცალკეული ელემენტარული პროცესორი შეიცავს

მეხსიერების ბლოკს, შემავალ ასოციაციურ ბლოკს, არითმეტიკულ-ლოგიკურ ასოციაციურ ბლოკს, გამო-
მავალ ასოციაციურ ბლოკს. მონაცემთა შეყვანის,
დამუშავების, გამოყვანის პროგრამები მუშაობს ერთდრო-
ულად და ერთმანეთისგან დამოუკიდებლად, ბლოკების
აქტივიზაცია ხდება შესაბამისი ასოციაციური
რეგისტრებით.

მატრიცული ასოციაციური პროცესორის მაღალი
მწარმოებლობა განპირობებულია მასში გამოყენებული
სხვადასხვა სახის პარალელიზმით, სახელდობრ,
პარალელურად და ერთდროულად მუშაობს ელემენ-
ტარული პროცესორების სიმრავლე, ახალი მონაცემები
პარალელურად შედარდება მონაცემების ყველა ნაკრებს
და ხდება მათი შეყვანა შერჩეულ ელემენტარულ
პროცესორებში, საჭირო შედეგების გამოყვანა ხდება
გამომავალი ასოციაციური ბლოკებიდან პარალელური
ძიების გზით, ხოლო ოპერაციები შეიძლება შესრულდეს
ერთდროულად მონაცემთა სხვადასხვა ქვესიმრავლეებზე.

პარალელური პროცესორების მიერ მაღალი
მწარმოებლობის მისაღწევად გადამწყვეტი მნიშვნელობა
ენიჭება პროცესორსა და გარე მოწყობილობებს შორის
ინფორმაციის გაცვლის ეფექტურ ორგანიზაციას, ხოლო
ინფორმაციის განახლების დრო მაღალმწარმოებლური
პარალელური კომპიუტერების ერთ-ერთი უმნიშვნე-
ლოვანესი ხარისხობრივი მაჩვენებელია.

ნაშრომებში [14, 15] გამოკვლეულია პარალელურ
პროცესორსა და გარე მოწყობილობებს შორის ინფორმა-
ციის გაცვლის სხვადასხვა (V1, V2, H1, H2) მეთოდები
ინფორმაციის გაცვლის ერთი და ორი რეგისტრის ბაზა-
ზე. მოცემულია ამ მეთოდებისათვის დროითი მახა-
სიათებლების გაანგარიშებები, აგრეთვე შესაბამისი აპა-
რატურის მოცულობების შეფასება. ამ მეთოდების ანალი-
ზის საფუძველზე მოცემულია პარალელური პროცესორ-
ის მუშაობის შესაძლო რეჟიმები, რომლებიც შესაძლებ-
ლობას იძლევა შემცირდეს პროცესორის მართვის აპარა-
ტურის მოცულობა.

ნაშრომში [16] შემოთავაზებულია სამგანზომილებიანი ასოციაციური პროცესორის არქიტექტურა, რომელიც მინიმუმამდე ამცირებს ასოციაციურ მატრიცაში ინფორმაციის განახლების დროს. ზოგად შემთხვევაში სამგანზომილებიანი პარალელური პროცესორის სტრუქტურაში შეიძლება შედიოდეს ასოციაციური მატრიცები, ბუფერული მეხსიერებისა და ერთგვაროვანი გადაწყობადი სტრუქტურის შრეები, აგრეთვე სხვა ტიპის (მათ შორის, ნეირონული ქსელების) შრეები. განსაკუთრებით აღსანიშნავია, რომ პროცესორის შიგა ფენებს შესაძლებელია არ ჰქონდეს კავშირები გარემოსთან, რაც მნიშვნელოვანია დიდი და ზედიდი ინტეგრალური სქემებით პროცესორის რეალიზებისას მათი გამოყვანების მინიმიზაციის თვალსაზრისით. შემოტანილია ინფორმაციის დამუშავების არხის ცნება, რომლის სიგრძესაც არა აქვს პრინციპული შეზღუდვა და, შესაბამისად, არ არის შეზღუდული სამგანზომილებიანი პარალელური პროცესორის შრეების რაოდენობაც. დამუშავების არხის ელემენტებს შეიძლება ჰქონდეს ლოგიკური სქემები და უკუკავშირები სხვადასხვა კომპონენტებს შორის, რაც შესაძლებლობას იძლევა შრეში მოთავსებული მასივები დამუშავდეს ერთი ან რამდენიმე ტაქტის განმავლობაში. შრეები ავტონომიურია, რაც განაპირობებს მასივების ერთდროული შეყვანის, დამუშავების და გამოყვანის ჯგუფური ოპერაციების განხორციელების შესაძლებლობას, აგრეთვე მუშა რეჟიმში მის პარალელურად შრეების კონტროლისა და დიაგნოსტიკის პროცედურების ჩატარებას. ყოველივე ეს უზრუნველყოფს ასეთი პროცესორის განსაკუთრებით მაღალ პარალელიზმს, მწარმოებლობასა და საიმედოობას.

ასოციაციური პარალელური პროცესორების ეფექტურობა კარგად განპარალელებადი ამოცანებისათვის ნაჩვენებია ნაშრომებში [17] და [18], სადაც შემოთავაზებულია სახეთა ამოცნობის ორსაფეხურიანი მაღალსაიმედო ალგორითმი და დროის რეალურ მასშტაბში მისი განხორციელების მეთოდика ასოციაცი-

ური პარალელური პროცესებით. ალგორითმის პირველი საფეხური ასრულებს წინასწარ ამოცნობას, ხოლო მეორე საფეხური საბოლოოდ ამოიცნობს პირველი საფეხურიდან მიღებულ რეალიზაციებს. პირველი საფეხურის ალგორითმი ეფუძნება რეალიზაციათა სასწავლო ნაკრებისაგან სუპერპოზიციის პრინციპისა და ლოგიკური ფუნქციების გამოყენებით მიღებულ ე.წ. მინი და მაქსი პორტრეტებს ანუ ეტალონურ აღწერებს, რომლებსაც შედარდება უცნობი რეალიზაციები და გაანალიზდება შესაძლო ცხრა სიტუაცია, რის საფუძველზეც მიიღება ერთ-ერთი გადაწყვეტილება:

უცნობი რეალიზაცია მიეკუთვნება ერთ სახეს;

უცნობი რეალიზაცია მიეკუთვნება რამდენიმე სახეს;

უარის თქმა ამოცნობაზე;

- რეალიზაცია გადაიგზავნება ამოცნობის მეორე საფეხურზე.

მეორე საფეხურის ალგორითმში გამოიყენება ე.წ. ინფორმაციული სახეთა აღწერები, რომლებიც, თავის მხრივ, მიიღება ალბათური ეტალონებისაგან. ამ ეტაპზე გადაწყვეტილების მისაღებად გამოიყენება უცნობი რეალიზაციის, ინფორმაციული ეტალონის და ალბათური ეტალონის ვექტორების სკალარული ნამრავლი. ექსპერიმენტებით დადგინდა უცნობი რეალიზაციების ამოცნობის მაღალი საიმედოობა.

აღსანიშნავია, რომ ასოციაციურ მატრიცებში შესაძლებელია ინფორმაციის სხვადასხვა სახით განლაგება, რაზეც მნიშვნელოვნად არის დამოკიდებული როგორც ამოცანის გადაწყვეტის დრო, ასევე საჭირო აპარატურის მოცულობა. ნაშრომში [17] განხილულია ზემოთ აღწერილი ალგორითმების განხორციელება ასოციაციურ მატრიცაში ინფორმაციის სხვადასხვა მეთოდით განლაგების შემთხვევაში და მოცემულია მათთვის საჭირო რესურსების შეფასებები.

ნაშრომში [18] მოცემულია ასოციაციურ პარალელურ პროცესორში მაქსიმალური სიტყვის ძიების ალგორითმი და მისი რეალიზების მეთოდიკა, რაც აუცილებე-

ლი პროცედურაა ზემოთ აღწერილი სახეთა ამოცნობის ორსაფეხურიანი ალგორითმის განხორციელებისთვის, გამოკლეულია ინდიკაციის ვექტორის მდგომარეობის განსაზღვრის საკითხი და მოცემულია ერთადერთი „ერთიანის“ იდენტიფიკაციის ალგორითმი. ყოველივე ეს იძლევა აგრეთვე იმის შესაძლებლობას, რომ შესრულდეს ასოციაციური მატრიციდან ინფორმაციის დალაგებული ამოკრეფის პროცედურები.

როგორც იყო აღნიშნული, მაღალი პარალელუზმის გამოთვლითი სისტემები ძალზე მრავალფეროვანია და მათი დეტალური განხილვა სცილდება ჩვენს მიზნებს.

ფონ ნეიმანის პარადიგმის მნიშვნელოვანი განვითარების მიუხედავად, რამაც ელემენტური ბაზის განვითარებასთან ერთად მიგვიყვანა თანამედროვე ზემადალი მწარმოებლობის სუპერკომპიუტერების შექმნამდე, მთლიანად მაინც ვერ გადაწყდა ისეთი პრობლემები, როგორცაა კომპიუტერის პროგრამული მართვის სირთულეები, მაღალმწარმოებლურ სისტემებში ინფორმაციის შეყვანა-გამოყვანის სიძნელეები, აგრეთვე, ადამიანისა და კომპიუტერის ეფექტური ურთიერთქმედების საკითხები. ამასთან, გამოიკვეთა, რომ სისტემის მწარმოებლობის და პროცესორების რაოდენობის ზრდას შორის არ არის წრფივი დამოკიდებულება, არამედ სისტემის მწარმოებლობა მიახლოებით იზრდება, როგორც პროცესორების რაოდენობის ლოგარითმი. ეს ნიშნავს, რომ რაც უფრო მეტად იზრდება პროცესორების რაოდენობა, მით ნაკლებად იზრდება სისტემის მწარმოებლობა, მით მეტია გამოთვლების ღირებულება, ხოლო ინფორმაციისა და ბრძანებების გაცვლის ეფექტური ორგანიზაცია მრავალპროცესორიან სისტემებში თავად ხდება რთული მეცნიერული პრობლემა.

მეოცე საუკუნის ბოლოს გამოთვლითი ტექნიკის მრავალი სპეციალისტისათვის ნათელი გახდა, რომ მრავალპროცესორიანი სისტემების განვითარების ამ გზას მრავალი რთული გამოთვლითი ამოცანისათვის აღარ აქვს პერსპექტივა. ცალკე პრობლემაა ტრადიციულ

კომპიუტერებზე ისეთი ამოცანების გადაწყვეტა, რომელთა ალგორითმების შექმნა გაძნელებულია ან ზოგჯერ შეუძლებელიც.

გამოსავალი ასეთი მდგომარეობიდან არის ის, რომ რთული ამოცანები, რომელთა გადასაწყვეტად იქმნება პარალელური პროცესორები და მრავალპროცესორიანი სისტემები, უნდა გადაწყდეს ინფორმაციის დამუშავების სხვა მეთოდებითა და ტექნოლოგიებით. ასეთი ალტერნატიული ტექნოლოგია არის ინფორმაციის დამუშავება ხელოვნური ნეირონული ქსელებით. სწორედ ამიტომ ბოლო წლებში მრავალი მიმართულებით მიმდინარეობს ინტენსიური კვლევები ახალი კომპიუტერული პარადიგმების შესაქმნელად.

ძნელად ფორმალიზებადი ამოცანების გადაწყვეტისა და კომპიუტერების მწარმოებლობის შემდგომი ამაღლების გზებს სულ უფრო მეტად უკავშირებენ ხელოვნურ ნეირონულ ქსელებს, კერძოდ, კონექციონისტურ ნეირო-კომპიუტერულ ტექნოლოგიებს.

1.4. სახეთა ამოცნობის თეორიის ელემენტები

ამოცნობის პროცესები ბუნებრივი და ხელოვნური ინტელექტუალური სისტემების ფუნქციონირების ერთ-ერთი უმნიშვნელოვანესი შემადგენელი ნაწილია. ამდენად, მნიშვნელოვანია ნეიროინფორმატიკის იდეოლოგიის მესამე წყაროს – სახეთა ამოცნობის თეორიის იმ განსაზღვრებების, დებულებებისა და მეთოდების განხილვა, რომლებიც ხშირად გამოიყენება ხელოვნური ნეირონული ქსელების თეორიასა და პრაქტიკაში.

სახეთა ამოცნობის თეორიის აქ წარმოდგენილი განსაზღვრებები, დებულებები და მეთოდები ძირითადად აღებულია [20,24,40,41] ნაშრომებიდან.

1.4.1. ამოცნობის პროცესების ძირითადი ცნებები და ტერმინები

სახე ეწოდება ყველა იმ ობიექტის სიმრავლეს, რომელიც აკმაყოფილებს ერთიანობისა და მთლიანობის თვისებებს. სახის სინონიმია ტერმინი კლასი. ამოსაცნობ სახეთა სიმრავლე უნდა შედგებოდეს ერთზე მეტი ელემენტისაგან – სხვა შემთხვევაში ამოცნობის პრობლემა არ იარსებებს. სახეთა სიმრავლის სინონიმია ტერმინი – სახეთა ანსამბლი.

სახეთა სიმრავლე შეიძლება აერთიანებდეს ორგანოზომილებიან ობიექტებს, რომელთაც ეწოდება გამოსახულება.

შტრიხული ეწოდება ისეთ გამოსახულებას, რომელიც შედგება წირების ან მონაკვეთებისაგან.

ლაქისებრი ეწოდება ისეთ გამოსახულებას, რომელიც ნებისმიერი ფორმის მთლიანი ლაქებისაგან შედგება.

სცენა ეწოდება სივრცის იმ ნაწილს, სადაც მოთავსებულია სამგანზომილებიანი ამოსაცნობი ობიექტები.

ამოცნობა ეწოდება პროცესს, რომლის შედეგად უცნობი ობიექტი მიეკუთვნება ამა თუ იმ სახეს. ამოცნობის სინონიმებია გარჩევა, კლასიფიკაცია.

ამოცნობის პროცესი აუცილებლად შეიცავს შედარების პროცედურას, რისთვისაც საჭიროა, სულ ცოტა, ორი ობიექტის არსებობა. ერთი არის უცნობი ობიექტი, ხოლო მეორე, შესადარებელი ობიექტი ანუ ეტალონი, რომლის სინონიმებია: პროტოტიპი, ნიმუში, სახის აღწერა, საყრდენი ობიექტი, ტიპური წარმომადგენელი.

თითოეულ სახეს შეიძლება ჰქონდეს ერთი ან მრავალი ეტალონი. უცნობი ობიექტის ეტალონებთან შედარების შედეგები შეადგენს ალტერნატივების სიმრავლეს, საიდანაც ამოცნობის მომდევნო ეტაპზე ხდება უცნობი ობიექტის მიკუთვნება ამა თუ იმ სახისადმი. ამ პროცედურას უწოდებენ საკუთრივ გადაწყვეტილებათა მიღების პროცესს.

სახეს აქვს თვისებათა გარკვეული სიმრავლე, რომლითაც ხდება ამ სახის წარმოდგენა და აღწერა, რასაც სახეთა ამოცნობაში ნიშნები ეწოდება.

ამოცნობის პროცესებში ნიშნების გამოსაყენებლად აუცილებელია მათი ანალიზი. თუ ნიშნები ზომვადია, ისინი უნდა გაიზომოს, ხოლო სხვა შემთხვევებში უნდა დაფიქსირდეს მათი მიმართებები მოცემული ანსამბლის სახეებისათვის.

ნიშნები შეიძლება იყოს რაოდენობრივი ან თვისებრივი. სახის ზომვად თვისებებს ეწოდება რაოდენობრივი ნიშნები, ხოლო თუ ნიშნების წარმოდგენა შეუძლებელია რიცხვების საშუალებით, მაშინ მათ ეწოდება თვისებრივი ნიშნები. სახეთა ანსამბლის ნიშანთა სიმრავლე ეწოდება იმ ნიშანთა ერთობლიობას, რომელსაც ვიყენებთ მოცემულ სახეთა სიმრავლის აღწერისათვის.

ნიშანთა საწყისი სიმრავლე (პირველადი) ეწოდება საწყის ეტაპზე ევრისტიკულად არჩეულ ნიშანთა ერთობლიობას.

ნიშანთა სიერცე ეწოდება რაოდენობრივ ნიშანთა დალაგებულ სიმრავლეს, სადაც განსაზღვრულია რაიმე მეტრიკული ზომა.

რეცეპტორული ველი ეწოდება ტექნიკურად განხორციელებულ ან გრაფიკულად მოცემულ ნიშანთა სიერცის ანალოგს, მისი სინონიმია რასტრი.

პიქსელი ეწოდება რასტრში შემავალი ერთი ნიშნის ანალოგს. ნიშნების ანალიზით მიღებულ შედეგთა ერთობლიობას ეწოდება რეალიზაცია. იგი წარმოადგენს ვექტორს, თუ ნიშნები ზომვადი და დალაგებულია, მაშინ ვექტორის განზომილება ნიშანთა რაოდენობის ტოლია. რეალიზაციის სინონიმებია: დაკვირვებათა შედეგები, გაზომვათა შედეგები, სახის გამოსახულება.

უცნობი რეალიზაციები ეწოდება იმ რეალიზაციებს, რომელთათვისაც არ არის გარკვეული, თუ რომელ სახეს მიეკუთვნებიან ისინი.

რეალიზაციათა ის სიმრავლე, რომლებისთვისაც ცნობილია, თუ რომელ სახეს მიეკუთვნებიან, შეიძლება გამოყენებულ იქნეს სწავლების პროცესში ანუ ეტალონური აღწერებისა და გადაწყვეტილებათა მიღების პროცესების ფორმირებისათვის ან ხელოვნურ ნეირონულ ქსელებში პარამეტრების (ძირითადად სინაფსური წონების) შეთანაწყობისათვის. ასეთ სიმრავლეს ეწოდება სასწავლო ნაკრები.

რეალიზაციათა სასწავლო ნაკრების ერთ ნაწილს არ იყენებენ სწავლებისათვის, არამედ ინახავენ სწავლების ხარისხის შესამოწმებლად – ტესტირებისათვის. რეალიზაციათა ასეთ სიმრავლეს საგამოცდო ნაკრები ეწოდება, მისი სინონიმია საკონტროლო ნაკრები.

რეალიზაციები შეიძლება იყოს თანაბარგანზომილებიანი ან არათანაბარგანზომილებიანი, რაც დამოკიდებულია სახეთა ანალიზის არჩეულ მეთოდზე. თანაბარგანზომილებიანია ის რეალიზაციები, რომელთაც აქვს ტოლი რაოდენობის ნიშნები. ცდილობენ შეარჩიონ სახეთა ანალიზის ისეთი მეთოდები, რომ სახეთა მოცემული ანსამბლისათვის მიიღონ თანაბარგანზომილებიანი

რეალიზაციები, ვინა-იდან არათანაბარგანსომილებიანი რეალიზაციების დამუშავება გაცილებით უფრო რთულია.

გაზომვებით მიღებული შედეგები ხშირად შეიცავს ხელშეშლებს, ჭარბ ინფორმაციას, არამკაფიო რეალიზაციებს, რომელთა უშალო გამოყენება იწვევს ამოცნობის საიმედოობის დაქვეითებას ამის თავიდან ასაცილებლად საჭირო ხდება რეალიზაციათა პირველადი დამუშავება. ინფორმაციის პირველადი გადამუშავების პროცესს რეალიზაციათა გაუმჯობესების მიზნით, პრეპარირება ეწოდება, რომლის პროცედურებია: ნორმირება, რომელიც გულისხმობს მასშტაბირებას, როდესაც იცვლება რეალიზაციის ერთი ან რამდენიმე განზომილება;

ფილტრაცია, როდესაც განზომილებების ცვლილებების გარეშე ხდება სხვადასხვა ხელშეშლებისა და დამახინჯებების მოცილება;

გამოსახულებათა დავიწროება ან შემოკონტურება, როდესაც განზომილებების ცვლილებების გარეშე იცვლება რეალიზაციათა პარამეტრები;

გამოსახულებათა პრეპარირების დროს მიმართავენ მასშტაბირებას, ფილტრაციას, დასუფთავებას, კონტურირებას, დავიწროებას, დაგლუვებას. სამეტყველო სიგნალების პრეპარირებისას ინფორმაციული სიჭარბის შესამცირებლად მიმართავენ სიგნალის ამპლიტუდის შეზღუდვას, რასაც კლიპირება ეწოდება. სიგნალის ხანგრძლივობის მასშტაბირებაში გამოიყენება არაწრფივი მასშტაბირება. აქ ყველაზე რთულია სამეტყველო სიგნალების სეგმენტაცია.

კლასტერი ეწოდება გარკვეული წესით შედგენილ რეალიზაციათა ერთობლიობას. კლასტერის სინონიმია ტერმინი — ჯგუფი.

კლასტერირება ეწოდება კლასტერის შედგენის (აგების) პროცესს, კლასტერირების სინონიმია ტერმინი — დაჯგუფება.

სახეთა ამოცნობის ერთ-ერთი უმნიშვნელოვანესი პროცესია ნიშანთა შეფასება და არჩევა, სადაც საწყის ნიშანთა სიმრავლისაგან განსხვავებით უნდა დაეტოვოთ მხოლოდ მნიშვნელოვანი და ვარგისი (ამოცნობის შედეგის თვალსაზრისით) ნიშნები. ნიშანთა შეფასების შემდეგ შეიძლება აღმოჩნდეს, რომ ნიშანთა მოცემული სიმრავლე ვერ უზრუნველყოფს საიმედო ამოცნობას, ამიტომ საჭირო ხდება ნიშანთა რაოდენობის გაზრდა ან შემცირება. ძალიან ხშირად ერთდროულად წყდება ეტალონური აღწერებისა და გადაწყვეტილებათა მიღების წესის ფორმირების საკითხები, ვინაიდან მათი შინაარსი და ფორმა მნიშვნელოვნად განსაზღვრავენ ურთიერთს. საკონტროლო ნაკრების ან უცნობი რეალიზაციების შედარების პროცედურა ეტალონებთან გეპაძლევს იმ ალტერნატივათა სიმრავლეს, საიდანაც აირჩევა ერთი, ანუ მოხდება გადაწყვეტილების მიღება.

1.4.2. სწავლების პროცესები ამომცნობ სისტემებში

ამომცნობ სისტემებში სწავლების პროცესები გამოიყენება სახეთა აღწერისა და გადაწყვეტილებათა მიმღები წესების ფორმირების დროს. სწავლების სამი ფორმა არსებობს: მასწავლებლით, უმასწავლებლოდ და შერეული. სწავლების პროცესები უმასწავლებლოდ (თვითსწავლება) გაცილებით რთულია, ამიტომ ტექნიკურ სისტემებში, როგორც წესი, ამ მეთოდს ერიდებიან.

სწავლების პროცესის პირველ ეტაპზე ხდება ინფორმაციის მოპოვება და დამახსოვრება, ხოლო მეორე ეტაპზე – მისი გადამუშავება.

სწავლების ისეთ ალგორითმს, სადაც გადამუშავებული ინფორმაციის ზრდასთან ერთად სწავლების მიზანი სულ უფრო უკეთ ხორციელდება და მიისწრაფვის გარკვეული ცნობილი ზღვარისაკენ, ეწოდება კრებადი ალგორითმი (ან პროცესი). სწავლების

დროს ხშირად გამოიყენება იტერაციული პროცედურები შესაბამისი მიზნობრივი ფუნქციებით. აგებული იტერაციული პროცედურების კრებადობის ხარისხის კვლევისათვის იყენებენ სხვადასხვა მათემატიკურ აპარატს, კერძოდ, ამ მიზნისათვის გამოიყენება ოპტიმიზაციის მეთოდების კარგად დამუშავებული აპარატი სწავლების როგორც დეტერმინირებული, ასევე სტოქასტიკური პროცესების კვლევისათვის. აქ იტერაციის ყოველ ბიჯზე ხდება მიზნობრივი ფუნქციის გამოთვლა და შედარება წინა ბიჯზე მიღებულ შედეგთან, რის მიხედვითაც ხორციელდება შემდეგ ბიჯზე შესასრულებელი მოქმედებები.

განასხვავებენ სწავლების დეტერმინირებულ (არაპარამეტრული) და სტოქასტიკურ ანუ სტატისტიკურ (პარამეტრული) მეთოდებს.

სწავლების დეტერმინირებულ პროცესებში გამოთვლებისას უშუალოდ გამოიყენება სასწავლო ნაკრების რეალიზაციები.

სწავლების სტოქასტიკური პროცესებისათვის აპრიორულად უნდა იყოს ცნობილი რეალიზაციათა განაწილების კანონები. თუ სახეთა ნიშნები ერთმანეთზე არის დამოკიდებული, მაშინ საჭიროა ერთობლივი სტატისტიკური განაწილების კანონების დადგენა, რაც სასწავლო ნაკრების რეალიზაციათა ძალზე დიდ რაოდენობას მოითხოვს, ამიტომ მრავალგანზომილებიანი სახეებისათვის ამ მეთოდების გამოყენება შეზღუდულია.

არასაკმარისი ინფორმაციის შემთხვევაში სწავლების პროცესებს განიხილავს ადაპტური სისტემების თეორია, სადაც შესწავლილია როგორც პარამეტრული, ასევე არაპარამეტრული სწავლების პროცესები. ადაპტურ სისტემებში ყოველი ახალი ინფორმაციის გამოჩენა წარმოადგენს იტერაციის ერთ ბიჯს, რომელზეც ხდება მიზნობრივი ფუნქციის ან მისი გრადიენტის მნიშვნელობის გამოთვლა და ამის საფუძველზე ფორმირდება ამოცნობისათვის საჭირო ერთი ან რამდენიმე პარამეტრი. სწავლება ჩაითვლება

დამთავრებულად, თუ გამოსათვლელი პარამეტრების მნიშვნელობების ცვლილება არ სცილდება დასაშვებ ზოლს.

1.4.3. გადაწყვეტილებათა მიღება ამომცნობ სისტემებში

გადაწყვეტილებათა მიღება ფაქტობრივად წარმოადგენს უშუალოდ ამომცნობის პროცესს. მის პირველ ეტაპზე შედარების პროცედურის შედეგად ვიღებთ ასარჩევ ალტერნატივათა სიმრავლეს, ხოლო მეორე ეტაპზე ამ სიმრავლიდან ვირჩევთ ერთ ალტერნატივას, რომლის მიხედვითაც ხდება გადაწყვეტილების მიღება (დასკენა) გადაწყვეტილების მიღების იმ წესის მიხედვით, რომელიც წინასწარ არის შემუშავებული.

უცნობი რეალიზაცია X და ეტალონური აღწერები $\{E\}$ წარმოადგენენ მსგავსების ზომის ფუნქციის არგუმენტებს $\rho(X, \{E\})$. ეტალონთა $\{E\}$ სიმრავლის ყოველი ელემენტისათვის $\rho(\bullet)$ ფუნქციის გამოთვლის შედეგად ვიღებთ ალტერნატივათა $\{a\}$ სიმრავლეს, $\text{Card}\{E\} = \text{Card}\{a\}$. ამის შემდეგ, გადაწყვეტილების მიღების წესის მიხედვით განისაზღვრება ალტერნატივათა სიმრავლის ერთ-ერთი პარამეტრი:

ა) $\min\{a\}$ ან ბ) $\max\{a\}$.

არჩეული ალტერნატივის მიხედვით განისაზღვრება გადაწყვეტილების მიღების შესაძლებლობა შემდეგი ნუსხიდან:

ა) საბოლოო და ცალსახა გადაწყვეტილების მიღება უცნობი რეალიზაციის მიკუთვნების შესახებ რომელიმე სახისათვის;

- ბ) შუალედური გადაწყვეტილების მიღება, ანუ უცნობი რეალიზაცია სავარაუდოდ მიეკუთვნება რამდენიმე სახეს;
- გ) საბოლოო და ცალსახა გადაწყვეტილების მიღება, რომ რეალიზაცია არ მიეკუთვნება სახეთა მოცემულ სიმრავლეს.
- დ) უარის თქმა ამოცნობაზე.
ყველა ამ სიტუაციის მოდელირება შესაძლებელია მსგავსების ზომის ზღურბლური მნიშვნელობის შერჩევით, რაც ხდება ექსპერიმენტების საფუძველზე ან ევრის-ტიკულად.

1.4.4. გადაწყვეტილებათა მიღების მეთოდების კლასიფიკაცია

გადაწყვეტილებათა მიღების მეთოდების დიდი რაოდენობა არსებობს. მათი კლასიფიკაცია შეიძლება სხვადასხვა ნიშან-თვისების მიხედვით, ამიტომ პრაქტიკულად შეუძლებელია ყველა ამ ფუნქციის ერთ ნუსხაში მოქცევა. ჩამოვაყალიბოთ ძირითადი საკლასიფიკაციო ნიშნები და მათი ზოგადი შეფასებები.

მიზეზ-შედეგობრივი კავშირების მიხედვით გადაწყვეტილებათა მიღების მეთოდების ორი კლასი არსებობს:

ალბათური (პარამეტრული) მეთოდები;

დეტერმინირებული (არაპარამეტრული) მეთოდები.

ალბათური მეთოდების გამოყენება აუცილებლად მოითხოვს ნიშანთა და რეალიზაციათა განაწილების კანონების აპრიორულად დადგენას ყოველი სახისათვის, რისთვისაც აუცილებელია სასწავლო ნაკრების დიდი სიმძლავრე. ამიტომ ალბათური მეთოდები ძირითადად გამოიყენება სახეებისა და ნიშნების შეზღუდული რაოდენობის შემთხვევაში.

ალბათური გადაწყვეტილების მიმღები წესებიდან ყველაზე ცნობილია ბაიესის კრიტერიუმი, ჭეშმარიტ მსგავსებათა თანაფარდობის მეთოდი, სტატისტიკური ჰიპოთეზების შემოწმება, ამოცნობის რისკის მინიმიზაციის მეთოდები და სხვ.

დეტერმინირებული მეთოდები კიდევ უფრო მრავალფეროვანია, მათ მიეკუთვნება: გადაწყვეტილებათა მიღების ლოგიკური ფუნქციები, გადაწყვეტილების მიღების მინი და მაქსი პორტრეტების მეთოდი, წრფივი და არა-წრფივი ფუნქციები (რომელთა ნაწილიც მეტრიკული თვისებების მქონეა), პოტენციალური ფუნქციების მეთოდი, კლასტერიზაციის მეთოდები, დისკრიმინანტული ფუნქციები და სხვ.

ფუნქციონირების მეთოდის მიხედვით გვაქვს გადაწყვეტილებათა მიღების შემდეგი წესები:

- დისკრიმინანტული (განმამხოლობელი) ფუნქციები;
- პოტენციალური ფუნქციების მეთოდი;
- „ხმის მიცემის“ (შეფასებათა გამოთვლის) მეთოდი;
- ლოგიკური მეთოდები.

დისკრიმინანტული ფუნქციების შემთხვევაში ვეძებთ კლასტერების განმამხოლობელ საზღვრებს წირების ან სიბრტყეების განტოლებების სახით. მრავალგანზომილებიანი ნიშანთა სივრცეებისათვის ამ მეთოდის გამოყენება პრობლემატურია.

პოტენციალური ფუნქციების გამოყენებით ხდება ნიშანთა სივრცისაგან განსხვავებული სივრცის აგება, სადაც შემუშავებული თეორიის მიხედვით უზრუნველყოფილი უნდა იყოს მოცემული სახეების განმხოლობეულობა და შესაბამისად, საიმედო ამოცნობა. მეთოდს ახასიათებს არაცალსახა გადაწყვეტილების მიღების მაღალი ალბათობა.

ხმის მიცემის მეთოდი გულისხმობს უცნობი რეალიზაციისა და ეტალონის ფრაგმენტების ცალ-ცალკე შედარებას. ყოველი ფრაგმენტისათვის ცალკე გადაწყვეტილების მიღების თითოეული სახისადმი მიკუთვნების

შესახებ საბოლოო გადაწყვეტილება მიიღება ხმების დათვლით.

გადაწყვეტილებათა მიღების ლოგიკური მეთოდებისათვის დამახასიათებელია რეალიზაციებისა და ნიშნების განხილვა, როგორც ლოგიკური ცვლადებისა, ხოლო სახეთა აღწერები და გადაწყვეტილების მიღების პროცესები წარმოდგენილია ბულის ალგებრის ფუნქციებით. ეს მეთოდები ფართოდაა გაერცელებული გადაწყვეტილების მიმღები წესის ალგორითმული განხორციელების მიხედვით, მეთოდები სამ კლასად ჯგუფდება:

1. გადაწყვეტილების მიღების ერთი მთლიანი პროცესი;
2. გადაწყვეტილების მიღების მიმდევრობითი პროცესები;
3. გადაწყვეტილების მიღება ერთზე მეტი გადაწყვეტილების მიმღები ფუნქციით.

გადაწყვეტილების მიღების ერთი მთლიანი პროცესი ყველაზე გაერცელებულია ამომცნობ სისტემებში - უცნობი რეალიზაცია და ეტალონი განიხილება როგორც მთლიანი, არადაყოფადი ობიექტები. შედარებისათვის საჭირო გამოთვლების თვალსაზრისით, აქ გადაწყვეტილება მიიღება მხოლოდ შედარების პროცედურების დამთავრების შემდეგ.

გადაწყვეტილების მიღების მიმდევრობით პროცესებში რეალიზაციებისა და ეტალონების შედარებისას გამოთვლების სტრუქტურული ერთეულია მათი ფრაგმენტები, რომლებიც კერძო შემთხვევაში შეიძლება ცალკეული ნიშნებიც იყოს. აქ საბოლოო გადაწყვეტილების მიღება შეიძლება შედარების პროცესის ნებისმიერ ეტაპზე.

გადაწყვეტილების მიღება ერთზე მეტი გადაწყვეტილების მიმღები ფუნქციით საჭირო ხდება, როდესაც ერთი გადაწყვეტილების მიმღები წესით არ ხერხდება ამოცნობის საჭირო საიმედოობის მიღწევა, მაშინ გამოიყენება სხვა ფუნქციებიც. ამასთან, თითოეული მათგანი დამოუკიდებლად იღებს გადაწყვეტილებას. საბოლოო გადაწყვეტილების ფორმირება „ხმის მიცემის“ წესით ხდება, ამასთან თითოეულ ხმას შეიძლება მიეკუთვნოს გარკვეული წონითი კოეფიციენტი.

თავი II. ხელოვნური ნეირონი და ნეირონული ქსელები

ხელოვნური ნეირონული ქსელები წარმოადგენს პარალელურ განაწილებულ გამოთვლით სისტემებს. ისინი დაკავშირებულნი არიან ნეიროფიზიოლოგიურ წარმოდგენებთან, ვინაიდან მათი შემადგენელი ელემენტების ხელოვნური ნეირონების ფუნქციური შესაძლებლობები ანალოგიურია ბიოლოგიური ნეირონების ზოგიერთი ელემენტარული ფუნქციისა, ხოლო სტრუქტურული ორგანიზაცია ნაწილობრივ ემსგავსება ტვინის ანატომიურ აგებულებას.

მიუხედავად ასეთი ზედაპირული ანალოგიისა, ხელოვნურ ნეირონულ ქსელებს აქვს ადამიანის ტვინის მსგავსი ზოგიერთი თვისება, კერძოდ, ისინი მაგალითებზე სწავლობენ, ახალი მაგალითების მიმართ ანზოგადებენ წარსულ გამოცდილებას, შეუძლიათ შემოსული ჭარბი ინფორმაციიდან ამოკრიბონ არსებითი თვისებრივი ნიშნები. ხელოვნურ ნეირონულ ქსელებს შეუძლია შეცვალონ თავისი რეაქცია გარემოს ცვლილებების ადევნატურად. სწორედ ეს ფაქტორი ღიდად აძლიერებს მათდამი ინტერესს.

დღეისათვის მეცნიერების მიერ შემოთავაზებული და შესწავლილია ნეირონებისა და ნეირონული ქსელების მრავალი განსხვავებული მოდელი, რომელთაგანაც ყველაზე მნიშვნელოვანი მოდელები განხილულია მომდევნო თავებში.

2.1. ისტორიული ცნობა ხელოვნური ნეირონული ქსელების შესახებ

ტერმინი - ხელოვნური ნეირონული ქსელი, მეოცე საუკუნის მეორე ნახევარში დამკვიდრდა. პირველი მნიშვნელოვანი შედეგები ნეიროინფორმატიკაში,

როგორც ბოლო წლებში ამ დარგს უწოდებენ, დაკავშირებულია მაკ-კალოკის, პიტსის, პების, როზენბლატის, მინსკის, პეიპერტის, პოპფილდის, უიდროუს და სხვების სახელებთან.

1943 წელს უ. მაკ-კალოკმა და უ. პიტსმა წამოაყენეს ბუნებრივი ნეირონის მოდელის ე.წ. ფორმალური ნეირონის, ანუ ხელოვნური ნეირონის იდეა, რომელიც ეყრდნობოდა თავის ტვინის ფუნქციონირების ძირითად დებულებებს.

1949 წელს დ. პებმა გამოთქვა იდეები თავის ტვინში ნეირონების კავშირებისა და ურთიერთქმედებების შესახებ (უჯრედული ანსამბლები, სინაფსური პლასტიკურობა). მანვე წამოაყენა ნეირონული ქსელების სწავლეუბის წესების შესახებ მოსაზრებები.

1957 წელს ფ. როზენბლატმა შეიმუშავა ე.წ. პერსეპტრონების ორგანიზაციისა და ფუნქციონირების პრინციპები. მანვე წამოაყენა მსოფლიოში პირველი ნეიროკომპიუტერის (Mark) ტექნიკური რეალიზაციის გეგმა.

1959 წელს დ. პიუბელმა და ტ. ვიზელმა აჩვენეს ინფორმაციის შენახვისა და დამუშავების პარალელური და განაწილებული ხასიათი ბიოლოგიურ ნეირონულ ქსელებში.

1969 წელს გამოქვეყნდა მ. მინსკის და ს. პეიპერტის წიგნი - „პერსეპტრონები“, სადაც დამტკიცებულია პერსეპტრონების (ერთშრიანი ნეიროქსელების) შესაძლებლობების პრინციპული შეზღუდვები. ამ ნაშრომის გამოქვეყნების შემდეგ, მეოცე საუკუნის 80-იან წლებამდე ინტერესი ხელოვნური ნეირონული ქსელებისადმი მკვეთრად შემცირდა, თუმცა ცალკეული მკვლევარები მაინც განაგრძობდნენ მუშაობას ამ მიმართულებით (კოპონენი, გროსბერგი, ანდერსონი). მიღწევებმა მიკროელექტრონიკაში, კომპიუტერულ ტექნოლოგიებში, ნეიროფიზიოლოგიაში ხელი შეუწყო ინტერესის აღორძინებას ხელოვნური ნეირო-ქსელებისადმი. განსაკუთრებით მნიშვნელოვანი იყო მრავალშრიანი

ნეიროქსელების სწავლების მეთოდების აღმოჩენა, რამაც განაპირობა მ. მინსკის მიერ დაყენებული რიგი პრობლემების გადაწყვეტა სტანდარტული პროცედურებით. ამჟამად, ნეირონული ქსელების თეორიისა და პრაქტიკის განვითარება აღმავალი ხაზით მიმდინარეობს:

1982 წელს ჯ. პოპფილდმა შეიმუშავა ნეირონული ქსელების სახეობები, რომელთაც შეუძლიათ ასოციაციური მექანიზმების მოდელირება;

1985 წელს გამოჩნდნენ პირველი კომერციული ნეიროკომპიუტერები (მაგალითად, Mark III);

1989 წლიდან პრაქტიკულად მსოფლიოს ყველა მსხვილი ელექტროტექნიკური და კომპიუტერული ფირმა მუშაობს ხელოვნური ნეირონული ქსელების სფეროში;

1991 წელს გაყიდვების მოცულობამ ნეიროკომპიუტერების ბაზარზე შეადგინა 140 მლნ დოლარი;

1996 წელს მსოფლიოს სხვადასხვა ქვეყანაში ჩატარდა ასამდე საერთაშორისო კონფერენცია ნეიროინფორმატიკის დარგში;

1997 წელს გაყიდვის მოცულობამ ნეიროკომპიუტერების ბაზარზე გადააჭარბა 2 მილიარდ დოლარს, ხოლო გაყიდვის ყოველწლიური ზრდა 50 პროცენტს შეადგენს;

2000 წლიდან ჩნდება პუბლიკაციები ნეიროკომპიუტერების ახალი არქიტექტურული და ტექნოლოგიური გადაწყვეტების შესაძლებლობის შესახებ მოლეკულური და ბიომოლეკულური ტექნოლოგიების, აგრეთვე ნანოტექნოლოგიების გამოყენებით.

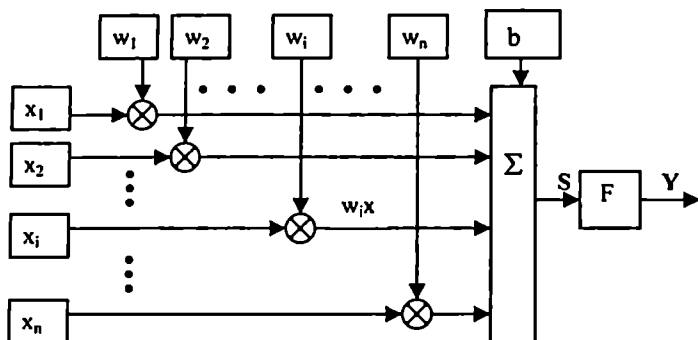
ამ მოკლე მიმოხილვიდანაც ნათლად ჩანს, რომ ნეიროინფორმატიკა დინამიკურად ვითარდება. ნეიროკომპიუტერები უკვე გამოიყენება სახეთა ამოცნობის, ფუნქციათა აპროქსიმაციის, კლასტერიზების, პროგნოზირების, ოპტიმიზაციის, ასოციაციური მექანიზმების, დინამიკური სისტემების მართვისა და მრავალი სხვა რთული და ძნელად ფორმალიზებადი ამოცანების გადასაწყვეტად.

2.2. ხელოვნური ნეირონის სტრუქტურა და თვისებები

ხელოვნური ანუ ფორმალური ნეირონი წარმოადგენს ცალკე აღებული ბიოლოგიური ნეირონის გამორჩეული ინტეგრალური თვისებების რეკონსტრუქციას, თუმცა ამ მოდელში უგულებელყოფილია ბიოლოგიური ნეირონის თვისებების გაცილებით უფრო რთული კომპლექსი. ყველაზე გავრცელებულია მაკ-კალოკისა და პიტსის ფორმალური ნეირონი, რომელიც ხელოვნური ნეირონული ქსელების ძირითადი შემადგენელი ნაწილია. თავის ტვინის ნერვული უჯრედების ანალოგიურად, თითოეული ფორმალური ნეირონი ხასიათდება თავისი მიმდინარე მდგომარეობით და შეიძლება იყოს აგზნებული ან დამუხრუჭებული. ფორმალურ ნეირონს აქვს სინაფსური კავშირების ჯგუფი, რომლითაც იგი უკავშირდება სხვა ნეირონების გამოსასვლელებს. მას აქვს აგრეთვე ერთი აქსონი - გამოსასვლელი, რომლითაც ამ ნეირონის სიგნალი (აგზნების ან დამუხრუჭების) გადაეცემა მომდევნო ნეირონს.

2.1 სურათზე ნაჩვენებია ფორმალური ნეირონის სტრუქტურა. ფორმალური ნეირონი შედგება სამი ტიპის ელემენტისაგან: გამრავლების ელემენტები (სინაფსები, მათი რაოდენობა შემაჯავლი ვექტორის განზომილების ტოლია); ამჯამავე ელემენტი (ზოგჯერ უწოდებენ ადაპტიურ სუმატორს); არაწრფივი გარდამქმნელი ელემენტი (აქტივაციის ელემენტი).

სინაფსები აძლიერებს ან ასუსტებს ნეირონში შემაჯავლ სიგნალს იმისდა მიხედვით, თუ როგორია შესაბამისი სინაფსური კავშირის ძალა ანუ წონითი კოეფიციენტი.



სურ. 2.1. ფორმალური ნეირონის სტრუქტურა.

ფორმალური ნეირონის მიმდინარე მდგომარეობა განისაზღვრება როგორც მისი შესასვლელების შეწონილი ჯამი (წინა ცვლილების გათვალისწინებით)

$$S = \sum_{i=1}^n w_i x_i + b, \quad (2.1)$$

ხოლო ფორმალური ნეირონის გამოსასვლელი (სიგნალი აქსონზე) არის მისი მიმდინარე მდგომარეობის ფუნქცია:

$$y = F(S) = F\left(\sum_{i=1}^n w_i x_i + b\right), \quad (2.2)$$

სადაც w_i არის სინაფსური წონა (weight), $i=1,2,\dots,n$;

n - შემავალი ვექტორის განზომილება;

x_i - შემავალი ვექტორის ელემენტი, $i=1,2,\dots,n$;

b - წანაცვლების მნიშვნელობა (bias);

S - შეჯამების შედეგი (sum) (იყენებენ NET აღნიშვნასაც);

y - ნეირონის გამომავალი სიგნალი (ხშირად იყენებენ OUT აღნიშვნას);

F - არაწრფივი გარდაქმნა (აქტივაციის ფუნქცია).

როგორც ფორმალური ნეირონის მათემატიკური მოდელი (2.1), (2.2) გვიჩვენებს, რომ არაწრფივი გარდამქმნელი ელემენტი ახორციელებს ერთი არგუმენტის – ამჟამავე ელემენტის გამოსასვლელის არაწრფივ ფუნქციას ანუ აქტივაციის ფუნქციას, ხოლო მთლიანად ფორმალური ნეირონი ახორციელებს ვექტორული არგუმენტის სკალარულ ფუნქციას.

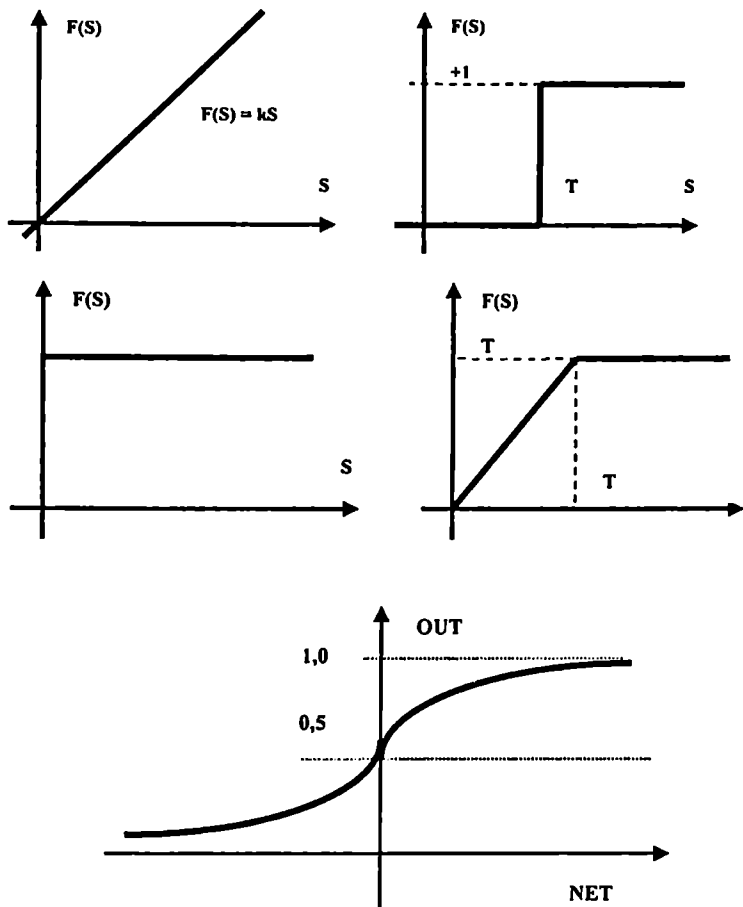
ფორმალური ნეირონის ზოგიერთი აქტივაციის ფუნქცია მოცემულია 2.1 ცხრილში, ხოლო აქტივაციის ფუნქციის რამდენიმე გრაფიკი 2.2 სურათზე.

ყველა ზემოთ ნაჩვენები აქტივაციის ფუნქცია არის ერთპარამეტრიანი, თუმცა ზოგჯერ გამოიყენება მრავალპარამეტრიანი აქტივაციის ფუნქციებიც, მაგალითად,

$$F(S) = a_1 \frac{S}{|S| + a_2} + a_3 \quad (2.3)$$

ცხრილი 2.1. ფორმალური ნეირონის აქტივაციის ფუნქციები

დასახელება	ფორმულა	არე
წრფივი	$F(S) = kS$	$(-\infty, \infty)$
ნახევრად წრფივი	$F(s) = \begin{cases} ks, & S > 0 \\ 0, & S \leq 0 \end{cases}$	$(0, \infty)$
ლოგ-სიგმოიდური	$F(S) = \frac{1}{1 + e^{-aS}}$	$(0, 1)$
ჰიპერბოლური ტანგენსი	$F(S) = \frac{e^{aS} - e^{-aS}}{e^{aS} + e^{-aS}}$	$(-1, 1)$
სიგმოიდურ-რაციონალური	$F(S) = \frac{S}{a + S }$	$(-1, 1)$
ექსპონენციალური	$F(S) = e^{-aS}$	$(0, \infty)$
სინუსოიდური	$F(S) = \sin(S)$	$(-1, 1)$
ბიჯური (წრფივი გაჯერებით)	$F(S) = \begin{cases} -1, & S \leq -1 \\ S, & -1 < S < 1 \\ 1, & S \geq 1 \end{cases}$	$(-1, 1)$
ზღურბლური (საფეხურისებრი)	$F(S) = \begin{cases} 0, & S < T \\ 1, & S \geq T \end{cases}$	$(0, 1)$
მოდულის	$F(S) = S $	$(0, \infty)$
სიგნატურული	$F(S) = \begin{cases} 1, & S > 0 \\ -1, & S \leq 0 \end{cases}$	$(-1, 1)$



სურ. 2.2. ფორმალური ნეირონის ზოგიერთი აქტივაციის ფუნქციის გრაფიკი.

ფორმალური ნეირონის არაწრფივი გარდამქმნელი ელემენტი ახდენს ამჯამავე ელემენტიდან გამოსული სიგნალის მოდიფიცირებას, რის შედეგადაც ნეირონის აქსონზე მიიღება ნეიროქსელში შემდგომი პროცე-

დურების ჩასატარებლად უფრო მოხერხებული მოდიფიცირებული სიგნალი.

თუ აქტივაციის F ფუნქცია S შეწონილი ჯამის სიდიდის ცვლილების დიაპაზონს ავიწროებს ისე, რომ S -ის ნებისმიერი მნიშვნელობისას აქსონზე მიღებული სიგნალი y ეკუთვნის გარკვეულ სასრულ ინტერვალს, მაშინ F ფუნქციას უწოდებენ „შემკუმშავ“ ფუნქციას. ყველაზე ხშირად შემკუმშავ ფუნქციად იყენებენ სიგმოიდურ ფუნქციებს, კერძოდ ლოგ-სიგმოიდურ ფუნქციას

$$F(S) = \frac{1}{1 + e^{-\alpha s}}. \quad (2.4)$$

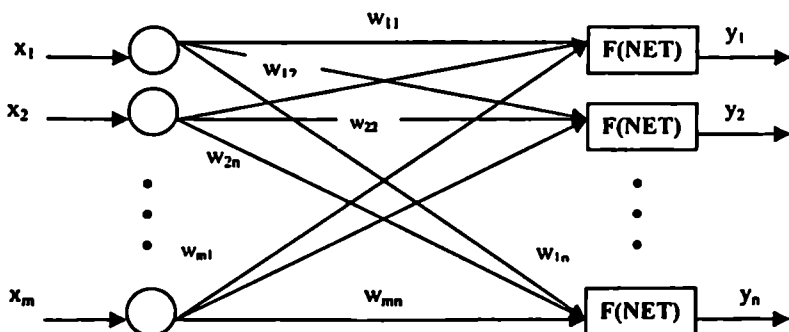
სიგმოიდური აქტივაციის ფუნქციას აქვს რამდენიმე სასარგებლო თვისება, α პარამეტრის შემცირებისას იგი უფრო დამრეცი ხდება, ხოლო როდესაც $\alpha = 0$, იგი გადაგვარდება ჰორიზონტალურ წრფედ 0.5-ის დონეზე. α -ს გაზრდისას სიგმოიდი ემსგავსება საფეხურისებრ (ზღურბლურ) ფუნქციას, ანუ იცვლება (0.1) დიაპაზონში. ამ ფუნქციას აქვს წარმოებულის მარტივი გამოსახულება, რომლის გამოყენებაც აღწერილი იქნება შემდგომში

$$F'(S) = \alpha F(S) [1 - F(S)]$$

გროსბერგმა (1973წ) აღმოაჩინა, რომ სიგმოიდური აქტივაციის ფუნქცია ნეიროქსელებში ხმაურით გაჯერების დილემასაც აგვარებს; სახელდობრ, იგი უზრუნველყოფს გაძლიერების ავტომატურ კონტროლს, ანუ ნეირონი ფუნქციონირებს დიდი გაძლიერებით შემაგალი სიგნალის ფართო დიაპაზონში, შედეგად ძლიერ სიგნალებს ქსელი აღიქვამს გაჯერების გარეშე, ხოლო სუსტი სიგნალები ქსელს გაივლის ზედმეტი შესუსტების გარეშე.

2.3. ერთშრიანი და მრავალშრიანი ხელოვნური ნეირონული ქსელები

ერთ ცალკე აღებულ ხელოვნურ ანუ ფორმალურ ნეირონს შეუძლია ამოცნობის უმარტივესი პროცედურების შესრულება, მაგრამ ნეირონული გამოთვლების ეფექტურობას განაპირობებს მრავალი ნეირონის გაერთიანება ქსელებში. უმარტივესი ნეირონული ქსელი შედგება ნეირონების ჯგუფისაგან, რომლებიც გაერთიანებულია ერთ შრეში (სურ. 2.3), სადაც წრეებით აღნიშნულია ე.წ. ნულოვანი შრე, რომელიც არ ასრულებს გამოთვლით ოპერაციებს, არამედ ანაწილებს შემომავალ სიგნალებს გამომთვლელ ნეირონებს, რომლებიც გამოსახულია კვადრატებით.

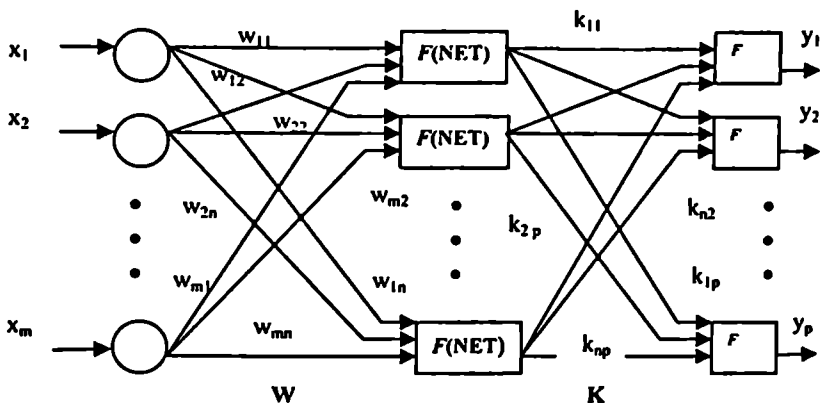


სურ. 2.3. ერთშრიანი ხელოვნური ნეირონული ქსელი.

X შემავალი ვექტორის ელემენტების ($x_1, x_2, \dots, x_n$) ნამრავლი შესაბამის წონით კოეფიციენტებზე (სინაფსურ წონებზე) მიეწოდება თითოეულ ნეირონს, ხოლო შესასვლელების შეწონილ ჯამს ($y_1, y_2, \dots, y_m$) თითოეული ნეირონი გადასცემს ნეირონული ქსელის მომდევნო კომპონენტს. საკმაოდ მოხერხებულია სინაფსური წონების წარმოდგენა W მატრიცის ელემენტებად. მატ-

რიცხვს აქვს m სტრიქონი და n სვეტი, სადაც m არის შესასვლელების რიცხვი, ხოლო n – ნეირონების რიცხვი. მაგალითად, მატრიცის ელემენტი w_{12} არის წონა, რომელიც აკავშირებს პირველ შესასვლელს მეორე ნეირონთან, ხოლო w_{21} არის წონა, რომელიც აკავშირებს მეორე შესასვლელს პირველ ნეირონთან. ამრიგად, გამომავალი N ვექტორის გამოთვლა, რომლის კომპონენტებიცაა ნეირონების გამოსასვლელები, დაიყვანება მატრიცულ გამრავლებაზე $N=XW$, სადაც N და X ვექტორ-სტრიქონებია. ხშირად ნეირონის გამოსასვლელს აღნიშნავენ OUT – ით, მაშინ $OUT = XW$.

აღსანიშნავია, რომ მრავალშრიან ხელოვნურ ნეირონულ ქსელებს ერთშრიანი ნეირონულ ქსელებთან შედარებით გაცილებით მეტი გამოთვლითი შესაძლებლობები აქვს. დამუშავებულია მრავალშრიანი ხელოვნური ნეირონული ქსელების მრავალი სახეობა და შექმნილია მათი სწავლების შესაბამისი ალგორითმები. თითოეულ ამ ნეიროქსელს ახასიათებს თავისი ძლიერი და სუსტი მხარეები. მრავალშრიანი ხელოვნური ნეირონული ქსელები შეიძლება შეიქმნას შრეების კასკადური შეერთებით, მაშინ ერთი შრის გამოსასვლელი იქნება მომდევნო შრის შესასვლელი. ითვლება, რომ სინაფსური წონები დაკავშირებულია მათი მომდევნო შრის ნეირონებთან. ამრიგად, შრე შედგება სინაფსური წონებისა და მათი მომდევნო ნეირონებისაგან, რომლებიც შეწონილ ჯამებს გამოითვლიან. 2.4 სურათზე ნაჩვენებია ორშრიანი ხელოვნური ნეირონული ქსელი.



სურ. 2.4. ორშრიანი ხელოვნური ნეირონული ქსელი.

ერთშრიან ქსელებთან შედარებით, მრავალშრიან ხელოვნურ ნეირონულ ქსელებს გაცილებით მეტი შესაძლებლობები აქვს მხოლოდ და მხოლოდ იმ შემთხვევაში, თუ შრეებს შორის აქტივაციის ფუნქცია იქნება არაწრფივი.

წრფივი აქტივაციის ფუნქციის მქონე ორშრიანი ქსელის გამოსასვლელი გამოითვლება მასში შემავალი ვექტორის გამრავლებით პირველ წონით მატრიცაზე, შემდეგ კი პირველი შრის მიერ გამომუშავებული შედეგის გამრავლებით მეორე წონით მატრიცაზე: $(XW_1)W_2$, იმის გამო, რომ მატრიცების გამრავლებას აქვს ასოციაციურობის თვისება, გვექნება $X(W_1W_2)$, ანუ ნათელია, რომ ორშრიანი ხელოვნური ნეირონული ქსელი წრფივი აქტივაციის ფუნქციით ეკვივალენტურია ისეთი ერთშრიანი წრფივი ქსელისა, რომლის წონითი კოეფიციენტების მატრიცა არის პირველი და მეორე წონითი კოეფიციენტების მატრიცების ნამრავლი. აქედან გამომდინარეობს შედეგი: ნებისმიერი მრავალშრიანი ხელოვნური ნეირონული ქსელი, რომლის ნეირონებსაც

აქვს წრფივი აქტივაციის ფუნქციები, შეიძლება შეიცვალოს მისი ეკვივალენტური ერთშრიანი ნეირონული ქსელით.

ქვემოთ ნაჩვენები იქნება ერთშრიანი წრფივი აქტივაციის ფუნქციის მქონე ნეირონული ქსელის გამოთვლითი შესაძლებლობების პრინციპული შესაძლებლობები, ამიტომაც ნეირონული ქსელის გამოთვლითი შესაძლებლობების გაზრდისათვის აუცილებელი პირობაა ქსელში არაწრფივი აქტივაციის ფუნქციის შეტანა.

2.4. ხელოვნური ნეირონული ქსელების კლასიფიკაცია

ხელოვნური ნეირონული ქსელი წარმოადგენს ფორმალური ნეირონების ერთობლიობას, რომლებიც ერთმანეთსა და გარემოს უკავშირდება ნეირონების პარამეტრებით განსაზღვრული კავშირებით. ხელოვნური ნეირონული ქსელების კლასიფიკაცია შესაძლებელია სხვადასხვა ნიშნის მიხედვით.

ნეირონულ ქსელში შესრულებული ფუნქციის მიხედვით თვით ნეირონები იყოფა შემავალ, საშუალებო (ფარულ) და გამომავალ ნეირონებად. შემავალ ნეირონებს მიეწოდება გარე ზემოქმედების შესაბამისი ვექტორი, რომლის ელემენტებსაც იგი გადასცემს სხვა ნეირონებს ცელილებების გარეშე, ანუ შემავალი ნეირონები ზეგავლენას არ ახდენენ ქსელის გამოთვლით შესაძლებლობებზე, არამედ ემსახურებიან სინაფსური წონების პირველი სიმრავლის განაწილებას პირველი შრის ნეირონებზე, ხოლო საშუალებო და გამომავალი ნეირონები (2.1) და (2.2) გამოსახულებების შესაბამისად, ასრულებენ გამოთვლით პროცედურებს და შედეგებს გადასცემენ სხვა ნეირონებს ან ქსელის გამოსასვლელს. ნეირონული ქსელების დაყოფა შეიძლება ფორმალური ნეირონის სტრუქტურის მიხედვით ჰომოგენურ და ჰეტეროგენურ ქსელებად. ჰომოგენური ქსელები (ერთგვაროვანი ქსელები) შედგება ერთი ტიპის

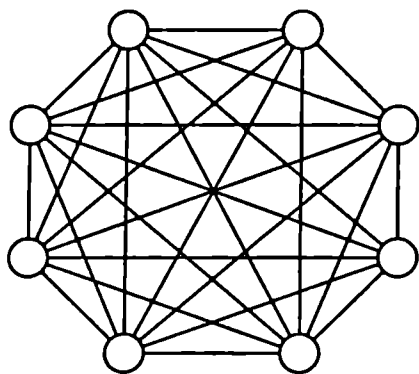
ნეირონებისაგან, რომელთაც ერთნაირი აქტივაციის ფუნქცია აქვს, ხოლო პეტეროგენურ ქსელებში შემავალ ნეირონებს აქვს სხვადასხვა აქტივაციის ფუნქციები.

ფორმალური ნეირონის აქტივაციის ფუნქცია განსაზღვრავს არის ნეიროქსელი დისკრეტული, (ბინარული) თუ უწყვეტი (ანალოგური). დისკრეტული ქსელები ოპერირებენ ბინარული სიდიდეებით, ანუ თითოეული ნეირონის გამოსასვლელმა შეიძლება მიიღოს ლოგიკური ნულის (დამმუხრუტეებული მდგომარეობა) ან ლოგიკური ერთის (აგზნებული მდგომარეობა) მნიშვნელობა. უწყვეტი ქსელები კი ოპერირებენ ნამდვილი რიცხვებით, მათი აქტივაციის ფუნქციები წარმოადგენს სხვადასხვა სახის სიგმოიდურ ფუნქციებს.

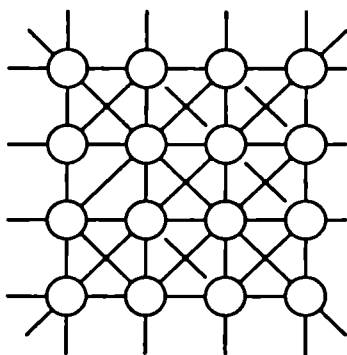
ტოპოლოგიის თვალსაზრისით შეიძლება ნეიროქსელების ოთხი სახეობის გამოყოფა:

- სრულკავშირებიანი ნეირონული ქსელები;
- მრავალშრიანი ნეირონული ქსელები;
- ლოკალურკავშირებიანი ნეირონული ქსელები;
- არასტრუქტურირებული ნეირონული ქსელები.

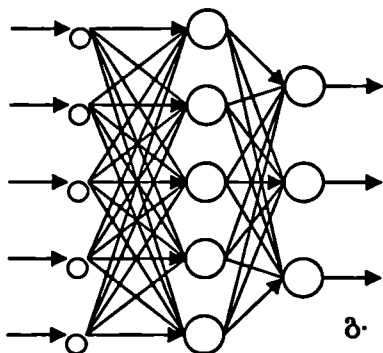
ნეირონული ქსელების ტოპოლოგიის სახეობები მოცემულია 2.5 სურათზე. სრულკავშირებიან ნეირონულ ქსელებში (სურ. 2.5, ა) თითოეული ნეირონი თავის გამომავალ სიგნალს გადასცემს ყველა ნეირონს, მათ შორის საკუთარ თავსაც, ყველა შემავალი სიგნალი მიეწოდება ყველა ნეირონს. ქსელის გამომავალ სიგნალებად შეიძლება გამოყენებული იყოს ქსელის ფუნქციონირების რამდენიმე ტაქტის შემდეგ ყველა ან ზოგიერთი ნეირონის გამომავალი სიგნალი. მრავალშრიან ნეირონულ ქსელებში (სურ. 2.5, გ) ნეირონები გაერთიანებულია შრეებად. შრეს აყალიბებენ ნეირონები, რომელთაც აქვთ საერთო შემავალი სიგნალები, ანუ საერთო რეცეპტორული ველი.



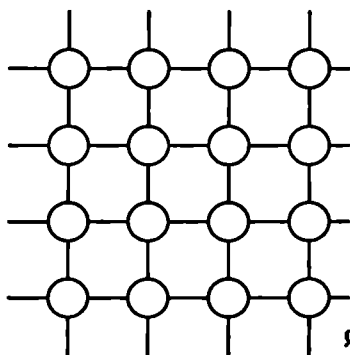
ა.



ბ.



ბ.



დ.

სურ. 2.5. ნეირონული ქსელების ტოპოლოგიის სახეობები.

ნეირონების რაოდენობა შრეში შეიძლება იყოს ნებისმიერი და არ არის დამოკიდებული მათ რაოდენობაზე სხვა შრეებში. ზოგადად, ნეიროქსელში შეიძლება იყოს რამდენიმე შრე, რომლებიც მარცხნიდან მარჯვნივ ინომრება. შემავალ შრეს ნულოვან ნომერს ანიჭებენ, ამით ხაზი ესმება იმას, რომ შემავალი შრე არ აწარმოებს გამოთვლით პროცედურებს. შემავალი და

გამომავალი შრეების გარდა, შეიძლება არსებობდეს რამდენიმე საშუალოდ, ანუ ფარული შრე (როგორც წესი, მათი რაოდენობა ხუთს არ აღემატება). რომელიმე k შრის ნეირონების გამოსასვლელებიდან $(k+1)$ მომდევნო შრის ნეირონების შესასვლელებთან კავშირებს ეწოდება მიმდევრობითი კავშირები.

თავის მხრივ, მრავალშრიანი ნეიროქსელები სამ ჯგუფად იყოფა:

- მონოტონური ნეირონული ქსელები;
- პირდაპირი გავრცელების ნეირონული ქსელები;
- უკუკავშირებიანი ნეირონული ქსელები.

მონოტონურ ქსელებში თითოეული შრე, გარდა უკანასკნელისა, შედგება ორი ბლოკისაგან, ანუ ამგზნები და მამუხრუჭებელი ნეირონების ჯგუფებისაგან. ბლოკებს შორის კავშირებიც ასევე იყოფა ამგზნებ და მამუხრუჭებელ კავშირებად. ერთი შრის შემადგენელ ნეირონებს შორის კავშირებს, ეწოდება ლატერალური კავშირები.

პირდაპირი გავრცელების ნეირონულ ქსელებში, როგორც წესი, თითოეული k შრის ნეირონის გამომავალი სიგნალი მიეწოდება $(k+1)$ შრის თითოეული ნეირონის შესასვლელს; თუმცა შესაძლებელია, რომ k შრე დაუკავშირდეს უშუალოდ რომელიმე $(k+p)$ შრეს.

2.5 სურათზე ნაჩვენებია ასეთი მრავალშრიანი (ორშრიანი) პირდაპირი გავრცელების ნეირონული ქსელის მაგალითი.

უკუკავშირებიან ნეირონულ ქსელებში სიგნალები გადაეცემა როგორც მომდევნო შრეს, ასევე რომელიმე წინამდებარე შრეს. ასეთ ქსელებში შეიძლება ორი ძირითადი სახეობის გამოყოფა:

ა) ციკლური ნეიროქსელები, სადაც უკანასკნელი შრე გადასცემს თავის გამომავალ სიგნალებს პირველ შრეს, ხოლო ყველა შრე თანაბრად არის უფლებამოსილი მიიღოს ქსელში შემოსული სიგნალები და ასევე გასცეს ქსელიდან გამომავალი სიგნალები;

ბ) რეკურენტული სრულკავშირებიანი ნეირონული ქსელები, სადაც თითოეული შრე არის სრულკავშირებიანი; ამასთან, სიგნალების გადაცემა ხდება როგორც შრიდან შრეზე, ასევე საკუთრივ შრის ნეირონებს შორის. ამასთან, თითოეულ შრეში სამუშაო ციკლი სამ ეტაპად იყოფა:

- სიგნალების მიღება წინამდებარე შრისაგან;
- სიგნალების გაცვლა შრის შიგნით;
- გამომავალი სიგნალების გამომუშავება და გადაცემა მომდევნო შრისათვის.

ლოკალურკავშირებიან ნეირონულ ქსელებში ნეირონები უკავშირდება თავის უახლოეს ოთხ, ექვს ან რვა ტოპოლოგიურ მეზობელს. ლოკალურკავშირებიანი ნეიროქსელების მაგალითები ნაჩვენებია 2.5 ბ და 2.5 დ სურათებზე.

არასტრუქტურირებული ნეირონული ქსელების ჯგუფს წარმოადგენს ისეთი ნეიროქსელები, რომელთა მიკუთვნებაც ზემოთ აღწერილი სამი ჯგუფისათვის (სრულკავშირებიანი ქსელები, მრავალშრიანი ქსელები, ლოკალურკავშირებიანი ქსელები) შეუძლებელია.

არსებობს სინქრონული და ასინქრონული ხელოვნური ნეიროქსელები. სინქრონულ ქსელებში დროის ყოველ მომენტში თავის მდგომარეობას იცვლის მხოლოდ ერთი ნეირონი, ხოლო ასინქრონულ ქსელებში თავის მდგომარეობას ერთდროულად იცვლის ნეირონთა ჯგუფი ან მთლიანად შრე.

2.5. ხელოვნური ნეიროქსელების სწავლება

ნეიროქსელების ყველაზე უფრო საინტერესო თვისება არის ის, რომ აქვს სწავლების უნარი. სწავლება იმდენად ემსგავსება ადამიანის ინტელექტუალური განვითარების პროცესს, რომ ნეიროქსელების პირველმკვლევარები დარწმუნებულნი იყვნენ, რომ მიღწეული იყო სწავლების პროცესების პრობლემის

გადაწყვეტა. შემდეგ აღმოჩნდა, რომ ხელოვნური ნეიროქსელების სწავლების შესაძლებლობები შესდურულია და ამ გზაზე გადასაწყვეტია მრავალი რთული ამოცანა, თუმცა, მეორე მხრივ, უკვე მიღებულია დამაიმედებელი შედეგები.

ქსელის სწავლების მიზანია ის, რომ შემავალი სიგნალებისათვის ქსელის გამოსასვლელებზე მიღებულ იქნეს სასურველი ან მასთან მიახლოებული გამომავალი სიგნალები. თითოეული შემავალი და გამომავალი სიგნალების სიმრავლე განიხილება, როგორც ვექტორი, სწავლება ხორციელდება ქსელისათვის შემავალი ვექტორების თანამიმდევრობითი წარდგენით და იმავედროულად, გარკვეული პროცედურის შესაბამისად, ქსელის სინაფსური წონების შეთანაწყობით. სწავლების პროცესში სინაფსური წონები იცვლება ისე, რომ თითოეული შემავალი ვექტორი გამოიმუშავეს შესაბამის გამომავალ ვექტორს.

განასხვავებენ სწავლების ალგორითმებს მასწავლებლით და მასწავლებლის გარეშე. სწავლება მასწავლებლით გულისხმობს, რომ თითოეული შემავალი ვექტორისათვის არსებობს მიზნობრივი ვექტორი. რომელიც წარმოადგენს სასურველ შედეგს. ორივეს ერთად უწოდებენ მასწავლებელ წყვილს. ამრიგად, ქსელს წარედგინება შემავალი ვექტორი, ქსელი გამოითვლის გამომავალ ვექტორს, ხოლო სხვაობა (შეცდომა) უკუკავშირით მიეწოდება ქსელს და სინაფსური წონები იცვლება იმ გარკვეული ალგორითმის თანახმად, რომელიც მიისწრაფვის შეცდომის მინიმიზაციისაკენ. სინაფსური წონების შეთანაწყობა გრძელდება მანამდე, ვიდრე შეცდომა არ გახდება მისაღები მცირე სიდიდე. ასეთი იტერაციული პროცედურები ხორციელდება მასწავლებელ წყვილთა მთელი სიმრავლისათვის.

კოჰონენისა და სხვა მკვლევარების მიერ განვითარებულია სწავლების მეთოდები მასწავლებლის

გარეშე. ეს მეთოდები არ საჭიროებს მიზნობრივ ვექტორებს და, მაშასადამე, კსელის გამომავალი ვექტორი აღარ შედარდება იდეალურ პასუხს. სასწავლო ნაკრები შედგება მხოლოდ შემავალი ვექტორებისაგან. სინაფსური წონების შეთანაწყობის ალგორითმი მუშაობს ისე, რომ კსელის გამოსასვლელზე მიღებულ იქნეს შეთანხმებული გამომავალი ვექტორები, ანუ კსელისათვის საკმარისად მსგავსი შემავალი ვექტორების წარდგენის შედეგად კსელმა უნდა გამოიმუშაოს ერთნაირი გამომავალი სიგნალები. ამრიგად, სწავლება მასწავლებლის გარეშე გამოჰყოფს სასწავლო ნაკრების სტატისტიკურ თვისებებს და მსგავს ვექტორებს კლასებად აჯგუფებს.

ნეიროქსელების სწავლების თანამედროვე ალგორითმების უმეტესობა ეფუძნება ჰების კონცეფციებს. ჰებმა წამოაყენა მასწავლებლის გარეშე სწავლების მოდელი, რომელშიც სინაფსური წონა იზრდება, თუ აგზნებულია ორივე ნეირონი – გადამცემი და მიმღები. ამრიგად, კსელის ის სინაფსები და შესაბამისი გზები, რომლებიც ხშირად გამოიყენება, ძლიერდება. ამ მოდელით კარგად აიხსნება მიჩვევისა და სწავლების ფენომენი ბიოლოგიურ ორგანიზმებში.

ხელოვნურ ნეიროქსელში, სადაც გამოიყენება სწავლების ჰების მოდელი, სინაფსური წონების გაძლიერება განისაზღვრება გადამცემი და მიმღები ნეირონების აგზნების დონეების ნამრავლით:

$$w_{ij}(n+1) = w_{ij}(n) + \alpha \text{OUT}_i \text{OUT}_j, \quad (2.6.)$$

სადაც $w_{ij}(n)$ არის შეთანწყობამდე სინაფსური წონის სიდიდე i ნეირონიდან j ნეირონამდე; $w_{ij}(n+1)$ – შეთანაწყობის შემდეგ სინაფსური წონის სიდიდე i ნეირონიდან j ნეირონამდე; α სწავლების სიჩქარის კოეფიციენტი; OUT_i i ნეირონის გამოსასვლელი და j

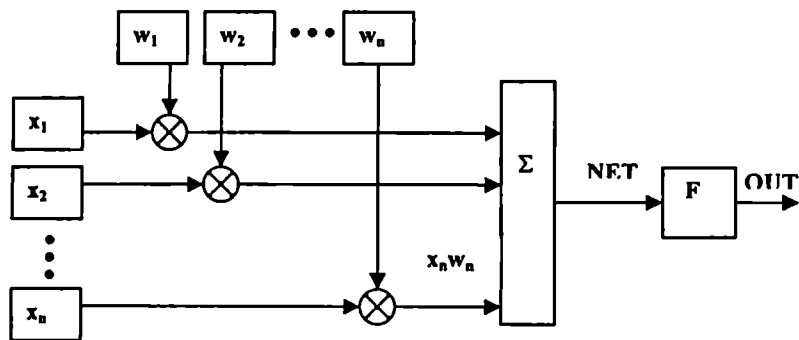
ნეირონის შესასვლელი; OUT_j j ნეირონის გამოსასვლელი.

სწავლების პეზის მოდელი შედარებით მარტივია. ბოლო წლებში დამუშავებულია სწავლების უფრო ეფექტური ალგორითმები, რომლებიც ხასიათდება სწავლების უფრო მაღალი სიჩქარით და სასწავლო სახეების მახასიათებლების შედარებით დიდი დიაპაზონით. დღეისათვის შემუშავებულია მრავალი ასეთი ალგორითმი, მათი ყველაზე ტიპური წარმომადგენლები განხილული იქნება მომდევნო თავებში. ხელოვნური ნეირონული ქსელების ყველაზე უფრო მნიშვნელოვანი თვისება არის ის, რომ აქვს განზოგადების (Generalization) უნარი; რაც ნიშნავს, რომ ისინი იძლევიან სწორ პასუხთან მიახლოებულ შედეგს მაშინაც, როდესაც წარედგინებათ ისეთი რეალიზაციები, რომლებიც არ შედიოდა მასწავლებელ სიმრავლეში. ნეიროქსელს რომ არ ჰქონდეს ასეთი თვისება, მაშინ ის იქნებოდა მხოლოდ და მხოლოდ დამახსოვრების მექანიზმი.

2.6. პერსეპტრონები

ხელოვნური ნეიროქსელების სისტემატური შესწავლა პირველებმა დაიწყეს მაკ-კალოკმა და პიტსმა (1943წ). 2.6 სურათზე ნაჩვენებია მათ მიერ შემუშავებული ნეირონის მოდელი, სადაც ამჟამავი $\sum$ ელემენტი თითოეულ შემავალ x სიგნალს ამრავლებს w სინაფსურ წონაზე და აჯამებს შეწონილ ნამრავლებს. თუ ჯამი მეტია წინასწარ დადგენილი ზღურბლის მნიშვნელობაზე, მაშინ გამოსასვლელზე არის ლოგიკური სიგნალი „ერთი“, თუ ნაკლებია ზღურბლის მნიშვნელობაზე, მაშინ გამოსასვლელზე იქნება სიგნალი „ნული“. ამ სისტემამ და მრავალმა მსგავსმა მიიღო პერსეპტრონის (perception - აღქმა) დასახელება. XX საუკუნის სამოციან წლებში პერსეპტრონმა დიდი ინტერესი და ოპტიმიზმი გამოიწვია

[43]. როზენბლატმა და უიდროუმ მიიღეს რიგი დამაიმედებელი შედეგები, მაგრამ მალე აღმოჩნდა, რომ

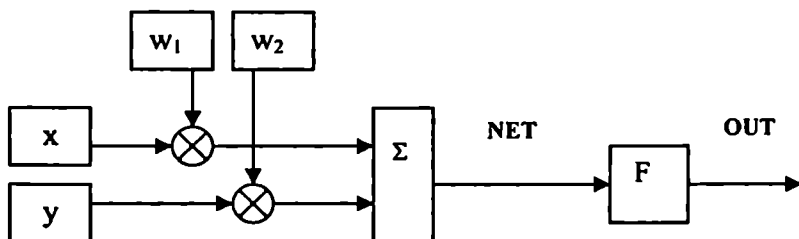


სურ. 2.6. მაკ-კალოკის და პიტსის პერსეპტრონული ნეირონი.

პერსეპტრონს არ შეუძლო ზოგიერთი მარტივი ამოცანის გადაწყვეტა. ნაშრომში [22] მინსკიმ და პეიპერტმა მაღალ თეორიულ დონეზე განიხილეს პერსეპტრონების პრობლემა და მკაცრად აჩვენეს თუ რისი შესრულება შეუძლია მას და რისი არა, შესაბამისად რისი სწავლება შეუძლია ერთშრიან პერსეპტრონებს. სწორედ მინსკის და პეიპერტის კლასიკური ნაშრომის გამოქვეყნების შემდეგ მკვლევართა უმრავლესობამ მრავალი წლით დაკარგა ინტერესი ნეიროქსელებისადმი. მდგომარეობა შეიცვალა მხოლოდ მას შემდეგ, როცა აღმოჩენილ იქნა მრავალშრიანი ნეიროქსელების სწავლების მეთოდები. მიუხედავად თავიანთი შეზღუდული შესაძლებლობისა, პერსეპტრონების თეორია წარმოადგენს საფუძველს მრავალი სხვა ტიპის ნეიროქსელებისათვის, რომლებიც შემდეგ შეიქმნა.

2.7. ლოგიკური არატოლმნიშვნელობის პრობლემა

მინსკის ნაშრომის [22] ერთ-ერთი შედეგი გვინვენებს, რომ ერთშიან პერსეპტრონებს არ შეუძლია ისეთი მარტივი ლოგიკური ფუნქციის განხორციელება, როგორიც არის მაგალითად, ლოგიკური არატოლმნიშვნელობა (შეკრება 2-ის მოდულით). ეს არის ორი ბინარული არგუმენტის ფუნქცია. იგი იღებს „ერთის“ მნიშვნელობას მხოლოდ მაშინ, როდესაც ერთ-ერთი არგუმენტი არის „ერთი“ და ნულის მნიშვნელობას, როდესაც ორივე არგუმენტი არის „ნული“ ან ორივე არგუმენტი არის „ერთი“. პრობლემის ილუსტრაცია შეიძლება ერთნეირონიანი ორშესასვლელიანი (x, y) პერსეპტრონით (სურ. 2.7).

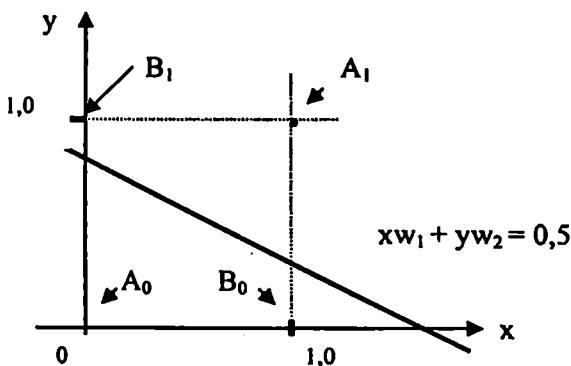


სურ. 2.7. ერთნეირონიანი პერსეპტრონი ორი შესასვლელით.

შესასვლელების ყველა შესაძლო კომბინაცია $x - y$ სიბრტყეზე წარმოდგენილია ოთხი წერტილით, როგორც ეს ნაჩვენებია 4.8 სურათზე. მაგალითად, წერტილი კოორდინატებით $x = 0$ და $y = 0$ სურათზე აღნიშნულია როგორც A_0 წერტილი.

წერტილები	x	y	გამოსასვლელი
A_0	0	0	0
B_0	1	0	1
B_1	0	1	1
A_1	1	1	0

22 ცხრილი წარმოადგენს ლოგიკური არატოლმნიშვნელობის ჭეშმარიტების ცხრილს, სადაც შემავალი კომბინაციები, რომლებმაც გამოსასვლელზე უნდა მოგვცეს „ნული“, აღნიშნულია როგორც A_0 და A_1 , ხოლო რომლებმაც გამოსასვლელზე უნდა მოგვცეს „ერთი“, აღნიშნულია როგორც B_0 და B_1 წერტილები.



სურ2.8. ლოგიკური არატოლმნიშვნელობის პრობლემის ილუსტრაცია

2.8 სურათზე F ფუნქცია არის ზღურბლი, ამიტომ OUT იღებს „ნულის“ მნიშვნელობას, როდესაც $NET < 0,5$,

და „ერთის“ მნიშვნელობას, როდესაც $NET \geq 0,5$. ნეირონი ასრულებს შემდეგ გამოთვლას:

$$NET = xw_1 + yw_2.$$

ორი სინაფსური წონის არანაირი მნიშვნელობის კომბინაციას არ შეუძლია მოგვცეს 2.2 ცხრილით განსაზღვრული შესაბამისობა შესასვლელსა და გამოსასვლელს შორის. ამის დასამტკიცებლად NET დავაფიქსირეთ 0,5 ზღურბლის სიდიდეზე. ამ შემთხვევაში პერსეპტრონი აღიწერება შემდეგი განტოლებით:

$$NET = xw_1 + yw_2 = 0,5.$$

ეს განტოლება არის წრფივი x და y წერტილების მიმართ, ანუ არგუმენტის ყველა მნიშვნელობა, რომელიც აკმაყოფილებს განტოლებას, განლაგებული იქნება x - y სიბრტყის რაიმე წრფეზე. არგუმენტების მნიშვნელობები წრფის ერთ მხარეზე უზრუნველყოფს ზღურბლზე მეტი სიდიდის NET-ს, ანუ შესაბამისად უზრუნველყოფს $OUT = 1$ და $OUT = 0$, ხოლო w_1 და w_2 წონების ცვლილებები იწვევს წრფის დახრისა და მდებარეობის ცვლილებას. იმისათვის, რომ ქსელმა შეასრულოს ლოგიკური არატოლმნიშვნელობის ფუნქცია, წრფე უნდა მოთავსდეს ისე, რომ მის ერთ მხარეს იყოს A წერტილები და მეორე მხარეს - B წერტილები; ეს კი შეუძლებელია. ამრიგად, რა მნიშვნელობებიც უნდა მივანიჭოთ ზღურბლსა და სინაფსურ წონებს, ქსელი ვერ შეძლებს ლოგიკური არატოლმნიშვნელობის ფუნქციის წარმოდგენას.

2.8. წრფივად განმხილოებადობის პრობლემა

როგორც ვნახეთ, შეუძლებელი აღმოჩნდა წრფის ისეთი განლაგება $x-y$ სიბრტყეზე, რომ წრფემ განამხოლოს A და B წერტილები ისე, რომ მის ერთ მხარეს იყოს A წერტილები და მეორე მხარეს B წერტილები, ანუ ქსელმა განახორციელოს ლოგიკური არატომნიშვნელობის (შეკრება 2-ის მოდულით) ფუნქცია. სამწუხაროდ, ეს მაგალითი არ არის ერთადერთი, არამედ არსებობს ფუნქციების დიდი კლასი, რომლებსაც ვერ ახორციელებს ერთშრიანი ხელოვნური ნეიროქსელები, ასეთ ფუნქციებს უწოდებენ წრფივად განუმხილოებელს.

წრფივად განმხილოებადობა ერთშრიანი ხელოვნური ნეიროქსელების შესაძლებლობებს ზღუდავს კლასიფიკაციისა და კლასტერირების ისეთი ამოცანებით, სადაც წერტილების სიმრავლეების (რომლებიც შეესაბამება ნეიროქსელში შემავალ სიგნალებს) განმხილოება შეიძლება გეომეტრიულად. ზემოთ განხილულ მაგალითში, სადაც ორი შესასვლელი იყო, გამყოფის როლს სიბრტყეზე ასრულებდა წრფე. თუ გვექნება სამი შესასვლელი, მაშინ გამყოფის როლს სამგანზომილებიან სივრცეში შეასრულებს სიბრტყე. ოთხი და მეტი შესასვლელის შემთხვევაში კი უნდა წარმოვიდგინოთ შესაბამისად ოთხი და მეტი განზომილების სივრცე, რომელშიც განმხილოებას გეომეტრიული ობიექტი – ჰიპერსიბრტყე ასრულებს.

ცნობილია, რომ კლასტერირების ამოცანების დაყოფა პირობითად შესაძლებელია სამ ჯგუფად: პირველ ჯგუფს შეადგენს ისეთი ამოცანები, სადაც მოცემული სიმრავლის არაგადამკვეთ ქვესიმრავლებად დაყოფა-დაჯგუფება (განმხილოება) შეიძლება ჰიპერსიბრტყეებით – ასეთი სიმრავლეები მიეკუთვნება წრფივად განმხილოებად სიმრავლეებს; მეორე ჯგუფს შეადგენს ისეთი ამოცანები, სადაც მოცემული სიმრავლის არაგადამკვეთ ქვესიმრავლებად

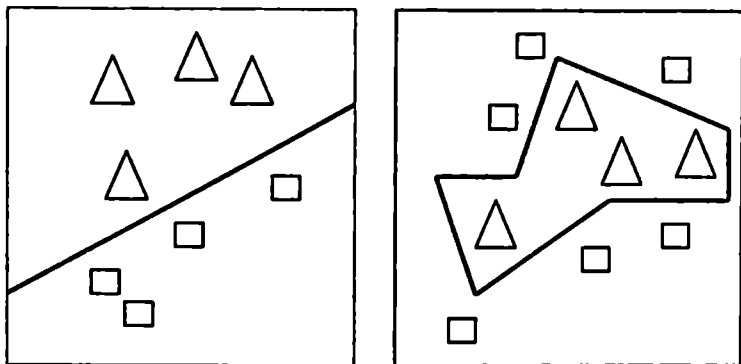
განმხოლოება ჰიპერსიბრტყეებით შეუძლებელია. თუმცა, დაყოფის უფრო რთული მეთოდების გამოყენებით ზოგიერთ ამოცანებში შესაძლებელი ხდება მათი განმხოლოება - ასეთი სიმრავლეები მიეკუთვნება არაწრფივად განმხოლოებად სიმრავლეებს; მესამე ჯგუფს შეადგენს ისეთი ამოცანები, სადაც ქვესიმრავლეები იკვეთება; ასეთ შემთხვევებში მსჯელობენ მხოლოდ ალბათურ განმხოლოებადობაზე. რეალური ამოცანების გადაწყვეტისას ხშირად ვერ ხერხდება საწყისი მონაცემების წინასწარი დამუშავება, რომელიც უზრუნველყოფს სახეთა წრფივად განმხოლოებადობას. ეს კი ართულებს არაგადამკვეთ ქვესიმრავლეებად დაჯგუფების პროცესს და შესაბამისად, კლასტერიზებისათვის საჭირო ნეირონული ქსელების სინთეზს. ამრიგად, მნიშვნელოვანია იმის გარკვევა არის თუ არა მოცემული სახეთა ნაკრები წრფივად განმხოლოებადი. ცნობილია, რომ, თუ სახეთა აღწერისათვის გამოყენებული ნიშანთა რაოდენობა დიდია, არ არსებობს წრფივად განმხოლოებადობის დადგენის საკმარისად მარტივი მეთოდები. უკვე ოთხზე მეტი ნიშნის არსებობის დროს იმის ალბათობა, რომ არჩეული ფუნქცია იქნება წრფივად განმხოლოებადი, ხდება მცირე სიდიდე. ასე, მაგალითად, n რაოდენობის ბინარული შესასვლელის მქონე ნეირონს შეიძლება მიეწოდოს 2^n რაოდენობის სხვადასხვა რეალიზაცია. იმის გამო, რომ თითოეული სახის რეალიზაცია შეიძლება შეესაბამებოდეს საფეხურისებრი აქტივაციის ფუნქციის მქონე ნეირონის გამოსასვლელის ორ მდგომარეობას (ნული ან ერთი), გვექნება n ცვლადის 2^n ფუნქცია.

ცხრილი 23. წრფივად განმხოლოებადი ფუნქციები

n	2^n	წრფივად განმხოლოებადი ფუნქციების რაოდენობა
1	4	4
2	16	14
3	256	104
4	65536	1882
5	$4,3 \times 10^9$	94572
6	$1,8 \times 10^{19}$	15 028 134

2.3 ცხრილიდან [21] ჩანს, რომ ორი არგუმენტის 16 ფუნქციიდან წრფივად განმხოლოებადი არის 14, სამი არგუმენტის 256 ფუნქციიდან კი – 104, ხოლო ოთხი ცვლადის 65536 ფუნქციიდან – მხოლოდ 1882 ფუნქცია.

ამრიგად, იმის აღბათობა, რომ შემთხვევით არჩეული ფუნქცია იქნება წრფივად განმხოლოებადი, არგუმენტების რიცხვის ზრდასთან ერთად ძალზე უმნიშვნელო ხდება. 2.9 სურათზე მოცემულია წრფივად განმხოლოებადი (მარცხნივ) და წრფივად განუმხოლოებადი (მარჯვნივ) სიმრავლეების მაგალითები. წრფივად განმხოლოებადობის პრობლემა კარგად არის შესწავლილი და დამტკიცებულია, რომ მისი დაძლევა შესაძლებელია ერთშრიანი ნეირონული ქსელისათვის დამატებითი შრეების კასკადური მიერთებით [2, 21]. ასეთ მრავალშრიან ნეირონულ ქსელებს შეუძლია რთული კლასტერიზების პროცესების განხორციელება. მართალია მრავალშრიანი ნეირონული ქსელების შესაძლებლობები ცნობილი იყო, მაგრამ ბოლო პერიოდამდე არ იყო შემუშავებული მათი სინაფსური წონების შეთანაწყობის თეორიულად დასაბუთებული ალგორითმები.



სურ. 2.9. წრფივად განმხილვადი (მარცხნივ) და წრფივად განუმხილვადი (მარჯვნივ) სიმრავლეების მაგალითები.

ასეთი ალგორითმები და მათი სხვადასხვა მოდიფიკაცია უკვე მრავალია დამუშავებული. მათ მომდევნო თავებში განვიხილავთ.

2.9. პერსპექტონის სწავლება

როგორც ნაჩვენები იყო, წრფივად განმხილვადობის კრიტერიუმის გამოყენებით დადგინდება შეუძლია თუ არა ერთშირიან ხელოვნურ ნეირონულ ქსელს საჭირო ფუნქციის შესრულება. თუ გაირკვევა, რომ ქსელი შეძლებს საჭირო ფუნქციის შესრულებას, იმისათვის, რომ ქსელს პქონდეს პრაქტიკული ღირებულება, უნდა ვიპოვოთ სინაფსური წონებისა და ზღურბლების საჭირო მნიშვნელობები. ამ მნიშვნელობების გამოთვლისათვის კი აუცილებელია შესაბამისი ალგორითმის არსებობა.

ასეთი ალგორითმი შექმნა ფ. როზენბლატმა [23], რომელიც 2.10 სურათზე არის წარმოდგენილი.

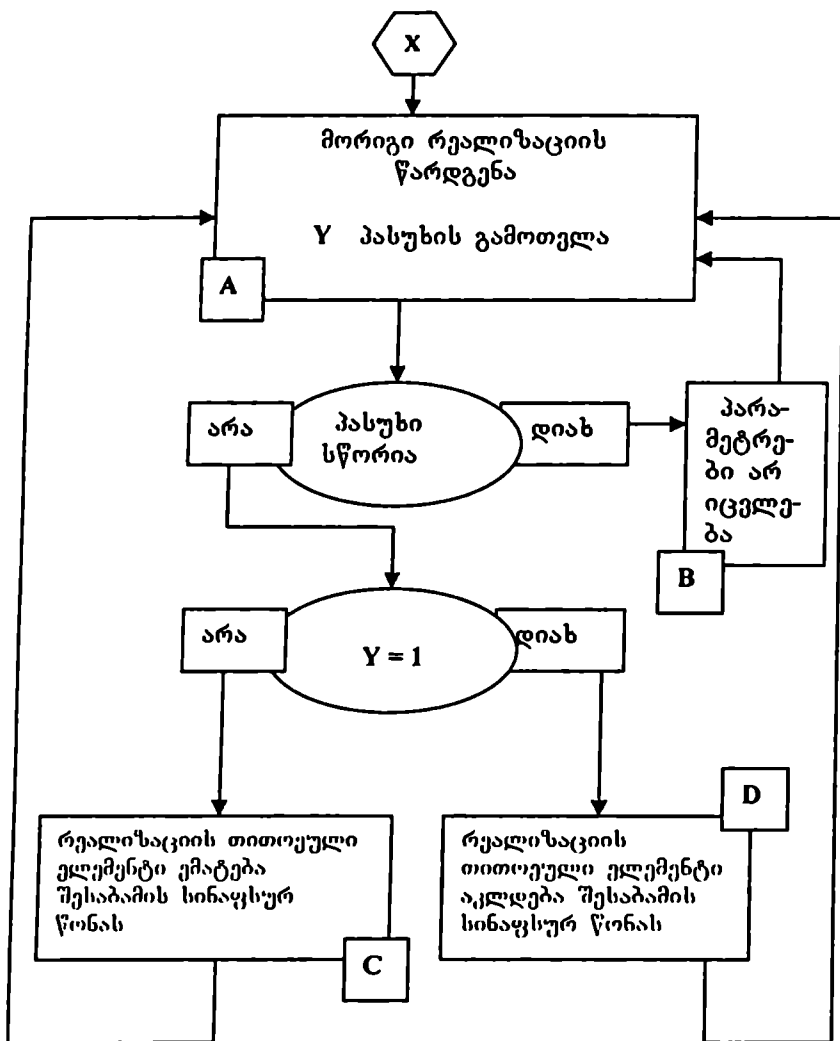
ზოგადად, ნეიროქსელებისათვის, როგორც აღწერილია, განასხვავებენ სწავლების ალგორითმებს მასწავლებლით (სუპერვიზორით) და მასწავლებლის გარეშე, ხოლო პერსექტრონის სწავლება შესაძლებელია მხოლოდ მასწავლებლით, რომელიც ნეიროქსელის მიმართ გარეშე ელემენტია. მასწავლებლის ფუნქციაა სისტემის ქცევის შეფასება და სისტემის მოდიფიცირების მართვა, ანუ ნეიროქსელის პარამეტრების შეთანაწყობა.

ფ. როზენბლატის გამოკვლევებმა ძლიერი სტიმული მისცა ნეიროქსელური თეორიის განვითარებას. მისი სწავლების ალგორითმის ელემენტები გვხვდება თანამედროვე ნეიროქსელურ პარადიგმებშიც.

ფ. როზენბლატის ალგორითმის მიხედვით პერსექტრონის სწავლება მიმდინარეობს შემდეგნაირად: სახეთა სიმრავლის რეალიზაციები რიგრიგობით მიეწოდება პერსექტრონის შესასვლელს და სინაფსური წონების შეთანაწყობა გრძელდება მანამ, სანამ პერსექტრონის გამოსასვლელზე მიიღწევა ყველა რეალიზაციისათვის სიგნალის სასურველი მნიშვნელობა.

განვიხილოთ პერსექტრონის სწავლების მაგალითი. ვთქვათ, მის შესასვლელს მიეწოდება ციფრების გამოსახულებები, ხოლო პერსექტრონს უნდა შევასწავლოთ კენტი და ლუწი ციფრების გარჩევა, ანუ გამოსასვლელზე უნდა მივიღოთ “ერთი”, თუ წარედგინება კენტი ციფრი და „ნული“, თუ წარედგინება ლუწი ციფრი. ვთქვათ, X შემავალი ვექტორი არის ციფრის რეალიზაცია, მაშინ მისი თითოეული ელემენტი ($x_1, x_2, \dots, x_n$) უნდა გამრავლდეს W სინაფსური წონების ვექტორის შესაბამის ელემენტზე ($w_1, w_2, \dots, w_n$), ხოლო ნამრავლები აიჯამოს. ამის შემდეგ უნდა ჩატარდეს ზღურბლური ოპერაცია: თუ შეწონილი ჯამი აღემატება ნეირონის ზღურბლს, მაშინ გამოსასვლელზე იქნება სიგნალი “ერთი” ($y = 1$); თუ შეწონილი ჯამი ნაკლებია ნეირონის ზღურბლზე, მაშინ გამოსასვლელზე იქნება სიგნალი „ნული“ ($y = 0$), ანუ პირველ შემთხვევაში

პერსპექტრონი გვინვენებს, რომ წარდგენილია კენტი ციფრი, ხოლო მეორე შემთხვევაში - ამოსაცნობად წარდგენილია ლუწი ციფრი. პერსპექტრონის სწავლებისათვის ციფრის რეალიზაცია X წარედგინება შესასვლელს და გამოითვლება y გამოსასვლელი. თუ პასუხი სწორია სინაფსური წონები არ იცვლება; თუ პასუხი არასწორია, მაშინ იმ სინაფსური წონების, რომლებიც არასწორი შედეგის გამაძლიერებელ შესასვლელებს უკავშირდება, ისეთი მოდიფიცირება უნდა მოხდეს, რომ შეცდომა შემცირდეს. შესაბამისი პროცედურები ასახულია პერსპექტრონის სწავლების როზენბლატის ალგორითმში (სურ. 2.10).



სურ. 2.10. პერსეპტრონის სწავლების როზენბლატის ალგორითმი.

თუ პასუხი სწორია სინაფსური წონები არ იცვლება, თუ პასუხი არასწორია, მაშინ იმ სინაფსური წონების, რომლებიც არასწორი შედეგის გამაძლიერებელ შესასვლელებს უკავშირდება, ისეთი მოდიფიკირება უნდა მოხდეს, რომ შეცდომა შემცირდეს. შესაბამისი პროცედურები ასახულია პერსეპტრონის სწავლების როზენბლატის ალგორითმში (სურ. 2.10).

2.10. უწყვეტი ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი

ფ. როზენბლატის სწავლების აღწერილი ალგორითმი, რომელიც ბინარული ნეირონული ქსელებისათვის არის განკუთვნილი, შეიძლება განსოგადდეს უწყვეტი ნეირონული ქსელებისათვის. ამისათვის საჭიროა δ სიდიდის შემოტანა, რომელიც წარმოადგენს სხვაობას ქსელის გამოსასვლელზე საჭირო, ანუ მიზნობრივი (target) სიგნალისა და გამოსასვლელზე რეალურად მიღებულ Y სიგნალის სიდიდეებს შორის: $\delta = (T - Y)$.

თუ $\delta = 0$, მაშინ $T = Y$, ანუ გამოსასვლელზე მიღებულია სწორი პასუხი. ეს შემთხვევა შეესაბამება 4.10 სურათზე წარმოდგენილი ალგორითმის B ბლოკს; თუ δ მეტია ნულზე, ეს შეესაბამება ალგორითმის C ბლოკს; ხოლო, თუ δ ნაკლებია ნულზე, ეს სიტუაცია შეესაბამება ალგორითმის D ბლოკს. ყველა ამ შემთხვევისათვის აღწერილი სწავლების ალგორითმი ძალაშია, თუ δ გამრავლდება შემავალი ვექტორის თითოეულ x_i ელემენტზე და ეს δx_i ნამრავლი აღემატება შესაბამის სინაფსურ წონას. განსოგადების მიზნით, უნდა შემოვიტანოთ სწავლების η სინქარის კოეფიციენტი, რომელიც უნდა გამრავლდეს δx_i სიდიდეზე, რაც მოგვცემს სინაფსური წონების საშუალო სიდიდის

მართვის შესაძლებლობას. ალგებრულ ფორმაში გვექნება:

$$\Delta_i = \eta \delta x_i$$
$$w_i(n+1) = w_i(n) + \Delta_i$$

სადაც Δ_i არის i -ური შესასვლელის კორექცია, ხოლო $w_i(n+1)$ – i -ური სინაფსური წონის მნიშვნელობა კორექციის შემდეგ; $w_i(n)$ არის i -ური სინაფსური წონის მნიშვნელობა კორექციამდე.

ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი წარმოდგენილია 2.11 სურათზე.

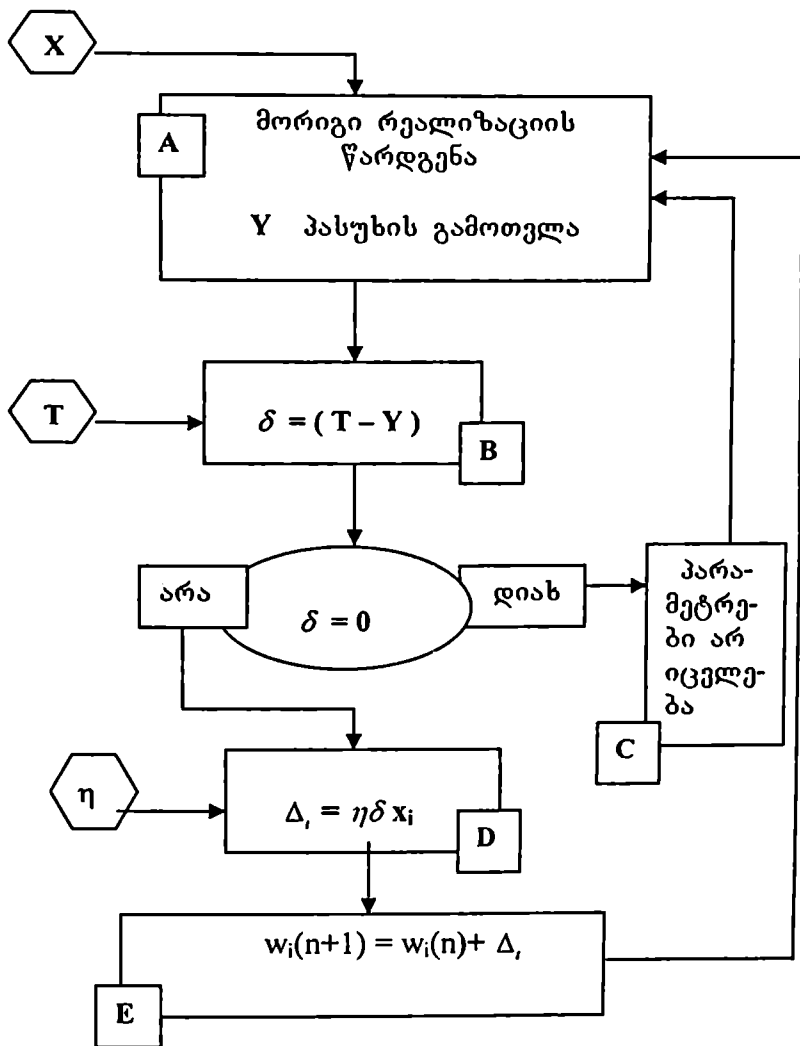
ამასთან ერთად, უნდა გავითვალისწინოთ, რომ პერსეპტრონის სწავლების ალგორითმის პრაქტიკული განხორციელების დროს შესაძლოა წავაწყდეთ შემდეგ სიძნელეებს:

ზოგჯერ ძნელია კონკრეტული სასწავლო ნაკრებისათვის იმის დადგენა, არის თუ არა შესრულებული წრფივად განმხილოებადობის კრიტერიუმი;

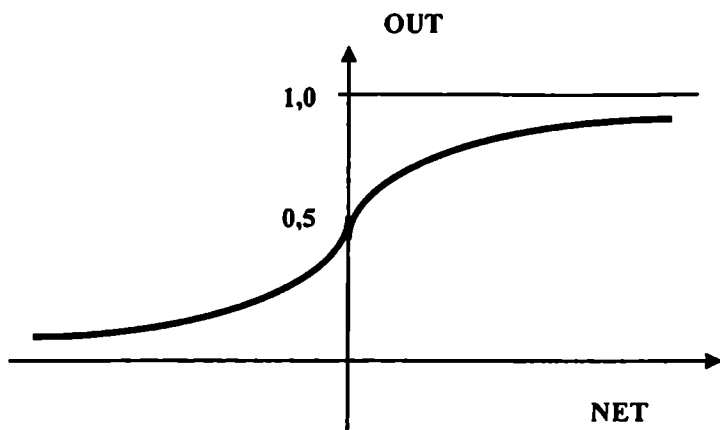
პრაქტიკულ ამოცანებში, როდესაც ქსელის შესასვლელზე სწრაფად იცვლება რეალიზაციები, შესაძლებელია კრიტერიუმი დროის ზოგ მომენტში შესრულდეს, ხოლო დროის სხვა მომენტში არა;

შესაძლებელია დამტკიცდეს, რომ ქსელის სწავლება დამთავრდება სასრული რაოდენობის ბიჯების შემდეგ. თუმცა, ეს სასრული რაოდენობის ბიჯები შეიძლება აღმოჩნდეს ასტრონომიული სიდიდეების თანაზომადი და ცხადია, პრაქტიკისათვის მიუღებელი.

არ არის დამტკიცებული, რომ ზემოთ განხილული სწავლების ალგორითმი უფრო სწრაფია, ვიდრე სინაფსური წონების შეთანაწყობა მათი ყველა შესაძლო მნიშვნელობის მარტივი გადასინჯვის მეთოდით. ყველა აქ ჩამოთვლილი პრობლემა თანამედროვე ნეიროქსელური კვლევებისათვისაც არის აქტუალური.



სურ. 2.11. ნეირონული ქსელების დელტა-წესით სწავლების ალგორითმი.



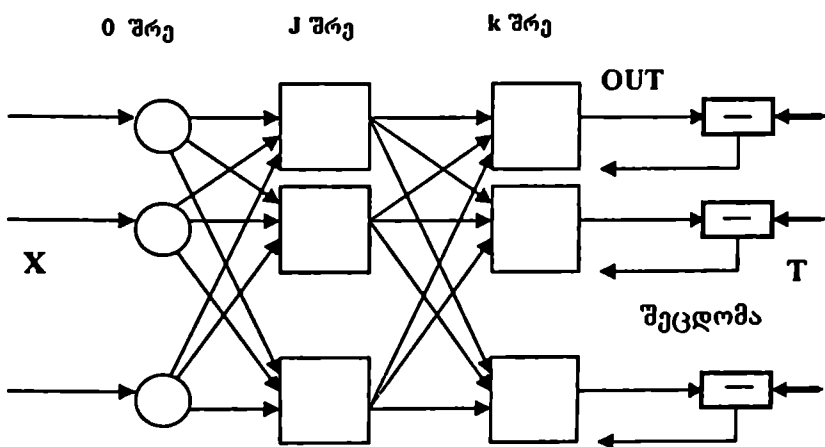
სურ. 2.13. სიგმოიდური აქტივაციის ფუნქცია.

სიგმოიდის გარდა, უკუგაპროცესების პროცედურაში შესაძლებელია სხვა არაწრფივი ფუნქციების გამოყენებაც, იმ პირობით, რომ ისინი, ისევე როგორც სიგმოიდი, დიფერენცირებადნი უნდა იყვნენ განსაზღვრის მოელს არეზე; თუმცა, სიგმოიდის გამოყენება განსაკუთრებით მოხერხებელია, ვინაიდან მისი წარმოებული, რომელიც აუცილებელია უკუგაპროცესების პროცედურის შესასრულებლად, მარტივი სახისაა

$$F'(NET) = \frac{\partial OUT}{\partial NET} = OUT(1 - OUT). \quad (2.9)$$

მრავალშრიანი ხელოვნური ნეირონული ქსელი, რომლის სწავლებაც შესაძლებელია უკუგაპროცესების პროცედურით, ნაჩვენებია 2.14 სურათზე. ნეირონების პირველი შრე ასრულებს მხოლოდ შემავალი სიგნალების მიღებისა და განაწილების პროცედურას, ანუ შემავალი სიგნალები გაივლის პირველი შრის ნეირონებს როგორც გამანაწილებელ წერტილებს შესაბამისი სინაფსური წონებისაკენ რაიმე გარდაქმნის გარეშე, ხოლო მომდევნო შრის ნეირონები გამოითვლის NET და OUT სიგნალებს ისე, როგორც ეს აღწერილია ზემოთ. სწორედ ამიტომ 2.14 სურათზე ნაჩვენებია

უკუგავრცელების ნეირონული ქსელი განიხილება როგორც ორშრიანი ქსელი, ხოლო გამანაწილებელ შრეს ეწოდება ნულოვანი. უკუგავრცელების პროცედურის გამოყენება შესაძლებელია ნებისმიერი რაოდენობის შრეების მქონე ხელოვნური ნეირონული ქსელებისათვის. ამასთან, ამ ალგორითმის გამოყენება შეიძლება როგორც პირდაპირი გავრცელების, ასევე უკუკავშირებიანი ნეირონული ქსელების სწავლებისათვის. ნეირონული ქსელების სწავლების მიზანია მისი სინაფსური წონების ისეთი შეთანაწყობა, რომ შემავალი სიგნალების გარკვეულ



სურ. 2.14. უკუგავრცელების ნეირონული ქსელი.

სიმრავლეს (შემავალ X ვექტორს) შეესაბამებოდეს სასურველი გამოძევალი სიგნალების სიმრავლე (გამომავალი, ანუ მიზნობრივი T ვექტორი). სწავლების პროცესის განხორციელებისას ეთვლით, რომ ყოველი შემავალი ვექტორისათვის არსებობს მისი მეწყვეილე მიზნობრივი ვექტორი, ორივე ვექტორს ერთად ეწოდება

მასწავლებელი წყვილი. იმ შემთხვევაში, თუ ქსელმა უნდა ამოიცნოს სახეთა რომელიმე სიმრავლე, საჭირო გახდება იმდენივე მასწავლებელი წყვილის ფორმირება, რამდენი სახეც არის ამოსაცნობი. ასეთ ჯგუფს ეწოდება მასწავლებელი სიმრავლე. სწავლების დაწყების წინ თითოეულ სინაფსურ წონას უნდა მიენიჭოს რაიმე საწყისი მნიშვნელობა. პრაქტიკაში ცდილობენ, რომ საწყისი სინაფსურ წონებს მიენიჭოს მცირე შემთხვევითი რიცხვითი მნიშვნელობები, რათა სწავლების პროცესში ნეირონული ქსელი არ გადაიტვირთოს სინაფსური წონების დიდი მნიშვნელობებით და თავიდან იქნეს აცილებული ზოგიერთი პათოლოგიური შემთხვევა. მაგალითად, თუ სინაფსურ წონებს მიენიჭება ტოლი საწყისი ნიშნელობები, ხოლო ქსელის ფუნქციონირებისათვის აუცილებელია არატოლი სინაფსური წონები, ქსელი შეიძლება არასწავლებადი აღმოჩნდეს.

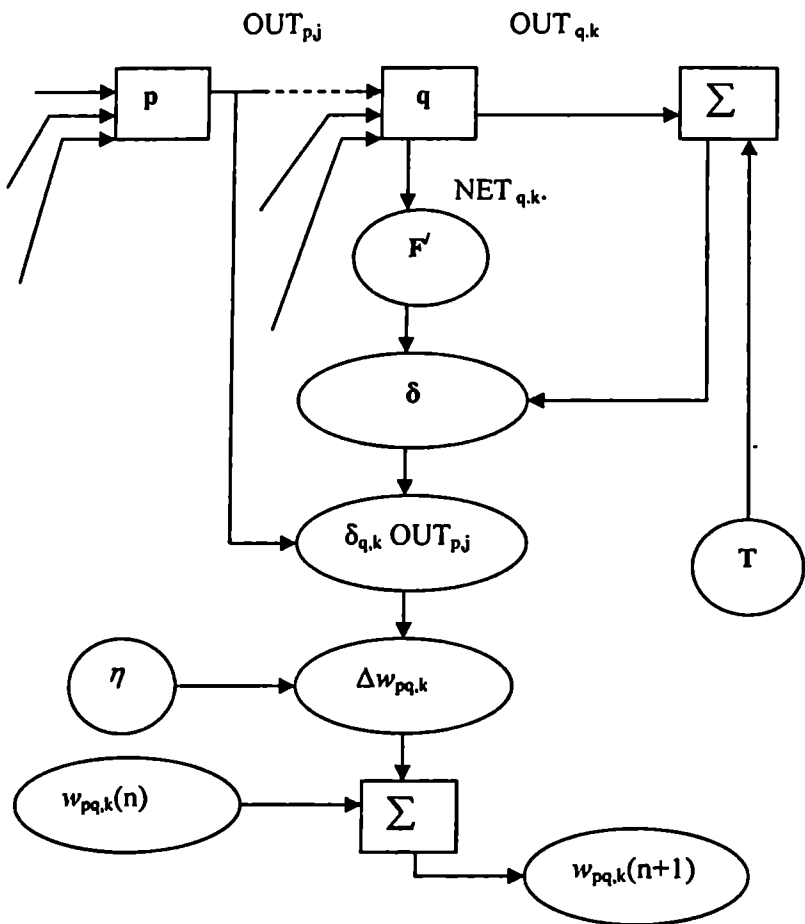
ნეირონული ქსელების სწავლებისათვის უკუგავრცელების ალგორითმის გამოყენებით საჭიროა შემდეგი პროცედურების ჩატარება:

1. მორიგი მასწავლებელი წყვილის არჩევა და შემავალი ვექტორის წარდგენა ქსელისათვის;
2. ქსელის რეალური OUT გამოსასვლელის გამოთვლა;
3. ქსელის რეალურ OUT გამოსასვლელსა და მიზნობრივ (Target) გამოსასვლელს შორის სხვაობის გამოთვლა;
4. ქსელის სინაფსური წონების კორექტირება ისე, რომ შემცირდეს შეცდომა (სხვაობა);
5. ბიჯების (1,2,3,4) განმეორება მასწავლებელი სიმრავლის თითოეული წყვილისათვის მანამ, სანამ შეცდომა არ გახდება მისაღები სიდიდის.

გამოთვლები სრულდება ნეირონული ქსელების შრეების მიხედვით, მაგალითად 2.14 სურათზე ნაჩვენები ქსელისათვის ჯერ უნდა გამოითვალოს j შრის გამოსასვლელები, რომლებიც k შრის შესასვლელებად

გამოიყენება და მათი მეშვეობით გამოითვალოს k შრის გამოსასვლელები, რომლებიც ქმნის ქსელის გამოსასვლელ OUT ვექტორს. მესამე ბიჯზე მიზნობრივ (Target) გამოსასვლელს უნდა გამოაკლდეს ქსელის შესაბამისი რეალური OUT გამოსასვლელი, რაც მოგეცემს ქსელის შეცდომას (Error), რომელიც მეოთხე ბიჯზე გამოიყენება სინაფსური წონების კორექციისათვის. ამასთან, წონების ცვლილებების ნიშანი და სიდიდე განისაზღვრება სწავლების ქვემოთ მოცემული ალგორითმის თანახმად. ამ ოთხი ბიჯის საკმარისი რაოდენობით განმეორების (იტერაციის) შემდეგ, ქსელის რეალურ OUT გამოსასვლელსა და მიზნობრივ (Target) გამოსასვლელს შორის სხვაობა უნდა შემცირდეს მისაღებ სიდიდემდე, მაშინ ითვლება, რომ ქსელის სწავლება დამთავრებულია. ამის შემდეგ სინაფსური წონების კორექცია აღარ ხდება და ქსელის გამოყენება შეიძლება პრაქტიკული ამოცანების გადასაწყვეტად. პირველი და მეორე ბიჯები განიხილება როგორც სიგნალის პირდაპირი გავრცელება, ხოლო მესამე და მეოთხე ბიჯები შეიძლება ჩაეთვალოთ სიგნალის უკუგავრცელებად, ანუ ამ დროს გამოთვლილი შეცდომის სიგნალი ქსელში ვრცელდება ძირითადი სიგნალის გავრცელების საპირისპირო მიმართულებით და გამოიყენება სინაფსური წონების შეთანაწყობისათვის.

სიგნალის პირდაპირი და უკუგავრცელება განვიხილოთ უფრო დაწვრილებით და ეს პროცედურები მათემატიკური ფორმით გამოვსახოთ.



სურ. 2.15. სინაფსური წონის შეთანაწყობა გამომავალ შრეში.

2.11.1. პირდაპირი გავრცელების პროცედურა

ვექტორული ფორმით პირველი და მეორე ბიჯები შეიძლება შემდეგნაირად წარმოვადგინოთ: ქსელს წარედგინება X შემავალი ვექტორი და გამოსასვლელზე მიიღება O გამოშვებული ვექტორი, ვექტორული წყვილი X და T უნდა ავიღოთ მასწავლებელი სიმრავლიდან. სინაფსური წონები განიხილება როგორც მატრიცა W , მაშინ ვექტორული ნოტაციით ქსელის შრისათვის გამოთვლითი პროცესი აღიწერება შემდეგი გამოსახულებით:

$$O = F(XW). \quad (2.10)$$

მოცემული შრის გამოშვებული ვექტორი იმედროულად არის მომდევნო შრისათვის შემავალი ვექტორი, ამიტომ უკანასკნელი შრის გამოსასვლელი ვექტორის გამოსათვლელად აუცილებელია (2.10) პროცედურის განხორციელება თითოეული შრისათვის, დაწყებული ქსელის შესასვლელიდან.

2.11.2. უკუგავრცელების პროცედურა

განვიხილოთ გამოშვებული შრის სინაფსური წონების შეთანაწყობა. ვინაიდან გამოშვებული შრის თითოეული ნეირონისათვის მოცემული გვაქვს მიზნობრივი მნიშვნელობა, ამიტომ სინაფსური წონების შეთანაწყობა ადვილად შეიძლება განვახორციელოთ მოდიფიცირებული დელტა-წესის მიხედვით, რაც განხილული იყო ზემოთ (პ. 2.10). შედარებით უფრო რთულია ფარული (საშუალო) შრეების სწავლების პროცედურები, ვინაიდან შედარების ოპერაციის ჩასატარებლად მათთვის არა გვაქვს მიზნობრივი მნიშვნელობები.

2.15 სურათზე ნაჩვენებია ერთი სინაფსური წონის სწავლების პროცესი j ფარული შრის p ნეირონიდან k გამომაველი შრის q ნეირონამდე. შეცდომის სიდიდეს მივიღებთ, თუ მიზნობრივ მნიშვნელობას (Target) გამოვაკლებთ k გამომაველი შრის ნეირონის OUT სიგნალს. შემდეგ შეცდომის სიდიდე უნდა გამრავლდეს აქტივაციის (შემკუმშავი) ფუნქციის წარმოებულზე, რომელიც k გამომაველი შრის ამ ნეირონისათვის არის გამოთვლილი, რაც მოგვცემს δ სიდიდეს

$$\delta = \text{OUT}(1 - \text{OUT})(\text{Target} - \text{OUT}). \quad (2.11)$$

ამის შემდეგ მრავლდება იმ j ნეირონის OUT სიდიდეზე, საიდანაც გამოდის ეს სინაფსური წონა. ნამრავლი თავის მხრივ უნდა გამრავლდეს η სწავლების სიჩქარის კოეფიციენტზე (ჩვეულებრივ, იგი 0,01-დან 1,0-მდე ფარგლებშია), ხოლო შედეგი უნდა დაემატოს სინაფსურ წონას. ასეთივე პროცედურები უნდა შესრულდეს თითოეული სინაფსური წონისათვის ფარული შრის ნეირონიდან გამომაველი შრის ნეირონამდე:

$$\Delta w_{pq,k} = \eta \delta_{q,k} \text{OUT}_{pj} \quad (2.12)$$

$$w_{pq,k}(n+1) = w_{pq,k}(n) + \Delta w_{pq,k}. \quad (2.13)$$

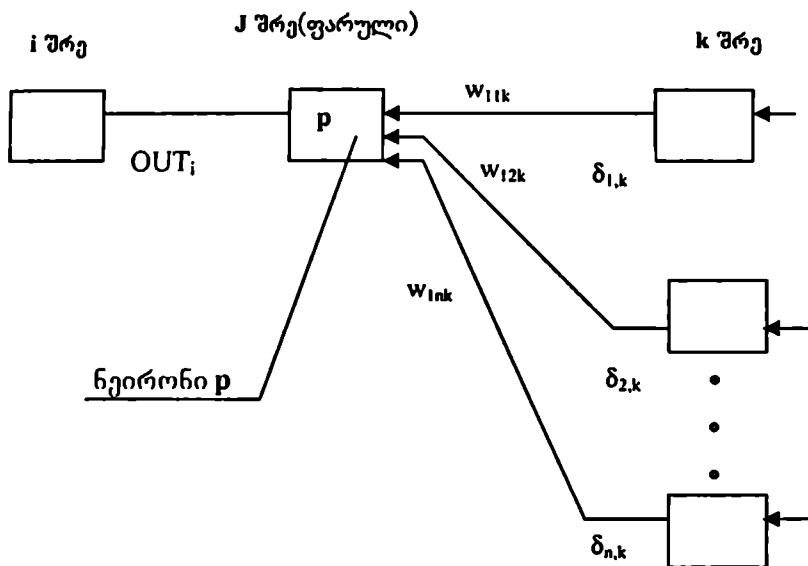
სადაც $w_{pq,k}(n)$ არის სინაფსური წონა ფარული შრის p ნეირონიდან გამომაველი შრის q ნეირონამდე n ბიჯზე კორექციამდე. ამასთან, ინდექსი k ეკუთვნის იმ შრეს, სადაც მთავრდება მოცემული წონა; $w_{pq,k}(n+1)$ – სინაფსური წონა $(n+1)$ ბიჯზე, ანუ კორექციის შემდეგ; $\delta_{q,k}$ არის δ სიდიდე გამომაველი k შრის q ნეირონისათვის; OUT_{pj} არის j ფარული შრის p ნეირონის OUT-ის მნიშვნელობა; η – სწავლების სიჩქარის კოეფიციენტი, რომელიც სინაფსური წონების ცვლილებების საშუალო სიდიდის მართვის შესაძლებლობას იძლევა.

2.11.3. სინაფსური წონების შეთანაწყობა ფარულ შრეში

განვიხილოთ ერთი ნეირონი, რომელიც კუთონის გამომავალი შრის წინამდებარე ფარულ შრეს. ეს ნეირონი პირდაპირი გავრცელების პროცესში თავის გამომავალ სიგნალს გადასცემს გამომავალი შრის ნეირონებს, მათი დამაკავშირებელი სინაფსური წონების საშუალებით. სწავლების პროცესში კი ეს სინაფსური წონები მუშაობს საპირისპირო მიმართულებით, ანუ სიდიდეს გაატარებს გამომავალი შრისაგან უკან, ფარული შრისაკენ, ამასთან, თითოეული ეს წონა უნდა გამრავლდეს გამომავალი შრის იმ ნეირონის სიდიდეზე, რომელსაც იგი უკავშირდება. ფარული შრის ნეირონის შეთანაწყობისათვის აუცილებელი სიდიდე მიიღება ყველა ასეთი ნამრავლის შეჯამებით და შედეგის გამრავლებით აქტივაციის (შემკუმშავი) ფუნქციის წარმოებულზე

$$\delta_{q,k} = \text{OUT}_{p,j} (1 - \text{OUT}_{p,j}) \left[\sum_q \delta_{q,k} w_{pq,k} \right]. \quad (2.14)$$

ფარულ შრეში სინაფსური წონის შეთანაწყობის პროცედურები ნაჩვენებია 2.16 სურათზე. მას შემდეგ, რაც დ სიდიდე მიღებულია, შეიძლება ჩატარდეს პირველი ფარული შრის კუთვნილი სინაფსური წონების შეთანაწყობის პროცედურები (2.12) და (2.13) განტოლებების მიხედვით, სადაც ინდექსები შრის შესაბამისად იქნება მოდიფიცირებული.



სურ. 2.16. სინაფსური წონის შეთანაწყობა ფარულ შრეში.

მოცემული ფარული შრის თითოეული ნეირონისათვის მიღებულ უნდა იქნეს δ სიდიდე და შესრულდეს ამ ფარულ შრესთან დაკავშირებული ყველა სინაფსური წონის შეთანაწყობა. ეს პროცესი მეორდება ყველა შრის მიმართ ნეიროქსელის შემოსასვლელის მიმართულებით, ვიდრე არ დამთავრდება ქსელის ყველა სინაფსური წონის შეთანაწყობა, ანუ ნეიროქსელის სწავლების მიზნით ხდება შეცდომის სიგნალის უკუგაერცვლება, რაც აისახა ალგორითმის დასახელებაში.

2.12. სწავლების უკუგავრცელების ალგორითმის დაჩქარების მეთოდები

ნეირონული ქსელების სწავლების უკუგავრცელების ალგორითმის მრავალი გაუმჯობესება და განზოგადება იქნა შემოთავაზებული მკვლევარების მიერ. განვიხილოთ მათ შორის ყველაზე პერსპექტიული მიდგომები. უკუგავრცელების ალგორითმით სწავლების დაჩქარების ერთ-ერთ მეთოდს ეწოდება იმპულს-მეთოდი. იგი აღწერილია [25] ნაშრომში, სადაც სინაფსური წონის კორექციის გამოსახულებას ემატება წევრი, რომელიც სინაფსური წონის წინა ცვლილების პროპორციულია. როგორც კი კორექციის ოპერაცია ჩატარდება, მისი მნიშვნელობა უნდა იქნეს დამახსოვრებული და გამოყენებული ყველა შემდგომი კორექციის მოდიფიკაციისათვის. კორექციის განტოლებების მოდიფიკაცია ხდება შემდგენიარად:

$$\Delta w_{pq,k}(n+1) = \eta \delta_{q,k} \text{OUT}_{pj} + \alpha \Delta w_{pq,k}(n); \quad (2.15)$$

$$w_{pq,k}(n+1) = w_{pq,k}(n) + \Delta w_{pq,k}(n+1), \quad (2.16)$$

სადაც α არის იმპულსის კოეფიციენტი (ჩვეულებრივ, იღებენ 0,9 სიდიდეს).

ლიტერატურაში [21] აღნიშნულია, რომ იმპულს-მეთოდი კარგ შედეგს იძლევა ზოგიერთი ამოცანის გადაწყვეტისას, მაგრამ სხვა ამოცანების შემთხვევაში შედეგები სუსტი ან ზოგჯერ უარყოფითიც კი არის.

ნაშრომში [26] აღწერილია მსგავსი მეთოდი. იგი ეფუძნება ექსპონენციალურ დაგლუვებას, რომელსაც ზოგიერთ შემთხვევაში შეიძლება პქონდეს უპირატესობა.

$$\Delta w_{pq,k}(n+1) = (1 - \alpha) \delta_{q,k} \text{OUT}_{pj} + \alpha \Delta w_{pq,k}(n). \quad (2.17)$$

შემდეგ გამოითვლება სინაფსური წონის ცვლილება

$$w_{pq,k}(n+1) = w_{pq,k}(n) + \eta \Delta w_{pq,k}(n+1), \quad (2.18)$$

სადაც α არის შესწორების კოეფიციენტი, რომლის ვარირება ხდება 0,0-დან 1,0-მდე დიაპაზონში. თუ α უდრის 1,0, მაშინ ახალი კორექცია იგნორირდება და მეორდება წინა კორექცია; ხოლო ნულსა და ერთს შორის არეში, სინაფსური წონის კორექცია გასწორდება α -ს პროპორციული სიდიდით. სწავლების სინქარის კოეფიციენტი η სინაფსური წონების ცვლილების საშუალო სიდიდის მართვის შესაძლებლობას იძლევა.

[27] ნაშრომში აღწერილია უკუგავრცელების ალგორითმის კრებადობის დანქარების მეთოდი, რომელსაც ეწოდა მეორე რიგის უკუგავრცელება. იგი სინაფსური წონების საჭირო კორექციის უფრო ზუსტი შეფასებისათვის იყენებს მეორე რიგის წარმოებულებს. ამავე ნაშრომში ნაჩვენებია, რომ მეორე რიგის უკუგავრცელება ოპტიმალურია იმ გაგებით, რომ შეფასების გაუმჯობესება შეუძლებელია კიდევ უფრო მაღალი რიგის წარმოებულების გამოყენების გზით.

მეთოდი მოითხოვს დამატებით გამოთვლებს, ამიტომ ცალკე გამოკვლევაა საჭირო იმის დასამტკიცებლად, რომ გამართლებულია ეს მიდგომა.

[28] ნაშრომში აღწერილია უკუგავრცელების ქსელების სწავლების მახასიათებლების გაუმჯობესების საინტერესო მეთოდი, სადაც მითითებულია, რომ ფარული ნეირონების შესასვლელებისა და გამოსასვლელების სოგადად მიღებული დინამიკური დიაპაზონი ნულიდან ერთამდე არ არის ოპტიმალური. ბინარული შემავალი ვექტორის შემთხვევაში საშუალოდ ნახევარი ამ ვექტორის ელემენტებისა იქნება ნული და შესაბამისად, ამ შესასვლელებთან დაკავშირებული სინაფსური წონების სწავლება არ მოხდება, ვინაიდან სინაფსური წონის კორექციის სიდიდე $\Delta w_{pq,k}$ პროპორციულია იმ ნეირონის გამომავალი სიგნალის დონისა, რომელიც OUT_{pj} -ს გამოიმუშავებს. მაშინ ნულოვანი დონე იწვევს იმას, რომ სინაფსური წონა არ იცვლება. ნაპოვნია შემდეგი გამოსავალი: ბინარული

მნიშვნელობები ქსელის შესასვლელზე შეცვლილია $+0,5$ და $-0,5$ მნიშვნელობებით, ხოლო აქტივაციის ფუნქციას ემატება წანაცვლება $-0,5$; მაშინ აქტივაციის ფუნქცია იღებს შემდეგ სახეს:

$$\text{OUT} = -\frac{1}{2} + \frac{1}{1 + e^{-\text{NET}}}. \quad (2.19)$$

შედგად უკუგაგრძელების ალგორითმის კრებადობის დრო მცირდება საშუალოდ 30-დან 50 პროცენტამდე. ეს არის პრაქტიკული მოდიფიკაციის კარგი მაგალითი, როელმაც არსებითად გააუმჯობესა ალგორითმის მახასიათებლები.

სწავლების უკუგაგრძელების ალგორითმი შეიძლება გამოყენებულ იქნეს აგრეთვე უკუაქშირებიანი ნეირონული ქსელებისათვის, რისთვისაც დამუშავებულია შესაბამისი მეთოდიკები.

სწავლების უკუგაგრძელების ალგორითმი წარმატებით გამოიყენება მრავალფეროვანი პრაქტიკული ამოცანების გადაწყვეტის დროს. აქვე უნდა აღინიშნოს ისიც, რომ ზოგიერთი რთული ამოცანის გადაწყვეტისას სწავლების პროცესი შეიძლება ძალზე დიდხანს გაგრძელდეს, ან საერთოდ ვერ მოხდეს ნეიროქსელის სწავლება. ასეთ წარუმატებლობას ჩვეულებრივ ორი მიზეზი აქვს, ეგრეთ წოდებული ქსელის დამბლა ან მოხედარა ლოკალურ მინიმუმში. სწავლების პროცესში კორექციების შედეგად სინაფსურმა წონებმა შეიძლება ძალზე დიდი მნიშვნელობები მიიღოს. ამის გამო, ნეირონების უმეტესობა OUT სიგნალის დიდი მნიშვნელობების პირობებში ფუნქციონირებს იმ არეში, სადაც აქტივაციის ფუნქციის წარმოებული ძალზე მცირე სიდიდეა. იმის გამო, რომ სწავლების პროცესში უკუმიმართულებით გაგზავნილი შეცდომის სიგნალი არის წარმოებული პროპორციული, სწავლების პროცესი შეიძლება განერდეს. ეს პრობლემა თეორიულად სუსტად არის შესწავლილი. მის გადალახვას

ცდილობენ ე ბიჯის შემცირებით, რაც, თავის მხრივ იწვევს სწავლების პროცესის გახანგრძლივებას, ამიტომ ამ პრობლემის შესწავლა გრძელდება.

ქსელის ლოკალურ მინიმუმში მოხვედრის პრობლემა და მისი დაძლევის ხერხები ნაწვენები იქნება სწავლების სტოქასტიკური მეთოდების განხილვის დროს.

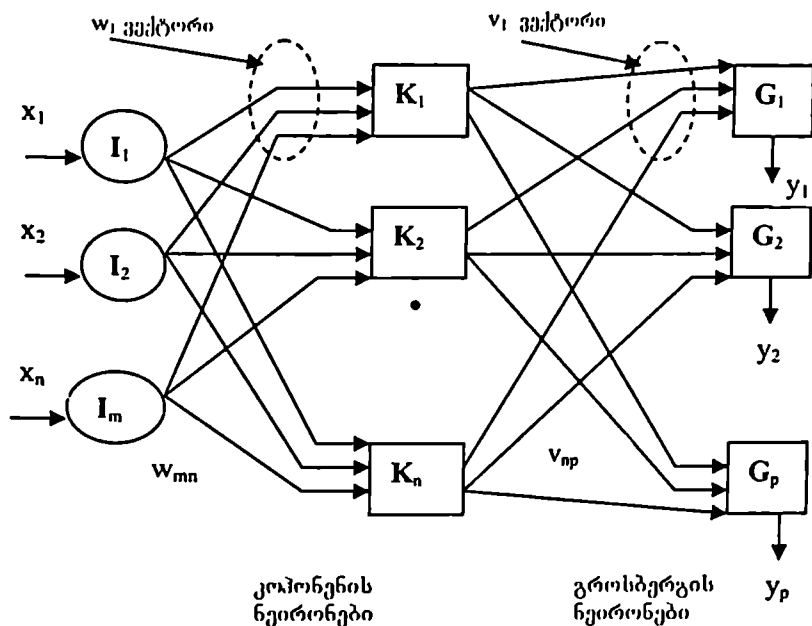
2.13. შემხვედრი გავრცელების ქსელები

შემხვედრი გავრცელების ნეიროქსელების [21] შესაძლებლობები გაცილებით აღემატება ერთშრიანი ნეიროქსელების შესაძლებლობებს; ხოლო უკუგავრცელების ნეიროქსელებთან შედარებით, სწავლებისათვის შეიძლება დასჭირდეთ ასჯერ ნაკლები დრო, თუმცა შემხვედრი გავრცელების ნეიროქსელების გამოყენების არეალი შეზღუდულია, ვიდრე უკუგავრცელების ნეიროქსელების. მაგრამ ისინი კარგ შედეგს იძლევიან ისეთი ამოცანების გადაწყვეტისას, სადაც შეუძლებელია სწავლების ხანგრძლივი პროცედურების ჩატარება. გარდა ამისა, მათ აქვთ სხვა საინტერესო თვისებები, რასაც ქვემოთ განვიხილავთ.

შემხვედრი გავრცელების ნეიროქსელებში გაერთიანებულია ორი ცნობილი ალგორითმი: კოჰონენის თვითორგანიზებადი რუკა და გროსბერგის ვარსკვლავი. მათი გაერთიანება გვაძლევს ისეთ თვისებებს, რომლებიც ცალ-ცალკე არც ერთს არა აქვს. შემხვედრი გავრცელების ნეიროქსელებს აქვს განზოგადების უნარი. სწავლების პროცესში შემავალი ვექტორები ასოცირდება შესაბამის გამომავალ ვექტორებთან. ვექტორები შეიძლება იყოს ბინარული ან უწყვეტი.

სწავლების პროცესის დამთავრების შემდეგ შემავალი ვექტორის წარდგენა გვაძლევს საჭირო გამომავალ ვექტორს. შემხვედრი გავრცელების ნეიროქსელების განზოგადების უნარი იძლევა იმის შესაძლებლობას, რომ მივიღებთ სწორ პასუხს მაშინაც

კი, როდესაც კსელს წარედგინება მცირედ დამახინჯებული ან არასრული შემავალი ვექტორი. ასეთი თვისებების გამო, შემხველარი გაერცელების ნეიროქსელებს იყენებენ სიგნალების გასაძლიერებლად, სახეთა ამოცნობისათვის, აღდგენისათვის და სხე.



სურ. 2.17. შემხველარი გაერცელების ნეიროქსელის გამარტივებული სტრუქტურა.

ქსელის ფუნქციონირების განხილვის გასაადვილებლად 2.17 სურათზე ნაჩვენებია შემხველარი გაერცელების პირდაპირი მოქმედების ნეიროქსელის გამარტივებული სტრუქტურა.

სურათზე ნაჩვენებია შემავალი (ნულოვანი) შრის ნეირონები გამოთვლებს არ ასრულებს, არამედ

წარმოადგენს განშტოების წერტილებს. ნულოვანი შრის თითოეული ნეირონი უკავშირდება პირველი შრის (კოჰონენის შრის) თითოეულ ნეირონს ცალკეული w_{00} სინაფსური წონით. ეს სინაფსური წონები განიხილება, როგორც W წონების მატრიცა. ანალოგიურად, კოჰონენის შრის თითოეული ნეირონი უკავშირდება მეორე შრის (გროსბერგის შრის) თითოეულ ნეირონს ცალკეული v_{00} სინაფსური წონით. ეს სინაფსური წონები შეადგენს V წონების მატრიცას. ყოველივე ეს ზემოთ განხილული ნეიროქსელების მსგავსია, თუმცა განსხვავება მდგომარეობს იმ ოპერაციებში, რასაც ასრულებენ კოჰონენის და გროსბერგის ნეირონები. შემხვედრი გავრცელების ნეიროქსელი, ისევე როგორც მრავალი სხვა ნეიროქსელები ფუნქციონირებს ორ რეჟიმში. პირველი რეჟიმი არის ნორმალური (მუშა) რეჟიმი, როდესაც ქსელს წარედგინება შემავალი ვექტორი X და მიიღება გამომავალი ვექტორი Y ; მეორე არის სწავლების რეჟიმი, როდესაც ქსელს წარედგინება შემავალი ვექტორი X და ხდება სინაფსური წონების ისეთი კორექტირება, რომ მივიღოთ სასურველი T გამომავალი ვექტორი.

2.13.1. შემხვედრი გავრცელების ნეიროქსელების ნორმალური ფუნქციონირების რეჟიმი

შემხვედრი გავრცელების ნეიროქსელების პირველი შრე (კოჰონენის შრე) ფუნქციონირებს ისე, რომ მისი მხოლოდ და მხოლოდ ერთი ნეირონის გამოსასვლელზე შეიძლება იყოს ლოგიკური ერთიანი მოცემული შემავალი ვექტორისათვის, ყველა დანარჩენი ნეირონი იძლევა ლოგიკურ ნულს (ამბობენ, რომ შრეში არის ერთი გამარჯვებული ნეირონი). ქსელის თითოეულ შესასვლელს კოჰონენის ნეირონი უერთდება მასთან დაკავშირებული სინაფსური წონების საშუალებით.

მაგალითად, 2.17 სურათზე ნაჩვენებია K_1 კოჰონენის ნეირონი, მისი სინაფსური წონები $w_{11}, w_{21}, \dots, w_{m1}$ შეადგენს W_1 წონების ვექტორს და ნულოვანი შრის მეშვეობით უკავშირდება X შემავალი ვექტორის შემადგენელ $x_1, x_2, \dots, x_m$ სიგნალებს. კოჰონენის ნეირონის NET გამოსასვლელი გამოითვლება ჩვეულებრივად

$$NET_j = w_{1j}x_1 + w_{2j}x_2 + \dots + w_{mj}x_m, \quad (2.20)$$

სადაც NET_j არის კოჰონენის j -ური ნეირონის NET გამოსასვლელი

$$NET_j = \sum_i x_i w_{ij}. \quad (2.21)$$

ვექტორული ფორმით იქნება

$$N = XW, \quad (2.22)$$

სადაც N არის კოჰონენის შრის NET გამომავალი ვექტორი. კოჰონენის შრის ის ნეირონი არის გამარჯვებული, რომელსაც აქვს NET-ის მაქსიმალური მნიშვნელობა. გამარჯვებული ნეირონის

გამოსასვლელზე იქნება ლოგიკური ერთიანი, ხოლო კოჰონენის შრის ყველა დანარჩენი ნეირონის გამოსასვლელზე იქნება ლოგიკური ნული.

გროსბერგის შრე ფუნქციონირებს მსგავსი წესით. მისი გამოსასვლელი NET არის კოჰონენის შრის $k_1, k_2, \dots, k_n$ გამოსასვლელების, რომლებიც შეადგენს K ვექტორს, შეწონილი ჯამი. სინაფსური წონები $v_{11}, v_{21}, \dots, v_{np}$ შეადგენს V ვექტორს, მაშინ გროსბერგის ნეირონის NET გამოსასვლელი გამოითვლება შემდეგნაირად:

$$NET_j = \sum_i k_i w_{ij}, \quad (2.23)$$

სადაც NET_j არის გროსბერგის j -ური ნეირონის NET გამოსასვლელი. იგი ვექტორული ფორმით ასე ჩაიწერება:

$$Y = KV, \quad (2.24)$$

სადაც Y არის გროსბერგის შრის გამომავალი ვექტორი; K კოჰონენის შრის გამომავალი ვექტორი;

V - გროსბერგის შრის სინაფსური წონების მატრიცა. იმის გამო, რომ კოჰონენის შრე ფუნქციონირებს ისე, რომ მოცემული შემავალი ექტორისათვის მხოლოდ ერთი ნეირონის გამოსასვლელზე შეიძლება იყოს ლოგიკური ერთიანი, ხოლო ყველა დანარჩენი ნეირონი იძლევა ლოგიკურ ნულს, ცხადია, K ექტორის მხოლოდ ერთადერთი ელემენტია ერთის ტოლი და ამიტომ გამოთვლები ძალზე მარტივდება. ფაქტობრივად გროსბერგის შრის თითოეული ნეირონი გვაძლევს იმ სინაფსურ წონას, რომელიც აკავშირებს ამ ნეირონს კოჰონენის შრის არანულოვანი სიგნალის მომცემ ერთადერთ ნეირონთან.

2.13.2. კოჰონენის შრის სწავლება

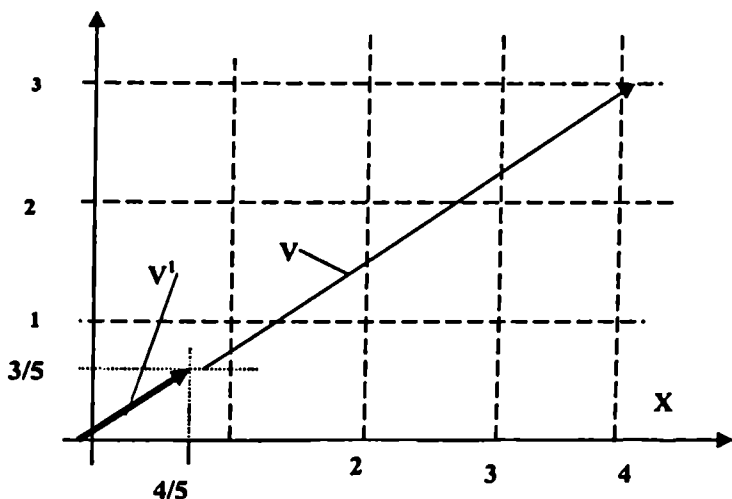
შემხვედრი გავრცელების ნეიროქსელებში კოჰონენის შრე ახდენს შემავალი ექტორების კლასიფიკაციას მსგავს ჯგუფებად, რისთვისაც სინაფსური წონების ისეთი შეთანაწყობა მიიღწევა, რომ მსგავსი შემავალი ექტორები იწვევს კოჰონენის შრის ერთი და იმავე ნეირონის აქტივიზაციას. ამის შემდეგ, გროსბერგის შრის ამოცანა არის სასურველი გამომავალი სიგნალების ფორმირება. კოჰონენის შრე არის თვითსწავლებადი ანუ სწავლების მეთოდია მაწავლებლის გარეშე. ამიტომ წინასწარ არ არის ცნობილი, თუ რომელი ნეირონის აქტივიზაცია მოხდება, ამის აუცილებლობა არც არის. აუცილებელია მხოლოდ იმის უზრუნველყოფა, რომ სწავლების შედეგად ქსელმა შეიძლოს მსგავსი ექტორების ერთად დაჯგუფება, ანუ არამსგავსი ექტორების განმხოლოება.

მიზანშეწონილია შემხვედრი გავრცელების ნეიროქსელებისათვის წარსადგენი ექტორების წინასწარი ნორმალიზება. ამისათვის წარსადგენი ექტორის თითოეული კომპონენტი უნდა გაიყოს ექტორის

სიგრძეზე. ვექტორის სიგრძე არის კვადრატული ფესვი ვექტორის კომპონენტების კვადრატების ჯამიდან, რომელიც აღგებრული ფორმით ასე გამოისახება:

$$x'_i = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}. \quad (2.25)$$

ამით წარსადგენი ვექტორი გარდაიქმნება იგივე მიმართულების ერთეულოვან ვექტორად, ანუ ერთეულოვანი სიგრძის ვექტორად n -გაზომილების სივრცეში.



სურ. 2.18. წარსადგენი ვექტორის ნორმალიზების მაგალითი.

(2.25) განტოლება განაზოგადებს ორი განზომილების შემთხვევას, სადაც ვექტორის სიგრძე გამოითვლება პითაგორას თეორემით. 2.18 სურათზე ნაჩვენებია ორგანზომილებიანი V ვექტორი x - y კოორდინატებით, სადაც $x = 4$ და $y = 3$.

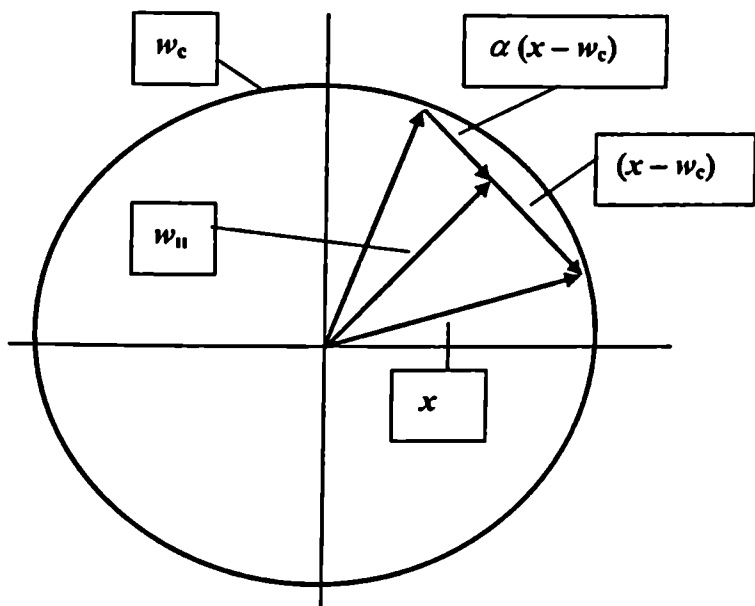
ნორმალიზების შედეგად, (2.25) განტოლების თანახმად მივიღებთ ახალ V' ვექტორს კოორდინატებით $x=4/5$ და $y=3/5$, რომელსაც ექნება იგივე მიმართულება, მაგრამ ერთეულოვანი სიგრძე. სამგანზომილებიანი ვექტორის შემთხვევაში, ის დამთავრდებოდა ერთეულოვანი სფეროს ზედაპირზე. ეს შეიძლება განვაზოგადოთ ნებისმიერი N განზომილების ვექტორებზე, მაშინ ისინი შემოსასღვრავენ ერთეულოვან ჰიპერსფეროს.

კოჰონენის შრის სწავლების დროს შესასვლელს წარედგინება შემავალი ვექტორი და გამოითვლება მისი სკალარული ნამრავლები იმ სინაფსურ წონებთან, რომლებიც უკავშირდება კოჰონენის შრის ყველა ნეირონს. სკალარული ნამრავლის მაქსიმალური მნიშვნელობის მქონე ერთადერთი ნეირონი იქნება შრეში გამარჯვებული და მოხდება მისი წონების შეთანაწყობა. იმის გამო, რომ NET სიდიდეების გამოსათვლელი სკალარული ნამრავლი წარმოადგენს მსგავსების ზომას შემავალ ვექტორსა და წონების ვექტორს შორის, ამიტომ სწავლების პროცესი მდგომარეობს კოჰონენის შრის ისეთი ნეირონის შერჩევაში, რომლის წონების ვექტორი ყველაზე უფრო ახლოს არის შემავალ ვექტორთან და ამ წონების ვექტორის შემდგომ მიახლოებაში შემავალ ვექტორთან. მიმდინარეობს სწავლება მასწავლებლის გარეშე, ანუ მიმდინარეობს ქსელის თვითორგანიზების ისეთი პროცესი, რომ კოჰონენის შრის მოცემული ნეირონის გამოსასვლელი იღებს მოცემული შემავალი ვექტორისათვის მაქსიმალურ მნიშვნელობას. სწავლების პროცესი აღიწერება შემდეგი განტოლებით:

$$w_{ii} = w_c + \alpha (x - w_c), \quad (2.26)$$

სადაც w_{ii} არის იმ წონის ახალი მნიშვნელობა, რომელიც აკავშირებს x შემავალ ელემენტს გამარჯვებულ ნეირონთან; w_c - აღნიშნული წონის მნიშვნელობა წინა ბიჯზე;

α - სწავლების სიჩქარის კოეფიციენტი, რომლის ცვლა შესაძლებელია სწავლების პროცესში.



სურ. 2.19. წონის ვექტორის მოძრაობა სწავლების პროცესში.

თითოეული წონა, რომელიც უკავშირდება კოჰონენის შრის გამარჯვებულ ნეირონს, იცვლება მისი მნიშვნელობისა და მასთან დაკავშირებული შესასვლელის მნიშვნელობას შორის სხვაობის $(x - w_c)$ პროპორციულად. ცვლილების მიმართულება ისეთია, რომ მცირდება ეს სხვაობა. 2.19 სურათზე მოცემულია ამ პროცესის ილუსტრაცია ორგანზომილებიანი ვექტორებისათვის. სურათზე W_{II} არის ახალი სინაფსური წონის ვექტორი, ხოლო W_c - ძველი სინაფსური წონის ვექტორი. ჯერ ეპოულობთ $X - W_c$ ვექტორს, რისთვისაც გავაგებთ მონაკვეთს ძველი სინაფსური წონის

ვექტორის ბოლოდან X ვექტორის ბოლოსკენ; მიღებულ ვექტორს დაეამოკლებთ ერთზე ნაკლები სიდიდის α სწავლების სინქარის კოეფიციენტზე გამრავლებით – შედეგად ვიღებთ ცვლილების δ ვექტორს. მაშინ ახალი სინაფსური წონის ვექტორი W_{α} იქნება მონაკვეთი მიმართული კოორდინატთა სათავედან δ ვექტორის ბოლოსაკენ. ამ ილუსტრაციიდან ნათლად ჩანს, რომ სწავლების პროცესის ეფექტი შეიძლება წარმოვიდგინოთ როგორც სინაფსური წონის ვექტორის მოძრაობა მისი სიგრძის უმნიშვნელო ცვლილებით შემავალი ვექტორის მიმართულებით. α სწავლების სინქარის კოეფიციენტს სწავლების დასაწყისში ჩვეულებრივ იღებენ 0,7 სიდიდეს, რომელიც სწავლების პროცესში შეიძლება თანდათან შემცირდეს, ანუ პროცესის დასაწყისში კეთდება მოზრდილი ნაბიჯები, რომლებიც საბოლოო სიდიდესთან მიახლოებისას თანდათან მცირდება.

2.13.3. წონითი ვექტორების საწყისი მნიშვნელობების შერჩევა

შემხვედრი გავრცელების ქსელის ყველა სინაფსურ წონას სწავლების პროცესის დაწყების წინ უნდა მიენიჭოს გარკვეული საწყისი მნიშვნელობა.

ნეირონული ქსელების სწავლების პრაქტიკაში ზოგადად, მიღებულია სინაფსური წონების საწყის მნიშვნელობებად მცირე შემთხვევითი სიდიდეების გამოყენება. კოჰონენის შრის სწავლების დაწყებამდე საჭიროა შერჩეული წონითი ვექტორების ნორმალიზება. სწავლების პროცესის დამთავრების შემდეგ წონითი ვექტორების საბოლოო მნიშვნელობები თანხვედრილი იქნება შემავალ ნორმალიზებულ ვექტორებთან. ამიტომ ნორმალიზება სწავლების დაწყებამდე ახდენს წონითი ვექტორების მიახლოებას მათ საბოლოო მნიშვნელობებთან და შედეგად ამოკლებს სწავლების

პროცესს. კოჰონენის შრის სწავლებისათვის სინაფსური წონების საწყის მნიშვნელობებად შემთხვევითი სიდიდეების გამოყენებამ (რანდომიზაციამ) შეიძლება სწავლებისას წარმოიქმნას პრობლემები, რადგანაც წონითი ვექტორები ჰიპერსფეროს ზედაპირზე ლაგდება თანაბრად, ხოლო შემავალი ვექტორები, როგორც წესი, განაწილებულია არათანაბრად და ჯგუფდება ჰიპერსფეროს ზედაპირის შედარებით პატარა უბნებზე. შედეგად, წონითი ვექტორების დიდი ნაწილი იმდენად იქნება დაშორებული ნებისმიერი შემავალი ვექტორიდან, რომ ისინი ვერ მოგვცემენ საუკეთესო შესაბამისობას. კოჰონენის შრის ეს ნეირონები უსარგებლონი აღმოჩნდებიან, ვინაიდან მათი გამოსასვლელები ეოველთვის ნულის ტოლი იქნება. შეიძლება წარმოიქმნას უარესი პრობლემაც: დარჩენილი წონითი ვექტორების რაოდენობა არასაკმარისი იყოს იმისათვის, რომ კლასებად იქნეს დაყოფილი ის შემავალი ვექტორები, რომლებიც ერთანეთთან ახლოს არიან განლაგებული ჰიპერსფეროს ზედაპირზე.

დამუშავებულია ამ პრობლემებთან ბრძოლის რამდენიმე მეტ-ნაკლებად ეფექტური მეთოდი. ერთ-ერთი მათგანია ამოზნექილი კომბინაციის მეთოდი (convex combination method), რომელიც წონებს ანიჭებს ერთნაირ საწყის მნიშვნელობებს

$$w_i = \frac{1}{\sqrt{n}}, \quad (2.27)$$

სადაც n არის თითოეული წონითი ვექტორის ელემენტების რაოდენობა. ამის გამო ყველა წონითი ვექტორი თანხვედრილია და აქვს ერთეულოვანი სიგრძე. შემავალი ვექტორის თითოეულ ელემენტს მიენიჭება შემდეგი მნიშვნელობა:

$$x_i = \alpha x_i + \frac{1 - \alpha}{\sqrt{n}}, \quad (2.28)$$

სადაც n არის შესასვლელების რაოდენობა. თავდაპირველად α ძალზე მცირე სიდიდეა, რის გამოც შემავალ ვექტორებს აქვს $1/\sqrt{n}$ -თან მიახლოებული სიგრძე და თითქმის ემთხვევა წონით ვექტორებს. ქსელის სწავლების პროცესში α თანდათანობით იზრდება და უახლოვდება ერთს. ეს იძლევა შემავალი ვექტორების განმსოლოების შესაძლებლობას და საბოლოოდ მათ ენიჭება ჭეშმარიტი მნიშვნელობები. ამოზნექილი კომბინაციის მეთოდი კარგია, მაგრამ ანელებს სწავლების პროცესს.

შემხვედრი გაერცელების ქსელებში სინაფსური წონებისათვის სწავლების პროცესის დაწყების წინ საწყისი მნიშვნელობების მინიჭების პრობლემა დამატებით შესწავლას მოითხოვს.

კოჰონენის სწავლების მეთოდს აქვს შემავალი მონაცემების სიმრავლიდან სტატისტიკური თვისებების გამოყოფის სასარგებლო უნარი.

2.13.4. გროსბერგის შრის სწავლება

გროსბერგის შრის სწავლება შედარებით მარტივია. სწავლების მეთოდია მასწავლებლით. გროსბერგის შრისათვის შემავალი ვექტორი არის კოჰონენის შრის გამოსასვლელი, რომელიც წარედგინება გროსბერგის შრის ნეირონებს და მათი გამოსასვლელები გამოითვლება ისევე, როგორც ნორმალური ფუნქციონირების რეჟიმში. ამის შემდეგ, თითოეული წონის კორექტირება ხდება მხოლოდ იმ შემთხვევაში, თუ ის უკავშირდება არანულოვანი გამოსასვლელის მქონე კოჰონენის შრის ნეირონს. წონის კორექციის სიდიდე პროპორციულია წონასა და გროსბერგის შრის იმ ნეირონის გამოსასვლელს შორის სხვაობისა, რომელსაც იგი უკავშირდება:

$$v_{ijn} = v_{ije} + \beta (y_j - v_{ije})k_i, \quad (2.29)$$

სადაც k_i არის i -ური კოპონენის ნეირონის გამოსასვლელი (ნულისაგან განსხვავდება მხოლოდ ერთი კოპონენის ნეირონისათვის); y_j არის სასურველი გამოსასვლელი ვექტორის j -ური ელემენტი. თავდაპირველად β -ს მნიშვნელობას ადებენ $\sim 0,1$ -ის ტოლს და სწავლების პროცესში იგი თანდათან მცირდება.

შემხვედრი გავრცელების ნეირონული ქსელების შემქმნელი რ. ჰეხტ-ნილსენი [29] თვითონ აღნიშნავს, რომ ამ ქსელებს აქვს შეზღუდული შესაძლებლობები უკუგავრცელების ქსელებთან შედარებით, თუმცა კარგ სტატისტიკურ მოდელებს იძლევა გარკვეულ ამოცანებში. შემხვედრი გავრცელების ქსელის სწავლება შედარებით სწრაფად ხდება და სწორად გამოყენებისას იგი ზოგავს მნიშვნელოვან მანქანურ დროს. ეს ქსელები სასარგებლოა განსაკუთრებით მაშინ, როდესაც საჭიროა სისტემების სწრაფი მოდელირება და ნაკლებ მნიშვნელოვანია დიდი სიზუსტე.

თავი III. ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკური მეთოდები

სტოქასტიკური მეთოდების გამოყენება კარგ შედეგებს იძლევა როგორც ხელოვნური ნეირონული ქსელების სწავლების, ასევე “ნასწავლი” ქსელებისაგან სასურველი გამოსასვლელი სიგნალების მიღების, ანუ ნორმალური ფუნქციონირების რეჟიმის დროს. სწავლების სტოქასტიკური მეთოდები შესაძლებლობას იძლევა სწავლების პროცესში დაეძლიოთ ლოკალური მინიმუმების პრობლემა, თუმცა ამ მეთოდების გამოყენება დაკავშირებულია ზოგიერთ სირთულეებთან.

ეს თავი ეთმობა ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკურ მეთოდებს, ხოლო სტოქასტიკური მეთოდების გამოყენება უკვე ნასწავლი ქსელებისაგან სასურველი გამოსასვლელი სიგნალების მისაღებად, განხილული იქნება მომდევნო თავში.

3.1. სწავლების ლოკალური მინიმუმების პრობლემა

ხელოვნური ნეირონული ქსელების სწავლება წარმოადგენს მისი პარამეტრების (როგორც წესი, სინაფსური წონების) მოდიფიცირების გარკვეულ პროცესს. იმ შემთხვევაში, თუ სწავლების პროცესი წარმატებული იყო, მაშინ ქსელისათვის შემავალი სიგნალების სიმრავლის წარდგენის შედეგად ქსელის გამოსასვლელზე უნდა გამოიმუშავდეს სასურველი სიგნალების სიმრავლე.

ცნობილია, რომ არსებობს სწავლების მეთოდების ორი ჯგუფი: დეტერმინირებული და სტოქასტიკური.

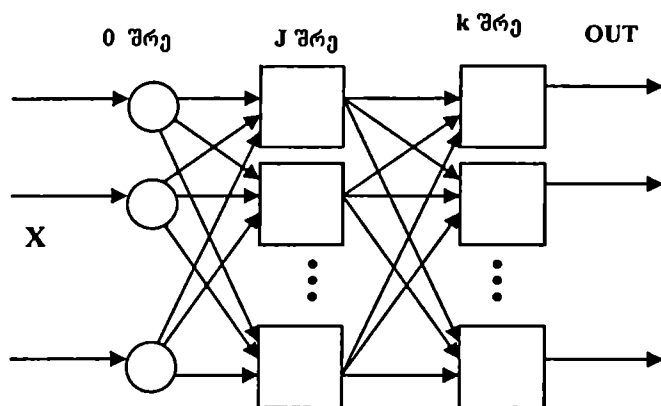
სწავლების დეტერმინირებული მეთოდები სინაფსური წონების მოდიფიცირების პროცედურას ახორციელებს ბიჯურად, ამასთან იყენებს წონების მიმდინარე მნიშვნელობებს, აგრეთვე შემავალი სიგნალების, სასურველი გამომავალი სიგნალების და ფაქტობრივი

გამომავალი სიგნალების მნიშვნელობებს. ასეთი დეტერმინირებული მიდგომის მაგალითი არის პერსპექტივის სწავლების პროცესები, რომლებიც განხილული იყო პ. 2.9 -ში.

სწავლების სტოქასტიკური მეთოდები ახორციელებს ქსელის სინაფსური წონების ფსევდოშემთხვევით ცვლილებებს; ამასთან, ინარჩუნებს ისეთ ცვლილებებს, რომლებიც ერთმანეთთან აახლოებს ფაქტობრივი გამომავალი სიგნალებისა და სასურველი გამომავალი სიგნალების მნიშვნელობებს. იმისათვის, რომ ვნახოთ, თუ როგორ ხორციელდება ეს პროცედურები, განვიხილოთ ტიპური ორშრიანი ხელოვნური ნეირონული ქსელი, რომელიც 3.1 სურათზე არის ნაჩვენები. იგულისხმება, რომ ქსელის ნეირონების აქტივაციის ფუნქცია არის არაწრფივი. ნეირონული ქსელის სწავლებისათვის შეიძლება გამოყენებულ იქნეს შემდეგი პროცედურა:

1. შევარჩიოთ სინაფსური წონის შემთხვევითი სიდიდე და მოვახდინოთ მისი კორექცია მცირე შემთხვევითი სიდიდით. წარვუდგინოთ ქსელს შემავალი სიგნალების სიმრავლე და გამოვითვალოთ გამომავალი სიგნალების ფაქტობრივი მნიშვნელობები.
2. შევადაროთ გამომავალი სიგნალების ფაქტობრივი მნიშვნელობები სასურველი სიგნალების მნიშვნელობებს და გამოვითვალოთ მათ შორის სხვაობა. ეს პროცედურა უნდა შესრულდეს მასწავლებელი წყვილის თითოეული ელემენტისათვის, შემდეგ ელემენტების სხვაობები აიყვანება კვადრატში და გამოითვლება მათი ჯამი. სწავლების მიზანი არის ამ სხვაობის მინიმიზაცია, რასაც ხშირად უწოდებენ მიზნობრივ ფუნქციას.
3. შევარჩიოთ სინაფსური წონის შემთხვევითი სიდიდე და მოვახდინოთ მისი კორექცია მცირე შემთხვევითი სიდიდით. თუ კორექცია სასარგებლოა, ანუ ამცირებს მიზნობრივ ფუნქციას, იგი უნდა იქნეს დამახსოვ-

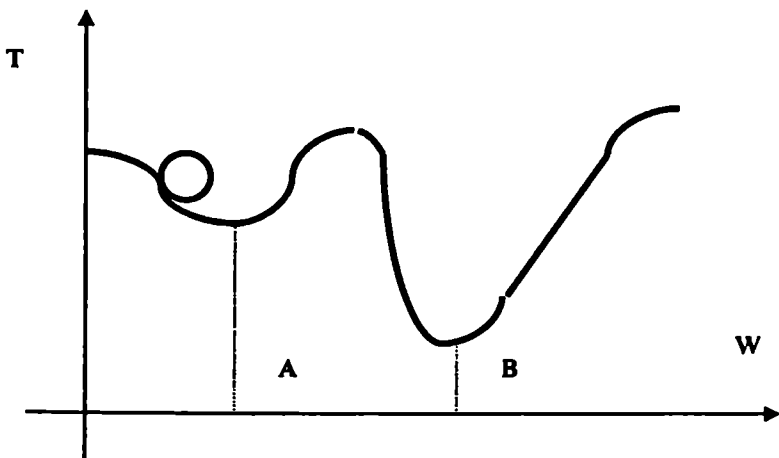
რებული, წინააღმდეგ შემთხვევაში უნდა დაეუბრუნდეთ სინაფსური წონის ძველ მნიშვნელობას.
 4. გაეიმეოროთ 1, 2, 3 ბიჯები მანამ, სანამ ნეირონული ქსელი საკმარისად ნასწავლი არ გახდება.



სურ. 3.1. ტიპური ორშრიანი ხელოვნური ნეირონული ქსელი.

აღწერილი პროცესი მიისწრაფვის მიზნობრივი ფუნქციის მინიმიზაციისაკენ, მაგრამ შეიძლება მოხდეს ლოკალურ მინიმუმში, როგორც მახეში, რადგანაც ქსელი ჩათვლის, რომ გადაწყვეტილება მიღებულია. თუმცა, ეს გადაწყვეტილება ფაქტობრივად წარუმატებელია. ლოკალური მინიმუმების პრობლემა ერთი სინაფსური წონის კორექციის შემთხვევისათვის ნაჩვენებია 3.1 სურათზე.

ვთქვათ, A წერტილი შეესაბამება სინაფსური წონის საწყის მნიშვნელობას, თუ წონის ცვლილებები მცირე სიდიდის იქნება, ნებისმიერი გადახრა A წერტილიდან გაზრდის T მიზნობრივი ფუნქციის მნიშვნელობას და



სურ. 32. ლოკალური მინიმუმების პრობლემა.

ამიტომ უარყოფილი იქნება, ხოლო სინაფსური წონის საუკეთესო მნიშვნელობა, რომელსაც B წერტილი შეესაბამება, ვერასოდეს ვერ იქნება მიღწეული. სისტემა A წერტილის ლოკალური მინიმუმის მახეში დარჩება, ნაცვლად B წერტილის გლობალური მინიმუმისა. იმ შემთხვევაში, თუ წონის შემთხვევითი ცვლილებები დიდი სიდიდის იქნება, წონა განიცდის იმდენად მკვეთრ ცვლილებებს, რომ სასურველ (გლობალურ) მინიმუმში მისი მოხვედრა პრაქტიკულად შეუძლებელი იქნება.

ამგვარი პრობლემების თავიდან აცილების სასარგებლო სტრატეგია მდგომარეობს დიდი საწყისი ბიჯებიდან საშუალო შემთხვევითი ბიჯის ზომის თანდათანობით შემცირებაში. ეს იძლევა ლოკალური მინიმუმის მახიდან თავის დაღწევისა და ამავე დროს, ქსელის საბოლოო სტაბილიზაციის შესაძლებლობას. ლოკალური მინიმუმის მახეები ემუქრება ყველა იმ ალგორითმს, რომლებიც აგებულია მინიმუმის ძიების პრინციპზე. სწავლების სტოქასტიკური მეთოდები კი ამ პრობლემის დაძლევის შესაძლებლობას იძლევა, ეინაიდან წონების კორექციის ისეთი სტრატეგია შეიძლება

მოინახოს, რომელიც აიძულებს წონებს მიიღონ გლობალური მინიმუმის მნიშვნელობა (B წერტილი).

მიღებულია, რომ გლობალური მინიმუმის ძიება წააგავს ფოლადის წრთობის პროცესს, ამიტომ ხშირად იყენებენ ტერმინს – წრთობის იმიტაცია.

გამდნარ ლითონში ატომები იმყოფებიან ძლიერ უწყესრიგო მოძრაობაში, ისინი ისწრაფვიან ენერგიის მინიმუმის მდგომარეობისაკენ (ამ შემთხვევაში ერთიანი კრისტალისაკენ), მაგრამ ამას მაღალ ტემპერატურებზე ეწინააღმდეგება ატომური მოძრაობების ენერგია. ლითონის თანდათანობითი გაცივების პროცესში წარმოიშობა სულ უფრო დაბალენერგეტიკული მდგომარეობები, ვიდრე არ იქნება მიღწეული უმცირესი ენერგეტიკული მდგომარეობა, ანუ გლობალური მინიმუმის მნიშვნელობა. ენერგეტიკული მდგომარეობების $P(e)$ განაწილებას აღწერს შემდეგი გამოსახულება:

$$P(e) = \exp(-e/kT), \quad (3.1)$$

სადაც $P(e)$ არის იმის ალბათობა, რომ სისტემა იმყოფება e ენერგიის მდგომარეობაში; k - ბოლცმანის მუდმივა; T – ტემპერატურა კელვინის სკალით.

მაღალ ტემპერატურებზე $P(e)$ უახლოვდება ერთის სიდიდეს ყველა ენერგეტიკული მდგომარეობისათვის. ტემპერატურის შემცირებასთან ერთად მაღალ-ენერგეტიკული მდგომარეობების ალბათობა მცირდება, ხოლო ნულთან მიახლოებისას იმის ალბათობა, რომ სისტემა იმყოფება მაღალენერგეტიკულ მდგომარეობაში, ძალზე უმნიშვნელო სიდიდე ხდება.

3.2. სწავლების ბოლცმანის მეთოდი

ბოლცმანის სტოქასტიკური მეთოდით ხელოვნური ნეირონული ქსელების სწავლებისათვის უნდა ჩატარდეს შემდეგი პროცედურები:

1. განისაზღვროს T ცვლადის მნიშვნელობა (ე.წ. ხელოვნური ტემპერატურა), თავდაპირველად მას ანიჭებენ დიდ რიცხვით მნიშვნელობას;

2. წარედგინოს ქსელს შესასვლელი სიმრავლე და გამოითვალოს გამოსასვლელები და მიზნობრივი ფუნქცია;

3. შემთხვევითი სიდიდით შეიცვალოს წონითი კოეფიციენტი და განისაზღვროს შესაბამისი ქსელის გამოსასვლელი და მიზნობრივი ფუნქციის ცვლილება, რომელიც გამოწვეულია წონითი კოეფიციენტის ცვლილებით;

4. მიზნობრივი ფუნქციის შემცირების შემთხვევაში წონითი კოეფიციენტის ცვლილებას ვიმახსოვრებთ, ხოლო, თუ წონითი კოეფიციენტის ცვლილება იწვევს მიზნობრივი ფუნქციის გაზრდას, მაშინ ამ ცვლილების შენახვის ალბათობა გამოითვლება ბოლცმანის განაწილების მიხედვით:

$$P(c) = \exp(-c/kT), \quad (3.2)$$

სადაც $P(c)$ არის მიზნობრივ ფუნქციაში c ცვლილების ალბათობა; k – ბოლცმანის მუდმივას ანალოგიური მუდმივა, რომელიც ამოცანის შესაბამისად შეირჩევა; T - ხელოვნური ტემპერატურა.

შეირჩევა რაიმე r შემთხვევითი რიცხვი ნულიდან ერთამდე, თუ $P(c)$ არის r -ზე მეტი, მაშინ ცვლილება შეინახება, წინააღმდეგ შემთხვევაში წონითი კოეფიციენტს უნდა დაუბრუნდეს წინა მნიშვნელობა. ეს საშუალებას აძლევს სისტემას გააკეთოს შემთხვევითი ბიჯი მიზნობრივი ფუნქციის გაუარესებისაკენ, რითაც

იგი ამოდის ლოკალური მინიმუმიდან, სადაც ნებისმიერი მცირე ბიჯი ზრდის მიზნობრივ ფუნქციას.

ბოლცმანის სტოქასტიკური მეთოდით ხელოვნური ნეირონული ქსელის სწავლების დასასრულებლად საჭიროა 3 და 4 ბიჯის გამეორება ქსელის თითოეული წონისათვის. ამასთან თანდათანობით მცირდება ტემპერატურა, ვიდრე არ მიიღწევა მიზნობრივი ფუნქციის დასაშვები დაბალი დონე. ამ მომენტში ქსელს წარედგინება მომდევნო შემავალი ვექტორი და სწავლების პროცესი მეორდება. ეს პროცედურები მეორდება სასწავლო სიმრავლის ყველა ვექტორისათვის, ვიდრე მიზნობრივი ფუნქცია ყველასათვის არ გახდება დასაშვები სიდიდის. მესამე ბიჯზე წონითი კოეფიციენტის ცვლილების შემთხვევითი სიდიდე შეიძლება განისაზღვროს სხვადასხვა მეთოდით, მაგალითად, გაუსის განაწილების შესაბამისად.

დადგენილია, რომ ტემპერატურის შემცირების სიჩქარე უნდა იყოს დროის ლოგარითმის უკუპროპორციული, რათა მიღწეულ იქნეს კრებადობა გლობალურ მინიმუმთან. გაცივების სიჩქარე იქნება

$$T(t) = \frac{T_0}{\log(1+t)}, \quad (3.3)$$

სადაც $T(t)$ არის ხელოვნური ტემპერატურა როგორც დროის ფუნქცია; T_0 - საწყისი ხელოვნური ტემპერატურა; t - ხელოვნური დრო.

(3.3) განტოლებიდან ჩანს, რომ მოსალოდნელია გაცივების დაბალი სიჩქარე, რაც დადასტურდა ექსპერიმენტულად. ბოლცმანის სტოქასტიკური მეთოდით ხელოვნური ნეირონული ქსელების სწავლება ხშირად ითხოვს მანქანური დროის ძალზე დიდ რესურსს.

3.3. სწავლების კოშის მეთოდი

ბოლცმანის სტოქასტიკური მეთოდის განვითარებას წარმოადგენს სწავლების კოშის მეთოდი. ამ მეთოდში ბოლცმანის განაწილება შეცვლილი არის კოშის განაწილებით, რომელსაც აქვს უსასრულო დისპერსია, ეს კი ზრდის დიდი ბიჯების ალბათობას. აღნიშნული მარტივი ცვლილების შედეგად ტემპერატურის შემცირების მაქსიმალური სინქარე იქნება დროის წრფივი სიდიდის უკუპროპორციული და არა დროის ლოგარითმის უკუპროპორციული, როგორც ეს იყო ბოლცმანის სტოქასტიკური მეთოდით სწავლებისას

$$T(t) = \frac{T_0}{1+t} \quad (3.4)$$

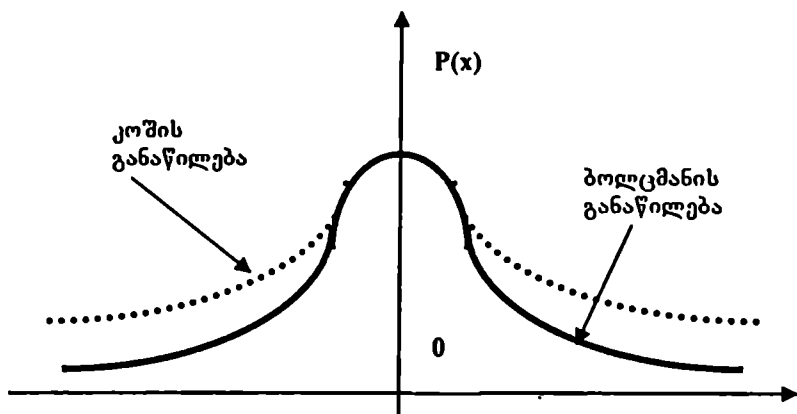
ეს გარემოება მკვეთრად ამცირებს ქსელის სწავლების დროს. კოშის განაწილებას აქვს შემდეგი სახე:

$$P(x) = \frac{T(t)}{T(t)^2 + x^2}, \quad (3.5)$$

სადაც $P(x)$ არის x სიდიდის ბიჯის ალბათობა.

ბოლცმანის განაწილება და კოშის განაწილება ნაჩვენებია 3.3 სურათზე.

მიუხედავად იმისა, რომ ბოლცმანის მეთოდთან შედარებით კოშის მეთოდი მკვეთრად ამცირებს ქსელის სწავლების დროს, ეს მანევრებელი შეიძლება მაინც დიდი აღმოჩნდეს. ქსელის სწავლების ხანგრძლივობის შემდგომი შემცირება შესაძლებელია ე.წ. ხელოვნური თბოტევადობის მეთოდით, რომელიც რთულია და ნაკლებად გამოიყენება.



სურ. 3.3. ბოლცმანისა და კოშის განაწილება.

3.4. ხელოვნური ნეირონული ქსელების სინთეზის პროცედურები

ზემოთ აღნიშნული იყო, რომ ხელოვნური ნეირონული ქსელები ძნელად ფორმალისებადი ამოცანების ეფექტურად გადაწყვეტის საშუალებაა. ამოცანების ამ კლასს განეკუთვნება კლასტერიზების, სახეთა ამოცნობის, ასოციაციური მეხსიერების, ფუნქციათა აპროქსიმაციის, პროგრესირების, ოპტიმიზაციის, ფილტრაციის, ინფორმაციის შეკუმშვის, დინამიკური სისტემების იდენტიფიკაციისა და მართვის მრავალფეროვანი პრობლემა [2,16,17,20,21,24,38].

ნეიროქსელური ტექნოლოგიით ძნელად ფორმალისებადი ამოცანების გადაწყვეტა ნიშნავს ამ ამოცანების ადეკვატური ხელოვნური ნეირონული ქსელების სინთეზის რეალისებას [31,32,36,37,42].

ნამოყაალებით ნეირონული ქსელის სინთეზის პროცედურები და მათი თანამიმდევრობა კლასტერიზების

ამოცანის გადაწყვეტის მაგალითზე, ვინაიდან კლასტერირება წარმოადგენს ფუძემდებლურ პროცედურას ძნელად ფორმალისებადი ამოცანების ისეთი ძირითადი პროცესების მოდელირებისათვის, როგორიც არის, მაგალითად, სახეთა აღწერა, ნიშანთა სივრცის შერწყვა და შეფასება, გადაწყვეტილებათა მიღების წესისა და დისკრიმინანტული ფუნქციების ფორმირება [33,34,35].

კლასტერი განსაზღვრებით არის ნებისმიერი სიმრავლის ისეთი ქვესიმრავლე, რომელშიც ობიექტები გაერთიანებულია რაიმე ნიშნის ან ნიშნების მიხედვით, ხოლო კლასტერირება წარმოადგენს ობიექტთა მოცემული სიმრავლის რაიმე ნიშნის ან ნიშნების მიხედვით არაგადაძვეთ ქვესიმრავლეებად დაყოფა-დაჯგუფების პროცესს.

უნდა აღინიშნოს, რომ თანამედროვე სამეცნიერო ლიტერატურაში ცნობილია კლასტერის ცნების განსხვავებული განმარტებებიც, რაც დაკავშირებულია კლასტერირების კონკრეტულ მეთოდთან ან ალგორითმთან. კლასტერის ცნების განხილვადებული, მკაცრად ფორმალისებული და კონსტრუქციული განმარტება რანგული კავშირების აპარატით მოცემულია ნაშრომში [24]. კლასტერირება მოიცავს სახეთა ამოცნობის პრობლემას, იგი ამოცნობასთან შედარებით უფრო ფართო ცნებაა, ანუ კლასტერირება შესაძლებელია მაშინაც, როდესაც საჭირო არ არის ან ვერ ხერხდება სახეთა ამოცნობა. კლასტერირების პროცესის შედეგები შეიძლება გამოყენებულ იქნეს ამომცნობ სისტემებში მასწავლებელ სისტემად, ხოლო რანგული კავშირების მეთოდით გადაწყვეტილების მიღების პროცესებისათვის საწყის ინფორმაციას სწორედ კლასტერირების შედეგები წარმოადგენს.

ვთქვათ, $\{X\}$ სახეთა ანსამბლის ნიშანთა სიმრავლეა, ხოლო K - კლასტერების რაოდენობა, მაშინ კლასტერირება და მისი შედეგები შესაძლებელია წარმოვადგინოთ შემდეგი გამოსახულებით:

$$\text{CLS}\{X\} = \{\text{CLX}_k\}, \quad k = 1, K, \quad (3.6)$$

სადაც $\text{CLS}\{X\}$ კლასტერიზაციის პროცესის ოპერატორია; $\{\text{CLX}_k\}$ - კლასტერების სიმრავლე.

ნაშრომში [2] დამტკიცებულია თეორემა, რომ მრავალი ცვლადის ნებისმიერი უწყვეტი ფუნქციის მიახლოება ნებისმიერი სიზუსტით შეიძლება წრფივი ფუნქციების წრფივი კომბინაციების და სუპერპოზიციების და ერთი არგუმენტის ნებისმიერი ერთი არაწრფივი ფუნქციის მეშვეობით. ამ თეორემიდან გამომდინარეობს, რომ კონექციონისტური ნეირონული ქსელები აპროქსიმაციის უნივერსალურ საშუალებას წარმოადგენს და შეუძლია ნებისმიერი უწყვეტი ავტომატის ნებისმიერი სიზუსტით იმიტაცია. კლასტერიზაციის ამოცანების გადაწყვეტის ერთ-ერთი ეფექტური ინსტრუმენტია მეორე თავში განხილული მრავალშრიანი პირდაპირი გავრცელების ნეირონული ქსელები.

კლასტერიზაციის ობიექტები ანუ რეალიზაციები წარმოადგენს ვექტორებს, რომელთა ელემენტები ასახავს უცნობი ობიექტის სხვადასხვა მახასიათებლებს და გავლენას ახდენს გადაწყვეტილების მიღების პროცესზე, ანუ მოცემულ რეალიზაციას მიაკუთვნებენ ამა თუ იმ კლასტერს.

ნეირონული ქსელის სინთეზისას განსაკუთრებული მნიშვნელობა ენიჭება ნიშანთა სივრცის სათანადო შერჩევას. იმ შემთხვევაში, როდესაც პარამეტრების რაოდენობა მცირეა, შესაძლებელია რეალიზაციების ვექტორები ერთნაირი ელემენტებისაგან შედგებოდეს, მაგრამ ეკუთვნოდეს სხვადასხვა კლასტერებს, ანუ ნეირონული ქსელის სწავლება შეუძლებელი იქნება. მეორე მხრივ, ნიშანთა სივრცის განზომილების გაზრდისას შეიძლება სასწავლო ნაკრების სიმძლავრე არასაკმარისი გახდეს ნეირონული ქსელის სწავლებისათვის, ასეთ დროს ქსელი მხოლოდ

დაიმახსოვრებს რეალიზაციებს მათი განსოგადების ნაცვლად და კლასტერიზებას ვერ მოახდენს.

კლასტერიზების ამოცანების დაყოფა შესაძლებელია სამ ჯგუფად: პირველ ჯგუფს შეადგენს ისეთი ამოცანები, სადაც ქვესიმრავლეებად დაჯგუფება შეიძლება პიპერსიბრტყეებით - ასეთი სიმრავლეები მიეკუთვნება წრფივად განმხოლოებად სიმრავლეებს; მეორე ჯგუფს შეადგენს არაწრფივად განმხოლოებადი სიმრავლეები; მესამე ჯგუფში შედის ისეთი ამოცანები, სადაც ქვესიმრავლეები იკვეთება.

რეალური ამოცანების გადაწყვეტისას ზოგჯერ ვერ ხერხდება საწყისი მონაცემების ისეთი წინასწარი დამუშავება, რომ უზრუნველყოფილი იყოს სახეთა წრფივად განმხოლოებადობა, რაც ართულებს არაგადამკვეთ ქვესიმრავლეებად დაჯგუფების პროცესს და შესაბამისად, საჭირო ნეირონული ქსელების სინთეზის პროცედურებს. ამრიგად, მნიშვნელოვანია იმის გარკვევა არის თუ არა მოცემული სახეთა ნაკრები წრფივად განმხოლოებადი. წრფივად განმხოლოებადობის პრობლემა განხილულია პ. 2.8 -ში და ცნობილია, რომ მისი დაძლევა შესაძლებელია ერთშრიანი ნეირონული ქსელისათვის დამატებითი შრეების კასკადური მიერთებით. ასეთ მრავალშრიან ნეირონულ ქსელებს შეუძლია რთული კლასტერიზების პროცესების განხორციელება.

ეთქვათ, ნეირონულ ქსელში გვაქვს ერთი ფარული შრე, მაშინ კლასტერიზების პროცესს შეიძლება მიეცეთ შემდეგი ინტერპრეტაცია: ნეირონულ ქსელში შესული ამოსაცნობი სახის რეალიზაცია ანუ შემავალი რაიმე N განზომილების ვექტორი ფარულ შრეში გარდაიქმნება სხვა L განზომილების ვექტორად, რომელსაც გამომავალი შრის ნეირონების შესაბამისი პიპერსიბრტყეები დაყოფს სასურველ კლასტერებად. აქვე უნდა აღინიშნოს, რომ რანგული კავშირებით [24] კლასტერიზების

პროცესებში პიპერსიბრტყეებით ოპერირება არ ხდება, ამიტომ რანგული კავშირებით კლასტერიზაციის მეთოდის გამოყენება ნეიროკომპიუტინგში მოითხოვს განსხვავებული არქიტექტურის მქონე ნეირონული ქსელების შემუშავებას.

ცნობილია, რომ კლასტერიზაციის ობიექტთა სიმრავლე შეიძლება წარმოდგენილი იყოს როგორც რაოდენობრივი, ასევე თვისებრივი სახით; ამასთან, რიცხვითი მონაცემები შეიძლება იკვლევოდეს ნებისმიერ დიაპაზონში, ნეირონული ქსელები კი ოპერირებენ ნამდვილი რიცხვებით ნულიდან ერთამდე ფარგლებში. ამის გამო აუცილებელია საწყისი მონაცემების წარმოდგენა რიცხვითი ფორმით და მათი ნორმირება. ნორმირების პროცესი გულისხმობს სახეთა ნიშნების რაოდენობრივი მნიშვნელობების მასშტაბირებას ისე, რომ შესაძლებელი იყოს მათი თანაბარმნიშვნელოვნად გამოყენება კლასტერიზაციის პროცესებში. არსებობს ნორმირების მრავალი მეთოდი, მარტივი წრფივი გარდაქმნებით დაწყებული და ექსპონენციალური შეფასებების ფორმირებით დამთავრებული, რომლებიც განხილულია შესაბამის ლიტერატურაში, მაგალითად, ნაშრომში [20].

გარკვეულ პრობლემას წარმოადგენს ნეირონული ქსელის მიერ გენერირებული გამოშვებული სიდიდეების (შედეგების) კოდირება იმ შემთხვევებში, როდესაც კლასტერების რაოდენობა ორზე მეტია. კოდირების მრავალ მეთოდს შორის ყველაზე მარტივი და გაერცხლებულია გამოშვებული სიდიდეების წარმოდგენა ვექტორის სახით, რომლის ელემენტებიც შეესაბამება კლასტერების ნომერს, ანუ ვექტორის j -ური ელემენტი შეესაბამება j -ურ კლასტერს. ჩვეულებრივ მიღებულია, რომ კლასტერის ნომერს განსაზღვრავს ნეირონული ქსელის იმ გამოსასვლელის ნომერი, რომელზეც არის სიგნალის მაქსიმალური მნიშვნელობა. მაგალითად, თუ გვაქვს ნეირონული ქსელი ოთხი გამოსასვლელით, ხოლო გამოშვებული ვექტორის ელემენტების მნიშვნელობები

არის (0,35; 0,29; 0,71; 0,44) - ცხადია, რეალიზაცია ეკუთვნის მესამე კლასტერს, ვინაიდან ვექტორის მესამე ელემენტის მნიშვნელობა არის მაქსიმალური გამოსახველვებს შორის. კოდირების ამ მეთოდის გამოყენებისას სასურველია სანდოობის ცნების შემოტანა, ანუ იმის რაოდენობრივად განსაზღვრა, თუ რამდენად სარწმუნოა ის, რომ რეალიზაცია ეკუთვნის ამორჩეულ კლასტერს. სანდოობის დადგენა მარტივად შეიძლება გამომაველი ვექტორის ელემენტების მაქსიმალურსა და მასთან უახლოეს სიდიდეებს შორის სხვაობის განსაზღვრით. ზემოთ მოყვანილ მაგალითში სანდოობის მანქანებელი იქნება $0,71-0,44=0,27$. იმის აღბათობა, რომ ნეირონულმა ქსელმა სწორი კლასტერირება მოახდინა მით უფრო მაღალია, რაც უფრო მაღალია სანდოობის მანქანებელი.

ნეირონული ქსელების სინთეზის ორი ძირითადი მიდგომა არის ცნობილი - კონსტრუქციული და დესტრუქციული. კონსტრუქციული მიდგომისას მინიმალური მოცულობის ნეირონული ქსელის მოცულობის თანდათანობითი გაზრდით მიიღწევა მოთხოვნილი სიზუსტე. აქ საჭირო ხდება ქსელის მოცულობის ყოველი გაზრდის შემდეგ სწავლების პროცესის ხელახალი ჩატარება. შესაძლებელია აგრეთვე კასკადური კორელაციის მეთოდის გამოყენებაც, რაც გულისხმობს შეცდომის მინიმიზაციის მიზნით ქსელის არქიტექტურის კორექტირებას ყოველი ეპოქის (Epoch) დამთავრების შემდეგ. დესტრუქციული მიდგომისას დიდი მოცულობის ქსელიდან ხდება ამ ნეირონებისა და კავშირების მოშორება, რომლებიც ნეირონული ქსელის მიერ გენერირებულ შედეგებზე უმნიშვნელოდ მოქმედებენ. აუცილებელია იმ პირობის შესრულებაც, რომ სასწავლო სიმრავლეში რეალიზაციების რაოდენობა უნდა აღემატებოდეს სინაფსური წონითი კოეფიციენტების რაოდენობას (ზოგიერთი შეფასებით ათჯერ), წინააღმდეგ შემთხვევაში, ქსელი მხოლოდ დაიმახსოვრებს მონაცემებს და ვერ მოახდენს იმ

რეალიზაციების კლასტერიზებას, რომლებიც არ შევიდა სასწავლო სიმრავლეში.

ნეირონული ქსელების არქიტექტურის შერჩევისას საჭირო ხდება სხვადასხვა კონფიგურაციის სტრუქტურისა და ტოპოლოგიის, სხვადასხვა რაოდენობის ნეირონიების მქონე ალტერნატივების შემოწმება შერჩეული კრიტერიუმების მიხედვით.

ნეირონების აქტივაციის ფუნქციის შერჩევა ამოცანის სპეციფიკაზეა დამოკიდებული. გავრცელებულია შემდეგი აქტივაციის ფუნქციები: უბან-უბან წრფივი, ლოგისტიკური, ჰიპერბოლური ტანგენსი, ბიჯური, ექსპონენციალური, ზღურბლური (საფეხურისებრი), სიგნატურული, მოდულის, კვადრატული (ცხრილი 2.1).

ნეირონის აქტივაციის ფუნქცია არის ერთ-ერთი ძირითადი ფაქტორი, რომელიც განსაზღვრავს ნეირონების სწავლების სინქარეს. ნაშრომში [2] აღწერილია გამოთვლითი ექსპერიმენტი, სადაც სხვადასხვა აქტივაციის ფუნქციები შედარებულია სწავლების სინქარის კრიტერიუმის მიხედვით. მიღებული შედეგების ანალიზის საფუძველზე გაკეთებულია შემდეგი დასკვნა:

ნეირონული ქსელების სწავლების სინქარე მნიშვნელოვნად არის დამოკიდებული შერჩეული აქტივაციის ფუნქციის სახეზე, ამასთან, სხვადასხვა ტესტური ამოცანებისათვის ოპტიმალურია სხვადასხვა აქტივაციის ფუნქცია;

- ექსპერიმენტის ტესტური ამოცანებისათვის საუკეთესო მაჩვენებლებთან მიახლოებული შედეგები უზრუნველყო აქტივაციის შემდეგმა სიგმოიდურმა ფუნქციამ:

$$OUT = \frac{1}{1 + e^{-NET}} = F(NET), \quad (3.7)$$

სადაც $NET = XW$ არის ნეირონის სომაში შემავალი სიგნალი; X - ამოსაცნობი სახის რეალიზაციის ვექტორი;

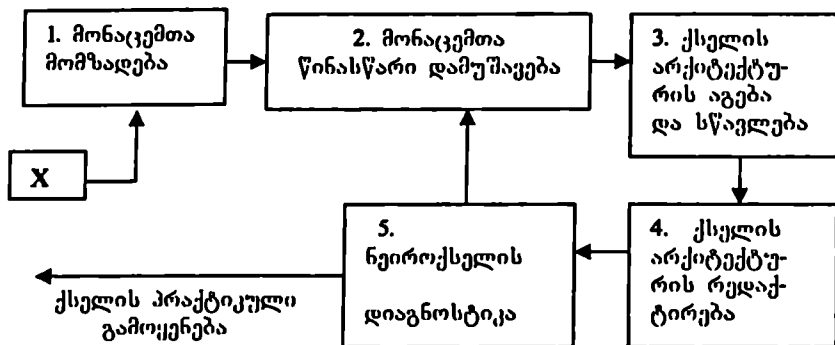
W – სინაფსური წონების ვექტორი; OUT – ნეირონის აქსონზე გენერირებული გამოძევალი სიგნალი.

ქსელების სწავლება შეიძლება მრავალი ალგორითმის საშუალებით. რიგი არქიტექტურისათვის შემუშავებულია სწავლების სპეციალური ალგორითმები, რომელთაც აქვს სპეციფიკური ეპირატესობები და ამავე დროს, ნაკლიც [2,21,31]. ყველაზე უფრო ხშირად ნეიროქსელების სწავლებისათვის გამოიყენება შეცდომის უკუგავრცელების ალგორითმი და მისი მოდიფიკაციები.

არსებობს ნეიროქსელების სწავლების ორი ძირითადი მიდგომა: მასწავლებლით და მასწავლებლის გარეშე. ამ პარადიგმების წარჩეობაში შეიძლება სწავლების ოთხი მთავარი წესის გამოყოფა, რაც განპირობებულია მათთან დაკავშირებული ნეიროქსელების არქიტექტურით, ესენია: შეცდომის კორექციის წესი, ბოლცმანის წესი, ჰების წესი, შეჯიბრის მეთოდი. თავის მხრივ, ამ წესების ბაზაზე შექმნილია სწავლების მრავალი ალგორითმი და მათი მოდიფიკაცია (ალგორითმების ნაწილი განხილულია ზემოთ, სხვები მომდევნო თავებში იქნება წარმოდგენილი)

განხილული მოსაზრებების საფუძველზე ჩამოყალიბებული ნეირონული ქსელის სინთეზის პროცედურები და მათი თანამიმდევრობა ბლოკ-სქემის სახით ნაჩვენებია 3.4 სურათზე.

ბლოკ-სქემაზე წარმოდგენილი ნეირონული ქსელის სინთეზის პროცედურების ელემენტები მოცემულია ქვემოთ.



სურ. 3.4. ნეიროქსელის სინთეზის პროცედურები

ბლოკი 1. მონაცემთა მომზადება.

1.1. მონაცემთა ბაზის შედგენა გადასაწყვეტი ამოცანის დამახასიათებელი მაგალითებისაგან;

1.2. მონაცემთა ბაზის დაყოფა სასწავლო და საკონტროლო სიმრავლეებად, აგრეთვე ვალიდაციური სიმრავლის ფორმირება;

ბლოკი 2. მონაცემთა წინასწარი დამუშავება.

2.1. ამოცანის შესაბამისი ნიშანთა სივრცის შერჩევა, მონაცემთა წარმოდგენა რიცხვითი ფორმით და მათი ნორმირება;

2.2. გამოჩევილი სიდიდეების (შედგენის) კოდირების სისტემის შერჩევა.

ბლოკი 3. ნეირონული ქსელის არქიტექტურის აგება, მისი სწავლება და ქსელის ხარისხის შეფასება.

3.1. ნეირონული ქსელის სტრუქტურისა და ტოპოლოგიის შერჩევა (შრეების რაოდენობა, ნეირონების რაოდენობა შრეში, კავშირები ნეირონებს შორის);

3.2. ნეირონების აქტივაციის ფუნქციის შერჩევა;

3.3. ნეირონული ქსელის სწავლების ალგორითმის შერჩევა;

3.4. ნეირონული ქსელის ფუნქციონირების კრიტერიუმების შერჩევა (სწავლების სინქარე, ამოცნობის დრო, გადაწყვეტილებათა სიზუსტე, სწავლების პროცესის გამჭვირვალობა და სხვ.);

3.5. ნეირონული ქსელის მუშაობის ხარისხის შეფასება შერჩეული კრიტერიუმებით.

ბლოკი 4. ნეირონული ქსელის არქიტექტურის რედაქტირება (სტრუქტურის, ტოპოლოგიის, ნიშანთა სივრცის კორექტირება).

4.1. ნეირონული ქსელის იმ ალტერნატივის საბოლოო ამორჩევა, რომელიც უზრუნველყოფს საუკეთესო შედეგებს;

4.2. ნეირონული ქსელის მუშაობის ხარისხის საბოლოო შეფასება.

ბლოკი 5. ნეირონული ქსელის პრაქტიკული გამოყენება და დიაგნოსტიკა.

5.1. გადაწყვეტილებათა მიღებაზე სხვადასხვა ფაქტორების გავლენის ევრისტიკული შეფასება;

5.2. ამოცანის გადაწყვეტის მოთხოვნილი სიზუსტის უზრუნველყოფა, რისთვისაც საჭიროების შემთხვევაში უნდა მოხდეს მეორე ეტაპზე დაბრუნება.

5.3. ნეირონული ქსელის პრაქტიკული გამოყენება დასმული ამოცანების გადასაწყვეტად.

ნეირონული ქსელის სინთეზის წარმოდგენილი მეთოდია, როგორც აღნიშნული იყო, შეიძლება გამოყენებულ იქნეს არა მარტო კლასტერიზების, არამედ მცირე მოდიფიცირებით სახეთა ამოცნობისა და სხვა ძნელად ფორმალიზებადი ამოცანების გადასაწყვეტად.

3.5. ნეირონული ქსელის სინთეზის მაგალითი

ხელოვნური ნეირონული ქსელის სინთეზის მეთოდის ილუსტრირების მიზნით, მაგალითად მოვიყვანოთ ქართული სტილიზებული ანბანის ამოცნობის ავტომატიზაციის ამოცანის გადაწყვეტის პროცედურები.

ანბანის სიმბოლოების ამოცნობის ამოცანა დაისმება შემდეგნაირად: მოცემულია რასტრული შავ-თეთრი სტილიზებული ქართული ანბანის სიმბოლო, საჭიროა განისაზღვროს ანბანის 33 სიმბოლოდან რომელია ამოსაცნობად წარმოდგენილი. ჩამოვყავლიბოთ ამოცანის მოცემულობა ნეირონული ქსელების ტერმინებით, ფორმულირება იქნება:

მოცემულია ნეირონულ ქსელში შემავალი N განზომილების ბინარული ვექტორი, საჭიროა N შესასვლელის და 33 გამოსასვლელის მქონე ნეირონული ქსელის სინთეზი. ამასთან, გამომავალი ვექტორის ელემენტები შეესაბამება ანბანის სიმბოლოებს და მაქსიმალური მნიშვნელობის მქონე ვექტორის ელემენტის ნომერი განსაზღვრავს ამოსაცნობ სიმბოლოს.

სტილიზებული ქართული ანბანის სკანირებული სიმბოლო არის რასტრული შავ-თეთრი გამოსახულება, ანუ ბინარულელემენტებიანი მატრიცა:

$$X = x_{hv} \quad \forall x_{hv} \in \{0,1\},$$

სადაც X არის N განზომილების მქონე რეალიზაციის ბინარული ვექტორი ანუ მატრიცა; $h=1, H$ სტრიქონის ნომერი; $v=1, V$ - სვეტის ნომერი; H - ჰორიზონტალური განზომილების მნიშვნელობა; V - ვერტიკალური განზომილების მნიშვნელობა. ექსპერიმენტული კვლევების შედეგად შერჩეულ იქნა ნიშანთა სივრცის პარამეტრები: $H=22$ პიქსელი და $V=27$ პიქსელი, მაშინ რეალიზაციის ვექტორის განზომილება იქნება $N=22*27=594$.

პარამეტრების ასეთი შერწყვა განპირობებულია შემდეგი მოსაზრებით: ერთი მხრივ, რეალიზაციის ვექტორის განზომილება უნდა იყოს იმდენად დიდი, რომ არ დაიკარგოს ქართული ანბანის ასოების სპეციფიკური წანაზარდები კბილები, აგრეთვე არ მოხდეს პრიმიტივების ზედმეტი მიახლოება, ხოლო მეორე მხრივ – იმდენად მცირე, რომ რასტრში არ იყოს არაინფორმაციული სტრუქტურები.

ამოცანის ნეიროქსელური ტერმინებით ფორმულირება განაპირობებს შედეგის (ქსელიდან გამოშავალი სიგნალების) კოდირების სისტემის შერწყმას, ანუ გამოშავალი შრის 33 ნეირონი ბუნებრივად განსაზღვრავს ანბანის თითოეული ასოსათვის უნიკალურ გამოშავალ ვექტორს.

ნეიროქსელის არქიტექტურის აგებისათვის შეირჩევა დესტრუქციული მეთოდი, ხოლო ქსელი განისაზღვრება როგორც ორშრიანი პირდაპირი გავრცელების მიმდევრობითკავშირებიანი ნეიროქსელი, მაშინ 594 ნეირონი შეადგენს პირველ შრეს და 33 ნეირონი - მეორე შრეს. ფორმალური ნეირონი წარმოადგენს მაკ-კალოკისა და პიტსის ნეირონის საფეხურისებრი აქტივაციის ფუნქციით.

ნეიროქსელის სწავლების პარადიგმა იქნება მასწავლებლით (ვინაიდან ამ ტიპის ნეიროქსელის სწავლება მასწავლებლის გარეშე შეუძლებელია ხოლო სწავლება განხორციელდება შეცდომის კორექციის წესის ალგორითმის მიხედვით. სასურველია, რომ თითოეული სიმბოლოსათვის სასწავლო და საკონტროლო ნაკრებების რეალიზაციების რაოდენობა გაცილებით მეტი იყოს სინაფსური წონების რაოდენობაზე. მეთოდიკის ეველა სხვა რეკომენდაციის შესრულებით შესაძლებელია ამოცნობის მაღალი საიმედოობის უზრუნველყოფა.

3.6. დასკვნა

“ნეიროინფორმატიკის საფუძვლების” (პირველი ნაწილი) პირველ თავში მოკლედ არის განხილული ნეიროინფორმატიკის იდეოლოგიის სამი წყაროს – ადამიანის ცენტრალური ნერვული სისტემის ფიზიოლოგიის, სახეთა ამოცნობის თეორიისა და მაღალი პარალელიზმის გამოთვლითი სისტემების მნიშვნელოვანი განსაზღვრებები, დებულებები და მეთოდები, რამაც ხელი უნდა შეუწეოს ძირითადი მასალის ათვისებასა და სწორად გააზრებას.

მეორე თავში მოცემულია ისტორიული ცნობა ხელოვნური ნეირონული ქსელების შესახებ, ხელოვნური ნეირონის სტრუქტურა და თვისებები, ერთშრიანი და მრავალშრიანი ხელოვნური ნეირონული ქსელები, დისკრეტული და უწყვეტი ნეირონული ქსელების სწავლების სხვადასხვა მეთოდები და ალგორითმები, სწავლების უკუგავრცელების ალგორითმის დანქარების გზები, კოჰონენისა და გროსბერგის ქსელების სწავლების მეთოდები, განხილულია ლოგიკური არატოლ-მნიშვნელობის, წრფივად განმხილვადობისა და სხვა პრობლემები.

მესამე თავში წარმოდგენილია ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკური მეთოდები, ლოკალური მინიმუმების პრობლემა, სწავლების ბოლცმანისა და კოშის მეთოდები. განხილული მოსაზრებებისა და მეთოდების განზოგადების საფუძველზე ჩამოყალიბებულია ნეირონული ქსელების სინთეზის პროცედურები და მათი თანამიმდევრობა ძნელადფორმალიზებადი ამოცანების გადასაწყვეტად.

“ნეიროინფორმატიკის საფუძვლების” მეორე ნაწილში განხილული იქნება პოპფილდის ნეირონული ქსელები, ასოციაციური ქსელების სხვადასხვა სახეობები, ადაპტიური რეზონანსული თეორია, მისი შესაბამისი ნეიროქსელები და სწავლების მეთოდები, კოგნიტრონი, ნეოკოგნიტრონი, აგრეთვე სხვადასხვა ნეიროემულატო-

რები და პრაქტიკული ამოცანების გადაწყვეტის მაგალითები.

ხელოვნური ნეირონული ქსელების თეორია ამჟამად ფორმირების ეტაპზეა, რაც განაპირობებს პრობლემების დასმისა და ძირითად განსაზღვრებათა მრავალფეროვნებას. უკვე არსებობს ნეიროქსელების მრავალი არქიტექტურა და სწავლების მრავალი ალგორითმი, ხოლო სხვადასხვა ავტორები ნეიროქსელებთან დაკავშირებულ ცნებებს იყენებენ განსხვავებული ინტერპრეტაციით ან მკაცრი განსაზღვრებების გარეშე, ამიტომ საჭირო ხდება ამ ცნებების აქსიომატიკურად განსაზღვრა ან დასუსტება, რასაც მეორე ნაწილში მიექცევა დიდი ყურადღება.

აღწერილი ხელოვნური ნეირონული ქსელები, როგორც აღნიშნული იყო, წარმოადგეს კონექციონისტურ სამეცნიერო-ტექნიკურ მიმართულებას. კონექციონისტური ხელოვნური ნეირონული ქსელები ძნელად ფორმალიზებადი მრავალი პრობლემის გადაწყვეტის ეფექტური საშუალებაა, თუმცა მათაც აქვთ გარკვეული შეზღუდვები. კონკრეტული პრაქტიკული ამოცანებისათვის მათი გამოყენების მკაცრი მათემატიკური დასაბუთება არსებობს მხოლოდ ნეირონული ქსელების მოდელების შეზღუდული ჯგუფისათვის, ნეირონული ქსელების მოდელების ეს ჯგუფიც იყოფა სხვადასხვა ტიპის ამოცანებზე ორიენტირებულ მოდელებად და სხვ.

არსებობს ფართო შესაძლებლობები ნეიროქსელური გადაწყვეტილებების ტექნიკური მახასიათებლების გაუმჯობესების თვალსაზრისით, მაგალითად, ხელოვნური ნეირონული ქსელებით მიღებული გადაწყვეტილებების ეფექტურობის გაზრდის მიზნით, ჩვენი აზრით, კვლევები შეიძლება წარიმართოს შემდეგი მიმართულებებით:

- ფორმალური ნეირონის მოდელის გამოთვლითი შესაძლებლობების გაზრდა, კერძოდ, ნეირონმა უნდა შეძლოს არამარტო შემაჯავლი ვექტორის წრფივი განმხილოება, არამედ არაწრფივი განმხილოებაც. ასეთი ნეირონის მოდელში აღარ იქნება წრფივი ელემენტი

$$F(x) = \sum_n w_n x_n ,$$

სადაც w არის წონითი კოეფიციენტი; x - შემავალი ვექტორის კოორდინატი; n - ვექტორის განზომილება, იგი შეიცვლება შემდეგი ფუნქციით:

$$F(x) = \sum_n w_n \varphi_n(x) ,$$

სადაც $\varphi_n(x)$ არის არაწრფივი ფუნქცია, რომელიც დაშლადია თავისი არგუმენტების მიმართ [34]. იგი შეიძლება დაეშალოს ორთოგონალური ფუნქციების სახით, მაგალითად, ფურიეს მწკრივად ან წარმოვადგინოთ პოლინომური ფორმით

$$\varphi_n(x) = a_{1n}x_n + a_{2n}x_n^2 + \dots a_{kn}x_n^k.$$

ასეთ ფორმალურ ნეირონს შეიძლება ეწოდოს მაღალი რიგის ანუ არაწრფივი ნეირონი, რომლისთვისაც არ არსებობს კლასიკური ლოგიკური არატოლმნიშვნელობის პრობლემა, ვინაიდან იგი არამარტო წრფივად განამხოლოებს ნიშანთა სივრცეს, არამედ შეეძლება ნებისმიერი განმხოლოება, რომელსაც რამდენიმე არგუმენტის პოლინომური ფუნქცია განაპირობებს. ნეირონული ქსელები, რომლებიც სინთეზირებულია არაწრფივი ნეირონებით, გაცილებით მაღალი გამთვლელითი სიმძლავრის და ამავე დროს, შედარებით მარტივი ტოპოლოგიის იქნება. საფიქრებელია, რომ სწავლების პროცედურებიც გამარტივდება და შესაბამისად, გაიზრდება სწავლების სიჩქარე;

ახალი ნეიროქსელური პარადიგმების დამუშავება ანუ კონკრეტული ამოცანისათვის სპეციალური ნეიროქსელური მოდელის სინთეზი. თანამედროვე ნეირონული დოქტრინისა და ახალი ბიომორფოლოგიური მონაცემების თანახმად, ადამიანის ნეოკორტექსში ფიქსირდება ნეირონთა კონგლომერატების სვეტისებრი სტრუქტურები დაახლოებით 600 მილიონი სვეტი 110 ნეირონით თითოეულ სვეტში. ანალოგიით შესაძლებელია, სტრუქტურულად ორგანიზებული ხელოვნური ნეიროქსელების

ფორმირება. პირველ რიგში, ეს აგვაცილებს ზედმეტ სინაფსურ კავშირებს ნეირონთა შორის, რაც დაზოგავს აპარატურულ და პროგრამულ რესურსებს. ამ მოსაზრების კარგი ილუსტრაციაა ქართული ანბანის სიმბოლოები, რომლებიც, როგორც ცნობილია, სამსართულიან ბადეში თავსდება. ანბანის სიმბოლოს რეალიზაცია (ვექტორი, მატრიცა) შეიძლება აგრეთვე წარმოდგენილ იქნეს სამი ფრაგმენტის სახით, რომელთაც შეესაბამება ნეირონთა დამოუკიდებელი ჯგუფები – კონგლომერატები ანუ ამ ჯგუფებს შორის არ იქნება სინაფსური კავშირები, ხოლო ნეირონების აქსონებს, რომლებიც, თავის მხრივ, ნეიროქსელის მომდევნო შრის სინაფსურ კავშირებს ქმნიან, შეეძლებათ მომდევნო შრის ნეირონების განსხვავებული დაჯგუფება. პირობითად, ნეირონების ასეთ დამოუკიდებელ ჯგუფს შეიძლება ეუწოდოთ პოლინეირონი, რომელიც ხელოვნური ნეირობიექტების იერარქიაში დაიკაუებს საშუალოდ პოზიციას ფორმალურ ნეირონსა და ნეირონულ შრეს შორის. შესაძლებელია პოლინეირონული ქსელებით შეიქმნას სტანდარტული მოდულები სხვადასხვა ამოცანების გადასაწყვეტად, მაშინ წინასწარ ნასწავლი სტანდარტული პოლინეირონებით შესაძლებელი გახდება „აწყობა“ რთული მრავალმოდულიანი სტრუქტურებისა, რომელთაც შეეძლება კონკრეტული ძნელად ფორმალიზებადი ამოცანების გადაწყვეტა;

-წრფივი და არაწრფივი ფორმალური ნეირონებისა და მათი შესაბამისი ნეირონული სტრუქტურების ახალი სწავლების წესებისა და ალგორითმების დამუშავება, მაგალითად სწავლების წესების დამუშავება ალტერნატიული ფორმალური ნეირონის მეთოდით.

წარმოდგენილი მასალებიდან ნათელია, რომ ხელოვნური ნეირონული ქსელების თეორიისა და პრაქტიკის ღრმად შესწავლისა და შემდგომი განვითარებისათვის აუცილებელია ერთი შეხედვით ძნელად-თავსებადი ისეთი სამეცნიერო დისციპლინების საფუძვლების ცოდნა, როგორიც არის ნეიროფიზიო-

ლოგია, ფსიქოლოგია, სახეთა ამოცნობის თეორია, მართვის თეორია, პარალელური გამოთვლები და პარალელური გამოთვლითი სისტემები, გამოთვლების თეორია, ხელოვნური ინტელექტის პრობლემატიკისა და შესაბამისი მათემატიკური აპარატის ფლობა, მეორე მხრივ, თვით ხელოვნური ნეირონული ქსელების თეორია გარკვეულ სტიმულებს აძლევს ამ დისციპლინებს ახალი ინსტრუმენტებისა და წარმოდგენების სახით.

თანამედროვე მეცნიერებისა და ტექნიკის უდიდესი პრობლემის – ხელოვნური ინტელექტის შექმნის გადასაწყვეტად, ჩვენი აზრით, აუცილებელია ისეთი ახალი ტიპის სამეცნიერო ჯგუფების შემოქმედებითი მუშაობა, რომელთა თითოეულ წევრს მრავალმხრივი და საფუძვლიანი ცოდნა ექნება.

1. Фон Нейман Дж. Теория самовоспроизводящихся автоматов. Москва, 1971.
2. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996.
3. ნანეიშვილი თ. ქცევის ფსიქოფიზიოლოგია. თბილისი, 1996.
4. <http://www.itcb.serpukhov.su/ctc/ivwork.htm>
5. Strachey C. Time-sharing in large fast computers, Proc. of ICIP UNECCO, 1959.
6. Wilks M.H. The Growth of interest in microprogramming; literature survey, Comp. Serv., V. 1, №3, 1969.
7. Slotnik D.L. The fastest computer, Sci. Amer., V.224, 1971.
8. Однородные микроэлектронные ассоциативные процессоры. Под ред. И.В. Прангишвили. Москва: Советское Радио, 1973.
9. Чоговадзе Р. А., Чигвинадзе О. Д. Структура ассоциативного параллельного процессора с расширенными внутримодульными функциями// Труды ГПИ № 8 (181). Тбилиси, 1975.
10. Чигвинадзе О.Д., Чоговадзе Р. А. Ассоциативный параллельный процессор с расширенными внутримодульными функциями, реализуемый на БИС // Сб. Исследования некоторых вопросов вычислительной техники. Тбилиси: ТГУ, 1973.
11. Чигвинадзе О.Д., Чоговадзе Р. А. Выполнение групповой операции умножения в АП процессоре с расширенными внутримодульными функциями// Труды ГПИ № 8 (181). Тбилиси, 1975.
12. Чигвинадзе О.Д., Чоговадзе Р. А., Чачанидзе В. Г. Выполнение групповой операции деления в АП процессоре с расширенными внутримодульными функциями // Сб. Исследования некоторых вопросов вычислительной техники. Тбилиси: ТГУ, 1973.
13. Чоговадзе Р. А., Чигвинадзе О. Д. Матричный ассоциативный процессор для параллельной обработки данных в реальном масштабе времени // Доклады на

республиканской н.- т. конференции по вопросам внедрения ВТ и АСУ. Тбилиси, 1974.

14. Чоговадзе Р. А. Организация ввода и вывода информации в ассоциативном параллельном процессоре // Доклады на республиканской н.- т. конференции по ВТ. Труды НПО ЭЛВА, ч.1. Тбилиси, 1973.

15. Чоговадзе Р. А. О способах обновления информации в ассоциативных процессорах // Труды всесоюзной школы-семинара по управлению большими системами. Тбилиси, 1976 г.

16. Чоговадзе Р. А. Трехмерный ассоциативный параллельный процессор // Сб. Вопросы кибернетики. Москва, 1973 г.

17. Чоговадзе Р. А., Верулава О. Г. О возможности использования ассоциативного параллельного процессора для распознавания образов в реальном масштабе времени // Доклады республиканской научно-технической конференции молодых специалистов и аспирантов по вычислительной технике. Труды НПО ЭЛВА, часть I. Тбилиси, 1973.

18. Чоговадзе Р. А. К вопросу использования АП процессора для распознавания образов в реальном масштабе времени // Тезисы докладов республиканской научно-технической конференции по вопросам разработки и внедрения вычислительной техники и АСУ для народного хозяйства. Тбилиси, 1974.

19. Галушкин А. И. О современных направлениях развития нейрокомпьютеров // Информационные технологии, №5, 1997.

20. ვერულავა ო., ხუროძე რ. ამომცნობი სისტემების თეორიის საფუძვლები. თბილისი: ტექნიკური უნივერსიტეტი, 2001.

21. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. Москва: Мир, 1992.

22. Минский М. Л., Пейперт С. Л. Перцептроны. – Москва: Мир, 1971.

23. Розенблатт Ф. Принципы нейродинамики. Перцептроны и теория механизмов мозга. Москва: Мир, 1965.

24. ვერულავა ო., ხუროძე რ. რანგული კავშირების თეორია-ამოცნობის პროცესების მოდელირება. თბილისი: ტექნიკური უნივერსიტეტი, 2004.
25. Rumelhart D. E., Hinton G. E., Williams R. J. Learning internal representations by error propagation. In Parallel distributed processing, vol. 1, Cambridge, MA: MIT Press. 1986.
26. Sejnowski T. J., Rosenberg C. R. Parallel networks that learn to pronounce English text. Complex Systems 1: 1987.
27. Parker D. B. Second order back propagation: Implementing an optimal $O(n)$ approximation to Newton's method as an artificial neural network. Manuscript submitted for publication. 1987.
28. Stornetta W. S., Huberman B. A. An improved three-layer, backpropagation algorithm. In Proceedings of the IEEE First International Conference on Neural Networks, eds. M. Caudill and C. Butler. San Diego, CA: SOS Printing. 1987.
29. Hecht-Nielsen R. Counterpropagation networks. In Proceedings of the IEEE First International Conference on Neural Networks, vol. 2, San Diego, CA: SOS Printing. 1987.
30. Горбань А.Н. Обучение нейронных сетей. Москва, СП ПараГраф, 1990.
31. Chogovadze R., Khurodze R., Verulava O., Martirosian I. Decisions making modeling on basis of Mini and Maxi portraits by neural networks. Georgian technical university, Transactions, Vol.3 (453), Tbilisi, 2004.
32. Круглов В.В., Длин М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. Москва, 2001.
33. ჩოგოვაძე რ. ხელოვნური ნეირონული ქსელების სინთეზი სახეთა ამოცნობის პროცესებისათვის. თბილისი, სტუ-ს შრომები № 4 (454), 2004.
34. ჩოგოვაძე რ. ნეირონული ქსელების სინთეზი გადაწყვეტილებათა მიღების პროცესებისათვის // "მართვისა და ენერგეტიკის პრობლემების" საერთაშორისო კონფერენციის მასალები. თბილისი, 2004.

35. ვერულაძე ო., ხუროძე რ., ჩოგოვაძე რ. მსგავსების ზომის ფორმირება სიმრავლეების შორის // "მართვისა და ენერგეტიკის პრობლემების" საერთაშორისო კონფერენციის მასალები. თბილისი, 2004.
36. ვერულაძე ო., ხუროძე რ., გრიგალაშვილი გ. ნეირონისა და ნეირონული ქსელის მოდელირების პრინციპების შესახებ // სტუ-ს შრომები № 4 (420), თბილისი, 1998.
37. ვერულაძე ო., ხუროძე რ., ქონლაძე თ., მაძალუა მ. ნეირონებში დენდრიტების კოეფიციენტების საწყისი მნიშვნელობების შერჩევა // სტუ-ს შრომები №3 (427), თბილისი, 1999.
38. ჩოგოვაძე რ. ელექტრონული გამოთვლითი მანქანების ტექნიკური ექსპლუატაციის საკითხები მართვის ავტომატიზებულ სისტემაში. თბილისი, 1975. - 124 გვ.
39. Элементы теории биологических анализаторов. Под. ред. Н.В. Позина. Москва: Наука, 1978.
40. Цыпкин Я. З. Основы теории обучающихся систем. Москва: Наука, 1970.
41. Бонгард М.М. Проблема узнавания. Москва: Наука, 1967.
42. Braun H. Neuronal Netze. Springer, Berlin, Heidelberg, Germany, 1997.
43. Ивахненко А. Г., Квиташвили А. А., Иванов В. И., Растрингин Л. А., Персептрон - Система распознавания образов. Киев: Наукова Думка, 1969.

წინასიტყვაობა.....	4
თავი I. შესავალი	5
1.1. ნეიროინფორმატიკის იდეოლოგიის წყაროები.....	5
1.2. ბუნებრივი ნეირონისა და ნეირონული ქსელების ბიომორფოლოგიური აღწერა.....	12
1.2.1. ბუნებრივი ნეირონი და მისი სტრუქტურული ელემენტები.....	12
1.2.2. ბიოლოგიური ნეირონის ძირითადი ფუნქციები	17
1.2.3. ნერვული უჯრედების ურთიერთქმედების მექანიზმები და სისტემები	22
1.2.4. ნეირონული დოქტრინა და პარადიგმების ცვლა.....	24
1.3. მაღალი პარალელულობის გამოთვლითი სისტემები.....	27
1.3.1. კომპიუტერების არქიტექტურის ევოლუციის ძირითადი ეტაპები	27
1.3.2. კომპიუტერების არქიტექტურის კლასიფიკაცია	31
1.3.3. ნეიროკომპიუტერები და პარალელური პროცესორები.....	34
1.3.4. პარალელური მატრიცული პროცესორები	36
1.4. სახეთა ამოცნობის თეორიის ელემენტები.....	44
1.4.1. ამოცნობის პროცესების ძირითადი ცნებები და ტერმინები	44
1.4.2. სწავლების პროცესები ამომცნობ სისტემებში.....	48
1.4.3. გადაწყვეტილებათა მიღება ამომცნობ სისტემებში.....	50
1.4.4. გადაწყვეტილებათა მიღების მეთოდების კლასიფიკაცია.....	51
თავი II. ხელოვნური ნეირონი და ნეირონული ხელეპი	54
2.1. ისტორიული ცნობა ხელოვნური ნეირონული ქსელების შესახებ	54
2.2. ხელოვნური ნეირონის სტრუქტურა და თვისებები.....	57

2.3. ერთშრიანი და მრავალშრიანი ხელოვნური ნეირონული ქსელები	63
2.4. ხელოვნური ნეირონული ქსელების კლასიფიკაცია	66
2.5. ხელოვნური ნეირონული ქსელების სწავლება	70
2.6. პერსეპტრონები	73
2.7. ლოგიკური არატოლმნიშვნელობის პრობლემა	75
2.8. წრფივად განმხილვადობის პრობლემა	78
2.9. პერსეპტრონის სწავლება	81
2.10. უწყვეტი ქსელების დელტა-წესით სწავლების ალგორითმი	85
2.11. ნეირონული ქსელების სწავლება უკუგავრცელების ალგორითმით	88
2.11.1. პირდაპირი გავრცელების პროცედურა	95
2.11.2. უკუგავრცელების პროცედურა	95
2.11.3. წონების შეთანაწყობა ფარულ შრეში	97
2.12. სწავლების უკუგავრცელების ალგორითმის დაჩქარების მეთოდები	99
2.13. შემხედრი გავრცელების ქსელები	102
2.13.1. შემხედრი გავრცელების ქსელების ნორმალური ფუნქციონირების რეჟიმი	104
2.13.2. კოჰონენის შრის სწავლება	106
2.13.3. წონითი ვექტორების საწყისი მნიშვნელობების შერჩევა	110
2.13.4. გროსბერგის შრის სწავლება	112
თავი III. ხელოვნური ნეირონული ქსელების სწავლების სტოქასტიკური მეთოდები	114
3.1. სწავლების ლოკალური მინიმუმების პრობლემა	114
3.2. სწავლების ბოლცმანის მეთოდი	119
3.3. სწავლების კოშის მეთოდი	121
3.4. ხელოვნური ნეირონული ქსელების სინთეზის პროცედურები	122
3.5. ნეირონული ქსელის სინთეზის მაგალითი	132
3.6. დასკვნა	134
ლიტერატურა	139
შინაარსი	143

რედაქტორი ი. სემიკინა

ტექნიკური რედაქტორი ნ. ცირეკიძე

კორექტორი ნ. დოლიძე

გადაეცა წარმოებას 01.11.2004. ხელმოწერილია დასაბეჭდად 01.02.2005.
ქალაქის ზომა 60X84 1/16. გარნიტურა აკადემიური. პირობითი ნაბეჭდი თაბახი
9,25. სააღრიცხვო-საგამომცემლო თაბახი 8,25. ტირაჟი 100 ეგზ. შეკვეთა № 14

გამომცემლობა „ტექნიკური უნივერსიტეტი“, თბილისი,
კოსტავას 77



სტუ-ს სტამბა, თბილისი, კოსტავას 75